

Box Classifier Guide

UM643301

November 18



Gold
Microsoft Partner



© Boldon James Ltd. All rights reserved.

Customer Documentation

This document is for informational purposes only, and Boldon James cannot guarantee the precision of any information supplied.
BOLDON JAMES MAKES NO WARRANTIES, EXPRESS OR IMPLIED, IN THIS DOCUMENT.

Contents

1	Introduction	3
2	Installation and Deployment options	3
2.1	Initial Configuration of the Box Enterprise	3
2.1.1	<i>Create a Developer Account</i>	3
2.1.2	<i>Set up Two-Factor Authentication</i>	3
2.1.3	<i>Create the Box Classifier App</i>	3
2.1.3.1	Basic Configuration	4
2.1.3.2	Public Key Management	4
2.2	Authorise Box Classifier for Your Enterprise	5
2.3	Create the Box Metadata Classification Policy.....	5
2.4	Installation of Box Classifier	6
2.5	Configuration of Box Classifier	7
2.6	Policy control	9
2.7	The Box Classifier Service	9
3	Functionality	10
3.1	Events and Actions.....	10
3.2	Uploading a document to Box.com	11
3.3	Uploading a new version of a document to Box.com	11
3.4	Synchronisation of documents	11
3.5	Adjusting the Classification of a document in Box.com.....	11
3.6	Unlocking a document in Box.com	11
3.7	Copying a document in Box.com.....	12
3.8	Restoring a document from Trash in Box.com	12
3.9	Preventing documents with specific Classifier labels being retained in Box.com	12
3.10	Quarantining documents with specific Classifier labels in Box.com.....	12
3.10.1	<i>Changing an Adjudicator</i>	13
3.11	Using Box Notes in Box.com	13
4	Supporting Users with Languages other than English	14
5	Load Balancing	15
6	Audit Events	16
7	Interworking with Classifier applications	18
8	Deployment and Infrastructure Considerations	19
8.1	Firewalls and Proxies	19
8.2	Load and Performance	19

1 INTRODUCTION

This document summarises the features supported by Box Classifier.

Box Classifier is an integration application that controls documents labelled using Classifier within the Box (www.box.com) environment. Box is a content management platform that helps people securely store, access, share, and collaborate on files across devices. The Box platform does not recognise the Classifier labels that may be attached to documents. Instead, it uses its own metadata system for classification and access control. Box Classifier automatically synchronises Classifier labels with the Box metadata system for classification.

The main product features are:

- ✓ Monitor events within the customer's Box Enterprise.
- ✓ For document uploads synchronise Classifier labels with the Box metadata system for classification.
- ✓ Optionally delete or quarantine documents with specific Classifier labels.
- ✓ Optionally scan existing documents within the customer's Box Enterprise before monitoring new events.

2 INSTALLATION AND DEPLOYMENT OPTIONS

2.1 Initial Configuration of the Box Enterprise

Box Classifier is a custom application that uses Box Platform. To enable Box Classifier to work with your Box Enterprise you will need to create a configuration for the custom application recognised by Box and then authorise those connection details within your Box Enterprise. Follow the instructions below to create the initial configuration and authorise Box Classifier for your Box Enterprise

The following instructions summarise the guidelines provided by Box for creating this initial configuration. Administrators should review the official Box Guidelines at <https://developer.box.com/docs/application-setup> before continuing.

Please note that these instructions and any screenshots are correct at time of writing but may change over time. For the definitive guide on configuring an App in Box.com refer to the Box documentation as it appears on their website.

2.1.1 Create a Developer Account

To create the initial configuration for Box Classifier using Box Platform, you will first need to access the Box Developer Console.

You can sign-up to create a developer account at: https://box.com/developer_signup.

If you already have an account, use: <https://account.box.com/developers/services>

2.1.2 Set up Two-Factor Authentication

Set up two-factor authentication as follows:

- Log in to Box, within your **Account Settings**, select **Security**.
- Check **Login Verification** and follow the instructions to enable 2-step verification.

2.1.3 Create the Box Classifier App

Use the Box Developer Console to create your new app.

- Log in to Box and select **Dev Console**.
- Select **Create New App**.
- Select the **Enterprise Integration** app type.
- Select the **OAuth 2.0 with JWT (Server Authentication)** option.
- Choose a unique name for the app, such as **Classifier Integration for [My Company Name]**.
- Press **Create App**.

2.1.3.1 Basic Configuration

Ensure that your new App has the following options set:

- Authentication Method
 - OAuth 2.0 with JWT (Server Authentication)
- Application Access
 - Enterprise
- Application Scopes
 - Read and write all files and folders stored in Box
 - Manage users
 - Manage enterprise properties
- Advanced Features
 - Perform Actions as Users
 - Generate User Access Tokens

2.1.3.2 Public Key Management

Box Classifier makes requests to the Box Cloud Storage System that are signed using a private key and are verified by the public key registered with the Box System.

Generate the RSA key pair using the <https://www.openssl.org/> toolkit.

The OpenSSL project does not distribute any code in binary form, and does not officially recommend any specific binary distributions. An informal list of third party products can be found on <https://wiki.openssl.org/index.php/Binaries>.

During production, Boldon James used the <http://www.cygwin.com/> package to generate the OpenSSL RSA key pair with the commands below.

Generating a Private Key

```
openssl genrsa -aes256 -out private_key.pem 2048
```

Generating a Public Key

```
openssl rsa -pubout -in private_key.pem -out public_key.pem
```

Submitting the Public Key

Once you have generated a public key, you will need to save it in your Box Classifier application configuration under **Add and Manage Public Keys**.

Further advice on generating the RSA key pair can be found at:

<https://developer.box.com/docs/setting-up-a-jwt-app#section-step-2-generate-a-public-private-keypair>

Note: Be sure to protect your private key, as this is the only means to verify your application's identity.

2.2 Authorise Box Classifier for Your Enterprise

Grant authorization for your Box Classifier application in the Box Enterprise Admin console.

- Log in to Box and select **Admin Console**.
- Go to the **Enterprise Settings** and click on **Apps**.
- Within the **Custom Applications** section, choose **Authorize New App**.
- Enter the **API Key** for your Box Classifier Application. The **API Key** is the **Client ID** (OAuth 2.0 Credentials) of your app, see **Dev Console | Classifier Integration for [My Company Name] | Configuration | OAuth 2.0 Credentials | Client ID**. Entering the **API Key** will display the requested **Applications Scopes and User Access Level** that determines which users the application has control over.

Once **authorized**, the Box Classifier Application will be enabled within your Enterprise.

2.3 Create the Box Metadata Classification Policy

Security classification is available to all Box Governance customers. If you don't see the classification options contact box.com to ask them to enable the Box Governance Service for your Enterprise. Security classification is a feature that enables customers to classify files based on their confidentiality and enforce security policies associated with that confidentiality level.

To create a new classification using the Classic Admin Console:

- Log in to Box and select **Admin Console**.
- Go to **Policies** and select **Content Controls**.
- Click on **Create**.

For the New Admin Console the option is located at **Governance/Classification**

Now complete the following steps:

1. Begin by naming your classification label. This name must be unique, and have a maximum of 40 characters.
2. You also have the option to apply shared link restrictions on classified files. Select one of the following three levels of shared link restrictions available:
 - Public - No restriction on sharing the classified file
 - Company and Collaborators only - No external sharing of the classified file

- Collaborators only – company-wide or external sharing of the classified file is prohibited.
3. Once you are finished modifying your classification policy settings, click **Create**.
 4. You cannot rename a classification once it has been created, and existing classifications cannot be deleted. Confirm that you understand these restrictions, then click **Create** to proceed.

For example:

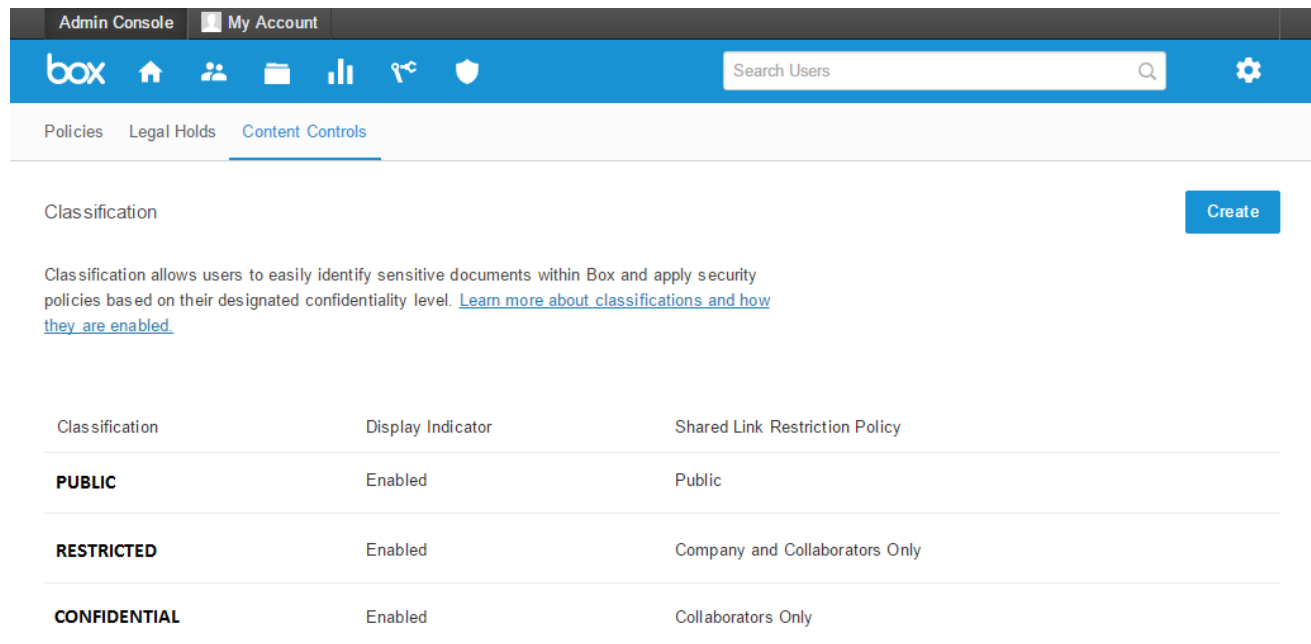


Figure 1: Example Box Classifications

2.4 Installation of Box Classifier

If there is an existing installation of the software, uninstall that version of the software before proceeding.

The product is installed into the server environment by running **BoxClassifier.msi** from the Classifier Bundle under the **Box Classifier** subfolder. The installation program presents a sequence of standard setup screens:

- A **Welcome** screen, where the **Next** button should be selected.
- An **End-User License Agreement** screen. Carefully read and accept the agreement then select the **Next** button.
- A **Select Installation Folder** screen. Confirm the installation folder by selecting the **Next** button.
- A **Ready To Install** screen, where the **Install** button should be selected.

The following steps are required to establish a working Box Classifier installation:

- Apply your Box Classifier licence to your Classifier Configuration using Classifier Administration **Global Settings / Licence** mechanisms.
- Configure Box Classifier [Policy control](#) (Rules).
- Ensure that the updated configuration is **Published** (using Classifier Administration / Publish) to a location accessible to your intended Box Classifier server(s).
- Ensure that the relevant Classifier configuration registry keys are set on the server that is to run Box Classifier. For example, if the Classifier Configuration is published via File Store, then set up the following registry entries with the relevant site values:
 - **HKEY_LOCAL_MACHINE\SOFTWARE\Boldon James\ConfigManager\ServiceMode**
 - ServerRootType REG_DWORD 0
 - ServerFileSystemRoot REG_SZ \\ICS-SVR-1\Classifier Configuration
 - LabelConfiguration REG_SZ Classifier Test
 - Policy REG_SZ All users
 - (You can find comprehensive details of the Classifier registry search algorithm in the Classifier Administration Guide **Configuration Deployment** section).
- Complete the [Configuration of Box Classifier](#).

2.5 Configuration of Box Classifier

After creating the initial configuration within Box and authorising that configuration for use with your Box Enterprise, configuration of the Box Classifier product can continue.

Note: For Box Classifier to connect to your Box Enterprise the UTC time on the machine must be synchronised with the Box servers. Check and if necessary synchronize the Box Classifier machine's time with a trusted, external resource. A difference of a few seconds in the times is allowed.

Box Classifier connects to your Box Enterprise through the two endpoints of api.box.com and dl.boxcloud.com. You may need to adjust Firewall configuration and Proxy Server configuration to allow this connection. See Section 8.1 for more detail.

Use the Box Classifier Configuration Console that is installed into your server's Apps repository to finalise configuration of Box Classifier.

The Configuration console, as shown in figure 2, writes configuration data to the application folder on the server where you deploy the Box Classifier.

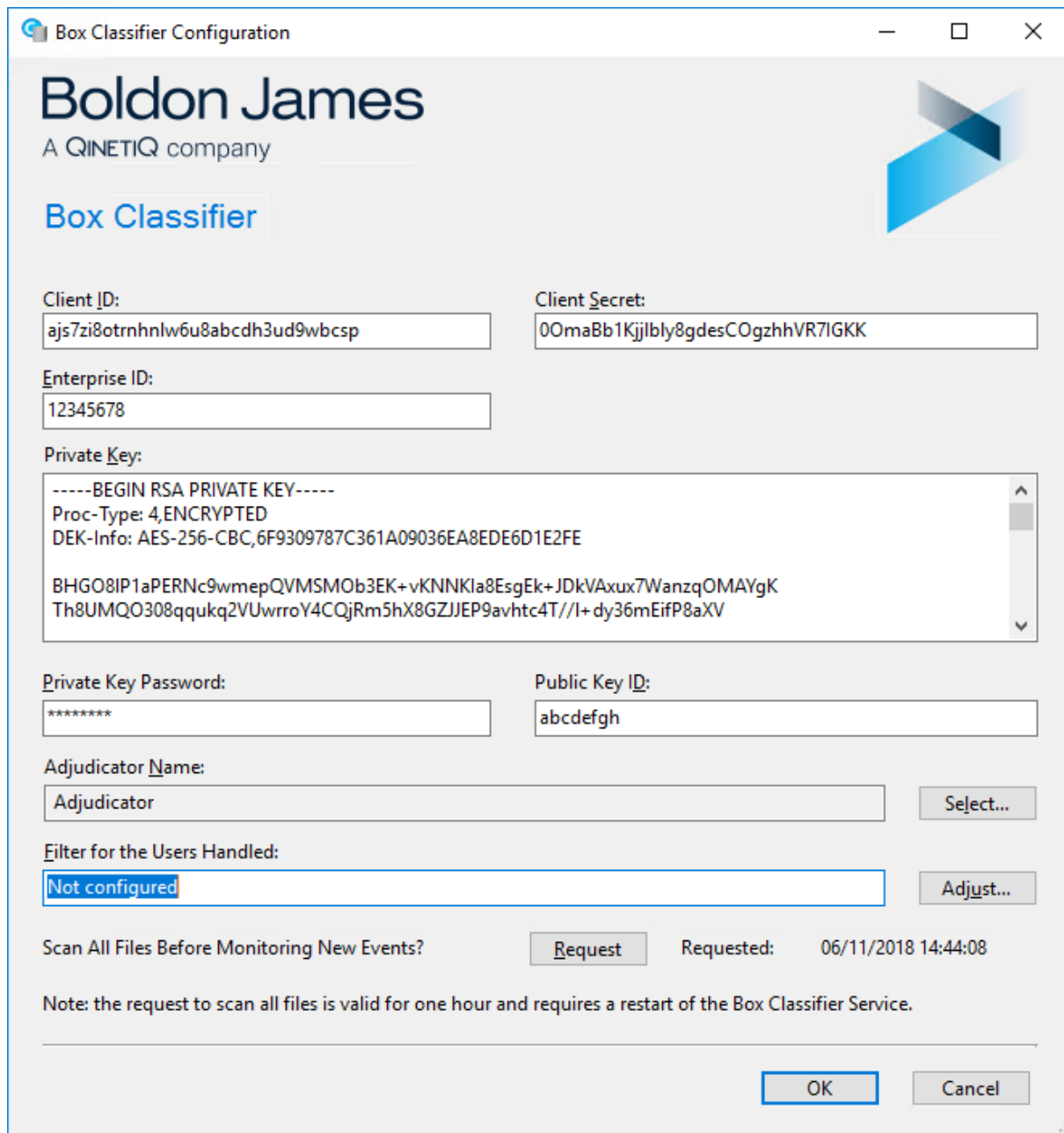


Figure 2: Box Classifier Configuration Console

Using the initial configuration created in section 2.1.3:

- Enter the **Client ID**, **Client Secret** and **Public Key ID**.
- Enter the **Private Key** and **Private Key Password** associated with the Public Key created in section 2.1.3.2
- Enter the **Enterprise ID** found in the Account Info page of your Box Enterprise Settings.

The **Adjudicator Name** field identifies the designated administrator or adjudicator that can act when there is a mismatch between classifier labels and box classification values. Click on the **Select...** button to show the Enterprise Users that Box Classifier can access.

The **Filter for the Users Handled** is only used for load balancing. Load balancing is only necessary if your Box Enterprise has a large number of users or usage is high. If you need to employ load balancing see section 5

The **Scan All Files Before Monitoring New Events** indicates whether existing files in the Box enterprise should be analysed for label inconsistencies. The request to scan all files is valid for one hour once the **Request** button is clicked and the configuration changes are committed using the **OK** button. The scan will take place when the Box Classifier Service is re-started. See [The Box Classifier Service](#) below for more details.

2.6 Policy control

Box Classifier operates using the Classifier Policy configured for the local machine. Refer to the Classifier Administration Guide → Rules Supported for details on creating rules appropriate to Box Classifier.

The Box Classification values for your enterprise were created in section 2.3 and may differ from those values available for Classifier Labels.

The Box Classifier product allows you to use rules to:

- Provide mappings from Classifier Labels to the metadata system used by the Box Cloud Storage system for classification and access control.
- Prevent documents with specific Classifier labels being retained in the Box Cloud Storage system.
- Provide Adjudicator access within your Box enterprise to a document with a specific Classifier label. This process is termed “Quarantine”. An Adjudicator is configured as described in Section 2.5.

For example, using the example Box Classification values shown in section 2.3 a set of rules could be:

Selectors and rules assigned to this policy.

Application	Name	Sensitivity	Distribution	Rule Option 2
Box Cloud Storage Documents	Delete the Box Document	COMMERCIAL-IN-CONFIDENCE	INTERNAL	Delete
Box Cloud Storage Documents	Quarantine the Box Document			Quarantine
Box Cloud Storage Documents	Apply a Box Classification of PUBLIC	UNMARKED		PUBLIC
Box Cloud Storage Documents	Apply a Box Classification of RESTRICTED	SENSITIVE		RESTRICTED
Box Cloud Storage Documents	Apply a Box Classification of CONFIDENTIAL	COMMERCIAL-IN-CONFIDENCE	EXTERNAL	CONFIDENTIAL

Figure 3: Example Policy Rules

2.7 The Box Classifier Service

The Box Classifier Service is started from the standard Microsoft Management Control for Services as shown in Figure 4.

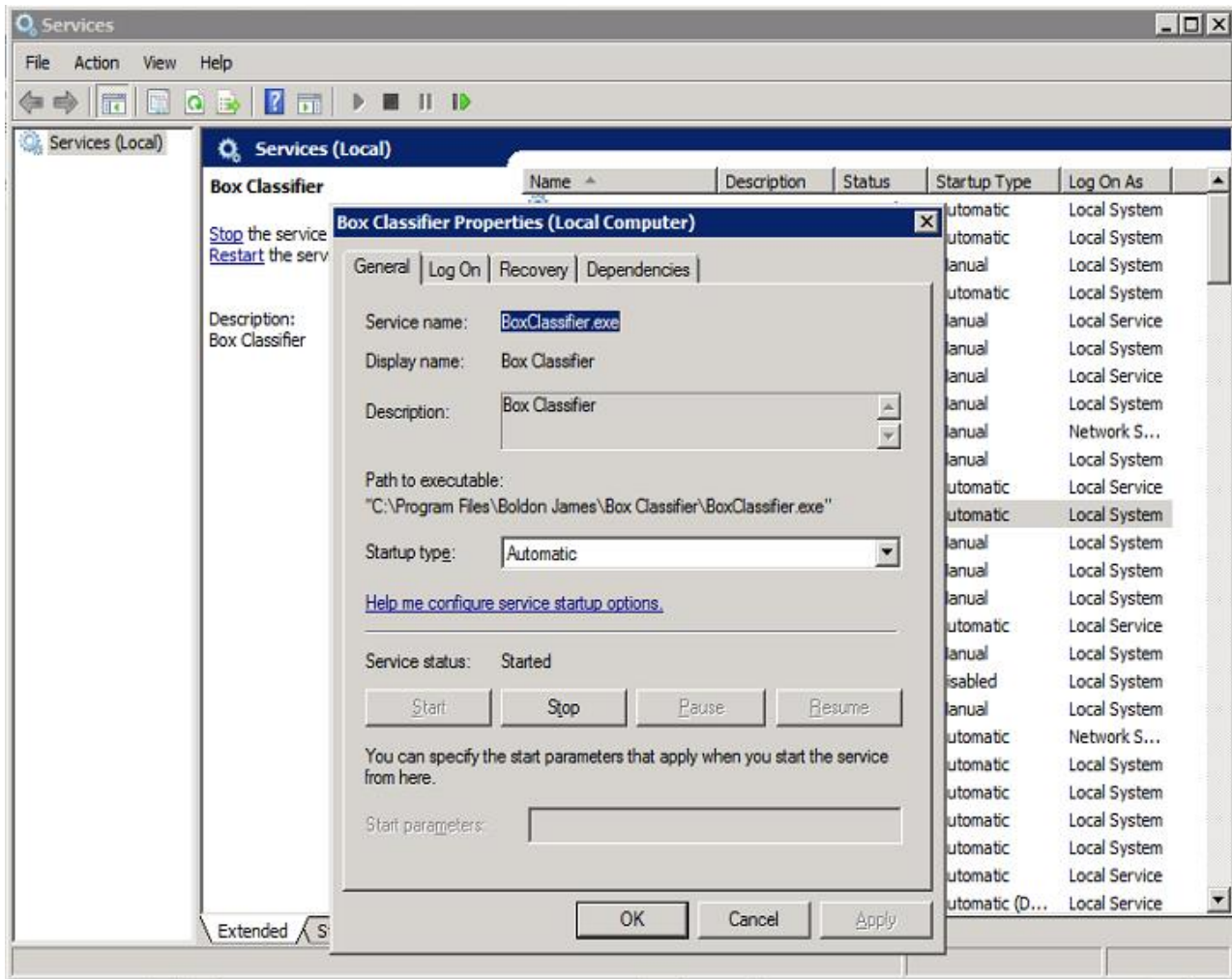


Figure 4: Box Classifier Service

Before using the Box Classifier Service, you should ensure that the Recovery options for the service are set appropriately. It is recommended that the **Restart the Service** option is used.

3 FUNCTIONALITY

3.1 Events and Actions

Various events are recorded by the Box Cloud Storage system which can then be queried by Box Classifier. The events identify the user and the document in question.

Box Classifier queries the Box Cloud Storage system for the following events:

- ✓ [Uploading a new document](#) to Box.com.
- ✓ [Uploading a new version](#) of an existing document to Box.com. [Synchronisation](#) actions may also give rise to such events.
- ✓ [Adjustment to the Classification](#) in Box.com.
- ✓ [Copying a document](#) in Box.com.
- ✓ [Unlocking a document](#) in Box.com.
- ✓ [Restoring a document from Trash](#) in Box.com.

In each of these instances, Box Classifier may perform actions according to any of the following Box Classifier Rules that have been defined through Classifier Administration:

- Check that the classification applied to the document in Box.com is appropriate to the Classifier Label. If the Box classification isn't appropriate, then it can be adjusted and a comment applied to the document recording the change.
- Based on a specific Classifier Label value or if a Classifier Label is not present, a rule may be created to [Delete the document](#) from Box.com.
- Based on a specific Classifier Label value or if a Classifier Label is not present, a rule may be created to Quarantine the document in Box.com. See [Quarantining documents](#) for more information on the Quarantine functionality.

Note: If there is no Classifier Label but there is a Classification value applied to the document, a rule to Quarantine the document will not be actioned. This rule exception is to allow documents without Classifier Labels to exist in the Box Cloud Storage system and yet still be subject to the access control facilities provided by the Classification system.

3.2 Uploading a document to Box.com

When a document is uploaded to the Box Cloud Storage system, an UPLOAD event is recorded. The UPLOAD event can take a short time, typically between 30 seconds and 90 seconds, before it is available to be queried by Box Classifier. A user might apply a Box classification during this time for example.

The [actions](#) described above can be performed on the document.

3.3 Uploading a new version of a document to Box.com

A new version of a document in Box.com will retain the classification value from a previous document version. The new version may, however, use a different Classifier Label. In this situation, Box Classifier may need to perform new [actions](#) on the revised document.

3.4 Synchronisation of documents

Synchronisation of documents between a user's desktop or mobile device and Box.com can introduce a new version similar to [Uploading a new version](#). The [actions](#) described above can be performed on the revised document:

3.5 Adjusting the Classification of a document in Box.com

When a document classification is created, deleted or changed in the Box Cloud Storage system, an event is recorded. The [actions](#) described above can be performed on the revised document:

3.6 Unlocking a document in Box.com

Downloading a locked document from Box.com is usually not possible. For Box Classifier to function correctly, it must be able to download documents to check the associated Classifier Label. For this reason, Box Classifier monitors events that record when a document is unlocked.

The [actions](#) described above can be performed on the revised document.

3.7 Copying a document in Box.com

When a document is copied in the Box Cloud Storage system, the classification value is not copied from the original version of the document. Box Classifier monitors the COPY event and will perform the [actions](#) described above on the document copy.

3.8 Restoring a document from Trash in Box.com

When a document is restored in the Box Cloud Storage system, Box Classifier checks the associated Classifier Label on the document.

The [actions](#) described above can be performed on the restored document.

3.9 Preventing documents with specific Classifier labels being retained in Box.com

Based on a specific Classifier Label value or if a Classifier Label is not present, a rule may be created to Delete the document from Box.com.

If this action is performed, then the document is **trashed** and then deleted permanently from the **Trash** folder in Box.

Note: If there are multiple versions of a document, a delete action on the current version will apply to all versions of the document. This is a Box Cloud Storage system approach.

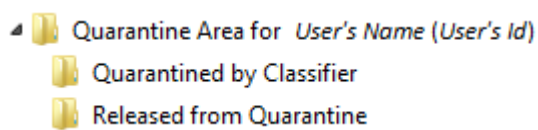
If desktop synchronization or mobile synchronization is enabled then this may also delete the document from those paired machines and/or devices.

3.10 Quarantining documents with specific Classifier labels in Box.com

Box Classifier requires the nomination of an Adjudicator to handle quarantined documents. The adjudicator is an enterprise user that is qualified to make corrections to documents to ensure they adhere to the company’s labelling policies. Typical document corrections are:

- ✓ Applying a Classifier Label.
- ✓ Applying a Classification
- ✓ Collaboration with the document owner to ensure policy compliance.

If a document in Box.com is to be quarantined, then Box Classifier will check if the user performing the action on the document has a quarantine folder structure created. If there is no quarantine folder structure created, then Box Classifier will create this folder structure in the Adjudicator’s storage area



The two sub-folders **Quarantined by Classifier** and **Released from Quarantine** are shared with two and only two collaborators: the quarantined document owner and the Adjudicator. Access permissions are applied to the **Quarantined by Classifier** sub-folder to allow only the Adjudicator to edit documents whereas the **Released from Quarantine** sub-folder allows full permissions for the document owner.

Once the quarantine folder structure is in place a document is moved into quarantine by:

1. Checking to see if a document with that name is already in quarantine. If it is that a new name is created for the document by appending a number in brackets to the document name. For example,

work proposal.docx → work proposal(1).docx
--

2. Moving the document into the **Quarantined by Classifier** folder.
3. Applying a comment to the document to describe what has happened
4. Locking the file

Once a document has been quarantined the Adjudicator can interrogate and edit the document within the Quarantine area. The Adjudicator can release the document from Quarantine by moving it to the **Released from Quarantine** sub-folder. Adjudicators should be careful to ensure the released document has an appropriate Classification, otherwise Box Classifier may quarantine the document for a second time.

3.10.1 Changing an Adjudicator

If a decision is made to adjust the Adjudicator then both the old adjudicator and new adjudicator will have co-ownership on the Quarantine folders. If the old adjudicator no longer requires access then that person can give ownership of the top level quarantine folder for a user to the new adjudicator and then remove their own collaboration on the folder. The top level quarantine folder for a user has a name similar to

📁 Quarantine Area for *User's Name (User's Id)*

3.11 Using Box Notes in Box.com

Box Classifier does not monitor Box Notes in Box.com

4 SUPPORTING USERS WITH LANGUAGES OTHER THAN ENGLISH

Box Classifier adds comments to documents within the Box Cloud Storage System whenever it automatically adjusts classification values or quarantines a document. The comments are, by default, in English.

Box Classifier can write the comments in the configured language for the Box User who edited the document. A Box User can change their language settings in the Account Settings section of Box.com. Box Classifier will use this setting along with information held in the **BoxClassifier.exe.config** file that is located in the installation folder for the product.

There are five comments that may be applied to Box Documents. These are described by the following section in the **BoxClassifier.exe.config** file:

```
<LanguageSettings>
  <BoxComments>
    <add index="1" ref="QuarantineString1" value="Box Classifier: the
item's document label requires the item be automatically quarantined."
languageCode="en" />
    <add index="2" ref="QuarantineString2" value="Box Classifier: there is no
document label for this item. The item has been automatically quarantined."
languageCode="en" />
    <add index="3" ref="QuarantineString3" value="The item has been renamed to
'{0}' because of a name conflict in the folder." languageCode="en" />
    <add index="4" ref="LabelString1" value="Box Classifier: document label of
{0} requires a classification of {1}. Automatically applying the classification
{1}." languageCode="en" />
    <add index="5" ref="LabelString2" value="Box Classifier: document label of
{0} does not match the classification of {1}. Automatically adjusting the
classification to {2}." languageCode="en" />
    <add index="6" ref="QuarantineString1" value="Box Classifier: Das
Dokument-Etikett des Elements erfordert, dass das Element automatisch unter
Quarantäne gestellt wird." languageCode="de" />
    <add index="7" ref="QuarantineString2" value="Box Classifier: Für dieses
Element ist kein Dokumentetikett vorhanden. Das Element wurde automatisch unter
Quarantäne gestellt." languageCode="de" />
    <add index="8" ref="QuarantineString3" value="Das Element wurde aufgrund
eines Namenskonflikts in den Ordner '{0}' umbenannt." languageCode="de" />
    <add index="9" ref="LabelString1" value="Box Classifier: Dokument-Label
von {0} erfordert eine Klassifizierung von {1}. Automatisches Anwenden der
Klassifizierung {1}." languageCode="de" />
    <add index="10" ref="LabelString2" value="Box Classifier: Dokument-Label
von {0} stimmt nicht mit der Klassifizierung von {1} überein. Automatische
Anpassung der Klassifizierung auf {2}." languageCode="de" />
```

```
</BoxComments>
```

```
</LanguageSettings>
```

The section above shows the five comments firstly in English and then in German. An administrator can choose to add further language variants so long as the following rules are adhered to:

1. Each XML `<add/>` node must have a unique `index` attribute.
2. The `ref` attribute must contain one of the values shown above. Box Classifier uses this attribute to understand the context of the comment.
3. The `value` attribute must retain any placeholders of `{0}`, `{1}` and `{2}`. These placeholders are replaced with appropriate text when the comment is applied.
4. The `languageCode` attribute is the two character country/region code taken from [ISO 639-1 Language Code](#)

Once the file has been edited to include additional languages, restart the Box Classifier Service.

5 LOAD BALANCING

Depending on the size of your Box Enterprise, it may be necessary to install Box Classifier onto multiple servers to provide load balancing. If you determine this is necessary, then simply adjust the load balancing filter using the Box Configuration Console.

Box Classifier will use the **Filter for the Users Handled** field in the configuration form to only handle events from users that match the Filter. For example, if you have two servers with Box Classifier installed then you can craft a filter to handle half the users on the first server and use another filter to handle the other half of the users on the second server.

The filter takes the form of a *regular expression*. To check that you have crafted the *regular expression* correctly or to receive a suggestion click the **Adjust...** button. The following dialog is presented.

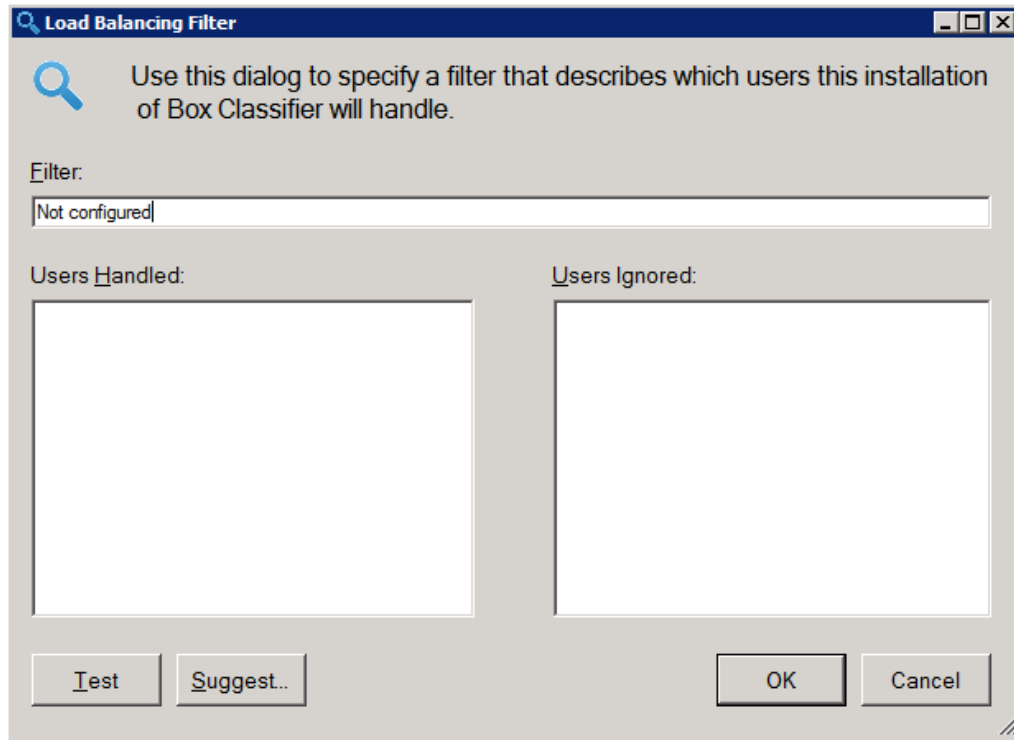


Figure 5: Load Balancing Filter

You can enter a *regular expression* in the **Filter** field and click **Test** to see which users the filter applies to. If you would like a suggested filter, click the **Suggest...** button to present the following dialog:

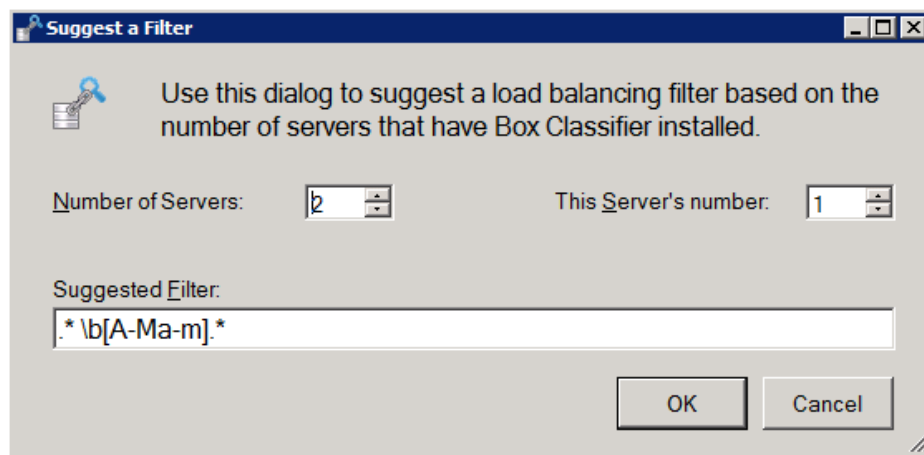


Figure 6: Suggest a Filter

Simply adjust the **Number of Servers** with Box Classifier installed and **This Server's Number** to suggest a filter. The filter suggested works on the first letter of Users' Surnames. The example shown in Figure 8, describes a filter that will handle surnames that begin with a letter in the first half of the alphabet, i.e. A to M.

6 AUDIT EVENTS

Important events are recorded in the Windows Application Event log under the log name Boldon James/Classifier. Events have the following general appearance

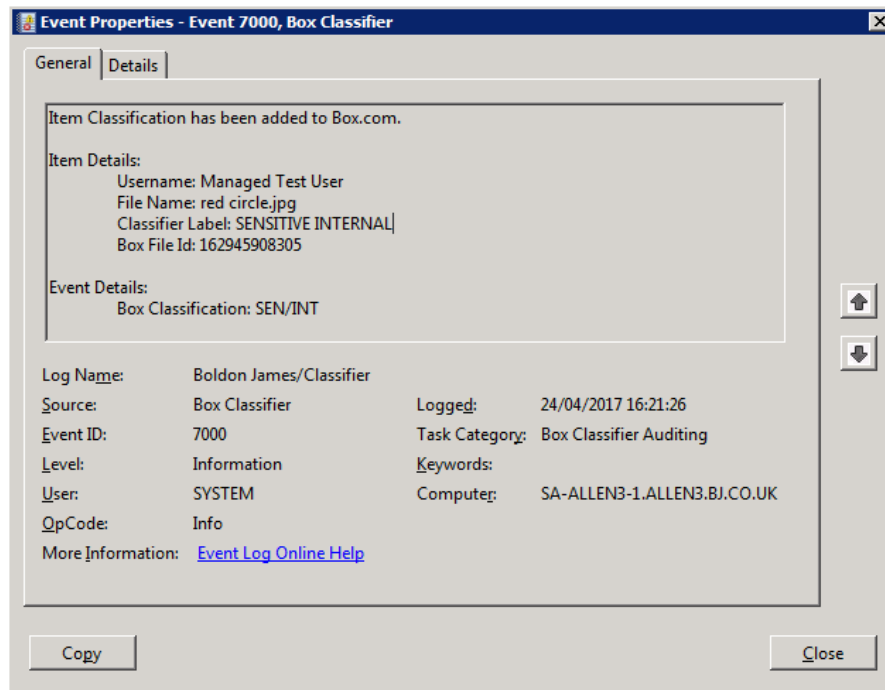


Figure 7: Example Audit Event

The following important events are recorded:

Level (Event ID)	Event Data	Meaning
Error (1005)	Licensing messages	There is a problem with the licensing for Box Classifier.
Information (2000)	Service started	Box Classifier Service started
Information (2000)	Service stopping	Box Classifier Service stopping
Error (2000)	The service encountered an issue that it cannot recover from and will be stopped. Further diagnostic information...	Box Classifier Service stopping
Information (7000)	Item Classification has been added to Box.com.	Box Classification added
<i>Item details and event details are also recorded</i>		
Information (7001)	Item Classification has been updated in Box.com. <i>Item details and event details are also recorded</i>	Box Classification updated

Information (7002)	Item has been deleted from Box.com. <i>Item details are also recorded</i>	Box Item deleted
Information (7003)	Item has been quarantined in Box.com. <i>Item details are also recorded</i>	Box Item quarantined
Warning (7006)	Quarantine of the item failed in Box.com. <i>Item details are also recorded</i>	Box Item quarantine failed
Warning (7007)	Deletion of the item failed in Box.com. <i>Item details are also recorded</i>	Box Item deletion failed
Warning (7008)	A required change to the item Classification failed in Box.com. <i>Item details are also recorded</i>	Box Classification change failed

7 INTERWORKING WITH CLASSIFIER APPLICATIONS

Key information on the interworking with Classifier applications is:

- Box Classifier only interrogates the source item to obtain the Classifier Label (if any) – it does not modify the original item.
- Box Classifier uses the same Label information as File Classifier, Office Classifier and CAD Classifier (Classifier ‘private’ label property and Label Locations) when reading labels from file types supported (e.g. Microsoft Word, Excel, PowerPoint and Project - note **VISIO** files must be labelled using the **VISIO** application).

Alternate Data Stream Limitation:

- Uploading a file labelled by Power or File Classifier (but not supported by Office Classifier) to the Box Cloud Storage System will lose the additional Alternate Data Stream (ADS) information. The file will then be unlabelled.

Note: More detailed information on the use of labels by, and propagation of labels between, Classifier applications can be found in **Administration Guide > Appendix – Label Propagation Mechanisms**.

8 DEPLOYMENT AND INFRASTRUCTURE CONSIDERATIONS

8.1 Firewalls and Proxies

Box Classifier connects to your Box Enterprise through the two endpoints of api.box.com and dl.boxcloud.com. You may need to adjust Firewall configuration and Proxy Server configuration to allow this connection.

Both the Box Classifier Service and Box Classifier Configuration Console make connections to the Box Enterprise. The Box Classifier Configuration Console connects under the current user's credentials whereas the Box Classifier Service runs as Local System account by default. You should change the Box Classifier Service to run as an appropriate user that is recognised by the Proxy Server.

The Box Classifier Service is started from the standard Microsoft Management Control for Services as shown in Figure 4.

8.2 Load and Performance

When deploying Box Classifier it is vital to understand the load that will be placed on the product. This depends on a number of factors including:

- The typical number of document uploads to Box.com and document changes in Box.com per day
- The typical size of documents that are uploaded or changed.

Details on these two factors can be obtained through the Reports facility that Box.com offers on the Box Admin Console.

Boldon James Box Classifier has the following performance characteristics in order of influence:

1. Box.com API Limits – Box.com impose usage restrictions on 3rd Party integrations. These standard enterprise restrictions impose an approximate handling limit for Box Classifier of:
 - 51000 document uploads and document changes per day of a 10KB file
 - 6900 document uploads and document changes per day of a 250MB file

Box.com can agree to a relaxation of these usage restrictions for an Enterprise. Interested customers should contact Box.com for further information.

2. Network Bandwidth – the Box Classifier service downloads files temporarily to memory to analyse the Classifier label of the file. It does not upload files. Network bandwidth plays a more significant role as document size increases. Testing indicates that a download speed of around 300 Mbps does not constrain performance for large files of 250MB; it is the API limits imposed by Box.com that constrain performance when such bandwidth is available. For smaller files of around 10KB then the download speed can be as low as 20 Mbps before Box.com API limits influence performance.
3. Memory – the Box Classifier service streams files directly to memory to check the associated Classifier label. This means that peak memory usage occurs when there are multiple large files being streamed simultaneously. Box Classifier uses ten channels to stream files and allows a simple heuristic of $10 \times \langle \text{maximum file size} \rangle$ to indicate memory required by the service. For instance, system testing shows peak private byte usage of 2.4GB when handling files of size 250MB.

Boldon James recommend that Box Classifier is installed on a standalone machine that does not share resources with other server applications such as Microsoft Exchange.

4. CPU - the Box Classifier service has an average processor utilisation of 0.37% when handling files of size 10 KB rising to 8.6% when handling files of size 250MB. This was determined using Intel i7-4770K CPU@3.5GHz
5. Disk – the Box Classifier service streams files directly to memory so disk size and disk performance is not a primary consideration.

The primary factor that affects the performance of Box Classifier is the Box.com API limits with a secondary influence of download speed. When Box Classifier approaches the Box.com API limits, user events that occur within Box.com are queued to be retried at a later time. From a user's perspective, a document will take longer to be processed. During peak usage this may be acceptable but if it occurs on a regular basis customers should consider deploying an additional instance of Box Classifier for load balancing as described in section 5.