



Kaltura Drop Folder Service for Content Ingestion

September, 2011



Contents

1. What Is This Document About?.....	3
2. Who Should Read This Document?	3
3. What is the Kaltura Drop Folder Service?	3
4. When Should You Use the Kaltura Drop Folder Service?	3
5. Drop Folder Service: Main Components	4
6. Drop Folder Location and File Transfer Options	5
7. Drop Folder Configuration Options	7
8. Drop Folder Control Panel.....	12

1. What Is This Document About?

This document describes the Kaltura Drop Folder service, including usage, deployment options, management and data flows, and the different workflow options it supports.

2. Who Should Read This Document?

This document is intended for Kaltura partners, community members, customers and prospective Kaltura users who want to become familiar with the Kaltura Drop Folder service. This document assumes you have a technical background.

3. What is the Kaltura Drop Folder Service?

The Kaltura Drop Folder service enables a publisher to automate the ingestion of content into the Kaltura platform. The service constantly watches each drop folder for new content, and activates automatic ingestion of new content to the publisher account.

Kaltura offers multiple configuration options for setting each drop folder to a specific workflow, from simple media upload tasks to bulk ingestion of complex media packages that include multiple media files and metadata items.

The Kaltura technical support and engineering teams set and enable the Kaltura Drop Folder service for a specific publisher account. Please contact Kaltura's sales or technical representative for more information on setting the service for your account.

4. When Should You Use the Kaltura Drop Folder Service?

Use the Kaltura Drop Folder service to:

1. Automate or partially automate an on-going content ingestion workflow.
2. Support bulk ingestion of content to Kaltura, either when you set up an account for migrating existing content or on an on-going basis.
3. Work with content providers and operational teams who do not have access to the Kaltura Management Console (KMC) or any other content management application.

5. Drop Folder Service: Main Components

The Kaltura Drop Folder service has the following main components:

1. **Drop Folder Configuration** – For each drop folder in the system, a configuration profile is assigned to the relevant publisher account. Kaltura drop folder configuration profiles store a combination of workflow, policies, location and access information. A drop folder configuration profile controls the way each of the publisher's drop folders is handled.
2. **Drop Folder Watchers** – A drop folder watcher watches for new content in a publisher's drop folders. A drop folder watcher is a scheduled batch job that runs on the Kaltura data cloud. There are multiple instances of drop folder watchers. Drop folder watchers monitor all drop folders configured in the Kaltura system and identify new content that is ready to be processed. The number of drop folder watchers and their activation frequency is configured in the Kaltura cloud. Kaltura adjusts the number and activation frequency to maintain timely and efficient service. By default, only few minutes pass after a file is completely uploaded to the drop folder until the file is processed.
3. **Drop Folder Handlers** – A drop folder handler processes new content that is available and ready for ingestion. A drop folder handler is a scheduled batch job that runs on the Kaltura data cloud. There are multiple instances of drop folder handlers. Drop folder handlers manage all of the steps needed to complete content ingestion according to the workflow and policy configured for each drop folder.
4. **Drop Folder File Management** – The drop folder file management system maintains details and status for files that are added to the drop folder.

6. Drop Folder Location and File Transfer Options

Kaltura supports the following options for the physical location of a drop folder:

1. Hosted on the Kaltura cloud
2. Hosted at the publisher's premises

6.1 Drop Folders Hosted on the Kaltura Cloud

With this option, account-specific drop folders are located on the Kaltura cloud and are fully hosted and managed by Kaltura. Publishers are granted FTP/SFTP access to one or more account-specific drop folders. Publishers are responsible for uploading their content to the drop folders using any standard or advanced FTP client. Kaltura recommends this option for publishers who prefer a fully-managed drop folder solution, and for publishers whose workflow and security needs are better met by this option.

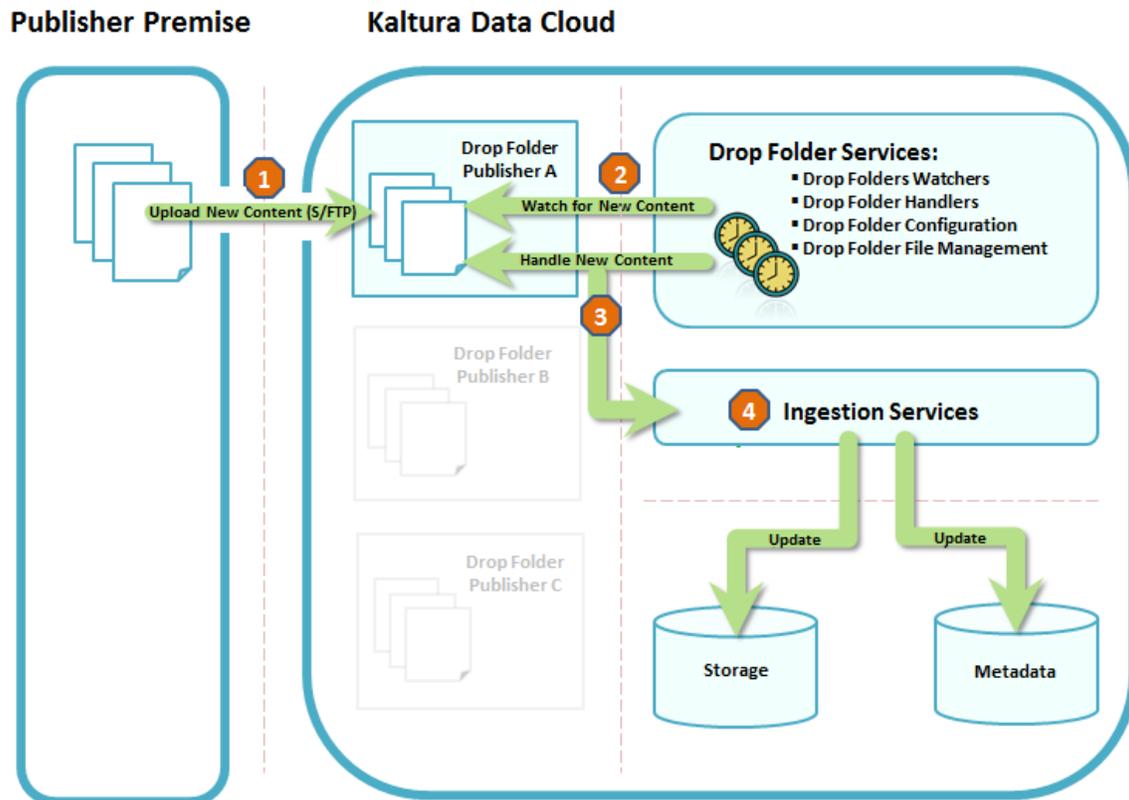


Figure 1 - Drop Folder Flow – Hosted on the Kaltura Cloud

6.2 Drop Folders Hosted at the Publisher's Premises

With this option, the publisher locates account-specific drop folders at the publisher's premises or on any FTP/SFTP -accessed location. Kaltura defines a drop folder configuration with the drop folder location and access details. A Kaltura drop folder watcher constantly checks the drop folder. When content is ready to be ingested, it is downloaded from the drop folder to the Kaltura cloud. Kaltura recommends this option for publishers who already have a network location that they want Kaltura to watch and use for ingestion, and for publishers whose workflow and security needs are better met by this option.

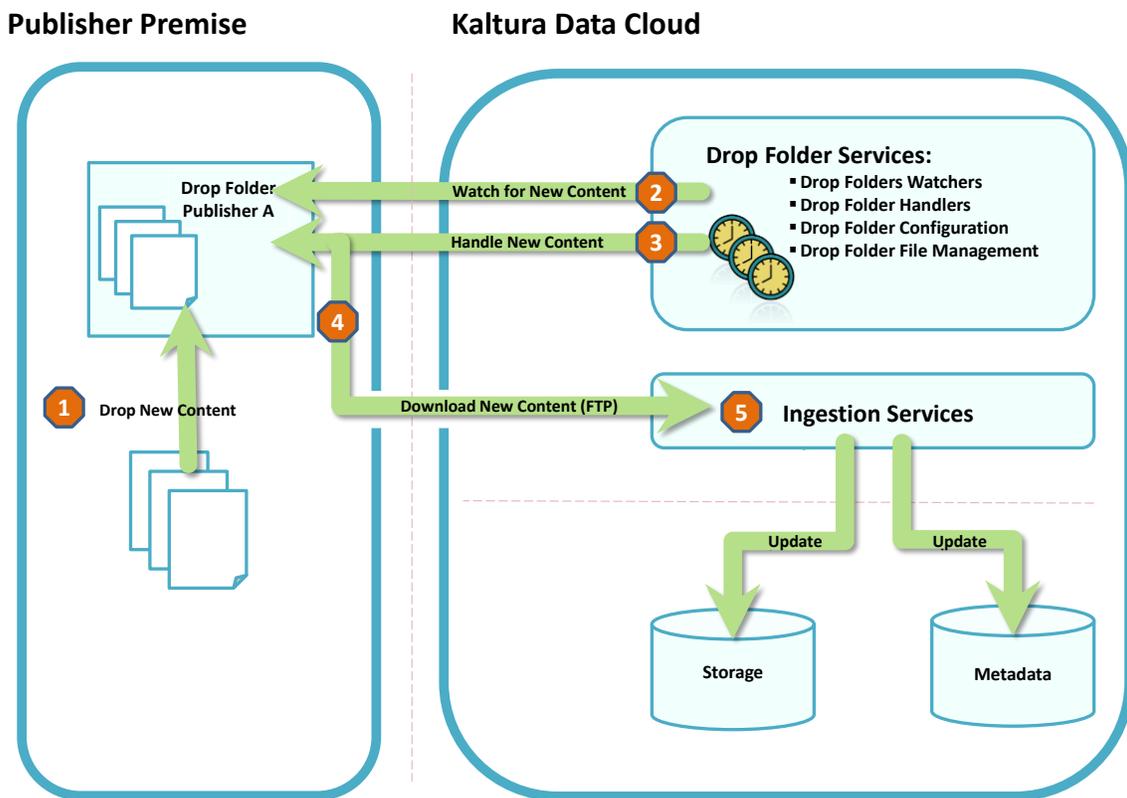


Figure 2 - Drop Folder Flow – Hosted at the Publisher Premises

7. Drop Folder Configuration Options

7.1 Transcoding Profile

The transcoding profile attached to the drop folder settings defines the transcoding settings required for content ingested from the drop folder. Consequently, the transcoding profile also determines whether media files in the drop folders are handled as a *single file media source* or as part of a *media package*. Kaltura fully transcodes a single file media source, while a media package consists of multiple bit rate transcoding flavors that are generated by a local transcoder. A media package is added from the drop folder to a single media entry as one unit.

7.2 File Name Patterns

By default, drop folders handle all file name patterns.

7.2.1 File name patterns to handle

A publisher can define a specific drop folder file name pattern to be handled by the drop folder (white list approach). The file name pattern should be configured when you need to apply a drop folder operation only to files of a specific type or naming structure.

7.2.2 File name patterns to ignore

A publisher can define a specific drop folder file name pattern to be ignored by the drop folder (black list approach). The file name pattern should be configured when you need a drop folder operation to ignore files of a specific type or naming structure. This option is useful when transferring *temporary* files created by transfer client applications when uploading files to the drop folder. The temporary files usually use a specific name pattern or extension. Such a specific name pattern or extension can be marked to be completely ignored by the drop folder service.

7.3 Ingestion Type

Content ingestion from a drop folder can be triggered from either of the following:

- Media files
- A Kaltura MRSS-formatted XML that includes both metadata and references to related media files

7.3.1 Ingestion from media files

With this ingestion type, only media files are dropped to the drop folder. An automatic ingestion process is triggered for uploading these files to Kaltura. You can conveniently edit the metadata in the KMC. This ingestion type requires setting additional configuration options for specific policies in the drop folder profile. The additional configuration options are described in the Ingestion from Media Files sections below.

7.3.2 Ingestion from XML

With this ingestion type, a Kaltura MRSS-formatted XML is dropped into the drop folder. The XML may include metadata information for multiple media items and references to media files related to each item. The referenced media files can reside within the drop folder itself or on a remote location. An automatic ingestion process is triggered that creates or updates media entries in Kaltura while populating the metadata fields of the entries from the XML. The ingestion process also automatically uploads the media files referenced in the XML and adds the media files to each related entry.

Note: When a transcoding profile is defined in the XML, it will be used for determining the content ingestion transcoding options for the job. In this case, the transcoding profile attached to the drop folder settings will be ignored.

7.4 Ingestion from Media Files - Matching Policies

When a drop folder is set to ingest content from media files, one of the following matching policies should be selected:

7.4.1 Policy 1 - Create an entry from media

With this drop folder policy, media files in the drop folder always trigger the creation of a new entry in Kaltura.

Use this policy when you need to automatically upload media files into Kaltura, and when metadata editing can/should be completed after files are uploaded to the system.

The typical workflow for this ingestion policy is:

1. Drop a media file into the drop folder.
2. A new entry in Kaltura is automatically created. This triggers an automatic conversion in Kaltura according to the transcoding profile settings attached to the drop folder. The title of the created entry is the media file name.
3. Edit the new entry metadata in the KMC as needed.

A process to automatically replace the media cannot be triggered from a drop folder configured to support this policy.

7.4.2 Policy 2 - Match media to an existing entry

With this policy, media files in the drop folder are automatically matched and added to an existing media entry in Kaltura. Automatic matching requires a *Reference ID*. Include the Reference ID in the media file name. You also must set the same Reference ID as a metadata item of a media entry in Kaltura.

Use this policy for fast-paced or complex operational environments that require parallel work on metadata and media.

The typical workflow for this ingestion policy is:

1. Prepare a new entry in the KMC.
2. Set metadata for the new entry in the KMC and set the **Reference ID** item in the entry metadata.
Reference ID example: *MyVideoClip123*
3. Drop a media file into the drop folder and include the **Reference ID** within the file name.
Media file name example: *MyVideoClip123.mp4*
4. *The MyVideoClip123.mp4* file in the drop folder is automatically added to the entry that was set with the string *MyVideoClip123* as a **Reference ID** metadata item. This triggers an automatic conversion in Kaltura according to the transcoding profile settings attached to the drop folder. When no entry exists in the system with a matching **Reference ID**, the media file is left in the drop folder until a matching entry is created.

When the matched entry already is already associated with media, a process is triggered that automatically replaces the media.

7.4.3 Policy 3 - Match media when entry exists or create a new entry

With this drop folder configuration, the drop folder service tries to match the media file in the drop folder to an existing entry based on a **Reference ID**, as described in [Policy 2](#). If no entry exists in the system with a matching Reference ID, a new entry is created in Kaltura. The media file in the drop folder is added to the new entry.

A process to automatically replace the media can be triggered from a drop folder configured to support this policy.

Use this policy when:

- You need to automatically upload media files into Kaltura.
- Metadata editing can/should be completed after files are uploaded to the system.
- You need to trigger a process from a drop folder to automatically replace media.

Note: (Applies to all policies) When an account is configured to support the use of a local transcoder, multiple bit rate media files also can be automatically added to the entry in the same manner when the files are handled *together* as a single media package. In this case, uploading the media files and adding them to the entry are triggered only when all media files that are part of the media package are available in the drop folder.

7.5 Ingestion from Media Files - Matching Regular Expression

When a drop folder is set to match media files to an existing media entry, the file name must include the **Reference ID** that is used as a matching key. By default, the file name itself is treated as a **Reference ID**.

A more specific naming convention can be defined as a textual regular expression that instructs the drop folder service how to extract the **Reference ID** from the file names.

Examples of naming conventions that can be defined by a regular expression (the **Reference ID** is red):

MyVideoClip123.mp4 (Default configuration)

01012011_**MyVideoClip123**.mp4

01012011_**MyVideoClip123**_provider-X.mp4

When an account is configured to support the use of a local transcoder and ingestion of multiple bit rate media files to a single media entry, file naming convention must be defined to also include the system name of a specific flavor in the transcoding profile that this file should be matched to.

In the following examples, all media files are added to a single entry and each file is matched to the appropriate transcoding flavor, according to its *system_name* definition in the transcoding profile (the **Reference ID** is red, the transcoding flavor system name is blue):

01012011_**MyVideoClip123**_HD.flv

01012011_**MyVideoClip123**_Basic-Small.flv

01012011_**MyVideoClip123**_Basic-Large.flv

7.6 Drop Folder File Management Policies

7.6.1 File Size Check Frequency

This policy defines how frequently a media file in the drop folder is checked to determine whether the file upload to the drop folder is complete. When the size of the media file stops growing, it is assumed that file upload is complete. Therefore, network conditions related to drop folder location should be taken into account when determining how frequently file size is checked. (The recommended value is 30 seconds.)

7.6.2 File Deletion Policy

This policy defines whether files in drop folders are deleted automatically by the Kaltura drop folder service or are deleted manually by the publisher.

When you select an **automatic deletion policy**, you must also define the number of days that the files are kept in the drop folder before being deleted. (The recommended value is 3 days.) This policy deletes only files that were processed successfully.

You can perform **manual deletion** from the drop folders control page in the KMC or by using the Kaltura API. Kaltura does not recommend direct manual deletion of files from the drop folder itself. This is because manual file deletion may impact the synchronization between the status of drop folder files and their representation, which is managed by Kaltura's drop folder file management service.

With both options, a change to an existing file in the drop folder (such as a change in the file creation date or file size) resets handling of the file. The drop folder service then treats the changed file as a new file.

7.7 Drop Folder Access Settings

When a drop folder is managed by Kaltura and is hosted on the Kaltura cloud, Kaltura's technical support team sets the drop folder location and provides full access and authentication information to the publisher.

When a drop folder is hosted and managed at the publisher's premises, the publisher must provide the following information, per preferred transfer protocol, to Kaltura in order to set up the drop folder configuration:

FTP:

Access Details: Host, Port, Storage path

Authentication Details: Username, Password

SFTP:

Access Details: Host, Port, Storage path

Authentication Details (according to the SSH authentication used): Username, Password, SSH private key file, SSH public key file, SSH pass phrase

8. Drop Folder Control Panel

Note: This section applies only after the Kaltura Drop Folder service is set up. For more information about the KMC, refer to the *KMC Quick Start Guide*.

In the KMC Content tab, publisher administrators can use a control panel to monitor and troubleshoot media files in drop folders. Access to the Drop Folders control panel can be granted to specific administrators through the administrator's user role definition.

The screenshot shows the Kaltura KMC interface. At the top, there is a navigation bar with tabs: Dashboard, Content, Studio, Analytics, Settings, Administration, and Upload. Below this is a sub-navigation bar with 'Manage', 'Moderate', 'Playlists', and 'Syndicate'. On the right side of this bar, there are links for 'Upload Control' and 'Drop Folders' (the latter is highlighted with a red box and an arrow). Below the navigation is a 'Filters' sidebar on the left, which includes a search box, a 'Drop Folders' filter, and a list of status options (All Statuses, Uploading, Pending, Matched, Not Matched, Done, Error, Delete Failed, Download Failed). The main content area is titled 'Drop Folders - Media Files' and contains a table with the following data:

File Name	Folder Name	Created On	File Size	Status	Error Description
FlvExample_Source.flv	dropfolder config1	05/19/2011 11:27	1.26 MB	Error	Failed to update an existing entry
titanicin5_seconds.flv	dropfolder config1	05/19/2011 14:14	0.48 MB	Error	Transcoding Flavor system name: seconds not fo...
media-ingestion.xml		05/22/2011 13:27	0.00 MB	Error	File name media-ingestion.xml does not comply wl...
nice_Flv.flv	dropfolder config2(auto delete)	05/22/2011 14:13	1.20 MB	Error	File name nice_Flv.flv does not comply with namin...
4_b_100909_cat_ra04.flv	dropfolder config2(auto delete)	05/22/2011 14:20	3.49 MB	Error	Transcoding Flavor system name: ra04 not found ...
4_b_100002_cat_Source.flv	dropfolder config2(auto delete)	05/22/2011 15:03	3.49 MB	Error	Transcoding Flavor system name: Source not fou...
4_b_100002_titanicin5_Source.flv	dropfolder config2(auto delete)	05/22/2011 15:03	0.48 MB	Error	Transcoding Flavor system name: Source not fou...
4_b_100001_titanicin5_BasicSmall.flv	dropfolder config2(auto delete)	05/22/2011 15:06	0.48 MB	Error	
4_b_100002_titanicin5_BasicSmall.flv	dropfolder config2(auto delete)	05/22/2011 15:08	0.48 MB	Error	Failed to create a new entry
nice_Flv3.flv	dropfolder config3	05/23/2011 14:45	1.20 MB	Error	Transcoding Flavor system name: Flv3 not found L...

At the bottom of the table, there is a pagination control showing '1 2 3 4 5 Next' and a 'Show rows: 50' dropdown. A 'Delete' button is located at the bottom right of the table area, highlighted with a red box and an arrow.

The Drop Folders control panel includes the status of each file in the drop folder, and an option to filter the file display for a specific drop folder, status and creation date range.

An additional option to manually delete files from the drop folder can be granted separately only to authorized administrators.