Reach Data privacy and Content deletion

This article focuses on the Data privacy and processing aspects when integrating with the Kaltura platform and with the Kaltura SaaS regions.

Introduction

Vendors who integrate with Kaltura Reach can be processing data from Kaltura customers in various regions of the world where Data privacy and content deletion policies need to be respected.

To enable this, Kaltura offers a number of vehicles which Reach vendors can choose to leverage. This article aims to help Reach vendors build their integrations accordingly.

There are two main vehicles to facilitate data download, storage and processing:

1. Processing region in Reach jobs
2. SaaS Regions API endpoint.

In this article we will describe two different models which leverage this information differently.

And we will describe how to handle content deletion.

Data privacy

Processing region in Reach jobs

Reach jobs and tasks always include a `vendorTaskProcessingRegion` parameter value which comes from the customer's Reach profile used to submit an order.

The value of `vendorTaskProcessingRegion` informs the Reach vendor on which region should the task processing take place.

Currently `vendorTaskProcessingRegion` supports three settings, see the following documentation: KalturaVendorTaskProcessingRegion - Kaltura VPaaS API Documentation

At the time of writing of this document, we support the following region settings:

Name	Value
US	1
EU	2
CA	3

It is important to note that this value will be submitted regardless of which PID or SaaS region a Reach order is submitted from.

SaaS Regions

During 2023 Kaltura rolled out what are called SaaS regions, a SaaS region is a set of cloud computing resources located in a specific geographical area, designed to meet the needs of that area in terms of performance, regulatory compliance, and other considerations.

At the time of writing this document, the available SaaS regions are the following:

SaaS Region	Region	SaaS Region API
NVP	United States	https://www.kaltura.com/api_v3/
IRP2	Europe (Ireland)	https://api.irp2.ovp.kaltura.com/api_v3/ or https://api.eu.kaltura.com/api_v3/
FRP2	Europe (Frankfurt)	https://api.frp2.ovp.kaltura.com/api_v3/ or https://api.de.kaltura.com/api_v3/

⚠ Note that each SaaS Region has a unique API endpoint. That means that in order to serve customers who chose to run their environments on a specific SaaS region, it is necessary that Reach vendors integrate with the API within those SaaS regions both to retrieve jobs and publish deliverables. Vendors who chose to integrate with different Kaltura SaaS regions will have multiple Vendor and Customer PIDs specific to each SaaS region.

Naturally, data privacy and sovereignty are critical considerations for customers when choosing specific SaaS regions. If a Reach vendor processes or stores customer data outside the enabled Kaltura SaaS region, it conflicts with the prime requirement of data localization. Therefore it is expected that Reach vendors who integrate with a SaaS region will also operate their infrastructure and code in the same region.

The possible architectures are as follows:

- **[MODEL 1] "Centralized job manager and decentralized regional task processors"**
    - The vendor will leverage a unique and central "Reach job manager" (for example in the US) to retrieve jobs from the multiple API endpoints of each SaaS region. The central "job manager" interfaces with all Kaltura API endpoints of every region (for example US and EU).
    - The central "Reach job manager" will leverage the `vendorTaskProcessingRegion` to decide in which region to store & process the data as indicated by the `vendorTaskProcessingRegion` (for example "EU").
    - The central "Reach job manager" will assign a decentralized and unique "Reach task processor" in the region (for example "EU") indicated to download the customer asset, store it and process it.
    - The "Reach task processor" in the region which complies with the requirement (EU) will then publish the deliverables to the Kaltura API endpoint in the region (for example EU) from which the reach job was retrieved.
        - For example there can be a case for orders submitted from a customer PID in SaaS region US to be processed in EU.
- **[MODEL 2] "Decentralized job manager and task processor"** - If vendors cannot leverage the vendorTaskProcessingRegion to decide in which region to store & process the data, then:
    - The vendor will leverage multiple decentralized "Reach job managers" in each region to retrieve jobs from the multiple API endpoints of each SaaS region. Each "Reach job manager" interfaces with only one Kaltura API endpoint of it's assigned region, the 1:1 assignment must be done in a way that each "Reach job

manager" is located in the same region as the Kaltura SaaS region.

- The "Reach job manager" ignores the `vendorTaskProcessingRegion` instruction.
- The "Reach job manager" will assign a local "Reach task processor" in the same region (for example "EU" Reach job manager assigns a local “EU” Reach task processor) to download the customer asset, store it and process it.
- The local "Reach task processor" in the region which complies with the requirement (EU) will then publish the deliverables to the Kaltura API endpoint in the region (EU) from which the reach job was retrieved

Note that in both models, Kaltura API endpoints of the various SaaS regions will be require authentication with different secrets.

⚠ A limitation of model 2 is that if a customer orders a job to be processed in a different region (ie. US) from which it is ordered (ie. EU), the order will be processed in the region from where it is ordered (ie. EU).

Here is a summary of how the two main vehicles are leveraged depending no the chosen model:

	Processing region in Reach jobs	**SaaS Regions API endpoint**
Indicates where data should be downloaded, stored and processed by vendor	Yes in Model 1 No in Model 2	No in Model 1 Yes in Model 2
Enables retrieving jobs from a specific SaaS region	No	Yes

Regardless of the chosen model, it is important that the vendor logs the Kaltura API source and the `vendorTaskProcessingRegion` instruction.

Content Deletion

Kaltura users are very sensitive about their content. Thus, each task will contain a deletion policy which the vendor is required to obey. Each Task will contain an attribute called `contentDeletionPolicy` (inside reach_vendor_profile) which will define if and for how long the vendor is allowed to store the media file obtained from Kaltura on their servers. The available deletion policies are:

https://developer.kaltura.com/api-docs/General_Objects/Enums/KalturaReachProfileContentDeletionPolicy

At the time of writing of this document, we support the following deletion settings:

Name	Value
DO_NOTHING	1
DELETE_ONCE_PROCESSED	2
DELETE_AFTER_WEEK	3
DELETE_AFTER_MONTH	4
DELETE_AFTER_THREE_MONTHS	5

The `contentDeletionPolicy` should be respected regardless of the model chosen to integrate with the Kaltura SaaS regions.

[template("cat-subscribe")]