Principles

Sending content to cold storage

WRITE_COLD_STORAGE permission

Status checked once per hour by default, changed with nuxeo.conf property

Retrieving content temporarily from cold storage

No specific permission, read permission seems to be enough

Status checked once per hour by default, changed with nuxeo.conf property

Restoring content from cold storage

WRITE_COLD_STORAGE permission

Status checked once per hour by default, changed with nuxeo.conf property

Making it work at scale

What options are available?

Bulk action

Using a scheduler

e.g., every night, send all claims that were solved more than x time ago and that are not under retention to cold storage

Triggered manually or using the REST API

e.g., to avoid deploying a new version of the app, you can configure a button only available to specific users that will send search results into cold storage

Using an event listener

e.g., when a document reaches a specific state or when a metadata is set to a specific value

How do I know how many documents will be impacted before triggering things?

Check the results count returned by a call to the search endpoint with your query

An audit trail is added by default for cold storage related actions. Details in the following Jira ticket:

https://jira.nuxeo.com/browse/NXP-30779

Time to retrieve content from cold storage is 3 to 5 hours. We leverage AWS Standard S3 Flexible Retrieval storage.

Time for restore is consistent no matter the content type. S3 Glacier is designed for 35 random restore requests per pebibyte (PiB) stored per day (source: https://docs.aws.amazon.com/amazonglacier/latest/dev/downloading-an-archive-two-steps.html), which should prove quite sufficient.

The architecture of the cold storage addon relies on the standard Nuxeo Platform principles, which makes the cold storage service customizable using code. It is possible to disable email notifications when content is retrieved or to apply a different behavior that way.
https://github.com/nuxeo/nuxeo-coldstorage/blob/lts-2021/nuxeo-coldstorage/src/main/java/org/nuxeo/coldstorage/service/ColdStorageServiceImpl.java

It is possible to prevent a document that was restored from being re-archived using some configuration or customization. How to achieve it will derive from the business rules you want to define in your application.

A low-resolution preview can be provided for any document type. By default, we rely on a thumbnail for office documents / pdfs, a low-resolution picture for pictures and a low-resolution video for videos. You can change these setting with configuration using Nuxeo Studio.

https://github.com/nuxeo/nuxeo-coldstorage/blob/lts-2021/nuxeo-coldstorage/src/main/resources/OSGI-INF/coldstorage-rendition-contrib.xml

There are both a permanent restore from cold storage option and a temporary restore from cold storage option available. We expose them in Web UI and as usual with Nuxeo Platform, it is possible to leverage these options in a variety of different ways as well if you want to.
https://github.com/nuxeo/nuxeo-coldstorage/tree/lts-2021/nuxeo-coldstorage/src/main/java/org/nuxeo/coldstorage/operations

When retrieving content from cold storage, annotations on the file will be available, exactly as you left them.

Annotations are stored as documents, so in MongoDB. There is no specific storage for them.

The low-resolution file is stored in standard S3 and counts towards your storage quota.