Docushare 4.0 Updates/Enhancements

Updated Look and Feel for WebUI:
- Paging UI for Collections and Search Results
- Sorting by column headers in listings
- Improved breadcrumbs
  shows you where you have been
- Improved SiteMap
  converted into to javascript based control for tree structure

Personal Home Page:
-My Docushare page
  static layout no way user can customize currently
- Favorites
- History
   links you have been to recently
- Personal Collection
-My tasks window
-Simple search options window (little triangle next to advanced search)
  to refine simple search with canned menus
  can be defined in VDF

Core Server :
- Enhanced performance (restructure schema and SQL)
- Add MailMessage support - not initially exposed in DS 4
- Add Instance Properties (for email, routing, and scanning)
  not exposed in UI available in next release
- Add Records Management support
-Expanded capacity
-Access logging expanded - not exposed in DS 4 UI
captures API requests
list of mapping in DSSession
-Display sequence
  re-arrange how things are displayed - custom props specifically
-Performance
handle IDs
new indexes
updated improved queries
enablement for different SQL server driver
connection tool management feature-specify the min number of active
and the max-will close the ones over the min after inactive for
specified time
-DB ping tests

Core Server Improvements (con't):
-Change object class – exposed only in API
-Versions and renditions are visible and accessible
-separate upgrade utility run by installer

Table changes:
-deleted bit in acl and link tables
-change size of integer column
-auto move preferred version-
-display new window
-paging increment value for user
-link desc table-links are dynamic-no UI added new links
-class display sequence table
-used defined order-menu props
-new indexes-8 new indexes-see sql

Performance Updates:
ACL change, delete expunge/restore, orphans, and trash display
-Config changes
ping stuff
sql hints – for experienced DBAs
enable pool management
batch size –batch size queries are performed
-DSCommonprops-enable summarize removed, smtp port
-DSServer-access added
-Logs-DSServer in trace shows begin and end

Known SQL driver problem that MS is working.

Search/Index:

Overview of object indexing and index_all process:
When an object is created, modified, deleted an event is generated. The Search server picks up the event and creates a job from it. For an indexing job, the Search server calls the DSServer to get the object's data (properties and content) and attaches them to the job. Verity is then called to submit the indexing job.

- Moved away from VDK implementation
- Move to latest K2
- Custom gateway (C++) has been eliminated. No need to recompile sources on different platforms to update to newer versions of Verity software.
- Added thesaurus support
When enabled, the operator <thesaurus> is automatically pre-pended to the search literal. This searches the original search literal including its synonyms. The Verity supplied supplied default thesaurus file is located in the locale directory being used. For englishx locale, it is located in <ds4>/lib/verity/k2/common/englishx/vdk30.syd. This file can be dumped to a readable text file, modified and converted back to binary using a Verity supplied tool, mksyd.
-ClassWeights
This is a list of classnames and the weights applied to them to modify the ordering of the search result. If a classname is not specified it gets a default weight of 10. For example, the classWeights entry could be:
ClassWeights=Document[100], Collection[95], User[80], Group[80], URL[80], Bulletin[50], BulletinBoard[50], Calendar[20], Event[0], SavedQuery[10]

Fixes made for performance and past common problems:
- uses bulk mode in getting the parent Collections of objects
- can now use more than 1 thread to get data from DSServer
- can now use more than 1 thread to process indexing job, even on a single Verity Collection
  For a 4-CPU system, use 2 threads to get data, 4 threads to index
- Object properties that are not 'system' properties are now mapped to an internally defined Verity field names.
For example, the Document.author is mapped to text_field0. There are extra 7 string fields, 5 date fields, 3 int fields and 2 float fields pre-created to be used for mapping newly added properties. This prevents re-indexing of the Verity Collection when some properties are added to the schema. You can see the property mapping in <ds4>/docucoll/property.map file.

dsindex:
- Documents are now processed last since they take the longest
Sets Search server state to R/W while running dsindex - enables searching with *partial* result and indexes objects that are currently uploaded or modified or deleted.
Search:
highlighted snippet is displayed with the search result on a Document objects instead of the abstract.
Search is called directly from the DSServer's search subsection to the k2server (search).
Config:
DSServerProperties.properties file additions for search:
searchWithThesaurus=false
classWeights=
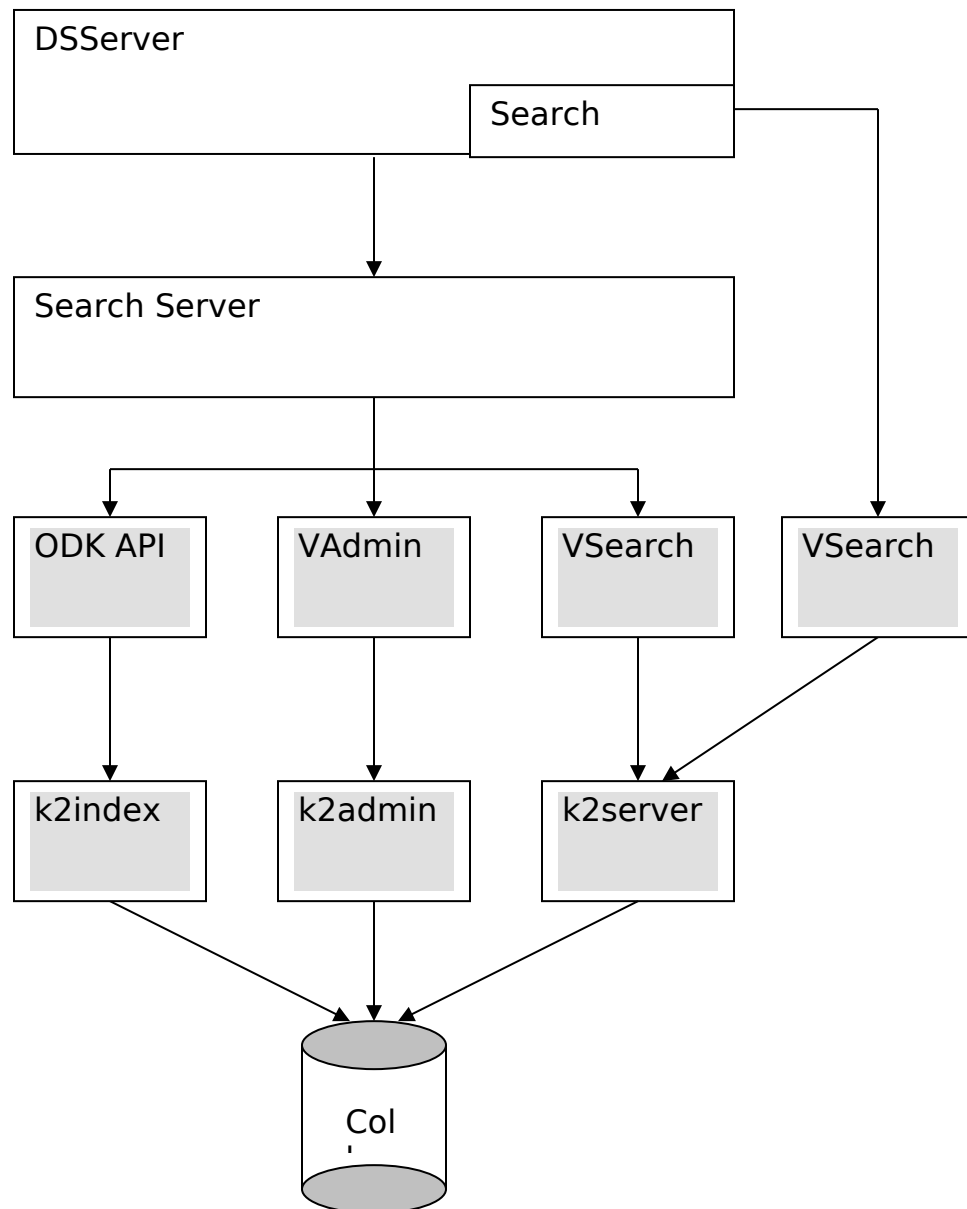SearchServer.properties changes
Removed:
- maxMemory - cannot set on K2 server
- maxFiles - cannot set on K2 server
- verityLogLevel
- verityLogFile
- maxIndexRetry
New:
- numberOfIndexingThreads
- numberOfDataThreads
- k2AdminAlias
- k2AdminPort
- k2IndexServerAlias
- k2IndexServerSpec
- k2SearchServerAlias
- k2SearchServerSpec

XEROX DocuShare®

DS4 Search and K2 Servers

```
DSServer
                                    Search

Search Server

   ODK API      VAdmin      VSearch      VSearch

   k2index      k2admin      k2server

                    Col
```

Search TroubleShooting Tips:

During initial installation and when DS4 is first started, I recommend that you check the processes running – make sure that all the DS4 processes are there and also the k2admin, k2index and k2server processes. Even if k2admin may have successfully started, it is possible that there could be a problem with the k2index or k2server and the monitor program cannot check that.

Problem:

Customers should no longer use the old workaround of deleting the docucoll or coll_<locale>_X directory before running dsindex.
When somebody deletes the contents of docucoll directory and then runs dsindex index_all, you will get an error from the k2admin server, saying that the collection already exists. This is because the Collection is still registered in the k2admin.

Solution:

Shutdown DS4.
Run config_k2.bat or sh on a new command window in <ds4>/bin directory. This will reset the k2admin data.
Start DS4 and run dsindex index_all again.

DS4.0 Notification/Subscription:

-Events – categorized into Object_Events, Document_Events, Login_Events, Link_Events, Class_Events, Subscribe_Events, Config_Change and Classification_Change.
For Subscription services, it listens to Object, Document, Link, and Subscribe events.
For Search server, it listens to Object, Document, Link, and Class events.
For any server or service that creates DSSession, it indirectly creates a listener that listens to Config_Change and Class events.
There are other services, which may listens to certain types of events. Such as, Routing, Sparrow, etc.
Classification_Change is a new event that is for Record Management service.
Object_Class_Change is a new event. It indicates the class got changed.
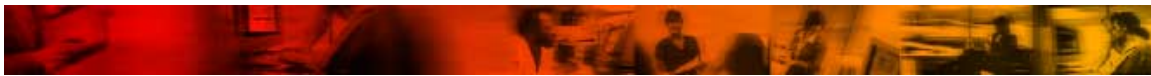
Subscription – Immediate, Daily and/or Weekly

Performance enhancements:
Some local caching to avoid extra traffic on server inquiries
Eliminated steps of user authentication and object inquiries for
Object_Delete events
A few minor message text changes

-IMQ – A SunOne MessageQueue product

Support IMQ 3.0.1. (IMQ 3.5 is planned) with file system.
It runs on a separated JVM and starts or stops through Monitor Control.
The IMQ is installed under <DOCUSHARE_HOME>/imq directory. Its
executables or scripts are existed in <DOCUSHARE_HOME>/imq/bin.
The instance of IMQ mailstore, named 'DSBroker', is in
<DOCUSHARE_HOME>/imq/var/instances.
Two server config files, .bindings and objstore.properties reside under
that folder. The mailstore specific config is in
<DOCUSHARE_HOME>/imq/var/instances/DSBroker/props/config.prope
rties.
The application loggings are recorded in imq.stdError. A separated
server logs are in
<DOCUSHARE_HOME>/imq/var/instances/DSBroker/log/log.txt.
Log and Debug Tips –
There are few logs that are related to IMQ/Notification/Subscription
Basically, Imq server has to be in running mode to have other services
running. When start Docushare server, it starts imq server first. Its
logging are recorded in imq.stdError. Imq starts normally when the end
of log contains ...
....Loading persistent data...
....Broker "DSBroker@<host>:7676" ready.
Other imq server logs can be found in log.txt
The monitor starts Notification that establishes connection with imq
server. If the connection request is not success, it will retry the
connection action for 30 times. The number of retry and time interval
between retries are runtime switches and can be added to the start
command.
Default is 30 reties and 5 seconds apart...commands are below
-Dcom.xerox.docushare.jmsretrycount=30
-Dcom.xerox.docushare.jmsretryinterval=5
Notification logs are in DSServer.stdOut and DSServer.stdError.
Subscription logs are in Subscription.stdOut and Subscription.stdError.
To get some quick investigation, logLevel=trace should be sufficient.
For detailed investigation, logLevel=debug may provide more
information. It should ONLY be used when the possible problem been

determined in a specific area. It has too much information to flood the real cause.

In DS4, most of false alarm messages have been downgraded to TRACE/DEBUG level. Most likely, you can skip the messages that have problem messages text but at TRACE or DEBUG level. Even it may contain exception information.

When there are Exceptions recorded, it may encounter a possible problem. Based its timestamp traced back to DSServer.std* or imq.stdError. You may find the related real causes. In the worst case, you may need to check imq\var\instances\DSBroker\logs\log.txt.

-Configuration
notification.properties – no new properties.
logLevel=trace - logs the data/work flow of notification service.
logLevel=debug – logs the detail debugging data of the service.
subscription.properties – new optional properties with associated helptext.
To skip processing a certain type(s) of event, specify a string of a comma separated events as value in a single line. Can be empty.
help-skipTypes=***valid types- [OBJECT_CREATED, OBJECT_PROP_CHANGED, OBJECT_PUBLISHED, OBJECT_DELETED, OBJECT_EXPUNGED, LINK_CHANGED, OBJECT_ACL_CHANGED, OBJECT_OWNER_CHANGED, OBJECT_CLASS_CHANGED, OBJECT_RESTORED, DOCUMENT_LOCKED, DOCUMENT_UNLOCKED, LOGIN_SUCCEEDED, LOGIN_FAILED, SUBSCRIBE_DAILY, SUBSCRIBE_WEEKLY, CLASSES_CHANGED, CLASS_LABEL_CHANGED, CONFIG_CHANGED]
skipTypes=
To eliminate subsequent events generated by a particular operation. eg. Add a new version may emit obj_prop_change, obj_expunge, obj_created, obj_acl_changed, link_change events, etc.
help-spamOK=***yes- allow multiple-subscription messages for each recipient per transaction, or no- ONLY one subscription message for each recipient per transaction
spamOK=yes
To accommodate the speed of a server. Same transaction may take a little bit longer on one machine than the other. It is to support when spamOK=yes.
help-timeGap=***deault 100 msec, The time interval that is considered the events from same emitter are from the same operating transaction.
timeGap=100.
To specify the time interval of gabage collection process. It is to support when spamOK=yes.
help-gcTime=***default 10 min, The time interval that garbage collection process got triggered.

gcTime=10
To activate the elapse time logging of internal processes. The logLevel has to be set to the level of trace or debug.

-Configuration (con't)
help-printDebugTime=***no- hide debug time intervals, yes- print debug time intervals
printDebugTime=no
The password for accessing authenticated SMTP session. Customer enters the password of the admin email account as plaintext. An encrypted string replaces it after docushare running. For smtp service that allows anonymous access, it shows an encrypted empty string.
smtpKey=
To indicate the main DS version that patches apply for. It was a required property for subsUpd2 patch.
patchVer=ds4

Troubleshoot and Known Problems –
Performance issue – mass acl changes populates or deletion on top hierarchy node object may involves long process time. It may eventually timeout the default max allocated time for timely subscription (default is 8 min). Use the runtime switch -Dcom.xerox.docushare.runtime=8 to make adjustment or using skipTypes to skip processing this type of event.
Multiple subscription messages per single transaction – see spamOK and timeGap settings to consolidate the events. Use with caution. If timeGap is set too large, it may eliminate events emitted by other transaction.
WARNING!! user may expect to receive certain message but it never gets materialized.
Customizable Subscription Message – the template settings in subscription.properties are for plaint text form recipient. To make it more comprehensive, customer needs to use subscription scripts with DSDE to implement more sophisticate message.
Imq server may encounter some internal broker exceptions. These situations may end up clogging the notification and cause it to stop functioning. The Imq 3.5 has some fixes to resolve this issue and some performance enhancement. Upgrade to imq 3.5 is planned for next release.

Docushare 4 Component Versions:
Verity KeyView = Version 8.1
DralaSoft DWE = Version 3.3
Sun Java JDK = 1.4.2_05-b04

SunONE IMQ = 3.0.1-SP2
Tomcat = 4.1.30
MSDE = 2000 SP3
PostgreSQL = 7.4.3

Col