Blue Coat® Systems SG™ Appliance

Volume 8: Managing Content

SGOS Version 5.1.x

Blue Coat

Contact Information

Blue Coat Systems Inc.
420 North Mary Ave
Sunnyvale, CA 94085-4121

http://www.bluecoat.com/support/contact.html

bcs.info@bluecoat.com
http://www.bluecoat.com

For concerns or feedback about the documentation: documentation@bluecoat.com

Document Number: 231-02844
Document Revision: SGOS 5.1.x—03/2007

Contents

Chapter 1: Introduction

Applying content filtering and virus scanning to requested and posted Web content in an enterprise is vital to securing the network and improving productivity.

- Content filtering allows you to regulate, based on content categories, which Web sites employees are allowed to access and which are restricted.
- Virus scanning allows you to scan both incoming content and content leaving the enterprise network for viruses and other malicious code, such as *drive-by* software that propagates spyware.

This document contains the following chapters:

- Chapter 2: "Content Filtering"
- Chapter 3: "ICAP"
- Chapter 4: "Configuring Service Groups"

Document Conventions

The following section lists the typographical and Command Line Interface (CLI) syntax conventions used in this manual.

Table 1-1. Document Conventions

Conventions	Definition
Italics	The first use of a new or Blue Coat-proprietary term.
`Courier font`	Command line text that appears on your administrator workstation.
`Courier Italics`	A command line variable that is to be substituted with a literal name or value pertaining to the appropriate facet of your network system.
`Courier Boldface`	A Blue Coat literal to be entered as shown.
{ }	One of the parameters enclosed within the braces must be supplied
[ ]	An optional parameter or parameters.
|	Either the parameter before or after the pipe character can or must be selected, but not both.

Chapter 2: Content Filtering

This chapter describes how to configure the SG appliance to process client Web requests and filter the returning content.

This chapter contains the following sections:

Section A: About Content Filtering

Content filtering allows you to block access to Web sites based on their perceived content.

Content Filtering Databases

A content filtering database is merely a list of sites, pages, and IP addresses organized by category. Depending on the vendor, a URL can belong only to one category or several categories. A content filtering database does not block any site or any category by default. The role of the database is to offer additional information to the proxy (and to the administrator) about the client request. Whether the Web site is blocked or allowed client access depends on the rules and policies implemented by the administrator in accordance with company standards. The challenge presented is that because of the dynamic nature of the Internet, there is a constant flow of new URLs (and URLs on lesser-known sites) that will not be in the content filtering database. As any URLs that are not in the database are marked as **none**, and you must created a policy to process these.

Content Filtering Categories

The infinite number of URLs can be reduced to a small number of categories. After the Web sites and content are categorized, access to that content can be controlled through policy by URL-based triggers.

Categories and their meanings are defined by the specific category providers. For third-party databases, the most up-to-date information on how categories are assigned to URLs can be obtained from the provider's Web site. You can request that specific URLs be reviewed for correct categorization, if your content-filtering provider supports this.

On-box vs. Off-box Solutions

Content filtering has two deployment options:

- On-box: The content filtering database exists on the proxy.
- Off-box: The proxy must contact another server to access the content filtering database.

The on-box solution provides better performance because the proxy does not require another network connection to perform the task; however, Blue Coat supports both methods (currently, Websense is the only supported off-box vendor).

The Blue Coat Content Filtering Solutions

The SG appliance offers the following content filtering options, any of which can be used separately or employed simultaneously:

- Employ the Blue Coat Web Filter (BCWF), an on-box content-filtering database maintained by Blue Coat that also offers dynamic category rating abilities.
- Employ your custom content filtering database (uploaded to the SG appliance), allowing or denying permission to URLs. You can create your own local database file and upload it to the SG appliance. This file is created in the same way that policy files are created, except that only `define category` statements are allowed in the local database.
- Employ a currently- supported third-party content filtering vendor database.

- Enable the Internet Watch Foundation (IWF) category.

The following diagram illustrates the process flow when content filtering (on-box or off-box) is employed in the network.

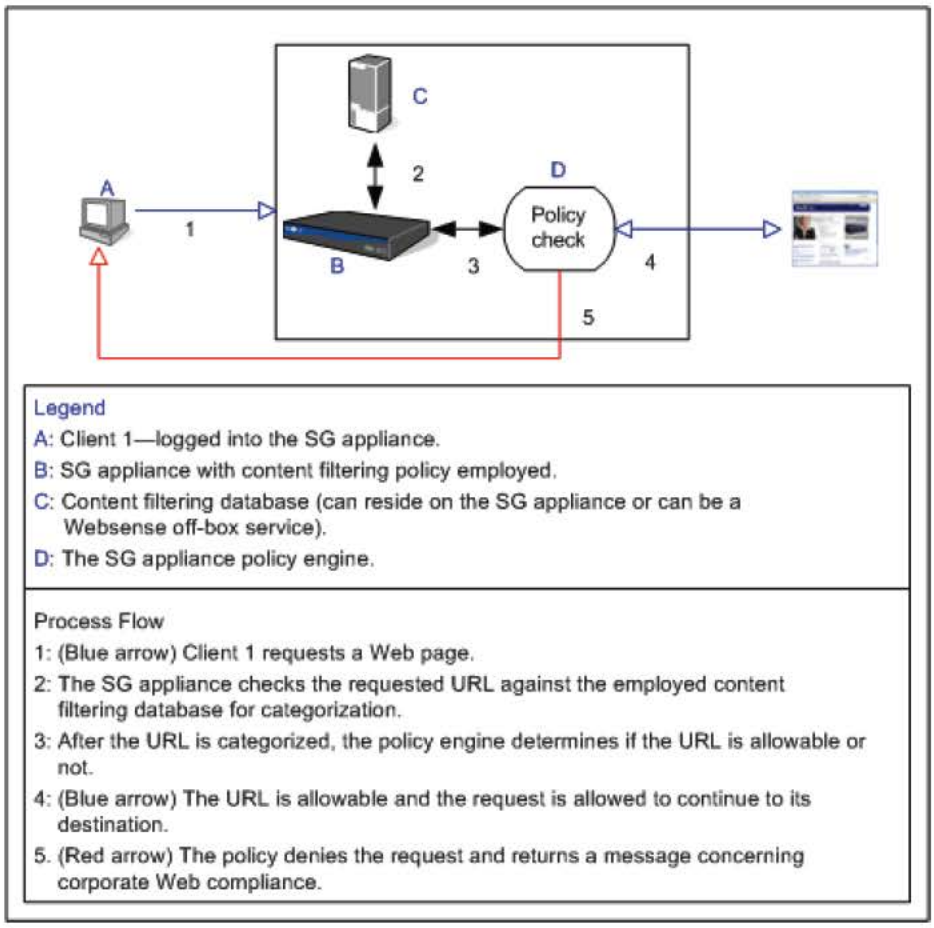


Figure 2-1. Basic Blue Coat Content Filtering Flow

The Blue Coat Web Filter Solution

The Blue Coat Web Filter (BCWF) is a hybrid solution: an extremely comprehensive URL list residing in a database on the SG appliance and the ability to provide real-time categorization of unlisted URLs. The URL list is updated daily.

About Blue Coat Web Filter

BCWF provides a comprehensive URL list with its on-box database and is consistent in how it categorizes resources and gives top priority to the most requested categories and Web sites. BCWF provides nearly 60 categories, which allows administrators substantial control when creating content filtering policy to allow or deny access content. Furthermore, a world-wide network of servers allows the SG appliance to expediently update the master BCWF database.

Supported Languages

BCWF supports many languages. Refer to the *Blue Coat Release Notes* for this release for the most recent list of supported languages.

About Dynamic Categorization

Used in conjunction with BCWF, *dynamic categorization* provides real-time analysis and content categorization of requested Web pages to deal with the problem of new and previously unknown uncategorized URLs—those not in the database. When a user requests a URL that has not already been categorized by the BCWF database (for example, a brand new Web site), the SG appliance *dynamic categorization service* analyzes elements of the requested content and assigns a category or categories. The dynamic service is consulted *only* when the installed BCWF database does not contain category information for an object.

If the category returned by this service is blocked by policy, the offending material never enters the network in any form.

Dynamic analysis of content is performed on a remote network service, and not locally on the SG appliance. Therefore, dynamic categorization incurs the following costs:

- Bandwidth: Represents the round trip request/response from the SG appliance to the service. Because the dynamic categorization protocol is compact, this cost is minimal.
- Latency: Represents the time spent waiting for the dynamic categorization service to provide a result.

While these costs are typically minute, certain conditions might require you to run dynamic categorization in the background or disable it.

The following diagram illustrates the BCWF content filtering flow when dynamic categorization is employed.

Section A: About Content Filtering

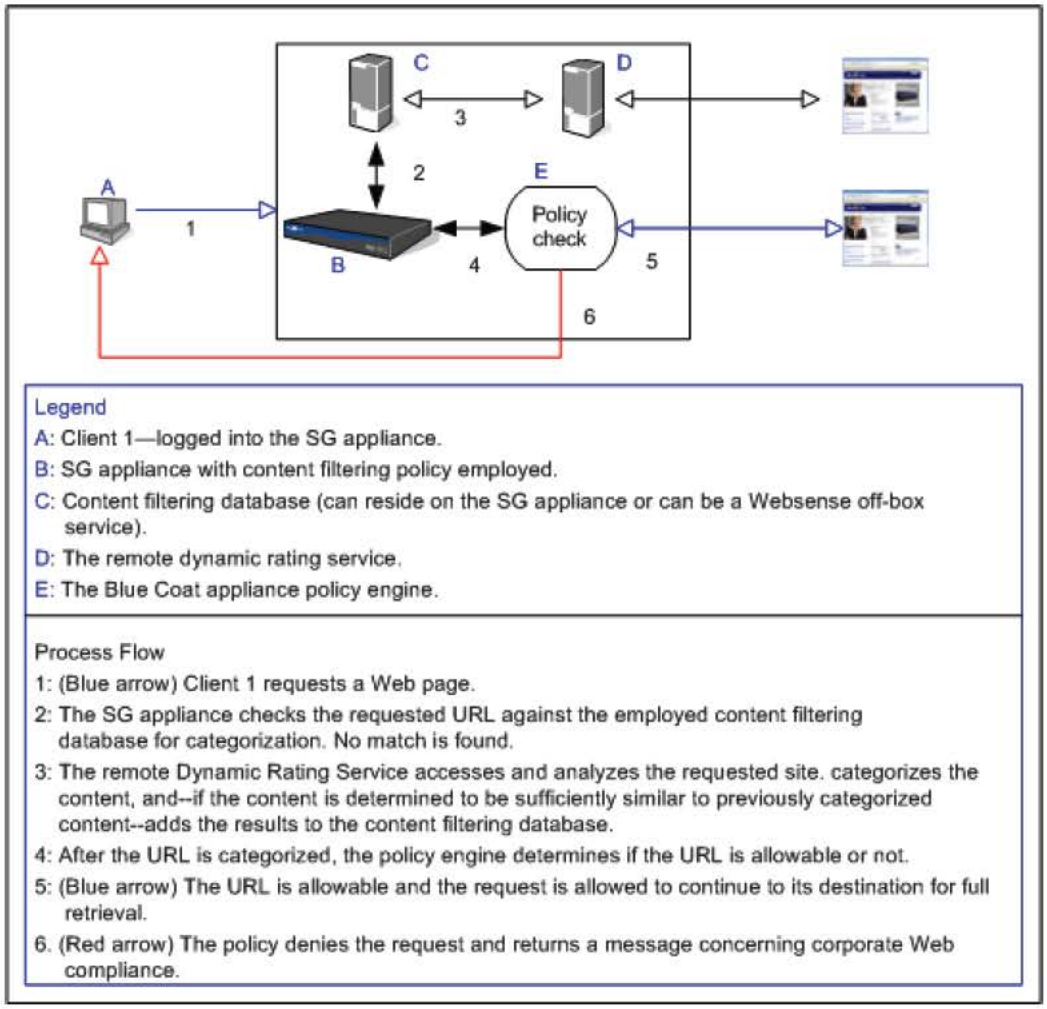


Figure 2-2. BCWF with Dynamic Categorization Content Flow

Supported Languages

The dynamic categorization system recognizes most of the languages supported in BCWF, however it only categorizes only a subset of those.

Section B: Configuring Blue Coat Web Filter

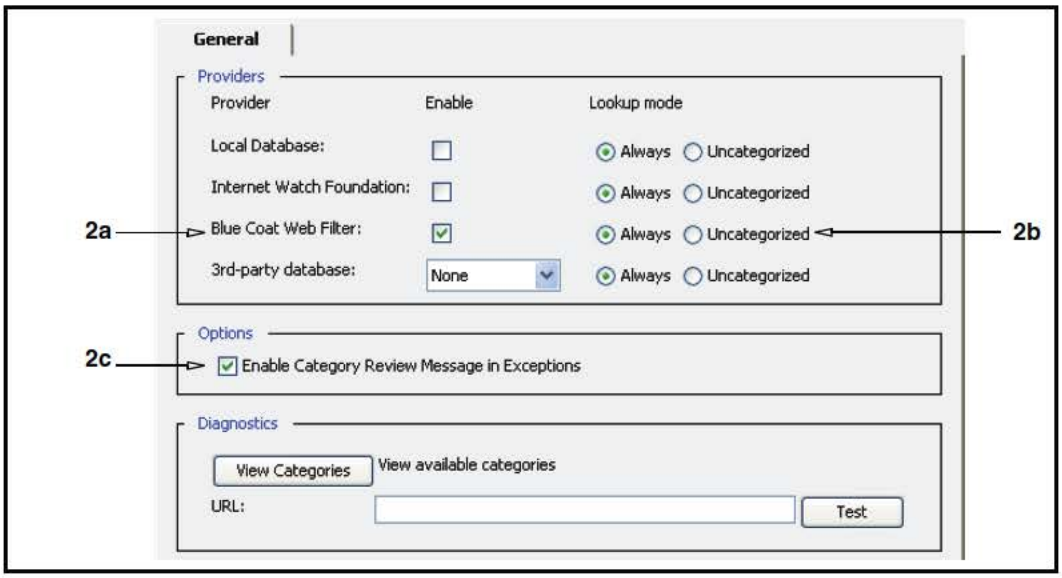
This section describes how to select and configure Blue Coat Web Filter (BCWF), how to schedule a the database update schedule, and how to change dynamic categorization settings.

Important: BCWF requires a valid license provided by Blue Coat. Refer to the Licensing chapter in *Volume 2: Getting Started*.

Selecting Blue Coat Web Filter and Downloading the Database

To configure Blue Coat Web Filter:

1. Select **Configuration > Content Filtering > General**.



2. Select BCWF as the provider:
   a. Select **Enable** for Blue Coat Web Filter.
   b. (Optional) Select the **Lookup Mode**. The default is **Always**, which specifies that the database is consulted on *every* categorization attempt. **Uncategorized** specifies that the database lookup is skipped if the URL already has categories assigned by another content filtering database (local, IWF, or third-party—in that order) employed on the SG appliance. The exception is if one of the content filtering database types is configured as **Uncategorized**, it is skipped.
   c. (Optional) Select **Enable Category Review Message in Exceptions**. Adds a link to the default content filter exception page that can be used to request review of the categories assigned to a blocked URL.

      Two substitutions (`$(exception_category_review_url)` and `$(exception_category_review_message)`) are automatically appended to the `help` element of all exception definitions. For information on using the `$(exception.help)` element, refer to *Volume 7: VPM and Advanced Policy*.

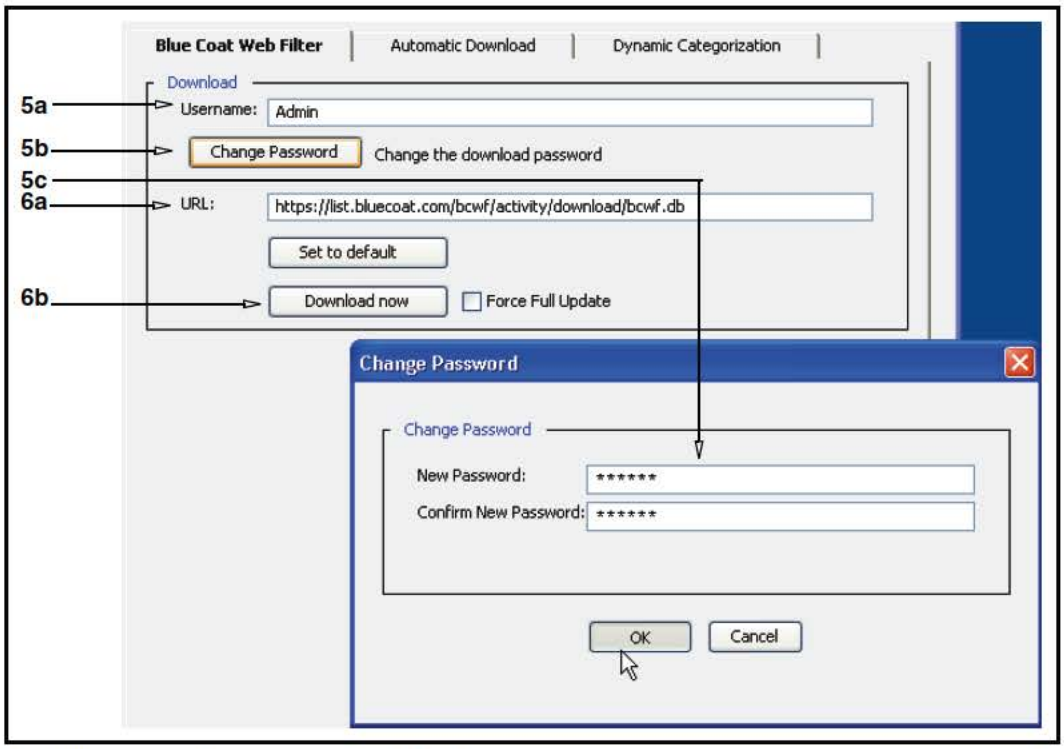
Note: The substitution values are empty if the database was not consulted for categorization, or if the categorization process failed due to an error.

3. Select **Apply** to commit the changes to the SG appliance.

Note: A small database that contains the category list is downloaded immediately, allowing immediate policy creation.

No username or password is required during the trial period (60 days). To download the database on demand or on a schedule, you must configure the BCWF service.

4. Select **Configuration > Content Filtering >** Blue Coat.



5. When you subscribed to the BCWF Service, you received a username and password for access to download updates.
   a. In the **Username** field, enter your username.
   b. Click **Change Password**; the Change Password dialog appears.
   c. Enter your password and click **OK**. (If you are in the trial period, no username or password is required.)
6. Download the database:
   a. The default database download location is displayed in the **URL** field.

Note: Only enter a new URL if instructed. Otherwise, accept the default.

b. Click **Download Now**. The Blue Coat Web Filter Installation status dialog box displays with the message Blue Coat **Web Filter download in progress**.

When the operation is complete, the dialog changes to indicate installation status.

c. Click **Results** to see the Blue Coat Web Filter download log:

```
Download log:
Blue Coat download at: Thu, 08 Jun 2006 00:04:06 UTC
Downloading from https://list.bluecoat.com/bcwf/activity/download/
bcwf.db
Requesting differential update
Differential update applied successfully
Download size:      84103448
Database date:      Wed, 07 Jun 2006 08:11:51 UTC
Database expires:   Fri, 07 Jul 2006 08:11:51 UTC
Database version:   2005040
```

d. Click **OK**.

7. Select **Apply** to commit the changes to the SG appliance.

Future Manual Downloads

You can return to this screen at any time and download a database on demand (independent of the automatic download feature, which is described in the next section). Ordinarily, the SG appliance checks if the database has changed before initiating a download. If the database is the most current, no download is performed. If an incremental update is available on the server, then it is downloaded (an incremental update contains only the changes between the current installed version and the latest published version of the database, and is much smaller than a full copy of the database). You can override this process and force a download of the full database by selecting **Force Full Update**.

Note: Because the incremental update process carefully verifies the update before and after applying it, forcing a full update is, rarely, if ever necessary. Routinely forcing a full update consumes excess download bandwidth and does not improve reliability.

Scheduling Automatic Downloads for Blue Coat Web Filter

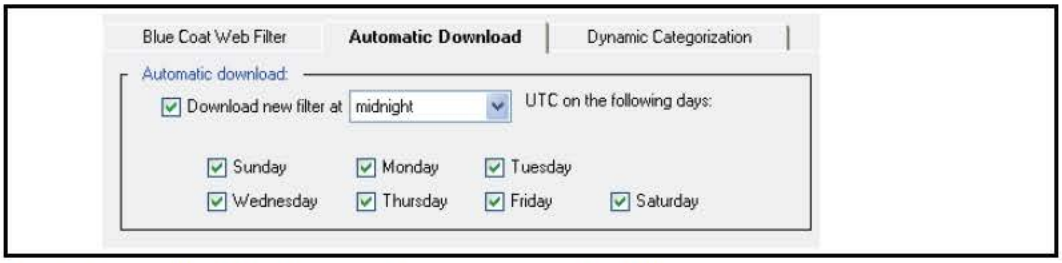
You can specify which days and times the BCWF database is downloaded. Because sites become stale quickly, Blue Coat recommends downloading on an automatic schedule daily.

When the database is downloaded, a log is available that includes the information about how the database was updated, but in a more detailed form. You can view the download log (Statistics>Advanced>Content Filter Service) or the CLI (SGOS#(config) `show content-filter status`).

Note: By default, the automatic download setting is enabled (for every day at midnight, UTC) and does not need to be configured unless you want to change the schedule or disable auto-download.

To schedule BCWF automatic download times:

1. Select **Configuration** > **Content Filtering** > Blue Coat > **Automatic Download**.



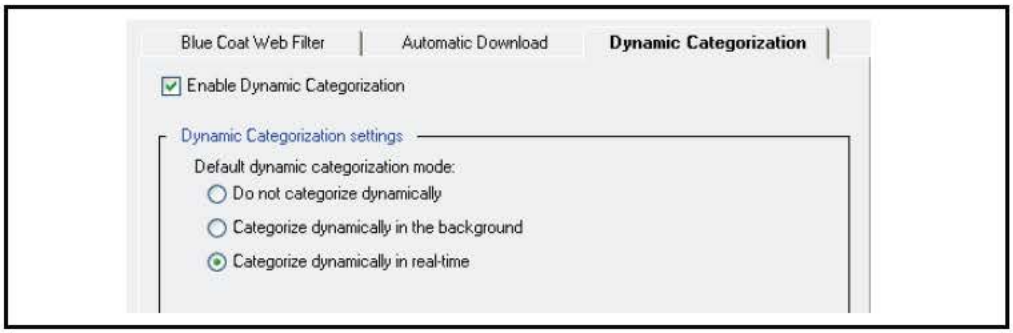
2. To change the default times (every day at midnight):
   a. From the **Download new filter at** drop-down list, select a time of day.
   b. Deselect days as required.
3. Select **Apply** to commit the changes to the SG appliance.

Configuring Dynamic Categorization

By default, dynamic categorization is enabled and configured to categorize un-categorized URLs. If this service is causing significant delays to enterprise Web communications, you can run the service in the background or disable the feature.

To configure dynamic categorization:

1. Select **Configuration** > **Content Filtering** > Blue Coat > **Dynamic Categorization**.



Dynamic Categorization is enabled by default. To disable it, clear the checkbox. If dynamic categorization is disabled, then the SG appliance does not contact the dynamic categorization service, even when no category is found for a URL in the database, and any dynamic categorization properties specified in policy are ignored. If dynamic categorization is enabled, it is only invoked while BCWF is in use

2. To change the Dynamic Categorization Settings, select one of the following:
   a. **Do not categorize dynamically**. The loaded database is consulted for category information. URLs not found in the database show up as category **none.** Dynamic categorization is still possible, but only occurs when explicitly invoked by policy.
   b. **Categorize dynamically in the background**. Background mode incurs only the bandwidth cost. In background mode once a call is made to the dynamic categorization service, the URL request immediately proceeds without waiting for the external service to respond. The system category *pending* is assigned to the request, indicating that the policy was evaluated with potentially incomplete category information.

      After they are received, the results of dynamic categorization are entered into a categorization cache (as are the results of real-time requests). This cache ensures that any subsequent requests for the same or similar URLs can be categorized quickly, without needing to query the external service again.
   c. **Categorize dynamically in real-time** (default). Real-time mode incurs both bandwidth and latency costs. The advantage of real-time mode dynamic categorization is that Blue Coat policy has access to the results of dynamic categorization, which means that policy decisions are made immediately upon receiving all available information.
3. Select **Apply** to commit the changes to the SG appliance.

About Dynamic Categorization States

Dynamic Categorization has three states:

- Enabled: The Dynamic Categorization service attempts to catgorize unrated Web sites.
- Disabled: If the Dynamic Categorization service is disabled, the SG appliance does not make any contact with the dynamic categorization service, regardless of any other installed policy.
- Suspended: If your BCWF filter expires and Dynamic Categorization is enabled, the service enters a suspended state. After the BCWF license is updated, the service returns to enabled status.

To view the Dynamic Categorization status (CLI only):

At the (`config`) prompt, enter the following command:

```
SGOS# (config content-filter) view
Provider:                       Blue Coat
.
.
.
Dynamic Categorization:
  Service:       Enabled/Disabled/Suspended <---one state is displayed
```

Diagnostics

Diagnostics allows you to see all categories available for use in policy or test a URL against the database. Categories are not displayed for a vendor or local database if no database has been downloaded.

To see all available categories:

1. **On the Configuration > Content Filtering > General page,** click **View Categories**.
2. To see what categories a Web site is assigned by your current configuration, enter the URL into the **URL field** and click **Test**.

Related CLI Syntax to Manage the BCWF Database

- To enter configuration mode:

```
SGOS#(config) content-filter
```

- The following subcommands are available:

```
SGOS#(config content-filter) provider bluecoat {enable | disable}
SGOS#(config content-filter) provider local lookup-node {always |
uncategorized}
SGOS#(config content-filter) categories
SGOS#(config bluecoat) download ?
SGOS#(config bluecoat) service {enable | disable | mode}
SGOS#(config bluecoat) no download
SGOS#(config bluecoat) {exit | view}
SGOS#(config content-filter) test-url url
```

Section C: Configuring a Local Database

This section describes how to select and refer to a local database and how to schedule the database update schedule.

Selecting the Local Database and Downloading the Database

Two main reasons to use a local database instead of a policy file for defining categories are:

- A local database is more efficient than policy if you have a large number of URLs.
- A local database separates administration of categories from policy. This separation is useful for three reasons:
  - It allows different individuals or groups to be responsible for administrating the local database and policy.
  - It keeps the policy file from getting cluttered.
  - It allows the local database to share categories across multiple boxes that have different policy.

However, some restrictions apply to a local database that do not apply to policy definitions:

- No more than 200 separate categories are allowed.
- Category names must be 32 characters or less.
- A given URL pattern can appear in no more than four category definitions.

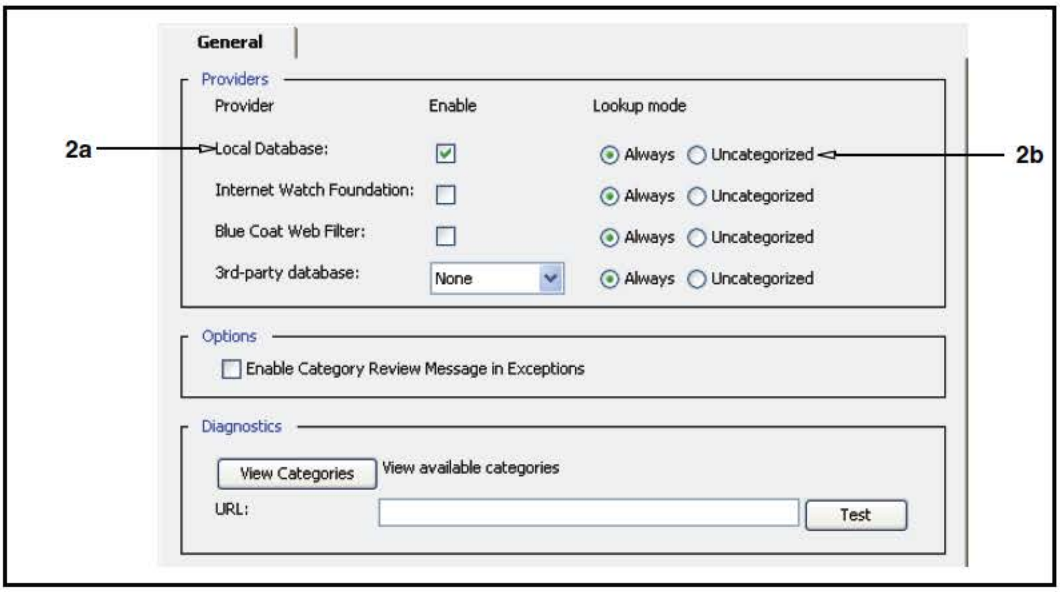
You can use any combination of the local database, policy files, or the VPM to manage your category definitions. See "Applying Policy to Categorized URLs" on page 37 for more information. You can also use both a local database and a third-party vendor for your content filtering needs.

Note: Blue Coat recommends locating your local database on the same server as any policy files you are using.

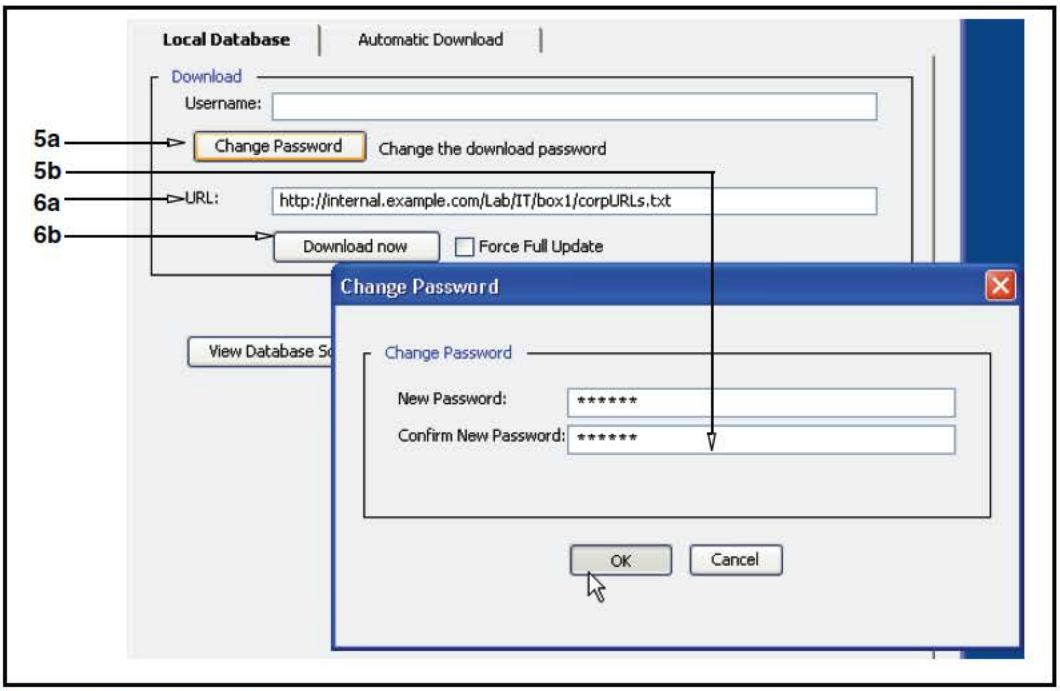
To configure local database content filtering:

1. Select **Configuration > Content Filtering > General**.

Section C: Configuring a Local Database



2. Select the local database as the provider:
   a. Select **Use Local Database**.
   b. (Optional) Select the **Lookup Mode**. The default is **Always**, which specifies that the database is consulted on *every* categorization attempt. **Uncategorized** specifies that the database lookup is skipped if the URL already has categories assigned by another content filtering database (IWF, BCWF, or third-party—in that order) employed on the SG appliance. The exception is if one of the content filtering database types is configured as **Uncategorized**, it is skipped.
3. Select **Apply** to commit the changes to the SG appliance.
4. Select **Configuration > Content Filtering > Local Database**.



5. If the database is located on a server that requires a password for access, you must configure the SG appliance to use that password when accessing the database:
   a. Click **Change Password**; the Change Password dialog appears.
   b. Enter your password and click **OK**.
6. Download the database:
   a. In the **URL** field, enter the location of the file to be downloaded.
   b. Click **Download Now**. The Commit Status dialog displays.

```
Download log:
  Local database download at: 11 Jun 2006 19:29:39 UTC
  Downloading from ftp://1.1.1.1/list-1000000-cat.txt
  Download size:      16274465
  Database date: 5 Jun 2006 19:31:58 UTC
  Total URL patterns: 1000000
  Total categories:   10
```

   c. Click **OK**.
7. Select **Apply** to commit the changes to the SG appliance.

Future Downloads

You can return to this screen at any time and download a database on demand (independent of the automatic download feature, which is described in the next section).

Ordinarily, the SG appliance checks to see if the database has changed before initiating a download. If the database is the most current, no download is performed. If an incremental update is available on the server, then it is downloaded (an incremental update contains only the changes between the current installed version and the latest

published version of the database, and is much smaller than a full copy of the database). You can override this process and force a download of the full database by selecting **Force Full Update**.

Scheduling Automatic Downloads for a Local Database

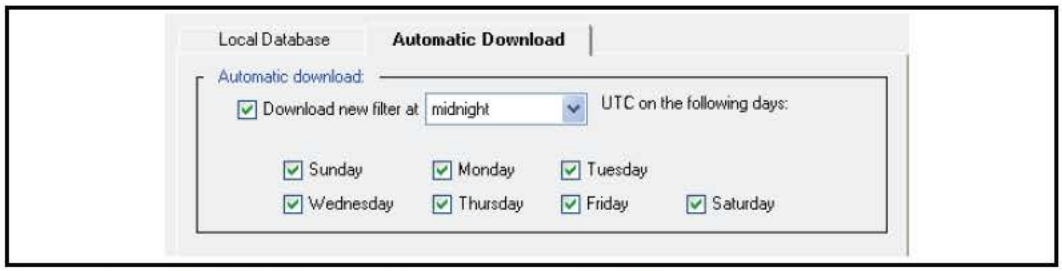
You can specify which days and times the Local Database is downloaded.

When the database is downloaded, a log is available that includes the information about how the database was updated, but in a more detailed form. You can view the download log (Statistics>Advanced>Content Filter Service) or the CLI (SGOS#(config) `show content-filter status`).

Note: By default, the automatic download setting is enabled (for every day at midnight, UTC) and does not need to be configured unless you want to change the schedule or disable auto-download.

To schedule local database automatic download times:

1. Select **Configuration > Content Filtering > Local Database > Automatic Download**.



2. To change the default times (every day at midnight):
   a. From the **Download new filter at** drop-down list, select a time of day.
   b. Deselect days as required.
3. Select **Apply** to commit the changes to the SG appliance.

Diagnostics

Allows you to see all categories available for use in policy or test a URL against the database. Categories are not displayed for a vendor or local database if no database has been downloaded.

To see all available categories:

1. **On the Configuration > Content Filtering > General page**, click **View Categories**.
2. To see what categories a Web site is assigned by your current configuration, enter the URL into the **URL field** and Click **Test**.

Related CLI Syntax to Configure Content Filtering

- To enter configuration mode:
  `SGOS#(config) content-filter`
- The following subcommands are available:
  `SGOS#(config content-filter) provider local {enable | disable}`

```
SGOS#(config content-filter) provider local lookup-node {always |
uncategorized}
SGOS#(config content-filter) categories
SGOS#(config local) download ?
SGOS#(config local) source
SGOS#(config local) clear
SGOS#(config local) {view | exit}
SGOS#(config content-filter) test-url url
```

Section D: Configuring Internet Watch Foundation

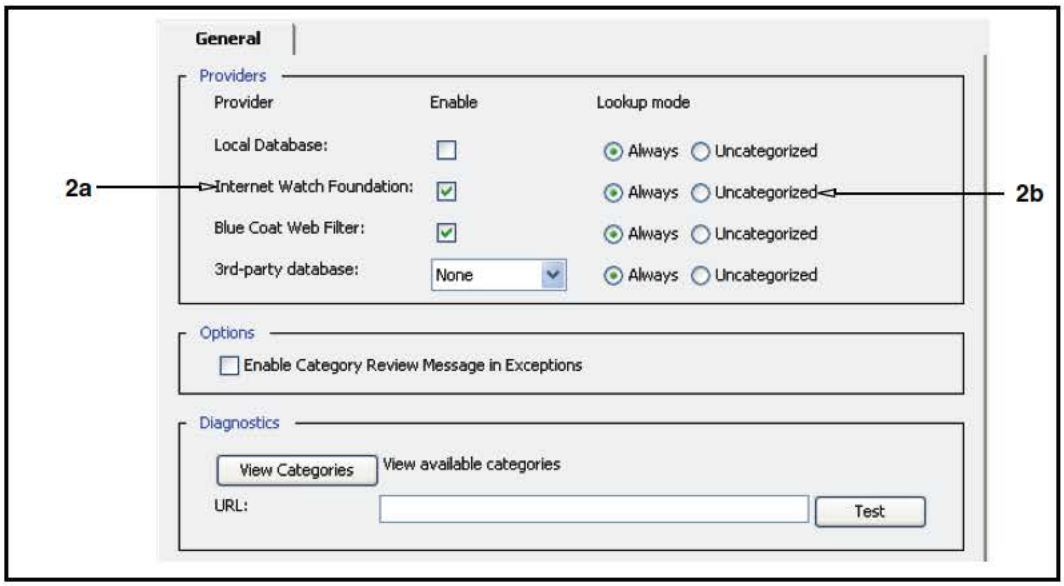
This section describes how to select the Internet Watch Foundation (IWF) database and how to schedule the database update schedule.

The IWF is a non-profit organization that provides to enterprises a list of known child pornography URLs. The IWF database features a single category called **IWF-Restricted**, which is detectable and blockable using policy. IWF can be enabled along with other content filtering services.

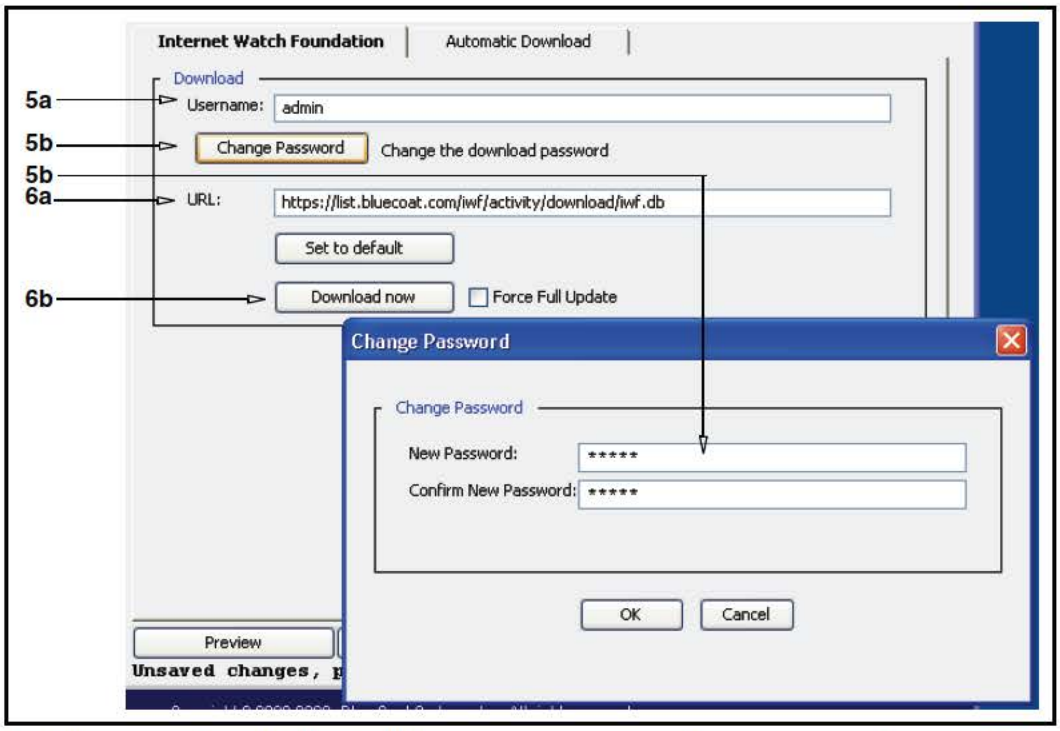
Selecting the IWF Database

To configure IWF content filtering:

1. Select **Configuration > Content Filtering > General**.



2. Select a provider:
   a. Select **Internet Watch Foundation**.
   b. (Optional) Select the **Lookup Mode**. The default is **Always**, which specifies that the database is consulted on *every* categorization attempt. **Uncategorized** specifies that the database lookup is skipped if the URL already has categories assigned by another content filtering database (local, IWF, or BCWF—in that order) employed on the SG appliance. The exception is if one of the content filtering database types is configured as **Uncategorized**, it is skipped.
3. Select **Apply** to commit the changes to the SG appliance.
4. Select **Configuration > Content Filtering > IWF**.



5. When you subscribed to the IWF Service, you received a username and password for access to download updates.
   a. In the **Username** field, enter your username.
   b. (Optional) Click **Change Password**; the Change Password dialog appears. Enter your password and click **OK**. (If you are in the trial period, no username or password is required.)
6. Download the database:
   a. The default database download location is displayed in the **URL** field.

   Note: Only enter a new URL if instructed. Otherwise, accept the default.

   b. Click **Download Now**. The IWF Installation status dialog box displays with the message: **IWF download in progress**.

      When the operation is complete, the dialog changes to indicate installation status.
   c. Click **Results** to see the IWF download log:

```
Download log:
Blue Coat download at: Thu, 08 Jun 2006 00:04:06 UTC
Downloading from https://list.bluecoat.com/iwf/activity/download/
iwf.db
Requesting differential update
Differential update applied successfully
Download size:      84103448
Database date:      Wed, 07 Jun 2006 08:11:51 UTC
Database expires:   Fri, 07 Jul 2006 08:11:51 UTC
Database version:   2005040
```

d. Click **OK**.

7. Select **Apply** to commit the changes to the SG appliance.

Future Manual Downloads

You can return to this screen at any time and download a database on demand (independent of the automatic download feature, which is described in the next section). Ordinarily, the SG appliance checks to see if the database has changed before initiating a download. If the database is the most current, no download is performed. If an incremental update is available on the server, then it is downloaded (an incremental update contains only the changes between the current installed version and the latest published version of the database, and is much smaller than a full copy of the database). You can override this process and force a download of the full database by selecting **Force Full Update**.

Note: Because the incremental update process carefully verifies the update before and after applying it, forcing a full update is, rarely, if ever necessary. Routinely forcing a full update consumes excess download bandwidth and does not improve reliability.

Scheduling Automatic Downloads for IWF

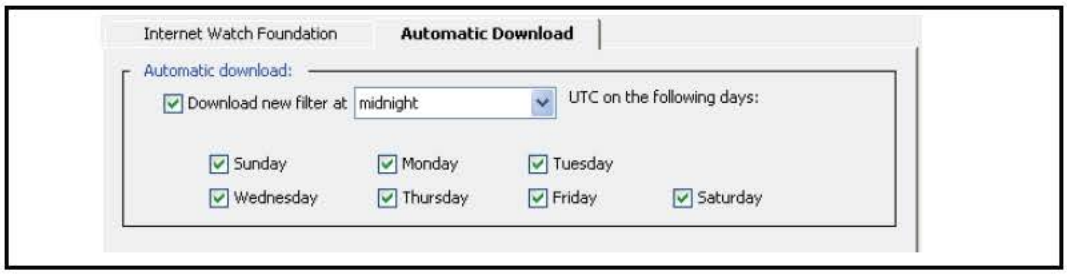
You can specify which days and times the IWF database is downloaded. Because sites become stale quickly, Blue Coat recommends downloading on an automatic schedule daily.

When the database is downloaded, a log is available that includes the information about how the database was updated, but in a more detailed form. You can view the download log (Statistics>Advanced>Content Filter Service) or the CLI (SGOS#(config) `show content-filter status`).

Note: By default, the automatic download setting is enabled (for every day at midnight, UTC) and does not need to be configured unless you want to change the schedule or disable auto-download.

To schedule IWF automatic download times:

1. Select **Configuration > Content Filtering > IWF > Automatic Download.**



2. To change the default times (every day at midnight):
   a. From the **Download new filter at** drop-down list, select a time of day.
   b. Deselect days as required.
3. Select **Apply** to commit the changes to the SG appliance.

Diagnostics

Allows you to test a URL against the database.

To test a URL:

1. **Select Configuration > Content Filtering > General**.
2. Enter the URL into the **URL field**.
3. Click **Test**.

Related CLI Syntax to Manage IWF

- To enter configuration mode:

```
SGOS#(config) content-filter
```

- The following subcommands are available:

```
SGOS#(config content-filter) provider iwf {enable | disable}
SGOS#(config content-filter) provider iwf lookup-node {always | uncategorized}
SGOS#(config bluecoat) download ?
SGOS#(config bluecoat) no download
SGOS#(config bluecoat) {exit | view}
SGOS#(config content-filter) test-url url
```

Section E: Configuring Third-Party Vendor Content Filtering

This section describes how to select and configure your preferred third-party vendor and how to schedule the database update schedule.

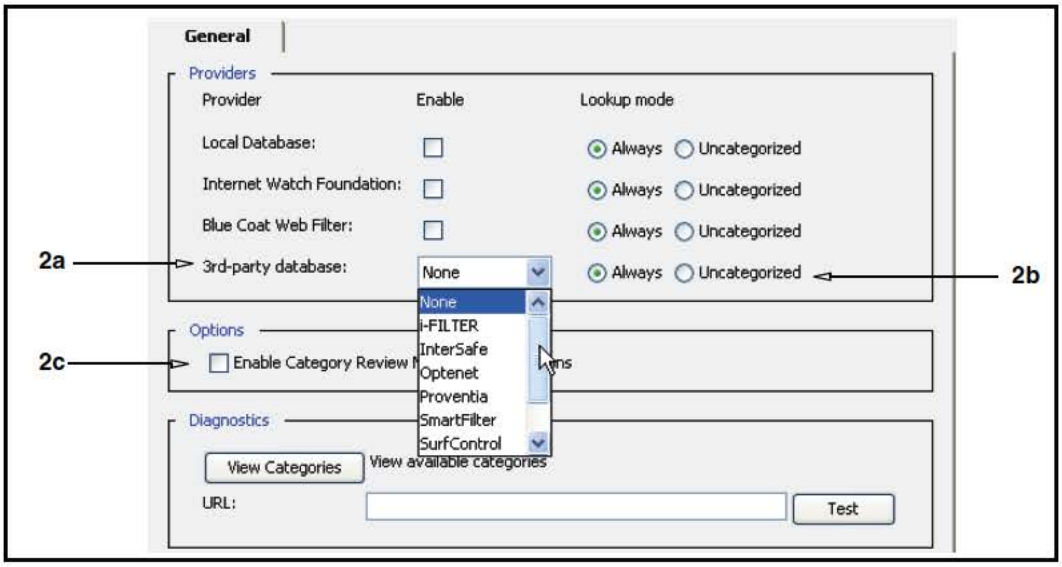
Most of the third-party vendor configuration tasks are identical, but there are a few with vendor-specific options. As you follow the procedures, you are prompted to proceed to another section for these vendors to continue the configuration.

Selecting the Provider and Downloading the Database

This procedure assumes you have a valid account with your preferred vendor.

To configure third-party content filtering:

1. Select **Configuration > Content Filtering > General**.



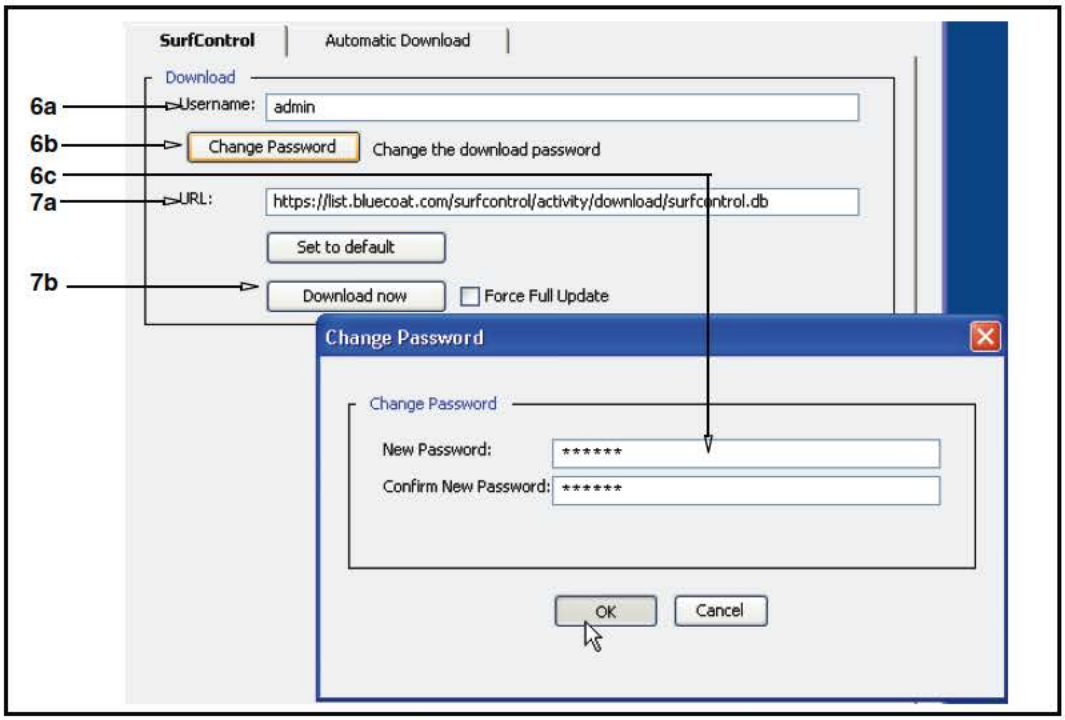
2. Select a provider:
   a. From the **Use 3rd Party Content Filters** drop-down list, select your preferred vendor.
   b. (Optional) Select the **Lookup Mode**. The default is **Always**, which specifies that the database is consulted on *every* categorization attempt. **Uncategorized** specifies that the database lookup is skipped if the URL already has categories assigned by another content filtering database (local, IWF, or BCWF—in that order) employed on the SG appliance. The exception is if one of the content filtering database types is configured as **Uncategorized**, it is skipped.

c. (Optional and applicable for SmartFilter *only*) Select **Enable Category Review Message in Exceptions**. Adds a link to the default content filter exception page that can be used to request review of the categories assigned to a blocked URL.

Two substitutions (`$(exception_category_review_url)` and `$(exception_category_review_message)`) are automatically appended to the `help` element of all exception definitions. For information on using the `$(exception.help)` element, refer to *Volume 7: VPM and Advanced Policy*.

Note: The substitution values are empty if the provider was not consulted for categorization, or if the categorization process failed due to an error.

3. Select **Apply** to commit the changes to the SG appliance.
4. Proceed accordingly:
   - **SmartFilter:** Continue with: "Configuring SmartFilter" on page 32.
   - **Websense**: Continue with : "Configuring Websense (on-box)" on page 33.
   - **i-Filter, InterSafe, Optenet, Proventia, SurfControl**, or **Webwasher**: Continue with Step 5.
5. Select **Configuration > Content Filtering >** *vendor*:



6. (This example uses SurfControl.) If the database is located on a server that requires a password for access, you must configure the SG appliance to use that password when accessing the database:
   a. Enter your third-party vendor username.

b. Click **Change Password**; the Change Password dialog appears.

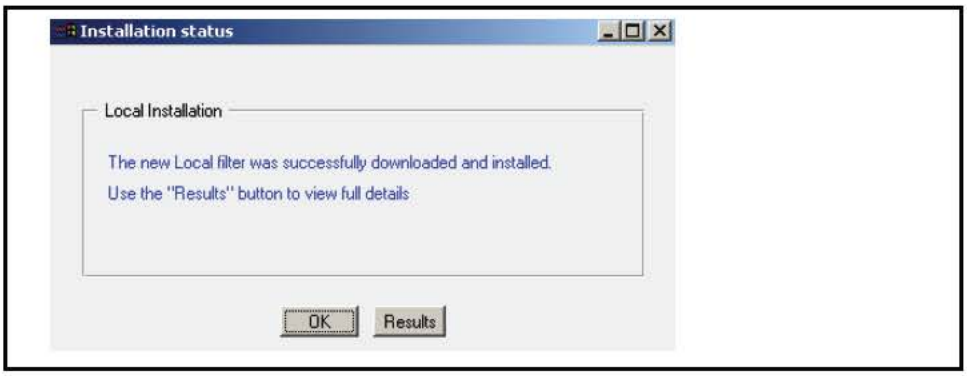
c. Enter your password and click **OK**. (If you are in the trial period, no username or password is required.)

7. Download the database:

a. The default database download location is displayed in the **URL** field. If you have been instructed to use a different URL, enter it here (optional: click **Set to default** to always use this location).

b. Click **Download Now**. The Installation Status dialog box displays with the message **Local filter download in progress**.

When the operation is complete, the dialog changes to indicate installation status.



c. Click **Results** to see the completion message:

```
Download log:
  IWF download at: 10 Jun 2006 20:16:16 UTC
  Downloading from https://list.bluecoat.com/.../download/iwf.db
    Warning: Unable to determine current database version; requesting
full update
  Download size:       8106572
  Database date: 08 Jun 2006 07:02:08 UTC
  Database expires: 10 Oct 2006 07:02:08 UTC
  Database version:  3
```

d. Click **OK**.

8. Select **Apply** to commit the changes to the SG appliance.

9. Continue with "Scheduling Automatic Downloads for a Third-Party Database" on page 35.

Future Downloads

You can return to this screen at any time and download a database on demand (independent of the automatic download feature, which is described in the next section).

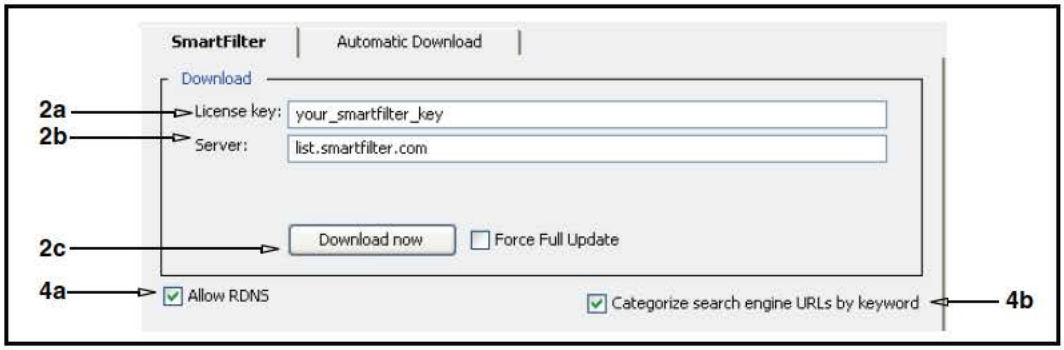
Ordinarily, the SG appliance checks to see if the database has changed before initiating a download. If the database is the most current, no download is performed. If an incremental update is available on the server, then it is downloaded (an incremental update contains only the changes between the current installed version and the latest published version of the database, and is much smaller than a full copy of the database). You can override this process and force a download of the full database by selecting **Force Full Update**.

Configuring SmartFilter

The SmartFilter database configuration screen contains unique options.

Configure SmartFilter:

1. Select **Configuration > Content Filtering > SmartFilter:**



2. Configure SmartFilter:
   a. In the **License key** field, enter the customer serial number assigned you by Secure Computing.
   b. In the **Server** field, the default server is displayed. If you have been instructed to use a different server, enter the hostname or IP address here.
   c. Click **Download now**. The SmartFilter Installation status dialog box displays with the message **SmartFilter download in progress**.

      When the operation is complete, the dialog changes to indicate installation status.
   d. Click **Results** to see the completion message:

```
Download log:
      SmartFilter download at: 06 Apr 2006 20:27:14 UTC
      Checking incremental update
        Warning: Unable to open input control list
        Warning: Unable to open installed control list
      Downloading full control file
        SmartFilter download at: 06 Apr 2006 20:27:14 UTC
        Downloading from http://example.com/...version=4.0
      Download size:      45854194
      Database version:   95
      Database date: 06 Apr 2006 07:05:01 UTC
      Database expires: 11 May 2006 07:05:01 UTC
```

> **Note:** The first time you download a SmartFilter database, warnings appear in the results message under `Checking incremental update`. These are expected, and represent the normal process of checking to see if an incremental update is possible. The next time you download a SmartFilter database, the SG appliance checks the previously downloaded database, and downloads only what is necessary to keep the database current.

3. Click **Apply**.
4. SmartFilter features the following optional configurations:

a. SmartFilter lookups can require use of reverse DNS to properly categorize a Web site. To disable the use of reverse DNS by SmartFilter, deselect **Allow RDNS**.

> **Important:** Disabling reverse DNS prevents SmartFilter from correctly classifying some sites and can increase the likelihood of the SG appliance serving inappropriate content.

b. By default, SmartFilter categorizes search engines based on keywords in the URL query. To disable this setting, deselect **Categorize search engine URLs based on keywords**.

> **Important:** Leaving keywords enabled can cause unexpected results. For example, the keyword *electoral college* falls into the educational category.

5. Select **Apply** to commit the changes to the SG appliance.
6. Continue with "Scheduling Automatic Downloads for a Third-Party Database" on page 35.

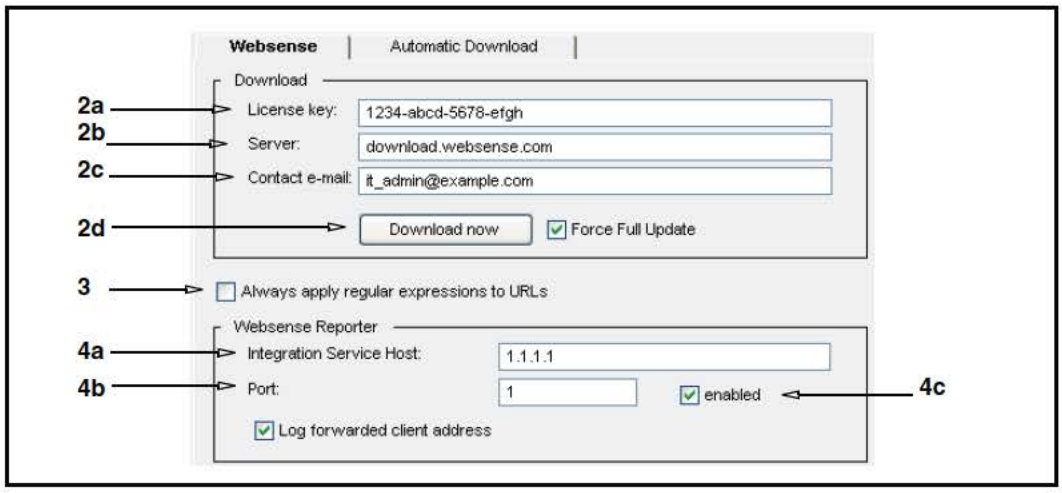
Configuring Websense (on-box)

The Websense database configuration screen contains unique options.

> **Note:** Websense databases contain a category called **User-Defined** to support locally-specified categorizations on other platforms. Do not use this category on the SG appliance. Instead, define your own categories through the ProxySG and assign URLs to them using Policy (see page "Defining Custom Categories in Policy" on page 40), or using a local category database (refer to *Volume 5: Securing the Blue Coat SG Appliance*).

To configure Websense (on-box):

1. Select **Configuration > Content Filtering > Websense**.



2. Configure Websense:

a. In the **License Key** field, enter the key assigned to you for downloading the Websense database.

b. In the **Server** field, the default server is displayed. If you have been instructed to use a different server, enter the hostname or IP address here.

c. (Optional) In the **Contact e-mail** field, enter an e-mail address by which Websense can contact you.

d. Click Download now. The Websense Installation status dialog box displays with the message **Websense download in progress**.

e. Click **Apply** to view the Websense download log:

```
Download log:
Websense download at: Fri, 09 Jun 2006 20:32:35 UTC
     No database is currently installed
     Attempting full download
     Downloading from download.websense.com
     Processing download file
        Retrieved full update
     Download size:      147079939
     Database version:   82300
     Database date:      2006-06-8
     License expires:    Sun, 05 Nov 2006 08:00:00 UTC
     License max users:  25
     Licenses in use:    0
     Library version:    3.2.0.0 [BCSI rev A]
```

f. Click **OK**.

3. (Optional) **Always apply regular expressions to urls**:

   Select this option to force an additional regular expression lookup for each URL to be categorized. Normally, regular expression lookups are done only when no category is found in the Websense database. If this option is selected, regular expression lookups always occur, even for categorized URLs. Selecting this option can cause a significant reduction in lookup performance, but allow certain sites (such as translation, search engine, and link-cache sites) to be categorized more accurately.

4. To use the Websense Reporter, you must enable the Websense Integration Service.

a. In the **Integration Service Host** field, enter the Integration Service Host IP (which has the same IP address as the Websense Log Server).

b. In the **Port** field, specify the port of the Websense Integration Service. It must be between 0 and 65535 and match the port selected on the Integration Service host.

c. Select **Enabled** to enable the service.

d. (Optional) Select **Log forwarded client address**. Normally, the SG logs the actual client IP address to the Websense Reporter log. You can configure the SG to log an address obtained from the X-Forwarded-For HTTP Header (if present and valid) instead. This is useful in some specific network topologies.

Note: The Policy Server, the Log Server, and Reporter must be installed and enabled on your PC before Reporter can be used. For information on Websense products, refer to: http://www.websense.com/support/documentation/integrationservice.

You must also set up access logging on the SG appliance with Websense as the client. For more information on configuring a Websense access logging client, refer to *Volume 9: Access Logging*.

---

5. Select **Apply** to commit the changes to the SG appliance.

6. Proceed to the "Scheduling Automatic Downloads for a Third-Party Database" on page 35.

Scheduling Automatic Downloads for a Third-Party Database

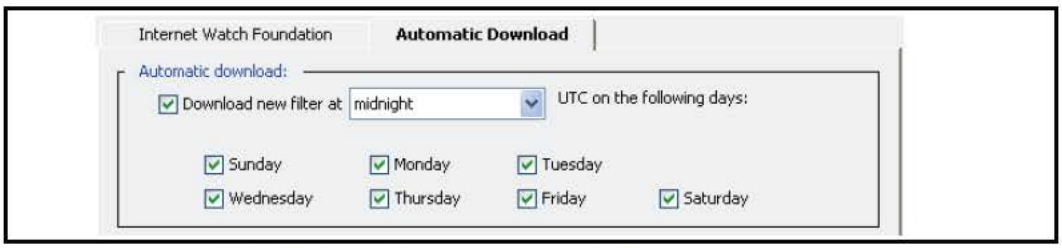
You can specify which days and times the database is downloaded. Because sites become stale quickly, Blue Coat recommends downloading on an automatic schedule frequently.

When the database is downloaded, a log is available that includes the information about how the database was updated, but in a more detailed form. You can view the download log (Statistics>Advanced>Content Filter Service) or the CLI (SGOS#(config) `show content-filter status`).

Note: By default, the automatic download setting is enabled (for every day at midnight, UTC) and does not need to be configured unless you want to change the schedule or disable auto-download.

To schedule local database automatic download times:

1. Select **Configuration > Content Filtering >** *vendor* **> Automatic Download**.



2. To change the default times (every day at midnight):

a. From the **Download new filter at** drop-down list, select a time of day.
b. Deselect days as required.

3. Select **Apply** to commit the changes to the SG appliance.

Diagnostics

Allows you to see all categories available for use in policy or test a URL against the database. Categories are not displayed for a vendor or local database if no database has been downloaded.

To see all available categories or test a URL:

1. **On the Configuration > Content Filtering > General page,** click **View Categories**.
2. To see what categories a Web site is assigned by your current configuration, enter the URL into the **URL field**.
3. Click **Test**.

Related CLI Syntax to Manage Third-Party Vendor Content Filtering

- To enter configuration mode:

```
SGOS#(config) content-filter
```

- The following subcommands are available:

```
SGOS#(config content-filter) {i-filter | intersafe | optenet |
proventia | smartfilter | surfcontrol | websense | webwasher}
SGOS#(config content-filter) provider 3rd-party lookup-mode {always |
uncategorized}
SGOS#(config vendor) download ?
SGOS#(config vendor) view
SGOS#(config smartfilter) download license license_key
SGOS#(config smartfilter) download server ip_address_or_hostname
SGOS#(config smartfilter) allow-rdns | no allow-rdns
SGOS#(config smartfilter) use-search-keywords
SGOS#(config websense) download email-contact e-mail_address
SGOS#(config websense) download server ip_address_or_hostname
SGOS#(config websense) download license license_key
SGOS#(config websense) {always-apply-regexes | no always-apply-
regexes}
SGOS#(config websense) integration-service {enable | disable}
SGOS#(config websense) integration-service host ip_address_or_hostname
SGOS#(config websense) integration-service port {0-65535}
```

Section F: Applying Policy

This section discusses the interactivity between content filtering categories and the application of control policies.

Applying Policy to Categorized URLs

Policy is applied to categories the same way as individual URLs: create policies that restrict, allow, and track access. Policy rules are created by composing Blue Coat Content Policy Language (CPL) or with the Visual Policy Manager (VPM).

Note: If you have extensive category definitions, Blue Coat recommends that you put them into a local database rather than into a policy file. The local database stores custom categories in a more scalable and efficient manner, and separates the administration of categories from policy. You can manage any combination of the local database, policy files, and Visual Policy Manager rules. See "Section C: Configuring a Local Database" on page 20.

The policy trigger `category=` is used to test the category or categories assigned to the request URL, and thus make a policy decision. For example, to block all requests for URLs that are categorized as Sports:

```
DENY category=Sports
```

The following example demonstrates a condition that is true when a request contains the Websense content categories Sexuality and Drugs:

```
<proxy>
  category=(sexuality, drugs)
```

You can block multiple categories with a single rule:

```
category=(Sports, Gambling, Shopping) exception(content_filter_denied)
```

In this example, three categories are blocked and instead the predefined exception page *content_filter_denied* is served; by default this indicates that the request was denied due to its content and specifies the categories found.

The following example shows a condition that includes an extensive number of categories:

```
category=(Abortion, Activist, Adult, Gambling, Illegal, Hacking,
Militancy, Racism, Shopping, Tasteless, Violence, Weapons)
```

URLs that are not categorized are assigned the system category `none`. This is *not* an error condition; many sites (such as those inside a corporate intranet) are unlikely to be categorized by a commercial service. Use `category=none` to detect uncategorized sites and apply relevant policy. The following example disallows access to uncategorized sites outside of the corporate network:

```
define subnet intranet
  10.0.0.0/8 ; internal network
  192.168.123.45; external gateway
end
```

```
<proxy>
  ; allow unrestricted access to internal addresses
  ALLOW url.address=intranet

  ; otherwise (internet), restrict Sports, Shopping and uncategorized
sites
  DENY category=(Sports, Shopping, none)
```

Such category tests can also be combined with other types of triggers to produce more complex policy, such as:

- Restrict access by category and time: block sports from 6 am to 6 pm:

  ```
  category=Sports time=0600..1800 DENY
  ```

- Restrict by category and user identity: only members of the group Sales are permitted to visit Shopping sites:

  ```
  category=Shopping group=!Sales DENY
  ```

- Require special authentication for access to certain categories:

  ```
  category=Hacking authenticate(restricted_realm)
  ```

  where *restricted_realm* is an authentication realm you have configured.

- Log certain types of access:

  ```
  category=Adult action.Log_adult_site_access(yes)
  ```

  where *Log_adult_site_access* is a policy action defined elsewhere that records extra information about this request in the event log.

Typically, `category=` can be used in policy anywhere that a basic URL test can be used. Refer to *Volume 11: Blue Coat SG Appliance Content Policy Language Guide* for more details.

Depending on which provider you have selected and whether you have defined any of your own categories in policy (see "Defining Custom Categories in Policy" on page 40), you have a number of possible category names that can be used with `category=`. To review the valid category names, use the `categories` CLI command or click **View Categories** in the Management Console: **Configuration > Content Filtering > General**.

The `category=` expressions are normally put in `<Proxy>` Layers (VPM: **Web Access Layers**) because the goal of content-filtering policy is to control requests from users. They can also be used in `<Cache>` (VPM: **Web Content Layers**) Layers. Either way, policy is enforced on all user requests.

It is possible for an attempt to categorize a URL to fail—for example, if no database is loaded, your license is expired, or if a system error occurs. In such a case, the category is considered *unavailable* and triggers such as:

```
category=Sports
```

are false, even if the URL is actually a sports site, because the SG appliance is unable to determine the category. When the policy depends on the category of a URL, you do not want such errors to inadvertently allow ordinarily restricted content to be served by the SG appliance. You can control how the SG appliance treats these situations with the condition:

```
category=unavailable
```

which is true in these cases. In continuing with the example, to make sure that Sports is always blocked, even when errors occur (this is a mode of operation called *fail-closed*), use a rule such as:

```
category=(sports, unavailable) exception(name_of_exception page)
```

This rule is true if the category is sports or if the category could not be determined, and in either case the proper exception page is served instead of the restricted content.

The category `unlicensed` is assigned in addition to *unavailable* when the failure to categorize occurred because of license expiry. That can be caused by the expiration of your Blue Coat license to use content filtering, or because of expiration of your license from the provider. You can use

```
category=unlicensed
```

to detect this situation as a distinct case from other causes of unavailability.

You can also use this feature with custom exception pages (refer to *Volume 7: VPM and Advanced Policy*):

```
<proxy>
  category=sports time=0800..1800 exception(sports_during_bus_hrs)
  category=unlicensed exception(contact_admin_re_license)
  category=unavailable exception(content_filter_unavailable)
```

where *sports_during_bus_hrs* is a custom exception page you have created to respond to requests for Sports pages between 8 am and 6 pm local time.

contact_admin_re_license is another page that instructs the user to inform the administrator about license expiry, and is served if a license check fails. When the category is unavailable for some other reason, the pre-defined exception (*content_filter_unavailable*) is served.

The most common reason (other than license expiry) why categories are unavailable is that a provider is selected but no database is installed. Barring hardware or network problems that might cause a downloaded database to become corrupted and unreadable, it is unlikely that the database will suddenly become unavailable.

To define policies on the SG appliance, use either the VPM or manually edit Policy files.

Content filtering policies are usually found in `<Proxy>` and `<Cache>` layers.

If you are using content filtering to manage a type of content globally, create these rules in the `<Cache>` layer.

However, if your content filtering policy is dependent on user identity or request characteristics, create these rules in the `<Proxy>` layer.

Using Content Filtering Vendors with Blue Coat Policies

The SG appliance provides the ability to define flexible Web access and control policies. With content filtering, you can set up policies to provide a customized level of Web-site access control. With vendor-based content filtering, these policies use and can supplement vendor categories. By supplementing content-filtering vendor categories, you can further refine the type of content filtering the SG appliance performs. For example, if **Travel** is a vendor-defined content category, you can define a policy that allows only Human Resources staff to access travel sites. You can define policies that filter by a variety of conditions, including category, protocol (including MMS and RTSP streaming protocols), time of day, and user or user groups.

Example

Policy: Limit employee access to travel Web sites.

The first step is to rephrase this policy as a set of rules. In this example, the model of a general rule and exceptions to that rule is used:

- Rule 1: All users are denied access to travel sites
- Rule 2: As an exception to the above, Human Resources users are allowed to visit Travel sites

Before you can write the policy, you must be able to identify users in the Human Resources group. You can do this with an external authentication server, or define the group locally on the SG appliance. For information on identifying and authenticating users, refer to *Volume 5: Securing the Blue Coat SG Appliance*.

In this example, a group called `human_resources` is identified and authenticated through an external server called `my_auth_server`.

This then translates into a fairly straightforward policy written in the local policy file:

```
<proxy>
; Ensure all access is authenticated
  Authenticate(my_auth_server)
<proxy>
; Rule 1: All users denied access to travel
  DENY category=travel
<proxy>
; Rule 2: Exception for HR
  ALLOW category=travel group=human_resources
  DENY category=sites
```

Example

Policy: Student access to Health sites is limited to a specified time of day, when the Health 100 class is held.

This time the policy contains no exceptions:

- Rule 1: Health sites can be accessed Monday, Wednesday, and Friday from 10-11am.
- Rule 2: Health sites can not be accessed at other times.

```
define condition Health_class time
  weekday=(1, 3, 5)  time=1000..1100
end
<proxy>
; 1) Allow access to health while class in session
   ALLOW category=health condition=health_class_time
; 2) at all other times, deny access to health
   DENY category=health
```

Defining Custom Categories in Policy

You can use CPL to create your own categories and assign URLs to them. This is done with the `define category` construct (for more complete information on the define category construct, refer to *Volume 11: Blue Coat SG Appliance Content Policy Language Guide*). To add URLs to a category, list them in the definition. You only need to specify a partial URL:

- hosts and subdomains within the domain you specify will automatically be included
- if you specify a path, all paths with that prefix are included (if you specify no path, the whole site is included)

Example:

```
define category Grand_Canyon
  kaibab.org
  www2.nature.nps.gov/air/webcams/parks/grcacam
  nps.gov/grca
  grandcanyon.org
end
```

Any URL at `kaibab.org` is now put into the `Grand_Canyon` category (in addition to any category it might be assigned by a provider). Only those pages in the `/grca` directory of `nps.gov` are put in this category.

Nested Definitions and Subcategories

You can define subcategories and nest category definitions by adding a `category=<name>` rule. To continue the example, you could add:

```
define category Yellowstone
  yellowstone-natl-park.com
  nps.gov/yell/
end
define category National_Parks
  category=Grand_Canyon; Grand_Canyon is a subcategory of
National_Parks
  category=Yellowstone; Yellowstone is a subcategory of National_Parks
  nps.gov/yose; Yosemite – doesn't have its own category (yet)
end
```

With these definitions, pages at `kaibab.org` are assigned TWO categories: `Grand_Canyon` and `National_Parks`. You can add URLs to the `Grand_Canyon` category and they are automatically added by implication to the `National_Parks` category as well.

Multiple unrelated categories can also be assigned by CPL. For instance, by adding:

```
define category Webcams
  www2.nature.nps.gov/air/webcams/parks/grcacam
end
```

the URL, `http://www2.nature.nps.gov/air/webcams/parks/grcacam/grcacam.htm`, will have three categories assigned to it:

- `Grand_Canyon` (because it appears in the definition directly)
- `National_Parks` (because `Grand_Canyon` is included as a subcategory)
- `Webcams` (because it also appears in this definition)

However, the other sites in the `Grand_Canyon` category are not categorized as `Webcams`. This can be seen by testing the URL (or any other you want to try) using the **Test** button on the Management Console or the `test-url` command in the CLI.

You can test for any of these categories independently. For example, the following example is a policy that depends on the above definitions, and assumes that your provider has a category called `Travel` into which most national park sites probably fall. The policy is intended to prevent access to travel sites during the day, with the exception of those designated `National_Parks` sites. But the `Grand_Canyon` webcam is an exception to that exception.

Example:

```
<proxy>
  category=Webcams DENY
  category=National_Parks ALLOW
  category=Travel time =0800..1800 DENY
```

Remember that you can use the **Test** button on the Management Console or the `test-url` command in CLI to validate the categories assigned to any URL. This can help you to ensure that your policy rules have the expected effect (refer to "Configuring Policy Tracing" in *Volume 11: Blue Coat SG Appliance Content Policy Language Guide)*.

If you are using policy-defined categories and a content-filter provider at the same time, be sure that your custom category names do not coincide with the ones supplied by your provider. You can also use the same names—this adds your URLs to the existing categories, and extends those categories with your own definitions. For example, if the webcam mentioned above was not actually categorized as Travel by your provider, you could do the following to add it to the Travel category (for the purpose of policy):

```
define category Travel  ; extending a vendor category
  www2.nature.nps.gov/air/webcams/parks/grcacam/ ; add the GC webcam
end
```

> **Note:** The policy definitions described in this section can also be used as definitions in a local database. See "Configuring a Local Database" on page 20 for information about local databases.

Notes

- When you use an expired database, the category **unlicensed** is assigned to all URLs and no lookups occur on the database. This can occur even if your download license with the database vendor is still valid, but you have not downloaded a database for a long time (databases expire after a certain number of days). You can view the date that your database expires (or expired) in the download log or by using the `view` command in the CLI.

  When you download a database, you can see the download log as soon as the download is complete. To see the download log when you download a database, click **Results** in the Installation Status dialog when the download is complete.

  To see the last download log without doing another download, enter the following CLI `(config)` commands:

  ```
  SGOS#(config) content-filter
  SGOS#(config content-filter) view
  ```

- When your license with the database vendor expires, you can no longer download. This does not have an immediate effect—you can still use the database you have for a period of time. But eventually, the database expires and you receive the category **unlicensed**, as described above.

- If a requested HTTPS host is categorized in a content filtering database, then filtering applies. However, if the request contains a path and the categorization relies on the host/relative path, content filtering only filters on the host name because the path is not accessible. This might result in a different categorization than if the host plus path were used.

Section F: Applying Policy

- If you receive an error message when downloading a content filtering database, check the error message (in the Management Console, click **Results** on the Installation status dialog; in the CLI, the results message appears in the event of an error). If you see an error message such as **ERROR: HTTP 401 - Unauthorized**, verify that you entered your username and password correctly. For example, the following error message was generated by entering an incorrect username and attempting to download a SmartFilter database:

```
Download log:
        SmartFilter download at: Thu, 08 Apr 2006 18:03:08 UTC
        Checking incremental update
           Checking download parameters
           Fetching:http://example.com/
           Warning: HTTP 401 - Unauthorized
        Downloading full control file
           SmartFilter download at: Thu, 08 Apr 2006 18:03:17 UTC
           Downloading from http://example.com/
           Fetching:http://example.com/
           ERROR: HTTP 401 - Unauthorized
           Download failed
        Download failed
Previous download:
...
```

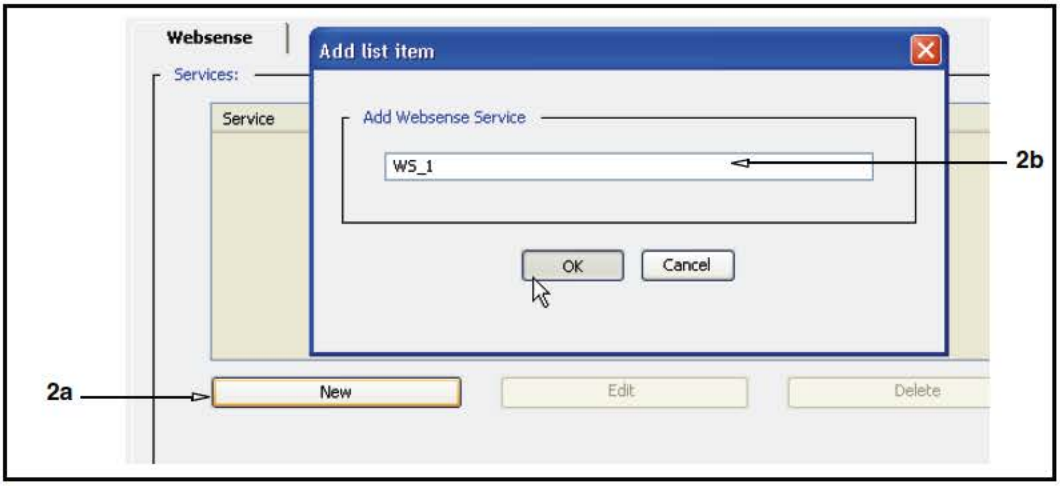
Section G: Configuring Websense Off-box Content Filtering

This section describes how to configure the SG appliance to communicate with a separate Websense server to perform content filtering tasks. This involves creating an external service on the SG appliance.

Note: The SG appliance supports Websense off-box server versions 4.4 and higher.

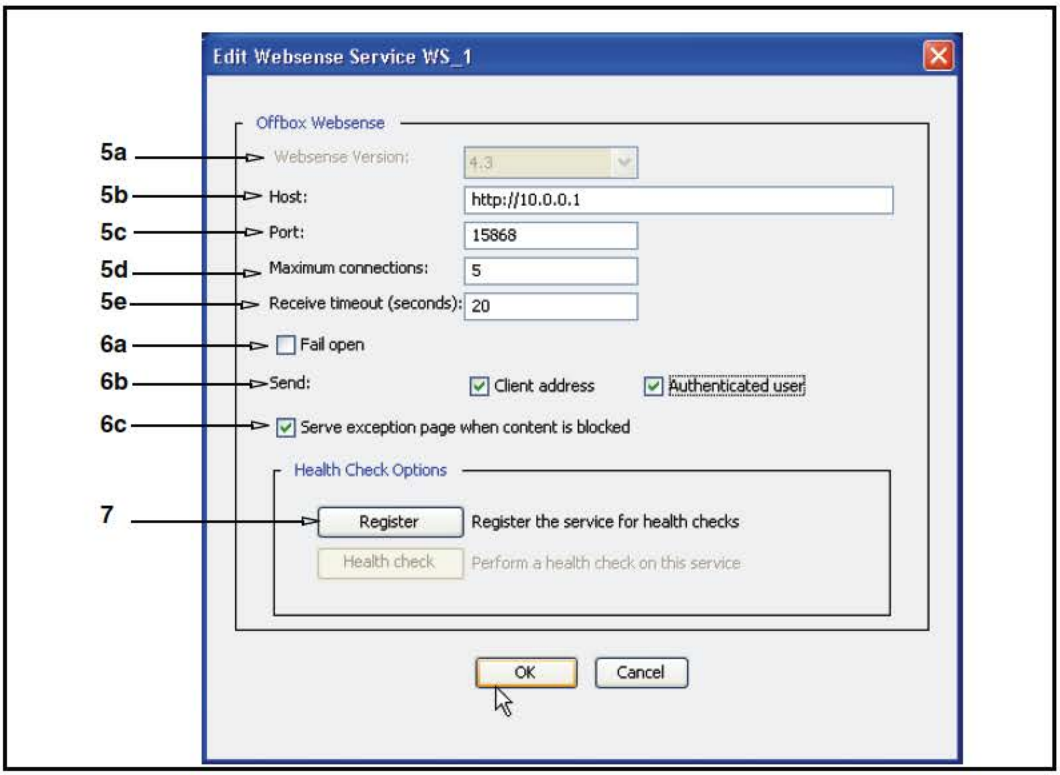
To configure Websense Off-box:

1. In the Management Console, select **Configuration > External Services > Websense**.



2. Add a new service:
   a. Click **New**. The Add list item dialog appears.
   b. Enter a name for the service. This example uses **WS_1**.
   c. Click **OK** to close the dialog.
3. Select **Apply** to commit the changes to the SG appliance.
4. Click **Edit**. The Edit Websense Service dialog appears.

Section G: Configuring Websense Off-box Content Filtering



5. Configure the service:
   a. Select the Websense server version: **4.3** or **4.4 and higher.**
   b. In the **Host** field, enter the hostname or IP address of the remote Websense server.
   c. In the **Port** field, enter the port number of the Websense server; or leave as is to accept the default (**15868**).
   d. In the **Maximum connections** field, enter the maximum number of connections. The range is a number from 1 to 65535. The default is **5**. Blue Coat recommends that the setting not exceed **200**.
   e. In the **Receive Timeout** field, enter the number of seconds the Proxy*SG* waits for replies from the Websense server. The range is 60 to 65535. The default timeout is **70** seconds.

6. The following are optional:
   a. **Fail open**—If a default Websense service is selected (from the **External Services > Websense** tab), a connection error with the Websense server results in requests and responses proceeding, as the default Websense service is subjected to policy.
   b. **Send client address**—Sends the client IP address to the Websense server.
   c. **Send Authenticated user**—Sends user information to the Websense server.
   d. **Serve exception page when content is blocked**—If the requested content is defined by Websense as inappropriate, the client receives a page with information stating the content is blocked. When this option is selected, the exception page originates from the Proxy*SG*; if not selected, the Websense server provides the exception page.
7. (Optional) For convenience, the Edit Websense Service dialog allows you to register a newly-created Websense service for health checking (this duplicates the functionality on the **Configuration > Health Checks > General** tab). Registering for health checking requires that a valid Websense server URL was entered.
   a. Click **Register**; a dialog prompts confirmation; click **OK**.
   b. You can also click **Health check** to perform an immediate health check on this service.

   For more information about health checks, refer to *Volume 6: Advanced Networking*.
8. Click **OK** to close the dialog.
9. Select **Apply** to commit the changes to the SG appliance.
10. (Optional) You can designate a default Websense service. On the **Configuration > External Services > Websense** tab, select a service from the **Default service to use** drop-down list.

Because this is an external service feature, you can create service groups that contain two or more Websense services. Then you can point the Proxy*SG* to the service group to allow for greater efficiency. See Chapter 4: "Configuring Service Groups" on page 71.

Related CLI Syntax to Configure Websense Off-box Content Filtering

- To enter configuration mode:

```
SGOS#(config) external-services
```

- The following subcommands are available:

```
SGOS# (config external-services) create websense service_name
SGOS# (config external-services) {edit | delete} service_name
SGOS# (config websense service_name) version {4.3 | 4.4}
SGOS# (config websense service_name) host {hostname | IP_address}
SGOS# (config websense service_name) port port_number
SGOS# (config websense service_name) max-conn number
SGOS# (config websense service_name) timeout timeout_seconds
SGOS# (config websense service_name) send {client-address |
authenticated-user}
SGOS# (config websense service_name) sense-categories
SGOS# (config websense service_name) apply-by-default
SGOS# (config websense service_name) fail-open
```

Section G: Configuring Websense Off-box Content Filtering

```
SGOS# (config websense service_name) test-url url
```

Chapter 3: ICAP

This chapter describes how to configure the SG appliance to interact with external ICAP and servers to provide content scanning and content transformation.

This chapter contains the following sections:

Section A: About Content Scanning

This section provides conceptual information regarding anti-virus (AV) scanning and the SG appliance solution.

When integrated with a supported ICAP server, such as the Blue Coat Blue Coat AV™, the SG appliance provides content scanning, filtering, and repair service for Internet-based malicious code. To eliminate threats to the network and to maintain caching performance, the SG appliance sends objects to the integrated ICAP server for checking, and saves the scanned objects in its object store. With subsequent content requests, the appliance serves the scanned object rather than rescan the same object for each request.

Supported ICAP Servers

The SG appliance with Blue Coat AV integration is a high-performance Web anti-virus (AV) solution.

The SG appliance also supports the following ICAP third-party ICAP implementations:

- Symantec AntiVirus Scan Engine (SAVSE)
- WebWasher
- Finjan Vital Security for Web

For the most current list of vendors and supported versions, refer to the *Blue Coat SGOS Release Notes* for this release.

Determining Which Files to Scan

In determining which files to scan, this integrated solution uses the content scanning server's filtering in addition to SG appliance capabilities. The following table describes the supported content types and protocols.

Table 3-1. Content Types Scanned By ICAP Server and the SG Appliance

ICAP Server supported content types	SG appliance supported protocols	Unsupported content protocols
All or specified file types, based on file extension, as configured on the server. Examples: `.exe` (executable programs), `.bat` (batch files), `.doc` and `.rtf` (document files), and `.zip` (archive files), or with specific MIME types.	• HTTP objects • FTP objects (uploads and downloads) • Transparent FTP responses	• Streaming content (for example, RTSP and MMS) • Live HTTP streams (for example, HTTP radio streams)
	HTTPS connections terminated at a Proxy*SG*.	HTTPS connections tunneled through a Blue Coat client-side SG appliance.

After the SG appliance retrieves a requested Web object from the origin server, it uses content scanning policies to determine whether the object should be sent to the ICAP server for scanning. If cached objects are not scanned or are scanned before a new pattern file was updated, the SG appliance rescans those objects upon:

- the next request for that object, or
- the object is refreshed.

With the SG appliance, you can define flexible, enterprise-specific content scanning policies. Consider the following example.

About Response Modification

The SG appliance sends the first part (a preview) of the object to the ICAP server that supports response modification. The object preview includes the HTTP request and response headers, and the first few bytes of the object. After checking those bytes, the ICAP server either continues with the transaction (that is, asks the SG appliance to send the remainder of the object for scanning) or sends a notification to the appliance that the object is clean and opts out of the transaction.

The ICAP server features and configuration determine how scanning works, including the following:

- Handling of certain objects, including those that are infected and cannot be repaired.
- Whether to attempt to repair infected files.
- Whether to delete infected files that cannot be repaired from the ICAP server's archive.

The following diagram illustrates the response modification process flow.

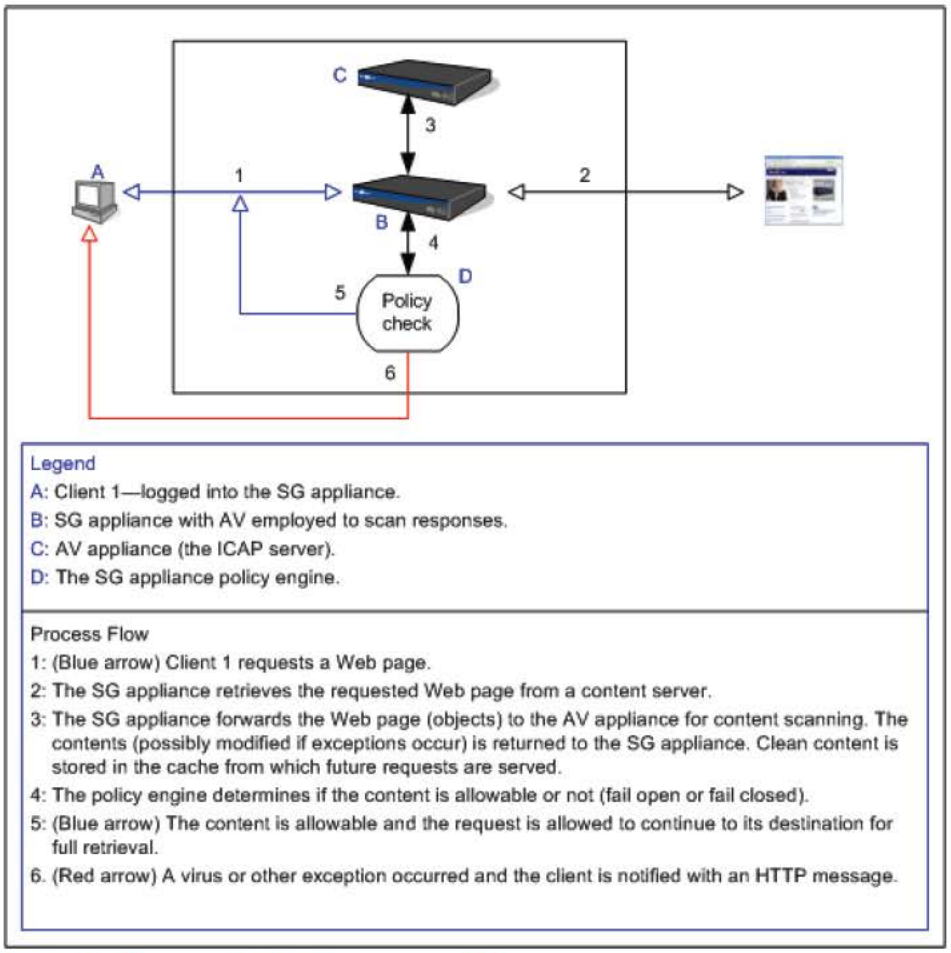


Figure 3-1. Response Modification Process Flow

About Request Modification

Request modification means the ICAP server scans contents that a client is attempting to send outside the network. This prevents perhaps unaware users from forwarding corrupted files or Webmail attachments. Request modification also is a method of content filtering and request transformation, which is used to protect network identity elements. Based on the results of the scan, the server might then return an HTTP response to the client (for example, sports not allowed); or the client request might be modified, such as stripping a referrer header, before continuing to the origin content server.

Note: Some ICAP servers do not support virus scanning for request modification, only content filtering.

The following diagram illustrates the request modification process flow.

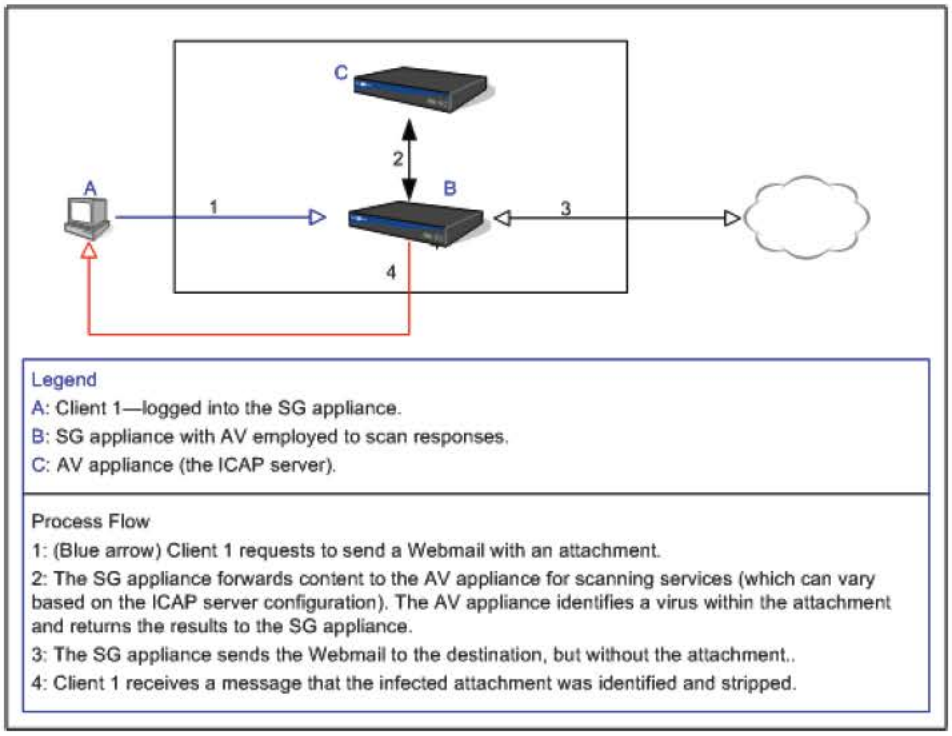


Figure 3-2. Request Modification Process Flow

Returning the Object to the Blue Coat Appliance

This object can be the original unchanged object, a repaired version of the original object minus a virus, or an error message indicating that the object contained a virus. Each of these responses is configured on the ICAP server, independent of the appliance and the ICAP protocol. If the appliance receives the error message, it forwards the error message to the client and does not save the infected file.

Caching and Serving the Object

After an object has been scanned and is determined cacheable, the SG appliance saves it and serves it for the subsequent content requests. When the appliance detects that the cached content has changed on the origin server, it fetches a fresh version, then forwards it to the ICAP server for scanning. If the SG appliance uses policies in the ICAP configuration, the policy applies to content fetches, distributes, and refreshes, as well as pipelining fetches.

For more information on policies, see "Section C: Creating ICAP Policy" on page 64. For more information on the `<Cache>` layer, refer to *Volume 11: Blue Coat SG Appliance Content Policy Language Guide*.

ICAP v1.0 Features

This section describes features of the ICAP v1.0 protocol.

Sense Settings

The Sense Settings feature allows the SG appliance to query any identified ICAP server running v1.0, detect the parameters, and configure the ICAP service as appropriate. See "Creating an ICAP Service" on page 55.

ISTags

An ICAP v1.0 server is required to return in each response an ICAP header ISTag indicating the current state of the ICAP server. This eliminates the need to designate artificial pattern version numbers, as is required in v0.95.

Note: Backing out a virus pattern on the ICAP server can revert ISTags to previous values that are ignored by the SG appliance. To force the SG appliance to recognize the old value, use the Sense Settings option, described in the configuration section.

Persistent Connections

New ICAP connections are created dynamically as ICAP requests are received (up to the defined maximum connection limit). The connection remains open to receive further requests. If a connection error occurs, the connection closes to prevent further errors.

Section B: Configuring SG Appliance ICAP Communications

This section describes how to configure the SG appliance to communicate with an ICAP server to perform content scanning tasks.

Configuration Tasks

Configuring ICAP on the SG appliance involves the following steps:

- Install the ICAP server.
- Configure the SG appliance to use ICAP and configure basic features.
- Create *patience pages*.
- Define scanning policies, then load the policy file on the SG appliance.

Installing the ICAP Server

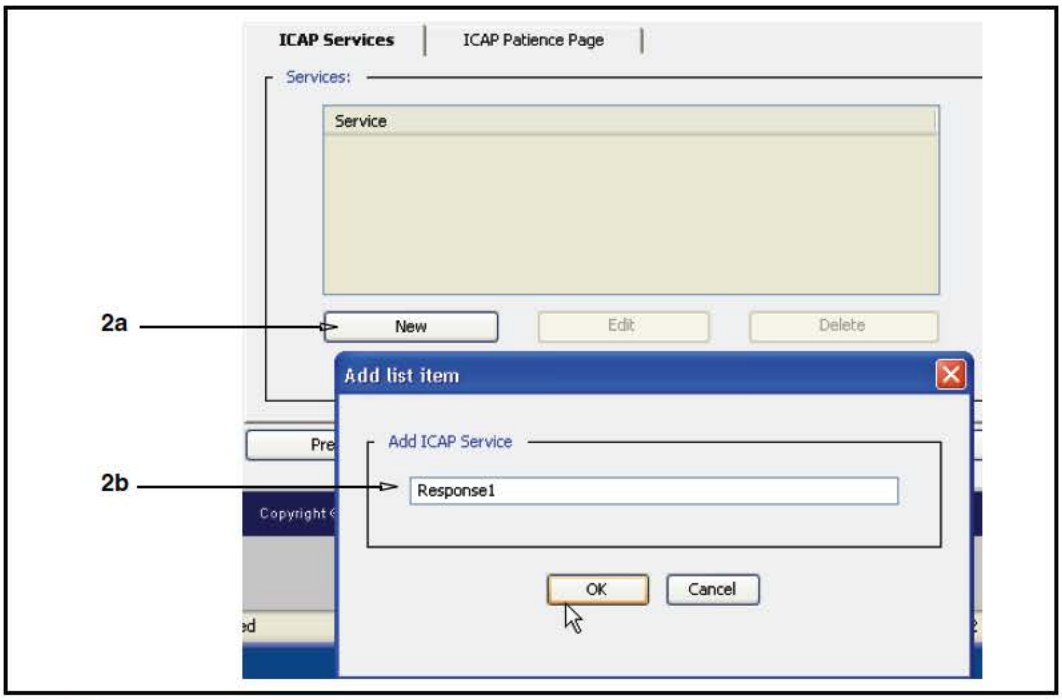
Follow the manufacturer instructions for installing the ICAP server, including any configuration necessary to work with the SG appliance. Based on your network environment, you might use the SG appliance with multiple ICAP servers or multiple scanning services on the same server. Configure options as needed, including the exception message displayed to end users in the event the requested object was modified or blocked.

Creating an ICAP Service

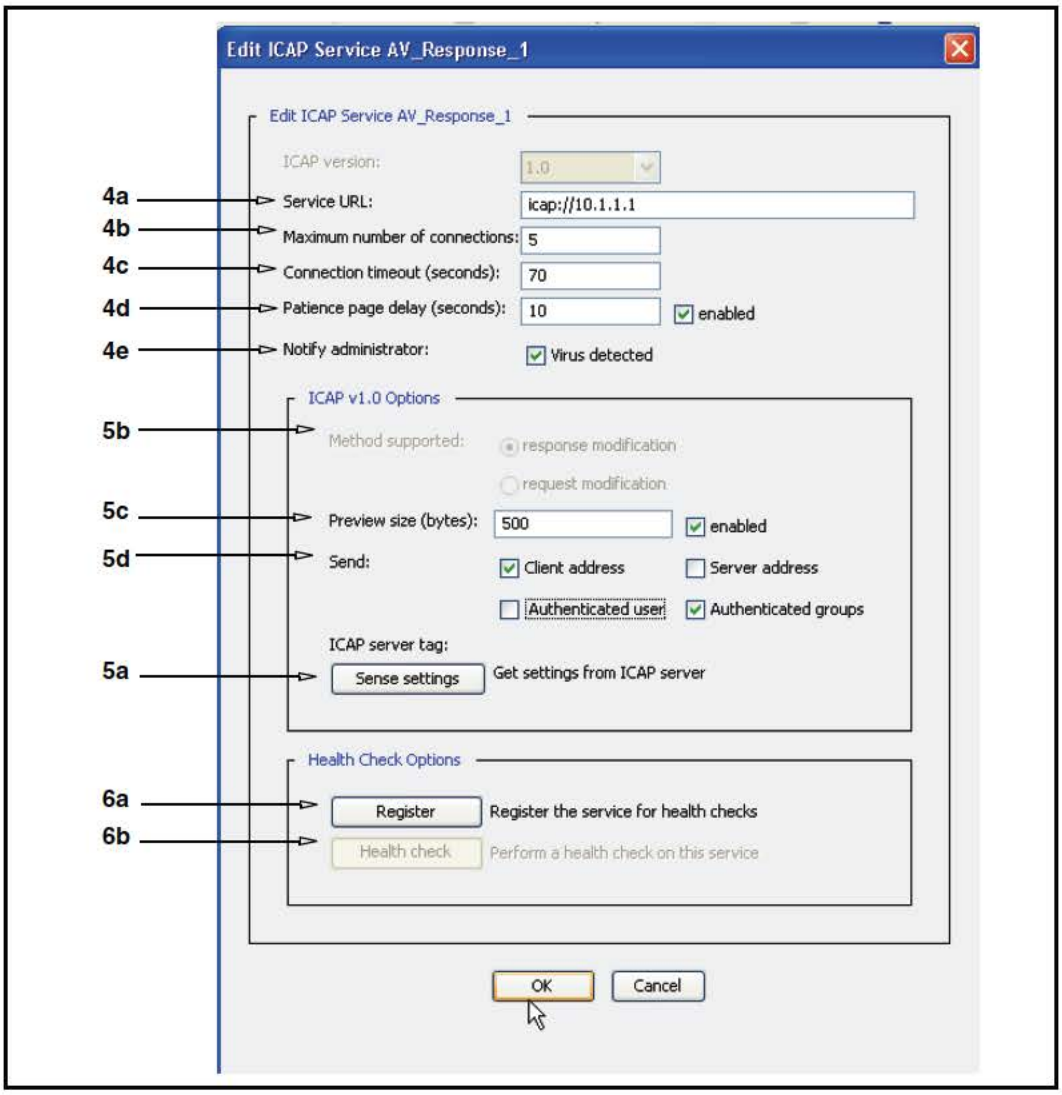
An ICAP service on the SG appliance is specific to the ICAP server and includes the server IP address or hostname, as well as the supported number of connections. If you are using the SG appliance with multiple ICAP servers or multiple scanning services on the same server, add an ICAP service for each server or scanning service.

To create and configure an ICAP service:

1. Select **Configuration > External Services > ICAP Services**.



2. Add a new service:
   a. Click **New**; the Add List Item dialog appears.
   b. In the **ICAP service name** field, enter an alphanumeric name. This example uses **Response1**.
   c. Click **OK** to close the dialog.
3. Highlight the new ICAP service name and click **Edit**; the Edit ICAP Service dialog appears.



4. Configure the service communication options:

Note: The default ICAP version is 1.0 and cannot be changed.

a. The service URL, which includes the URL schema, ICAP server hostname or IP address, and the ICAP port number. For example:

```
icap://10.x.x.x/
```

The default port number is `1344`, which can be changed; for example: `icap://10.x.x.x:99`. You can also give an HTTP URL, but you must define a port number.

> **Note:** An ICAP service pointing to a WebWasher server must use `icap` as the protocol in the URL. Blue Coat also recommends that you review your specific ICAP server documentation, as each vendor might require additional URL information.

b. The maximum number of connections possible at any given time between the SG appliance and the ICAP server. The range is a number from 1 to 65535. The default is 5. The number of recommended connections is dependent on the capabilities of the ICAP server. Refer to the vendor's product information.

c. The number of seconds the SG appliance waits for replies from the ICAP server. The range is 1 to 65536. The default timeout is 70 seconds.

d. Optional: You can enable the SG appliance to display a default patience page when an ICAP server is processing a relatively large object. The page informs users that a content scan is in process. If enabled, the patience page is displayed after the designated time value is reached for scanned objects.

> **Note:** Patience pages display regardless of any pop-up blocking policy that is in effect. Customizing patience pages is described in "Customizing ICAP Patience Text" on page 59.

To enable the patience page, in the **Patience page delay** field, enter the number of seconds the SG appliance waits before displaying the page. The range is 5 to 65535. Select **Enable**.

e. Select **Notify administrator: Virus detected** to send an e-mail to the administrator if the ICAP scan detects a virus. The notification is also sent to the Event Log and the Event Log e-mail list.

5. The following steps configure ICAP v1.0 features:

a. (Optional) Clicking **Sense Settings** automatically configures the ICAP service using the ICAP server parameters.

b. Select the ICAP method: response modification or request modification.

> **Note:** An ICAP server might have separate URLs for response modification and request modification services.

c. Enter the preview size (in bytes) and select **preview size enable**. The ICAP server reads the object up to the specified byte total. The ICAP server either continues with the transaction (that is, receives the remainder of the object for scanning) or opts out of the transaction.

The default is **0**. Only response headers are sent to the ICAP server; more object data is only sent if requested by the ICAP server.

d. (Optional) The **Send** options specify additional information that is forwarded to the ICAP server: **Send: Client address**, **Server address**, **Authenticated user**, or **Authenticated groups**.

e. Click **OK** to close the dialog.

6. For convenience, the Edit ICAP Service dialog allows you to register a newly-created ICAP service for health checking (this duplicates the functionality on the **Configuration > Health Checks > General** tab). Registering for health checking requires that a valid ICAP server URL was entered.
   a. Click **Register**; a dialog prompts confirmation; click **OK**.
   b. You can also click **Health check** to perform an immediate health check on this service.
7. Click **Apply**.

Monitoring ICAP Health Checks

In a browser, enter one of the following URLs to list health check information.

- To list all health check entries and their configuration parameters, enter:
  `http://SG_appliance_IP_address:8081/health_check/view`
- To list the statistics for all currently active entries, enter:
  `http://SG_appliance_IP_address:8081/health_check/statistics`

For more information about health checks, refer to *Volume 6: Advanced Networking*.

Deleting an ICAP Service

The following steps describe how to delete an ICAP service.

Note: You cannot delete an ICAP service used in a SG appliance policy (that is, if a policy rule uses the ICAP service name) or that belongs to a service group.

To delete an ICAP service:

1. Select **Configuration>External Services>ICAP**.
2. Select the service to be deleted.
3. Click **Delete**; click **OK** to confirm.
4. Click **Apply**.

Customizing ICAP Patience Text

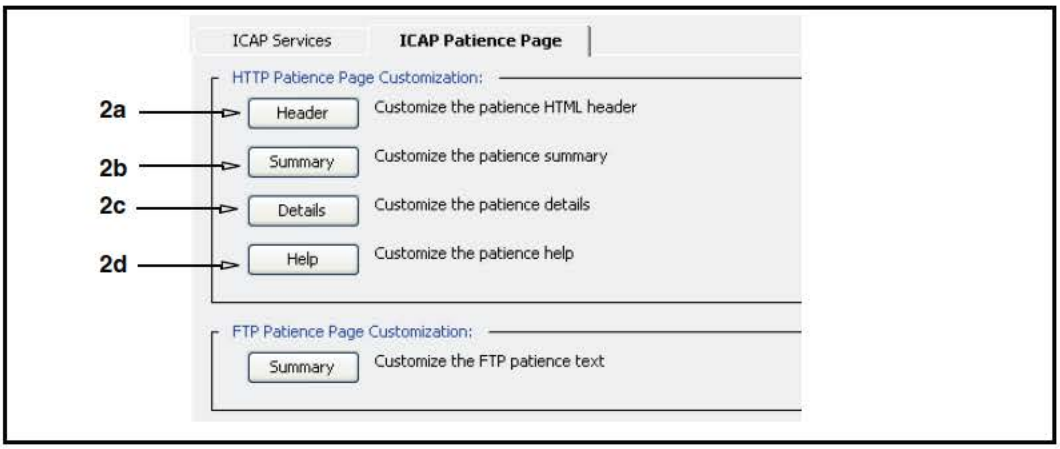
This section describes how to customize text displayed during ICAP scanning.

HTTP Patience Text

The SG appliance allows you to customize the patience page components and text that are displayed to users when HTTP clients experience delays as Web content is scanned.

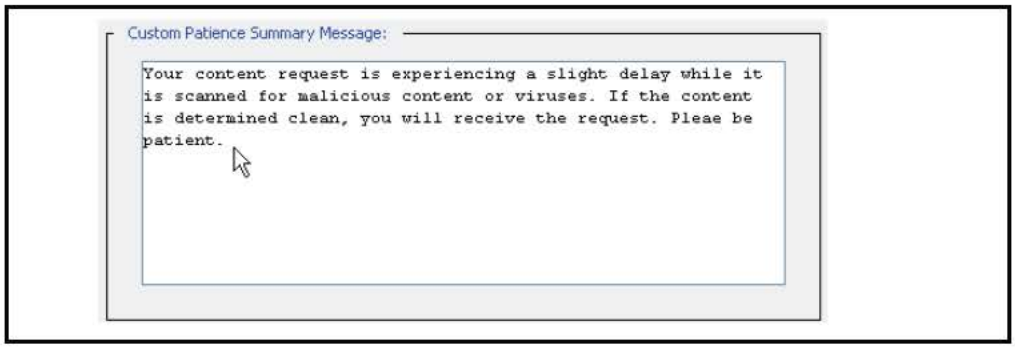
To customize HTTP patience pages:

1. Select **Configuration > External Services > ICAP > ICAP Patience Page**.



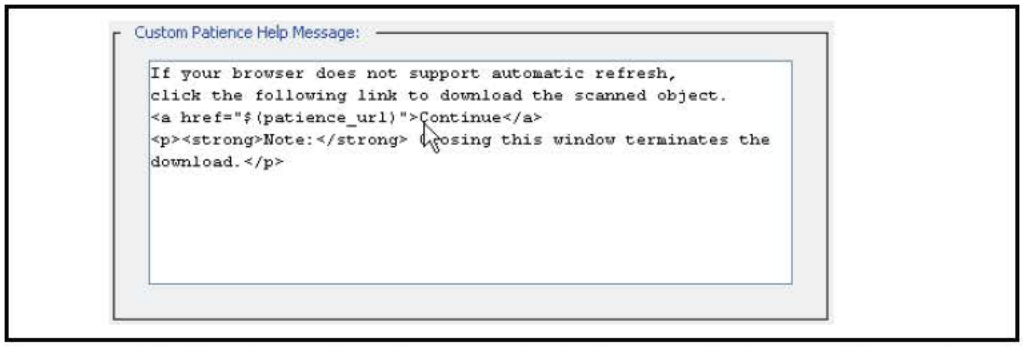
2. In the **HTTP Patience Page Customization** field, click **Header, Summary, Details**, or **Help**; the appropriate customize dialog appears. Customize the information as appropriate.

a. **Header**—Contains HTML tags that define what appears in the dialog title bar. This component also contains the `<meta http-equiv>` tag, which is used to specify a non-English character set.



b. **Summary**—HTML and text that informs users that a content scan is occurring.

c. **Details**—Uses data to indicate scanning progress. The information includes the URL currently being scanned, the number of bytes processed, and the elapsed time of the scan.



d. **Help**—Displays instructions for users should they experience a problem with the patience page.

3. Click **Apply**.

All of these components are displayed on the patience page.

Windows XP, Service Pack 2 Behavior

With Windows XP, Microsoft is continually updating the security measures, which impacts how the SG appliance manages patience pages.

- Browsers running on Windows XP, Service Pack 2 (XP SP2), experience slightly different patience page behavior when pop-up blocking is enabled.
  - If pop-up blocking is not enabled, patience page behavior should be normal.
  - If pop-up blocking is enabled (the default), the SG appliance attempts to display the patience page in the root window.
  - If the download triggers an invisible Javascript window, the user can track the scanning progress with the progress bar at the bottom of the window; however, if other policy blocks Javascript active content, this bar is also not visible.
- If Internet Explorer blocks all downloads initiated by Javascript, the user must click the yellow alert bar to download the scanned object.
- Users experience two patience page responses for non-cacheable objects.

Interactivity Notes

- When ICAP scanning is enabled and a patience page is triggered, a unique URL is dynamically generated and sent to the browser to access the patience page. This unique URL might contain a modified version of the original URL. This is expected behavior.
- Patience pages and exceptions can only be triggered by left-clicking a link. If a user right-clicks a link and attempt to save it, it is not possible to display patience pages. If this action causes a problem, the user might see browser-specific errors (for example, an Internet *site not found*); however, ICAP policy is still in effect.
- A patience page is not displayed if a client object request results in an HTTP 302 response and the SG appliance pipelines the object in the `Location` header. After the SG appliance receives the client request for the object, the client enters a waiting state because a server-side retrieval of the object is already in progress. The wait status of the client request prevents the patience page from displaying. To prevent the SG appliance from pipelining these requests (which decreases performance) and retain the ability to provide a patience page, configure HTTP:

  ```
  #SGOS (config) http no pipeline client redirects
  ```

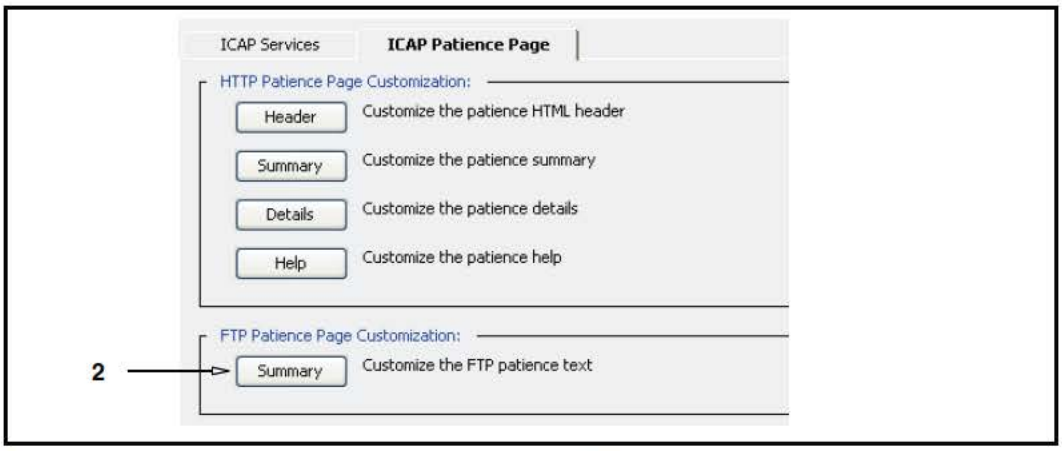
- The status bar update does not work if it is disabled or if the Javascript does not have sufficient rights to update it.
- Looping: Certain conditions cause browsers to re-spawn patience pages. For example, a site states it will begin a download in 10 seconds, initiates a pop-up download window, and returns to the root window. If the download window allows pop-ups, the patience page is displayed in another window. The automatic return to the root window initiates the download sequence again, spawning another patience page. If unnoticed, this loop could cause a system hang. The same behavior occurs if the user clicks the back button to return to the root window. For known and used download sites, you can create policy that redirects the page so that it doesn't return to the root window after a download starts.

FTP Patience Text

For content over FTP, the patience text displayed to FTP clients during an ICAP scan can be modified.

To customize FTP patience text:

1. Select **Configuration>External Services > ICAP > ICAP Patience Page**.



2. In the **FTP Patience Page Customization** field, click **Summary**; the Customize FTP Patience Text dialog appears. Customize the FTP client patience text as appropriate.
3. Click **OK**.
4. Click **Apply**.

Related CLI Syntax to Manage ICAP Communications

- To enter configuration mode:

  ```
  SGOS# (config) external-services
  ```

- The following subcommands are available:

  ```
  SGOS# (config external-services) create icap service_name
  SGOS# (config external-services) edit service_name
  SGOS# (config icap service_name) url icap://url
  SGOS# (config icap service_name) max-conn number
  SGOS# (config icap service_name) timeout timeout_seconds
  SGOS# (config icap service_name) notify virus-detected
  SGOS# (config icap service_name) methods {REQMOD | RESPMOD}
  SSGOS# (config icap service_name) preview-size bytes
  SGOS# (config icap service_name) send {client-address | server-address}
  SGOS# (config icap service_name) send {authenticated-user | authenticated-groups}
  SGOS# (config icap services service_name) sense-settings
  SGOS# (config icap services service_name) patience-page seconds
  SGOS# (config external-service) delete service_name
  SGOS# (config external-services) inline http icap-patience {details | header | help | javascript | summary} eof
  SGOS# (config external-services) inline ftp icap-patience-text eof
  ```

Section C: Creating ICAP Policy

Defined ICAP policy dictates the anti-virus behavior for your enterprise. You can either use the Visual Policy Manager (VPM) or you can manually edit policy files. For more information on the VPM and defining policies, refer to *Volume 7: VPM and Advanced Policy*.

Use the `request.icap_service()` (request modification) or `response.icap_service()` (response modification) properties to manage the SG appliance ICAP services.

VPM Objects

The VPM contains the following objects specific to AV scanning (linked to their descriptions in the VPM chapter).

Table 3-2. AV Scanning Objects

Object	Layer>Column
Virus Detected	Web Access>Service
ICAP Error Code	Web Access>Service
Return ICAP Patience Page	Web Access>Action
Set ICAP Request Service	Web Access>Action
Set ICAP Request Service	Web Content>Action
Set ICAP Response Service	Web Content>Action

Note: For CPL policy, refer to *Volume 11: Blue Coat SG Appliance Content Policy Language Guide*.

Example ICAP Policy

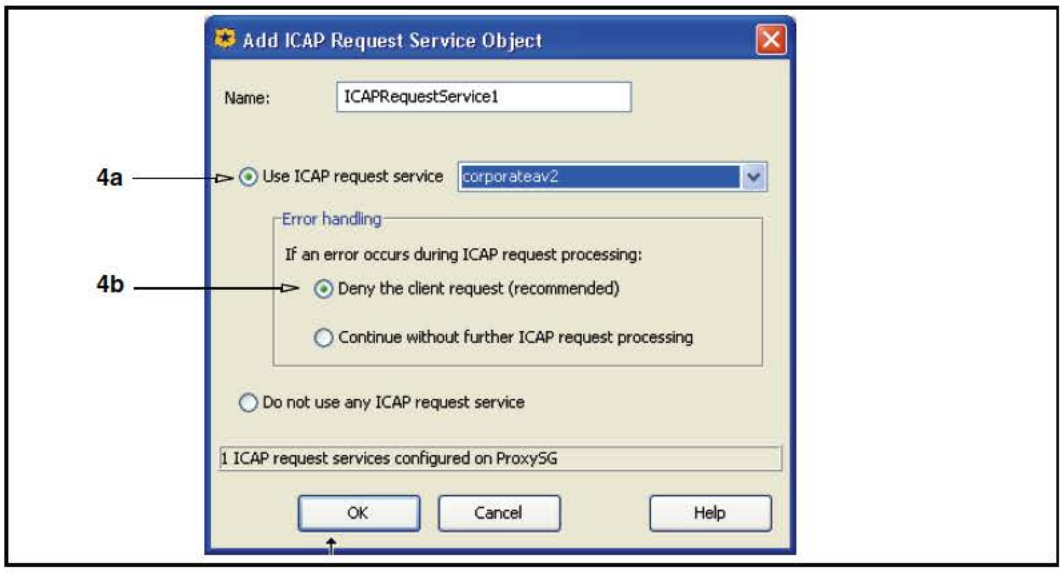
The following VPM example demonstrates the implementation of an ICAP policy that performs virus scanning on both client uploads (to prevent propagating a virus) and responses (to prevent the introduction of viruses).

For this example:

- The SG appliance has configured ICAP services. The response service is **corporateav1** and the request service is **corporateav2**.
- The Blue Coat AV is the virus scanner and is configured to serve password-protected files.
- A group named IT is configured on the SG appliance.
- The IT group wants to be allowed to download password protected files, but deny everyone else.

To perform virus scanning, protecting both the server side and the client side:

1. In the VPM, select **Policy > Web Access Layer**. Name the layer **RequestAV**.
2. Right-click the **Action** column; select **Set**. The Set Action Object dialog appears.
3. Select **Set ICAP Request Service**; the Add ICAP Request Service Object dialog appears.



4. Configure the request service object:
   a. From the **Use ICAP request service** drop-down list, select **corporateav2**.
   b. Accept the default: **Deny the client request**. This prevents a client from propagating a threat. If a virus is found, the content is not uploaded. For example, a user attempts to post a document that has a virus.
   c. Click **OK**; click **OK** again to add the object to the rule.

Figure 3-3. Request

5. In the VPM, select **Policy > Web Content Rule**. Name the rule **ResponseAV**.
6. Right-click the **Action** column; select **Set**. The Set Action Object dialog appears.
   a. Select **Set ICAP Response Service**; the Add ICAP Response Service Object dialog appears.
   b. From the **Use ICAP response service** drop-down list, select **corporateav1**.
7. Select **Deny the client request**. This scans the responses for viruses before the object is delivered to the client. If a virus is found, the content is not served.

To log a detected virus:

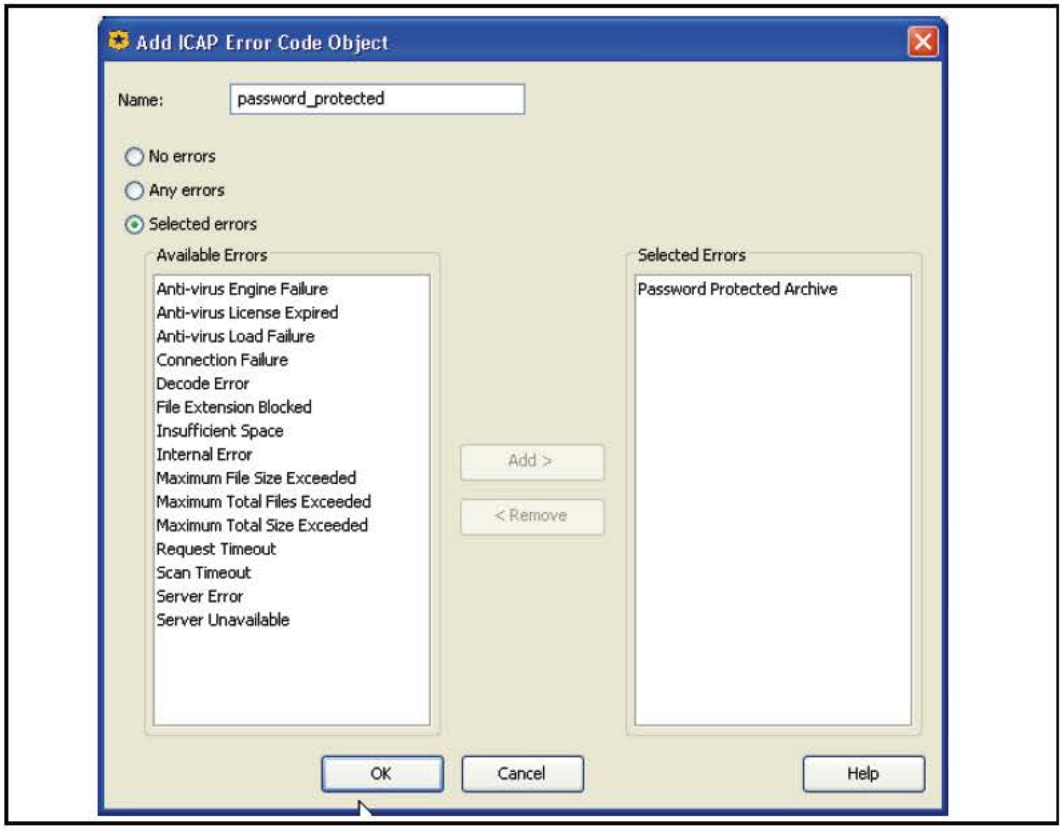
1. In the VPM, select **Policy > Web Access Layer**. Name the layer **AVErrors**.
2. Right-click the **Service** column; select **Set**. The Set Service Object dialog appears.
   a. Select **Virus Detected** (static object).
   b. Click **OK** to add the object to the rule.
3. Right-click the **Action** column. Select **Delete**.

4. Right-click the **Track** column. Select **Set**; the Set Track Object dialog appears.
   a. Click **New**; select **Event Log**. The Event Log dialog appears.
   b. In the **Name** field, enter **VirusLog1**.
   c. From the scroll-list, select `icap_virus_details`, `localtime`, and `client-address`. Click **Insert**.
   d. Click **OK**; click **OK** again to add the object to the rule.

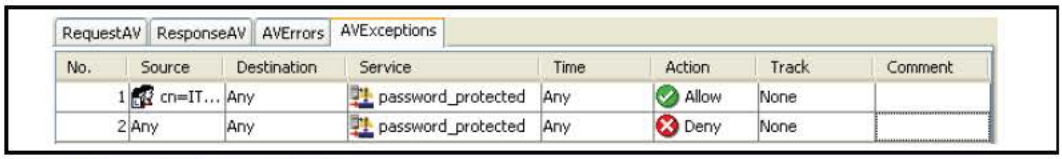
Figure 3-4. The AVErrors rule

To create an exception for IT group:

1. In VPM, select **Policy > Add Web Access Layer**. Name the rule **AVExceptions**.
2. Add the **IT** group object to the **Source** column.
3. Right-click the **Service** column; select **Set**. The Set Service Object dialog appears.
4. Click **New**; select **ICAP Error Code**. The **Add ICAP Error Code** Object appears.



5. Add the error code:
   a. Select **Selected Errors**
   b. From the list of errors, select **Password Protected Archive**; click **Add**.
   c. Name the object **password_protected**.
   d. Click **OK**; click **OK** again to add the object to the rule.
6. Right-click the **Action** column and select **Allow**.
7. Click Add **Rule**.
8. In the **Service** column, add the **password_protected** object.
9. Right-click the **Action** column; select **Deny**.



After this policy is installed:

- Virus scanning is performed for client attempts to upload content and content responses to client requests.
- If a virus is detected and there were no scanning process errors, a log entry occurs.
- As the Blue Coat AV is configured to serve password-protected objects, only the IT group can download such files; everyone else is denied.

Exempting HTTP Live Streams From Response Modification

The following CPL examples demonstrate how to exempt HTTP live streams from response modification, as they are not supported by ICAP. The CPL designates user agents that are bypassed.

```
<proxy>
url.scheme=http request.header.User-Agent="RealPlayer G2"
  response.icap_service(no)
url.scheme=http request.header.User-Agent="(RMA)"
response.icap_service(no)
url.scheme=http request.header.User-Agent="(Winamp)"
  response.icap_service(no)
url.scheme=http request.header.User-Agent="(NSPlayer)"
  response.icap_service(no)
url.scheme=http request.header.User-Agent="(Windows-Media-Player)"
  response.icap_service(no)
url.scheme=http request.header.User-Agent="QuickTime"
  response.icap_service(no)
url.scheme=http request.header.User-Agent="(RealMedia Player)"
  response.icap_service(no)
```

Streaming Media Request Modification Note

Some HTTP progressive download streaming media transactions are complex enough to disrupt ICAP request modification services. If such behavior is noticed (most common with RealPlayer), implement the following workaround policy to bypass the ICAP request modification service for HTTP progressive downloads:

```
<proxy>
url.scheme=http request_header.User-Agent="user_agent"
  request.icap_service(no)
url.scheme=http request_header.User-Agent="user_agent"
  request.icap_service(no)
```

where *user_agent* specifies a media player attribute that is disrupting service. For example:

```
<proxy>
url.scheme=http request_header.User-Agent="(RealMedia Player)"
  request.icap_service(no)
url.scheme=http request_header.User-Agent="RMA"
request.icap_service(no)
```

CPL Notes

- If policy specifies that an ICAP service is to be used, but the service is not available, the default behavior is to fail closed—that is, deny the request or response. The following CPL allows the serving of objects without ICAP processing if the server is down.

  ```
  request.icap_service(service_name, fail_open)
  response.icap_service(service_name, fail_open)
  ```

  When the ICAP service is restored, these objects are scanned and served from the cache if they are requested again.

  Note: Blue Coat recommends this CPL to be used for internal sites; use with caution.

- To provide an exception to a general rule, the following CPL negates ICAP processing:

  ```
  request.icap_service(no)
  response.icap_service(no)
  ```

Section D: Managing Virus Scanning

You might need to perform additional SG appliance maintenance concerning virus scanning, particularly for updates to the virus definition on the ICAP virus scanning server.

Advanced Configurations

This section summarizes more-advanced configurations between the SG appliance and multiple ICAP servers. These brief examples provide objectives and suggest ways of supporting the configuration.

Using Object-Specific Scan Levels

You can specify different scanning levels for different types of objects, or for objects from different sources.

This requires a service group of ICAP servers, with each server configured to provide the same level of scanning. For more information, refer to Chapter 4: "Configuring Service Groups" on page 71.

Improving Virus Scanning Performance

You can overcome request-handling limitations of ICAP servers. Generally, SG appliances can handle many times the volume of simultaneous user requests that ICAP servers can handle.

This requires multiple ICAP servers to obtain a reasonable performance gain. On the SG appliance, define policy rules that partition requests among the servers. If you are going to direct requests to individual servers based on rules, configure in rule conditions that only use the URL. Note that you can increase the scale by using a service group, rather than use rules to partition requests among servers. For more information on using multiple ICAP servers, refer to Chapter 4: "Configuring Service Groups" on page 71. For more information about defining policies, refer to the Managing Policy Files chapter in *Volume 7: VPM and Advanced Policy*, as well as *Volume 12: Blue Coat SG Appliance Command Line Reference.*

When the virus definitions are updated, the SG appliance stores a signature. This signature consists of the server name plus a virus definition version. If either of these changes, the SG appliance checks to see if the object is up to date, and then rescans it. If two requests for the same object are directed to different servers, then the scanning signature changes and the object is rescanned.

Updating the ICAP Server

If there is a problem with the integration between the SG appliance and a supported ICAP server after a version update of the server, you might need to configure the preview size the appliance uses. For information, see “Creating an ICAP Service” on page 55.

Replacing the ICAP Server

If you replace an ICAP server with another supported ICAP server, reconfigure the ICAP service on the SG appliance:

```
SGOS# (config) external-services
SGOS# (config external-service) edit service_name
SGOS# (config service_name) url url
```

For information about these commands, see "Creating an ICAP Service" on page 55.

Access Logging

The SG appliance provides access log support for Symantec and Finjan ICAP 1.0 server actions (**Management > Access Logging**). The following sections describe access logging behavior for the various supported ICAP servers.

Symantec AntiVirus Scan Engine 4.0

When this Symantec server performs a scan, identifies a problem (for example, a virus), and performs a content transformation, the action is logged. For example:

```
"virus-id: Type=number; Resolution=[0 | 1 | 2]; Threat=name;"
```

where:

`Type=number`	Specifies the numeric code for the virus.
`Resolution=`	Specifies an integer value that indicates what action was taken to fix the file. Zero (`0`) defines the file is unrepairable, one (`1`) specifies that the file was repaired, and two (`2`) specifies that the file was deleted.
`Threat=`	Specifies the name of the virus.

Finjan SurfinGate 7.0

When this Finjan ICAP server performs a scan, identifies a problem (for example, a virus), and performs a content transformation, the action is logged. For example:

```
"virus-id: name, response-info: Blocked, response-desc: virus_name was
detected"
```

Finjan ICAP servers also log occurrences malicious mobile code.

Note: The access log string cannot exceed 256 characters. If the header name or value extends the length over the limit, then that string does not get logged. For example, if the `x-virus-id` header value is 260 characters, the access log displays `"x-virus-id: "` with no value because the value is too long to display. Also, if the access log string is already 250 characters and the SG appliance attempts to append a `"Malicious-Mobile-Type: "` string, the string is not appended

---

Access log entries might vary depending upon the type of ICAP scan performed and the custom log formats. For information about Access Logging, refer to *Volume 9: Access Logging*.

Chapter 4: Configuring Service Groups

This chapter describes how to create and manage ICAP or Websense service groups. In high-traffic network environments, a service group accelerates response time by a performing a higher volume of scanning.

About Weighted Load Balancing

The SG appliance supports weighted load balancing in forwarding requests to service groups. By default, the SG appliance performs typical round-robin load balancing and evenly forwards requests sequentially to servers as defined within the service group. Manually assigning weights takes advantage of round-robin load balancing in service groups that are not homogeneous, or where the servers have different capacities.

Weighting determines what proportion of the load one server bears relative to the others. If all servers have either the default weight (1) or the same weight, each share an equal proportion of the load. If one server has weight 25 and all other servers have weight 50, the 25-weight server processes half as much as any other server.

Before configuring weights, consider the relative weights to assign to each server. Factors that could affect assigned weight of a ICAP server include the following:

- The processing capacity of the server hardware in relationship to other servers (for example, the number and performance of CPUs or the number of network interface cards)
- The maximum number of connections configured for the service. The maximum connections setting pertains to how many simultaneous scans can be performed on the server, while weighting applies to throughput in the integration. While these settings are not directly related, consider both when configuring weighted load balancing.

Note: External services (ICAP, Websense off-box) have a reserved connection for health checks (if you created health check services). This means that as the load goes up and the number of connections to the external service reaches the maximum, with additional requests being queued up and waiting, the maximum simultaneous connections is actually one less than the limit.

The following diagram provides an example of how weighting works with a service group of three Blue Coat AV ICAP servers.

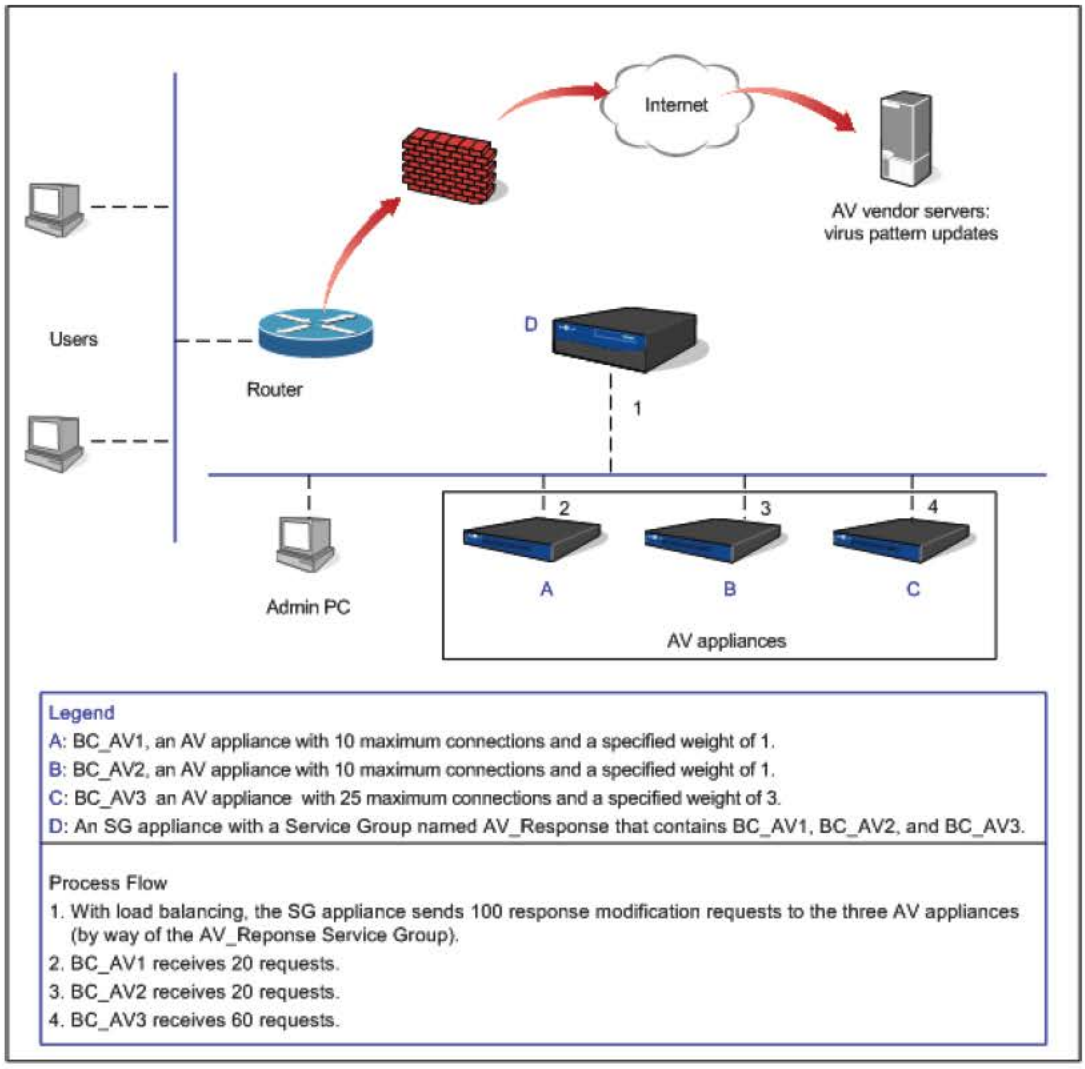


Figure 4-1. Service Group Process Flow

Note: Setting the weight value to **0** (zero) disables weighted load balancing for the ICAP service. Therefore, if one ICAP server of a two-server group has a weight value of **1** and the second a weight value of **0**, should the first server go down, a communication error results because the second server cannot process the request.

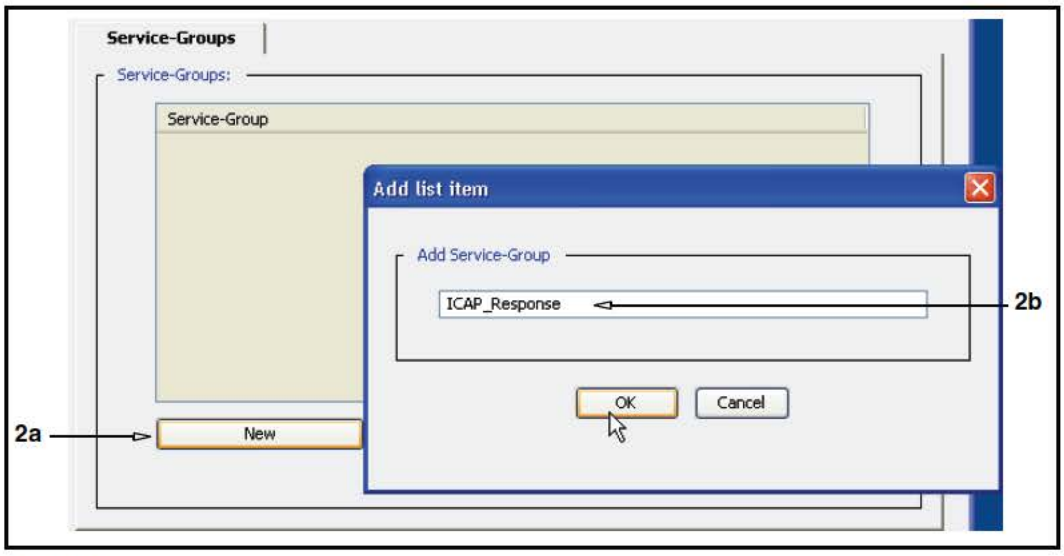
While you cannot specifically designate an ICAP server in a group as a backup, you can specify weight values that create a large differential between a server that is used continuously and one that is rarely used, thus simulating a backup server.

Creating a Service Group

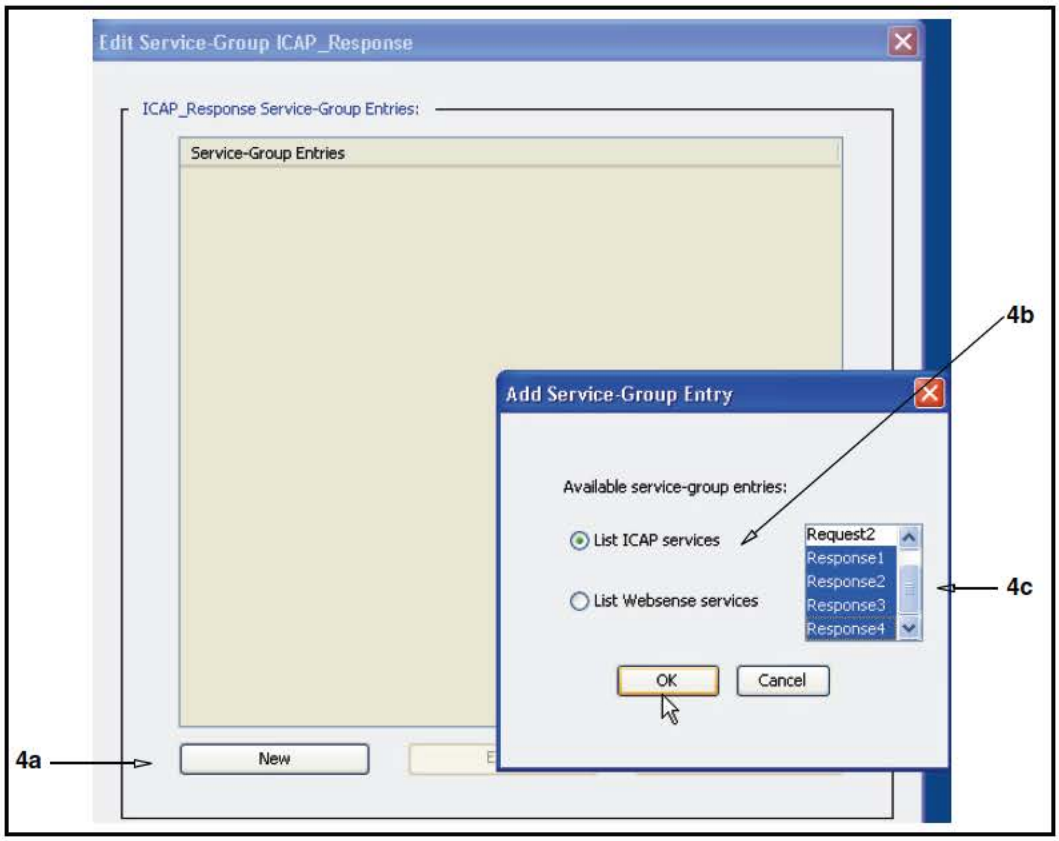
Create the service group and add the relevant ICAP or Websense services to the group. Services within group must be the same type (ICAP or Websense).

To configure a service group:

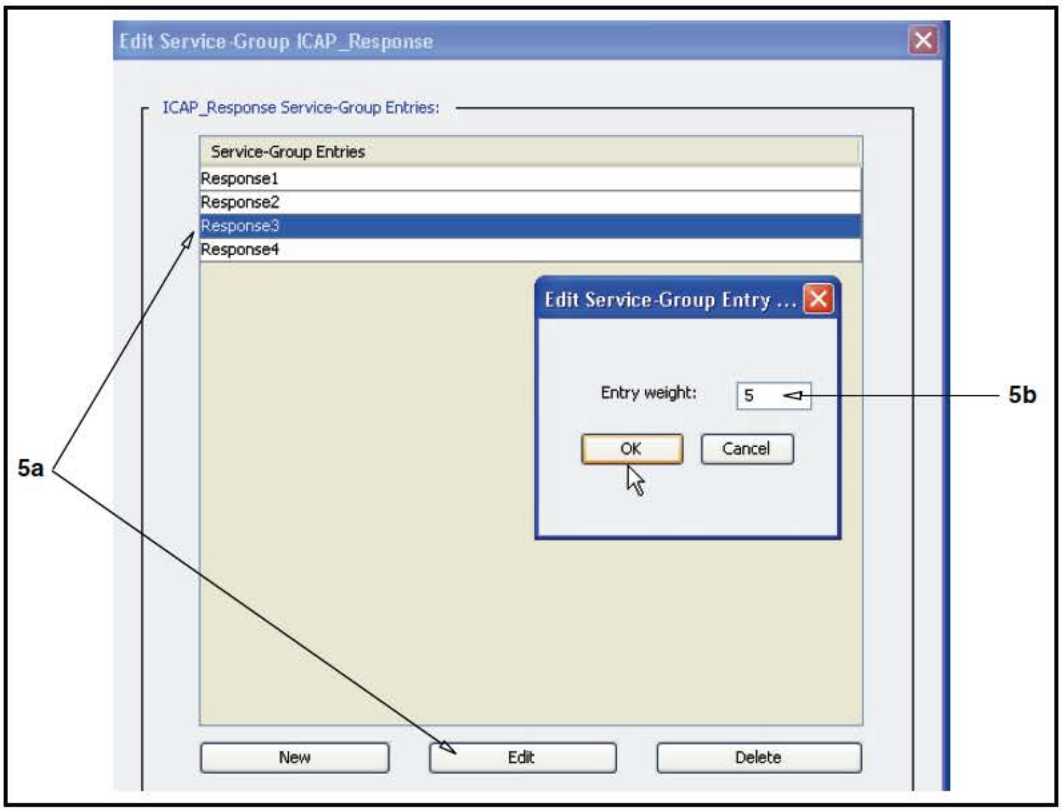
1. Select **Configuration > External Services > Service-Groups**.



2. Add a new group:
   a. Click **New**; the Add List Item dialog appears.
   b. In the **Add Service Group** field, enter an alphanumeric name. This example creates a group called **ICAP_Response**.
   c. Click **OK**.
3. Highlight the new service group name and click **Edit**; the Edit Service Group dialog appears.



4. Select existing services:
   a. Click **New**; the Add Service Group Entry dialog appears.
   b. If this SG appliance contains many configured ICAP or Websense (off-box) services, you can narrow the viewable list by selecting **List ICAP services** or **List Websense services**.
   c. From the list of existing services, select the ones to add to this group. Hold the Control or Shift key to select multiple services.
   d. Click **OK** to add the selected services to group.



5. Assign weights to services:
   a. Select a service and click **Edit**; the Edit Service Group Entry weight dialog appears.
   b. In the **Entry Weight** field, assign a weight value. The valid range is 0-255. For conceptual information about service weighting, see "About Weighted Load Balancing" on page 71.
   c. Repeat steps a and b for other services, as required.
   d. Click **OK** to close the dialog.
   e. Click **OK** again to close the Edit Service Group Entry dialog
6. Click **Apply**.

Result: When instructed by created policies, the SG appliance sends ICAP response modification requests to ICAP servers in the service group. The load carried by each service in the group is determined by the weight values.

Deleting a Service Group or Group Entry

You can delete the configuration for an entire service group from the SG appliance, or you can delete individual entries from a service group.

Note: A service or service group used in a SG appliance policy (that is, if a policy rule uses the entry) cannot be deleted; it must first be removed from the policy.

To delete a service group:

1. Select **Configuration > External Services > Service-Groups**.
2. Select the service group to be deleted.
3. Click **Delete**; click **OK** to confirm.
4. Click **Apply**.

To delete a service group entry:

1. Select **Configuration > External Services > Service-Groups**.
2. Select the service group to be modified.
3. Click **Edit**.
4. Select the service entry to be deleted; click **Delete**.
5. Click **OK**.
6. Click **Apply**.

Displaying External Service and Group Information

After configuring a service group, you can display aggregate service group (and other External Services) information.

To display information about all external services and groups:

At the (`config`) command prompt, enter the following commands:

```
SGOS# (config) external-services
SGOS# (config external-services) view
```

Individual service information is displayed first, followed by service group information. For example:

```
; External Services
icap4
ICAP-Version:        1.0
  URL:                 icap://10.1.1.1
  Max-conn:            5
  Timeout(secs):       70
  Health-checks:       no
  Patience-page(secs): disabled
  Notification:        never
  Methods:             RESPMOD
  Preview-size:        0
  Send:                nothing
  ISTag:
websense4
  Version:             4.4
  Host:                www.websense.com/list
  Port:                15868
  Max-conn:            5
  Timeout(secs):       70
  Send:                nothing
  Fail-by-default:     closed
  Apply-by-default:    no
  Serve-exception-page:yes
```

```
; External Service-Groups
CorpICAP
  total weight 5
entries:
  ICAP1
   weight 4
  ICAP2
   weight 1
BranchWebsense
  total weight  2
entries:
  Websense1
   weight    1
  Websense2
   weight    1
```

Related CLI Syntax to Manage External Services

- To enter configuration mode:

  `SGOS# (config) external-services`

- The following commands are available:

  `SGOS# (config external-services) create service-group name`

  `SGOS# (config service-group name) add service_name`

  `SGOS# (config service-group name) edit service_name`

  `SGOS# (config service-group name) weight value`

  `SGOS# (config external-services) delete service_group_name`

  `SGOS# (config type name) remove entry_name`

  `SGOS# (config external-services) view`

  `SGOS# (config type name) view`

Appendix A: Glossary

Term	Description
ADN Optimize Attribute	Controls whether to optimize bandwidth usage when connecting upstream using an ADN tunnel.
Asynchronous Adaptive Refresh (AAR)	This allows the Proxy*SG* to keep cached objects as fresh as possible, thus reducing response times. The AAR algorithm allows HTTP proxy to manage cached objects based on their rate of change and popularity: an object that changes frequently and/or is requested frequently is more eligible for asynchronous refresh compared to an object with a lower rate of change and/or popularity.
Asynchronous Refresh Activity	Refresh activity that does not wait for a request to occur, but that occurs *asynchronously* from the request.
Attributes (Service)	The service attributes define the parameters, such as explicit or transparent, cipher suite, and certificate verification, that the SG appliance uses for a particular service. .
Authenticate-401 Attribute	All transparent and explicit requests received on the port always use transparent authentication (cookie or IP, depending on the configuration). This is especially useful to force transparent proxy authentication in some proxy-chaining scenarios
authentication	The process of identifying a specific user.
authorization	The permissions given to a specific user.
Bandwidth Gain	A measure of the difference in client-side and server-side Internet traffic expressed in relation to server-side Internet traffic. It is managed in two ways: you can enable or disable bandwidth gain mode or you can select the Bandwidth Gain profile (this also enables bandwidth gain mode)..
Bandwidth Class	A defined unit of bandwidth allocation. An administrator uses bandwidth classes to allocate bandwidth to a particular type of traffic flowing through the SG appliance.
Bandwidth Class Hierarchy	Bandwidth classes can be grouped together in a class hierarchy, which is a tree structure that specifies the relationship among different classes. You create a hierarchy by creating at least one parent class and assigning other classes to be its children.
Bandwidth Policy	The set of rules that you define in the policy layer to identify and classify the traffic in the SG appliance, using the bandwidth classes that you create. You must use policy (through either VPM or CPL) in order to manage bandwidth.
Bypass Lists	The bypass list allows you to exempt IP addresses from being proxied by the SG appliance. The bypass list allows either <All> or a specific IP prefix entry for both the client and server columns. Both UDP and TCP traffic is automatically exempted.

Term	Description
Byte-Range Support	The ability of the Proxy*SG* to respond to byte-range requests (requests with a `Range:` HTTP header).
Cache-hit	An object that is in the Proxy*SG* and can be retrieved when an end user requests the information.
Cache-miss	An object that can be stored but has never been requested before; it was not in the Proxy*SG* to start, so it must be brought in and stored there as a side effect of processing the end-user's request. If the object is cacheable, it is stored and served the next time it is requested.
Child Class (Bandwidth Gain)	The child of a parent class is dependent upon that parent class for available bandwidth (they share the bandwidth in proportion to their minimum/maximum bandwidth values and priority levels). A child class with siblings (classes with the same parent class) shares bandwidth with those siblings in the same manner.
Client consent certificates	A certificate that indicates acceptance or denial of consent to decrypt an end user's HTTPS request.
Compression	An algorithm that reduces a file's size but does not lose any data. The ability to compress or decompress objects in the cache is based on policies you create. Compression can have a huge performance benefit, and it can be customized based on the needs of your environment: Whether CPU is more expensive (the default assumption), server-side bandwidth is more expensive, or whether client-side bandwidth is more expensive.
Default Proxy Listener	See " Proxy Service (Default)" .
Detect Protocol Attribute	Detects the protocol being used. Protocols that can be detected include: HTTP, P2P (eDonkey, BitTorrent, FastTrack, Gnutella), SSL, and Endpoint Mapper.
Directives	Directives are commands that can be used in installable lists to configure forwarding. See also *forwarding Configuration*.
Display Filter	The display filter is a drop-down list at the top of the Proxy Services pane that allows you to view the created proxy services by service name or action.
Early Intercept Attribute	Controls whether the proxy responds to client TCP connection requests before connecting to the upstream server. When early intercept is disabled, the proxy delays responding to the client until after it has attempted to contact the server.
Emulated Certificates	Certificates that are presented to the user by ProxySG when intercepting HTTPS requests. Blue Coat emulates the certificate from the server and signs it, copying the subjectName and expiration. The original certificate is used between the Proxy*SG* and the server.
ELFF-compatible format	A log type defined by the W3C that is general enough to be used with any protocol.
Encrypted Log	A log is encrypted using an external certificate associated with a private key. Encrypted logs can only be decrypted by someone with access to the private key. The private key is not accessible to the SG appliance.

Term	Description
explicit proxy	A configuration in which the browser is explicitly configured to communicate with the proxy server for access to content. This is the default for the SG appliance, and requires configuration for both browser and the interface card.
Fail Open/Closed	Failing open or closed applies to forwarding hosts and groups and SOCKS gateways. Fail Open/Closed applies when the health checks are showing sick for each forwarding or SOCKS gateway target in the applicable fail-over sequence. If no systems are healthy, the SG appliance fails open or closed, depending on the configuration. If closed, the connection attempt simply fails. If open, an attempt is made to connect without using any forwarding target (or SOCKS gateway). Fail open is usually a security risk; fail closed is the default if no setting is specified.
Forwarding Configuration	Forwarding can be configured through the CLI or through adding directives to a text file and installing it as an installable list. Each of these methods (the CLI or using directives) is equal. You cannot use the Management Console to configure forwarding.
Forwarding Host	Upstream Web servers or proxies.
forward proxy	A proxy server deployed close to the clients and used to access many servers. A forward proxy can be explicit or transparent.
Freshness	A percentage that reflects the objects in the Proxy*SG* cache that are expected to be fresh; that is, the content of those objects is expected to be identical to that on the OCS (origin content server).
Gateway	A device that serves as entrance and exit into a communications network.
Global Default Settings	You can configure settings for all forwarding hosts and groups. These are called the global defaults. You can also configure private settings for each individual forwarding host or group. Individual settings override the global defaults.
FTP	See Native FTP; Web FTP.
Host Affinity	Host affinity is the attempt to direct multiple connections by a single user to the same group member. Host affinity is closely tied to load balancing behavior; both should configured if load balancing is important.
Host Affinity Timeout	The host affinity timeout determines how long a user remains idle before the connection is closed. The timeout value checks the user's IP address, SSL ID, or cookie in the host affinity table.
Inbound Traffic (Bandwidth Gain)	Network packets flowing into the SG appliance. Inbound traffic mainly consists of the following: • Server inbound: Packets originating at the origin content server (OCS) and sent to the SG appliance to load a Web object. • Client inbound: Packets originating at the client and sent to the SG appliance for Web requests.

Term	Description
Installable Lists	Installable lists, comprised of directives, can be placed onto the SG appliance in one of several methods: through creating the list through the SG text editor, by placing the list at an accessible URL, or by downloading the directives file from the local system.
Integrated Host Timeout	An integrated host is an Origin Content Server (OCS) that has been added to the health check list. The host, added through the `integrate_new_hosts` property, ages out of the integrated host table after being idle for the specified time. The default is 60 minutes.
IP Reflection	Determines how the client IP address is presented to the origin server for explicitly proxied requests. All proxy services contain a reflect-ip attribute, which enables or disables sending of client's IP address instead of the SG's IP address.
Issuer keyring	The keyring that is used by the SG appliance to sign emulated certificates. The keyring is configured on the appliance and managed through policy.
Listener	The service that is listening on a specific port. A listener can be identified by any destination IP/subnet and port range. Multiple listeners can be added to each service.
Load Balancing	The ability to share traffic requests among multiple upstream targets. Two methods can be used to balance the load among systems: `least-connections` or `round-robin`.
Log Facility	A separate log that contains a single logical file and supports a single log format. It also contains the file's configuration and upload schedule information as well as other configurable information such as how often to rotate (switch to a new log) the logs at the destination, any passwords needed, and the point at which the facility can be uploaded.
Log Format	The type of log that is used: NCSA/Common, SQUID, ELFF, SurfControl, or Websense. The proprietary log types each have a corresponding pre-defined log format that has been set up to produce exactly that type of log (these logs cannot be edited). In addition, a number of other ELFF type log formats are also pre-defined (im, main, p2p, ssl, streaming). These can be edited, but they start out with a useful set of log fields for logging particular protocols understood by the SG appliance. It is also possible to create new log formats of type ELFF or Custom which can contain any desired combination of log fields.
Log Tail:	The access log tail shows the log entries as they get logged. With high traffic on the SG appliance, not all access log entries are necessarily displayed. However, you can view all access log information after uploading the log.
Maximum Object Size	The maximum object size stored in the Proxy*SG*. All objects retrieved that are greater than the maximum size are delivered to the client but are not stored in the Proxy*SG*.
NCSA common log format	A log type that contains only basic HTTP access information.

Term	Description
Negative Responses	An error response received from the OCS when a page or image is requested. If the Proxy*SG* is configured to cache such negative responses, it returns that response in subsequent requests for that page or image for the specified number of minutes. If it is not configured, which is the default, the Proxy*SG* attempts to retrieve the page or image every time it is requested.
Native FTP	Native FTP involves the client connecting (either explicitly or transparently) using the FTP protocol; the SG appliance then connects upstream through FTP (if necessary).
Outbound Traffic (Bandwidth Gain)	Network packets flowing out of the SG appliance. Outbound traffic mainly consists of the following: • Client outbound: Packets sent to the client in response to a Web request. • Server outbound: Packets sent to an OCS or upstream proxy to request a service.
Origin Content Server (OCS)	
Parent Class (Bandwidth Gain)	A class with at least one child. The parent class must share its bandwidth with its child classes in proportion to the minimum/maximum bandwidth values or priority levels.
PASV	Passive Mode Data Connections. Data connections initiated by an FTP client to an FTP server.
proxy	Caches content, filters traffic, monitors Internet and intranet resource usage, blocks specific Internet and intranet resources for individuals or groups, and enhances the quality of Internet or intranet user experiences. A proxy can also serve as an intermediary between a Web client and a Web server and can require authentication to allow identity based policy and logging for the client. The rules used to authenticate a client are based on the policies you create on the SG appliance, which can reference an existing security infrastructure—LDAP, RADIUS, IWA, and the like.
Proxy Service	The proxy service defines the ports, as well as other attributes. that are used by the proxies associated with the service.
Proxy Service (Default)	The default proxy service is a service that intercepts all traffic not otherwise intercepted by other listeners. It only has one listener whose action can be set to bypass or intercept. No new listeners can be added to the default proxy service, and the default listener and service cannot be deleted. Service attributes can be changed.
realms	A realm is a named collection of information about users and groups. The name is referenced in policy to control authentication and authorization of users for access to Blue Coat Systems SG services. Multiple authentication realms can be used on a single SG appliance. Realm services include IWA, LDAP, Local, and RADIUS.
Reflect Client IP Attribute	Enables the sending of the client's IP address instead of the SG's IP address to the upstream server. If you are using an Application Delivery Network (ADN), this setting is enforced on the concentrator proxy through the Configuration>App. Delivery Network>Tunneling tab.

Term	Description
Refresh Bandwidth	The amount of bandwidth used to keep stored objects fresh. By default, the Proxy*SG* is set to manage refresh bandwidth automatically. You can configure refresh bandwidth yourself, although Blue Coat does not recommend this.
reverse proxy	A proxy that acts as a front-end to a small number of pre-defined servers, typically to improve performance. Many clients can use it to access the small number of predefined servers.
rotate logs	When you rotate a log, the old log is no longer appended to the existing log, and a new log is created. All the facility information (headers for passwords, access log type, and so forth), is re-sent at the beginning of the new upload. If you're using Reporter (or anything that doesn't understand the concept of "file," such as streaming) the upload connection is broken and then re-started, and, again, the headers are re-sent.
serial console	A device that allows you to connect to the SG appliance when it is otherwise unreachable, without using the network. It can be used to administer the SG appliance through the CLI. You must use the CLI to use a serial console. Anyone with access to the serial console can change the administrative access controls, so physical security of the serial console is critical.
Server Certificate Categories	The hostname in a server certificate can be categorized by BCWF or another content filtering vendor to fit into categories such as banking, finance, sports.
Sibling Class (Bandwidth Gain)	A bandwidth class with the same parent class as another class.
SOCKS Proxy	A generic way to proxy TCP and UDP protocols. The SG appliance supports both SOCKSv4/4a and SOCKSv5; however, because of increased username and password authentication capabilities and compression support, Blue Coat recommends that you use SOCKS v5..
SmartReporter log type	A proprietary ELFF log type that is compatible with the SmartFilter SmartReporter tool.
Split proxy	Employs co-operative processing at the branch and the core to implement functionality that is not possible in a standalone proxy. Examples of split proxies include : Mapi Proxy SSL Proxy
SQUID-compatible format	A log type that was designed for cache statistics.
SSL	A standard protocol for secure communication over the network. Blue Coat recommends using this protocol to protect sensitive information.
SSL Interception	Decrypting SSL connections.
SSL Proxy	A proxy that can be used for any SSL traffic (HTTPS or not), in either forward or reverse proxy mode.

Term	Description
static routes	A manually-configured route that specifies the transmission path a packet must follow, based on the packet's destination address. A static route specifies a transmission path to another network.
SurfControl log type	A proprietary log type that is compatible with the SurfControl reporter tool. The SurfControl log format includes fully-qualified usernames when an NTLM realm provides authentication. The simple name is used for all other realm types.
Traffic Flow (Bandwidth Gain)	Also referred to as *flow*. A set of packets belonging to the same TCP/UDP connection that terminate at, originate at, or flow through the SG appliance. A single request from a client involves two separate connections. One of them is from the client to the SG appliance, and the other is from the SG appliance to the OCS. Within each of these connections, traffic flows in two directions—in one direction, packets flow out of the SG appliance (outbound traffic), and in the other direction, packets flow into the SG (inbound traffic). Connections can come from the client or the server. Thus, traffic can be classified into one of four types: • Server inbound • Server outbound • Client inbound • Client outbound These four traffic flows represent each of the four combinations described above. Each flow represents a single direction from a single connection.
transparent proxy	A configuration in which traffic is redirected to the SG appliance without the knowledge of the client browser. No configuration is required on the browser, but network configuration, such as an L4 switch or a WCCP-compliant router, is required.
Variants	Objects that are stored in the cache in various forms: the original form, fetched from the OCS; the transformed (compressed or uncompressed) form (if compression is used). If a required compression variant is not available, then one might be created upon a cache-hit. (Note: policy-based content transformations are not stored in the Proxy*SG*.)
Web FTP	Web FTP is used when a client connects in explicit mode using HTTP and accesses an ftp:// URL. The SG appliance translates the HTTP request into an FTP request for the OCS (if the content is not already cached), and then translates the FTP response with the file contents into an HTTP response for the client.
Websense log type	A proprietary log type that is compatible with the Websense reporter tool.

Term	Description
Wildcard Services	When multiple non-wildcard services are created on a port, all of them must be of the same service type (a wildcard service is one that is listening for that port on all IP addresses). If you have multiple IP addresses and you specify IP addresses for a port service, you cannot specify a different protocol if you define the same port on another IP address. For example, if you define HTTP port 80 on one IP address, you can only use the HTTP protocol on port 80 for other IP addresses. Also note that wildcard services and non-wildcard services cannot both exist at the same time on a given port. For all service types except HTTPS, a specific listener cannot be posted on a port if the same port has a wildcard listener of any service type already present.

Index