techarrow

contentACCESS

EMAIL ARCHIVE BEST PRACTICES

in contentACCESS version 7.0

Table of Contents

techarrow

Use case: Email Archive configuration for companies with up to 2,000 mail-boxes

This document gives you an overview about how to configure Email archive for the company that is similar to the company in this use case.

> *Important:* Please keep in mind that the names, configurations and units we chose in this use case have an informative character only.

What is contentACCESS?

contentACCESS is an information management platform for migration, collaboration, access, and transformation of data from different sources, any time, from any device.
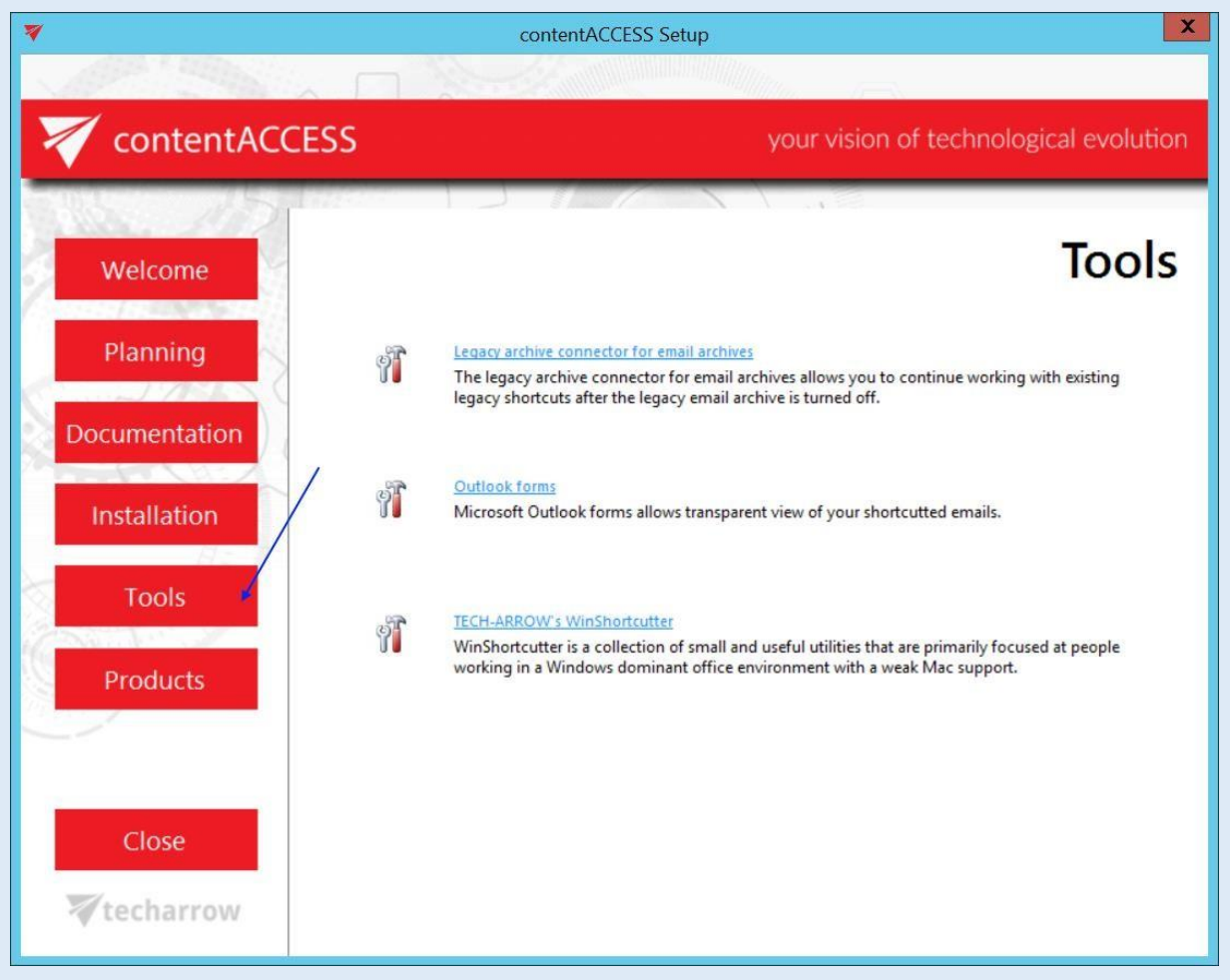
USE CASE

Goal:	**TECH-ARROW's mailbox archive**
Company size:	up to 2,000 mailboxes
Exchange server:	Exchange 2010 SP2 in this use case (possible to configure also Microsoft 365)
Group to be archived:	Archiving **Mes-**
sage classes to be archived:	Mails, Documents
Archiving method:	HTML shortcut
contentACCESS single server:	tacatest **conten-**
tACCESS server	
externally accessible over:	https://tacatest.ta.internal (to be configured in FQDN settings)
CPU:	2.4 GHz quad core
Memory:	24 GB
Storage:	single disk store
Database:	single database
Retention time:	10 years
Scheduler:	Provisioning job running at Saturday from 10 PM; Email archive process running every day from 6 PM
Monitoring feature:	Troubleshooting of potential system and job related misconfigurartion.

Installing contentACCESS

Install contentACCESS (server, central administration, and contentACCESS Portal) on a single server (in our case, *tacatest*) and navigate to the *Central administration* configuration interface of the software (for more details, refer to the **contentACCESS Manual**). Log in with the system user credentials. During the installation, assign 12GB of memory to the Elastic search server.

Note: It is also recommended to install proxy server over HTTPS (secure) connection and to publish contentACCESS Outlook forms in MS Outlook (to ensure HTML shortcut transparency). Outlook forms are accessible in the **Tools** part of the contentACCESS setup package.

External accessibility settings

Navigate to *System* ➐ *Services => System,* configure the **Resource storage** and specify the fully qualified domain name (FQDN) under *External accessibility settings* – in this use case, https://tacatest.ta.internal, otherwise the links in the HTML shortcuts will not work.

System

System > Services > System

General settings

Is system configured:	☑
Plugins directory:	C:\Program Files\TECH-ARROW\contentACCESS\GA

Globalization settings

Time zone:	(UTC) Coordinated Universal Time
Date format:	31/01/2019 5:50:50 (24h)

User search settings

Disable user searching:	☐

Tenant level resource settings

Disable resource settings:	☐

Resource storage settings

Storage path:	C:\TestingStorage
Use explicit credentials:	☐
User name:	
Password:	
Confirm password:	

Test

Database settings

Database type:	MSSQL
Server name:	TACATEST
Database:	MNEtestDB
Schema:	dbo
User:	sa

Configure ...

Proxy settings

- ⦿ Do not use proxy
- ○ Use system default proxy settings
- ○ Use explicit proxy server settings

Proxy server URL:	
Username:	
Password:	

Search server settings

Search server url:	net.tcp ：// localhost ： 8732

Test

Monitoring log storage settings

- ⦿ Use the system database
- ○ Use the default search server
- ○ Use a different search server

External accessibility settings

Retrieve server FQDN:	https ：// tacatest.ta.internal ： 443

Enter an externally accessible alias for the contentACCESS server.This value is used for shortcut links too.

Test

Enable link security:	☑
Central login:	
	http ：// tacatest.ta.internal ： 80

Provide necessary information, where central login is deployed.

Test

In case you are using online Exchange server, the Lightweight Directory Access Protocol (LDAP) settings do not need to be configured. If you are using an on-premise Exchange server and in all other cases, the LDAP settings need to be filled in, too. For LDAP settings, navigate to *System -> Services -> System* and configure it under the *Active directory integration settings*.

Adding a tenant

> *Important*: If you have a single tenant-license, you can skip this step, as the **Default** tenant will be automatically created.

Open the **Tenants** page (*System -> Tenants -> Tenants)* and click **+ new**. **Add and Configure** a new tenant & tenant administrator into the system.
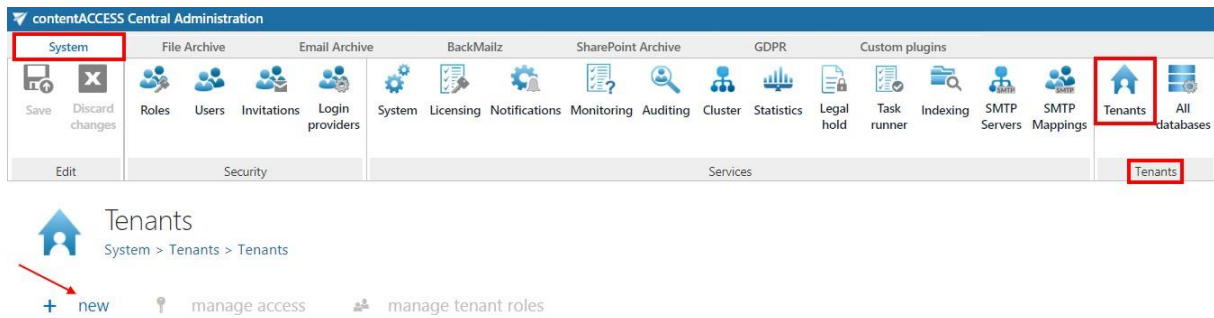
The newly created tenant appears in the upper right corner of the **Central Administration** user interface, like on the picture below:

Note: If you want to read the detailed description of tenant creation and configuration, please, check this section of our online documentation.

Creating a database to your tenant

Open page **All databases** (*System => Tenants => All databases)*, click **+ create new** and specify the required values in the *Database repository* window. Create a single SQL database to TECH-ARROW tenant with connection name "Email Archive DB" and Database user "Administrator". We won't use the system database server (TACATEST) as it is recommended to use a separate one, so we leave the "Use system database server (TACATEST)" check-box unchecked. The application will automatically set this first created database to default.

tech**arrow**

Configuration details

Connection name: the display name of the database in contentACCESS

Option "Use system database server": if the database must be created on the system database server, select this option; otherwise, specify the required database values (**Server name**, **Database**, **Schema**) manually.

In section **Database user,** you need to specify a user for the newly created database:

Enter the credentials of the existing MSSQL user or create a new user and password, clicking on the **Generate user and password** button. contentACCESS will give access rights to the database on MSSQL for this user. This newly created user will have access permissions only to

the created database. If you specify no explicit user in this section, then the service user will be used to access the database.

In section **Connect with user,** you must specify the MSSQL admin credentials**:**

Use explicit credentials: if this checkbox is not checked, then contentACCESS will automatically connect to the database (MSSQL) under the contentACCESS service account. If you wish to connect with an explicit user, then specify an admin user with the rights to create a new database on the database server.

Tenants: select the database owner tenant from the dropdown list (in our case **TECH- AR-ROW**).

> *Note:* When creating databases, always consider the size of the company and the number of mailboxes archived. For bigger companies with a high number of mailboxes (more than 2,000), it is always recommended to create separate databases, e.g., one for the store and one for the email archive to ensure the transparency of processes. The database created at first will be the default database. This option is configurable on the page of **Databases** using the context menu.

Email Archive settings

Navigate to the **Email Archive** tab on the ribbon.
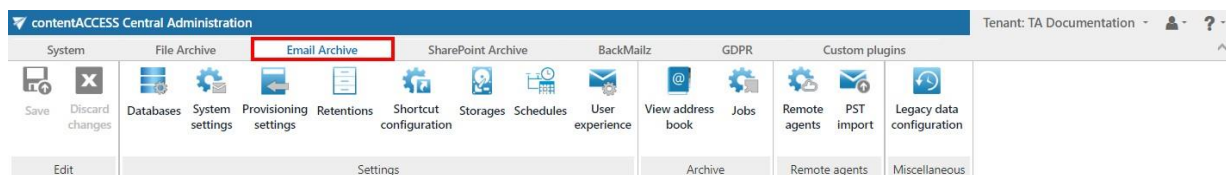
> *Note:* When opening the **Email archive** tab for the first time, you will have to activate Email Archive by clicking on the **Activate** button. With this action, a provisioning job will be automatically created.
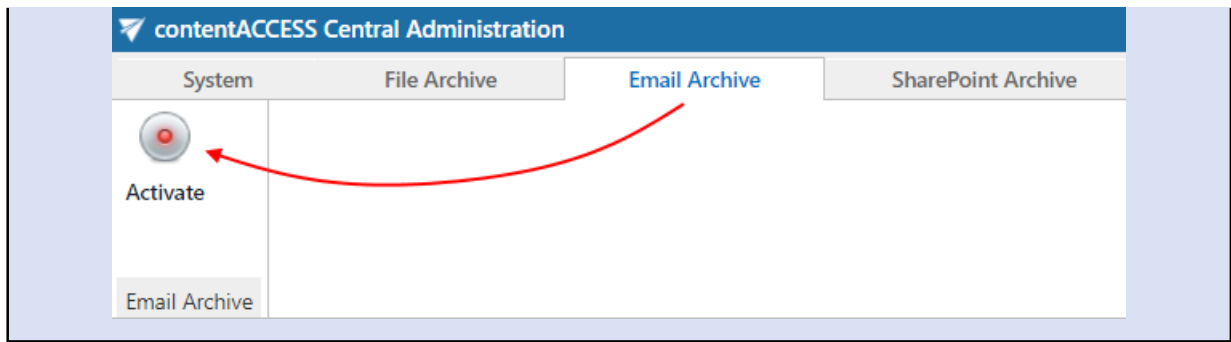
System settings configuration

Open the **System setting** page (by navigating to *Email Archive => Settings => System* **settings** on the ribbon) and fill in: a) "Email Archive DB" under "Database settings", b) "On-premise" under "EWS settings", check the "Keep connection alive" option, and c) set the global exclude rules of email archiving in section "Message class handling". Now, we exclude everything, except "Mails" and "Documents".

Important: The user specified in *Email Archive → System settings→ EWS Settings* must have full access to the mailbox(es) to be archived and must have "View-Only Organization Management" rights on the Exchange. If contentACCESS MailApp will be also used, then *Organization management rights* are required.

Important: In case of **custom Exchange solution(s)** used for emails, the **custom email properties will not be archived!** The recommended solution: do not archive these emails and/or use the exclude feature for these custom message classes.

It is also important to remember the *Message class handling* settings configured on *System settings page* always take priority over the Exclude settings of the Email archiving job.

a) Database settings. Here, the database will be selected, which will be used by the provisioning job. The provisioning job will store the metadata (mailbox permissions, group memberships, etc.) during the synchronization process.

b)   **EWS Settings.** Under **Exchange server type** and **Exchange server interface version,** select the type and version of the Exchange server, where your email provider is running. Then, enter the **EWS URL, PowerShell URL,** and the **PowerShell authentication**. These data must be requested from the email provider you use. The maximum count of connection-attempts can be determined under **Max. connection count** option.

> **Note:** It is possible to configure Microsoft 365 connection here. If you select Microsoft 365 from the **Exchange server type** dropdown list, **Exchange server interface version**, **EWS URL**, **PowerShell URL**, and the **PowerShell authentication** will be automatically pre-filled.

It is also possible to disregard any SSL errors by checking the checkbox **Ignore SSL errors**.

Keep connection alive: If it is turned on, then the connection between contentACCESS and the Exchange server is cached, which increases the performance.

▪ **If you are communicating** with **one Exchange server**: **turn it on** to increase performance.

▪ **If you have a load balanced Exchange environment** and there are **communication problems** with the Exchange: it is recommended to **turn it off.**

Enter the **User name** and **Password** for explicit credentials or log in under service-credentials and run a test connection via **Test** button.

c)   **Message class handling.** This feature can be useful for companies with a policy that certain message classes, e.g., Contacts, Tasks etc. mustn't be archived. Here, you can also exclude custom message classes, e.g., if your company already has emails archived by a product from another company, then you can exclude these already archived emails from the re- archiving process.

Adjusting retentions

Open the **[Retentions](#)** page (*Email Archive => Settings => Retentions)* and **configure the duration of storing the archived item to 10 years**. Click on **+ new** and the "Retention repository" window opens. Name your retention ("10 years") and specify the retention time in months (120).

> *Note:* During this retention time, the archived items will not be deleted from the selected storage. The first created retention will be automatically set to default. This option is configurable on the

Retentions page, using the given retention's context menu. Default retentions are used by the contentACCESS MailApp and OWA integration (by archiving/restoring from the mailbox directly). For more information, refer to the online guide here.

HTML shortcut configuration

Open the Shortcut configuration page (*Email Archive => Settings => Shortcut configuration*). In section HTML Templates, click **+ new** to create a default HTML template file. Name this HTML "Default" (Screenshot A). Then click **+ new** option under the breadcrumbs, and according to Screenshot B, create an HTML shortcut with name "Shortcut HTML". In the repository window, select the "Default" template you created a while ago.

If you check the **Keep formatted body** checkbox, the original HTML body won't be converted to plain text during the shortcut creation, but the system will try to inject the attachment links into the existing message body. A template, which supports this formatting option, must be selected.
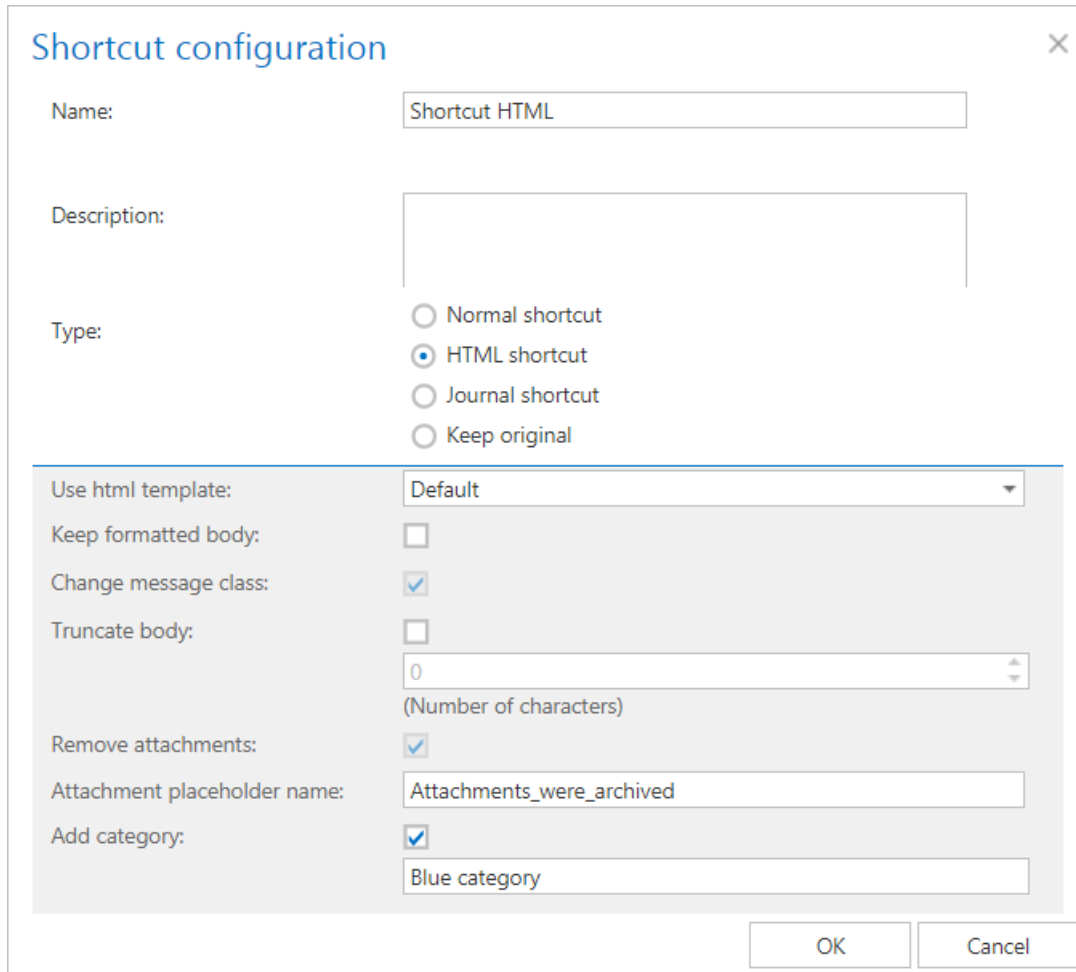
It will also become possible to keep inline attachments in the mail body by checking the **Keep in-line attachments** checkbox.

The **Remove attachments** checkbox is checked by default for HTML shortcuts. Removed attachments can be later retrieved by double clicking on the archived email in the compose window. It is also possible to specify here the **Attachment placeholder name**.

If you want a category to be added to the archived emails, check the **Add category** checkbox and type the name of the category into the textbox.

Screenshot A

techarrow
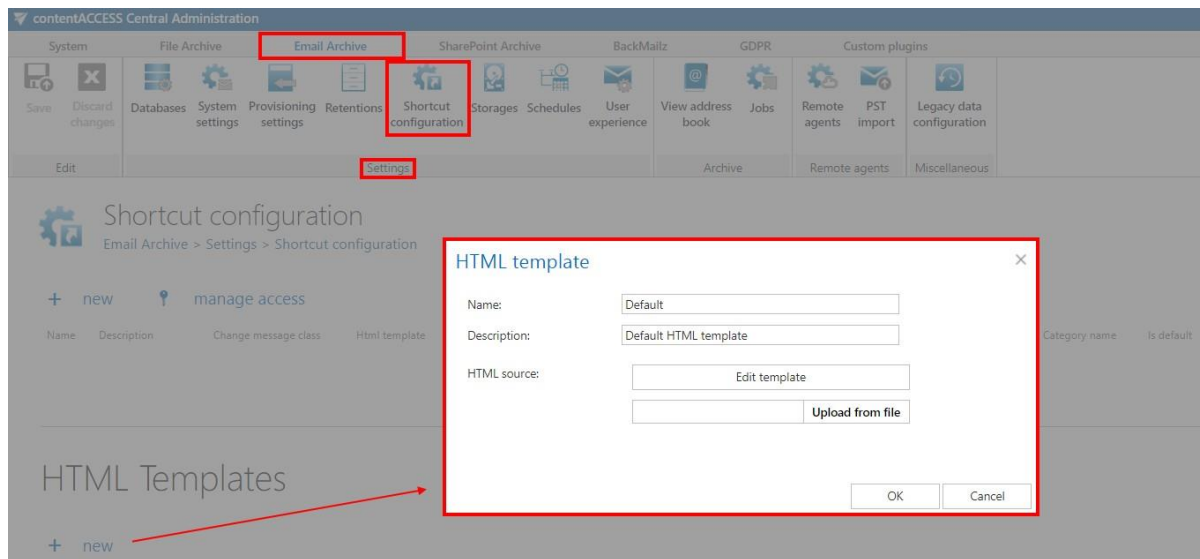
Screenshot B
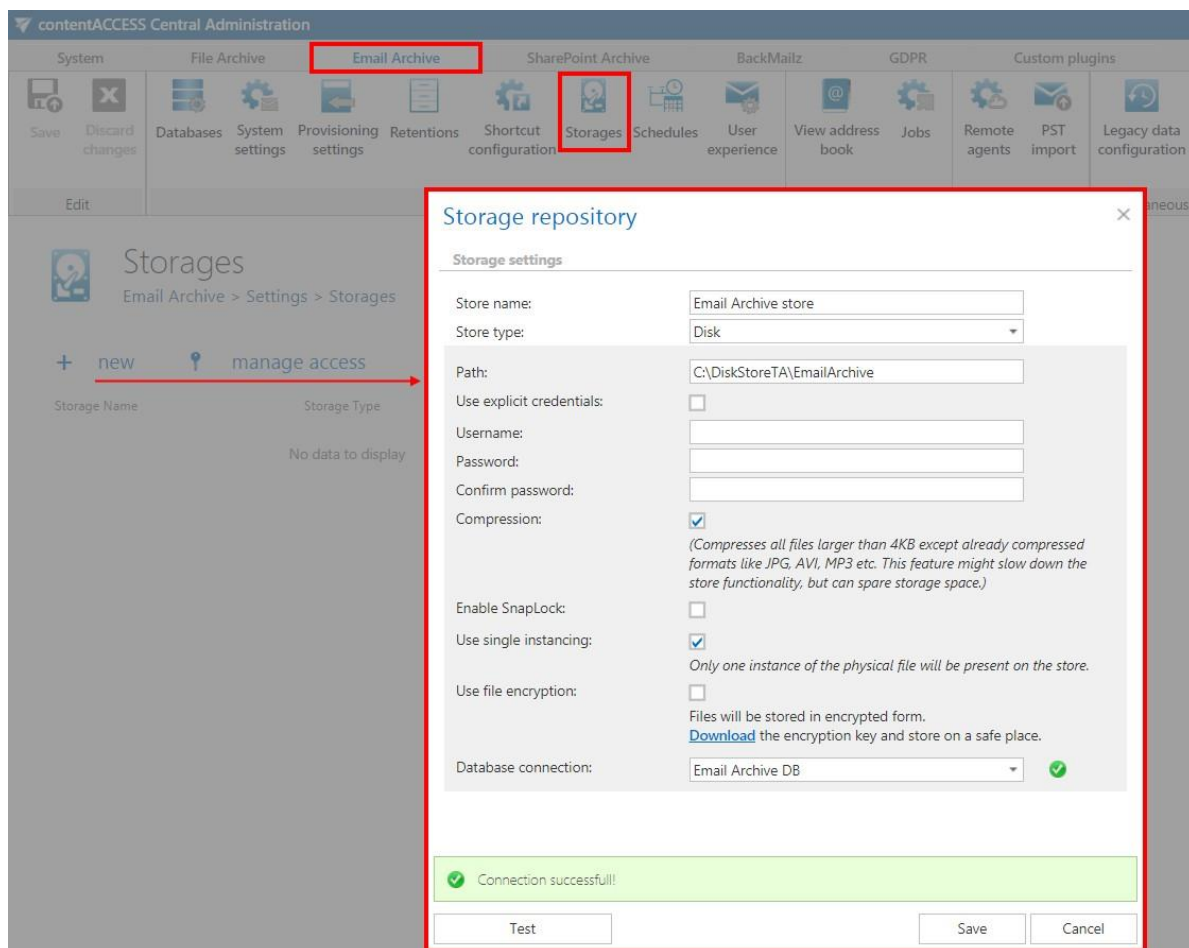
contentACCESS uses shortcuts to retrieve the archived items. It is recommended to use **HTML shortcut** instead of **Normal shortcut**. Normal shortcut works only in MS Outlook if contentACCESS Outlook forms are installed. The HTML shortcut will work with MS Outlook and OWA. If HTML shortcut is used, then the archived email message and the attachment are accessible using the links in the HTML. The archived emails will be marked with the category you select under 'Add category' option. Thanks to the **Remove attachments** option, the attachments will be stored separately from the body of the email message.

The shortcut created, at first, is automatically set to default. Default shortcut is used by manual archiving by the contentACCESS MailApp and OWA integration. The option is configurable on the **Shortcut configuration** page using the context menu.

The HTML Templates are free for editing. The size of a HTML-template cannot exceed **32 kB**; for this reason, it is not recommended to use MS Word for editing HTML templates.

Creating a disk store with name "Email Archive store"

Open the **Storages** page (*Email archive => Settings =>Storages)*, click **+ new** and fill the required values in the "Storage repository" window. Name it ("Email Archive store"), select the "Disk" type store, and fill in the path: "C:\DiskStoreTA\EmailArchive". By checking the **Compression** checkbox, all files larger than 4 kilobytes will be compressed, except of already compressed file formats such as JPG, MP3 etc. If you wish to keep only one copy of the physical file in the store, check the **Use single instancing** checkbox. **Compression** and **Single instancing** can spare a lot of space in the database. Under Database connection, select single **"Email Archive DB"** that we created in Step 4 above.

The location you set here will be used to store your binaries. If an explicit user will not be selected for this storage, then the contentACCESS service user will be used automatically. The store must also have **enough free space** and the data on it must be secured against data loss (a backup must be created).

The storage created first will be set to default automatically.

tech**arrow**

Configuring the provisioning and archive scheduler

Configure 2 schedulers, one for the provisioning job, the second for the email archive job. Open the [**Schedules**](#) page (*Email Archive => Settings =>Schedules*) and click option **+ new**. Create a "Week schedule with repeat" scheduler with name "Every day at 6 PM". Check the **Run only once per interval** checkbox (we do not need to run the archiving multiple times in the interval). Fix the start time of the email archive job running in the week map (Screenshot A).
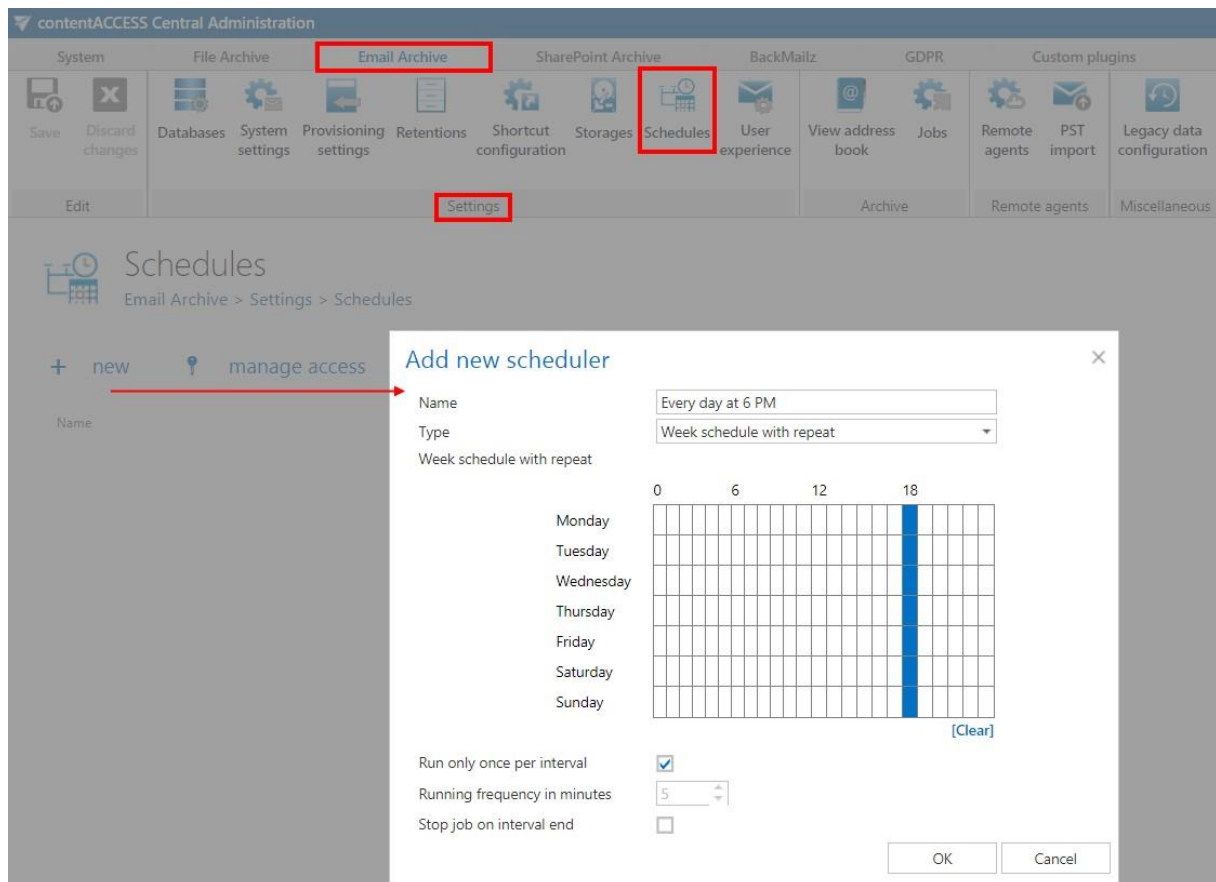
Screenshot A

For the provisioning scheduler: on the same page, create a "Week schedule with repeat" scheduler with name "Provisioning scheduler". Check the **Run only once per interval** checkbox (we do not need to run the job multiple times in the interval). Fix the start time 1 AM – of the provisioning running in the week map (Screenshot B).

Screenshot B

Provisioning settings

Open the **Provisioning settings** page on the ribbon (*Email Archive => Settings => Provisioning settings)*. Configure the provisioning as follows: a) set the "Provisioning scheduler" from the list,

b) under Notification settings, choose "Never" from the dropdown list (monitoring job will check potential system and job misconfigurations), and c) under Exchange groups, click **+ select** and add "Archiving" to the list of objects to be provisioned.

> *Important!!!* The provisioning job synchronizes the Exchange with the contentACCESS Address book. Therefore, the **Provisioning job must run before the archiving process is started**. We configured a daily scheduler starting at 1 AM for the provisioning job and a daily scheduler from 6 PM for the email archiving job. It is also a solution to start the provisioning job manually to ensure all the selected objects are synchronized into contentACCESS.

The provisioning job is an elementary job to one or more email archive jobs, which should extract Exchange groups and subgroups and synchronize these changes and mailbox permissions with contentACCESS. If the whole Exchange server is selected, then it is redundant to select Exchange groups forprocessing.

Under **Notification settings**, you can also set where and when to send emails about the potential errors, warnings that might occur during the running of the synchronization. It is recommended to send these emails when errors occur, but if you use the monitoring feature, then this step can

techarrow

be disregarded. If you use notifications, then it must be first configured in System => Services => Notifications.

Notification settings

Send when:	Never
Recipient list:	john.doe@example.com,peter.white@example.com,

Under **Archive settings**, check the **Assign to mailbox** checkbox. 3 dropdowns and the **Create contentACCESS Portal user** checkbox will appear. With the dropdowns it is possible to select the default **Database**, **Store** and **Index zone** for the mailboxes. During the run of the provisioning job, if there are new mailboxes provisioned, the selected **Database**, **Store** and **Index zone** will be automatically assigned to them. Check also the **Create contentACCESS Portal user** checkbox (this must be enabled to automatically give access to end users). Thanks to this, a contentACCESS user will be created with the selected **Mailbox owner role** in case that **Create contentACCESS Portal access** is set to **Inherit** in the **Address book** for that mailbox **AND**:

- the mailbox is **not** in any group and **Create contentACCESS Portal access is not set to Forbidden** for the server in Address book, **OR**
- for all the groups the mailbox is a member of, the **Create contentACCESS Portal access** is set to **Inherit** in Address book and **Create contentACCESS Portal access is not set to Forbidden** for the server in Address book, **OR**
- for at least one of the groups the mailbox is a member of, the **Create contentACCESS Portal access** is set to **Enable**

Assign to mailbox:	☑
Database connection:	Email Archive DB ✓
Storage name:	Email Archive store
Index zone:	Default index zone
Create contentWEB user:	☑

Creating a Content indexing job

Now we create a **Content indexing** job, which will assign the selected index zone to emails of the selected entities.

a) On the page of **Indexing** (*System → Services -> Indexing*) click on **+ new** option under **Indexing**.
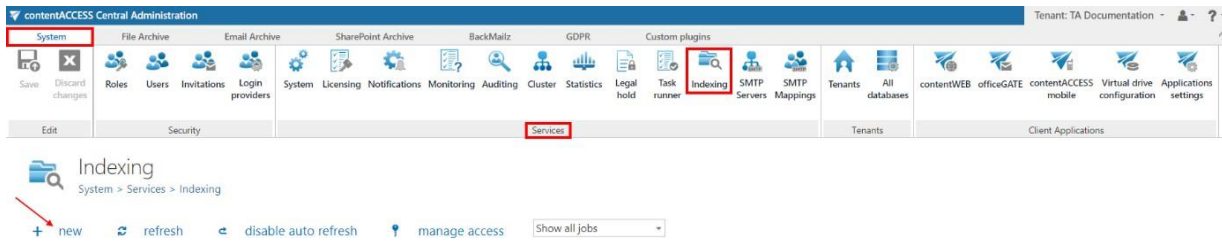
b) **Name** the indexing job (in our case, EmailArchive indexing) and select the node it will run on. Click on **Add**. When the system asks if you would like to configure this instance now, select **Yes**.

The administrator must configure these sections on the job's configuration page:

c) **Scheduling settings** – run every 5 minutes

d) **Resource settings** – specify the number "4"

e) **Filtering settings** – set here the file types that should and shouldn't be processed

f) **Entities to index** – select the entities that will be processed by the job (in our case the EmailArchive group)

g) **Index zone settings** - set the index zone that will be assigned to items by the indexing job (in our case, we will use the default one)

h) **Notification settings** - you can set where and when to send emails about the potential errors

Creating an Email archive job

Now we create an Email Archive job.

a) On the page of **Jobs** (*Email Archive => Archive => Jobs*) click **+ new** option.

b) Enter the **Display name** (Daily mailbox archiver) into the **Add new job instance** window and click **Add**.

techarrow

c) Click **OK** in the pop-up window to open the job's configuration page and configure it.

The administrator must configure these sections on the job's configuration page: **Database settings**, **Scheduling settings**, **Storage settings**, **Shortcutting**, **Retention settings**, **Address book objects to process**, **Include folders**. Filtering settings are optional and allow to set specific criteria for the email archiving (age, size, emails to be excluded etc.).

d) Set "Email Archive DB" that we created in step 4 above. The metadata will be stored here.

e) Select the "Every day at 6 PM" scheduler we created in step 10 above.

f) Select the "Email Archive store" storage we configured in step 9 above.

g) In "Shortcut settings", select "Shortcut HTML" that we configured in step 8 above. By clicking on the **Change** button(s), it is also possible to set filters for the archiving job (more information here).

> *Note:* In the same section, you may set age and size filters that should be used by the archive job. If everything needs to be archived (like in this use case), you can skip these filters.
>
> The shortcut settings allow to combine more shortcut types. If you want to archive your messages, but need to keep them in the mailbox during a certain time (e.g., 6 months) and replace

them with a shortcut (e.g., HTML), after this time, you can combine the "Keep original" and a second, e.g., HTML shortcut. The screenshot below shows how to combine shortcuts for this use case (for more information, refer to the online guide [here](#)):

h) Select **retention** "10 years", which we configured in step 7.

i) In the next step, select **Archiving** group to be archived. Click **+ select** option in section "Address book objects to process" and search for the desired group in the dialog.

techarrow

Address book objects to process

Mailbox selection ✕

↻ refresh

		Name	Type	Email address
☑		Archiving		
☑	👥	Archiving	Group	

☑ ⌖ Contains([Name], 'Archiving')   Clear

OK   Cancel

j) Select the mailbox folder that must be archived in the "Include folders" section. By default, the job is set to process the entire mailbox. Now, we will archive only subfolder "All documents" on the INBOX root, so we set it here. (Refer to the picture below.)

k) The "Exclude folders" section allows to set folders to be excluded from the archiving.

> *Important:* Exclude rules, defined in **System settings** in step 6 above, take priority over the **Item filtering** settings of the **Archive job**.

l) Under "Notification settings", select "Never", as we will use the monitoring feature to check the details of possible system or job-related failures.

> *Note*: If you wish to get notification emails if any errors/warning occur during the archiving process, you can set it here. Specify, to be informed about errors, warnings or both, and enter the email address where these notification emails must be sent.

m) Under "Resource settings", set value "4".

> *Note:* This value determines the number of simultaneously processed items by the archive job.

Saving the changes

Save your changes.

The job can be started also manually by clicking on "start immediately" in the header bar of its configuration page:

Archive job events

Now, we will check the "Archive job" events to see if everything has run properly. Open the **System logs** page (*System => Logs and auditing => System logs),* using the "logs" option in the job's header bar. Under "Events", you can see the job has finished the task successfully. Use the

Export function to export the events into an HTML, XML or CSV file format:

Email archive's Addressbook

The Archive job will assign the "Email Archive DB" database and "Email Archive store" during its first run. These configurations can be checked on the page of View address book (*Email Archive => Archive =>View address book)*, in the corresponding column of the archived mailboxes in the **Exchange mailboxes** tab:

Enjoy it!

Enjoy it! Your users can now log in to contentACCESS Portal and access their emails.