

Workshare Protect Server

Email Data Loss Prevention

Table of Contents

Chapter 1: Introducing Workshare Protect Server	4
What is Workshare Protect Server?	5
Workshare Protect Server Functionality	5
Securing external email	5
Workshare Protect Routing Agent	6
Profiles determine cleaning and conversion	7
Rule-based email blocking	7
Policies determine access	7
Additional functionality	8
Web console for configuration	9
Chapter 2: Creating Policies with the Policy Editor	10
Introducing the Policy Editor	11
Accessing the Policy Editor	11
Creating Policies	12
Defining rules	14
Grouping rules	18
Configuring document identification	19
Activating policies	20
Blocking Emails	21
Monitoring blocked emails	21
Appendix A. Regular Expressions	23
Regular Expressions and Workshare Protect Server	24
Character classes	24
Character escapes	25
Anchors	25
Grouping constructs	26
Quantifiers	27
Backreference constructs	27

Alternation constructs	28
Substitutions.....	28
Miscellaneous	28

Chapter 1: Introducing Workshare Protect Server

This chapter introduces Workshare Protect Server, providing an overview of how it works as well as a summary of the key features and benefits. It includes the following sections:

- **What is Workshare Protect Server?**, page 5, introduces Protect Server.
- **Workshare Protect Server Functionality**, page 5, describes the different functionality available with Protect Server.

What is Workshare Protect Server?

Workshare Protect Server provides server-side metadata cleaning and document processing. Protect Server processes all emails passing through the corporate mail server, including those that originate from webmail and mobile mail clients. In corporate email scenarios, email is always routed through the corporate email server - and consequently Protect Server - ensuring complete protection. By locating this processing effort on the server, email send performance on the originating device is not impacted, and users are not affected.

Protect Server is a mail gateway that removes metadata from Microsoft Office attachments (Word, Excel and PowerPoint) as well as PDF attachments. It can also automatically convert Microsoft Office attachments to PDF. A web application - the Workshare Protect Server web console - is provided to enable administrators to configure which metadata elements to remove and view a history of what was previously removed.

Additionally, Protect Server can include the Policy Editor which provides a new policy framework to determine whether an email should be sent or bounced. It provides for control at the matter level through policies that associate email whitelists with matter-IDs.

Workshare Protect Server Functionality

Workshare Protect Server can secure attachments before they are sent by cleaning them of metadata or converting to PDF. Additionally, Protect Server can block emails and prevent them from being sent at all. An overview of the functionality is provided in the following sections.

Securing external email

Protect Server processes emails leaving an organization according to the profile applied to the email. This processing could be cleaning the attachments by removing metadata from them or converting the attachments to PDF, or both.

Where emails are sent to both internal and external recipients and the Workshare Protect Routing Agent is **not** installed, the internal recipient will receive the unprocessed document and the external recipient will receive a processed version. When the Protect Routing Agent **is** installed, both internal recipients and external recipients will receive a processed version.

Protect Server processes Microsoft Office and PDF attachments. The following file formats can be processed:

File format	Cleans	Converts to PDF or PDF/A
Microsoft Word 97 or later (DOC, DOT, DOCX, DOTX, DOCM, DOTM)	✓	✓
Microsoft Excel 97 or later (XLS, XLT, XLSX, XLTX)	✓	✓
Microsoft PowerPoint 97 or later (PPT, POT, PPTX, POTX)	✓	✓
RTF	✓	✓
Word 2003 XML	✓	✓
Open Document Text (both ODT 1.1 and 1.2 are cleaned, but files saved in ODT 1.1)	✓	✓
PDF	✓	✗

Protect Server processes the following types of attachments:

- Password-protected attachments (when the Workshare Protect Portal is installed).

Note: *Protect Server does not process password-protected PowerPoint files.*

- The contents of archive (ZIP) attachments.
- Attachments of embedded emails.
- Attachments to meeting requests and other Microsoft Exchange-specific features, such as polls and forms.

Note: *Protect Server does not process digitally signed documents or corrupt documents and does not check the attachments of digitally signed emails.*

Workshare Protect Routing Agent

With the installation of the Protect Routing Agent, Protect Server processes attachments of emails that are sent to both internal and external recipients. When an email has a relevant attachment and includes both external and internal recipients, the Protect Routing Agent will ensure that both the internal recipient and the external recipient receive the same processed version of the attachment.

The installation and setup of the Protect Routing Agent is described in the *Workshare Protect Routing Agent Admin Guide*.

Profiles determine cleaning and conversion

Profiles specify what metadata to remove from an email attachment and whether to convert the attachment to PDF. Every profile has an email address and this is how Protect Server determines which profile to apply to any given email. When a sender adds the email address of a profile as a recipient in an email then this profile will be used to process the email.

When a sender does not specify a profile email address in an email, the following occurs:

- If you have enabled the Active Directory cache feature, then Protect Server will look to see what AD group the sender belongs to and apply whichever profile you have defined for that group. If no profile has been defined for the group, then Protect Server will apply whichever profile you defined as the Default profile.
- If you have **not** enabled the Active Directory cache feature, then Protect Server will apply whichever profile you have defined as the Default profile.

If the sender specifies more than one profile email address in an email, then Protect Server will apply whichever profile you have defined as the Fallback profile.

Users allocated an Administrator role can create and manage multiple metadata cleaning/PDF conversion profiles. Configuring profiles is described in the *Workshare Protect Server Metadata Removal guide*.

Rule-based email blocking

Protect Server can prevent emails from being sent, based on business rules. The decision to block emails is driven by policies configured in the Protect Server Policy Editor. The Policy Editor provides for control at the matter level through policies that associate email groups with client engagements.

Policies determine access

You can set a policy that tells Protect Server to look for specific custom properties in the email attachments together with specific email addresses. If found, the email will be bounced.

The Policy Editor supports both whitelisting and blacklisting with policies that ensure access to certain information is to pre-approved recipients only as well as policies that ensure certain documents are never sent to untrustworthy recipients.

This granular approach to policies ensures confidential files pertaining to confidential matters are not accidentally sent to the wrong people.

This functionality can be enabled during or after installation (described in the *Workshare Protect Server Installation guide*). Configuring policies is described in this guide.

Additional functionality

Additional functionality can be configured on the Protect Server web console.

- **Clean reports for senders**

Senders receive a clean receipt email with a Clean Report PDF attached providing details of the exact metadata cleaned from the document. The clean receipt email can also include the original email and the processed attachments. Administrators configure clean receipts on the Protect Server web console.

- **Bouncing emails**

Protect Server can be configured to prevent emails with attachments that include comments or track changes or that cannot be processed from being delivered. There are several reasons why Protect Server may not be able to process an attachment. For example, the attachment may be corrupt or digitally signed. When Protect Server is configured this way, it bounces the email back to the sender with a non-delivery report. Administrators configure which emails to bounce on the Protect Server web console.

- **Preview functionality for senders**

Protect Server provides previews of the cleaned/converted attachments to the sender. The sender can request a preview of what the processed attachments will look like before sending them to the recipients. This is done by sending an email to a profile email address only. Protect Server will treat such an email as a preview request and send the processed attachments back to the sender.

- **Synchronization**

You can configure synchronization so that where any email attachment is processed by Protect Server, the original copy of the email found in the sender's "Sent Items" folder will be updated with the processed attachments. This update occurs for email destined for external recipients only as well as emails destined for both internal and external recipients.

Additionally, to ensure that internal recipients always have access to the same version of attachments that are received by external recipients, where emails are sent internally and externally, the internal recipients can receive a clean receipt with the processed attachment included.

Web console for configuration

The functionality available in the Protect Server web console depends on the type of user:

- Users allocated an **Administrator** role can do the following:
 - View information about the performance and current health of Protect Server. For example, whether Protect Server services are up and running, whether the database is connected, whether the Protect Server license is expired and details of any emails queued on Protect Server.
 - Search through all emails processed by Protect Server.
 - Configure profiles.
 - Define policies (when the Policy Editor is enabled)
 - Specify Protect Server configuration settings, such as whether clean reports should be sent, how to override cleaning settings as well as configuring alert settings and email templates.
- Users allocated a **Business** role can do the following:
 - Access detailed statistics on the activities of Protect Server. For example, how many emails were processed using each profile and how many emails were processed in Microsoft Word format.
 - Search through and view emails that they themselves have sent through Protect Server. They cannot search or view emails sent by other users.
 - View profiles configured for Protect Server. They cannot create new profiles or modify or delete existing profiles.
- Users allocated a **User** role can do the following:
 - Search through and view emails that they themselves have sent through Protect Server. They cannot search or view emails sent by other users.
 - View profiles configured for Protect Server. They cannot create new profiles or modify or delete existing profiles.

Note: Roles are allocated during installation.

The web console and configuration of Protect Server is described in the *Workshare Protect Server Installation guide*.

Chapter 2: Creating Policies with the Policy Editor

This chapter introduces the Policy Editor, an add-on to Workshare Protect Server. It includes the following sections:

- **Introducing the Policy Editor**, page 11, introduces Policy Editor and describes how to install and access it.
- **Creating Policies**, page 12, describes how to create policies using the Policy Editor.
- **Blocking Emails**, page 21, describes how to monitor which emails have been blocked by policies.

Introducing the Policy Editor

The Policy Editor provides a new policy framework to determine whether an email should be sent or bounced. It provides for control at the matter level through policies that associate email whitelists with matter-IDs.

You can set a policy that tells Workshare Protect Server to look for specific custom properties in the email attachments together with specific email addresses. If found, the email will be bounced.

The Policy Editor supports both whitelisting and blacklisting with policies that ensure access to certain information is to pre-approved recipients only as well as policies that ensure certain documents are never sent to untrustworthy recipients.

This granular approach to policies ensures confidential files pertaining to confidential matter are not accidentally sent to the wrong people.

Policies are applied first before profiles for cleaning or converting attachments. Thus, if an email is bounced by a Protect Server policy, it will not be cleaned of metadata or converted to PDF by a Protect Server profile.

Accessing the Policy Editor

The Policy Editor is accessed (by a user with an Administrator role) via the Protect Server web console which runs in a browser. It is recommended to use the latest version of Google Chrome or Mozilla Firefox.

To access the Policy Editor:

1. Open a web browser and enter `http://[URL to Protect Server website]` in the address bar. Generally this will be in the format:

`http://<server-name/ip-address>.<domain>/protect.`

The login page is displayed. In the example below, the IP address 10.10.9.69 was used, with the URL formed as **`http://10.10.9.69/protect.`**



2. Enter your login credentials and click **OK**. The **Messages** page of the Protect Server web console is displayed.
3. Select the **Settings** page.
4. Click **Policy** in the left menu. The Policy Editor is displayed on the right.

Creating Policies

Workshare Protect Server checks all emails by looking at the email headers (for the addresses the email is sent to and the address the email is sent from) and the attachment contents. In particular, Protect Server looks at custom properties and the header, footer and body of the attachment.

If an email matches an active policy, Protect Server will block it and notify the sender of the action.

Using the Policy Editor, you can create multiple policies and select to activate them as necessary. Policies can detail the people or things considered to be acceptable or trustworthy (whitelist) or conversely, they can detail the people or groups regarded as unacceptable or untrustworthy (blacklist).

Workshare provides two policies to illustrate how to blacklist and whitelist. By default, both of these policies are inactive.

A policy is made up of one or more rules. A rule is a statement about the sender or recipients, or custom properties, or content in the headers, footers or body of the attachment. This statement may be required to be true or false.

For example, a rule could state that the recipient email address ends with “mintonslaw.com”. If this is set to true then when an email is sent to any email address with the domain “mintonslaw.com, it will be blocked. If the rule is set to false, then when an email is sent to any email address including “mintonslaw.com, it will be sent and all other emails will be blocked.

Defined rules are grouped and a policy can include several groups of rules. An “and” or “or” operator connects the rules within a group and connects the groups.

To create a new policy:

1. In the Policy Editor, click **Add new policy**.

The screenshot shows the 'Policy builder' interface. At the top left, there is a back arrow and the text 'Back to policies Policy builder'. The form contains several fields: 'Policy name:' with an empty text input; 'Description:' with a larger text area; and 'Remediation:' with another text area. Below these is an 'Is enabled:' checkbox, which is currently unchecked. A note states: 'Emails meeting the following conditions will be blocked and bounced back to the sender.' Below this note is a dropdown menu with the placeholder text 'Select a rule' and a downward arrow. To the right of the dropdown is a small information icon and the text 'Drag and drop rules for reordering and creating groups'. At the bottom right of the form are two buttons: 'Cancel' and 'Save policy'.

2. In the **Policy name** field, add a name for the policy. This is the name that users see when an email is bounced because it triggers the policy.
3. In the **Description** field, describe the scope and purpose of the policy. Again, this is the explanation that users see when they are informed that their email has triggered a policy so it should be relevant and helpful.
4. In the **Remediation** field, enter advice for users or action they can take when their email has been blocked by a policy. This information is displayed to the user in the notification email they receive when an email has been blocked by the policy.
5. Select the **Is enabled** checkbox if you want the policy to be active once created.
6. Define rules for your policy. Refer to [Defining rules](#).

7. Group the rules in your policy and set the relationship between them. Refer to [Grouping rules](#).
8. Click **Save policy**. Your policy is saved and appears in your list of policies. If you activated it during the creating process, it will be active immediately. If not, you can activate it at a later time. Refer to [Activating policies](#).

Note: *It may take a minute for a newly activated policy to become live in your environment.*

Defining rules

A policy is made up of one or more rules. A rule describes a condition and Protect Server checks to see if an email and its attachments match that condition. Where a policy has multiple rules, there may be an “and” or an “or” operation that links them. This means that an email and its attachments must match ALL or ANY of the rules in a policy in order for the policy to be triggered and an email to be bounced.

A rule is a statement about recipients or about custom properties in the attachment or about content in the headers, footers or body of the attachment that may be required to be true or false.

To define email address rules:

1. From the rule dropdown list, select the type of rule to add to your policy:
 - **Recipient email address:** This rule is a statement about who an email is being sent to – the email address. For example, recipient equals elizabeth.morris@mintonlaw.com. Another example, recipient contains palmerpartners.com.
 - **Sender’s email address:** This rule is a statement about who is sending an email – the email address. For example, sender equals elizabeth.morris@mintonlaw.com. Another example, sender contains palmerpartners.com.
 - **File includes custom property:** This rule is a statement about the value of a custom property within an attachment to an email. For example, a file that contains a “text” type custom property called “company” that equals “Palma”. Another example, a file that contains a “number” type custom property called “docID” that is greater than “4000”.
 - **File includes a header that:** This rule is a statement about the content in the header of an email attachment.
 - **File includes a footer that:** This rule is a statement about the content in the footer of an email attachment.
 - **Document body:** This rule is a statement about the content in the body of an email attachment.

2. If you selected a **Recipient email address** or **Sender's email address** rule, select the required qualifier:
 - **Equals:** The email address exactly matches the value.
 - **Contains:** The email address includes the value.
 - **Starts with:** The email address starts with the exact value.
 - **Ends with:** The email address ends with the exact value.
 - **Starts with (case insensitive):** The email address starts with the value (ignoring the case).
 - **Ends with (case insensitive):** The email address ends with the value (ignoring the case).
 - **Equals (case insensitive):** The email address exactly matches the value (ignoring the case).
 - **Includes the regular expression:** The email address includes the regular expression specified in the value field.
3. Enter the value.
4. Select whether the condition you have defined should be met or not by selecting **Condition true** or **Condition false** from the dropdown list above the rule.
 For example, if your condition is recipient email address equals elizabeth.morris@mintonslaw.com and you select **Condition true**, then only emails to elizabeth.morris@mintonslaw.com will be blocked. If you select **Condition false**, then all emails will be blocked except those to elizabeth.morris@mintonslaw.com.
5. Click **confirm**.

Recipient email address is "elizabeth.morris@mintonslaw.com"

[Add rule below](#) | [Edit](#) | [Delete](#)

To define custom property rules:

1. If you selected a **File includes custom property** rule, enter the name of the custom property.
2. Select the type of custom property – **Boolean**, **Number**, **Text** or **Date**. This refers to the type of value the property could have.
3. Select the required qualifier. This will vary according to the type selected.

When **Boolean** is selected, only **Equals** is available as the qualifier.

When **Number** is selected:

- **Equals:** The custom property exactly matches the number value.
- **Ceiling equals:** Rounding the custom property up to the nearest whole number matches the number value.
- **Floor equals:** Rounding the custom property down to the nearest whole number matches the number value.
- **Greater than:** The custom property is more than the number value.

- **Greater than or equal:** The custom property is more than or the same as the number value.
- **Less than:** The custom property is less than the number value.
- **Less than or equal:** The custom property is less than or the same as the number value.
- **Round equals:** Rounding the custom property to the nearest whole number (up or down) matches the number value.

When **Text** is selected:

- **Equals:** The custom property exactly matches the text value.
- **Contains:** The custom property includes the text value.
- **Starts with:** The custom property starts with the exact text value.
- **Ends with:** The custom property ends with the exact text value.
- **Starts with (case insensitive):** The custom property starts with the text value (ignoring the case).
- **Ends with (case insensitive):** The custom property ends with the text value (ignoring the case).
- **Equals (case insensitive):** The custom property exactly matches the text value (ignoring the case).
- **Includes the regular expression:** The custom property includes a regular expression that includes the text value.

When **Date** is selected:

- **Equals:** The custom property exactly matches the date value.
- **Less than:** The custom property is less than the date value.
- **Greater than:** The custom property is more than the date value.
- **Less than or equal:** The custom property is less than or the same as the date value.
- **Greater than or equal:** The custom property is more than or the same as the date value.

4. Enter the value.
5. Select whether the condition you have defined should be met or not by selecting **Condition true** or **Condition false** from the dropdown list above the rule.

For example, if your condition is that a file contains a “text” type custom property called “company” that equals “Palma” and you select **Condition true**, then only emails with attachments that include a custom property with the value Palma will be blocked. If you select **Condition false**, then all emails with attachments will be blocked except those with attachments that include a custom property with the value Palma.

6. Click **confirm**.

File doesn't include custom property "company" that is "Palma"

[Add rule below](#) | [Edit](#) | [Delete](#)

To define header, footer, body rules:

1. If you selected a **File includes a header that** or **File includes a footer that** or **Document body** rule, select the required qualifier:
 - **Equals:** The header/footer/body of the attachment exactly matches the value.
 - **Contains:** The header/footer/body of the attachment includes the value.
 - **Starts with:** The header/footer/body of the attachment starts with the exact value.
 - **Ends with:** The header/footer/body of the attachment ends with the exact value.
 - **Starts with (case insensitive):** The header/footer/body of the attachment starts with the value (ignoring the case).
 - **Ends with (case insensitive):** The header/footer/body of the attachment ends with the value (ignoring the case).
 - **Equals (case insensitive):** The header/footer/body of the attachment exactly matches the value (ignoring the case).
 - **Includes the regular expression:** The header/footer/body of the attachment includes the regular expression specified in the value field.
2. Enter the value.
3. Select whether the condition you have defined should be met or not by selecting **Condition true** or **Condition false** from the dropdown list above the rule.

For example, if your condition is that the file includes a header that contains “private and confidential” and you select **Condition true**, then only emails with attachments with “private and confidential” in the header will be blocked. If you select **Condition false**, then all emails with attachments will be blocked except those with attachments that have “private and confidential” in the header.

4. Click **confirm**.

File includes a header that contains "private and confidential"

[Add rule below](#) | [Edit](#) | [Delete](#)

In order to add further rules to your policy, click **Add rule below** and define further rules. When a rule is added below, it is automatically added to the same group as the rule above. You can easily drag rules from one group to another, as described in [Grouping rules](#).

Use the **Edit** and **Delete** links to the right of a rule to edit it or delete it from the policy.

Grouping rules

Rules are created in groups and a policy can have multiple groups, each one containing multiple rules. Within a group, each rule is connected by an “and” or “or” operation. Additionally, the entire group can be inverted (much like each rule can be) by being set to false. In a group set to false, all the rules in the group must NOT be matched in order for the policy to be met and the email blocked.

In the above example, there are three groups defined in the policy as follows:

- **Group 1:** This group has 2 rules connected by “or”. The rules specify emails sent to elizabeth.morris@mintonlaw.com or jeremy.philips@mintonlaw.com. But since the group condition is set to **false**, this group is matched by emails sent to ANY email addresses except elizabeth.morris@mintonlaw.com or jeremy.philips@mintonlaw.com.
- **Group 2:** This group has 2 rules connected by “and”. The rules specify documents with two particular custom properties will match this group. Since the group condition is set to **true**, only documents that match both rules will match the group.
- Group 1 is connected to Group 2 by “and” so both groups must be matched.
- **Group 3:** This group includes groups 1 and 2 and has a group condition of true. This means that the conditions in both groups 1 and 2 must be matched.

Once this policy is active, if a document with custom properties **company=Palma** and **isconfidential=true** is sent to sarah.smith@mintonslaw.com, it will be blocked. If it is sent to elizabeth.morris@mintonslaw.com or jeremy.philips@mintonslaw.com, it will be sent.

To move rules:

Simply drag and drop the rules between groups as required. You must set a connection between rules in a group – either “and” or “or”.

To create a new group:

Drag a rule below the existing group and a new group is created. You must set a connection between groups – either “and” or “or”. You must also specify the condition of each group - whether the rules in it must be matched (**true**) or not matched (**false**).

Configuring document identification

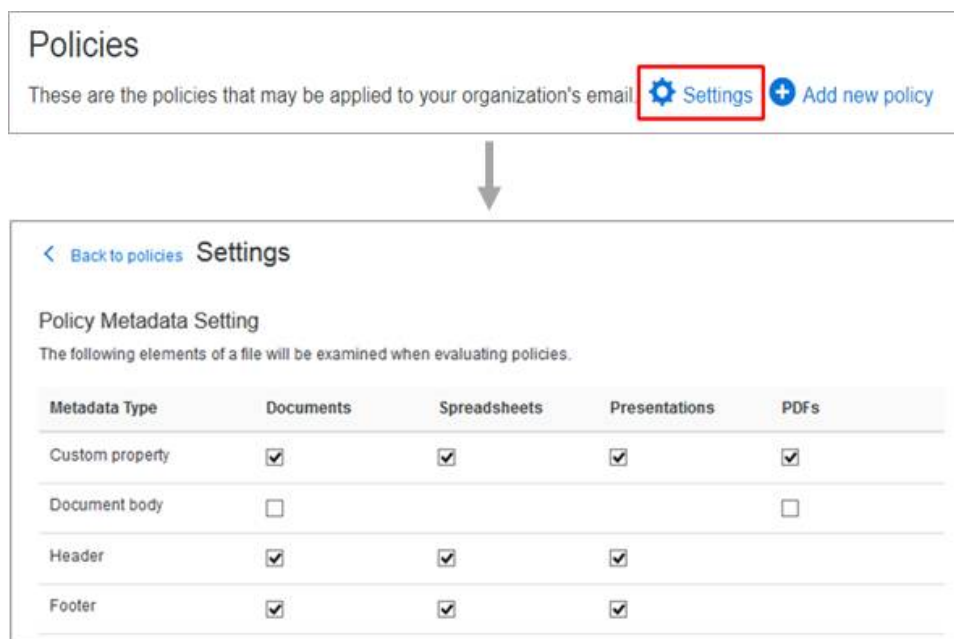
Documents are marked as being of interest using markers within them. Workshare Protect Server looks through the following sets of data (in increasing order of complexity):

- Custom properties
- Headers and footers
- The body of the document

The settings screen can be used to configure which document elements will be checked for each document type.

To configure document identification:

1. In the Policy Editor, click **Settings**.



- For each file format, select which areas you want Workshare Protect Server to check.
- Click **Save**.

Activating policies

Policies need to be active in order for Protect Server to apply them. You can see in your list of policies whether a policy is active or not.

Policies			
These are the policies that may be applied to your organization's email. Settings + Add new policy			
Name	Description	Is active	
Recipient not present on iManage SPM Client Whitelist	The matter files you have attempted to share on email belong to JP Morgan and the recipients are not on the approved list to receive the file. Please only send the information individuals on the whitelist as maintained in iManage SPM. To update settings either contact the Matter Administrator or login directly at https://10.90.4.71:8080/admin/login	<input checked="" type="checkbox"/>	Edit Delete
Recipient present on iManage SPM Client Blacklist	The matter files you have attempted to share on email belong to JP Morgan and the recipients are not on the approved list to receive the file. Please only send the information individuals on the whitelist as maintained in iManage SPM. To update settings either contact the Matter Administrator or login directly at https://10.90.4.71:8080/admin/login	<input checked="" type="checkbox"/>	Edit Delete
Sample Whitelist Policy	Demonstration of a white list. Emails will be bounced to sender if: 1. email contains an attachment where 'COMPANY=Palma Investments' and 'CONFIDENTIAL=TRUE' 2.	<input type="checkbox"/>	Edit Delete

You can activate a policy from this list by selecting the **Is active** checkbox. The policy is activated straightaway without need for confirmation.

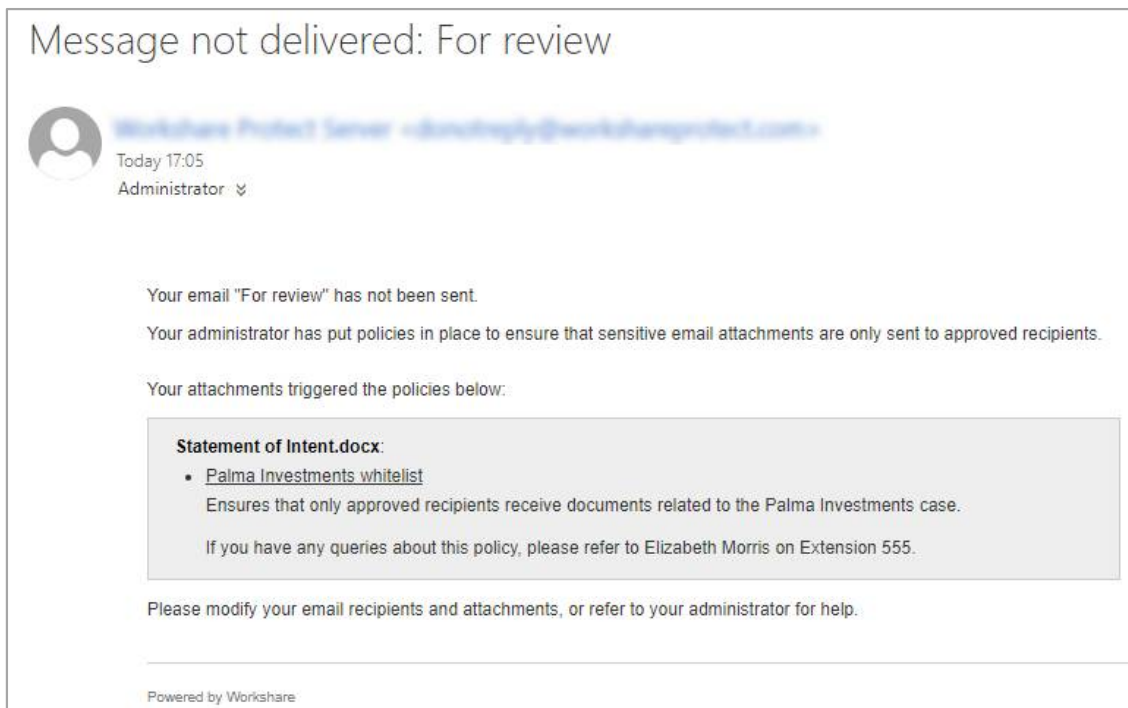
You can also create an active policy by selecting the **Is enabled** checkbox when creating it. If you do this, the policy becomes active when you create and save the policy.

Note: It may take a minute for a newly activated policy to become live in your environment.

Blocking Emails

Workshare Protect Server checks all emails by looking at the email headers (for the addresses the email is sent to and the address the email is sent from) and the attachment contents. In particular, Protect Server looks at custom properties and the header, footer and body of the attachment.

When a policy is matched or triggered, Protect Server will block the email and bounce it back to the sender. The sender will receive notification that the email has not been sent. For example:



This email provides information about the policies that have been triggered and the sender can click the policy link and view the read-only details about the policy. The email can also include useful advice as to what the sender can do. For example, who to contact.

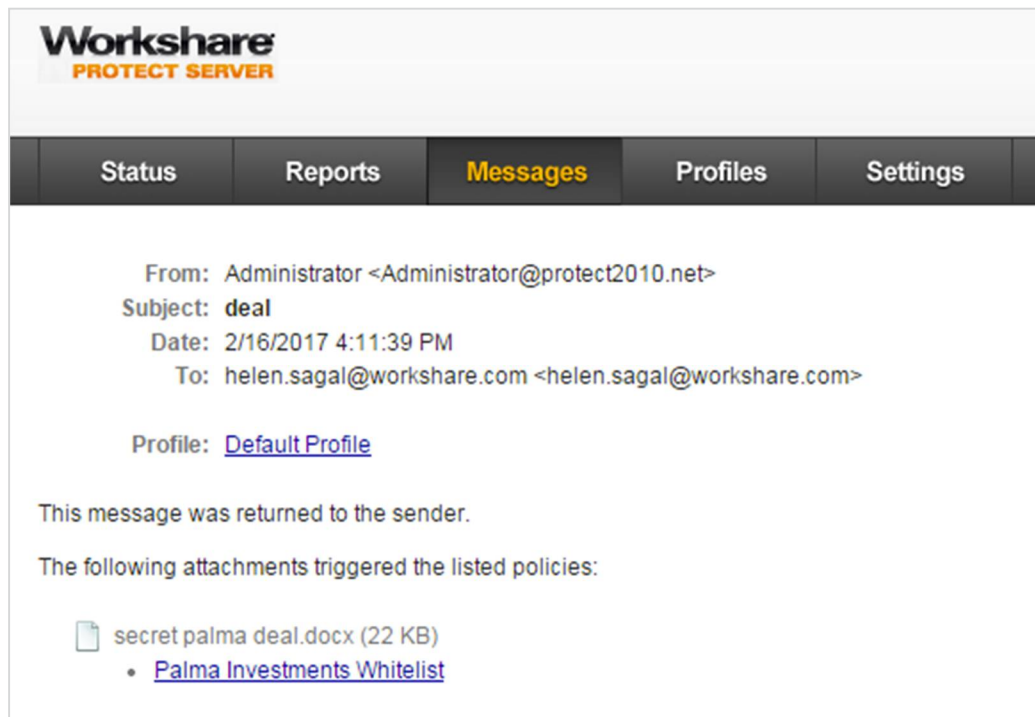
Monitoring blocked emails

As an administrator, you can monitor the blocked emails from the Protect Server web console. The **Messages** tab in the Protect Server web console provides information about all the emails processed by Protect Server.

To review message logs:

1. Log into the Protect Server web console and select **Messages**.
2. Search for the email that's been blocked.

3. Click the **Subject** of the email to display further details. For example:



The screenshot shows the Workshare interface with the 'Messages' tab selected. The email details are as follows:

Workshare
PROTECT SERVER


Status Reports **Messages** Profiles Settings

From: Administrator <Administrator@protect2010.net>
Subject: **deal**
Date: 2/16/2017 4:11:39 PM
To: helen.sagal@workshare.com <helen.sagal@workshare.com>

Profile: [Default Profile](#)

This message was returned to the sender.

The following attachments triggered the listed policies:

-  secret palma deal.docx (22 KB)
 - [Palma Investments Whitelist](#)

You can now see which policy was triggered and you can click the policy link and view the read-only details about the policy.

Appendix A. Regular Expressions

This appendix describes the definition syntax for regular expressions.

Regular Expressions and Workshare Protect Server

Microsoft.Net regular expressions can be used in policies and Workshare Protect Server will search for the specified string or pattern in the email address, or within the attachment. Further information can be found at <http://regexstorm.net/reference>.

Character classes

	Definition
[characters]	Matches any character found in <code>characters</code> .
[^characters]	Matches any character not found in <code>characters</code> .
[first-last]	Matches any character in the range of characters from <code>first</code> to <code>last</code> .
.	Wildcard. Matches any character except <code>\n</code> .
\p{category}	Matches any character in a category of unicode characters, specified by <code>category</code> .
\P{category}	Matches any character not in a category of unicode characters, specified by <code>category</code> .
\w	Matches any letter, decimal digit, or an underscore.
\W	Matches any character except a letter, decimal digit, or an underscore.
\s	Matches any white-space character.
\S	Matches any character except a white-space character.
\d	Matches any decimal digit.
\D	Matches any character except a decimal digit.

Character escapes

	Definition
<code>\r</code>	Matches a carriage return.
<code>\n</code>	Matches a newline.
<code>\t</code>	Matches a tab.
<code>[\b]</code>	Matches a backspace. Note that it must be enclosed in brackets to have this meaning.
<code>\f</code>	Matches a form feed.
<code>\e</code>	Matches an escape.
<code>\v</code>	Matches a vertical tab.
<code>\a</code>	Matches the bell character.
<code>\octal</code>	Matches a character, where <code>octal</code> is the octal representation of that character.
<code>\x hex</code>	Matches a character, where <code>hex</code> is the two digit hexadecimal representation of that character.
<code>\u unicode</code>	Matches a unicode character, where <code>unicode</code> is the four digit hexadecimal representation of that unicode character.
<code>\c character</code>	Matches an ASCII control character specified by <code>character</code> .

Anchors

	Definition
<code>^</code>	Matches the beginning of the input.
<code>\$</code>	Matches the end of the input, or the point before a final <code>\n</code> at the end of the input.
<code>\A</code>	Matches the beginning of the input. Identical to <code>^</code> , except it is unaffected by the multiline option.
<code>\Z</code>	Matches the end of the input, or the point before a final <code>\n</code> at the end of the input. Identical to <code>\$</code> , except it is unaffected by the multiline option.
<code>\z</code>	Matches the end of the input, without exception.
<code>\G</code>	Matches the point that the previous match ended. Used to find contiguous matches.
<code>\b</code>	Matches any word boundary. Specifically, any point between a <code>\w</code> and a <code>\W</code> .

	Definition
<code>\B</code>	Matches any point that is not a word boundary. Specifically, any point not between a <code>\w</code> and a <code>\W</code> .

Grouping constructs

	Definition
<code>(subpattern)</code>	Captures <code>subpattern</code> as an unnamed group.
<code>(?<name>subpattern)</code>	Captures <code>subpattern</code> as a named group specified by <code>name</code> .
<code>(?<name-previous>subpattern)</code>	Balancing group definition. This allows nested constructs to be matched, such as parentheses or HTML tags. The previously defined group to balance against is specified by <code>previous</code> . Captures <code>subpattern</code> as a named group specified by <code>name</code> , or <code>name</code> can be omitted to capture as an unnamed group. For more information, check out Morten Maate's tutorial on matching nested constructs.
<code>(?:subpattern)</code>	Noncapturing group. Allows the use of parentheses without <code>subpattern</code> being captured into a group.
<code>(?enabled-disabled:subpattern)</code>	Allows <code>subpattern</code> to be matched with different options than the rest of the pattern. Any inline option characters in <code>enabled</code> or <code>disabled</code> will enable or disable specific options, respectively.
<code>(?=subpattern)</code>	Zero-width positive lookahead assertion. Continues matching only if <code>subpattern</code> matches on the right.
<code>(?!subpattern)</code>	Zero-width negative lookahead assertion. Continues matching only if <code>subpattern</code> does not match on the right.
<code>(?<=subpattern)</code>	Zero-width positive lookbehind assertion. Continues matching only if <code>subpattern</code> matches on the left.
<code>(?<!=subpattern)</code>	Zero-width negative lookbehind assertion. Continues matching only if <code>subpattern</code> does not match on the left.

	Definition
<code>(?>subpattern)</code>	Prevents backtracking over <code>subpattern</code> , which can improve performance.

Quantifiers

	Definition
<code>*</code>	Matches previous element zero or more times.
<code>+</code>	Matches previous element one or more times.
<code>?</code>	Matches previous element zero or one times.
<code>{n}</code>	Matches previous element exactly <code>n</code> times.
<code>{n,}</code>	Matches previous element at least <code>n</code> times.
<code>{n,m}</code>	Matches previous element at least <code>n</code> times and at most <code>m</code> times.
<code>*?</code>	Matches previous element zero or more times, but as few times as possible.
<code>+?</code>	Matches previous element one or more times, but as few times as possible.
<code>??</code>	Matches previous element zero or one times, but as few times as possible.
<code>{n,}??</code>	Matches previous element at least <code>n</code> times, but as few times as possible.
<code>{n,m}??</code>	Matches previous element at least <code>n</code> times and at most <code>m</code> times, but as few times as possible.

Backreference constructs

	Definition
<code>\number</code>	Matches the value of a previously captured group, specified by <code>number</code> .
<code>\k<name></code>	Matches the value of a previously captured named group, specified by <code>name</code> .

Alternation constructs

	Definition
	Functions as a logical or. Matches any elements that it separates.
(?(subpattern)yes no)	Treats <code>subpattern</code> as a zero-width assertion to check if it matches. If so, attempts to match with the <code>yes</code> subpattern. Otherwise, tries the <code>no</code> subpattern. The <code> no</code> part is optional.
(?(group)yes no)	Checks if a previously defined group was successfully captured, specified by <code>group</code> , which can be a number or a name for a named group. If so, attempts to match with the <code>yes</code> subpattern. Otherwise, tries the <code>no</code> subpattern. The <code> no</code> part is optional.

Substitutions

	Definition
<code>\$number</code>	Substitutes the value of a group, specified by <code>number</code> .
<code>\${name}</code>	Substitutes the value of a named group, specified by <code>name</code> .
<code>\$\$</code>	Substitutes the <code>\$</code> character.
<code>\$&</code>	Substitutes the entire match.
<code>\$`</code>	Substitutes all input text found before the match.
<code>\$'</code>	Substitutes all input text found after the match.
<code>\$+</code>	Substitutes the last group that was captured.
<code>\$_</code>	Substitutes the entire input.

Miscellaneous

	Definition
(?enabled-disabled)	Changes options in the middle of a pattern. Any inline option characters in <code>enabled</code> or <code>disabled</code> will enable or disable specific options, respectively.
(?# comment)	Inline comment, not evaluated as part of the pattern.
# comment	End of line comment, not evaluated as part of the pattern. The Ignore Whitespace option must be enabled to use this.

 Workshare Ltd.

© 2018. Workshare Ltd. All rights reserved.

Copyright

Workshare Professional and Workshare DeltaView are registered trademarks of Workshare Ltd. Workshare Compare, Workshare Protect, Workshare 3, Workshare DeltaServer, SafetyGain, and the Workshare logo are trademarks of Workshare Ltd. All other trademarks are those of their respective holders.

Trademarked names may appear throughout this guide. Instead of listing these here or inserting numerous trademark symbols, Workshare wishes to state categorically that no infringement of intellectual or other copyright is intended and that trademarks are used only for editorial purposes.

Disclaimer

The authors/publishers of this guide and any associated help material have used their best efforts to ensure accuracy and effectiveness. Due to the continuing nature of software development, it may be necessary to distribute updated help from time to time. The authors would like to assure users of their continued best efforts in supplying the most effective help material possible.

The authors/publishers, however, make no warranty of any kind, expressed or implied, with regard to Workshare programs or help material associated with them, including this guide. The authors/publishers shall not be liable in the event of incidental or consequential damages in connection with, or arising out of, the programs or associated help instructions.

Revisions

Published for Workshare Protect Server 3.10: 29/06/18

Published for Workshare Protect Server 3.11: 24/08/18

Sold under a license for U.S. Patent Nos. 7,895,276 and 8,060,575 and 8,977,697.

Workshare Ltd., 20 Fashion Street, London E1 6PX www.workshare.com