DJIGZO EMAIL ENCRYPTION

DJIGZO DLP Setup Guide

May 29, 2012, Rev: 5445

Contents

1   Introduction

Data Leak Prevention (DLP) is a feature that prevents certain information to leave the organization via email. What information this is, is defined in the configuration of the DLP system. Typically, it includes credit card numbers, bank account numbers, excessive amounts of email addresses or other personal information in one email message, etc. DLP is implemented as a filter on outgoing email. DLP can be a separate system or product, or it can be integrated with another email related product or system. DJIGZO has integrated DLP with our Email Encryption Gateway.

DLP can monitor email at various levels:

- email body content

- email headers

- email attachments of various types

- nested attachments of various types

DJIGZO DLP currently filters email bodies, attachments and nested attachments of type text, html, xml and other text-based formats. Filtering attachments of type pdf, doc, xls etc. will be part of a future offering of DJIGZO DLP.

Configuring DLP is done via the DJIGZO Web GUI. You can specify keywords and sentences that outgoing email messages should not contain. More elaborate filtering is achieved via *regular expressions*, a specification format that allows you to specify virtually any combination of characters, words or sentences that should be filtered. Sample regular expression configuration files can be downloaded from our web site. DLP can be configured on three levels, similar to how encryption is configured: at gateway level, at domain level and at individual user level. The latter is useful in specific cases where some users can send out information via email that other users cannot.

2   Express setup

In order to set up DLP for simple keyword filtering, use the following steps. In these steps you will create a filtering pattern and assign that pattern to global settings, so all domains on your djigzo server will filter using that pattern.

1. Log in to your DJIGZO server.

2. Select DLP tab. The title of the page reads *DLP Policy Patterns*.

3. The list of patterns is empty. Click *create pattern*.

4. The first field is the name of the pattern. Type "my first pattern", or any pattern name you find appropriate.

5. Under Reg. Exp. ("Regular Expression") type the following: `credit|card|number`. or any sequence of words, separated by a vertical bar. The vertical bar is

important, it separates the keywords. If you omit the bars, the words will be used as a sentence to filter for. We will cover sentences later, because there's a few caveats there.

6. Under "Description" type a description of your pattern, or leave empty.

7. Under "Threshold", type a number. This number specifies how many times a pattern should occur in a message before it is filtered out. We will type "1", meaning that any occurrence of any of the words will cause a message to be filtered. If you type 6, only messages that contain six occurrences of the pattern will be filtered.

8. Under "Priority", select "Quarantine". This means that if the message is filtered, it will be place in quarantine and will not be sent. The other options will be explained later.

9. Under "Match filter", select nothing. If you select "Mask", every occurrence of a pattern will be replaced by asterisks.

10. Now click "Add". What you now have is one entry in the list of patterns, named "my first filter". The red cross indicates that you can delete the pattern. As soon as the pattern is in use, you cannot delete it without first removing all uses of the pattern.

11. Now, go to the "domains" tab. Select one of your domains, then select "DLP". The title reads "DLP settings for domain: yourdomain.com". The sub title is "DLP patterns". The "Enabled" checkbox is checked, the "inherit" checkbox to the right of that is checked also. This means that the pattern you have entered, applies to this domain.

12. In the "Quarantine URL" field, type the url of your DJIGZO administration interface, followed by "external/dlp/quarantine/view". If your DJIGZO url is "https://djigzo.mydomain.com", the quarantine url should now read "https://djigzo.mydomain.com/external/dlp/quarantine/view". You need to specify this url if you want notification emails that DJIGZO DLP sends to contain a link to manage the quarantined message.

Now go to your email program and type a new message to an email address outside your organization, like a gmail or yahoo email address that you use for testing. The body of the message should contain one or more of the words we used in the pattern, like "credit card number". Send the message. Watch your Inbox. You should receive a message with subject:

`*** DLP quarantine warning ***`

The body of the messages tells you that your email has been put in quarantine because it contained one or more of the keywords that you specified in the pattern. Also, the message contains a link. Click that link. It takes you to a page in the DJIGZO server with title "Quarantined email info". It shows the message, the DLP policies that have been violated, and some action buttons. Press "release" to release the message. Then check your external email address to verify that the message has been released.

In a real world setup, you don't want all users to have access to a release link when a message has been quarantined. In the remainder of this guide we will show you how you can make this link available to system administrators or specific users only. For now, go to the domain tab, select "DLP" and remove the Quarantine URL, by emptying the field. Now, the notification email will not contain a link. You can release messages via the DJIGZO administration interface.

3 Patterns

In order to filter email, you need to specify what you want to filter on. Typically, there are a number of things you filter on. First, most companies don't want outgoing email messages to contain credit card numbers, collections of email addresses, bank account numbers, collections of addresses, and information in some other well defined format. Next, there may be keywords or combinations of keywords that are off limit. If you're a big beverages company, you don't want an employee to email to the outside world the centuries old and well kept secret recipe of your companies strategic product.

If you are familiar with regular expressions, you can skip most of this section. DJIGZO policy patterns are regular expressions, you can enter as many of them as you like in the page under the DLP tab. If you are not, and you think you will want to do more than filter on simple keywords or use predefined policy files from DJIGZO, please read a tutorial about regular expressions. The web site http://www.regular-expressions.info/ provides tutorials and examples of regular expressions. We will explain all regular expressions we will use here, so if you like, you can continue reading and read the turorials later.

Go to the DLP tab of your DJIGZO. If you went through the express setup section, you will find the pattern that you created, otherwise you'll see an empty list titled "DLP Policy Patterns". We will now create a number of sample policy patterns.

Note: The patterns shown in the next section are simplified versions of existing patterns and are only provided as example patterns. For example the matcher for creditcard numbers does not detect numbers with spaces and dashes ('-'). More robust patterns can be downloaded from our website.

3.1 List of keywords

The simplest policy pattern is a single keyword or a list of keywords. You create a simple one keyword pattern by clicking "create pattern" and type in "Reg. Exp" the single keyword. Enter a name, press "Add", done. A pattern with multiple keywords is created by entering keywords, separated by a vertical bar. The vertical bar means "or", so `keyword1|keyword2|keyword3` means that the message is filtered out if any of the keywords is present in the message. The vertical bar is the only separator you can use for this. A space or a comma or a period is not a separator. If you specify `credit card|number`, DJIGZO DLP will filter messages that contain the word "number", and messages that

contain the text "credit card". It will not filter for the words "credit", "card", or "creditcard".

There is one caveat here. DJIGZO has a list of "skip words", words that will not participate in the pattern matching process. The reason for this is that words that occur frequently are not likely to be essential in filtering, while skipping them will increase performance of the system. You can view the list of skip words under the DLP tab, in the "skip list" menu option on the left. You can edit the skip list to match your needs, or you can delete all words to disable the skip words list. Now, if you want to filter for a sentence like "rock the boat", DJIGZO DLP will never find a match, because the word "the" is removed from the message before a pattern is applied. You have to remove the word "the" from your pattern also. The pattern "rock boat" will filter for "rock the boat" as well as for "rock boat". This will almost always be desired behavior. If it's not, however, than you have to remove the word "the" from the skip words list. Then you can filter for "rock the boat" and "rock boat" separately. Please check out the section in this guide about the skip words list. There's another caveat. In order to streamline the policy pattern matching process and to increase performance, DJIGZO DLP normalizes email message text before matching a pattern. The whole message is converted to lower case text, with all subsequent space-like characters replaced by one space. You should not use capital letters in a pattern. For example, if you match for "credit card", every message will be filtered out that contains "credit card", or "Credit Card", or "Credit card", or "creDit cArd", or "credit card". Please check out the section on message normalization in this guide.

3.2 Numbers

A Master Card credit card number has the format `9999 9999 9999 9999` where "9" stands for any digit, any number from 0 to 9. The regular expression for this should read something like "four groups of four numeric characters, separated by dots". A regular expression (or short, "regexp"), for four numbers is `[0-9]{4}`. The regexp for a space is `\s`. The regexp for an MC credit card is thus `[0-9]{4}\s[0-9]{4}\s[0-9]{4}\s[0-9]{4}` or `([0-9]{4}\s){3}[0-9]{4}`. Or, realizing that an MC card number starts with 51 through 55, the regexp is `5[1-5][0-9]{2}(\s[0-9]{4}){3}`. This regexp says "a 5, a number from 1 to 5, two digits, then three groups consisting of a space and four digits".

Similarly, you create a regexp for Visa cards which have 14 digits and start with 4, or American Express which have 12 digits and start with 34 or 37.

3.3 Email addresses

A regexp that matches any email address is (from regular-expressions.info):

`\b[A-Z0-9._%+-]+@[A-Z0-9.-]+\.[A-Z]{2,4}\b`

This regexp doesn't check the top level domain, it matches .com domains but also .con domains or any fictitional tld. If you want to filter messages that contain four or more email addresses, you add this regexp in the pattern page and set "threshold" to 4.

3.4 Sentences

The sentence "the fox jumps over the lazy dog" is matched by

```
the fox jumps over the lazy dog
```

or alternatively,

```
the\sfox\sjumps\sover\sthe\slazy\sdog
```

where \s is a space, if there were no skip words list. By default, the words "the" and "over" are skipped, so we should have
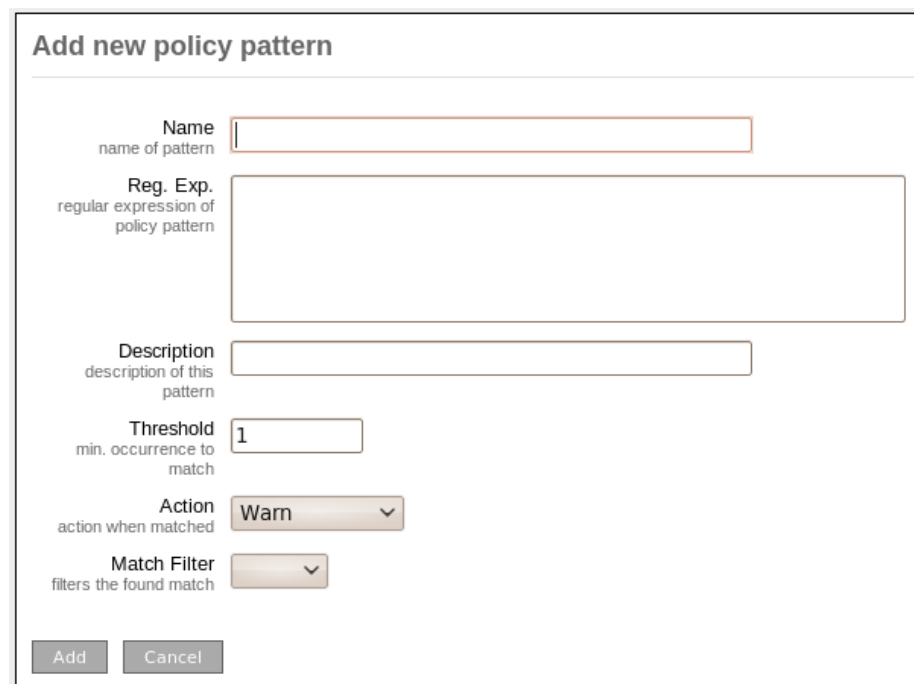
```
fox jumps lazy dog
```

or,

```
fox\sjumps\slazy\sdog
```

You can make it more interesting by putting

```
fox jumps lazy dog|dog jumps lazy fox|lazy fox jumps dog
```

4 Configuring a pattern

Go to DLP, then "create policy pattern". The *Add new policy pattern* page will be shown (see figure 1).

Figure 1: Add new policy pattern

The policy parameters will be briefly explained.

Name   The name for your pattern. This is the name with which it will show up in the pattern list, it needs to be unique.

Reg. exp.   The regular expression, as described in the previous section. Keep in mind that there is a skip word list.

Description   Enter anything that clarifies the pattern or helps you remember why you created it.

Threshold   is the number of times the pattern should occur in a message to be filtered. Default is 1, which means that any occurrence of the pattern will cause the message to be filtered. Entering 5 means that only messages that contain your regexp at least five times will be filtered.

Action   determines the action Djgizo DLP takes when the pattern is matched.

- "Warn" means that the user and the DLP managers get a warning message that the user has sent out a message with sensitive content that's not severe enough for the message to be blocked. The message is sent normally.

- "Must Encrypt" means that if this pattern is detected, the message can go out but it should be encrypted. If DJIGZO can encrypt it, it will encrypt and send the message, otherwise it will notify the user that the message can't be sent because it can't be encrypted.

- "Quarantine" means that DJIGZO DLP will place the message in the quarantine queue. The DLP admin can release the message from the queue, or prevent it from being sent.

- "Block" means that a message matching this pattern will be blocked. The DLP admin cannot change that. The user and the admin will be notified.

If a message violates multiple rules, the actual action taken depends on the action with the highest priority. The priority of the actions are as follows (from highest to lowest priority): *a*) Block *b*) Quarantine *c*) Must encrypt *d*) Warn

Example: A message violates two rules, one rule with *Quarantine* action and another rule with *Block* action. Because the *Block* action has the highest priority, the message will be blocked.

Note: It depends on the DLP notification settings whether the user and/or the DLP managers receive a notification message when one of the policies is violated (see section 7.2 for more information).

Match filter   currently contains one option. If you select "Mask", DJIGZO DLP
will replace the occurrences of the pattern with asterisks in notifications and
bounce messages. This means that if a message is blocked or quarantined,
that the user and the admin get a copy of the original message without the
content that triggered DLP. This way, the blocked credit card numbers do not
get proliferated in notification messages.

Press "Add" to add your pattern to the list of patterns.

4.1   Pattern groups

Multiple patterns can be combined into a pattern group. This makes it easier to
assign multiple patterns to a user or domain or to the global settings. For ex-
ample, suppose there are different patterns for different creditcard companies,
a pattern for Mastercard, a pattern for Visa etc. All the separate creditcard
patterns can be combined into one group named *Credicards*. If outgoing email
should be scanned for creditcard numbers, selecting the *Credicards* group will
be sufficient to scan for all types of credicard numbers. A group can be created
by clicking *create group* on the DLP page. The name of the group has to be
specified and applied before patters can be added to the group.

Note: A group can also contain other groups. By using groups within groups it
makes it easier to create different groups for different departments (for example
the financial department) without having to duplicate any patterns.

4.2   Importing and exporting patterns

DLP patterns can be imported from and exported to xml.

4.2.1   Exporting patterns

Patterns can be exported by opening the DLP patterns page, selecting the
patterns and clicking *export selected* (see figure 2). The selected DLP patterns
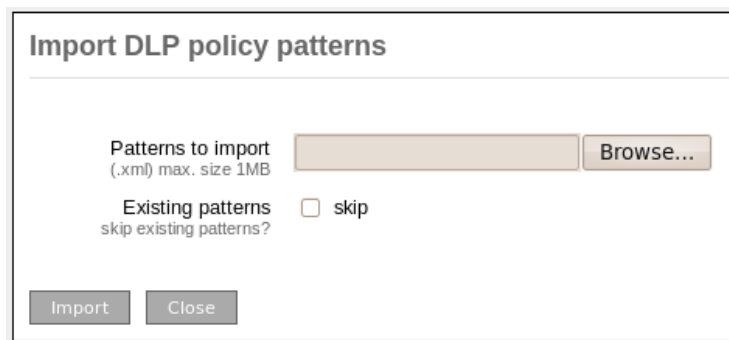will be exported to a downloadable .xml file.

Figure 2: Exporting DLP patterns

4.2.2   Importing patterns

Patterns can be imported from the DLP patterns page by clicking *Import patterns* on the left-hand side menu. On the *Import DLP policy patterns* page (see figure 3), the xml file from which the patterns should be imported can be selected. The name of a pattern should be unique. If a pattern with the imported name already exists and *Existing patterns* is not selected, none of the patterns will be imported. If *Existing patterns* is selected, duplicate patterns will be skipped during import.

Note: You should be careful when importing duplicate patterns with *Existing patterns* enabled. If a pattern group and a pattern that's part of the group is imported, and a pattern with the name already exists the existing pattern will be part of the group after the import (if *Existing patterns* was enabled). If the existing pattern was different from the pattern in the xml file, the group might use a different pattern (i.e., different regexp) with a similar name than what was expected.

Figure 3: Importing DLP patterns

5   Skip list

To improve scanning of messages, certain words are removed from the extracted text before scanning. The list of words that are removed can be edited by clicking *Sip list* on the left-hand side menu of the DLP page. The default list of words to skip is a list of the top 100 mostly used English words (see figure 4). New words can be added are removed. If no word should ever be skipped, all words can be removed.

5.1   Text normalization

To make writing patterns easier, all extracted text is normalized using the following procedure:

1. All carriage returns and line feeds are replaced with spaces.

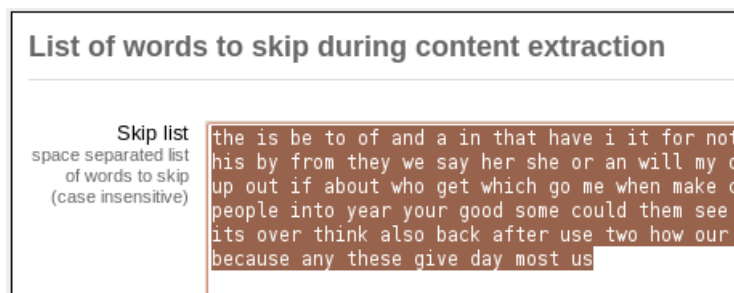2. Consecutive spaces are trimmed to one space.

List of words to skip during content extraction

Skip list
space separated list
of words to skip
(case insensitive)

the is be to of and a in that have i it for not
his by from they we say her she or an will my o
up out if about who get which go me when make c
people into year your good some could them see
its over think also back after use two how our
because any these give day most us

Figure 4: Skip list

3. All characters are converted to lowercase.

4. The resulting text is Unicode normalized (NFC).

Note: Because all text is converted to lowercase, any literal text used in one of the patters should be written in lowercase. For example, if the text to match is "this is converted to LOWERCASE", the pattern matching the word "lowercase" should be written in lowercase capitals.

6   Selecting patterns

Patterns are not automatically used. Patterns are only used if they are explicitly selected. A pattern can be selected for the global settings, for a domain or for a user. Patterns can be inherited just like any other setting. Domains inherit from the global settings and users inherit from the domain settings. A DLP pattern can be selected for the global settings, for a domain or for a user by opening the settings page for the global settings, the domain or the user and clicking the *DLP* link. The DLP settings page will be opened (see 6). The DLP patterns can be selected by clicking the *DLP patterns* link. The selected patterns page will be opened (see figure 5).
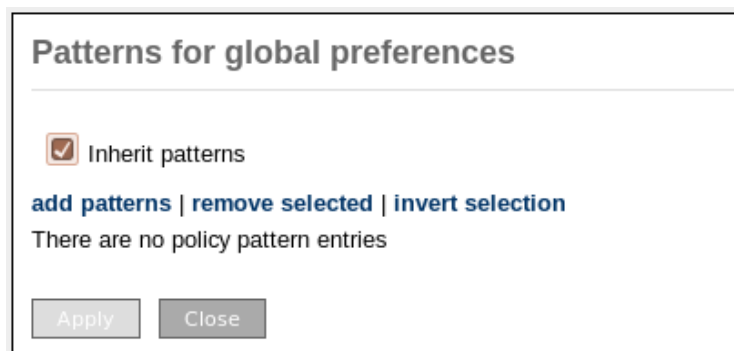

Patterns for global preferences

☑ Inherit patterns

add patterns | **remove selected** | **invert selection**
There are no policy pattern entries

Apply   Close

Figure 5: Selected DLP patterns

11

With "add patterns", new patterns can be added to the domain. Before new patterns are being added, all inherited patterns will be copied and patterns will no longer inherit from the upper level DLP settings, i.e., if new patterns are added to the upper level settings, those pattern will no longer be inherited.

Note: if you uncheck "inherit" and press *Apply*, all inherited patterns will be copied to the list of domain patterns and the domain will no longer inherit patterns from the global settings.

If you want to add a pattern to this domain that is not defined yet, you have to go back to the page where you define patterns, under *DLP main tab*, and create the pattern before you can add it to your domain. Select the patterns that you want to apply to this domain and press "Add Selected".

What if you have a pattern that you want to give different actions for different domains? Like, the credit card pattern should block for one domain but quarantine for another. The way to do this is to create two separate patterns, with the same regexp but different names and different actions, and apply one of the patterns to each domain.

7   DLP settings

The DLP settings can be specified for a user, a domain or for the global settings. The DLP settings can be edited by clicking the *DLP* link on the settings page. DLP settings can be inherited just like any other settings (see figure 6). By unchecking the "inherit" checkbox, inherited settings can be overridden. A brief introduction of the most important DLP settings will be given.

7.1   General settings

Enable pattern scanning   If set, DLP scanning is enabled for this particular user, domain or for the global settings. This is a sender and receiver property. This means that if the sender has DLP enabled, but the recipient has DLP disabled, email sent to that recipient will not be content scanned. Typically, DLP is enabled for the global settings to make sure that email sent to any recipient will be content scanned.

Quarantine URL   The quarantine url is the url via which the quarantine queue is visible to release individual messages. If a message is quarantined, the sender and/or DLP managers receive a notification. If the quarantine URL is provided, the message contains a link, via the *quarantine URL*, that takes the sender to a page with options, to release, or download, or block the message. The actual URL to use, depends on whether a firewall or proxy is used to access the gateway. For example, if the gateway is directly accessible on "https://djigzo.mydomain.com", the quarantine URL should be:

```
https://djigzo.mydomain.com/external/dlp/quarantine/view
```

Figure 6: DLP Settings

The default *Quarantine URL* is based on the portal *Base URL* (see DJIGZO administration guide). It is therefore advised to change the *Base URL* of the portal and only change the *Quarantine URL* if the quarantine service runs separately from the portal.

Note: The quarantined email will receive a randomly generated number which uniquely identifies the quarantined email. The randomly generated identifier

will be added to the quarantine URL. This ensures that the final URL cannot be guessed.

DLP Managers   The DLP managers, is a comma separated list of email addresses to which DLP notifications will be sent when a policy is violated. Whether or not a notification is actually sent to the DLP managers when a specific rule is violated, depends on the notification settings (see below).

Quarantine on failed encryption   If set and a message cannot be encrypted while encryption is mandatory, the email will be quarantined.

Quarantine on error   If a message is corrupt, it can happen that the message cannot be scanned and the scanning routine fails. If the scanning routine fails and cannot locally handle the error, and *Quarantine on error* is true, the message will be quarantined.

7.2   Notification settings

If a policy is violated, the gateway can send a notification message to the originator of the message and/or to the DLP managers. For every violation, there is a setting that determines whether a notification will be sent to the originator of the message (i.e., the sender) and a setting that determines whether a notification will be sent to the DLP managers (see figure 6). For example if *Quarantine to Originator* and *Quarantine to DLP managers* is enabled, a notification of the violation will be sent to the sender of the message and to the DLP managers. If both the notification settings for a violation action are disabled, no notification will be sent.

The following policy violation notifications are available: *a*) Warning to originator; *b*) Warning to DLP Managers; *c*) Quarantine to originator; *d*) Quarantine to DLP Managers; *e*) Block to originator; *f*) Block to DLP Managers; *g*) Error to originator; and *h*) Error to DLP Managers.

Besides the eight policy violation notifications, there are six notification messages which are sent when the message is released or deleted from quarantine or when a quarantined message expires. See section 7.3 for more info on the quarantine actions.

Note: If your mail system does not contain mail forwards, it is ok to leave all notifications on. However, if some email addresses have mail forwarding to an external email address, you need to be careful. If an external user sends an email to such an address, DJIGZO sees the forwarded message as an outgoing message and puts it through DLP. If the message hits a pattern, and the notification setting to originator is on, the outside party sending the email gets a notification about the message being quarantined by DJIGZO DLP. This outside party doesn't know what that means and shouldn't receive these messages. You are advised to leave the notifications to originator off for the global settings and only enable them for your internal domains.

7.3 Authorizations

If a message has been quarantined and the sender or one the DLP managers clicks on the link from the quarantine notification message, the quarantine email information page will be opened in the web browser (see figure 7). Besides some basic information as to why the email was quarantined, the quarantine email information page also allows the user to take certain actions regarding the quarantined email. The following actions can be taken: *a*) Allow download; *b*) Allow release; *c*) Allow release encrypt; *d*) Allow release as-is or, *e*) Allow delete. Whether or not a user is allowed to execute an action depends on whether the user is authorized for that particular action (see Authorizations in figure 6). The five authorization will briefly be discussed.

Allow download   If enabled, the user is authorized to download the message. The message will be downloaded in *.eml* format. The message can be opened with an email client for further inspection.

Allow release   If enabled, the user is allowed to release the message from quarantine. Whether or not the message is encrypted depends on the gateway settings.

Allow release encrypt   If enabled, the user is allowed to release the message from quarantine. The difference between *Allow release encrypt* and *Allow release* is that with *Allow release encrypt* the message is released with encryption forced. If the message cannot be encrypted, the message will not be sent and the originator will receive a notification that the message has not been sent.

Allow release as-is   If enabled, the user is allowed to release the message from quarantine. The difference between *Allow release as-is* and *Allow release* is that with *Allow release as-is* the message is immediately delivered without further processing.

Allow delete   If enabled, the user is allowed to delete the message from the quarantine.

Note: the authorizations of the quarantined email information page, are taken from the authorizations of the originator of the message. If a DLP manager receives a notification message and clicks on the quarantine information link, the allowed actions are based on the authorizations of the original sender of the message and not the authorizations of the DLP manager. If the DLP manager wishes to execute an action for which the originator was not authorized, the DLP manager has to login to the Web GUI and manage the quarantined message directly from the quarantine queue.

15

Figure 7: Quarantined email information

8   Quarantine

Any administrator with the correct roles can manage the DLP queue. The DLP queue can be opened by selecting *Queues* in the main menu of the Web GUI and then selecting *DLP quarantine*. The DLP queue contains all quarantined items ordered from oldest to newest (see figure 8).

Figure 8: Quarantine

The administrator can download, delete and release the message from the quarantine. Whether or not a notification will be sent when the message is deleted or released depends on the notification settings (see 7.2).

8.1   Expire

If a message from the quarantine is not deleted or released within 5 days, the message will expire and the expire notification will be sent.

16

9   Notification templates

The DLP specific notification messages are based on a configurable message template. The message template can be modified by clicking the settings (either global, domain or user settings) and selecting *templates*. The following templates can be edited: *a*) DLP warning; *b*) DLP quarantine; *c*) DLP block; *d*) DLP error; *e*) DLP release notification; *f*) DLP delete notification; and *g*) DLP expire notification (see figure 9)
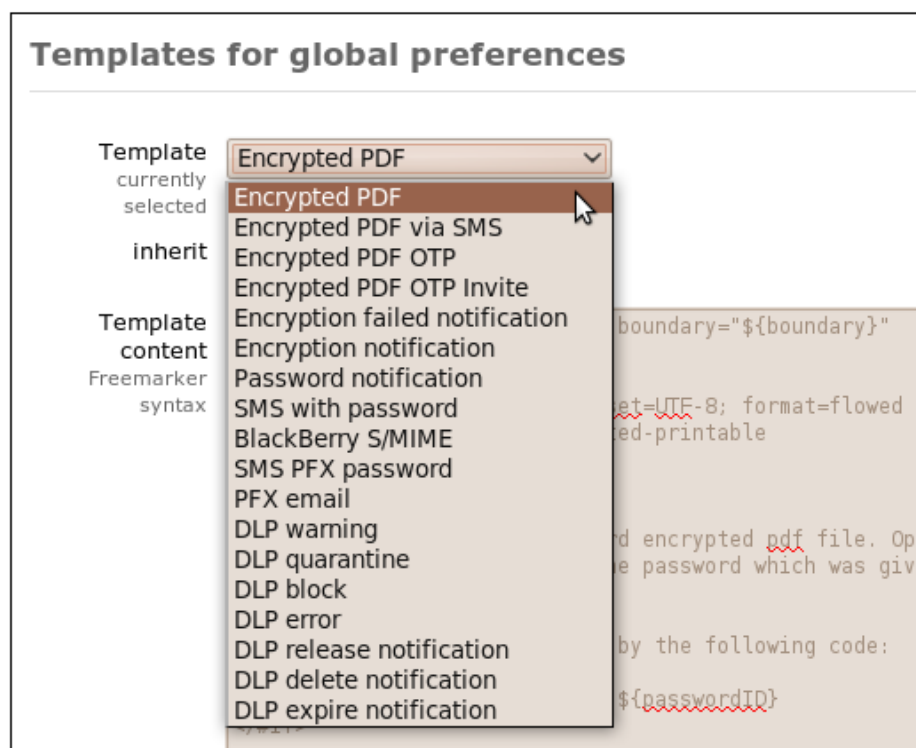

Figure 9: Templates

17