



Cisco WSA Splunk Application Installation, Setup, and User Guide

Splunk Application for Cisco WSA

Version 2.0.0 #
May 3, 2013

Cisco Systems, Inc. #
www.cisco.com

Cisco has more than 200 offices worldwide. #
Addresses, phone numbers, and fax numbers #
are listed on the Cisco website at #
www.cisco.com/go/offices.

THE SPECIFICATIONS AND INFORMATION REGARDING THE PRODUCTS IN THIS MANUAL ARE SUBJECT TO CHANGE WITHOUT NOTICE. ALL STATEMENTS, INFORMATION, AND RECOMMENDATIONS IN THIS MANUAL ARE BELIEVED TO BE ACCURATE BUT ARE PRESENTED WITHOUT WARRANTY OF ANY KIND, EXPRESS OR IMPLIED. USERS MUST TAKE FULL RESPONSIBILITY FOR THEIR APPLICATION OF ANY PRODUCTS.

THE SOFTWARE LICENSE AND LIMITED WARRANTY FOR THE ACCOMPANYING PRODUCT ARE SET FORTH IN THE INFORMATION PACKET THAT SHIPPED WITH THE PRODUCT AND ARE INCORPORATED HEREIN BY THIS REFERENCE. IF YOU ARE UNABLE TO LOCATE THE SOFTWARE LICENSE OR LIMITED WARRANTY, CONTACT YOUR CISCO REPRESENTATIVE FOR A COPY.

The Cisco implementation of TCP header compression is an adaptation of a program developed by the University of California, Berkeley (UCB) as part of UCB's public domain version of the UNIX operating system. All rights reserved. Copyright © 1981, Regents of the University of California.

NOTWITHSTANDING ANY OTHER WARRANTY HEREIN, ALL DOCUMENT FILES AND SOFTWARE OF THESE SUPPLIERS ARE PROVIDED "AS IS" WITH ALL FAULTS. CISCO AND THE ABOVE-NAMED SUPPLIERS DISCLAIM ALL WARRANTIES, EXPRESSED OR IMPLIED, INCLUDING, WITHOUT LIMITATION, THOSE OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE AND NONINFRINGEMENT OR ARISING FROM A COURSE OF DEALING, USAGE, OR TRADE PRACTICE.

IN NO EVENT SHALL CISCO OR ITS SUPPLIERS BE LIABLE FOR ANY INDIRECT, SPECIAL, CONSEQUENTIAL, OR INCIDENTAL DAMAGES, INCLUDING, WITHOUT LIMITATION, LOST PROFITS OR LOSS OR DAMAGE TO DATA ARISING OUT OF THE USE OR INABILITY TO USE THIS MANUAL, EVEN IF CISCO OR ITS SUPPLIERS HAVE BEEN ADVISED OF THE POSSIBILITY OF SUCH DAMAGES.

Cisco and the Cisco logo are trademarks or registered trademarks of Cisco and/or its affiliates in the U.S. and other countries. To view a list of Cisco trademarks, go to this URL: www.cisco.com/go/trademarks. Third-party trademarks mentioned are the property of their respective owners. The use of the word partner does not imply a partnership relationship between Cisco and any other company. (1110R)

Any Internet Protocol (IP) addresses and phone numbers used in this document are not intended to be actual addresses and phone numbers. Any examples, command display output, network topology diagrams, and other figures included in the document are shown for illustrative purposes only. Any use of actual IP addresses or phone numbers in illustrative content is unintentional and coincidental.

Cisco WSA Splunk Application Installation, Setup, and User Guide #
© 2013 Cisco Systems, Inc. All rights reserved.



CHAPTER 1

Installation and Setup	1-1
Introduction	1-1
Task Overview	1-2
Splunk Software	1-2
Download	1-2
Supported and Unsupported Splunk Features	1-2
Documentation	1-2
System Requirements	1-3
Splunk Instances	1-3
Splunk Application for Cisco Web Security Appliance	1-3
Sizing & Scaling Recommendations	1-3
Install and Configure Splunk	1-4
Notes about Authentication\Authorization	1-5
Splunk Basic Authentication	1-5
Install Splunk Application for Cisco Web Security Appliance	1-6
Installing Splunk Application for Cisco Web Security Appliance	1-6
Configuration Best Practices	1-6
Create the Folder Structure for Log Files	1-7
Import and Index Historical Data	1-7
(Optional) Customizing the Summary Script	1-7
(Optional) Estimating the Import Time	1-7
Importing and Indexing Historical Data	1-8
Set Up Ongoing Data Transfers	1-9
Configuring Data Inputs in Splunk	1-9
Establish Log Transfers from Cisco Web Security Appliance	1-10
(Optional) Setup Department Membership Query	1-10
Setting up Department Membership Reporting	1-11
Restricting Access to Department Reports by Role	1-11
Troubleshooting	1-12
(Optional) Setup Scheduled PDF Reporting	1-12
Setup Scheduled Reports	1-12
Additional Reading	1-13

Text Part Number:

CHAPTER 2

Reports 2-1

- Overview of Reports 2-1
- Data Formats 2-1
- Time Range 2-2
 - Timing of Data Availability 2-2
 - Timing of Summary Index Generation 2-3
- Export 2-4
 - Exporting to a .CSV File 2-4
 - Exporting to a PDF File 2-4
- General Versus Specific Data 2-4
 - Viewing Specifics 2-4
- Search 2-4
 - Department 2-5
- Predefined Reports 2-5
 - List of General Reports 2-5
 - List of Specific Reports 2-6
- Usage Scenarios 2-6
 - User Investigation 2-6
 - Viewing Web Usage Trends 2-7
 - Viewing Transaction History 2-7
 - URLs Visited 2-7
 - Viewing Most Visited Web Sites 2-7
 - URL Categories Visited 2-8
 - Viewing Most Common URL Categories 2-8

CHAPTER 3

Field Extractions 3-1

- Overview of Field Extractions 3-1
- Access Logs 3-1
- Traffic Monitor Logs 3-2



Installation and Setup

- [Introduction, page 1-1](#)
- [Task Overview, page 1-2](#)
- [Splunk Software, page 1-2](#)
- [System Requirements, page 1-3](#)
- [Sizing & Scaling Recommendations, page 1-3](#)
- [Install and Configure Splunk, page 1-4](#)
- [Install/Update Splunk Application for Cisco Web Security Appliance, page 1-6](#)
- [Create the Folder Structure for Log Files, page 1-7](#)
- [Import and Index Historical Data, page 1-7](#)
- [Set Up Ongoing Data Transfers, page 1-9](#)
- [\(Optional\) Setup Department Membership Query, page 1-10](#)
- [\(Optional\) Setup Scheduled PDF Reporting, page 1-12](#)
- [Additional Reading, page 1-13](#)

Introduction

Splunk for Cisco Web Security Appliance includes a customized Splunk application and a Splunk server that polls log data collected from a Cisco Web Security Appliance. The application provides reports and dashboards designed to provide insight into very large volumes of data from the Cisco Web Security Appliance.

The application receives data from a Cisco Web Security Appliance and stores the data in the default/main index. It generates summaries and stores them in the summary index. Customers can view these data using predefined reports. Customers can also perform ad hoc searches using the flashtimeline view and the web tracking forms.



Tip

Improve performance by choosing smaller time ranges and crafting searches to be as precise as possible.

Task Overview

-
- Step 1** [Install and Configure Splunk, page 1-4](#)
 - Step 2** [Install/Update Splunk Application for Cisco Web Security Appliance, page 1-6](#)
 - Step 3** [Create the Folder Structure for Log Files, page 1-7](#)
 - Step 4** [Import and Index Historical Data, page 1-7](#)
 - Step 5** [Set Up Ongoing Data Transfers, page 1-9 \(Including setup of Cisco Web Security Appliance.\)](#)
-

Splunk Software

Download

Splunk software is available as a free download from www.splunk.com.

Comprehensive documentation for installing and using Splunk software is available on the Splunk website at docs.splunk.com.

Related Topics

- [Install and Configure Splunk, page 1-4](#)

Supported and Unsupported Splunk Features

Component	Supported	Not Supported
Reports	Reports included in Splunk Application for Cisco Web Security Appliance.	Custom reports.
Search	Form-based search/web tracking tool included with the Splunk Application for Cisco Web Security Appliance	Native Splunk search engine.
Server	Single-server deployments	Multiple-server deployments
Transport Method	FTP (Files and Directories)	TCP
Virtualization	N/A	Virtualization of any core function of Splunk referenced within this document.
PDF Server Application	On Linux	On Windows

Documentation

This guide documents use of the Splunk Application for Cisco Web Security Appliance. Documentation for the Splunk product itself is available on the Splunk web site: <http://docs.splunk.com>.

System Requirements

Splunk Instances

Find the latest system requirements for Splunk on their web site:
<http://docs.splunk.com/Documentation/Splunk/latest/Installation/Systemrequirements>.

Splunk Application for Cisco Web Security Appliance

Operating system requirements:

- Red Hat Linux
- Windows

Platform Requirements: Reference hardware can be commodity-grade, and must have the following minimum specifications to be eligible for Cisco support.

- Intel x86-64-bit chip architecture with (2) CPU's, 4 cores per CPU, 2.5-3Ghz per core
- 16GB RAM
- (4) 300GB SAS hard disks at 10,000 rpm each in RAID10 (800 IOPS or better)
- Standard 1Gb Ethernet NIC, optional 2nd NIC for a management network

**Note**

Splunk is often constrained by disk I/O first, so always consider that first when selecting the storage hardware.

The file system will be assumed to be running on local disk volumes formatted as NTFS or EXT2/3. A separate OS volume should be created per industry best practices. The Splunk installation should reside on its own logical volume whenever possible.

Sizing & Scaling Recommendations

- The base configuration is a single-tier architecture with one server offering all 3 parts of the core functionality of a typical Splunk deployment:
 - a search head
 - an indexer
 - a monitor for data sources. a
- If the estimated requirements for indexed data volume exceed 100k/Users (estimate: 100GB/day,) the Splunk infrastructure should be adjusted.
- By adding another Splunk instance and adjusting the configuration, the new infrastructure would offer an increase in aggregate indexing and search performance (once the data is load-balanced), and an increase in storage and retention capacity.
- A dedicated forwarder server would also be added to the Splunk infrastructure and configured to monitor the WSA log files and forward the log data across multiple indexers using load balancing.

- To facilitate the implementation and configuration of an environment that exceeds 100k users, it is recommended that Cisco engage Splunk professional services on behalf of the Cisco Web Security Appliance customer.

Based upon log volume estimates against an Cisco Web Security Appliance device with 10k users, the amount of data collected is 10GB/day uncompressed. Once indexed, the data compresses to an estimated 2.5GB/day indexed storage used. The Splunk instance would retain approximately 200 days of indexed data based upon a volume size of 500GB.

Cisco Web Security Appliance Users	Estimated Log Volume (2,500 transactions/user/day)	Estimated Indexed Volume	Estimated retention (500GB volume)
10K	10GB/day	2.5GB	200 days
50K	50GB/day	13GB	40 days
100K	100GB/day	25GB	20 days

**Note**

Guidelines based upon estimated log volumes and mid-capacity drives in an array.

Daily Volume	77 GB/day	140 GB/day	180GB/day
Total Transactions	172 Million	325 Million	417 Million
Predefined Report Load time	<5 seconds	<10 seconds	<15 seconds

Total Volume	2.3 TB
Business days retention @70GB/day	33
Predefined Report Loading time	<20 seconds

Install and Configure Splunk

These tasks are out of the scope of this document but must be performed to use Splunk Application for Cisco Web Security Appliance. See Splunk documentation on the Splunk web site for help performing these tasks.

Task	More Information
Download and install the free Splunk software.	www.splunk.com
Login to Splunk using the admin account and change the password.	docs.splunk.com

Task	More Information
Licensing: <ol style="list-style-type: none"> 1. Consider the quantity of data to be indexed both during initial historical data upload and on a daily basis ongoing. 2. Acquire and upload a Splunk evaluation license sufficient for the historical data upload. 3. Acquire and upload a Splunk enterprise licence sufficient for the anticipated data of the applicable source type to be indexed. 4. Change the licence type from Trial to Evaluation or Enterprise. 5. Edit license pool to ensure that the index is reporting to the correct pool. First, customer may need an evaluation license good for a large volume of data to handle historical data input. Then,	docs.splunk.com See also: Additional Reading, page 1-13 .
Set the Cisco WSA Splunk application as the default app for all users/roles.	docs.splunk.com
(Optional) Enable SSL within Splunk.	docs.splunk.com
(Optional) Prepare associations with AD/LDAP: <ol style="list-style-type: none"> 1. Configure Splunk to use AD/LDAP for authentication. 2. Verify that Splunk can connect to your AD/LDAP server. 3. Map Existing AD/LDAP groups to Splunk roles 4. Add and edit roles within Splunk as needed. 5. (Optional) Enable SSL on your AD/LDAP server. 	Notes about Authentication\Authorization, page 1-5 docs.splunk.com
(Best practice) Verify Splunk services are set to restart automatically and test.	docs.splunk.com

Notes about Authentication\Authorization

- Splunk basic authentication
- AD/LDAP

Splunk Basic Authentication

Local Splunk authentication supersedes any other authentication option configured.

Default setup:

- Three roles: Define user privileges.
- One user account: admin, which is permanent. Use this account to configure, test, and troubleshoot.



Tip

- Add users to a splunk-specific group in the directory services.

- Import that group DN into Splunk.
- Either map the most appropriate default Splunk role to that group DN or create and map to a more appropriate role.

If your requirements are simple, for example, only a few people can view Splunk data, then using local authentication may be sufficient.

Install/Update Splunk Application for Cisco Web Security Appliance

Installing/Updating Splunk Application for Cisco Web Security Appliance

Before You Begin

- [Install and Configure Splunk, page 1-4](#)
- If you are upgrading from a previous version of Splunk Application for Cisco Web Security Appliance, uninstall the older version of the application.
- Receive the zip or tar file for the Splunk Application for Cisco Web Security Appliance.
- Open Splunk Web.

-
- Step 1** Within Splunk Web, navigate to **Manager> Apps>Install App from File**.
- Step 2** Browse to and select the zip or tar .file for the Splunk Application for Cisco Web Security Appliance.
- Step 3** Watch for notification that the application was imported successfully.
- Step 4** Restart Splunk: **Manager>Server Controls>Restart**.
- Step 5** Log into Splunk Web.
- Step 6** Navigate to **Manager>Apps** to verify that the Cisco WSA Splunk application is visible and enabled.
-

What to Do Next

- [Import and Index Historical Data, page 1-7](#).

Configuration Best Practices

- Set time zones consistently across Cisco Web Security Appliance appliances.
The time displayed in the search results reflects the 'local' time of the Splunk instance. By default, all Splunk inputs for the Cisco Web Security Appliance logs are set to TZ = GMT.
- Document the local admin account password (regardless of the chosen authentication method).

Create the Folder Structure for Log Files

Log	Default Path	Variables
Traffic Monitor	/\$Input_base/wsa_hostname/trafmonlogs/	\$Input_base=Splunk deployment host_name=WSA device
Access	/\$Input_base/wsa_hostname/accesslogs/	\$Input_base=deployment host_name=WSA device

Import and Index Historical Data

The default for the summary script is to summarize up to 90 days of history. By default, the summary script uses 8 cores.

(Optional) Customizing the Summary Script

- Step 1** Open the summary script for editing:
- Linux: `$SPLUNK_HOME/etc/apps/CiscoWSA/bin#/summary.sh`
 - Windows: `X:\$SPLUNK_HOME\etc\apps\CiscoWSA\bin\summary.vbs`

- Step 2** Search for this string:

```
time $Spath/bin/splunk cmd python $Spath/bin/fill_summary_index.py -app
SplunkforCiscoIronportWSA -namefile
$Spath/etc/apps/SplunkforCiscoIronportWSA/bin/summary.jobs -et -90d -lt now -j 8 -dedup true
```

- Step 3** Customize the start and end dates and the number of cores used by the summary script:

Setting	Default	Description
-et	-90d	Start day. Number of historical days at which to begin summarizing. The default value of -90d begins at 90 days prior to the current day.
-lt	now	End day. Number of historical days at which to stop summarizing. The default value of now stops with the current day. A default of -1d would stop with yesterday's data.
-j	8	Number of cores to be used by the summary script.

(Optional) Estimating the Import Time

The historical summary can take up to 9 hours to complete

- Step 1** Allow 4 minutes per 5M events (2GB of raw data) per summary job based upon the platform hardware recommendations.

Example: Expect a 10GB file representing 25M historical events to take 20 minutes to run against each summary job.

Step 2 Allow for the 27 summary jobs used by the Splunk Application for Cisco Web Security Appliance.

Importing and Indexing Historical Data

Before You Begin

- Complete configuration tasks listed in [Install and Configure Splunk, page 1-4](#).
 - Verify that field extractions are correct. See [Chapter 3, “Field Extractions”](#).
 - Know the folder structure. See [Create the Folder Structure for Log Files, page 1-7](#).
 - (Optional) See [\(Optional\) Estimating the Import Time, page 1-7](#).
-

Step 1 Copy the historical log files into the folder structure for log files.



Note By default, these logs will be deleted after the data is indexed.

Step 2 From a command prompt run the summary script:

Linux: `$SPLUNK_HOME/etc/apps/CiscoWSA/bin#/summary.sh`

Windows: `X:\$SPLUNK_HOME\etc\apps\CiscoWSA\bin\summary.vbs`

Step 3 Navigate to the Splunk folder and enter the local Splunk administrator credentials when prompted.



Note You may not see immediate results.

Step 4 Verify data is being imported:

- In Splunk Web, login as admin.
- Go to the search app.
- Go to Status>Index Activity>Index Activity Overview report.
- Look for summary index growth.

Step 5 If the historical data import was run under a Splunk evaluation license, install the Enterprise default license downloaded for the account and remove any non-Production licenses.

What to Do Next

- [Configuring Data Inputs in Splunk, page 1-9](#).

Set Up Ongoing Data Transfers

Configuring Data Inputs in Splunk

Before You Begin

- [Import and Index Historical Data, page 1-7](#)
- Know the path to your log files: [Create the Folder Structure for Log Files, page 1-7](#).
- Open Splunk Web.

-
- Step 1** In Splunk Web, navigate to **Manager>Data Inputs>Files and Directories**.
- Step 2** Disable any inputs labeled CiscoWSA.
- Step 3** Copy the file: `$SPLUNK_HOME/etc/apps/CiscoforIronportWSA/default/inputs.conf` # to the folder: `$SPLUNK_HOME/etc/apps/CiscoforIronportWSA/local/`
- Step 4** Using a text editor, open `$SPLUNK_HOME/etc/apps/CiscoforIronportWSA/local/inputs.conf`.
- Step 5** Locate the appropriate stanza for the input method and log source and edit the path.

Input Method	Stanza in inputs.conf File	More Information
Batch	<pre>sourcetype=wsa_accesslogs interval=60 move_policy = sinkhole</pre>	<p>This is the default. Reads and deletes the data.</p> <p>Only add <code>move_policy = sinkhole</code> if you want the original data to be deleted.</p> <p>Do not use Splunk as the primary log storage with batch input configuration.</p>
Monitor	<code>[monitor://<path>]</code>	<p>Splunk monitors a file or directory for changes.</p> <p><code>[batch:///data1/splunklogs/*]</code> (folder that is being monitored.)</p>

- Step 6** Within the same stanza, edit the value for disabled: `disabled#=#false`.
- Step 7** Save the file.
- Step 8** Restart Splunk.
- Step 9** In Splunk Web, navigate to **Manager>Data Inputs>Files and Directories**.
- Step 10** In Splunk Web, verify that the inputs are listed, enabled, and have the correct path.
- Step 11** In Splunk Web, for each input, set the source type manually to `wsa_accesslogs`.

Source Type	File/Directory name
More settings>Set the source type>Manual	<code>wsa_accesslogs</code>

- Step 12** In Splunk Web, for the input, select **More settings>Index>Default**.
-

Establish Log Transfers from Cisco Web Security Appliance

Before You Begin

- Know the path to your log files: [Create the Folder Structure for Log Files, page 1-7](#).
- Determine the frequency of transfers, no more than 60 minute increments.
- Open the web interface for the Cisco Web Security Appliance.

Step 1 In the web interface for the Cisco Web Security Appliance, # navigate to **System Administration>Log Subscriptions**.

Step 2 Click **Add Log Subscription...**

Step 3 Configure the subscription

Setting	Log Type	Value
Log directory	Access	accesslogs
	Traffic Monitor	trafmonlogs
Rollover by File Size [WSA version 7.5 and newer]	Either	Recommend no more than 500 Mb.
Maximum File Size [WSA version 7.1 and older]		
Rollover by Time [WSA version 7.5 and newer]	Either	Recommend custom rollover interval of one hour (1h) or more frequent rollovers.
Log Style	Access	Squid
	Traffic Monitor	N/A
(Optional) Custom Fields	Either	%XK (Adds a web reputation threat reason.)
Filename	Either	<user defined>
Retrieval Method	Either	FTP on <hostname_splunk_instance>



Note

Accessing online help from the Add Log Subscription page brings up detailed information about all the the Settings.

(Optional) Setup Department Membership Query

Perform the setup procedure for department membership requirements under these conditions:

- You will use AD/LDAP groups bound to roles in Splunk.
- You will run reports on data that is based on organizational roles.

Related Topics

- [Restricting Access to Department Reports by Role, page 1-11](#)

Setting up Department Membership Reporting

Before You Begin

- Linux users: Install ldapsearch tool using the following command:

```
sudo#yum#install#openldap-clients
```

Step 1 Identify the AD/LDAP Group Base DN's in the Membership Script:

- Open the appropriate membership script in a text editor:
 - Linux: `$SPLUNK_HOME/etc/apps/CiscoWSA/bin/discovery.py`
 - Windows: `X:\$SPLUNK_HOME\etc\apps\CiscoWSA\bin\discovery.vbs`
- Edit the first four fields at the top of the header:


```
strComputer#=# 'ad_ldap_host'
strUser#=# 'cn=service_account,cn=Users,dc=my_directory,dc=net'
strPassword#=# 'service_account_password'
strGroupOUs#=# 'Group#base#DN;Group#base#DN;Group#base#DN'
```
- Save the file.

Step 2 Enable use of the membership script by the inputs.conf Script:

- Open the inputs.conf script in a text editor:
`$SPLUNK_HOME/etc/apps/CiscoforIronportWSA/local/inputs.conf`
- Search for the appropriate string:
 - # membership script Windows
 - # membership script Linux
- Set disabled to false: `disabled#=#false`

Step 3 Verify that the script populated departments.csv with the user data:

```
$SPLUNK_HOME/etc/apps/CiscoWSA/lookups/departments.csv#
```

The membership script is set to run every day by default. The interval is set in seconds and can be changed as per the deployment requirements.

Restricting Access to Department Reports by Role

Before You Begin

- Understand that if users are restricted to viewing data from specific departments or groups, Layer 4 Transport Monitor (L4TM) data will only be available to administrators because L4TM data is not linked to a department or role.
- Open Splunk Web

Step 1 In Splunk Web, navigate to **Manager>Access controls>Roles**.

Step 2 Click New or edit an existing role.

Step 3 Define search restrictions for the role.

Example: To restrict a role to viewing data for the Sales Department, in the Restrict search terms field, type “department=sales”.

Step 4 Click **Save**.

Troubleshooting



Tip

- Linux users: Verify that `ldapsearch` tool is in the Splunk user’s path.
- Verify that the `departments.csv` file exists in the application’s lookup folder.
- Windows users: Comment out “option explicit” to reveal more specific information the origin and cause of an error.
- Verify the LDAP paths are syntactically correct.
- Verify the bind service account name is correct.
- Verify the correct bind password is entered.
- Test connection to the remote machine over port 389.
- Verify the correct attribute was configured for the member name.
- Verify the correct attribute was used for group membership.
- Verify the correct attribute was configured for group name.

(Optional) Setup Scheduled PDF Reporting



Note

Scheduled PDF reporting requires a Linux-based instance of Splunk running on the network. For a minimal installation. However, a standard Linux image with an installation of Splunk configured as a forwarder (no indexing or web interface required) can serve multiple Splunk instances for PDF generation.

Splunk Web users can generate a scheduled PDF output from any dashboard, view, search or report. To enable this functionality, the PDF Report Server app must be downloaded from Splunkbase and installed into a Splunk instance on a single Linux host. In addition, an internal email server will be configured in Splunk to allow it to send the PDF reports.

Setup Scheduled Reports

Step 1 Download and install the PDF Report Server add-on from Splunk:

<http://splunk-base.splunk.com/apps/22348/pdf-report-server-install-on-linux-only>

Step 2 Ensure that the Xvfb X server, xauth and fonts for your Linux distribution are installed. These are included with most Linux distributions, but not installed by default. On Red Hat, type:

```
yum#install#Xvfb#xauth#bitstream-vera-fonts#
```


- Step 3** Launch Splunk Web on the Linux host.
- Step 4** Navigate to **Manager>System Settings > Email Alert Settings**.
- Step 5** Check the **Use PDF Report Server** box.
- Step 6** Click **Save**.
- Step 7** Under Mail server settings, enter or update information related to the SMTP server that Splunk interacts with in order to send out alert emails.
- Step 8** Identify the SMTP mail host server.
- Step 9** Provide an authentication username/password if the SMTP server requires them.
- Step 10** (Optional) Specify that Splunk uses SSL or TLS when it communicates with the SMTP server.
- Step 11** Under Email format, configure the format of the emails that Splunk sends.

You can define the name that appears in the "sender" field (by default it is Splunk), and you can set up the format of the email subject line (by default it is Splunk Alert: \$name\$, where \$name\$ is the name of the search that the alert is based upon). You can also set at the Manager level the default email format for all alerts and whether or not alert emails provide inline results.

If the hostname of the Splunk Web instance that this PDF Report Server will talk to is not resolvable in DNS, enter its IP address or a hostname that resolves to that IP in the Link hostname field. This will ensure that Splunk Web can contact the PDF Report Server, and that links sent in emailed PDF reports work correctly. If the field is left empty, Splunk will try to autodetect the hostname.

- Step 12** To change the splunk core service port: From the %SPLUNK_HOME%\bin directory: splunk set splunkd-port #####
-

Additional Reading

- Splunk License Installation:
<http://www.splunk.com/base/Documentation/latest/Admin/Installlicense>
- Splunk License Violations:
<http://www.splunk.com/base/Documentation/latest/Admin/Aboutlicenseviolations>
- Backup of Splunk data:
<http://www.splunk.com/base/Documentation/latest/admin/Backupindexeddata>
- How to archive data:
<http://www.splunk.com/base/Documentation/latest/Admin/Automatearchiving>



Reports

- [Overview of Reports, page 2-1](#)
- [Data Formats, page 2-1](#)
- [Time Range, page 2-2](#)
- [Export, page 2-4](#)
- [Export, page 2-4](#)
- [General Versus Specific Data, page 2-4](#)
- [Predefined Reports, page 2-5](#)
- [Usage Scenarios, page 2-6](#)

Overview of Reports

Splunk Application for Cisco Web Security Appliance includes a set of predefined reports. As much as possible the reporting is consistent with the native reporting of Cisco Web Security Appliance.



Note

Reports generated using Splunk Application for Cisco Web Security Appliance may show more data than is available through Splunk due to the use of a summary index, which speeds the loading of reports.



Tip

Splunk administrators can control the hosts you see on the Overview report and Web Tracking report. Contact your Splunk administrator with details of any hosts you would like to add, remove, or rename.

Data Formats

In some cases, the formatting of data available through Cisco WSA Splunk application differs from the formatting of data available through native reporting functionalities.

Data	Format
Large numbers (greater than seven digits)	2E11 means 2×10^{11}
Time	d+hh:mm:ss.ms Example: 1+03:22:36.00 1 day, 3 hours, 22 minutes, # 36 seconds, 0 milliseconds

Time Range



Tip

Select a smaller time range to return results more quickly.

Timing of Data Availability

Range	Indexing Begins	Data Appears in Reports
Hour	Just past the hour	60-90 minutes after indexing begins
Day	After midnight daily	One day after indexing begins
Week	After midnight Saturday # (early Sunday morning)	One week after indexing begins
90 Days	After midnight of the 90th day.	90 days after indexing begins.
Custom: Less than hourly	Just past the hour	60-90 minutes after indexing begins
Custom: Less than daily	After midnight daily	One day after indexing begins
Custom: Less than weekly	After midnight Saturday # (early Sunday morning)	One week after indexing begins



Tip

Use the “jobs” menu to verify that scheduled searches are not running too long. “Too long” is in excess of its frequency, for example, a weekly search running more than a week.



Tip

Select the Jobs menu. Each report’s search(es) will have a marker denoting search description & interval summary. For example, search `_dashboard_users_base-search(*,1d)` is leveraging the user’s 1 day summary.

Timing of Summary Index Generation

Summary indexes speed report generation. Summaries are generated once per hour. Each night, hourly summaries are aggregated into daily summaries. Each week, daily summaries are aggregated into weekly summaries.

Summary Search	Frequency
[_dashboard SOCKS_base-sum-search-top-1h]	Hourly at 10 minutes past
[_dashboard SOCKS_base-sum-search-top-1d]	Daily at 2:30 AM
[_dashboard SOCKS_base-sum-search-top-1w]	Weekly at 1:45 AM
[_dashboard_anti-malware_base-sum-search-1h]	Hourly at 35 minutes past
[_dashboard_anti-malware_base-sum-search-1d]	Daily 1:30 AM
[_dashboard_anti-malware_base-sum-search-1w]	Sunday 2:15 AM
[_dashboard_application-visibility_base-sum-search-1h]	Hourly at 15 minutes past
[_dashboard_application-visibility_base-sum-search-1d]	Daily at 12:30 AM
[_dashboard_application-visibility_base-sum-search-1w]	Sunday at 2:45 AM
[_dashboard_overview_base-sum-search-bottom-1h]	Hourly at 50 minutes past
[_dashboard_overview_base-sum-search-bottom-1d]	Daily at 6:00 AM
[_dashboard_overview_base-sum-search-bottom-1w]	Weekly at 3:15 AM
[_dashboard_overview_base-sum-search-top-1h]	Hourly at the top of the hour.
[_dashboard_overview_base-sum-search-top-1d]	Daily at 5:00 AM
[_dashboard_overview_base-sum-search-top-1w]	Weekly at 3:45 AM
[_dashboard_overview_base-sum-search-uid-1h]	Hourly at 40 minutes past
[_dashboard_overview_base-sum-search-uid-1d]	Daily at 4:00 AM
[_dashboard_overview_base-sum-search-uid-1w]	Weekly at 4:15 AM
[_dashboard_url-categories_base-sum-search-1h]	Hourly at 30 minutes past
[_dashboard_url-categories_base-sum-search-1d]	Daily at 3:00 AM
[_dashboard_url-categories_base-sum-search-1w]	Weekly at 5:15 AM
[_dashboard_users_base-sum-search-1h]	Hourly at 20 minutes past
[_dashboard_users_base-sum-search-1d]	Daily at 2:00 AM
[_dashboard_users_base-sum-search-1w]	Weekly at 5:45 AM
[_dashboard_web-reputation-filters_base-sum-search-1h]	Hourly at 45 minutes past
[_dashboard_web-reputation-filters_base-sum-search-1d]	Daily at 1:00 AM
[_dashboard_web-reputation-filters_base-sum-search-1w]	Weekly at 4:45 AM
[_dashboard_web-sites_base-sum-search-1h]	Hourly at 55 minutes past
[_dashboard_web-sites_base-sum-search-1d]	Daily at Midnight
[_dashboard_web-sites_base-sum-search-1w]	Sunday at 1:15 AM

Export

Exporting to a .CSV File

Step 1 Generate the report.

Step 2 Select **Export**.

Exporting to a PDF File

Before You Begin

- Verify that the Splunk administrator has enabled PDF output.
-

Step 1 Generate the report.

Step 2 Select **Save as PDF**.

Related Topics

- [\(Optional\) Setup Scheduled PDF Reporting, page 1-12](#)

General Versus Specific Data

Predefined general reports provide hyperlinks to predefined specific reports.

Viewing Specifics

Step 1 Select the most appropriate predefined general report.

For example, if you want specific information about a user, begin with the predefined Users report.

Step 2 Click on the hyperlink for the subject for which you want specifics.

For example, click on the hyper-linked user name or IP address for an individual user.

Related Topics

- [Export, page 2-4](#)

Search

Simple and advanced search options are available using the Web Tracking Report.

**Timesaver**

Make the search as specific as possible and narrow the time range.

**Tip**

Splunk Application for Cisco Web Security Appliance uses a set of files to populate menus for the Web Tracking page. If you are experiencing problems with the Web Tracking page menus, verify that these files are in the application's lookups folder:

- malware_categories.csv
- transaction_types.csv
- url_categories.csv

**Tip**

The Splunk administrator can edit the list of URL categories visible within Splunk. When a category appears within the access log, but is not present in the lookup file, Splunk Application for Cisco Web Security Appliance displays "Custom Category".

Department

The departments.csv is a file used as part of the role based security functionality. This file may be edited manually or by configuring one of the role discovery scripts (available in the application's bin folder) as a scripted input. There is a script for both Linux and Windows.

- Ensure the file exists in the application's lookup folder
- If the Linux version is used, ensure the CLI ldapsearch is installed and in the Splunk user's path
- If the Windows version is used "option explicit" may be commented out to reveal more specific information regarding from where and why an error may have originated.
- Verify the LDAP paths are syntactically correct
- Verify the bind service account name is correct
- Verify the correct bind password is entered
- Test connection to the remote machine over port 389
- Verify the correct attribute was configured for the member name
- Verify the correct attribute was used for group membership
- Verify the correct attribute was configured for group name

**Tip**

Splunk administrators can control the options available in the dropdown fields in the Web Tracking form.

Predefined Reports

List of General Reports

- Overview

- Users
- Web Sites
- URL Categories
- Application Visibility
- Anti-Malware
- Client Malware Risk
- Web Reputation Filters
- L4 Traffic Monitor

List of Specific Reports

- Malware Category
- Malware Threat
- Application
- Application Type
- Domain
- URL Category
- User
- Reports by Location
 - Overview by Location
 - URL Categories by Location
 - Anti-Malware by Location
 - Web Reputation Filters by Location
 - Application Visibility by Location
 - Users by Location
 - Websites by Location

Related Topics

- [Search, page 2-4](#)

Usage Scenarios

User Investigation

This example demonstrates how a system administrator would investigate a particular user at a company. In this scenario, a manager has received a complaint that an employee is visiting inappropriate web sites at work. To investigate this, the system administrator now needs to look at the employee's web usage trends and transaction history:

- URL Categories by Total Transactions

- Trend by Total Transactions
- URL Categories Matched
- Domains Matched
- Applications Matched
- Malware Threats Detected
- Policies Matched for a particular User ID or Client IP.

Using these reports, the system administrator can discover whether, for example, user “johndoe” was trying to access blocked URLs, which can be viewed in the Transactions Blocked column under the Domains section.

Viewing Web Usage Trends

Step 1 Select **Users** from the Cisco WSA Splunk application dropdown menu.

Step 2 Click on the User ID or Client IP address.



Note If you do not see the User ID or Client IP address you want to investigate in the Users table, click on any User ID or Client IP. Then search for all or part of the User ID or Client IP address.

Step 3 (Optional) Select **Actions >Print**.

Viewing Transaction History

Step 1 Select **Web Tracking** from the Cisco WSA Splunk application dropdown menu.

Step 2 **Search** for the User/Client IP Address.

Step 3 Click **Pick fields** above the transaction list to change the information displayed for each transaction.

Step 4 (Optional) Click **Export** to export the data to a CSV file.

URLs Visited

In this scenario, a Sales manager wants to discover the top five visited web sites at their company for the last week. Additionally, the manager wants to know which users are going to those websites.

Viewing Most Visited Web Sites

Step 1 Select **Web Sites** from the Cisco WSA Splunk application dropdown menu.

Step 2 Select **Week** from the Time Range drop-down list.

Step 3 View the top 25 domains in the Domains Matched table.

Step 4 Click on a domain to view the users who have visited that domain in order of frequency.

URL Categories Visited

In this scenario, the Human Resources manager wants to know what the top three URL categories her employees have visited over the past 30 days. Additionally, a network manager wants to get this information to monitor bandwidth usage, to find out what URLs are taking up the most bandwidth on her network. The example below is to show how you can gather data for several people covering several points of interest, while only having to generate one report.

Viewing Most Common URL Categories

Before You Begin

- [\(Optional\) Setup Scheduled PDF Reporting, page 1-12](#)

-
- Step 1** Select **URL Categories** from the Cisco WSA Splunk application dropdown menu.
 - Step 2** View the top ten URL Categories by Total Transactions graph.
 - Step 3** Select **Actions>Schedule for PDF Delivery**.
 - Step 4** Send the PDF to the Human Resources manager.
 - Step 5** View the Bytes Allowed column in the URL Categories Matches table.
 - Step 6** Select **Actions>Schedule for PDF Delivery**.
 - Step 7** Send the PDF file to the Network Manager.
 - Step 8** For finer granularity, click on a specific URL Category.
-



Field Extractions

- [Overview of Field Extractions, page 3-1](#)
- [Access Logs, page 3-1](#)
- [Traffic Monitor Logs, page 3-2](#)

Overview of Field Extractions

This application relies heavily on field extractions. As most reports are generated from summary data, – it is important to ensure fields are being extracted correctly to enable successful and accurate reporting.

Access Logs



Tip

- Ensure timestamps are correctly being indexed
- Search for “*” and ensure app-specific fields are populated in the field picker. The next bullet item contains a more thorough examination of extracted fields
- Copy and paste the below search. You should not see any results and especially not very many results. If 1000 results are returned – the transforms.conf will need to be adjusted for the unique log format being indexed...

```
sourcetype=wsa_accesslogs#|#head#1000#|#fillnull#value="!!!!"##  
x_webcat_code_abbrev#x_wbrs_score#x_webroot_scanverdict#  
x_webroot_threat_name#x_webroot_trr#x_webroot_spyid#  
x_webroot_trace_id#x_mcafee_scanverdict#x_mcafee_filename#  
x_mcafee_scan_error#x_mcafee_detecttype#x_mcafee_av_virustype#  
x_mcafee_virus_name#x_sophos_scanverdict#x#x_sophos_filename#  
x_sophos_virus_name#x_ids_verdict#x_icap_verdict#  
x_webcat_req_code_abbrev#x_webcat_resp_code_abbrev#  
x_resp_dvs_threat_name#x_wbrs_threat_type#x_avc_app#x_avc_type#  
x_avc_behavior#x_request_rewrite#x_avg_bw#x_bw_throttled#  
x_user_type#  
x_resp_dvs_verdict#x_req_dvs_threat_name#x_suspect_user_agent#  
x_wbrs_threat_reason#dvc_time#duration#dvc_ip#result#http_status#  
bytes_in#http_method#dest_url#user_id_dom#hierarchy#hierarchy_domain#  
mime_type#acl_tag#user_id#user_domain#dest_domain#|#stats#count#by#
```

```
x_webcat_code_abbr#x_wbrs_score#x_webroot_scanverdict#
x_webroot_threat_name#x_webroot_trr#x_webroot_spyid#
x_webroot_trace_id#x_mcafee_scanverdict#x_mcafee_filename#
x_mcafee_scan_error#x_mcafee_detecttype#x_mcafee_av_virustype#
x_mcafee_virus_name#x_sophos_scanverdict#x_sophos_filename#
x_sophos_virus_name#x_ids_verdict#x_icap_verdict#
x_webcat_req_code_abbr#x_webcat_resp_code_abbr#
x_resp_dvs_threat_name#x_wbrs_threat_type#x_avc_app#x_avc_type#
x_avc_behavior#x_request_rewrite#x_avg_bw#x_bw_throttled#
x_user_type#
x_resp_dvs_verdict#x_req_dvs_threat_name#x_suspect_user_agent#
x_wbrs_threat_reason#dvc_time#duration#dvc_ip#result#http_status#
bytes_in#http_method#dest_url#user_id_dom#hierarchy#hierarchy_domain#
mime_type#acl_tag#user_id#user_domain#dest_domain#|#convert#
ctime(dvc_time)#|#search#user_id="!!!!"#AND#host="!!!!"#AND#
src_ip="!!!!"#AND#cause="!!!!"#AND#action="!!!!"#AND#
dest_domain="!!!!"
```

- Verify the host extractions are correct. This is part of the inputs strategy discussed in the installation guide. The folder structure should be appropriately established to allow proper host extractions to occur.
- Hosts may be renamed per the section of this guide that discusses the host lookup file

Traffic Monitor Logs

The L4TM reports are generated from L4TM data (not summary data). Field extractions will still need to be operable for those reports to function. Though the format is not as versatile as accesslogs, they may still be verified with the same technique.



Tip

Use this search to verify there are few or no results:

```
sourcetype=wsa_trafmonlogs#|#head#1000#|#fillnull#value="!!!!"#
dvc_time#log_level#action#proto#src_ip#src_port#dest_ip#dest_host#
dest_port#|#stats#count#by#dvc_time#log_level#action#proto#src_ip#
src_port#dest_ip#dest_host#dest_port#|#search#src_ip="!!!!"
```



Numerics

2E11 [2-2](#)

B

best practices [1-5, 1-6](#)

C

Custom Category [2-5](#)

D

data formats [2-1](#)

E

Export [2-4](#)

H

hosts [2-1](#)

L

L4TM [1-11](#)

Layer 4 Transport Monitor data [1-11](#)

M

menus, missing items in [2-5](#)

P

PDF [2-4](#)

S

search [2-4](#)

T

time format [2-2](#)

time zones [1-6](#)

U

URL categories [2-5](#)

W

Web Tracking Report [2-4](#)

