White Paper

EMC® Documentum® Interactive Delivery Services Accelerated: Step-by-Step Setup Guide

A Detailed Review

Abstract

This white paper is a step-by-step setup guide for users who would like to utilize the replication capability of EMC Documentum Interactive Delivery Services Accelerated (IDSx) in a multi-server environment. This paper guides users on how to publish content/metadata to multiple servers through IDSx. The paper also helps in understanding how to effectively administer the IDSx deployment.

July 2011

EMC²

Table of Contents

Executive summary

In most scenarios, large websites must be deployed on multiple servers to achieve a higher throughput, and high availability in a load balanced environment. In addition, content published by these websites may be managed in a central content management repository using content management tools such as Documentum Web Publisher, Documentum Page Builder, Documentum Webtop, Documentum CenterStage, Documentum Digital Asset Manager, Documentum Media Workspace, and so on.

As a result, there is an implicit need for a tool that can publish content on demand or on schedule from a central repository to multiple servers.

This paper discusses installation, configuration, usage, and troubleshooting details of the Interactive Delivery Services Accelerated (IDSx) software.

This paper details the step by step procedure to set up the IDSx deployment with multiple target servers, and helps to effectively manage and leverage all the rich sets of IDSx features.

Audience

This paper is intended for System Administrators, System Integrators, Technical Support staff, and others who are involved in managing the IDSx product.

Terminology

IDSx

Interactive Delivery Services Accelerated(IDSx) is an advanced version of the Interactive Delivery Services(IDS) product. The IDS products was formerly known as Site Caching Services (SCS).

The architecture of IDSx is the same as the architecture of the SCS product. It utilizes similar concepts from a usability perspective such as configuration management in Documentum Administrator, Source and Target combination, and so on.

IDSx Source

IDSx Source is a web application hosted on a JBOSS instance and is co-located with Content Server. The IDSx Source is primarily responsible for accepting requests pertaining to publishing and replication, exporting data to a temporary location, and transferring exported data to the IDSx target using an embedded UDP-based accelerated file transfer technology.

IDSx Staging Target and IDSx Replication Target

A stand-alone Java program running on a specified port. It accepts various requests from IDSx source such as login, file transfer commit, and so on. It also applies publish batch to the target directory and synchronized the target database.

Documentum Administrator (DA)

An EMC Documentum product used to administer various other products in the Documentum product stack (including IDSx).

Replication Manager

Replication Manager is a part of the IDSx target, and is an underlying process initiated by an IDSx source for a webc config id, whenever there is a request for replication.

Installation

This section provides a snapshot of a typical IDSx multi-server production deployment, and outlines the step-by-step procedure to install the various components of IDSx.
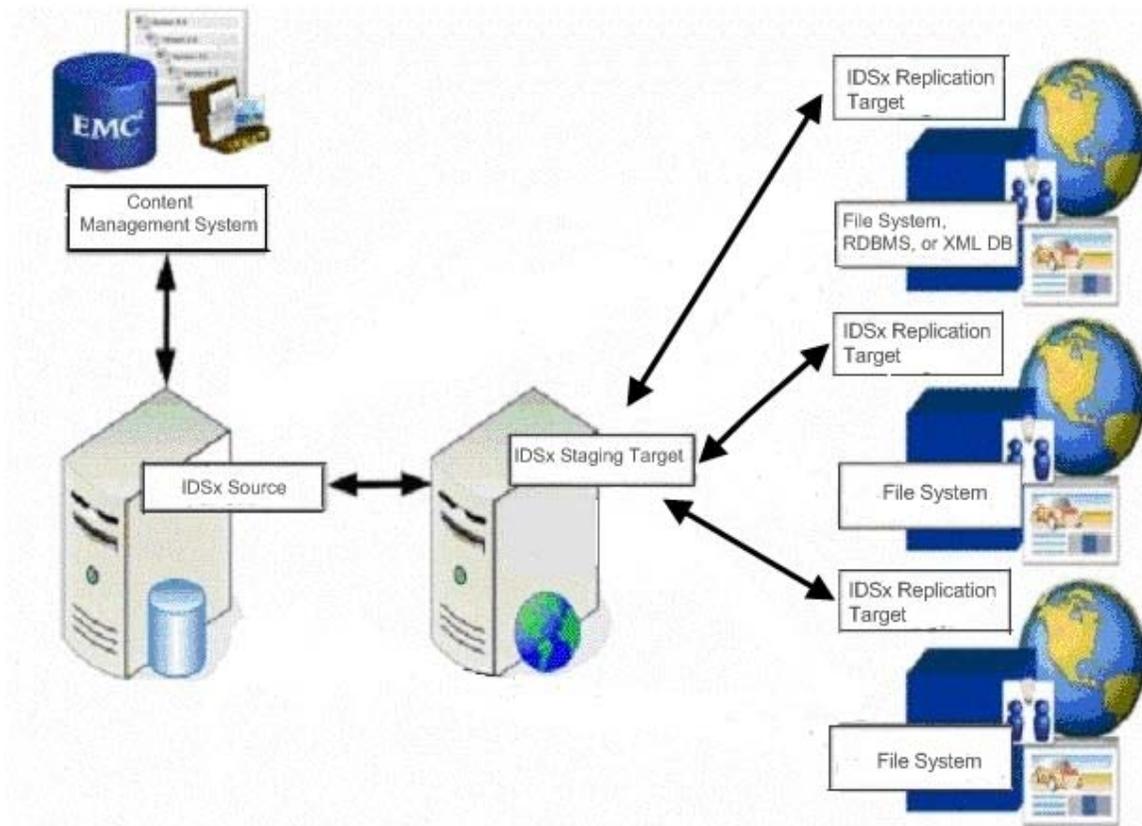
Figure 1. Replication configuration setup in a production environment

This configuration has been implemented by one of the largest financial companies in North America. The customer has deployed IDSx in the production environment with 14 replication targets and transactional capabilities enabled.

The process of installing the IDSx environment includes the following tasks:

1. Installing the IDSx Source on the Content Server machine.

2. Installing the IDSx Target on the Staging Target.

3. Installing the IDSx Target on the Replication Target(s).

Configuration

This section describes the procedure to create a typical IDSx configuration that allows users to start replicating content to multiple target servers .

IDSx configuration is created using Documentum Administrator(DA) for replicating content to multiple servers.

The process of configuring the IDSx environment includes the following tasks:

- Creating the IDSx configuration using Documentum Administrator
- Invoking the replication operation using Documentum Administrator

Creating an IDSx configuration using Documentum Adminstrator

Perform the following steps to create an IDSx configuration using Documentum Administrator:

1. Log in to the repository using the Documentum Administrator client.

2. Navigate to **Administration > Content Delivery > IDS Configurations**.

3. Select **File > New > Content Delivery Configuration** to create a Content Delivery configuration.

4. Enter the IDSx Staging Target information in the **Info** tab of the Content Delivery Configuration dialog box as illustrated in **Figure 2**.

EMC²
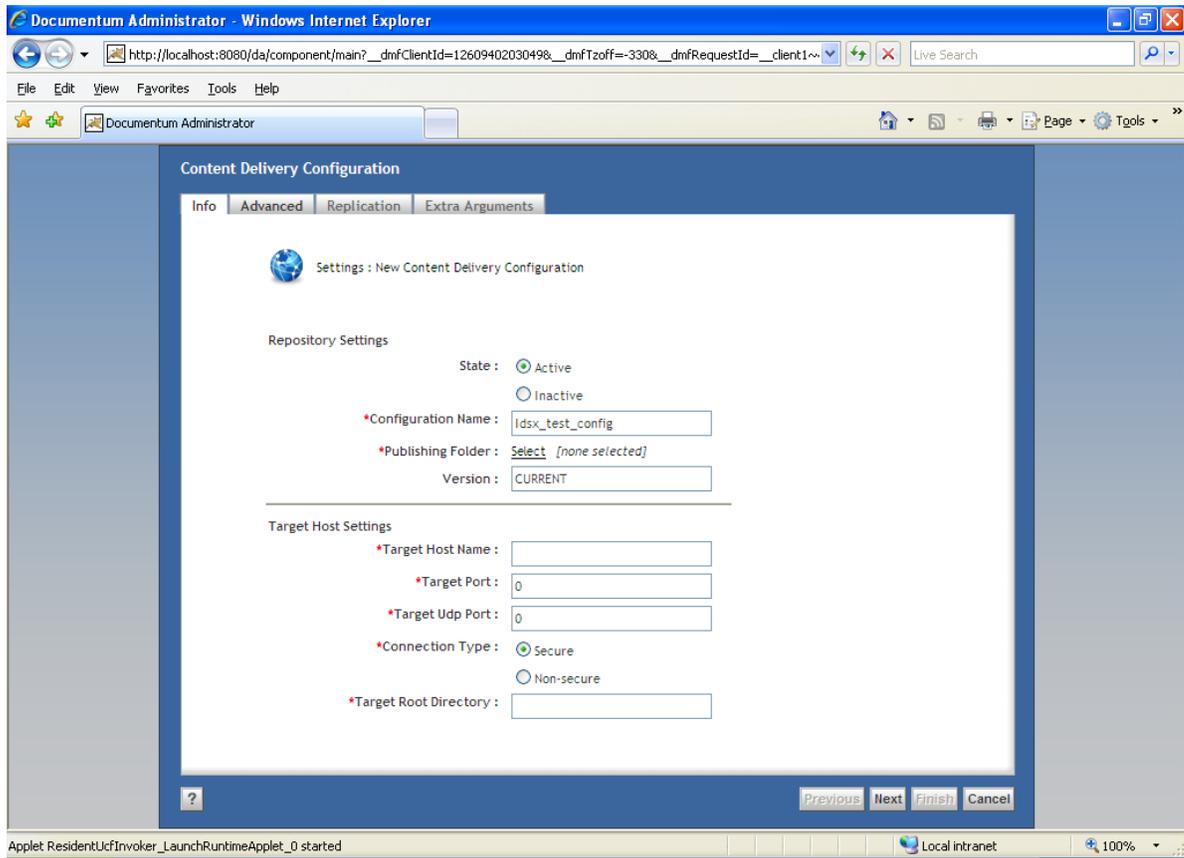
Figure 2. Info tab in the Content Delivery Configuration dialog box

5. Enter additional content delivery configuration information in the **Advanced** tab. You can choose to specify parameters for Export Properties, select Formats, and so on, as illustrated in **Figure 3**.

Figure 3. Advanced tab in the Content Delivery Configuration dialog box

6. Enter details for one or more Replication target(s) in the **Replication** tab as illustrated in **Figure 4.**
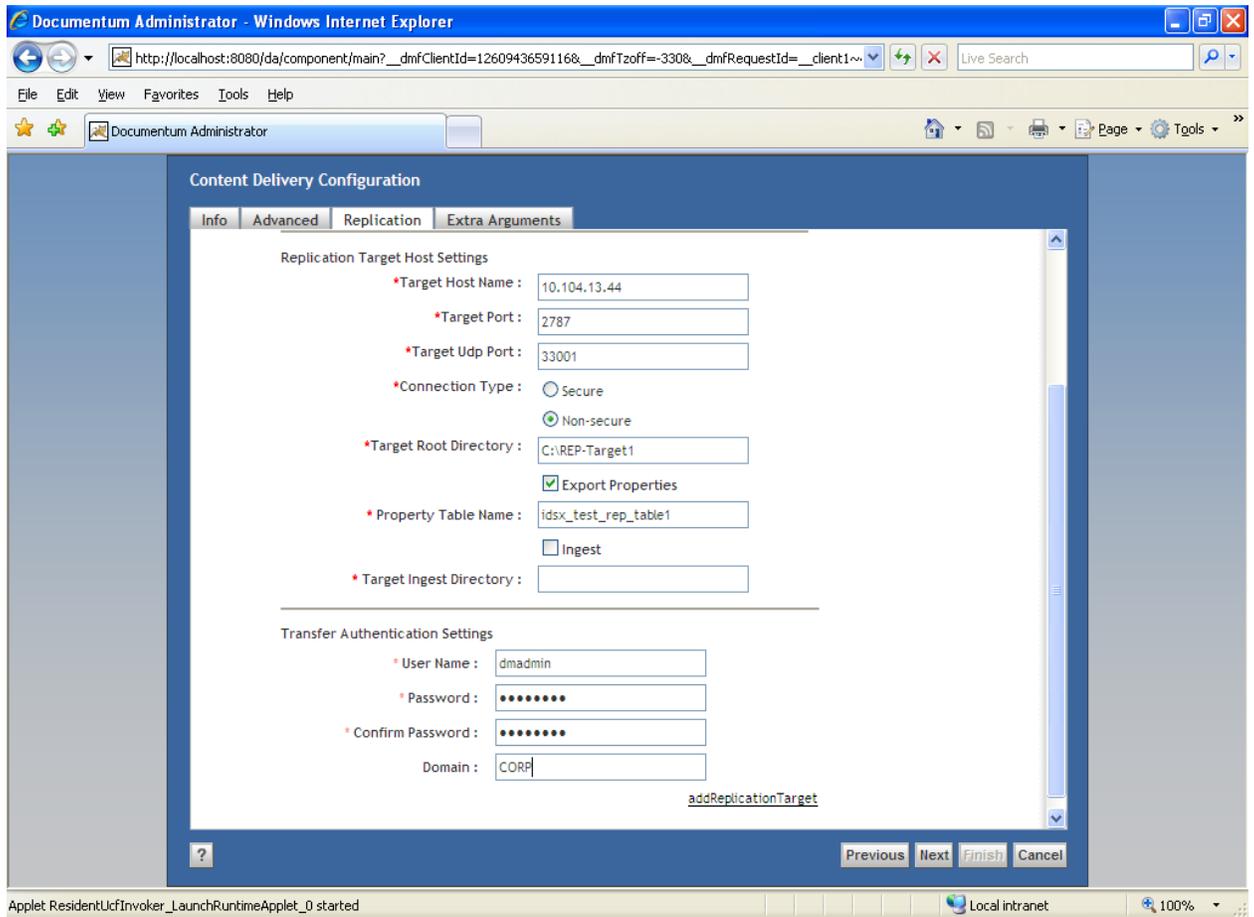
Figure 4. Replication tab in the Content Delivery Configuration dialog box

7. In the **Extra Arguments** tab , you can choose to configure any additional parameters for modifying the replication process as illustrated in **Figure 5**.
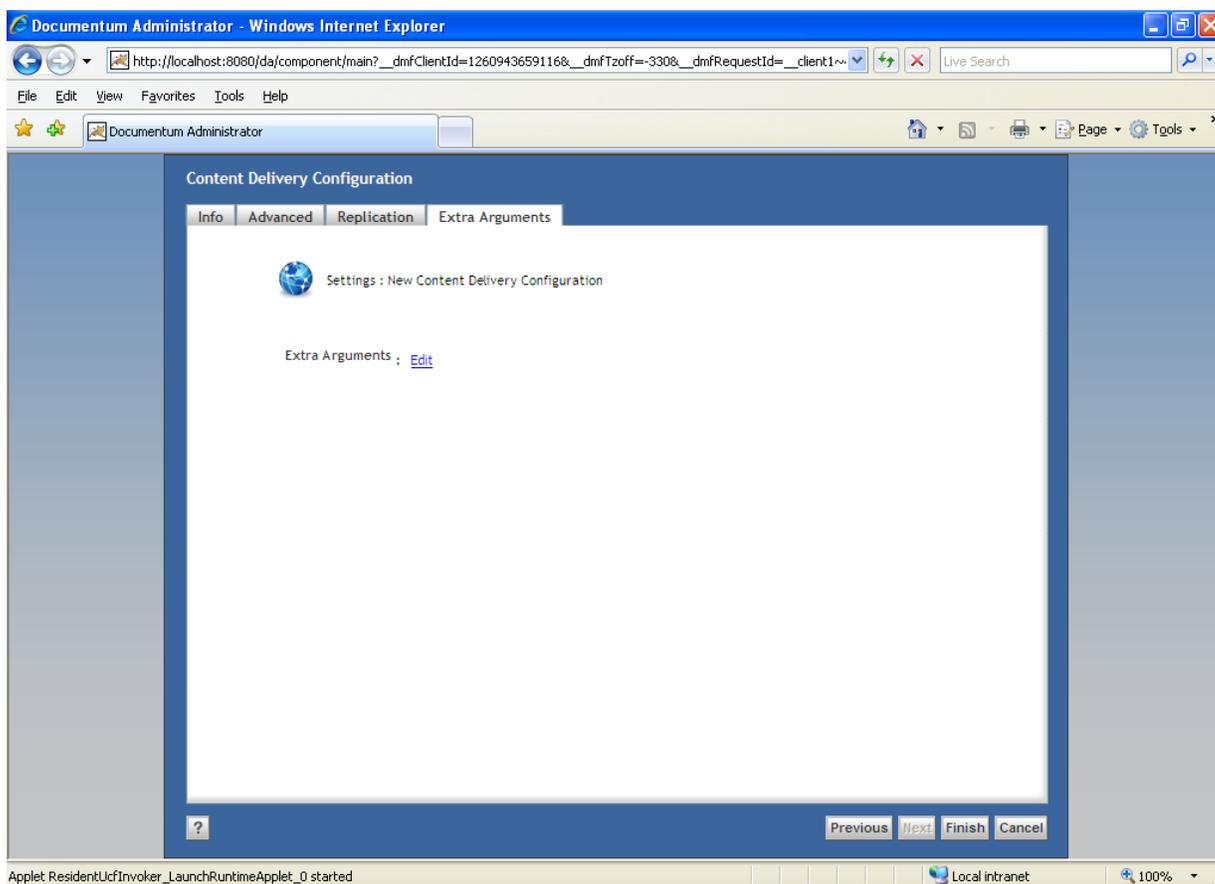
Figure 5. Extra Arguments tab in the Content Delivery Configuration dialog box

Invoking the replication operation using Documentum Administrator

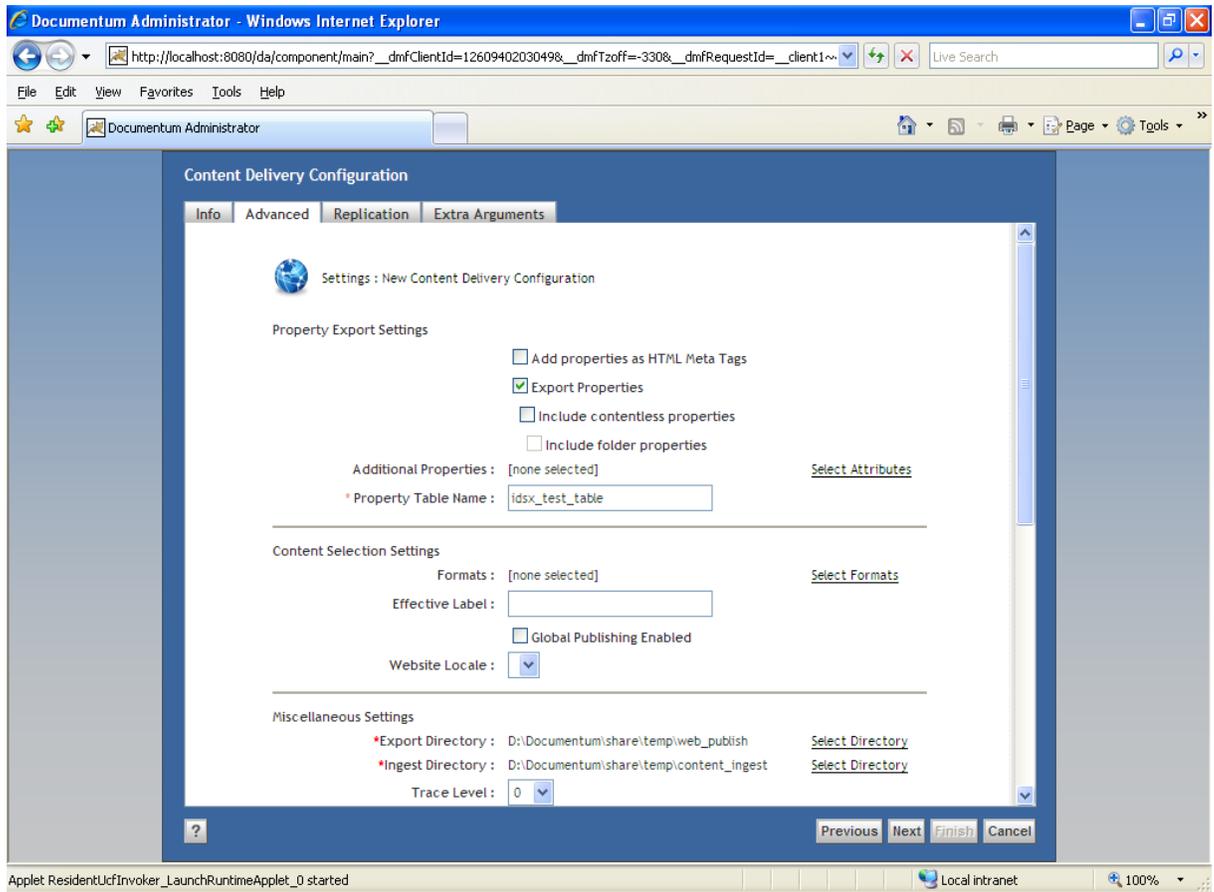1. Select the content delivery configuration and click **Tools > Replicate**. The Content Delivery Configuration Replicate page appears.

2. Select the trace level, and click **OK** to run replication. The final result of the replication operation is listed.

   Note: For more details, click the link in the Content Delivery Configuration Replicate page that provides a pointer to the replication log file.

Usage

This section provides an insight into the parameters used to control the replication operation, and manage the replication operation effectively based on specific requirements.

You can effectively manage the following IDSx replication features using extra arguments:

- Transaction capability

- Auto Replication
- Purging
- Email notification

Transaction capability

Transaction capability is an aspect of replication that ensures all replication targets remain in a consistent state. This capability allows IDSx to commit or rollback data (content and metadata) on all replication targets or none, based on the success or failure of the replication targets.

Setting values for transaction capability

The threshold value of transaction capability is based on the decision_commit_rollback(DCR) value. If the number of successful replication targets is less than the decision_commit_rollback value, a rollback is initiated on the replication targets that were successful. The rollback ensures all replication targets are reverted to the previous state.

The Replication Manager replicates all content and metadata in a transactional manner on each replication target if all the following conditions are true:

1. All replication targets must be in the same state before the instance of this replication. However, there is no state of a replication target at the first replication.

2. Size of the replication batch (tc_file_size) and the number of files to replicate (tc_file_count) must be less than the configured value.

3. The Replication Batch does not contain a full refresh publish batch in the current replication process. The decision_commit_rollback parameter can be set explicitly in the IDSx configuration in the Extra Arguments tab. The default DCR value is the total number of replication targets, and must be a positive integer.
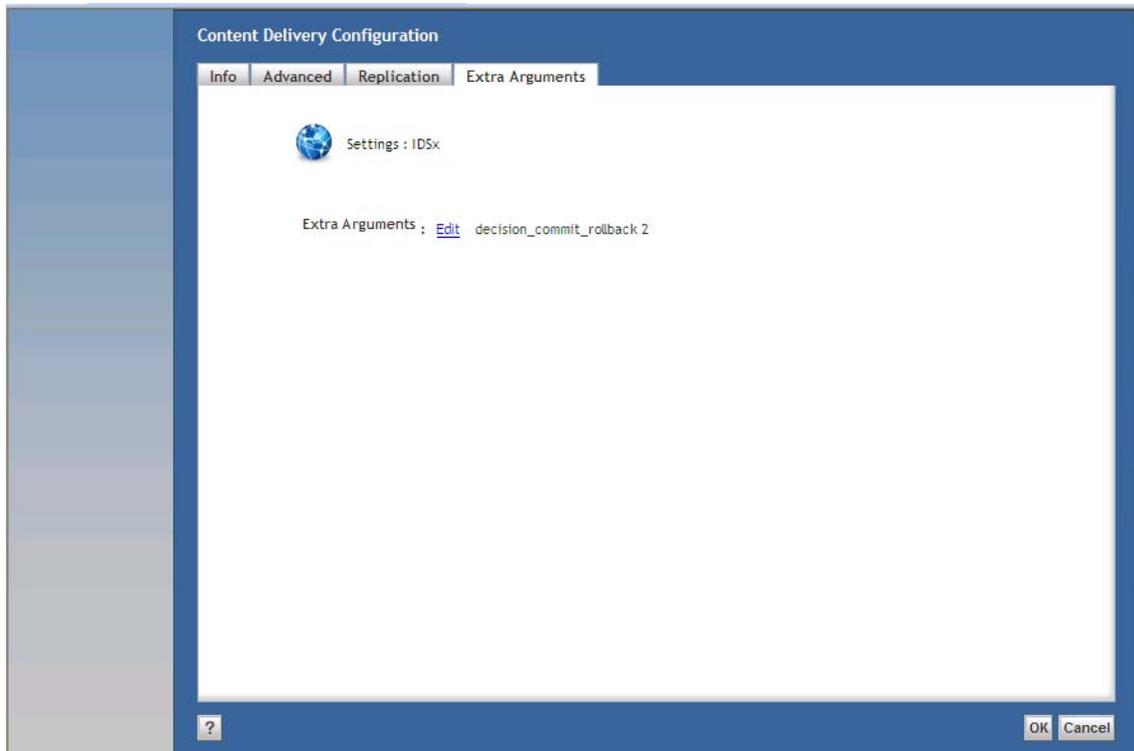
Figure 6. Setting the decision_commit_rollback parameter in the Extra Arguments tab

Auto Replication

Auto Replication is a feature that enables users to replicate content automatically after publishing. Replication can be invoked automatically after a publish operation is performed by setting the value of the auto_replication extra argument to true in Documentum Administrator.
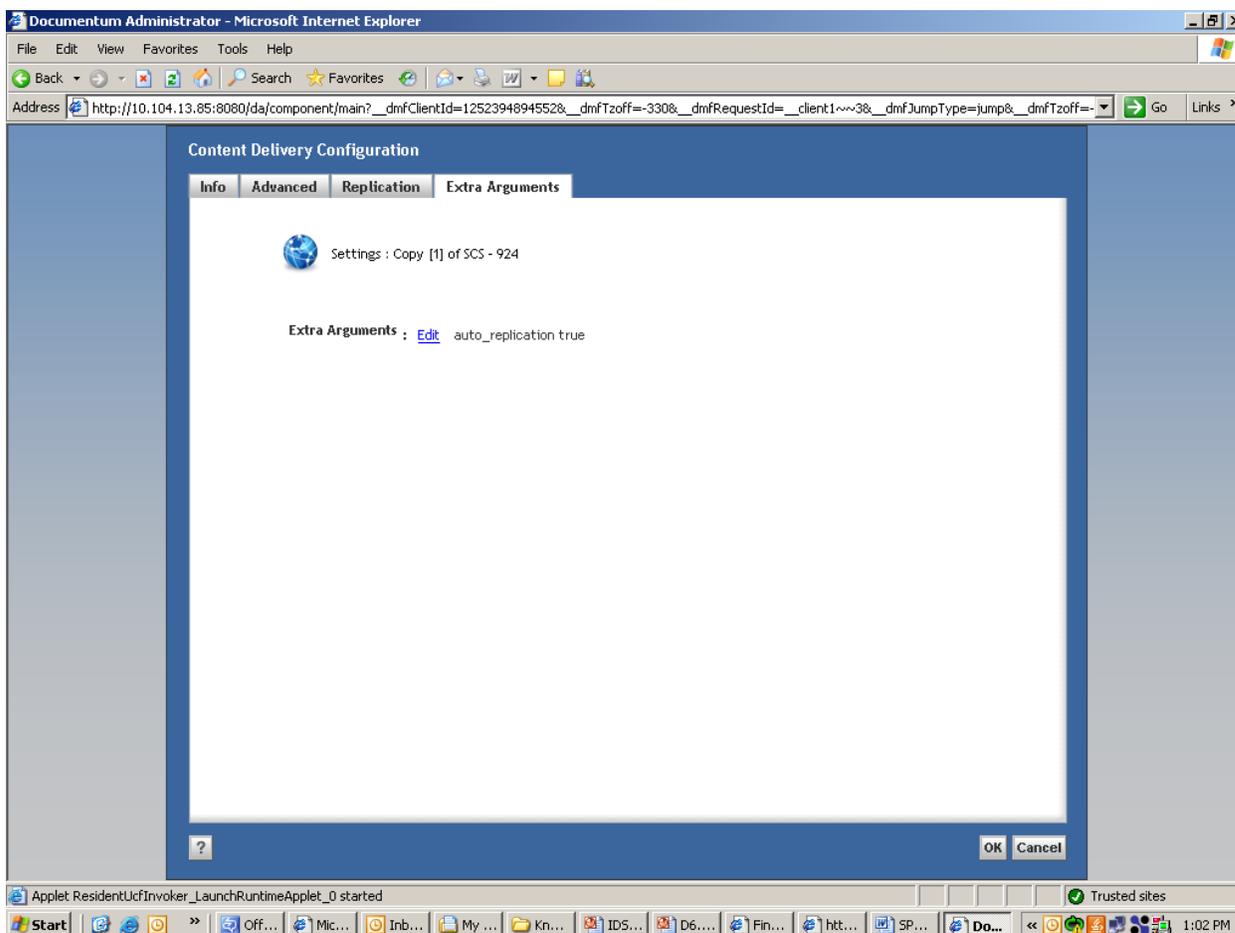
Figure 7. The extra arguments auto_replication value is set to true

The following extract is taken from a sample publish log file that is generated when auto_replication is enabled:

```
May 14 09:44:23.326:S:   INFO:   Webcache publish completed successfully.

May 14 09:44:23.332:S:   INFO:   Invoking Auto Replication

May 14 09:45:16.045:S:   INFO:   Auto Replication log file name : rep
rajiv_config1  20090514  09.43.12

May 14 09:45:16.045:S:   INFO:   Auto Replication completed successfully.
```

Purging

The Purge parameter allows you to clean up unwanted publish and replication batch directories from the IDSx staging target.

The purge operation is performed by setting the value of the, purge_publish_and_replication_batch parameter to true in the agent.ini file at the following location on the staging target:

IDSx Staging Target/admin/config/‹port_number›/
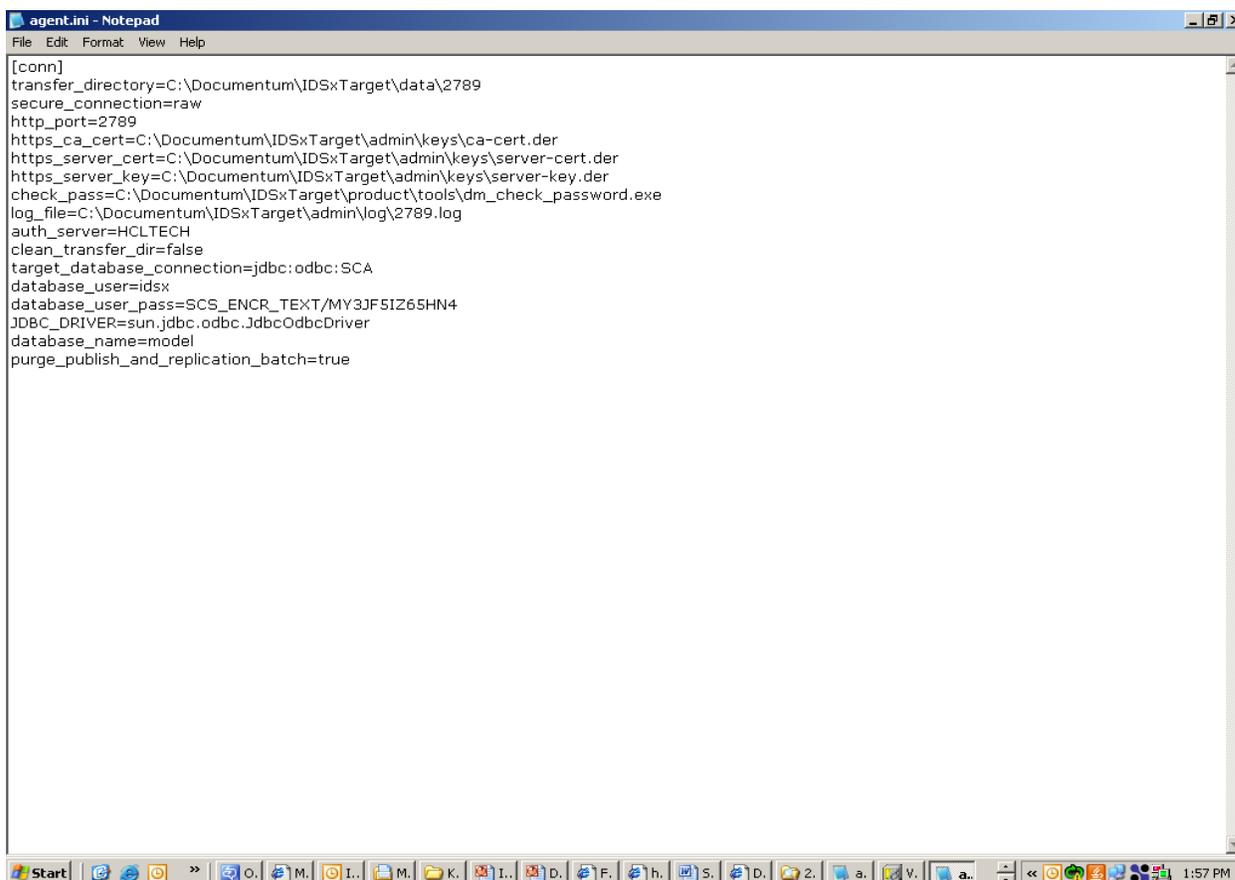
```
agent.ini - Notepad

File  Edit  Format  View  Help

[conn]
transfer_directory=C:\Documentum\IDSxTarget\data\2789
secure_connection=raw
http_port=2789
https_ca_cert=C:\Documentum\IDSxTarget\admin\keys\ca-cert.der
https_server_cert=C:\Documentum\IDSxTarget\admin\keys\server-cert.der
https_server_key=C:\Documentum\IDSxTarget\admin\keys\server-key.der
check_pass=C:\Documentum\IDSxTarget\product\tools\dm_check_password.exe
log_file=C:\Documentum\IDSxTarget\admin\log\2789.log
auth_server=HCLTECH
clean_transfer_dir=false
target_database_connection=jdbc:odbc:SCA
database_user=idsx
database_user_pass=SCS_ENCR_TEXT/MY3JF5IZ65HN4
JDBC_DRIVER=sun.jdbc.odbc.JdbcOdbcDriver
database_name=model
purge_publish_and_replication_batch=true
```

Figure 8. The content of the agent.ini file

Email notification

You can examine the IDSx log files to determine whether a publish or replication operation succeeded or failed. Alternatively, you can enable email notification for the status of each publish or replication operation. This notification is set as an extra argument in the IDSx configuration.

The mail_notification_on and mail_notification_user parameters are related to email notification.

Replication status

The following table categorizes replication results for email notifications based on the value of the mail_notification_on parameter. Specify the repository-username-value as mail_notification_user = repositoryusername if you want to receive email notifications by default. Otherwise, the emails are sent to the system user name, usually the server install owner.

Value	Description	Replication Status	Comments
Success	Email notifications are sent on successful, warning, and failed publishing operations	REPLICATION_COMPLETE	On successful replication operation
Warning	Email notifications are sent on warning and failed publishing operations	REPLICATION_ALREADY_SYNC	When a replication is run against a target that is already in sync
		REPLICATION_NO_TRANSFER_SYNC_ COMPLETE	Indicates that there is no new content that requires to be syndicated or transferred to a replication target. This status also implies that the synchronize operation is successful while trying to synchronize the contents transferred during the previous replication run.
		REPLICATION_COMPLETE_POSTSCRIP T_NOTDONE	Indicates that replication is successful but there is a postscript failure on the replication target.
Failure	Email notifications are sent only on failed publish operations	REPLICATION_FAIL_TRANSFER	Indicates failure while transferring contents from the staging target to the replication targets.
		REPLICATION_FAIL_SYNC	Indicates failure while synchronizing the files or the database content for a replication target.
		REPLICATION_NO_TRANSFER_SYNC_ FAIL	Indicates that there is no new content that requires to be syndicated or transferred to a replication target. This status implies that no content was transferred from the staging target to the data directory of the replication target. This status also implies that a failure occurred while trying to synchronize the content transferred during the previous replication run.

Troubleshooting

This section describes troubleshooting scenarios that pertain to interpreting replication operation results from log files, and explains the behavior of the replication operation when exception scenarios occur.

Interpreting replication results from log files

The log file enables you to understand the completion of the replication process on each replication target. The status of all replication targets is displayed at the end of the log file as follows:

```
Dec 11 12:35:10.476:  10.104.13.16:T:Thread-17: INFO:---------------
Replication details ---------------------------:

Dec 11 12:35:10.476:  10.104.13.16:T:Thread-17: INFO: ---- 10.104.13.85
: 2500  replication details -----:

Dec 11 12:35:10.476:  10.104.13.16:T:Thread-17: INFO: Last Publish Batch
Directory for 10.104.13.85 : 2500 Replication Target:
webcache_dir_80006034_16

Dec 11 12:35:10.476:  10.104.13.16:T:Thread-17: INFO: Transfer of
10.104.13.85 : 2500 Replication Target : true

Dec 11 12:35:10.476:  10.104.13.16:T:Thread-17: INFO: Synchronization of
10.104.13.85 : 2500 Replication Target: true

Dec 11 12:35:10.476:  10.104.13.16:T:Thread-17: INFO: Replication Status
REPLICATION_COMPLETE

Dec 11 12:35:10.476:  10.104.13.16:T:Thread-17: INFO:-----------------
--------------------------------------------------

Dec 11 12:35:10.476:  10.104.13.16:T:Thread-17: INFO:---- 10.104.13.86 :
2020  replication details -----:

Dec 11 12:35:10.476:  10.104.13.16:T:Thread-17: INFO: Last Publish Batch
Directory for 10.104.13.86 : 2020 Replication Target:

Dec 11 12:35:10.476:  10.104.13.16:T:Thread-17: INFO: Transfer of
10.104.13.86 : 2020 Replication Target : false

Dec 11 12:35:10.476:  10.104.13.16:T:Thread-17: INFO: Synchronization of
10.104.13.86 : 2020 Replication Target: false

Dec 11 12:35:10.476:  10.104.13.16:T:Thread-17: INFO: Replication Status
REPLICATION_FAIL_TRANSFER

Dec 11 12:35:10.476:  10.104.13.16:T:Thread-17: INFO:-----------------
--------------------------------------------------
```

Understanding exception scenarios

This section describes the behavior of the replication operation in the following exception scenarios:

When the Post Sync Script execution fails

Scenario 1: Transaction Capability Enabled

If the Post Sync Script execution fails when the transaction capability feature is enabled and while replication is in progress, the replication status is REPLICATION_COMPLETE_POSTSCRIPT_NOTDONE and Documentum Administrator treats this scenario as successful replication. However, a message indicating that a post sync script failure has occurred, is displayed.

Since data is committed on all replication targets, the post sync script must not be run again during a re-run of the replication operation. The user must run the correct post sync script manually, and re-run the post sync script on the replication target where the script failed.

Scenario 2: Transaction Capability Disabled

When a post sync script failure occurs during replication, the behavior is similar to the Publish operation. A synchronization failure exception occurs during a publish operation. On the next publish run, the publish operation starts synchronizing the last failed data set.

There are two scenarios in the context of a Post Sync Script failure because of the limited control available on data (content and metadata) that is committed in a transactional manner, to the file system and the database.

When the Pre Sync Script execution fails

If a pre sync script failure occurs during the replication operation the Replication Manager considers the scenario as "Transfer Complete, Sync Failure". If email notification is enabled a message indicating that a pre sync script failure has occurred. The replication status is set as REPLICATION_FAIL_SYNC and Documentum Administrator considers this event as a replication failure.

Note: When a pre sync failure occurs, Documentum Administrator (DA) does not warn the user with a Pre Sync Script failure message. Instead, DA considers this failure as a general failure. No separate status is indicated in the replication log file for such failure, because it is considered as a synchronization failure. The user must review the replication log file manually to determine the cause of the problem and take corrective action.

When the user has corrected the pre sync script failure, the replication operation can be re-run. If the replication operation is successful, the replication status is set to "No Transfer Sync Completed".

Note:
- It is not mandatory for the user to run any pre sync operation manually, as compared to the post sync operation. The Replication Manager runs the pre sync operation automatically, during a subsequent pre sync run.
- The user must ensure that the cause of the pre sync failure is corrected.

Conclusion

This paper has discussed in detail, the process of setting up Interactive Delivery Services Accelerated (IDSx), and provides insight to users on leveraging the replication capability to replicate content from the central Documentum repository to multiple replication targets.

References

For more information about Documentum Interactive Delivery Services Accelerated, see the following documents on Powerlink:

- EMC Documentum Interactive Delivery Services Accelerated 6.5 SP1 User Guide
- EMC Documentum Interactive Delivery Services Accelerated 6.5 SP1 Installation Guide
- EMC Documentum Interactive Delivery Services Accelerated 6.5 SP2 Release Notes