

Comparing IBM® FileNet® BatchIt
CE
with
IBM® Content Engine Bulk Import

Unique IBM® FileNet® BatchIt CE Capabilities

- ✓ Import Annotations, multiple Versions or Compound Documents
- ✓ File documents into Folder
- ✓ Report about archived documents for post processing or auditing
- ✓ Integrate document archival synchronously into a more complex document entry process
- ✓ Automatic file cleanup after archiving of documents
- ✓ Extended Error Handling / Automatic Retry
- ✓ Highly compatible with IBM® FileNet® BatchIt IM
- ✓ Support for generic XML import format
- ✓ Container ready
- ✓ Support Cloud Pak for Automation based FileNet Backends environments

More Details

Introduction

Both tools

- ✓ Import medium or large volumes of documents into Content Platform Engine on Prem or for SaaS (BACSoC and BAWoC)
- ✓ Use batch concepts to efficiently import documents into a Content Platform Engine
- ✓ Controlled by command line.
- ✓ Use text based input files to describe document meta data and content.
- ✓ Support import of single and multiple content elements for a document (e.g. like pdf file or single-page tiff files)
- ✓ Supports single and multi-valued property types String, DateTime, Float and Integer.
- ✓ Using files representing the content of documents from the file system.

Features 1/2

IBM® FileNet® BatchIt CE	IBM® Content Engine Bulk Import
Uses a descriptive input job format to describe a set of documents.	Uses a comma separated file (CSV) format to describe a document.
Provides various options to collect and process import jobs.	Works as daemon, watching input folders for new import jobs
Import documents using a configurable phase model.	Import documents using a fixed processing model.
Failed imports can be automatically or manually recovered.	Failed imports need manual intervention.
Configurable reporting for post processing or auditing purposes.	Journal file with static information.
Configurable File crawler feature.	

Features 2/2

IBM® FileNet® BatchIt CE	IBM® Content Engine Bulk Import
Optional cleanup phase to remove content and annotation files after successful archiving	
Support for properties of type binary and object	
Compatible with IBM FileNet BatchIt IM	Compatible with HP II
Support of versioning documents	
Support of compound documents	
Support to file documents into folders	
Extensible concept to implement customer requirements as part of a service engagement	

The Key Benefits IBM® FileNet® BatchIt CE

- ✓ The descriptive input format, compared to the CEBI format is less error-prone.
- ✓ The various options to run BatchItCE support synchronous integration into document entry processes or running unattended in the background.
- ✓ Configurable reporting allow to control the amount of information about a job and the imported documents for further processing or auditing issues.
- ✓ The richer feature set gives more options to cover business requirements (filing documents to folder, versioning of documents, compound documents).
- ✓ Optional cleanup reduces operating times and help to preserve storage resources.
- ✓ Easy migration from BatchItIM to BatchItCE.

The Key Benefits IBM® FileNet® BatchIt Input Format

- A descriptive input format:
 - ✓ Easy to generate by scripts or programs.
 - ✓ Easy to review, in case of errors, using standard editor programs.
 - ✓ Self containing, not requiring a configured class or index mapping.
 - ✓ Each line specifies a single aspect of a document like
 - the document class name,
 - the files representing the content,
 - the property name and values,
 - the folders a document should be filed in,
 - and more ...

The Key Benefits IBM® FileNet® BatchIt Input Format Sample

COrder

FSignedOrder_94849.pdf

Mapplication/pdf

AAnnotation01.txt

I_DocumentTitle=Signed Order 94849

I_ContractId= 94849

I_DateOfSignature=12.02.2019

I_CustomerName=Jon Doe Consulting Inc.

I_RelatedContracts=94851|949128

I_CustomerDataObject={1AB100E0-CD52-4962-BA46-CB2E7ED5AA51}

P/Jon Doe Consulting Inc/Orders ← File the document to a folder

D

• ...

BatchItCE Input Format Benefits

- ✓ The descriptive format is less error-prone compared to complex csv files.
- ✓ Self-containing in regard of describing the documents. Requires no class and index mapping configuration.
- ✓ Easy to generate by scripts or programs.
- ✓ Support of annotations, naming content elements and its mime types.
- ✓ Support to create and add new versions of documents.
- ✓ Support to create compound documents. #
- ✓ "Could use as input stream data different types of XML format files, transformed in a first pre-phase to a standard BatchItCE XML file"

BatchItCE Job Processing Options

- BatchItCE provides various options to collect and process input jobs:
 - Processing single input jobs explicitly.
 - Collect and process all input jobs located in a specific folder optionally including sub folders.
 - Periodically collect and process all input jobs and failed jobs located in a specific folder and optionally all sub folders.
 - Processed input jobs, related report and log files moved or renamed for further processing or auditing in a very flexible manner.

BatchItCE Job Processing Options Benefits

- ✓ Synchronous integration of BatchItCE as part of a more complex document import workflow.
- ✓ Unattended background processing of new input jobs reduce operative efforts.
- ✓ Automatic retry of failed jobs in case of unexpected backend system outages or network issues reduce operative efforts.

BatchItCE Phase Concept

BatchItCE use a configurable phase-model to process a single input job:

- ✓ During the mandatory first phase, the data of an input job gets validated and transformed to an intermediate format.
 - Class and meta-data is validated.
 - In case of validation errors the input job is rejected.
 - The input job is transformed into an intermediate format.
- ✓ The mandatory archive phase physically imports the documents to the Content Engine object store.
 - Uses the intermediate format for tracking and retry puporposes.
- ✓ The optional report phase generates report files containing meta-data and run-time information from an input job.
 - The format and information contained in report can be configured in a flexible way.
- ✓ The optional cleanup phase removes successfully imported document

BatchItCE Phase Concept Benefits

- ✓ Input Jobs containing invalid meta-data or file references completely rejected to avoid automatic reprocessing causing the same error again and again.
 - The configuration of BatchItCE supports flexible options to rename or move failed input jobs.
- ✓ In case of unexpected errors the intermediate file used to retry failed job from the point of failure.
 - Especially avoid archiving the same documents multiple times due to failures.
- ✓ The flexible configuration options for reports allows to create report files for auditing or post processing by foreign systems.
- ✓ The cleanup phase helps to save disk space and reduces the operating time.
- ✓ Phase model could be extended, based on customer requirements (e.g. a folder phase allowing to create folder hierarchies).

BatchItCE Other Features

- ✓ Support for properties of type binary and object
 - Import of binary data related to a document.
 - Link documents with other objects (e.g. Security Proxies).
- ✓ Filing of documents to one or more folder.
- ✓ Support to import annotations.
 - Import and assigning annotations to document content elements (pages).
- ✓ Support of versioning of documents.
 - Create series of versions by a single import job.
 - Add versions to existing documents.
- ✓ Compound document support.
 - Creating or extending parent-child relations for compound documents.

BatchItCE Other Feature Benefits

- ✓ Support for binary valued properties useful for storing binary data (e.g. a signature)
- ✓ Support for object valued properties useful for linking documents with other objects (e.g. security proxies).
- ✓ Import of annotations as part of DMS migration projects.
- ✓ Adding new Versions of automatically generated documents, like new revisions of technical documentations created by another system.
- ✓ Compound document useful to create complex documentation consisting of related documents.

Glossary

Verb	Description
CEBI	Abbreviation for IBM® Content Engine Bulk import
HPII	Abbreviation for High Performance Image Importer. A program to efficiently import documents into IBM® FileNet® Image Services system.
BatchItCE	Abbreviation for IBM® FileNet® BatchItCE
BatchItIM	Abbreviation for IBM® FileNet® BatchIt IM. A program to efficiently import documents into IBM® FileNet® Image Services system.
Input job	Describes a collection of document to be imported. A single Input Job delivers the class information, meta-data and references to the files making up the content for a set of documents.