An eXist XML Database for a Database Project
Draft 2010-10-25

Introduction and Context

This project will require the student to create multiple linked XML documents, load them into a native XML database and then query the collection to obtain printed output. The suggested environment is *eXist* (open source and free), a native XML DB. The choice of the XML context area is largely up to the individual student and will be student/instructor approved.

The intent here is to become familiar with a crucial data representation called *eXtended Markup Language* (XML). This language, its syntax, protocols, and extensions, has taken over the internet and the web, and is essential to know about and be able to use intelligently. XML is the fundamental basis of *web services* and the web services extensions to business enterprise level architectures called the *Service Oriented Enterprise* (SOE).

Why Use the 'eXist' Database as an Exemplar?

To illustrate the use of XML in a database context, which is what we are working on, I have chosen an open source, freely downloadable 'native' database named *eXist*. Created by Wolfgang Meier in early 2000, it is now in version 1.5, and it has won a number of honors in its category of native XML bases. For database students, the appeal will be its ease of installation and its support of the XQuery Language. Being able to query a database, as you know, is the whole point of having one and this DB has an XQuery processor built in. Later in this document I show how a DB can be built and then queried.

Getting Started

Download the eXist Distribution

You can go to the eXist web site, *http://exist-db.org/,* and download a zip file. Unzip it into your hard drive. Double click to install. The install should place an icon on your desktop. If not, you can create one by rt-clk on your desktop window > new shortcut> navigate to the eXist bin directory and select the startup.bat (for windows).

You will also be asked for a password, pick something you will remember. (I simply use *admin-admin*).

Optional (you may not need to do this as the bat file sets these environmental variables)

To set up the path to the executable that is within the distribution, do the following

1. Go to control panel>system'
2. Go to the advanced tab in the system properties dialog box> environment variables
3. In your personal user panel, set up EXIST_HOME. For example, if you extracted it to your hard drive as: C:\exist
   Then create a new variable EXIST_HOME and set it to C:\exist

Start the eXist Program

If you don't have an icon for the *startup.bat,* create one as described above. NOTE: You will also

need to create a shutdown icon, *shutdown.bat*, also from the bin directory, since failing to properly shut down the DB will result in subsequent startup errors.

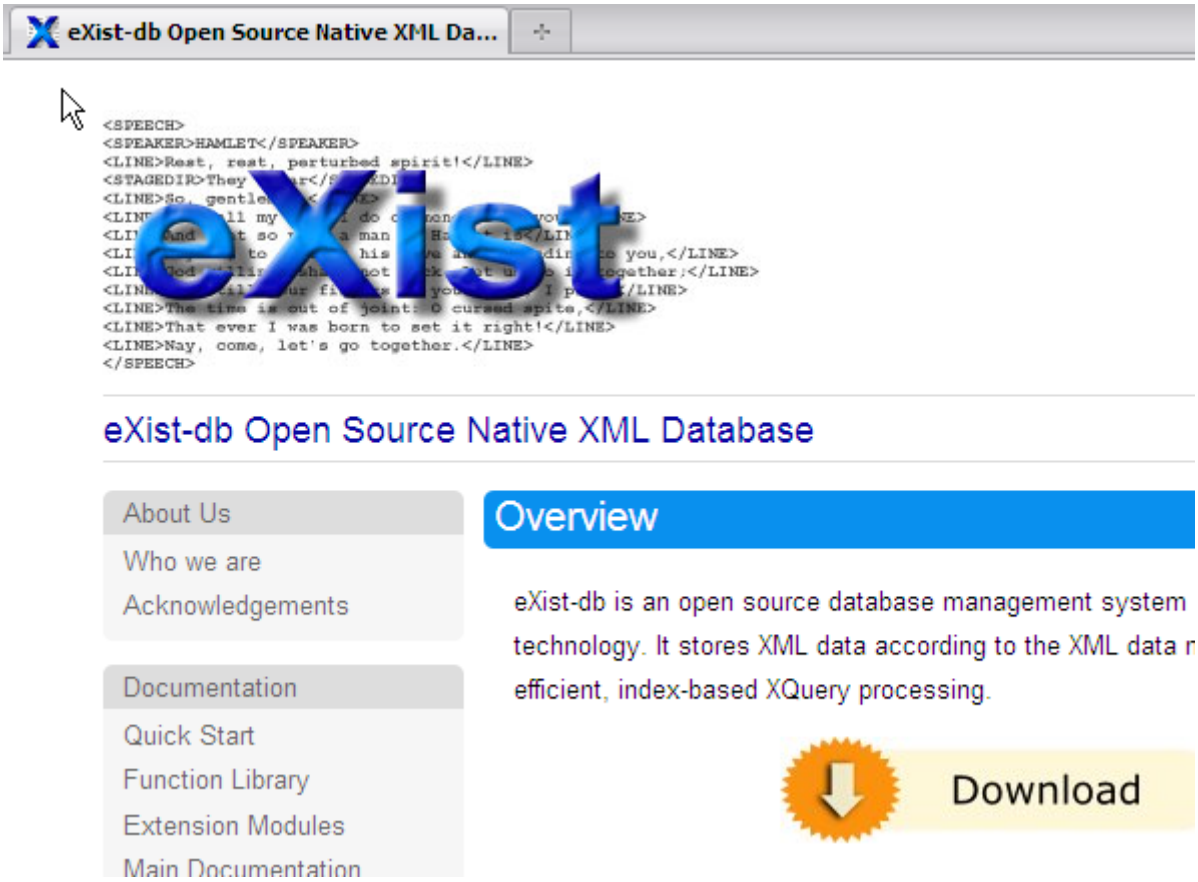
Double click the *startup* batch file and a command/shell window should open and begin scrolling log information output similar to the below example, as it builds the DB code. Note that the distribution has a built in web server (Jetty) that makes this a 'web application' running a server on localhost (127.0.0.1)

```
25 Oct 2010 10:20:15,796 [main] INFO  (JettyStart.java [run]:90) - Con-
figuring e
Xist from C:\aaPrograms\exist\conf.xml
. . .
. . .*lots of log output in here. The ending log output is shown next
25 Oct 2010 10:20:22,578 [main] INFO  (JettyStart.java [run]:175) -
eXist-db has
 started on port 8080. Configured contexts:
25 Oct 2010 10:20:22,578 [main] INFO  (JettyStart.java [run]:177) -
http://local
host:8080/exist
25 Oct 2010 10:20:22,578 [main] INFO  (JettyStart.java [run]:179) - ---
```

At this point the Jetty web server is running and is waiting to accept browser input on port 8080. So, open a browser (Firefox preferred) and type:

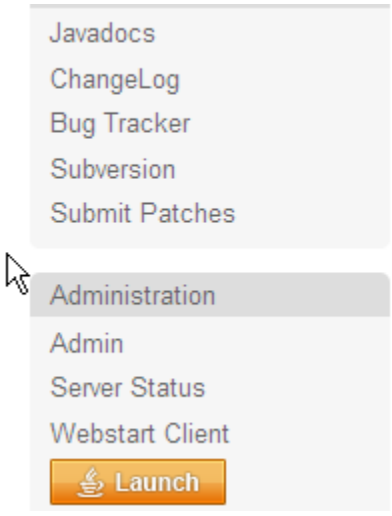
http://localhost:8080/exist/index.xml

The screen below is from the local distribution. If you see this, you are good to go.

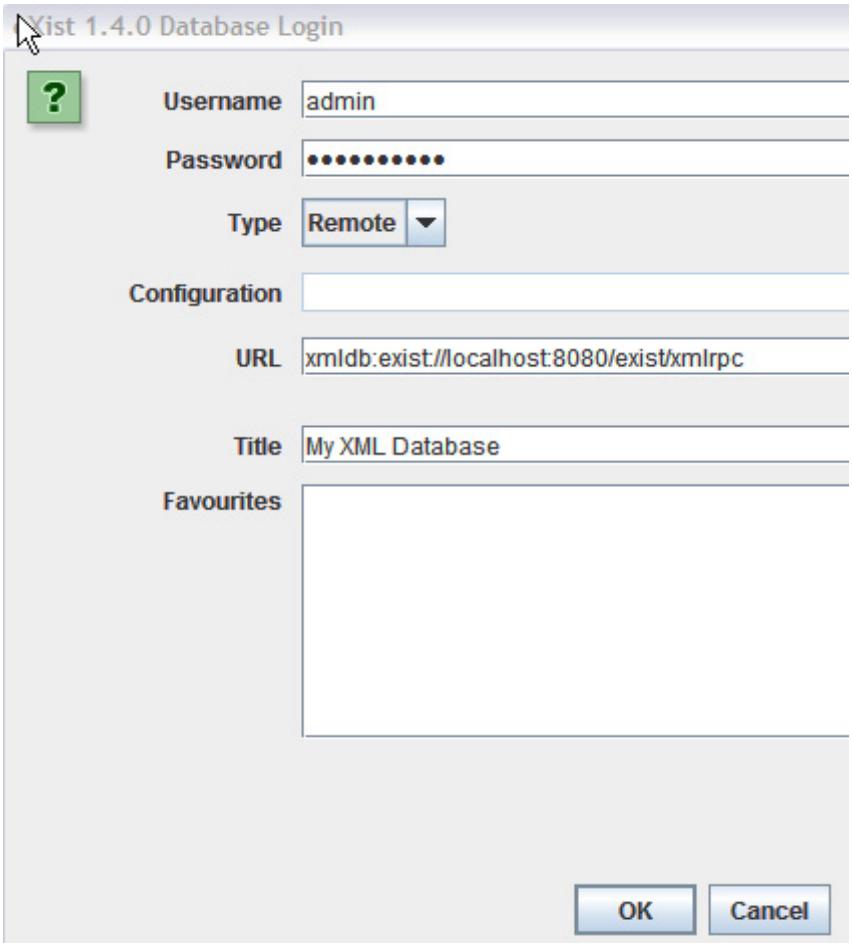


Launching a Graphical Client Program

Scrolling down the menu options on the left, you will see a *Webstart Client*, that, when launched, will allow you to graphically administer your database. Click the *Launch* bar.



Below is a login panel that comes up. When I installed eXist I gave it a password ‘adminadmin’ which is a generic password, useful for experimental work.



Creating and Viewing XML Document Collections.

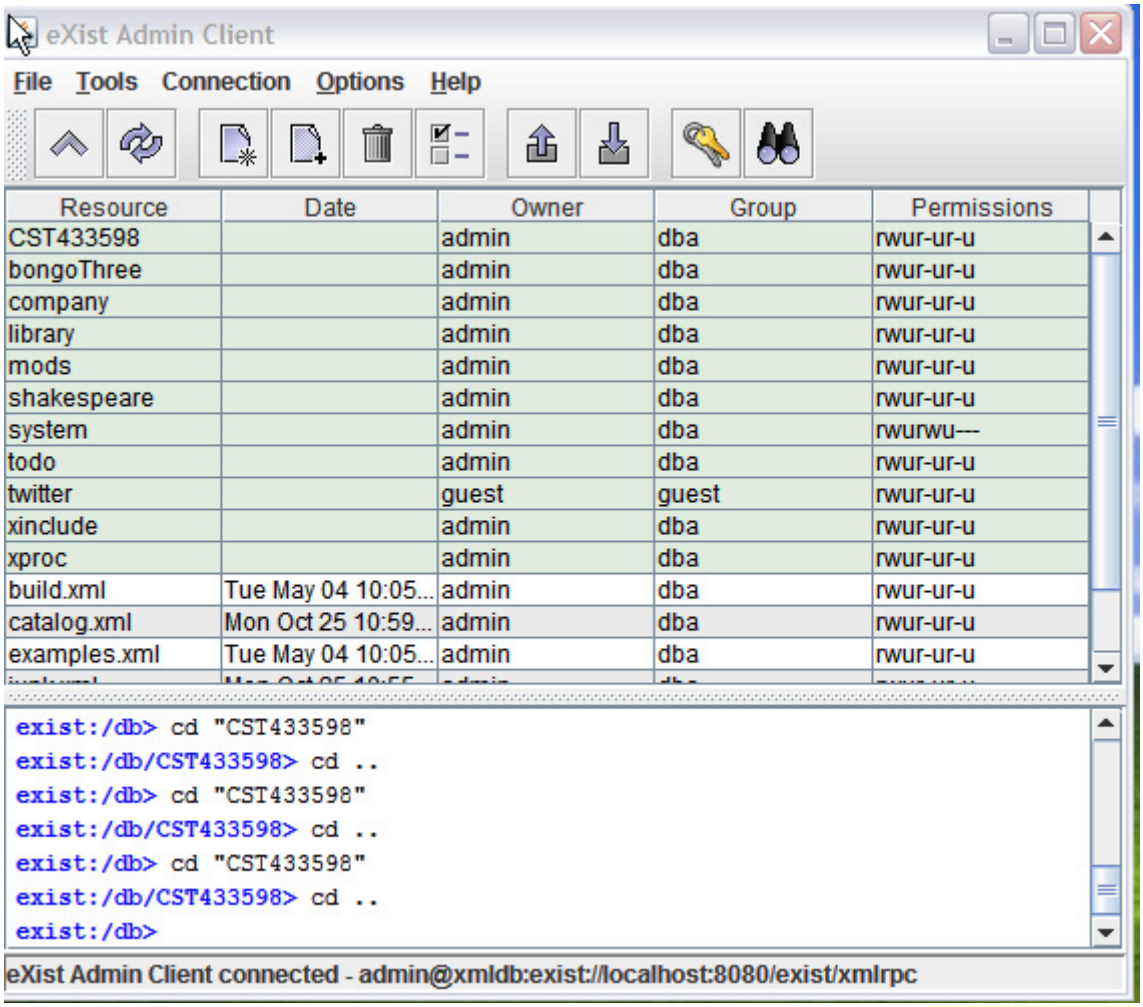
Once you are logged in, you will see a window as below. This is a view of the top level of xml *document collections* as shown by the Web Start Client GUI launched as above. The one of interest here is one I just built, *CST433598*. This will be *container* collection for three files (initially), *catalog.xml*, *prices.xml*, and *order.xml.* (You can upload any XML files you have created).

For the relational model recall that we had database schemas (such as APP) that included individual table, views, and indexes. In this XML context, we have collections of documents instead of tables.

Top Level View of Document Collections

If you double click on these entries, their children will appear. Subsequent double clicks will open the base documents as is shown below. Notice that the GUI operations are mirrored by shell commands. For example, accessing the CST433598 collection resulted in a

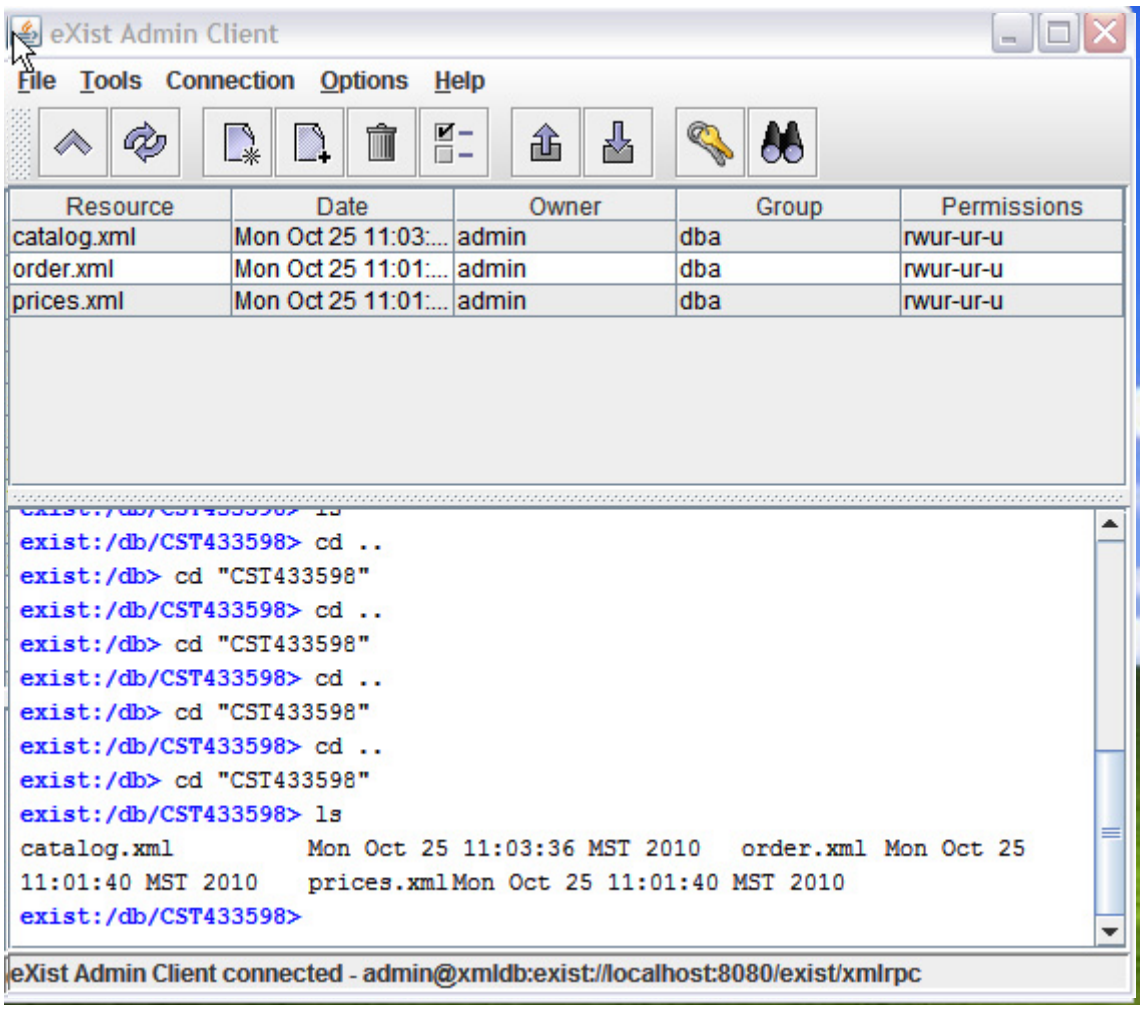
```
>cd "CST433598" command. Where 'cd' means change directory.
```



Viewing the Children Documents of Top Level Containers (Collections)

If you double click on the CST433598 entry, you get the next screen shot. This shows the child documents of the container CST433598, namely the three documents: *catalog.xml*, *prices.xml*, and *order.xml*.

If you double click on one of these, it will open in an editor that you can use to modify and then reload the file back to the DB.

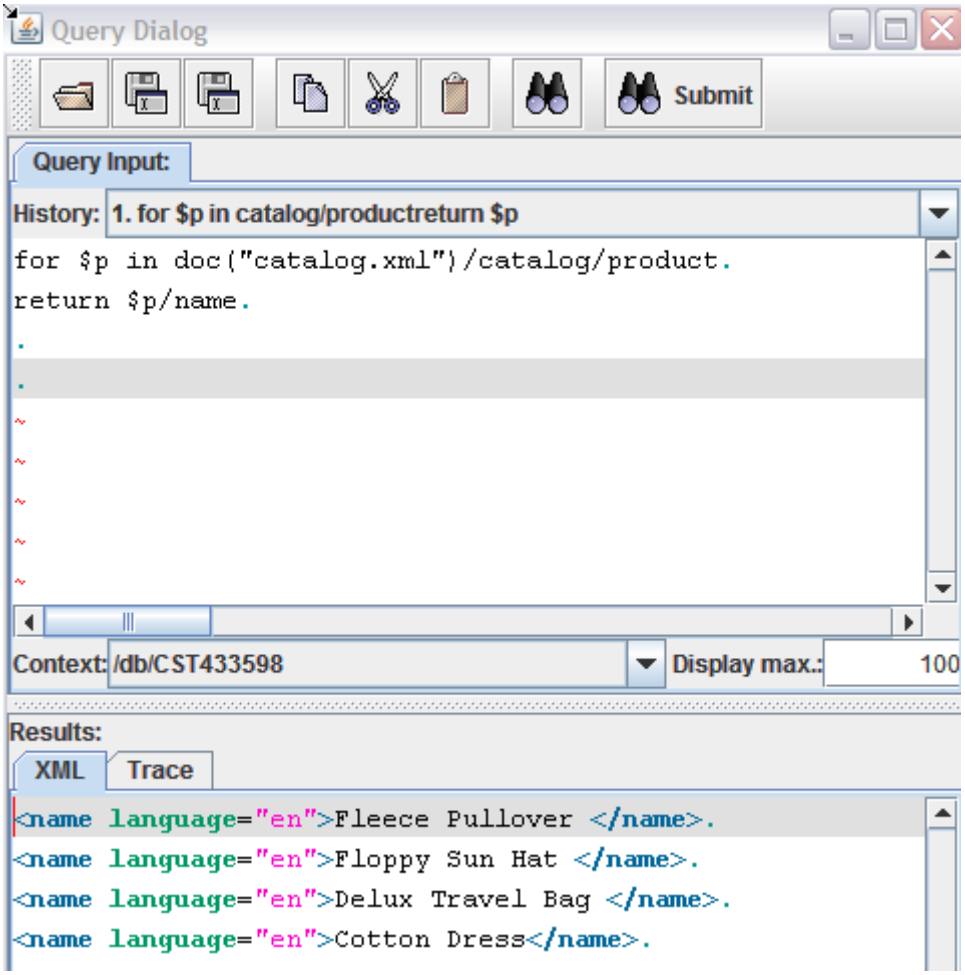

eXist Admin Client
File Tools Connection Options Help
Resource Date Owner Group Permissions
catalog.xml Mon Oct 25 11:03:... admin dba rwur-ur-u
order.xml Mon Oct 25 11:01:... admin dba rwur-ur-u
prices.xml Mon Oct 25 11:01:... admin dba rwur-ur-u
exist:/db/CST433598> cd ..
exist:/db> cd "CST433598"
exist:/db/CST433598> cd ..
exist:/db> cd "CST433598"
exist:/db/CST433598> cd ..
exist:/db> cd "CST433598"
exist:/db/CST433598> cd ..
exist:/db> cd "CST433598"
exist:/db/CST433598> ls
catalog.xml Mon Oct 25 11:03:36 MST 2010 order.xml Mon Oct 25
11:01:40 MST 2010 prices.xml Mon Oct 25 11:01:40 MST 2010
exist:/db/CST433598>
eXist Admin Client connected - admin@xmldb:exist://localhost:8080/exist/xmlrpc

The Document: catalog.xml

```
<!--.
    Document   : catalog.xml.
    Created on : Oct 25 2010 .
    Author     : rrucker.
    Description:.
        A simple exammple of a clothing catalog to illustrat
        XQuery. .
-->.
<catalog>.
    <product dept="WMN">.
        <number>557</number>.
        <name language="en">Fleece Pullover </name>.
        <colorChoices> navy black </colorChoices>.
    </product>.
    <product dept="ACC">.
        <number>563</number>.
        <name language="en">Floppy Sun Hat </name>.
    </product>.
    <product dept="ACC">.
        <number>443</number>.
        <name language="en">Delux Travel Bag </name>.
    </product>.
    <product dept="MEN">.
        <number>784</number>.
        <name language="en">Cotton Dress</name>.
        <colorChoices>white gray </colorChoices>.
        <desc>Our <i>best</i> brand of shirt!</desc>.
    </product>.
</catalog>.
```

Writing the XQuery

The code below shows an *XQuery* against the *catalog.xml* document. The *for* syntax begins a for loop as in other programming languages, and the *$p* is a dummy iterator variable that holds each encountered *product* element as the loop executes. So, the processor starts with the *catalog.xml* document, finds the *catalog* element and then sequentially examines each *product* element and returns the name associated with that particular *product* element.



Summary

This tutorial showed you how to set up and start the eXist DB. A few pre-written XML files were uploaded to the DB and then an XQuery script was shown to query one of the XML files and return parts of it.

References

eXist-db.org manuals are the best source

Walmsley, Pricilla, *XQuery* (2009) Prentice Hall is the best book on XQuery.

w3Schools is a good source for quick introductions to many XML topics.