

Migrating emails (from files) to contentACCESS Email archive with contentGATE

Import connector

The **Import connector** allows to migrate emails from files and zip files into any target supported by **contentGATE**. The files can be stored in various folder structures and the **Import connector** is able to leverage this information as mailbox folders.

The **Import connector** can also be used to migrate journal mailbox exports.

Email files

Emails can be saved in various formats; the most common formats are:

- MSG (Microsoft Outlook message format)
- EML (RFC-822 or MIME)

The files can be part of a zip file or can be located in a complicated folder structure.

Note: The EML format usually contains less information than the MSG format. The conversion from MSG to EML will lose user specific properties (like category, follow up flag,...) in the mail.

Where does the email belong?

The biggest challenge when working with email files is to discover which mailbox it must be assigned to and under which folder in the mailbox.

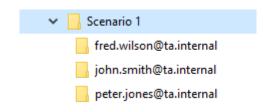
In the simplest scenario, the emails are organized in folders on the disk, where a folder is belonging to exactly one mailbox. In a more advanced scenario, the emails are dumped into a big pool of files where the folder does not define the mailbox where the mail belongs.

Folder structures

The import connector is able to migrate various folder structures:

- a) Top level folder containing the whole content of a mailbox
 - there is one dedicated folder for each mailbox
 - the name of the folder can be used to identify the mailbox

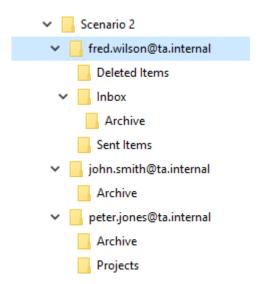




← → ▼ ↑ → This PC → Windows (C:) → Temp → Scenario 1 → fred.wilson@ta.internal				
Name	Date modified	Туре	Size	
1a9902500a9345a492325b3d5eb3bd19.msg	15.3.2017 12:35	Outlook Item	80 KB	
1f1fc10e4d5a444b888ddc9a329b1cd4.msg	15.3.2017 12:35	Outlook Item	72 KB	
2a45e108990448dabd51e8218e8916e1.msg	15.3.2017 12:36	Outlook Item	104 KB	
2ab910f4e5b2411690d8a6812bd272ba.msg	15.3.2017 12:38	Outlook Item	108 KB	
2abd91599c2e48c4ae8f19746029d126.msg	15.3.2017 12:37	Outlook Item	88 KB	
2bcfde1a76c54e42bdc2c5b93225f79f.msg	15.3.2017 12:37	Outlook Item	72 KB	
2f2533f376c6413bb40568d4db3026b4.msg	15.3.2017 12:38	Outlook Item	184 KB	
2fe160a5d6bd41c1810d74cd12467831.msg	15.3.2017 12:37	Outlook Item	504 KB	
3b1873aa8ae94e1f8f273d3bf1f25a9a.msg	15.3.2017 12:39	Outlook Item	288 KB	
3d13dce8739d4baab195c9c470e4b671.msg	15.3.2017 12:37	Outlook Item	96 KB	
3df4421a3abe4fedaa4c16d4f7b7e091.msg	15.3.2017 12:36	Outlook Item	80 KB	
4ab541de40064617b6ec0a8d75e922a9.msg	15.3.2017 12:37	Outlook Item	124 KB	
05a9279f6f4c4aef9e7a06d0ce34c580.msg	15.3.2017 12:37	Outlook Item	152 KB	
5d47haa7de4f45c7h7adchfd7cahcd47.msg 102 items	15.3.2017 12:39	Outlook Item	144 KB	

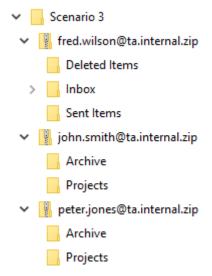
b) Folder containing subfolders

- similar to a)
- there are subfolders, which are corresponding to mailbox folders (i.e. Inbox, Sent Items, etc.)

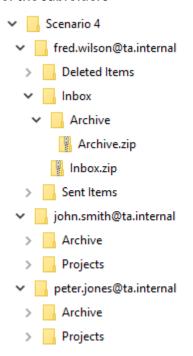




- c) Zip files containing the whole content of a mailbox
 - one Zip folder is belonging to exactly one mailbox
 - the Zip file can contain subfolders, which are corresponding to mailbox folders
 - multiple Zip files can belong to the same mailbox



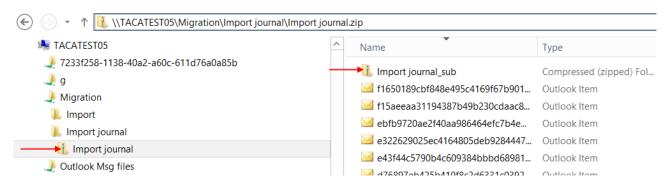
- d) Zip files in subfolders
 - similar to b)
 - there can be a Zip file in any of the subfolders



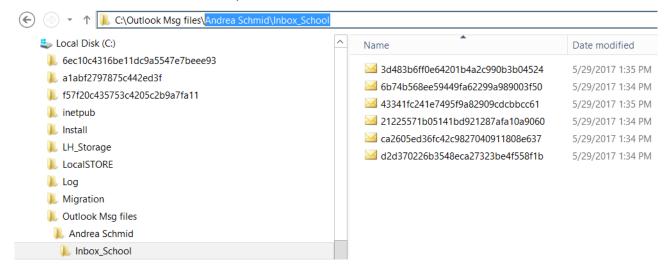
However, some rules/limitations still apply:

• The zip file can't contain further Zip files. If it does (like on screenshot below), it is ignored.



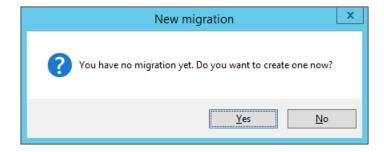


 The relative folder path (blue highlight on the screenshot) together with the file name can't reach 240 characters. If so, the file path must be shortened.



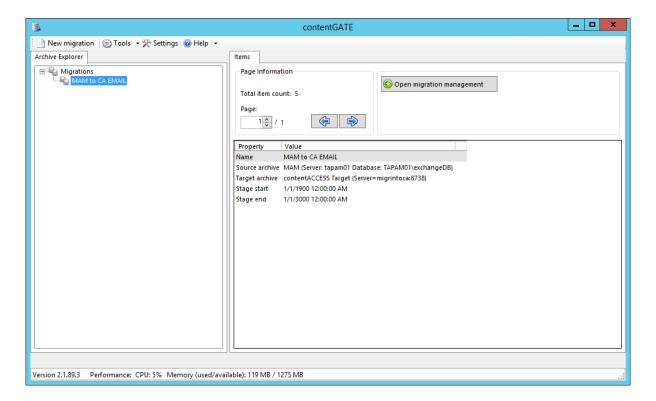
contentGATF welcome interface

When contentGATE is started for the first time, no migration is defined. The application prompts you to create one. You can click **No**. New migration unit can be created anytime by clicking on the **New migration** option in the toolbar.





The contentGATE welcome interface features a toolbar, offering easy access to general settings and functions:



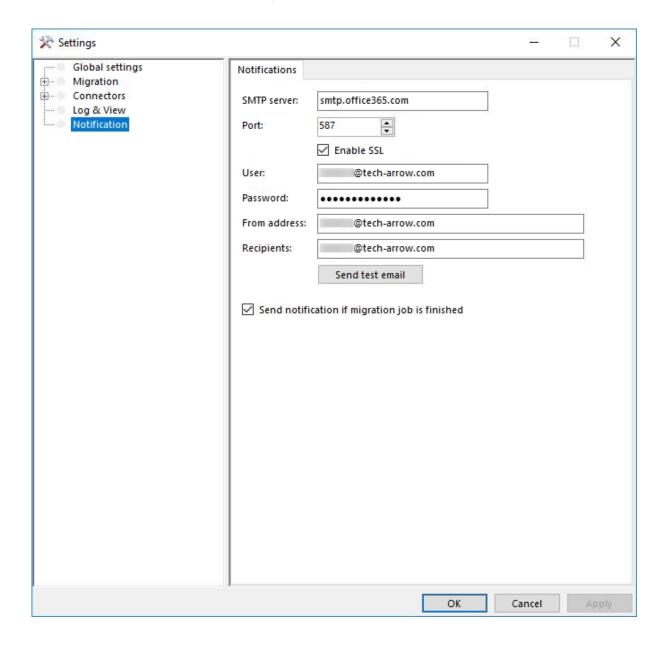
- **New migration** creates a new migration unit under which the source archive, target archive, user mapping etc. can be configured
- Tools connector specific tools are accessible from here

Note: These tools can be disregarded for this migration.

Settings – the general contentGATE settings and individual connectors (archives) settings are
accessible from here.

If you would like contentGATE to send you notification email when migration job is finished, you can configure it in this section.





SMTP server: the SMTP server that you want to connect to

Port: 25 or 587 (for secure connection)

Enable SSL: check this checkbox if you want your secure connection to use the SSL (Secure Socket Layer) protocol to encrypt communication

User and Password: credentials of the user you want to authenticate with

From address: email address from which the notification emails will be sent (the email address of the authenticated user)

Recipients: email addresses that will be receiving notification emails from the job; multiple addresses need to be separated by semicolon



Send notification emails if migration job is finished: do not forget to check this checkbox, otherwise notifications won't be sent. No matter if you stop it manually or it finishes by itself, the email will be sent.

• Help - information about license, version etc.

The user interface below the toolbar is divided into two panes:

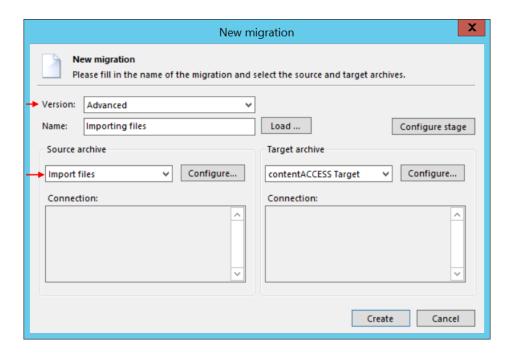
- Archive Explorer pane with a list of created migrations. Initially the pane is empty.
- Items pane that displays the content of a node selected in the Archive Explorer pane. Initially the pane is empty.

Use the **Open migration management** button to open the management interface of an advanced migration project that is selected in the **Archive Explorer** pane.

The **Page information** section in the **Items** pane allows you to adjust and handle items per page functionality.

Import connector configuration

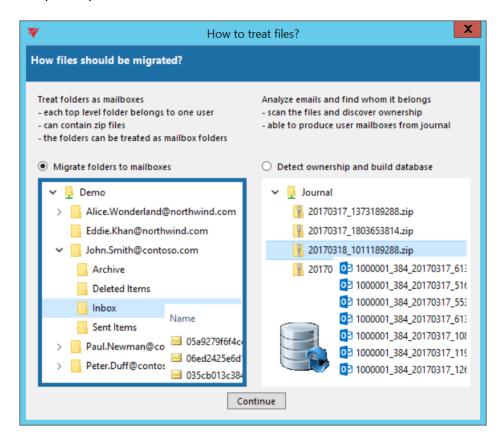
If you want to configure an **Import connector**, **Advanced** needs to be selected in **Version** and **Import files** needs to be selected in **Source archive**.





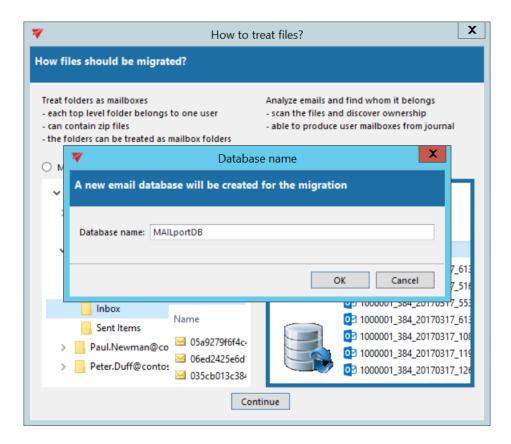
After clicking on **Configure** button on the left side, you have to choose from two options: **Migrate folders to mailboxes** or **Detect ownership and build database**.

- This selection can't be changed later
- See <u>5.1</u> or <u>5.2</u> for more info about the options
- Know your input data structure select based on that

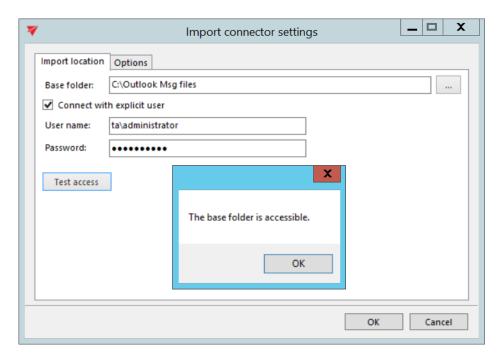


If you chose **Detect ownership and build database**, you need to name your database first.





The configuration requires a base folder, where the email files are located. This can be any local or UNC path. If explicit credentials are required to access the base folder, configure an account, which has at least read permissions on the folder. This can be configured in the **Connect with explicit user** part. The validity of the parameters can be tested by the **Test access** button.



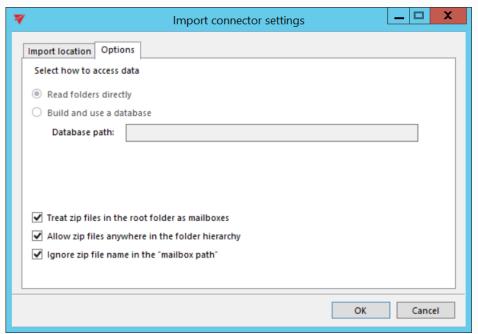


On the **Options** tab the type of the processing is already selected based on your previous selection – it cannot be changed in this step

- Read folders directly if you chose Migrate folders to mailboxes
- Build and use a database if you chose Detect ownership and build database

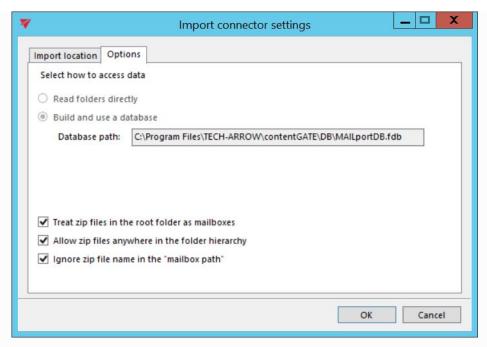
Additional options are available here to configure zip file usage:

- Treat zip files in the root folder as entities if you have zip files in the base folder, they can appear as individual mailboxes
- Allow zip files anywhere in the folder hierarchy use this option if zip files are located in subfolders
- Ignore zip file name in the "mailbox path" use this option to remove the zip file name from the path



Screenshot - Options when picked Migrate folders to mailboxes





Screenshot - Options when picked Detect ownership and build database

The **Source** archive part is now configured.

Read folders directly

Configure an Import source connector, pointing to the folder or Zip file, where the mailboxes to be migrated are exported.

In **How to treat files?** pick **Migrate folders to mailboxes**. contentGATE will recognize the top-level folders as entities (i.e. mailboxes) and treat them like source mailboxes (with or without deeper folder structure). The email files are usually the result of an export from a system, for example a legacy archive one.

The export might be already finished, meaning that all emails are already in the base folder. In this case the migration can be performed immediately without any limitations.

There might be situations when the export is not finished yet, meaning that new email files are exported continuously. In this case the migration must be planned carefully. **contentGATE** allows a continuous migration also in this situation, but some actions need to be taken multiple times based on the availability of the exported data.

The general rule is that the content of the base folder should not be changed: no files should be deleted or moved, however, new files and folders can be added. The migration is picking up only the files which



were not migrated yet – this is called **delta migration**. If continuous export is used, the migration can be performed parallel to the export. It is necessary to run the **delta migration** after the export is definitely finished.

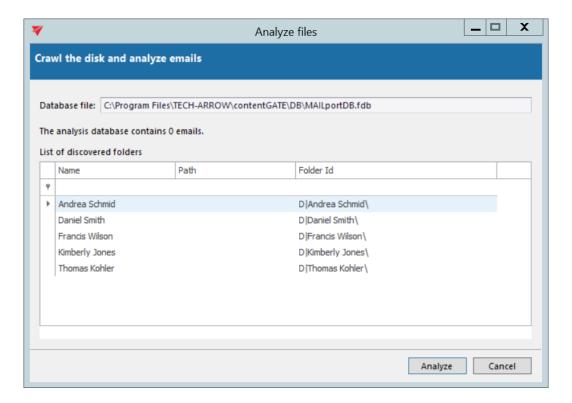
Build and use a database

In situations where the email ownership is not known, this information must be extracted from the emails itself. In this case a database must be built in order to categorize and assign mails to mailboxes and all emails must be present before the database is built.

To build the database, pick **Detect ownership and build database** in **How to treat files?**, name your database – this will create it and it will be used to hold the email metadata. The created database is empty and must be filled. The database is filled by "Discover emails...".

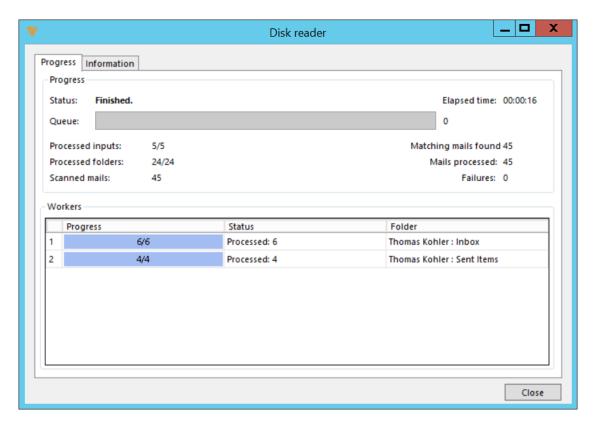


The discovery process displays the top-folders of the base folder and if it already contains emails, displays the number of emails as well.

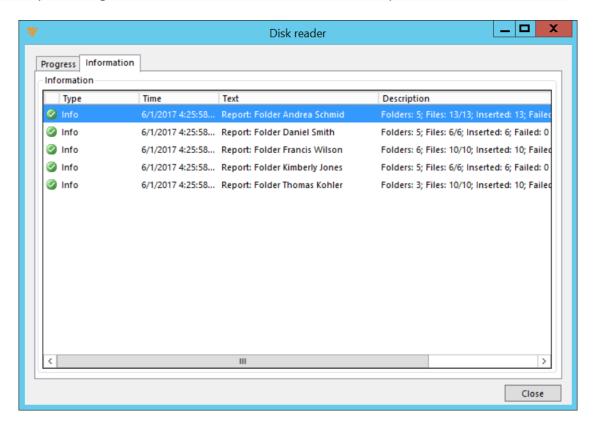




Start the analysis process with the Analyze button.



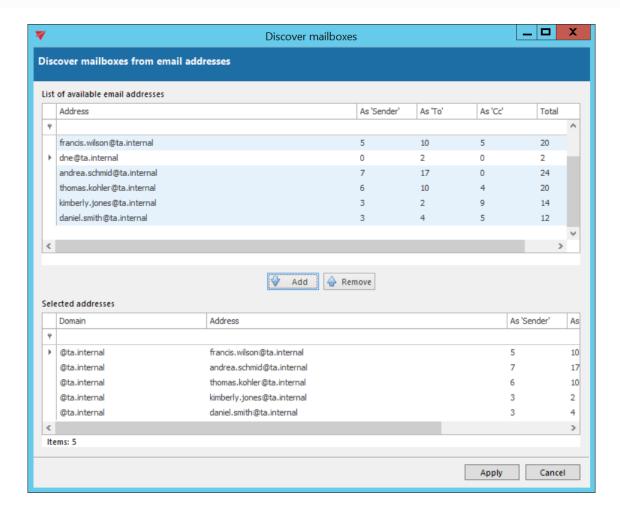
Once the processing is finished, the database has all the necessary metadata about the emails.





The next step is to select which mailboxes or which email addresses to use. This can be done by the "Discover mailboxes..." button.

This dialog loads the distinct email addresses from the whole database and groups them by domain name. The administrator can select which email addresses are important:



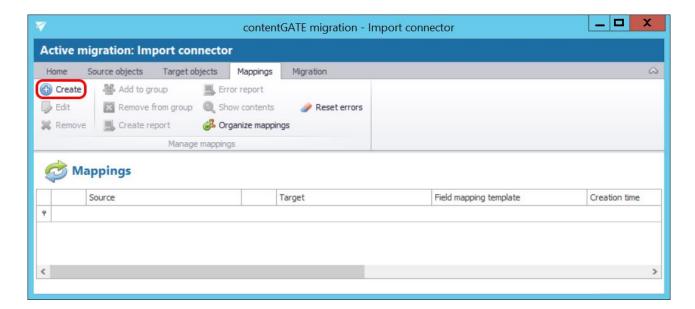
Once the addresses are selected and applied, the database is ready to be used. It is possible to add more addresses at any time – they will appear as new source entities, which can be migrated.

contentGATE organizes the emails into 2 fixed folders in each discovered mailbox: Incoming and Outgoing. If the user was present on the **From** address of the email, the email will be put into the **Outgoing** folder. If the user was present on **To**, **Cc** or **Bcc** list, the email will be put into the **Incoming** folder.

Mappings

Open the "Mappings" tab and click "Create" in the toolbar to create the desired object mapping(s).





The "Mappings" dialog features the

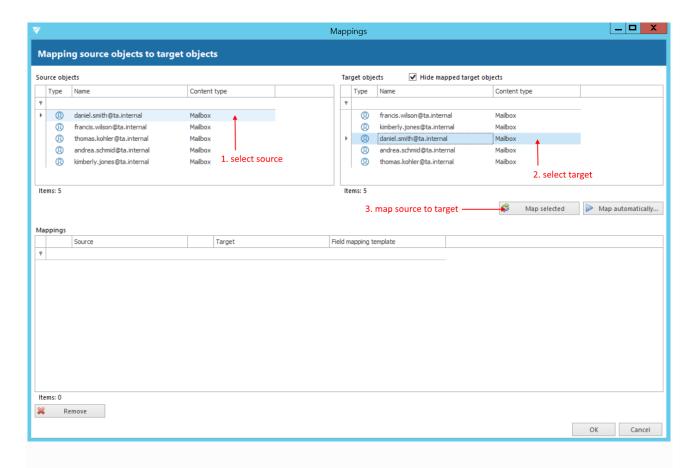
- source object(s) TO BE MIGRATED on the source side, and
- target object(s) WHERE the source objects can be migrated on the target side.

Note: One source object can be mapped to one target only, but one target object may be associated to multiple source objects. This action is called merging. Wish you associate multiple source objects to the same target, uncheck the "Hide mapped target object" option above the grid.

Mappings may be specified either manually or automatically.

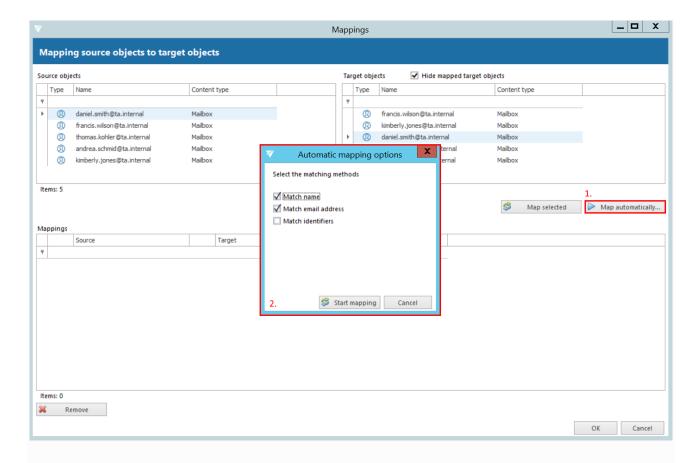
By **manual selection** select each source and target one-by-one, and map them together with clicking on the "Map selected" option.





Wish you use **automatic mapping**, 1) click on the "Map automatically" button and 2) select the mapping criteria.





There are 3 auto-mapping methods available:

Match name - maps the objects based on name

Match email address – maps the source to target based on the email address included in the object properties

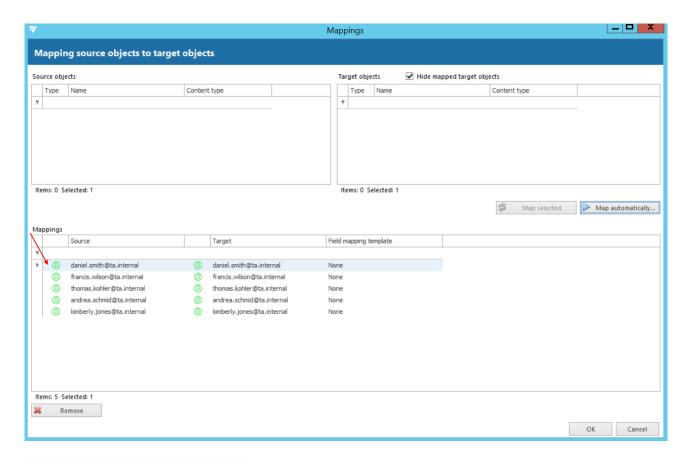
Match identifiers – maps the source to target based on the IDs included in the object properties (e.g. object ID, entity ID etc.)

Once the criteria are selected, click on the "Start mapping" button, and the objects will be paired automatically.

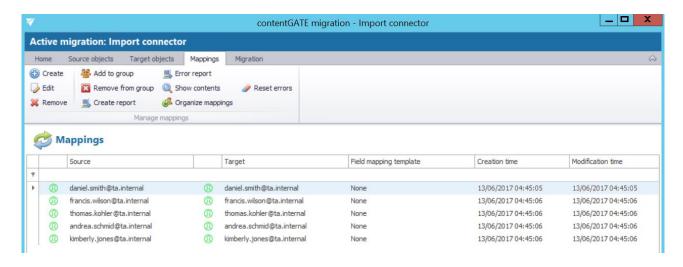
Note: If no matches have been found by the automap, the application will warn about it. In this case the user needs to map the objects manually.

Added mappings are shifted to the lower part of the **Mappings** dialog.



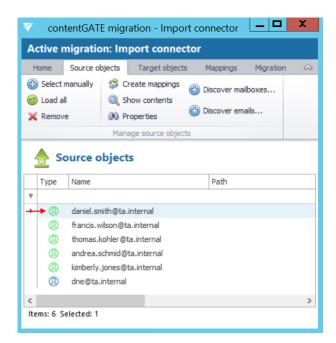


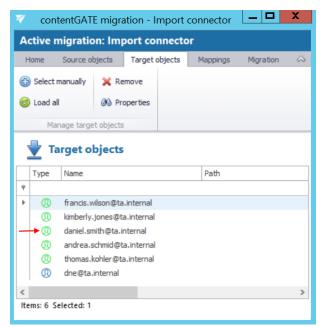
Click "OK" to add them to the grid.



Note: Mapped objects are marked with green color in the source and target grids, while unmapped objects remain blue.







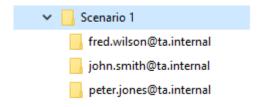
The contents of the mapped source objects can be checked using the "Show contents" button in the toolbar, too.

Folder containing unsorted content of one mailbox

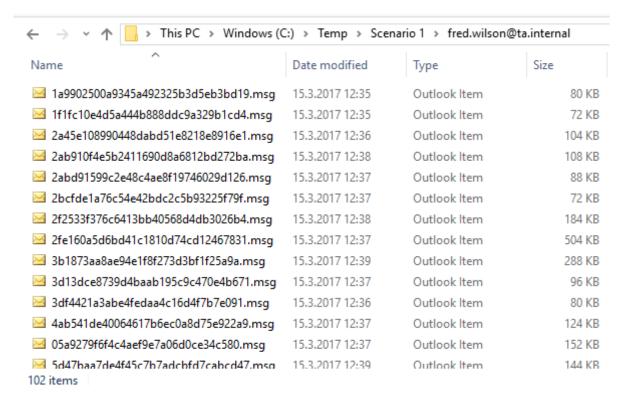
Use this scenario if you have folder structure as following:

Each top-level folder is one mailbox.

To make the mapping easier, the top-level folder names should contain the email address or name of the mailbox. In which case the folders can be auto-mapped to mailboxes.



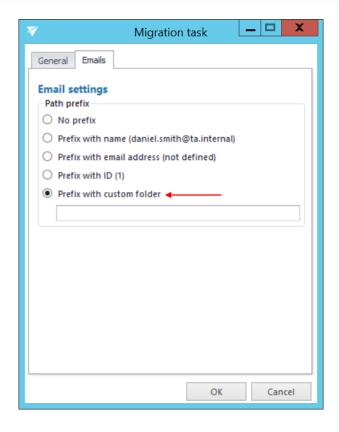




Screenshot 1. The contents of one folder: Scenario 1\fred.wilson@ta.internal

On the screenshot above the messages are not organized into subfolders.

In situation like this, it is highly recommended to specify a fixed folder path (see migration mapping parameters) for the migration.





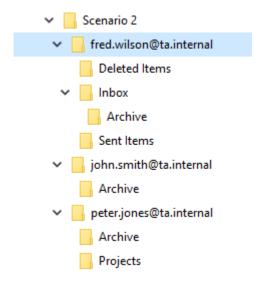
Screenshot 2. Migration Mapping parameters

If no fixed folder is assigned, the emails will be migrated under the root folder of the mailbox, which is not supported by all target connectors.

Folder containing sorted content (subfolders) of one mailbox

This scenario is a slightly modified version of <u>6.1</u>:

The emails are not located under the top-level folder, but are organized into subfolders.



In this scenario, the subfolder path will decide the folder path during the migration.

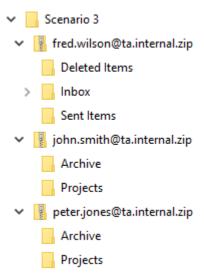
If no mapping parameter is defined, the folders will be used without modification: Inbox will be Inbox, Inbox\Archive, etc.

By defining the migration mapping parameters, it is possible to transform the folder path for example by prefixing the folder: Inbox will be transformed to Migration\Inbox\Archive will be transformed to Migration\Inbox\Archive, etc.

Zip file containing the content of one mailbox

The top-level folders can be replaced by zip files. In such situations, the zip file is containing a whole mailbox or part of the mailbox. The zip file might contain the subfolders (similar to $\underline{6.2}$) or it does not need to contain any subfolders (similar to $\underline{6.1}$).



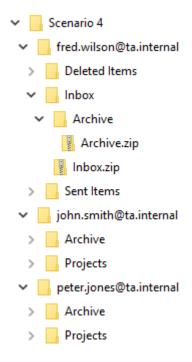


The processing is similar to both 6.1 and 6.2.

Note: Very large zip files will have a large impact on the migration performance.

Zip files as subfolders

A modified <u>6.2</u> where the subfolders are containing zip files:

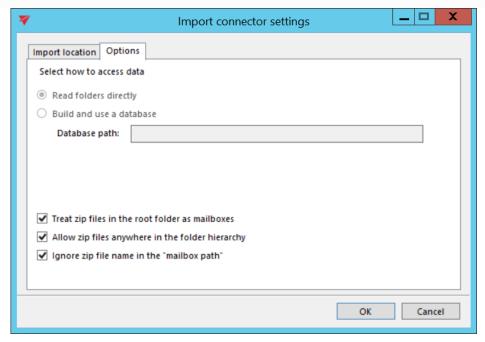


It does not matter whether the emails are stored as single files or are compressed – they are treated as equal.

In this scenario, it is recommended to use the following settings:



- Allow zip files anywhere in the folder hierarchy use this option if zip files are located in subfolders
- Ignore zip file name in the "mailbox path" use this option to remove the zip file name from the path

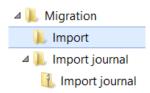


Migrating emails from journal mailbox to user mailboxes

The first step is to configure the **Import connector** (source archive), please refer to section <u>Import connector configuration</u>.

The source folders can have one of the following structures:

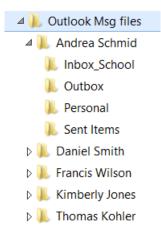
■ Top level folder/zip file containing the whole content of all journal mailboxes (similar to <u>6.1</u>)



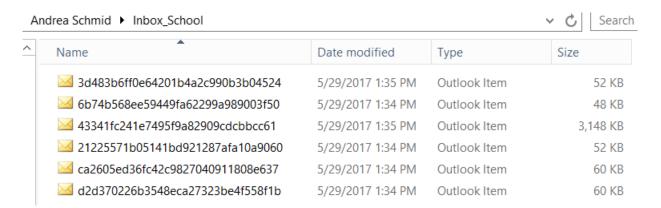


Of3c80e606094d468e037b86143307d9	5/29/2017 1:34 PM	Outlook Item	52 KB
1a0e50a02ac94f69b858518f3bedf91f	5/29/2017 1:34 PM	Outlook Item	48 KB
2fdacca9929e4b3c8f1d8e95c3c69c25	5/29/2017 1:35 PM	Outlook Item	3,144 KB
3d483b6ff0e64201b4a2c990b3b04524	5/29/2017 1:35 PM	Outlook Item	52 KB
3f4705fd47f540b28b1247e048430823	5/29/2017 1:35 PM	Outlook Item	3,152 KB
4ad5531e6fea4fdcaa33da54f561c836	5/29/2017 1:35 PM	Outlook Item	48 KB
6b74b568ee59449fa62299a989003f50	5/29/2017 1:34 PM	Outlook Item	48 KB
7a3744136ee44ea286821d49364a41fa	5/29/2017 1:35 PM	Outlook Item	44 KB
7c2e71a5ab774e61ba0b85f4f34b8f7b	5/29/2017 1:35 PM	Outlook Item	44 KB
7e305b5de2f74191881f2f844e9f03a7	5/29/2017 1:34 PM	Outlook Item	52 KB
8bb25e3fe57d4880a25a28dda2f48aeb	5/29/2017 1:35 PM	Outlook Item	52 KB
8fb2031c08d5423bb5d590a892da18da	5/29/2017 1:35 PM	Outlook Item	64 KB
9f43d2b8b4ea4b46969c902877b73c80	5/29/2017 1:35 PM	Outlook Item	64 KB
20a6045a5cba4c50a963f352b665f91d	5/29/2017 1:35 PM	Outlook Item	40 KB
31e1e5ef365b4d019939a68a2d03774c	5/29/2017 1:35 PM	Outlook Item	44 KB
64c83c52577948c285ef773344045395	5/29/2017 1:35 PM	Outlook Item	64 KB
85ca475119ca4db99c3b3da81cf76495	5/29/2017 1:35 PM	Outlook Item	44 KB
95bdbe79c998485c90edfe893559f97e	5/29/2017 1:35 PM	Outlook Item	40 KB
139bc574b76d477ebe1b3f2e80bde408	5/29/2017 1:35 PM	Outlook Item	52 KB
804ded406342443a902dd90d4af9a526	5/29/2017 1:35 PM	Outlook Item	44 KB

 Top level folder/zip file containing sorted content of all journal mailboxes (similar to 6.2, 6.3 and 6.4)



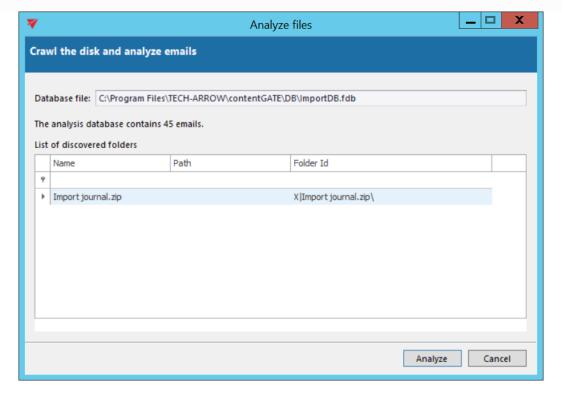




In How to treat files? pick Detect ownership and build database.

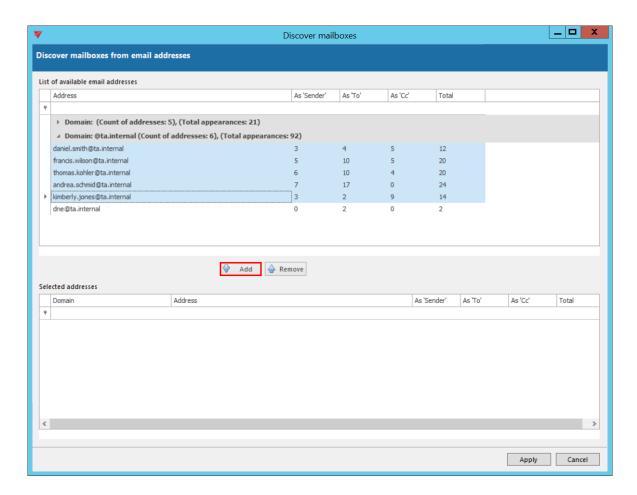
Next, you need to configure the **contentACCESS** (target archive). Please refer to <u>this</u> section, where the configuration is explained.

In the Migration management window select **Source objects** tab and click on **Discover emails**. A popup will show up, where you will be able to see the base folder that you selected when configuring the Import connector. Click on Analyze – this will discover all the email addresses that were used from the files in the base folder.

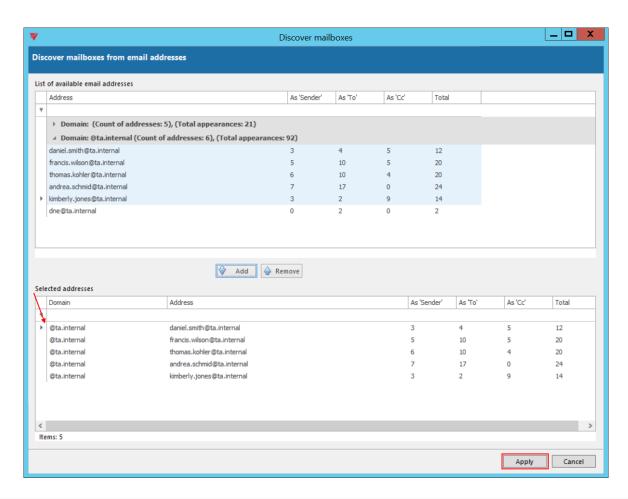


Then click on **Discover mailboxes**, again on the **Source objects** tab. Select the addresses that will be used as source objects and click on **Add**. This will add the addresses to the grid. Then click on **Apply**.

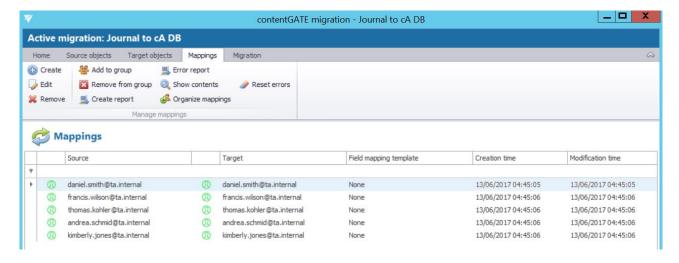






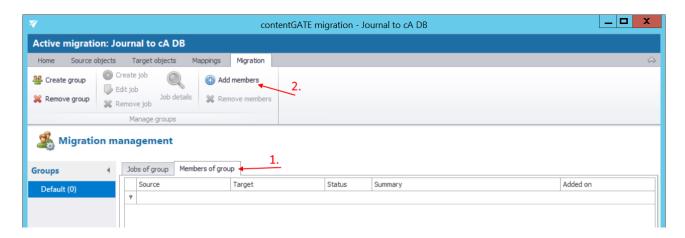


Create a mapping and map the source objects with target objects, either way manually or automatically.



Click on the **Migration** tab, then choose the **Members of group** tab and then click on **Add members** (if you wish to add members to a different group than Default, first you need to create one – click on Create group in the left corner).

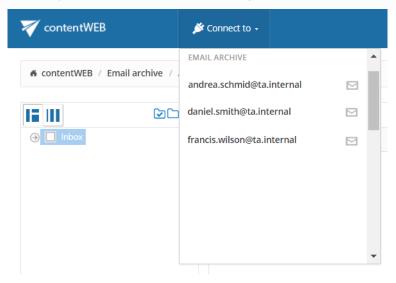




Click on **Jobs of group** and start the migration.

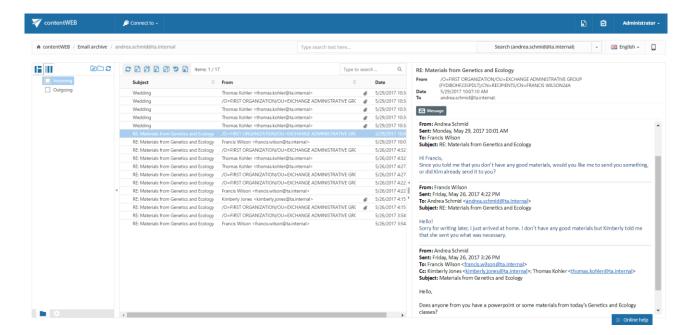
After the migration is finished and successful, log in to **contentWEB** with the user that has the rights to view the mailboxes you migrated. In the upper left part of the header click on **Connect to** menu to unfold the associated entities that you can connect to (to which an access permission is granted).

Under EMAIL ARCHIVE the journal mailboxes that you migrated can be seen.



After clicking on an entity (mailbox), you will be able to see its content.





Screenshot - content of a migrated journal mailbox