Endeca Community Solutions

ITL Process

Dear Endeca Community user,

On July 31, 2015 we closed EndecaCommunity.com, and asked users to instead take advantage of the collaboration tools and communities now provided by Oracle.

Why we did this? Because EndecaCommunity.com has fulfilled its purpose.

We created EndecaCommunity.com to fill a temporary void, to support teams working with Endeca technology in both Commerce and Data Insight capacities. The intention was to provide continuity through the transition period that followed Oracle's acquisition of Endeca, until Oracle had ramped up its own Endeca support community.

Now that Oracle's Endeca communities and formal user groups are well established, it's better for you as an Endeca user to find all the support you need from OAUG Special Interest Group and Oracle sources, rather than split your time having to cross-post to multiple disparate sites to find the answers you need.

We recommend joining the [Endeca Community on LinkedIn](#) or the [OAUG Endeca Special Interest Group](#). More support for Endeca can be found on the Oracle community website at [Technical Questions](#) and [Endeca Information Discovery](#).

On the following pages you will find the most frequently asked questions about ITL Process on EndecaCommunity.com and the best solutions as determined by members of the Community.

Thank you for taking advantage of EndecaCommunity.com and for trusting RealDecoy as your Endeca authority.

Sincerely,

Richard Isaac
President & CEO, RealDecoy

Table of Contents

ITL Process

Limiting DGIDX log file size

Q: Is there a way to reduce/limit the file size of the DGIDX log?

A: You're only able to control the number of log directory backups to store via the numLogBackupscomponent property, not the size of the actual logging done. You can however include a post dgidx script to tar/zip the generated log file after the dgidx is executed in an attempt to save space.

Remove application from Endeca

Q: I am trying to entirely remove an application from Endeca. I know that it has to be removed from the following components:
- EAC
- Workbench (IFCR)
- Remove any Record stores (if exist)
- Delete project directory

How would I be able to do that?

A: For the latest Oracle Endeca Commerce 3.1, you should be able to use the following command to remove an application definition
./control/recommend.sh --remove-app {YOUR_APP_NAME}
This command should remove your application configuration and definition from both the EAC and Wrorkbench (IFCR).
To remove your record stores, you would need to run the following command
$CAS_ROOT/bin/component-manager-cmd.sh delete-component -n {RECORD_STORE_NAME}
You can get the list of record stores using the following:
$CAS_ROOT/bin/component-manager-cmd.sh list-components
NOTE: you should look for all component types equal 'RecordStore'
The rest should be easy.

Initialize services failure

Q: I am getting the following exception while creating new Endeca application on starting initialize_services.sh. Please help me sort this out.

    Finished initializing record stores
    Importing sample content...
    [01.17.13 11:41:40] INFO: Checking definition from AppConfig.xml against existing EAC provisioning.
    [01.17.13 11:41:41] INFO: Definition has not changed.
    [01.17.13 11:41:41] INFO: Packaging contents for upload...
    [01.17.13 11:41:41] SEVERE: Unable to close output stream
    Caused by java.util.zip.ZipException
    java.util.zip.ZipOutputStream finish - ZIP file must have at least one entry
    [01.17.13 11:41:41] SEVERE: Caught an exception while invoking method 'importNode' on object 'IFCR'. Releasing locks.
    Caused by java.lang.reflect.InvocationTargetException
    sun.reflect.NativeMethodAccessorImpl invoke0 - null
    Caused by java.util.zip.ZipException
    java.util.zip.ZipOutputStream finish - ZIP file must have at least one entry
    Failed to import sample content.

A: If you used Discover app's deployment template to create a new application, then you have to remove (or comment-out) importing sample content & media in the initialize_services.sh file.

Dgidx error (property and dimension with same name)

Q: I am seeing this error during baseline update (Dgidx:

ERROR   04/09/13 13:50:14.936 UTC (1365515414936) DGIDX {dgidx,baseline} Property "Typeahead.suggestions" and dimension (id=8) have the same name. A property may not have the same name as any dimension. FATAL   04/09/13 13:50:14.937 UTC (1365515414937)   DGIDX {dgidx,baseline} Unable to initialize dimension [8] "Typeahead.suggestions" WARN   04/09/13 13:50:14.937 UTC (1365515414937)   DGIDX {dgidx,baseline}   Lexer/OLT log: level=-1: 2013/04/09 09:50:14 | INFO| Disabling log callback

I only have the "Typeahead.suggestions" dimension defined (no property with that name is defined) not sure what the issue might be?

A: Found it! FCM uses a file called "fcm.conflicts.xml" that is stored in "data/state" directory. This file defines which attributes (properties or dimensions) are controlled by FCM and which of them are controlled by Dev Studio. I made sure that the "Typeahead.suggestions" dimension is controlled by FCM so that the schema.csv definition is what takes effect.

How do I check which Hot Fix has been installed for the MDEX?

Q: Is there a way that I can check which version of a hot fix has been installed for the MDEX? I'm working with MDEX 6.1.3 and there are 3 hot fixes that I know about, but I need to know which of those (if any) has been installed so I can replicate the environment on a new server.

A: Execute the dgraph executable with the version flag
C:\Oracle\Endeca\Server\2.3.0\endeca-server\dgraph\bin>dgraph.exe --version
Oracle Endeca Server version 2.3.0
Endeca dgraph version <strong>7.4.0.637212
Interface specification: READS: 7.4.0e

Unexpected end of stream error (forge input record adapter)

Q: I am using an Input Record Adapter with Fixed Width. I'm getting the below error in forge when it tries to read the file:

> FATAL   06/18/13 14:23:15.532 UTC (1371565395532)   FORGE   {baseline}: Unexpected end of stream. Aborting
> ERROR   06/18/13 14:23:15.667 UTC (1371565395666)   FORGE   {baseline}: Unknown Exception. Exiting.

I have OEC 3.1.1 installed. Any help would be appreciated.

A: I would start by checking the file encoding and making sure that it's UTF-8 with BOM if not, convert it to that. Seems like your width is possibly varying in some parts of the file. You can always try pre-processing the file with AWK if you can access to Bash, convert it to a vertical file and just forget about fixed width.

How do I use web service as a data source for forge or Endeca?

Q: Is there any way to use webservice as data source in pipeline? I have requirement to get data from web service and dump it in Endeca. For this, I need some idea on how can be this achieved. Is it possible to do so?

A: You can try implementing a record adapter. But I'm not sure why you want to access the webservice directly from pipeline. You can dump the webservice output to a flat file and use that as an input for pipeline. It will be a cleaner implementation and easier maintenance wise.

Size of Indexed Data in Endeca

Q: Is it possible to get the size of indexed data(generated indices) in Endeca ? What is the location of the indexed data (indices)? Does this size of generated data have any impact on the performance of Endeca?

A: Related information can be found by clicking on the stats link in the Endeca JSP reference implementation. Details can be found about performance summary, disk usage etc.

How to change application name

Q: I have created an application called "EndecaIsComplex" and I was able to access the same at localhost:xxxx/EndecaIsComplex/. Now I need to change the name to say "EndecaIsFun" (localhost:xxxx/EndecaIsFun/). Can I change this name without creating a new application? If yes, what are the changes I need to make this possible?

A: From the above problem statement, looks like you would like to change the context root while accessing the application. This can be done by replicating the discover.xml found at the following location:
\Endeca\ToolsAndFrameworks\3.1.2\server\workspace\conf\Standalone\localhost.

The newly created xml should contain the context path as EndecaIsFun. If you are planning to change the name of the Endeca application, then go to the configuration folder and change the app name wherever applicable in the following location \config\script. Then re-initialize the application and perform the baseline update.

Baseline Update - Update some properties only, not entire index

Q: I have a requirement from my client related to update a property in Endeca every 30 minutes. The property is the availability, set to true or false, because this searchable property is used to filter the result set when navigating with ATG. The requirement is to update all products (over 30,000) but they don't want to run the baseline update, they need something faster only for this property.
A partial update cannot be used because it didn't change description or price, so ATG retrieves no records. Any suggestions? Is a baseline update the only solution?

A: Endeca does support this functionality in Partial update using a custom pipeline. Unfortunately when using ATG and CAS this is not available to you by default. I have the same problem with Hybris that also communicates with Endeca through CAS and uses its own partial pipeline.
Yes you can do it, but you would have to replace a lot of the export work ATG does for you with your own code as well as write your own custom partial export pipeline in Developer Studio and probably also have to do all the data transformation manually.

Dynamically add properties to the pipeline config (auto generate properties)

Q: I'm wondering if anyone has tried programmatically adding properties to the pipeline config. When I add a property with Endeca Dev Studio I can see that the `.render_config.xml` and `.proper_refs.xml` files are updated with additional property. Is there any reason I couldn't rewrite these files programmatically outside of dev studio and deploy them?

A: You can create any of the xml produced in Dev studio. Endeca facilitates this with an xml api reference, giving you a programming reference to all the features of Dev Studio. http://docs.oracle.com/cd/E28912_01/DeveloperStudio.612/pdf/XMLReference.pdf

Figuring out what xml files you need to change for a given purpose is not that difficult. First, create a template project and copy the folders to a temp directory. Then make the change in Dev studio and save the project. Then use a folder comparison tool like DiffMerg to compare the two directories and see what files changed.

Invalid input: Invalid rollup key "product.repositoryId".

Q: My problem is about the connection with MDEX engine via Endeca cartridges. In Endeca Experience Manager I created a home page collection, and created only one default page template and applied to all locations. Here's my index.jsp:

```
<%@ include file="/includes/context.jspf"%>
<dsp:page>

<dsp:importbean bean="/atg/dynamo/Configuration" />
<dsp:importbean
bean="/atg/endeca/assembler/cartridge/manager/DefaultActionPathProvider" />
<dsp:importbean bean="/atg/endeca/assembler/droplet/InvokeAssembler" />

  <head&gt;&lt;title&gt;MWS Home Page&lt;/title></head>
  <body>

 <dsp:droplet name="InvokeAssembler">
    <dsp:param name="contentCollection" value="/content/Web/Home Pages" />
    <dsp:oparam name="output">
      <dsp:getvalueof var="homePageContent"
vartype="com.endeca.infront.assembler.ContentItem" param="contentItem" />
    </dsp:oparam>
  </dsp:droplet>

  <c:out value="${homePageContent}"/>

  <c:if test="${not empty homePageContent}">
    <dsp:renderContentItem contentItem="${homePageContent}" />
  </c:if>
  </body>
</dsp:page>
```

RealDecoy

Chrome Browser shows me:

{contents=[{@type=OneColumnPage, name=Content, HeaderContent=[], MainContent=[{sortOption=sku.margin|1, subRecordsPerAggregateRecord=ALL, relRank=nterms,maxfield,glom,exact,static(sku.margin,descending), fieldNames=[common.id], boostStrata=[], @type=ResultsList, name=Results List, buryStrata=[], @error=com.endeca.infront.navigation.NavigationException: com.endeca.navigation.ENEException: HTTP Error 404 - Navigation Engine not able to process request 'http://localhost:15000/graph?node=0&merchrulefilter=endeca.internal.nonexistent &sort=sku.margin|desc&groupby=product.repositoryId&offset=0&nbin s=12&allbins=2&autophrase=1&autophrasedwim=1&filter=OR%28produ ct.catalogId%3acatalog10001%29&irversion=640'., recordsPerPage=12}]}], @type=ContentSlot, ruleLimit=1, contentCollection=/content/Web/Home Pages}

I tried to pass this URL in browser and I got the error message:
http://localhost:15000/graph?node=0&merchrulefilter=endeca.internal.nonexistent&sort=sku.margin|desc&groupby=product.repositoryId&offset=0&nbins=12&allbins=2&autophrase=1&autophrasedwim=1&filter=OR%28produc t.catalogId%3acatalog10001%29&irversion=640">http://localhost:15000/graph?node= 0&merchrulefilter=endeca.internal.nonexistent&sort=sku.margin|desc&gro upby=product.repositoryId&offset=0&nbins=12&allbins=2&autophrase =1&autophrasedwim=1&filter=OR%28product.catalogId%3acatalog10001%29&am p;irversion=640
404: Not Found
Invalid input : Invalid rollup key "product.repositoryId".

A: Problem was solved. Solution: I needed to set the attribute <property name="$repositoryId" text-searchable="true" is-rollup-key="true"/> in **/atg/commerce/endeca/index/product-sku-output-config.xml**

Deleting all the records/data from Endeca

Q: I want to know the correct steps for deleting all the data from a particular app in Endeca. I am using MDEX 6.4.x, CAS 3.1.x , Tools and Framework 3.1.x and platform service 6.1.x

Here are the reasons why I want to delete all the records from Endeca. We are using the integrated ATG-Endeca 10.2 platform and use ProductCatalogSimpleIndexingAdmin to trigger baseline. The admin first does a repository export and then triggers a baseline. Yesterday, due to some additional filters the admin exported 0 records and indexing was executed successfully thereafter. After the successful indexing, we expected to see zero products in jspref but it still showed the previous data. Now we want to confirm the course of events that is happening, i.e. if baseline was successfully executed with 0 products why is Endeca still showing products? Is there any fail-safe mechanism wherein if there are zero products Endeca will retain the older data?

That is why we want to clear the data once from Endeca, and then again run indexing with 0 products.

A: ProductCatalogSimpleIndexingAdmin can also show success messages with zero records.. It doesn't mean that zero records will be written to Endeca but you need to validate the time stamp of CAS record to store binary files /opt/endeca/CAS/workspace/state/ApplicationName_en_data/data/storage/index-xxxxx.db

Also check storage time stamp of ApplicationName_en_schema,ApplicationName_en_dimvals,ApplicationName_en_prules

The ATG layer will not generate product related data with zero records. I observed the same with categories that don't have any products associated with them. ATG Extraction layer will take care of this. The last generated indices were not updated. The best way to refresh data is to update the indices file in MDEX and restart the dgraphs. Not sure from Endeca side. Keep an empty file in apps/ApplicationName/test_data/baseline and execute baseline update from Endeca /control folder andrun initialize_services.sh to clean application/data .

How do I clean the Indexed data in Endeca (Dgraph data)?

Q: I need to clean my indexed data and start from fresh data, in my development environment.
I tried with cleaning the data directory "<appname>/data/dgraphs/DgraphA1/dgraph_input", But cleaning this will bring another issues like Dgraph not being able to start. What is the legal way to clean whole data?

A: Why are you cleaning? What are you trying to achieve? As a best practice: DO NOT DELETE FOLDERS. This is not a supported method of cleaning in my humble opinion.

When you run a new baseline it wipes out the previous index so there isn't a need to 'Clean'.

That said: sometimes you need to prune the dimensions to get rid of old dimension values that were 'cached'. This can be achieved by adding a flag in AppConfig.xml. Google: "Oracle Endeca flag prune" and you should get a hit on the required flag.
I get concerned when people ask about cleaning, as it normally means some data they didn't expect is showing up. This can be caused by dirty data in the source system. It is not uncommon for Endeca to highlight the dirt in ones data. You should actually expect it and 'clean' it in the source system instead of manipulating it in Endeca.

How will the Forge pipeline definition ('/data/processing/pipeline.epx') be created?

Q: I have created an application, using the CRS deployment template as given in
http://docs.oracle.com/cd/E36434_01/CRS.10-1-2/ATGCRSInstall/html/s0401appendixacreatingtheendecaapplic01.html But when I start the Forge, it is saying that the pipeline.epx is not there. Has anyone experienced this kind of issue?

A: You should use ATG Layer (atg/commerce/endeca/index/ProductCatalogSimpleIndexingAdmin) to index data in CRS. Once the initial baseline is successful then you can initiate the same from /control/baseline_update. Remember to always use ATG jobs because you have to extract data from ATG repositories to Endeca.

Endeca Load Balancing

Q: I have configured 2 MDEX hosts in AppConfig.html using the administrators guide. Also I have configured 1 Digraph for each of the MDEX server. Here's my AppConfig.xml file: https://drive.google.com/file/d/0B39N8ztL-yN3enRIOXJpZy0xQm8/view?usp=sharing">AppConfig.xml I have set up MDEX servers on 2 different hosts (not localhost). I have a Droplet that uses presentation API to query the MDEX servers. I am using a rest client to fire the query from my localhost. My problem is I am able to fetch the data from the 1st host but not the second. Can someone see if my configuration is correct? Or do we have to do more configuration?

If I have these servers set up and running, how can I do a load balancing of these servers (the servers are running on different systems)? I don't want to use a hardware load balancer.

A: Nothing looks horrible in your AppConfig.xml. Did you run initialize services after updating the definition to have 2 hosts?
If you don't want to use a hardware load balancer, (I would suggest you do use dedicated hardware unless you are 'just testing'.) apache servers can be used to do load balancing. A quick Google search will show you how to set it up.

Changing baseline and partial scripts used by EndecaScriptService

Q: We are trying to modify the baseline and partial update scripts used by the EndecaScriptService in the ATG Endeca integration modules. However, we're not sure exactly what scripts are being called and where those scripts are to make the modifications. We're assuming it's the same ones found under the ToolsAndFrameworks>11.1.0>deployment-template directories, correct?
Could someone point us in the right direction on where we can find and make the changes to the scripts?

A: You want to update AppConfig.xml in your application folder. (/path/to/apps/MyApplication/config/script/) This is written in bean shell and can be customized to call various java class and/or shell scripts. This is where you want to make the change and where all the components for your baseline are configured. You won't actually change the AppConfig.xml file but likely one of its child files, as it imports several of the files found in that folder.

https://docs.oracle.com/cd/E28906_01/Common.622/pdf/OracleEndecaGuidedSearchAdminGuide.pdf
This document describes several of the components in the script and will likely be helpful to get an understanding of it. There is another document that actually shows you how to write a custom component for appconfig.xml but I did not have time to look for it. Just read up on BeanShell (what appconfig is written in) and you will be able to do what you need to do.

Load Balance digraph cluster

Q: I have setup an Endeca cluster as per this Endeca cluster setup article:
http://mandarshinde.com/oracle-endeca-cluster-setup/
I need to do a load balance in the cluster using apache. Can someone direct me to a blog/article for the configuration? I'm really confused regarding where I should place the load balancer. I use an ATG-Endeca integration environment and I'm using presentation API to query the mdex (something like http://ourownjava.com/endeca/endeca-navigation-query-example-using-presentation-api/).
Now that I have the Endeca cluster how/where should I put the load balancer? I use a property file to define which mdexhost the presentation API should use to query. How do I direct the query to a particular mdex (through load balancer) so that the presentation API can use it?

A: All your queries should be sent to the load balancer, you don't want to specify a particular mdex. I think it might help you if you spent some time understanding what load balancers do, that might help you get a better understanding of how and where they fit.

To give you a code explanation:

```
MDEX_HOST = [load balancer host]
MDEX_PORT = [load balancer port]
private static ENEConnection createConnection() {
        return new HttpENEConnection(MDEX_HOST, MDEX_PORT);
    }
</code></pre>
```

I'm not sure why you are using the Presentation API instead of using the assembler, but I'm sure you have your reasons.