



SPLUNK

Part 2

Introduction To Splunk

splunk>

Table of contents

Module 3: Understanding Splunk and the Data Pipeline

- An Introduction to Splunk Indexes	2
- A First Look at the Splunk Web GUI	3
- Enabling the Receiver on Splunk Enterprise	9
- Understanding Key Configuration Files	12
- Understanding and Installing Splunk Apps	16
- Shipping Windows Event Logs into Splunk	20

Module 4: Getting External Data into Splunk

- Establishing a Dedicated Index for Testing Data	24
- Ingesting Exported Windows Event Log Files	30
- Ingesting Dynamic Windows Registry Data	44
- Onboarding Generic Linux System Logs	46
- Onboarding Authentication Logs from Ubuntu UF	51
- Configuring Ingestion for Apache Web Server Access Logs	53
- Comprehensive Guide to Onboarding CSV Files	62
- Custom Data Sources and the "Great 8" Best Practices	67
- Extracting Data Fields using the EXTRACT Command	71
- Extracting Data Fields using the REPORT Command	76

An Introduction to Splunk Indexes

1. What is an Index in Splunk?

An index is like a big folder where Splunk stores all the data you send to it. It organizes the data so you can search it quickly later.

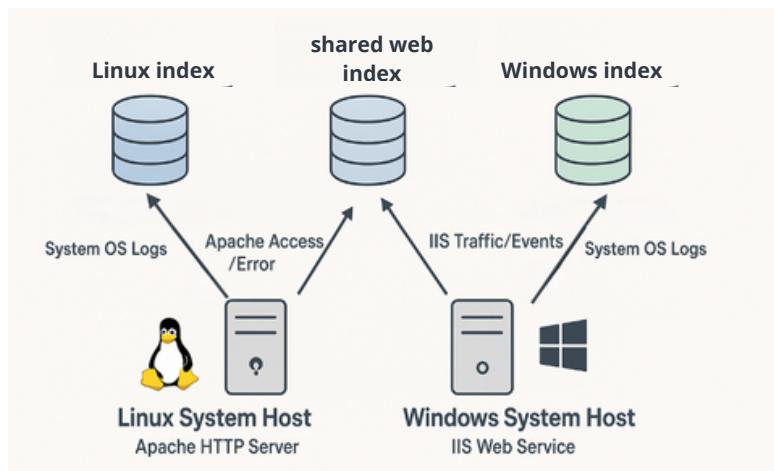
2. What does an Index do?

- Stores Data:** It's the main place where Splunk saves your logs and information.
- Groups Data:** You can keep similar data together (e.g., Linux logs in one index, web logs in another).
- Controls Access:** You decide who can see which index (security).
- Keeps Data for a Time:** You set how long Splunk keeps the data before deleting it.

3. Types of Indexes:

- Events Index:** For normal logs and messages with time stamps.
- Metrics Index:** For numbers like CPU usage or network speed. It's faster and uses less space for this kind of data.

4. How Indexes Work (Example):



Indexes help organize your data so searches are faster and easier.

Here's how it works:

- Linux logs → go into an index called `linux_logs`.
- Windows logs → go into an index called `windows_logs`.
- Web server logs → can go into a shared index called `web_access`.

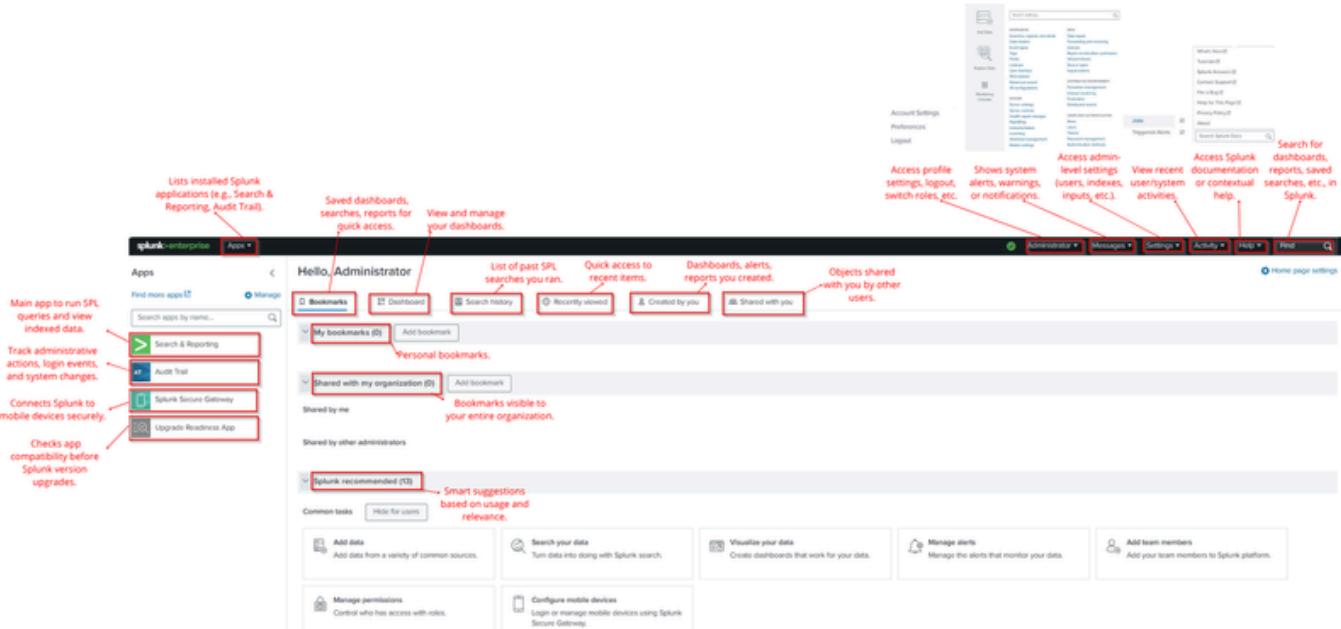
This structure makes searching and analysis simple because you know exactly where to look.

A First Look at the Splunk Web GUI

The Splunk Web interface is your main workspace for searching and analyzing data. It's designed to be simple and powerful.

1. The Splunk Enterprise Home Page

When you log in to Splunk, the Home Page is your main dashboard. It's the central hub for everything you do in Splunk.



Key Features on the Home Page:

- Apps Panel:** Access installed apps like Search & Reporting (for running queries), Audit Trail, and more.
- Bookmarks & Dashboards:** Quickly open saved searches, dashboards, and reports.
- Navigation Bar:**
 - Administrator Menu:** Manage your account settings.
 - Messages & Settings:** View system alerts and configure Splunk settings.
 - Activity & Help:** Check recent user activity and access documentation.
- Search Bar:** Find dashboards, reports, and saved searches easily.

This page helps you start searches, manage data, and access dashboards all in one place.

2. Performing a Basic Search

The Search & Reporting app is where you'll spend most of your time in Splunk. It's the main tool for exploring and analyzing data.

Understanding Splunk and the Data Pipeline

2.1 Accessing the Search Interface

- From the Home Page, click Search & Reporting (usually the first app listed).
- You'll see the Search screen with:
 - A search bar for your queries.
 - A time range picker (default: Last 24 hours).

The screenshot shows the Splunk Home Page. At the top, there's a navigation bar with 'splunk+enterprise' and 'Apps'. Below it, a secondary navigation bar includes 'Search', 'Analytics', 'Datasets', 'Reports', 'Alerts', and 'Dashboards'. On the right side of the page, a green button labeled 'Search & Reporting' is highlighted with a red box. A search bar labeled 'enter search here...' is also highlighted with a red box. To the right of the search bar is a tooltip: 'To type your SPL (Splunk Search Language) queries.' Further to the right is a time range selector set to 'Last 24 hours', which is also highlighted with a red box. A tooltip for this selector says: 'Sets the time range for the data to be searched.' At the bottom right of the main search area is a green search icon.

2.2 Constructing Your Search Query

- Search Bar:** Type your SPL (Splunk Search Language) queries here.
- Time Range Picker:** Choose the time frame for your search.

The screenshot shows the Splunk Search interface. At the top, there's a navigation bar with 'splunk+enterprise' and 'Apps'. Below it, a secondary navigation bar includes 'Search', 'Analytics', 'Datasets', 'Reports', 'Alerts', and 'Dashboards'. The 'Search' tab is active. On the right side, there are buttons for 'Save As', 'Create Table View', and 'Close'. A search bar contains the query 'index=_internal'. To the right of the search bar is a time range selector set to 'Last 7 days'. Below the search bar, a message indicates '420,736 events (6/27/25 7:00:00.000 PM to 7/4/25 7:20:29.000 PM)' and 'No Event Sampling'. The main area shows a timeline visualization with several bars representing event counts over time. Below the visualization, a table lists search results. The table has columns for 'Time' and 'Event'. The first few rows of the table are as follows:

Time	Event
7/4/25 7:20:27.066 PM	192.168.1.5 - admin [04/Jul/2025:19:20:27.066 +0000] "GET /en-US/splunkd/_raw/services/server/health/splunkd?output_mode=json&_=1751655885622 HTTP/1.1" 200 413 "-" "Mozilla/5.0 (Windows NT 10.0; Win64; x64) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/138.0.0.0 Safari/138.0.0.0" - b91dc96499b1d82780dc8fc328e5bcad 5ms host = splunk source = /home/splunk/splunk/var/log/splunk/splunkd_ui_access.log sourcetype = splunkd_ui_access
7/4/25 7:20:25.727 PM	07-04-2025 19:20:25.727 +0000 INFO PeriodicHealthReporter - feature="System Check" color=green due_to_stanza="feature:wlm_system_check" node_type=feature node_path=splunkd.workload_management.system_check host = splunk source = /home/splunk/splunk/var/log/splunk/health.log sourcetype = splunkd
7/4/25 7:20:25.727 PM	07-04-2025 19:20:25.727 +0000 INFO PeriodicHealthReporter - feature="Configuration Check" color=green due_to_stanza="feature:wlm_configuration_check" node_type=feature node_path=splunkd.workload_management.configuration_check host = splunk source = /home/splunk/splunk/var/log/splunk/health.log sourcetype = splunkd
7/4/25 7:20:25.727 PM	07-04-2025 19:20:25.727 +0000 INFO PeriodicHealthReporter - feature="Admission Rules Check" color=green due_to_stanza="feature:admission_rules_check" node_type=feature node_path=splunkd.workload_management.admission_rules_check host = splunk source = /home/splunk/splunk/var/log/splunk/health.log sourcetype = splunkd

2.3 Understanding Search Results

i	Time	Event	
✓	7/4/25 7:20:27.066 PM	192.168.1.5 - admin [04/Jul/2025:19:20:27.066 +0000] "GET /en-US/splunkd/_raw/services/server/health/splunkd?output_mode=json&_=1751655885622 HTTP/1.1" 200 413 "-" "Mozilla/5.0 (Windows NT 10.0; Win64; x64) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/138.0.0.0 Safari/537.36 Edg/138.0.0.0" - b91dc96499b1d82780dc8fc328e5bcad 5ms	
Event Actions ▾			
Type	Field	Value	Actions
Selected	<input checked="" type="checkbox"/> host	splunk	▼
	<input checked="" type="checkbox"/> source	/home/splunk/splunk/var/log/splunk/splunkd_ui_access.log	▼
	<input checked="" type="checkbox"/> sourcetype	splunkd_ui_access	▼
Event	<input type="checkbox"/> bytes	413	▼
	<input type="checkbox"/> clientip	192.168.1.5	▼
	<input type="checkbox"/> file	splunkd	▼
	<input type="checkbox"/> ident	-	▼
	<input type="checkbox"/> method	GET	▼
	<input type="checkbox"/> other	- b91dc96499b1d82780dc8fc328e5bcad 5ms	▼
	<input type="checkbox"/> referer	-	▼
	<input type="checkbox"/> req_time	04/Jul/2025:19:20:27.066 +0000	▼
	<input type="checkbox"/> root	en-US	▼
	<input type="checkbox"/> spent	5	▼
	<input type="checkbox"/> status	200	▼
	<input type="checkbox"/> uri	/en-US/splunkd/_raw/services/server/health/splunkd?output_mode=json&_=1751655885622	▼
	<input type="checkbox"/> url_path	/en-US/splunkd/_raw/services/server/health/splunkd	▼
	<input type="checkbox"/> url_query	output_mode=json&_=1751655885622	▼
	<input type="checkbox"/> user	admin	▼
	<input type="checkbox"/> useragent	Mozilla/5.0 (Windows NT 10.0; Win64; x64) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/138.0.0.0 Safari/537.36 Edg/138.0.0.0	▼
	<input type="checkbox"/> version	HTTP/1.1	▼
Time	<input type="checkbox"/> _time	2025-07-04T19:20:27.066+00:00	▼
Default	<input type="checkbox"/> index	_internal	▼
	<input type="checkbox"/> linecount	1	▼
	<input type="checkbox"/> punct[!/:_]+_*_-!//!!/?=&=/_.*_*_-!/_	▼
	<input type="checkbox"/> splunk_server	splunk	▼

After running a search, results appear in a table with:

- **Timestamp:** When the event happened.
- **Event Actions:** Options for that event.
 - **Columns:** Type: Category of the field.
 - **Field:** Name of the data point.
 - **Value:** Actual data.
 - **Actions:** Filter or add to search.

Common fields include:

- host (server name)
- source (file path)
- sourcetype (data format)
- clientip (IP address)
- _time (Splunk timestamp)
- index (where the event is stored)

2.4 Managing Displayed Fields

Splunk extracts numerous fields from your data. You can control which fields are shown in the event list.

< Hide Fields		All Fields	i Time	Event
SELECTED FIELDS			> 7/25 4:49:09.959 PM	192.168.1.5 - admin [05/Jul/2025:16:49:09.959 +0000] "GET /en-US/splunkd/_raw/services/search/jobs/1751734145.98/timeline?offset=0&count=10000_=1751734118658 HTTP/1.1" 200 400 "-" "Mozilla/5.0 (Windows NT 10.0; Win64; x64) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/138.0.0.0 Safari/537.36 Edg/138.0.0.0" - 8e9078f6e76fd54fe88821ffc2ce5927 6ms host = splunk source = /home/splunk/splunk/var/log/splunk/splunkd_ui_access.log sourcetype = splunkd_ui_access
INTERESTING FIELDS			> 7/25 4:49:09.956 PM	192.168.1.5 - admin [05/Jul/2025:16:49:09.956 +0000] "GET /en-US/splunkd/_raw/servicesNS/nobody/search/search/v2/jobs/1751734145.98/summary?output_mode=json&min_freq=0_=1751734118657 HTTP/1.1" 200 22512 "-" "Mozilla/5.0 (Windows NT 10.0; Win64; x64) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/138.0.0.0 Safari/537.36 Edg/138.0.0.0" - 8e9078f6e76fd54fe88821ffc2ce5927 93ms host = splunk source = /home/splunk/splunk/var/log/splunk/splunkd_ui_access.log sourcetype = splunkd_ui_access
			> 7/25 4:49:09.926 PM	192.168.1.5 - admin [05/Jul/2025:16:49:09.926 +0000] "GET /en-US/splunkd/_raw/servicesNS/nobody/search/search/v2/jobs/1751734145.98/output_mode=json&min_freq=0_=1751734118656 HTTP/1.1" 200 3261 "-" "Mozilla/5.0 (Windows NT 10.0; Win64; x64) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/138.0.0.0 Safari/537.36 Edg/138.0.0.0" - 8e9078f6e76fd54fe88821ffc2ce5927 9ms host = splunk source = /home/splunk/splunk/var/log/splunk/splunkd_ui_access.log sourcetype = splunkd_ui_access
			> 7/25 4:49:08.938 PM	192.168.1.5 - admin [05/Jul/2025:16:49:08.938 +0000] "GET /en-US/splunkd/_raw/services/search/jobs/1751734145.98/timeline?offset=0&count=10000_=1751734118655 HTTP/1.1" 200 407 "-" "Mozilla/5.0 (Windows NT 10.0; Win64; x64) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/138.0.0.0 Safari/537.36 Edg/138.0.0.0" - 8e9078f6e76fd54fe88821ffc2ce5927 3ms host = splunk source = /home/splunk/splunk/var/log/splunk/splunkd_ui_access.log sourcetype = splunkd_ui_access
		647 more fields	> 7/25 4:49:08.936 PM	192.168.1.5 - admin [05/Jul/2025:16:49:08.936 +0000] "GET /en-US/splunkd/_raw/servicesNS/nobody/search/search/v2/jobs/1751734145.98/summary?output_mode=json&min_freq=0_=1751734118654 HTTP/1.1" 200 15695 "-" "Mozilla/5.0 (Windows NT 10.0; Win64; x64) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/138.0.0.0 Safari/537.36 Edg/138.0.0.0" - 8e9078f6e76fd54fe88821ffc2ce5927 31ms host = splunk source = /home/splunk/splunk/var/log/splunk/splunkd_ui_access.log sourcetype = splunkd_ui_access
		+ Extract New Fields		

- Click “647 more fields” to see all fields.

Select Fields

Select All Within Filter	Deselect All	Coverage: 1% or more	Filter	Q	+ Extract New Fields
<input type="checkbox"/>	<input checked="" type="checkbox"/>	Field		# of Values	Event Coverage
> <input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	host		1	100%
> <input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	source		41	100%
> <input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	sourcetype		22	100%
> <input type="checkbox"/>	<input type="checkbox"/>	active		5	1.02%
> <input type="checkbox"/>	<input type="checkbox"/>	average_kbps		>100	1.75%
> <input type="checkbox"/>	<input type="checkbox"/>	avg_age		>100	14.57%
> <input type="checkbox"/>	<input type="checkbox"/>	blocked_count		20	3.18%
> <input type="checkbox"/>	<input type="checkbox"/>	bytes		>100	18.7%
> <input type="checkbox"/>	<input type="checkbox"/>	c		20	1.09%
> <input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	clientip		4	7.85%
> <input type="checkbox"/>	<input type="checkbox"/>	component		>100	89.38%
> <input type="checkbox"/>	<input type="checkbox"/>	ctx		70	1.09%

- Use Select Fields dialog to choose which fields to show.

Understanding Splunk and the Data Pipeline

Select Fields

Select All Within Filter		Deselect All	Coverage: 1% or more ▾	Filter <input type="text"/>	+	Extract New Fields
i	✓ ▾	Field ▾	# of Values	Event Coverage	Type	⋮
>	<input checked="" type="checkbox"/>	clientip	4	7.85%	String	
>	<input checked="" type="checkbox"/>	host	1	100%	String	
>	<input checked="" type="checkbox"/>	source	41	100%	String	
>	<input checked="" type="checkbox"/>	sourcetype	22	100%	String	
>	<input type="checkbox"/>	active	5	1.02%	Number	
>	<input type="checkbox"/>	average_kbps	>100	1.75%	Number	
>	<input type="checkbox"/>	avg_age	>100	14.57%	Number	
>	<input type="checkbox"/>	blocked_count	20	3.18%	Number	

- Example: Selecting clientip makes IP addresses visible.

The screenshot shows the Splunk Enterprise interface with a search titled "New Search". The search bar contains the query "index=_internal". The results table has the following columns: Time, Event, and a header row with "Format", "Show: 20 Per Page", and "View: List". The "Selected Fields" section on the left includes "clientip" (3 values). The "Interesting Fields" section includes "component", "date_hour", "date_mday", and "date_minute". The results table shows three log entries from November 25, 2025, at 14:15:04.734 PM. Each entry includes the host, index, source, and sourcetype.

Time	Event
11/25 2:15:04.734 PM	INFO PeriodicHealthReporter - feature="Scheduler Suppression" color=green due_to_stanza="feature:scheduler_suppression" node_type=feature node_path=splunkd.search.scheduler.scheduler_suppression host = splunk index = _internal source = /home/splunk/splunk/var/log/splunk/health.log sourcetype = splunkd
11/25 2:15:04.734 PM	INFO PeriodicHealthReporter - feature="Search Scheduler" color=green node_type=category node_path=splunkd.search.scheduler host = splunk index = _internal source = /home/splunk/splunk/var/log/splunk/health.log sourcetype = splunkd
11/25 2:15:04.734 PM	INFO PeriodicHealthReporter - feature="IOWait" color=yellow indicator="single_cpu_max_perc_last_3m" due_to_threshold_value=5 measured_value=5 reason="Maximum per-cpu iowait reached yellow threshold of 5" node_type=indicator node_path=splunkd.resource_usage.iowait.single_cpu_max_perc_last_3m host = splunk index = _internal source = /home/splunk/splunk/var/log/splunk/health.log sourcetype = splunkd

- We can see clientip in the Selected Field

2.5 Filtering and Analyzing Fields

Once fields are displayed, you can further refine your search or explore field values.

< Hide Fields		All Fields	Time	Event	< Prev	1	2	3	4	5	6	7	8	...	Next >
SELECTED FIELDS	<input checked="" type="checkbox"/>	clientip													
# host	<input type="checkbox"/>	host													
# index	<input type="checkbox"/>	index													
# source	<input type="checkbox"/>	source													
# sourcetype	<input type="checkbox"/>	sourcetype													
INTERESTING FIELDS															
# component	<input type="checkbox"/>	component													
# date_hour	<input type="checkbox"/>	date_hour													
# date_mday	<input type="checkbox"/>	date_mday													
# date_minute	<input type="checkbox"/>	date_minute													

- Click a field (e.g., clientip) to see:
 - Summary: Unique values and coverage.
 - Value Distribution: Counts and percentages.

Understanding Splunk and the Data Pipeline

clientip X

3 Values, 5.847% of events Selected Yes No

Reports

Top values Top values by time Rare values

Events with this field

Values	Count	%
127.0.0.1	1,047	91.043%
192.168.1.10	102	8.87%

- Use Action Links like:
 - Top values
 - Rare values
 - Events with this field
- Click a value (e.g., 192.168.1.10) to filter your search:

The screenshot shows the Splunk search interface with the following details:

- Search Bar:** Index=_internal "clientip=192.168.1.10"
- Time Range:** Last 24 hours
- Results:** 131 events (11/1/25 2:00:00.000 PM to 11/2/25 2:14:11.000 PM)
- Event Sampling:** No Event Sampling
- Panel Tabs:** Events (131), Patterns, Statistics, Visualization
- Event View:** Shows two log entries for the specified IP address.

Time	Event
11/2/25 2:14:07.358 PM	192.168.1.10 - admin [02/Nov/2025:14:14:07.358 +0000] "POST /en-US/splunkd/_raw/servicesNS/nobody/search/search/v2/jobs/1762092754.20/control HTTP/1.1" 200 59 "-" "Mozilla/5.0 (Windows NT 10.0; Win64; x64) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/142.0.0.0 Safari/537.36 Edg/142.0.0.0" - bdf5fed1b71364782a85ad7dc68c6fb2 2ms
11/2/25 2:13:51.877 PM	clientip = 192.168.1.10 host = splunk index = _internal source = /home/splunk/splunk/var/log/splunk/splunkd_ui_access.log sourcetype = splunkd_ui_access

This makes your search more precise and focused.

Enabling the Receiver on Splunk Enterprise

Splunk uses receivers to accept data from forwarders. A receiver is simply a port that listens for incoming data.

Most setups only need one receiver, but you can configure multiple for advanced environments.

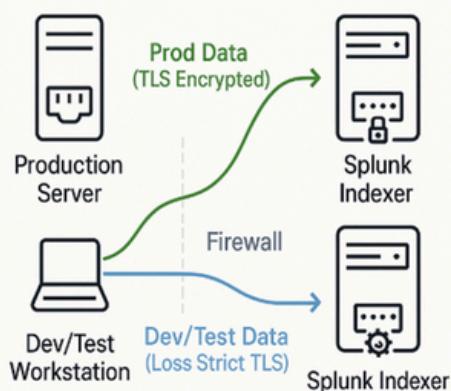
Splunk Data Ingestion: Single and Multi-Receiver Architectures

Enabling a Single Splunk Receiver

- 1 Navigate to Forwarding & Receiving Settings
- 2 Add a New Receiving Port
- 3 Configure Port (e.g.. 9997)
- 4 Verify New Port is Enabled



Advanced: Multiple Receivers for Segregation



Step-by-Step Guide

1. Navigate to Settings:

- From the Splunk Home page, click **Settings** → under **Data**, select **Forwarding and Receiving**.

A screenshot of the Splunk Home page. The top navigation bar includes "splunk>enterprise", "Apps", "Administrator", "Messages", "Settings" (which is currently selected), "Activity", "Help", and a search bar. The main content area shows the "Apps" sidebar with various icons and links like "Search & Reporting", "Audit Trail", "Splunk Secure Gateway", and "Upgrade Readiness App". To the right, a "Hello, Admin" dashboard is displayed with sections for "Bookmarks", "Explore Data", "Monitoring Console", and "Common tasks". A large central panel titled "Forwarding and Receiving" is shown, with its URL "splunk://splunk:8088/services/forwarding/receiving" visible in the address bar. The "Forwarding and Receiving" link is highlighted with a red box.

2. Add a New Receiving Port:

- On the **Forwarding and Receiving** page, find **Receive data** and click **+ Add new** next to **Configure receiving**.

The screenshot shows the 'Forwarding and receiving' page in Splunk. The 'Receive data' section is highlighted. A red box surrounds the '+ Add new' button next to the 'Configure receiving' link.

3. Configure the Receiving Port:

- Enter the port number in **Listen on this port** (commonly **9997**).
- Click **Save**.

The screenshot shows the 'Add new' configuration dialog for receiving data. The 'Listen on this port' field is set to '9997'. A red box surrounds the 'Save' button at the bottom right of the dialog.

4. Verify:

- You'll see a confirmation message like: "Successfully saved '9997'."
- The port will appear with **Status: Enabled**.

The screenshot shows the Splunk Enterprise web interface. At the top, there's a navigation bar with links for 'splunk>enterprise', 'Apps', 'Administrator', 'Messages', 'Settings', 'Activity', 'Help', 'Find', and a search icon. Below the navigation is a header for 'Receive data' with a sub-header 'Forwarding and receiving > Receive data'. A green button labeled 'New Receiving Port' is visible. A blue banner at the top of the main content area says 'Successfully saved "9997".' Below it, a message says 'Showing 1-1 of 1 item'. There's a search bar with a 'filter' placeholder and a magnifying glass icon, followed by a dropdown for '25 per page'. The main content is a table with three columns: 'Listen on this port', 'Status', and 'Actions'. The first row shows '9997' in the 'Listen on this port' column, 'Enabled | Disable' in the 'Status' column, and a 'Delete' link in the 'Actions' column. Navigation arrows for '«' and '»' are at the bottom of the table.

Listen on this port	Status	Actions
9997	Enabled Disable	Delete

Understanding Key Configuration Files

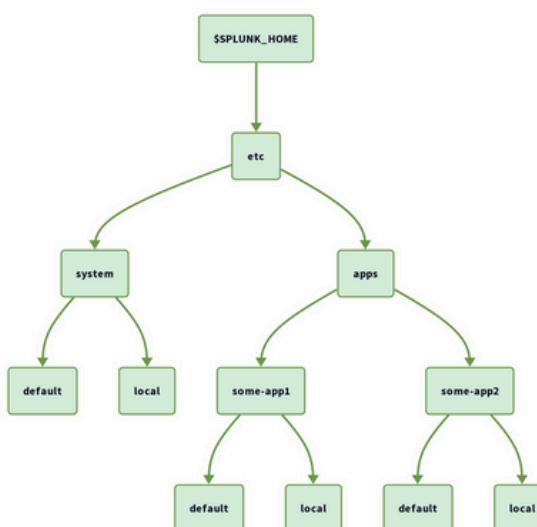
Splunk uses configuration files to control how it works. These files let you manage things like data inputs, parsing rules, server settings, and the web interface. Knowing how they are organized and which file takes priority is essential for administration.

Splunk Configuration File Overview

- Splunk settings are stored in **.conf files**, each for a specific purpose:
 - inputs.conf → Data inputs
 - props.conf → Parsing rules
 - transforms.conf → Data transformations
 - server.conf → Server settings
 - web.conf → Web interface
 - indexes.conf → Index settings
- These files are spread across the **\$SPLUNK_HOME/etc** directory, not in one place.
- Each file uses **stanzas** (sections in square brackets []) to organize settings.

Navigating Splunk Configuration Directories

- Main directory: \$SPLUNK_HOME/etc
- Inside, you'll find:
 - system/ → Global settings
 - apps/ → App-specific settings
- Each has:
 - default/ → Original settings (don't edit here)
 - local/ → Your custom changes (safe to edit)



For instance, within a Splunk installation (e.g., at /home/splunk/splunk), navigating to \$SPLUNK_HOME/etc/system/default reveals a collection of default configuration files:

```
splunk@splunk:~$ sudo su
[sudo] password for splunk:
root@splunk:/home/splunk# ls
splunk  splunk-9.4.2-e9664af3d956-linux-amd64.tgz  splunk-9.4.3-237ebbd22314-linux-amd64.tgz
root@splunk:/home/splunk# cd ./splunk
root@splunk:/home/splunk/splunk# ls
bin      include      LICENSE.txt  quarantined_files  splunk-9.4.2-e9664af3d956-linux-amd64-manifest
copyright.txt  lib      openssl      README-splunk.txt  swidtag
etc      license-eula.txt  opt      share      var
root@splunk:/home/splunk/splunk# cd etc
root@splunk:/home/splunk/splunk/etc# ls
anonymizer    init.d      log-debug.cfg      myinstall      splunk-launch.conf.default
apps          instance.cfg  login-info.cfg  openldap      splunk.version
auth          licenses      log-searchprocess.cfg  packages
copyright.txt  log-btool.cfg  log-tlsproxy.cfg  passwd      system
datETIME.xml   log-btool-debug.cfg  log-utility.cfg  prettyprint.xsl
deployment-apps log.cfg      manager-apps      shcluster
disabled-apps  log-cmdline.cfg  master-apps      splunk-enttrial.lic
findlogs.ini   log-cmdline-debug.cfg  modules      splunk-launch.conf
root@splunk:/home/splunk/splunk/etc# cd system/
root@splunk:/home/splunk/splunk/etc/system# ls
bin [default] local  lookups  metadata  README  static
root@splunk:/home/splunk/splunk/etc/system# cd default/
root@splunk:/home/splunk/splunk/etc/system/default# ls
agent_management.conf  default-mode.conf  ipc_broker.conf  savedsearches.conf  ui-tour.conf
alert_actions.conf     distsearch.conf  limits.conf    searchbnf.conf  viewstates.conf
app.conf               eventdiscoverer.conf literals.conf  segmenters.conf  visualizations.conf
audit.conf              event_renderers.conf  livetail.conf  serverclass.conf  web.conf
authentication.conf   eventtypes.conf   messages.conf  server.conf    web-features.conf
authorize.conf         federated.conf   metric_alerts.conf  source-classifier.conf  workflow_actions.conf
collections.conf       field_filters.conf metric_rollups.conf sourcetypes.conf  workload_policy.conf
commands.conf          fields.conf     multikv.conf   telemetry.conf  workload_pools.conf
conf.conf               global-banner.conf outputs.conf   times.conf    workload_rules.conf
data                  health.conf     procmon-filters.conf  transactiontypes.conf  transforms.conf
datamodels.conf         indexes.conf   props.conf     restmap.conf  ui-prefs.conf
datatypebsnf.conf      inputs.conf
```

Understanding Configuration File Precedence

When Splunk starts, it merges all files with the same name. If there's a conflict, this is the priority (highest to lowest):

1. \$SPLUNK_HOME/etc/system/local
2. \$SPLUNK_HOME/etc/apps/<app>/local
3. \$SPLUNK_HOME/etc/apps/<app>/default
4. \$SPLUNK_HOME/etc/system/default

Tip: Always make changes in local/ directories, not default.

Example: web.conf

The web.conf file controls **Splunk's web interface settings**.

```
root@splunk:/home/splunk/splunk/etc/system/default# nano web.conf
```

```
GNU nano 6.2                                         web.conf
# Version 9.4.2
# DO NOT EDIT THIS FILE!
# Changes to default files will be lost on update and are difficult to
# manage and support.
#
# Please make any changes to system defaults by overriding them in
# apps or $SPLUNK_HOME/etc/system/local
# (See "Configuration file precedence" in the web documentation).
#
# To override a specific setting, copy the name of the stanza and
# setting to the file where you wish to override it.
#
# This file contains possible attributes and values you can use to configure Splunk's web interface.
#
[default]

[settings]

# enable/disable the appserver
startwebserver = 1

# port number tag is missing or 0 the server will NOT start an http listener
# this is the port used for both SSL and non-SSL (we only have 1 port now).
httpport = 8000

# this determines whether to start SplunkWeb in http or https.
enableSplunkWebSSL = false

# location of splunkd; don't include http[s]:// in this anymore.
mgmtHostPort = 127.0.0.1:8089

# list of ports to start python application servers on (although usually
# one port is enough)
#
# In the past a special value of "0" could be passed here to disable
# the modern UI appserver infrastructure, but that is no longer supported.
appServerPorts = 8065

# default timeout, in seconds, when communicating with splunkd
splunkdConnectionTimeout = 30

# enable/disable custom netloc when using http client
enableSplunkWebClientNetloc = False

# SSL certificate files.
privKeyPath = $SPLUNK_HOME/etc/auth/splunkweb/privkey.pem
serverCert = $SPLUNK_HOME/etc/auth/splunkweb/cert.pem
```

What this means:

- startwebserver = true → Splunk Web will start.
- httpport = 8000 → The web interface runs on port 8000.
- enableSplunkWebSSL = false → SSL is disabled (not secure).

Different locations might have different values:

Location	Settings
\$SPLUNK_HOME/etc/system/local/web.conf	enableSplunkWebSSL = true
\$SPLUNK_HOME/etc/apps/some-app/local/web.conf	(none)
\$SPLUNK_HOME/etc/apps/some-app/default/web.conf	httpport = 443
\$SPLUNK_HOME/etc/system/default/web.conf	startwebserver = true httpport = 8000 enableSplunkWebSSL = false

Final Effective Settings

- startwebserver = true (from system/default, not overridden)
- httpport = 443 (from app/default, overrides default)
- enableSplunkWebSSL = true (from system/local, highest priority)

Why this matters:

- Always make changes in local/ directories, not default.
- This ensures your settings persist after upgrades and follow best practices.

Understanding and Installing Splunk Apps

Splunk apps are packages that extend Splunk's functionality. They can add dashboards, reports, integrations, and more.

1. What Are Splunk Apps?

- **Self-contained packages** (e.g., .tgz, .tar.gz) that enhance Splunk Enterprise or Universal Forwarder.
- **What they do:**
 - **Enhance Data Processing:** Add field extractions and parsing rules.
 - **Provide Visualizations:** Dashboards, charts, and reports.
 - **Enable Integrations:** Connect Splunk with external tools.
 - **Streamline Operations:** Pre-built alerts and dashboards.

App Directory Structure

- bin/ → Scripts and custom code.
- default/ → Default configs (don't edit here).
- local/ → Your custom changes (safe to edit).
- lookups/ → Lookup tables for enrichment.
- metadata/ → Permissions and metadata.
- static/ → Icons and images.

2. Installing Splunk Apps

Two main methods:

2.1 Installation via Splunk Web Interface

Steps:

1. **Access App Management:** Go to **Apps** from the Splunk homepage.

The screenshot shows the Splunk Web interface with the 'Apps' tab selected in the top navigation bar. The main content area is titled 'Hello, Administrator'. It features a 'Bookmarks' section with two collapsed categories: 'My bookmarks (0)' and 'Shared with my organization (0)'. Below these are sections for 'Recently viewed', 'Created by you', and 'Shared with you'. On the left side, there is a sidebar with a search bar and links to other Splunk apps like 'Search & Reporting', 'Audit Trail', 'Splunk Secure Gateway', and 'Upgrade Readiness App'.

Understanding Splunk and the Data Pipeline

2. Choose Installation Method:

- Browse More Apps (Splunkbase): Search and install directly.
- Install from File: Upload a .tgz file.

The screenshot shows the Splunk Enterprise dashboard. On the left, there's a sidebar with various links like Home, Search & Reporting, Audit Trail, etc. A red box highlights the 'Find More Apps' button under the 'Search apps' section. The main area shows a 'Hello, Administrator' greeting and a 'Bookmarks' section with 'My bookmarks (0)' and 'Add bookmark' button. Below it are sections for 'Shared with my organization (0)', 'Shared by me', and 'Shared by other administrators'.

This screenshot shows the 'Browse More Apps' page. At the top, there's a search bar with 'Splunk Add-on for Microsoft Windows' and a green 'Install' button highlighted with a red box. To the left is a 'CATEGORY' filter sidebar with checkboxes for IT Operations, Security, Fraud & Compliance, Business Analytics, Utilities, Artificial Intelligence, IoT & Industrial Data, DevOps, Directory Service, Email, Endpoint, Firewall, Generic, Identity Management, Information, Investigative, and Network Access Control. The main list shows several add-ons: 'Splunk Add-on for Microsoft Windows' (selected), 'Microsoft Defender for Identity Add-on for Splunk', 'Splunk Add-on for Microsoft SQL Server', and 'Splunk Add-on for Microsoft Hyper-V'. Each item has a brief description, category, author, download count, release date, and last update.

Login and Install

Enter your Splunk.com username and password to download the app.

[Forgot your password?](#)

The app, and any related dependency that will be installed, may be provided by Splunk and/or a third party and your right to use these app(s) is in accordance with the applicable license(s) provided by Splunk and/or the third-party licensor. Splunk is not responsible for any third-party app (developed by you or a third party) and does not provide any warranty or support. Installation of a third-party app can introduce security risks. By clicking "Agree" below, you acknowledge and accept such risks. If you have any questions, complaints or claims with respect to an app, please contact the applicable licensor directly whose contact information can be found on the Splunkbase download page.

Splunk Add-on for Microsoft Windows is governed by the following license: sgt

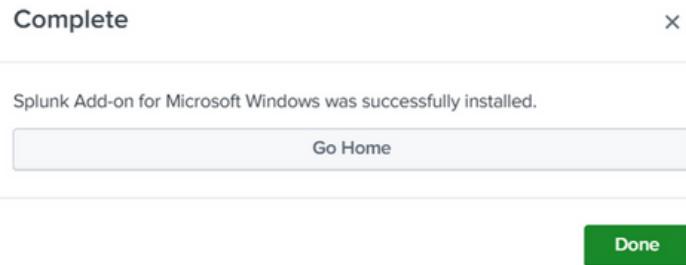
I have read the terms and conditions of the license(s) and agree to be bound by them. I also agree to Splunk's [Website Terms of Use](#).

[Cancel](#)

[Agree and Install](#)

3. Upload and Install: Splunk processes the package.

4. Completion: Confirmation message appears. Some apps may require a restart.



2.2 Verify Installation

- On the server:

```
splunk@splunk:~/splunk/etc/apps$ cd ./splunk/etc/apps
```

- Use ls -al to list installed apps.

```
splunk@splunk:~/splunk/etc/apps$ ls -al
total 124
drwxr-xr-x 31 splunk splunk 4096 Jul  6 20:29 .
drwxr-xr-x 18 splunk splunk 4096 Jul  6 19:42 ..
drwxr-xr-x  7 splunk splunk 4096 Mar 18 22:42 alert_logevent
drwxr-xr-x  7 splunk splunk 4096 Mar 18 22:42 alert_webhook
drwxr-xr-x  4 splunk splunk 4096 Mar 18 22:42 appsbrowser
drwxr-xr-x  6 splunk splunk 4096 Mar 18 22:42 audit_trail
drwxr-xr-x  5 splunk splunk 4096 Mar 18 22:42 introspection_generator_addon
drwxr-xr-x  4 splunk splunk 4096 Mar 18 22:42 journald_input
drwxr-xr-x  7 splunk splunk 4096 Mar 18 22:42 launcher
drwxr-xr-x  5 splunk splunk 4096 Jun 28 13:49 learned
drwxr-xr-x  4 splunk splunk 4096 Mar 18 22:42 legacy
drwxr-xr-x  9 splunk splunk 4096 Jun 28 20:12 python_upgrade_readiness_app
drwxr-xr-x  6 splunk splunk 4096 Mar 18 22:42 sample_app
drwxr-xr-x  8 splunk splunk 4096 Mar 18 22:42 search
drwxr-xr-x  6 splunk splunk 4096 Jun 28 19:41 splunk_archiver
drwxr-xr-x  9 splunk splunk 4096 Jun 28 13:49 splunk-dashboard-studio
drwxr-xr-x  4 splunk splunk 4096 Mar 18 22:42 SplunkDeploymentServerConfig
drwxr-xr-x  4 splunk splunk 4096 Mar 18 22:42 SplunkForwarder
drwxr-xr-x  7 splunk splunk 4096 Mar 18 22:45 splunk_gdi
drwxr-xr-x  3 splunk splunk 4096 Mar 18 22:42 splunk_httpinput
drwxr-xr-x  9 splunk splunk 4096 Jun 28 13:49 splunk_instrumentation
drwxr-xr-x  4 splunk splunk 4096 Mar 18 22:42 splunk_internal_metrics
drwxr-xr-x  4 splunk splunk 4096 Mar 18 22:42 SplunkLightForwarder
drwxr-xr-x  8 splunk splunk 4096 Mar 18 22:44 splunk_metrics_workspace
drwxr-xr-x  8 splunk splunk 4096 Jun 28 13:49 splunk_monitoring_console
drwxr-xr-x 10 splunk splunk 4096 Jul  5 16:43 splunk_rapid_diag
drwxr-xr-x  8 splunk splunk 4096 Jul 19  2024 splunk-rolling-upgrade
drwxr-xr-x 10 splunk splunk 4096 Feb 18 16:04 splunk_secure_gateway
drwxr-xr-x 10 splunk splunk 4096 Jul  6 20:29 Splunk_TA_windows
drwxr-xr-x  5 splunk splunk 4096 Nov 12  2024 splunk-visual-exporter
drwxr-xr-x  4 splunk splunk 4096 Mar 18 22:42 user-prefs
```

- Explore default/ for original configs and local/ for custom changes.

```
splunk@splunk:~/splunk/etc/apps$ cd Splunk_TA_windows/default/
splunk@splunk:~/splunk/etc/apps/Splunk_TA_windows/default$ ls -al
total 252
drwxr-xr-x  3 splunk  splunk   4096 Jul  6 20:29 .
drwxr-xr-x 10 splunk  splunk   4096 Jul  6 20:29 ..
-rw-----  1 splunk  splunk    558 Jul  6 20:29 app.conf
drwxr-xr-x  3 splunk  splunk   4096 Jul  6 20:29 data
-rw-r--r--  1 splunk  splunk  26329 Jul  6 20:29 eventtypes.conf
-rw-r--r--  1 splunk  splunk  20145 Jul  6 20:29 inputs.conf
-rw-r--r--  1 splunk  splunk  2407 Jul  6 20:29 macros.conf
-rw-r--r--  1 splunk  splunk 119677 Jul  6 20:29 props.conf
-rw-r--r--  1 splunk  splunk 12264 Jul  6 20:29 tags.conf
-rw-r--r--  1 splunk  splunk 40173 Jul  6 20:29 transforms.conf
-rw-r--r--  1 splunk  splunk  4450 Jul  6 20:29 wmi.conf
-rw-r--r--  1 splunk  splunk  1589 Jul  6 20:29 workflow_actions.conf
```

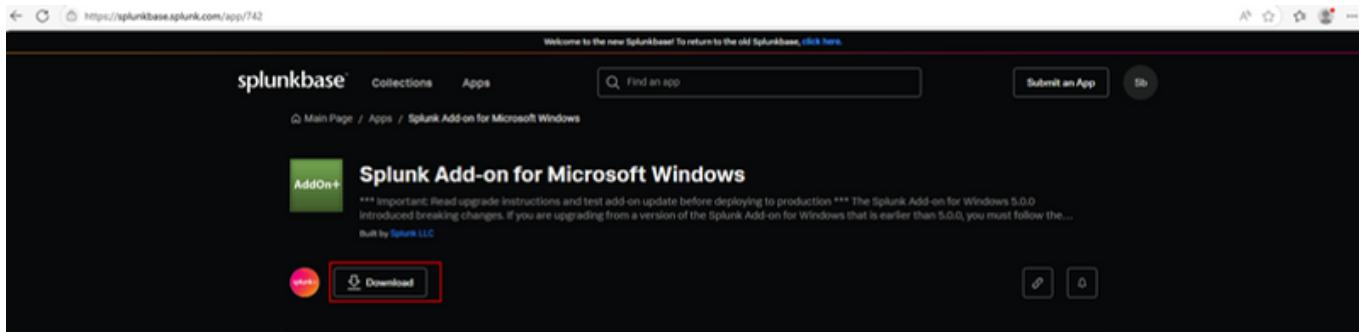
Best Practice: Always edit in local/, never in default/.

Shipping Windows Event Logs into Splunk

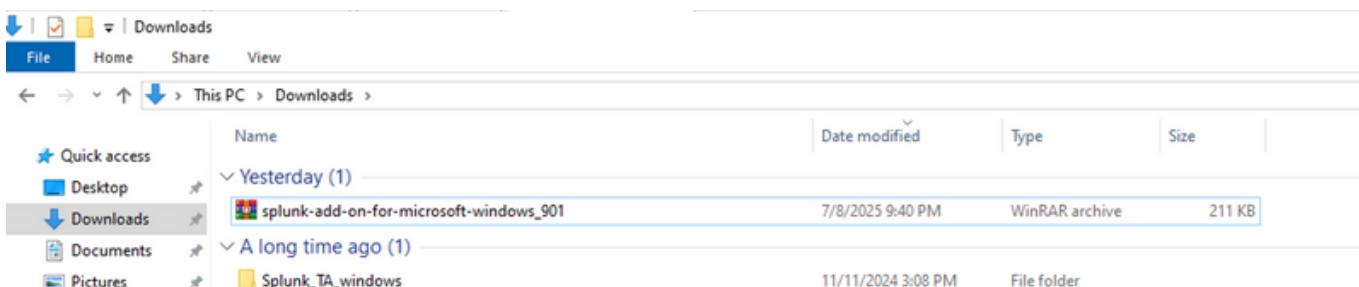
To collect Windows event logs and send them to Splunk, you use a Splunk Universal Forwarder with the Splunk Add-on for Microsoft Windows.

Step 1: Download the Add-on

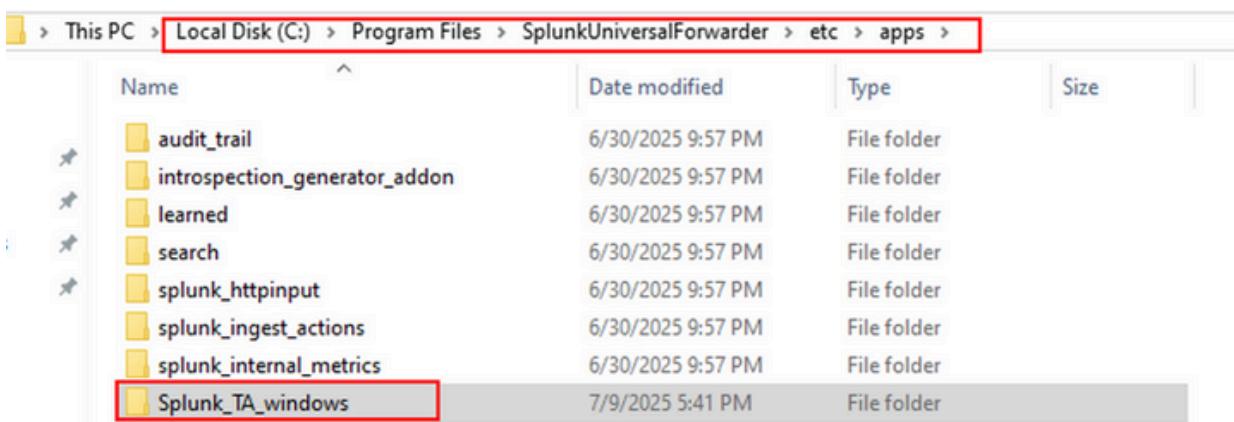
- Go to Splunkbase (<https://splunkbase.splunk.com/app/742>) and download **Splunk Add-on for Microsoft Windows**.



- Extract it to:



- Make sure the folder name is:



- Restart

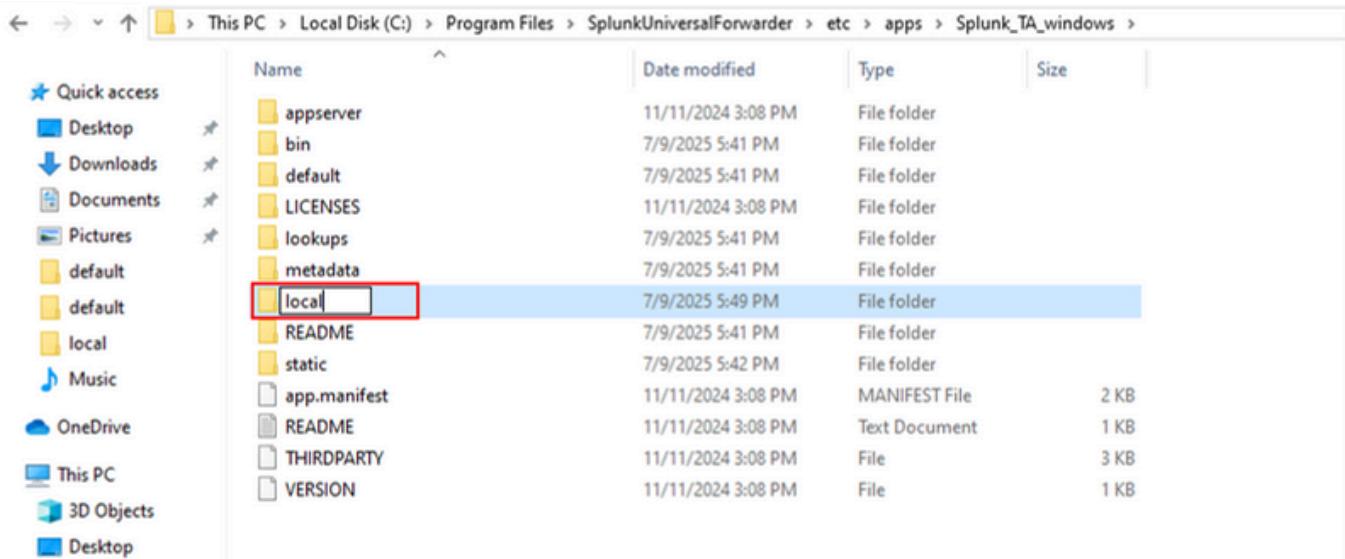
```
Microsoft Windows [Version 10.0.19045.5965]
(c) Microsoft Corporation. All rights reserved.

C:\Windows\system32:cd "\Program Files\SplunkUniversalForwarder\bin"
C:\Program Files\SplunkUniversalForwarder\bin\splunk restart
```

Step 2: Configure inputs.conf

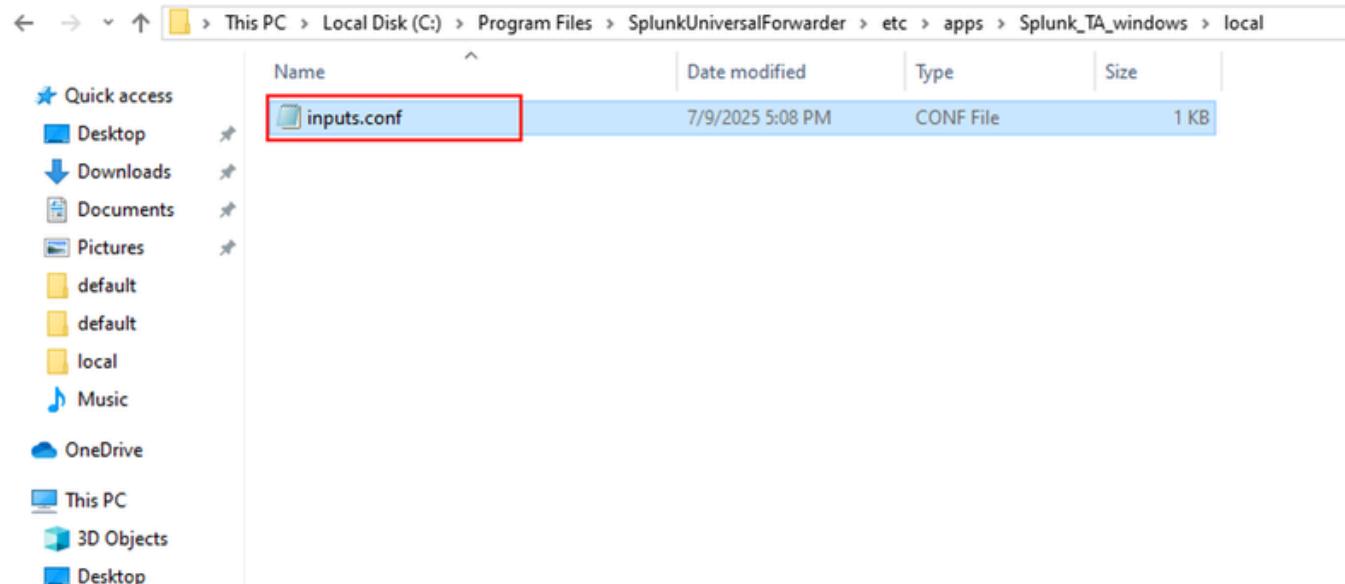
- Path:

C:\Program Files\SplunkUniversalForwarder\etc\apps\Splunk_TA_windows\local\



	Name	Date modified	Type	Size
Quick access	appserver	11/11/2024 3:08 PM	File folder	
Desktop	bin	7/9/2025 5:41 PM	File folder	
Downloads	default	7/9/2025 5:41 PM	File folder	
Documents	LICENSES	11/11/2024 3:08 PM	File folder	
Pictures	lookups	7/9/2025 5:41 PM	File folder	
default	metadata	7/9/2025 5:41 PM	File folder	
default	local	7/9/2025 5:49 PM	File folder	
local	README	7/9/2025 5:41 PM	File folder	
Music	static	7/9/2025 5:42 PM	File folder	
OneDrive	app.manifest	11/11/2024 3:08 PM	MANIFEST File	2 KB
This PC	README	11/11/2024 3:08 PM	Text Document	1 KB
3D Objects	THIRDPARTY	11/11/2024 3:08 PM	File	3 KB
Desktop	VERSION	11/11/2024 3:08 PM	File	1 KB

- Open inputs.conf and add:



	Name	Date modified	Type	Size
Quick access	inputs.conf	7/9/2025 5:08 PM	CONF File	1 KB
Desktop				
Downloads				
Documents				
Pictures				
default				
default				
local				
Music				
OneDrive				
This PC				
3D Objects				
Desktop				

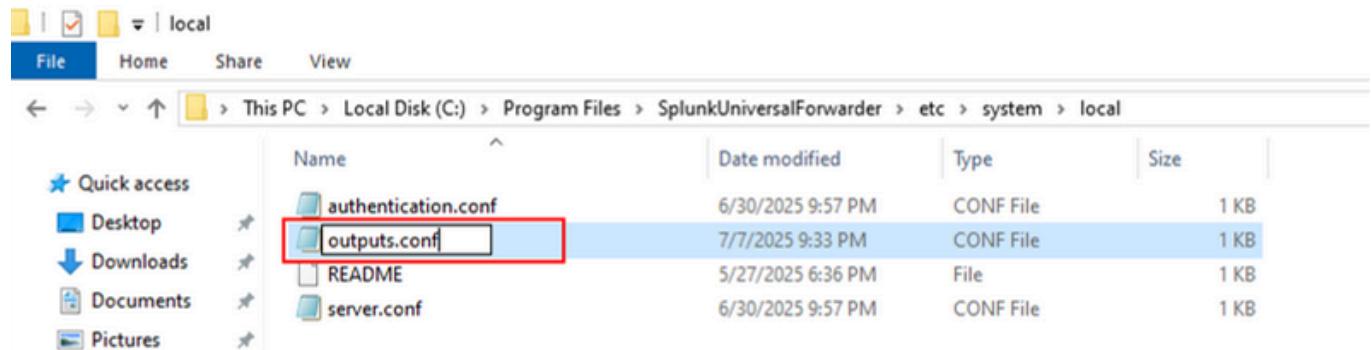


```
[WinEventLog://Security]
disabled = 0
```

- This enables collection of Windows Security event logs.

Step 3: Configure outputs.conf

- Path:
C:\Program Files\SplunkUniversalForwarder\etc\system\local\



- Open outputs.conf and add:

```
[tcpout]
defaultGroup = default-autolb-group

[tcpout:default-autolb-group]
server = 192.168.1.13:9997

[tcpout-server://192.168.1.13:9997]
```

- Replace 192.168.1.13 with your Splunk indexer IP and 9997 with the receiving port.

Step 4: Restart the Forwarder

- Open Command Prompt (Admin):

```
Administrator: Command Prompt

Starting splunk server daemon (splunkd)...
splunkForwarder: Unable to start the service: The specified service does not exist as an installed service.

C:\Program Files\SplunkUniversalForwarder\bin>splunk restart
splunkForwarder: Unable to stop the service: The specified service does not exist as an installed service.
```

Step 5: Verify in Splunk

- In Splunk Web, search:

Time	Event
7/10/25 4:29:45.000 PM	<pre><event xmlns="http://schemas.microsoft.com/win/2004/08/events/event"><System><Provider Name="Microsoft-Windows-Security-Auditing" Guid="54849625-5478-4934-a5ba-3e3b8328c3b0">/><EventID>4672</EventID><Version>0</Version><Level>4</Level><Task>12548</Task><Opcode>0</Opcode><Keywords>0x8200000000000000</Keywords><TimeCreated SystemTime="2025-07-10T14:29:45.0000000Z" /><EventRecordID>9642</EventRecordID><Correlation ActivityID="7263f239-f1b6-0001-4f72-42726ff1dbf1" /><Execution ProcessID="652" ThreadID="728" /><Channel>Security</Channel><Computer>DESKTOP-SR7SNFG</Computer><Security><System><EventID>4672</EventID><Data Name="SubjectUserName">SplunkForwarder</Data><Data Name="SubjectDomainName">INT SERVICE</Data><Data Name="SubjectLogonId" /><Data Name="LogonType">5</Data><Data Name="LogonProcessName">Advapi</Data><Data Name="AuthenticationPackageName">Negotiate</Data><Data Name="WorkstationName"></Data><Data Name="LogonGuid" />(00000000-0000-0000-0000-000000000000)</Data><Data Name="TransitedServices"></Data><Data Name="LogonPackageName" /></Data><Data Name="KeyLength" /><Data Name="ProcessId" />4278</Data><Data Name="ProcessName" >C:\Windows\System32</Data></System></EventRecord><Correlation ActivityID="7263f239-f1b6-0001-4f72-42726ff1dbf1" /><Execution ProcessID="652" ThreadID="728" /><Channel>Security</Channel><Computer>DESKTOP-SR7SNFG</Computer><Security><System><EventID>4672</EventID><Data Name="SubjectUserName">AUTORITE NTSystem</Data><Data Name="SubjectDomainName">WORKGROUP</Data><Data Name="SubjectLogonId" /><Data Name="TargetUserName">SplunkForwarder</Data><Data Name="TargetDomainName">INT SERVICE</Data><Data Name="TargetLogonId" />(0x124840)</Data><Data Name="LogonType" />5</Data><Data Name="LogonProcessName" />Advapi</Data><Data Name="AuthenticationPackageName" />Negotiate</Data><Data Name="WorkstationName" /></Data><Data Name="LogonGuid" />(00000000-0000-0000-0000-000000000000)</Data><Data Name="TransitedServices" /></Data><Data Name="LogonPackageName" /></Data><Data Name="KeyLength" /></Data><Data Name="ProcessId" />4278</Data><Data Name="ProcessName" >C:\Windows\System32</Data></System></EventRecord></pre>
7/10/25 4:29:45.000 PM	<pre><event xmlns="http://schemas.microsoft.com/win/2004/08/events/event"><System><Provider Name="Microsoft-Windows-Security-Auditing" Guid="54849625-5478-4934-a5ba-3e3b8328c3b0">/><EventID>4624</EventID><Version>2</Version><Level>4</Level><Task>12548</Task><Opcode>0</Opcode><Keywords>0x8200000000000000</Keywords><TimeCreated SystemTime="2025-07-10T14:29:45.0000000Z" /><EventRecordID>9642</EventRecordID><Correlation ActivityID="7263f239-f1b6-0001-4f72-42726ff1dbf1" /><Execution ProcessID="652" ThreadID="728" /><Channel>Security</Channel><Computer>DESKTOP-SR7SNFG</Computer><Security><System><EventID>4624</EventID><Data Name="SubjectUserName">AUTORITE NTSystem</Data><Data Name="SubjectDomainName">WORKGROUP</Data><Data Name="SubjectLogonId" /><Data Name="TargetUserName">SplunkForwarder</Data><Data Name="TargetDomainName">INT SERVICE</Data><Data Name="TargetLogonId" />(0x124840)</Data><Data Name="LogonType" />5</Data><Data Name="LogonProcessName" />Advapi</Data><Data Name="AuthenticationPackageName" />Negotiate</Data><Data Name="WorkstationName" /></Data><Data Name="LogonGuid" />(00000000-0000-0000-0000-000000000000)</Data><Data Name="TransitedServices" /></Data><Data Name="LogonPackageName" /></Data><Data Name="KeyLength" /></Data><Data Name="ProcessId" />4278</Data><Data Name="ProcessName" >C:\Windows\System32</Data></System></EventRecord><Correlation ActivityID="7263f239-f1b6-0001-4f72-42726ff1dbf1" /><Execution ProcessID="652" ThreadID="728" /><Channel>Security</Channel><Computer>DESKTOP-SR7SNFG</Computer><Security><System><EventID>4624</EventID><Data Name="SubjectUserName">TARGET</Data><Data Name="SubjectDomainName">INT SERVICE</Data><Data Name="SubjectLogonId" /><Data Name="TargetUserName" /><Data Name="TargetDomainName" />INT SERVICE</Data><Data Name="TargetLogonId" />(0x124840)</Data><Data Name="LogonType" />5</Data><Data Name="LogonProcessName" />Advapi</Data><Data Name="AuthenticationPackageName" />Negotiate</Data><Data Name="WorkstationName" /></Data><Data Name="LogonGuid" />(00000000-0000-0000-0000-000000000000)</Data><Data Name="TransitedServices" /></Data><Data Name="LogonPackageName" /></Data><Data Name="KeyLength" /></Data><Data Name="ProcessId" />4278</Data><Data Name="ProcessName" >C:\Windows\System32</Data></System></EventRecord></pre>

If events appear, your setup is successful.

Establishing a Dedicated Index for Testing Data

When working with custom data or lab environments, it's best to create a dedicated index. This keeps test data separate from production and makes cleanup easier.

1. Creating a New Splunk App

To better organize your custom configurations, it's recommended to create a dedicated Splunk app.

1. Go to **Manage**: In Splunk Web (browser interface), click Manage.

The screenshot shows the Splunk Web interface with the title 'Hello, Administrator'. On the left, there is a sidebar with a 'Find more apps' search bar and a 'Manage' button, which is highlighted with a red box. Below the search bar, there are several app icons: 'Search & Reporting', 'Audit Trail', 'Splunk Secure Gateway', and 'Upgrade Readiness App'. The main content area shows sections for 'My bookmarks (0)', 'Shared with my organization (0)', 'Shared by me', and 'Shared by other administrators'.

2. Click "Create App".

The screenshot shows the Splunk Web 'Apps' page. At the top, there is a green 'Create app' button highlighted with a red box. The main area displays a table of 29 items, each with columns for Name, Folder name, Version, Update checking, Visible, Sharing, Status, and Actions. The actions column includes links like 'Edit properties' and 'View objects'.

Name	Folder name	Version	Update checking	Visible	Sharing	Status	Actions
SplunkDeploymentServerConfig	SplunkDeploymentServerConfig		Yes	No	App Permissions	Enabled Disable	Edit properties View objects
SplunkForwarder	SplunkForwarder		Yes	No	App Permissions	Disabled Enable	
SplunkLightForwarder	SplunkLightForwarder		Yes	No	App Permissions	Disabled Enable	
Splunk Add-on for Microsoft Windows	Splunk_TA_windows	9.0.1	Yes	No	Global Permissions	Enabled Disable	Edit properties View objects View details on Splunkbase
Log Event Alert Action	alert_logevent	9.4.2	Yes	No	App Permissions	Enabled	Edit properties View objects
Webhook Alert Action	alert_webhook	9.4.2	Yes	No	App Permissions	Enabled	Edit properties View objects
Apps Browser	appsbrowser	9.4.2	Yes	No	App Permissions	Enabled	Edit properties View objects
Audit Trail	audit_trail	1.0.0	Yes	Yes	App Permissions	Enabled	Launch app Edit properties View objects
introspection_generator_addon	introspection_generator_addon	9.4.2	Yes	No	App Permissions	Enabled Disable	Edit properties View objects
journald_input	journald_input		Yes	No	App Permissions	Enabled Disable	Edit properties View objects
Home	launcher		Yes	Yes	App Permissions	Enabled	Launch app Edit properties View objects
learned	learned		Yes	No	App Permissions	Enabled Disable	Edit properties View objects
legacy	legacy		Yes	No	App Permissions	Disabled Enable	
Upgrade Readiness App	python_upgrade_readiness_app	4.6.2	Yes	Yes	App Permissions	Enabled Disable	Launch app Edit properties View objects View details on Splunkbase
sample data	sample_app		Yes	No	App Permissions	Disabled Enable	

3. Fill the form like this:

- Folder name: indexes
- Version: 1.0.0
- Visible: No (because it has no views/pages)
- Author: Thomas Fellinger
- Description: Our own indexes to test
- Template: barebones
- Upload asset: Leave it empty

Getting External Data into Splunk

Name Give your app a friendly name for display in Splunk Web.

Folder name * This name maps to the app's directory in \$SPLUNK_HOME/etc/apps/.

Version App version.

Visible No Yes Only apps with views should be made visible.

Author Name of the app's owner.

Description Enter a description for your app.

Template These templates contain example views and searches.

Upload asset No file chosen Can be any html, js, or other file to add to your app.



4. Click **Save** to finish creating the app.

Apps								Actions		
Name	Folder name	Version	Update checking	Visible	Sharing	Status	Actions	Browse more apps	Install app from file	Create app
indexes	indexes	1.0.0	Yes	No	App Permissions	Enabled Disable	Edit properties View objects	25 per page		

2. Check the App Configuration

After creating the app, you can check its settings using the command line (terminal).

Steps:

1. Go to the app folder on your server.

```
splunk@splunk:~/splunk/etc/apps$ ls -al | grep indexes
drwx--x--- 6 splunk splunk 4096 Jul  8 19:42 indexes
splunk@splunk:~/splunk/etc/apps$ cd indexes/
splunk@splunk:~/splunk/etc/apps/indexes$ ls -al
total 24
drwx--x--- 6 splunk splunk 4096 Jul  8 19:42 .
drwxr-xr-x 33 splunk splunk 4096 Jul  8 19:42 ..
drwx--x--- 2 splunk splunk 4096 Jul  8 19:42 bin
drwx--x--- 3 splunk splunk 4096 Jul  8 19:42 default
drwx----- 2 splunk splunk 4096 Jul  8 19:42 local
drwx--x--- 2 splunk splunk 4096 Jul  8 19:42 metadata
splunk@splunk:~/splunk/etc/apps/indexes$ cd default/
splunk@splunk:~/splunk/etc/apps/indexes/default$ ls -al
total 16
drwx--x--- 3 splunk splunk 4096 Jul  8 19:42 .
drwx--x--- 6 splunk splunk 4096 Jul  8 19:42 ..
-rw----- 1 splunk splunk 193 Jul  8 19:42 app.conf
drwx--x--- 3 splunk splunk 4096 Jul  8 19:42 data
```

2. Open **app.conf** file, this file shows the details you entered when creating the app.

```
splunk@splunk:~/splunk/etc/apps/indexes/default$ cat app.conf
#
# Splunk app configuration file
#
[install]
is_configured = 0

[ui]
is_visible = 0
label = indexes

[launcher]
author = Salma Benmina
description = Our own indexes to test
version = 1.0.0
```

3. Create a New Index (Using Configuration File)

An index is where Splunk stores data.

Steps:

1. Inside your app folder, go to the default directory.
2. Create or edit a file called indexes.conf.

```
splunk@splunk:~/splunk/etc/apps/indexes/default$ nano indexes.conf
```

3. Add this configuration to define a new index called test:

```
GNU nano 6.2                                         indexes.conf *
[test]
homePath = $SPLUNK_DB/${_index_name}/db
coldPath = $SPLUNK_DB/${_index_name}/colddb
thawedPath = $SPLUNK_DB/test/thaweddb
```

- **[test]**: Defines a new index named test.
- **homePath**: Where hot/warm data is stored.
- **coldPath**: Where cold data is stored.
- **thawedPath**: Where restored (thawed) data is stored.

4. Restart Splunk

To apply the new index settings, you need to restart Splunk.

```
splunk@splunk:~/splunk/bin$ ./splunk restart
Stopping splunkd...
Shutting down. Please wait, as this may take a few minutes.
==== AUTHENTICATING FOR org.freedesktop.systemd1.manage-units ====
Authentication is required to stop 'Splunkd.service'.
Authenticating as: splunk
Password:
==== AUTHENTICATION COMPLETE ===

Stopping splunk helpers...

Done.
splunkd.pid doesn't exist...

Splunk> Winning the War on Error

Checking prerequisites...
    Checking http port [8000]: open
    Checking mgmt port [8089]: open
    Checking appserver port [127.0.0.1:8065]: open
    Checking kvstore port [8191]: open
    Checking configuration... Done.
    Checking critical directories...      Done
    Checking indexes...
        Validated: _audit _configtracker _dsappevent _dsclient _dsphonehome _internal _introspection _metrics _metrics_rollup _telemetry _thefishbucket history main summary test
        Done
    Checking filesystem compatibility...  Done
    Checking conf files for problems...
|
```

- You'll be asked for your password to confirm the restart.

5. Check the New Index in Splunk Web

After restart:

1. Go to **Settings > Indexes** in Splunk Web.

Getting External Data into Splunk

The screenshot shows the Splunk web interface with the 'Settings' menu open. The 'Indexes' option under the 'DATA' section is highlighted with a red box and a red arrow pointing to it. Other options visible in the 'DATA' section include 'Data inputs', 'Forwarding and receiving', 'Report acceleration summaries', 'Virtual indexes', 'Source types', and 'Ingest actions'. The 'Indexes' section also contains a sub-section for 'DISTRIBUTED ENVIRONMENT' with options like 'Forwarder management', 'Indexer clustering', and 'Federation'. The 'SYSTEM' section includes 'Server settings', 'Server controls', 'Health report manager', 'RapidDiag', 'Instrumentation', 'Licensing', 'Workload management', and 'Mobile settings'. The 'USERS AND AUTHENTICATION' section includes 'Roles', 'Users', 'Tokens', 'Password management', and 'Authentication methods'.

2. You should see the new index called test listed there.

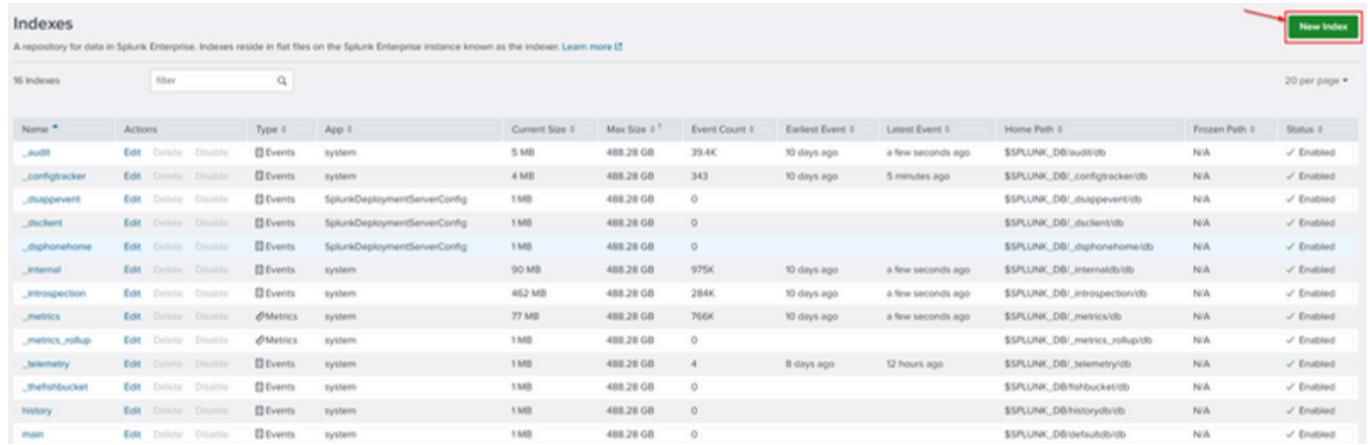
Indexes													New Index
A repository for data in Splunk Enterprise. Indexes reside in flat files on the Splunk Enterprise instance known as the Indexer. Learn more [?]													
Name	Actions	Type	App	Current Size	Max Size	Event Count	Earliest Event	Latest Event	Home Path	Frozen Path	Status		
_audit	Edit Delete Disable	Events	system	5 MB	488.28 GB	39.4K	10 days ago	a few seconds ago	\$SPLUNK_DB/audit/db	N/A	✓ Enabled		
_configtracker	Edit Delete Disable	Events	system	4 MB	488.28 GB	343	10 days ago	5 minutes ago	\$SPLUNK_DB/_configtracker/db	N/A	✓ Enabled		
_disappevent	Edit Delete Disable	Events	SplunkDeploymentServerConfig	1MB	488.28 GB	0			\$SPLUNK_DB/_disappevent/db	N/A	✓ Enabled		
_dsclient	Edit Delete Disable	Events	SplunkDeploymentServerConfig	1MB	488.28 GB	0			\$SPLUNK_DB/_dsclient/db	N/A	✓ Enabled		
_dphonenum	Edit Delete Disable	Events	SplunkDeploymentServerConfig	1MB	488.28 GB	0			\$SPLUNK_DB/_dphonenum/db	N/A	✓ Enabled		
_internal	Edit Delete Disable	Events	system	90 MB	488.28 GB	975K	10 days ago	a few seconds ago	\$SPLUNK_DB/_internal/db	N/A	✓ Enabled		
_introspection	Edit Delete Disable	Events	system	462 MB	488.28 GB	284K	10 days ago	a few seconds ago	\$SPLUNK_DB/_introspection/db	N/A	✓ Enabled		
_metrics	Edit Delete Disable	Metrics	system	77 MB	488.28 GB	766K	10 days ago	a few seconds ago	\$SPLUNK_DB/_metrics/db	N/A	✓ Enabled		
_metrics_rollup	Edit Delete Disable	Metrics	system	1MB	488.28 GB	0			\$SPLUNK_DB/_metrics_rollup/db	N/A	✓ Enabled		
_telemetry	Edit Delete Disable	Events	system	1MB	488.28 GB	4	8 days ago	12 hours ago	\$SPLUNK_DB/_telemetry/db	N/A	✓ Enabled		
_tefshbucket	Edit Delete Disable	Events	system	1MB	488.28 GB	0			\$SPLUNK_DB/_tefshbucket/db	N/A	✓ Enabled		
history	Edit Delete Disable	Events	system	1MB	488.28 GB	0			\$SPLUNK_DB/history/db	N/A	✓ Enabled		
main	Edit Delete Disable	Events	system	1MB	488.28 GB	0			\$SPLUNK_DB/main/db	N/A	✓ Enabled		
splunklogger	Edit Delete Disable	Events	system	0 B	488.28 GB	0			\$SPLUNK_DB/splunklogger/db	N/A	✗ Disabled		
summary	Edit Delete Disable	Events	system	1MB	488.28 GB	0			\$SPLUNK_DB/summary/db	N/A	✓ Enabled		
test	Edit Delete Disable	Events	indexes	1MB	488.28 GB	0			\$SPLUNK_DB/\$.index_name/db	N/A	✓ Enabled		

6. Create a New Index Manually (Alternative Method)

You can also create an index directly from the Splunk Web interface — no need to edit files.

Steps:

1. Go to **Settings > Indexes** in Splunk Web.
2. Click "**New Index**".

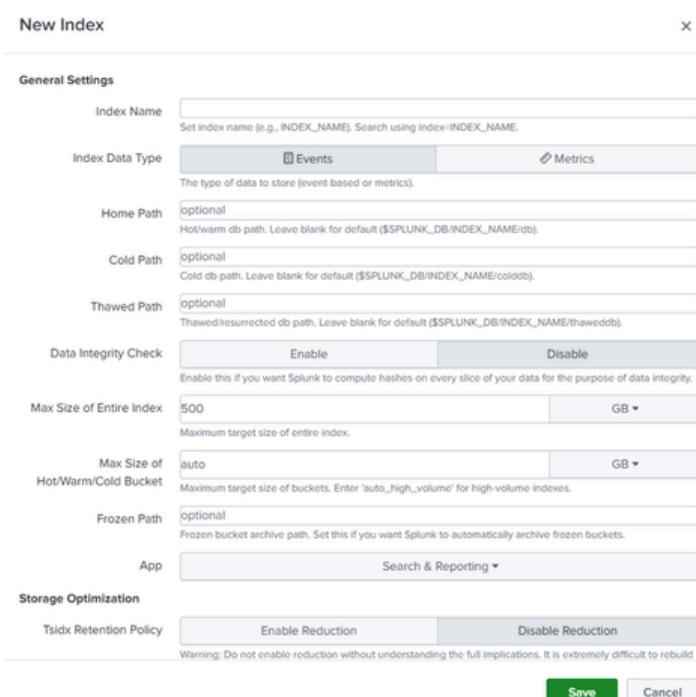


The screenshot shows the 'Indexes' page in Splunk Web. At the top, there's a header with the title 'Indexes'. Below it, a sub-header says 'A repository for data in Splunk Enterprise. Indexes reside in flat files on the Splunk Enterprise instance known as the indexer. Learn more [\[link\]](#)'. There are 16 indexes listed in a table. The columns include: Name, Actions, Type, App, Current Size, Max Size, Event Count, Earliest Event, Latest Event, Home Path, Frozen Path, and Status. Most indexes have a status of 'Enabled'. The last row, 'main', has a status of 'Disabled'. In the top right corner of the table area, there is a green button labeled 'New Index' with a red arrow pointing to it.

3. Fill in the form:

- **Index Name:** Type the name you want (e.g., test)
- **Index Data Type:** Choose Events or Metrics
- **Home Path / Cold Path / Thawed Path:** You can leave them empty to use default paths, or set custom ones.
- **Data Integrity Check:** Turn it on or off depending on your needs.

4. Click **Save** to create the index.



The screenshot shows the 'New Index' dialog box. The 'General Settings' tab is active. It contains the following fields:

- Index Name:** A text input field with placeholder text 'Set index name (e.g., INDEX_NAME). Search using index:INDEX_NAME.'
- Index Data Type:** A radio button group where 'Events' is selected, with a note below: 'The type of data to store (event based or metrics)'.
- Home Path:** An optional text input field with placeholder text 'Hot/warm db path. Leave blank for default (\$SPLUNK_DB/INDEX_NAME/db)'.
- Cold Path:** An optional text input field with placeholder text 'Cold db path. Leave blank for default (\$SPLUNK_DB/INDEX_NAME/colddb)'.
- Thawed Path:** An optional text input field with placeholder text 'Thawed/resurrected db path. Leave blank for default (\$SPLUNK_DB/INDEX_NAME/thaweddb)'.
- Data Integrity Check:** A radio button group where 'Enable' is selected, with a note below: 'Enable this if you want Splunk to compute hashes on every slice of your data for the purpose of data integrity.'
- Max Size of Entire Index:** A text input field with '500' and a dropdown menu with 'GB' selected.
- Max Size of Hot/Warm/Cold Bucket:** A text input field with 'auto' and a dropdown menu with 'GB' selected.
- Frozen Path:** An optional text input field with placeholder text 'Frozen bucket archive path. Set this if you want Splunk to automatically archive frozen buckets.'
- App:** A dropdown menu currently set to 'Search & Reporting'.

 At the bottom, there are 'Save' and 'Cancel' buttons. A warning message above the 'Save' button says: 'Warning: Do not enable reduction without understanding the full implications. It is extremely difficult to rebuild'.

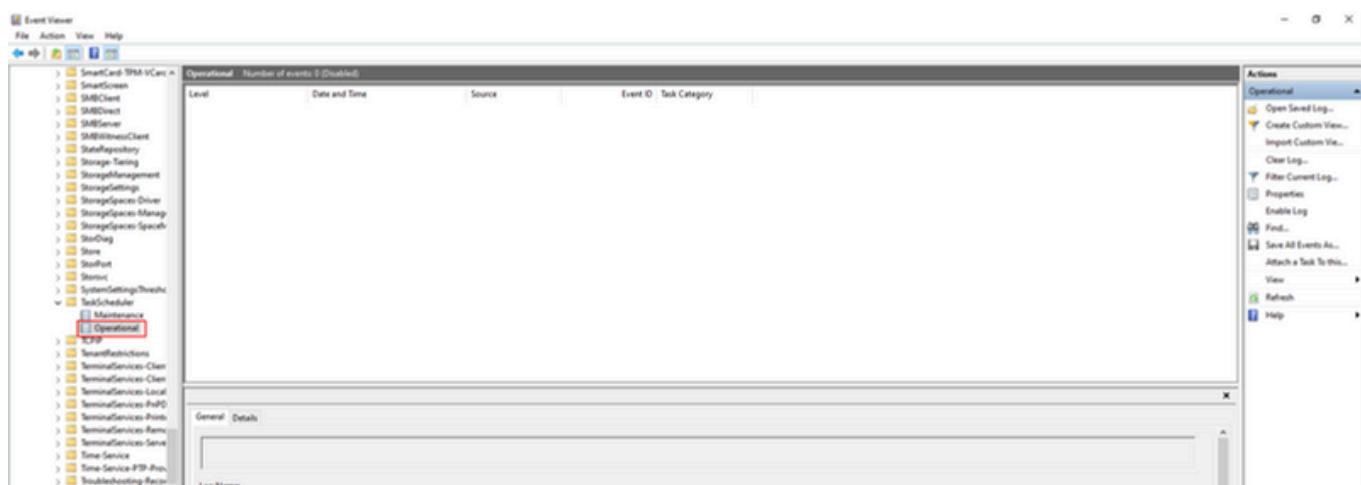
Ingesting Exported Windows Event Log Files

This guide shows how to collect Windows Task Scheduler logs and send them to Splunk using the Universal Forwarder.

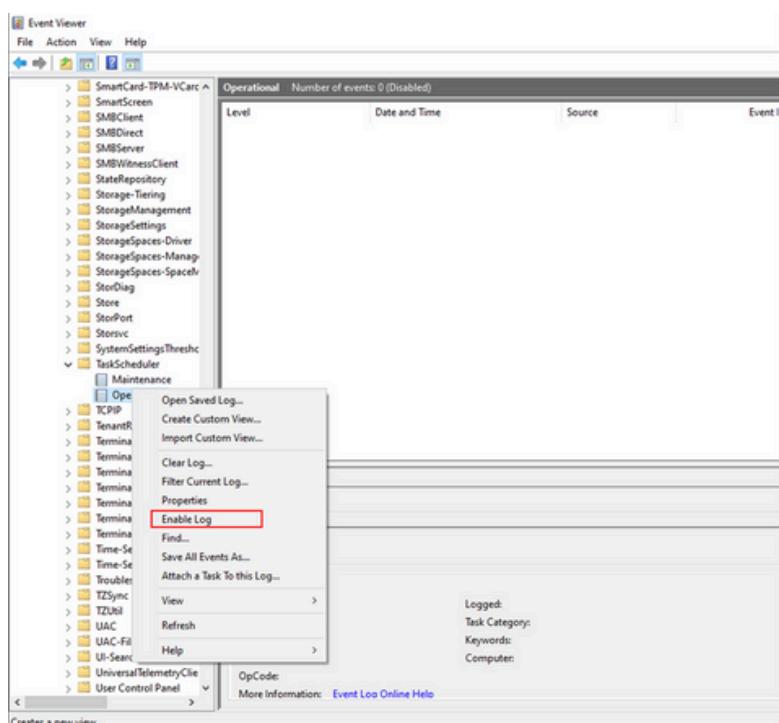
1. Enable Task Scheduler Logs

To collect detailed logs from Task Scheduler:

- Open Event Viewer (search for it in Windows).
- Go to: Applications and Services Logs > Microsoft > Windows > TaskScheduler



- Right-click **Operational** and choose **Enable Log**.

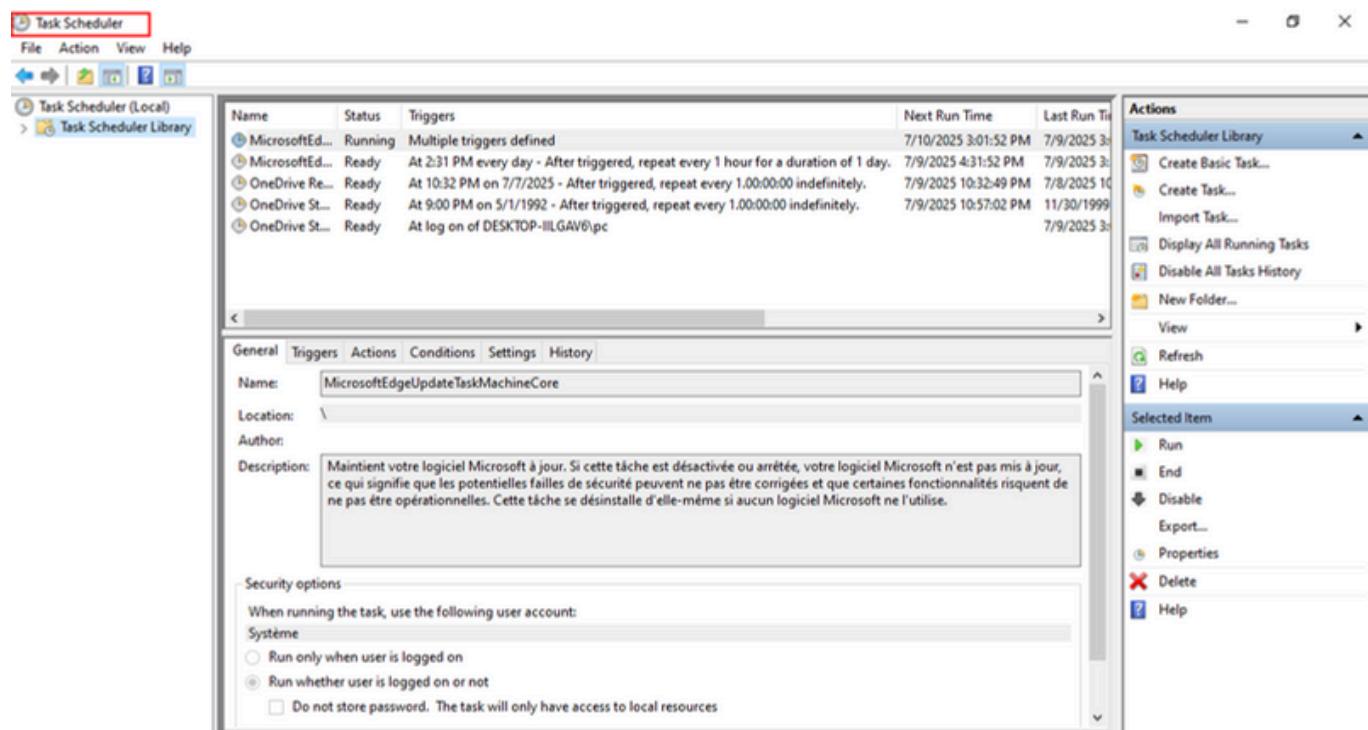


2. Simulate Persistence with a Scheduled Task

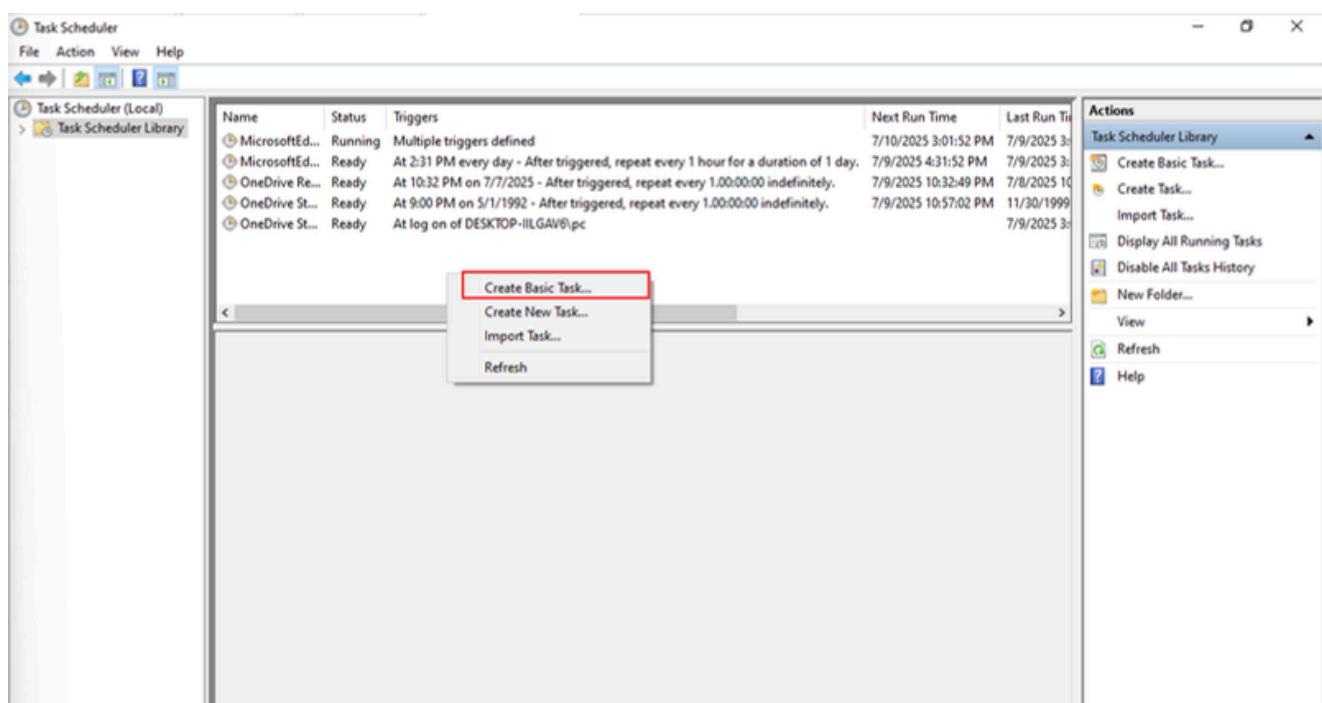
This step creates a fake scheduled task to simulate persistence (like malware staying active).

Steps:

1. Open **Task Scheduler**.

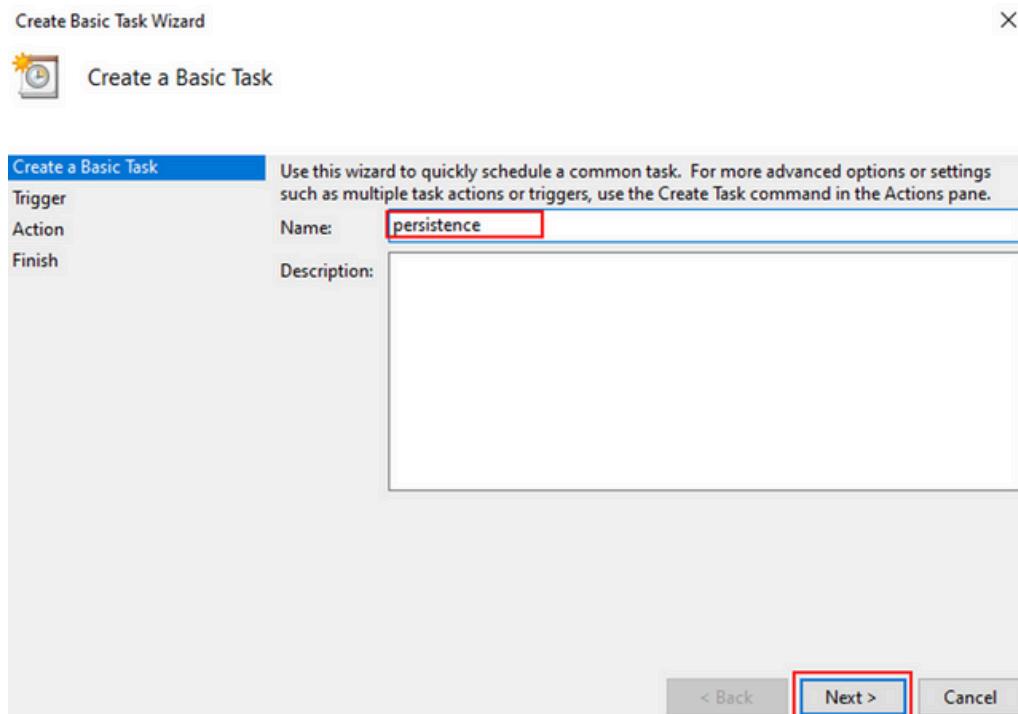


2. Click **Create Basic Task**.

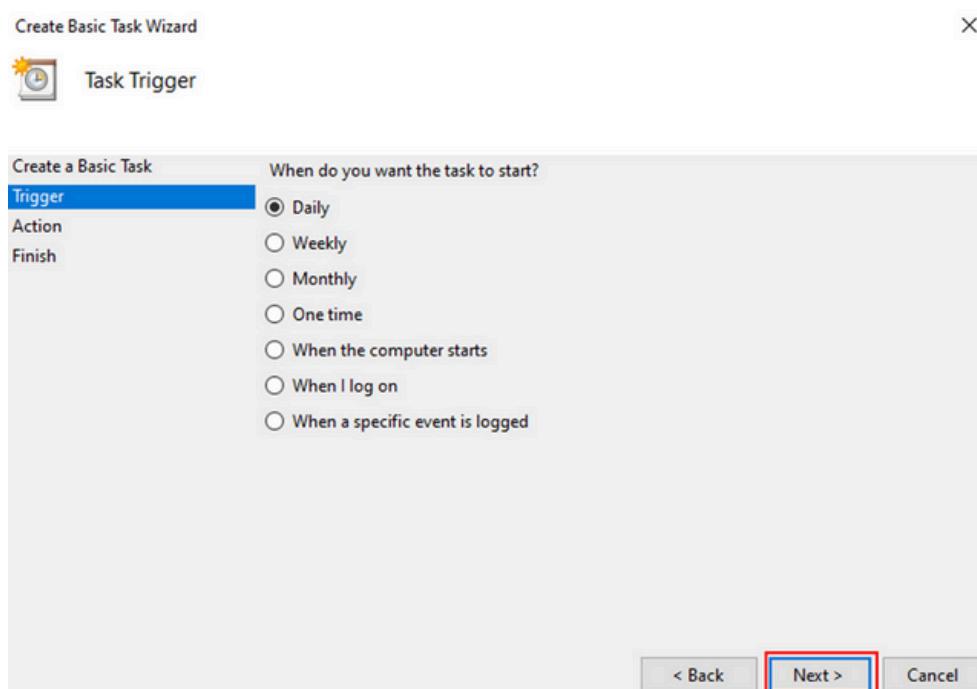


3. In the wizard:

- **Name:** persistence

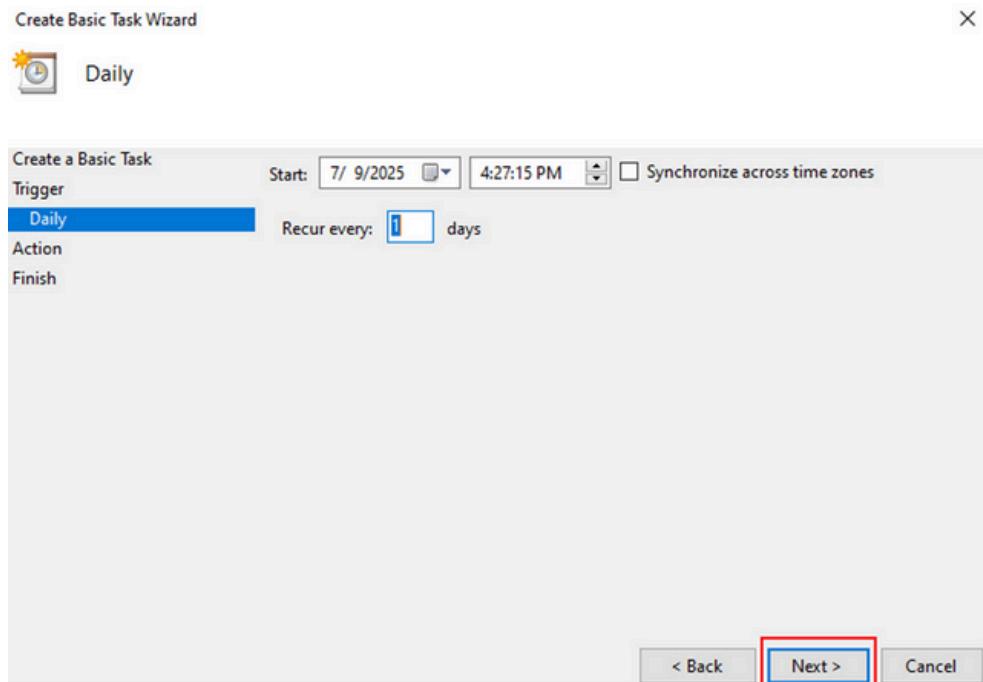


- **Trigger:** Daily

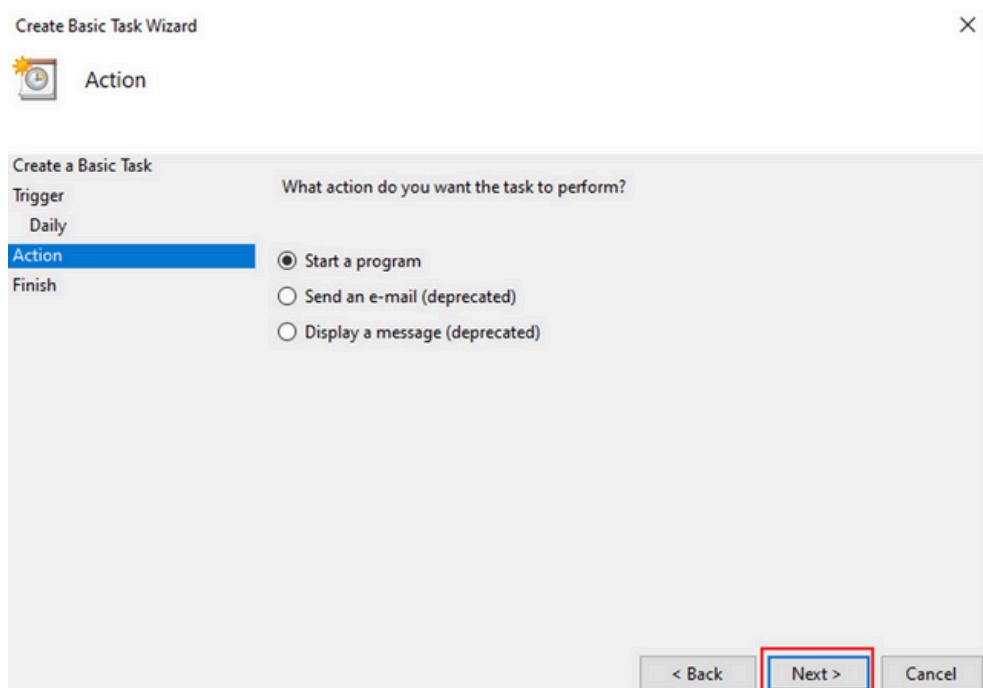


- **Start time:** Choose any time

Getting External Data into Splunk

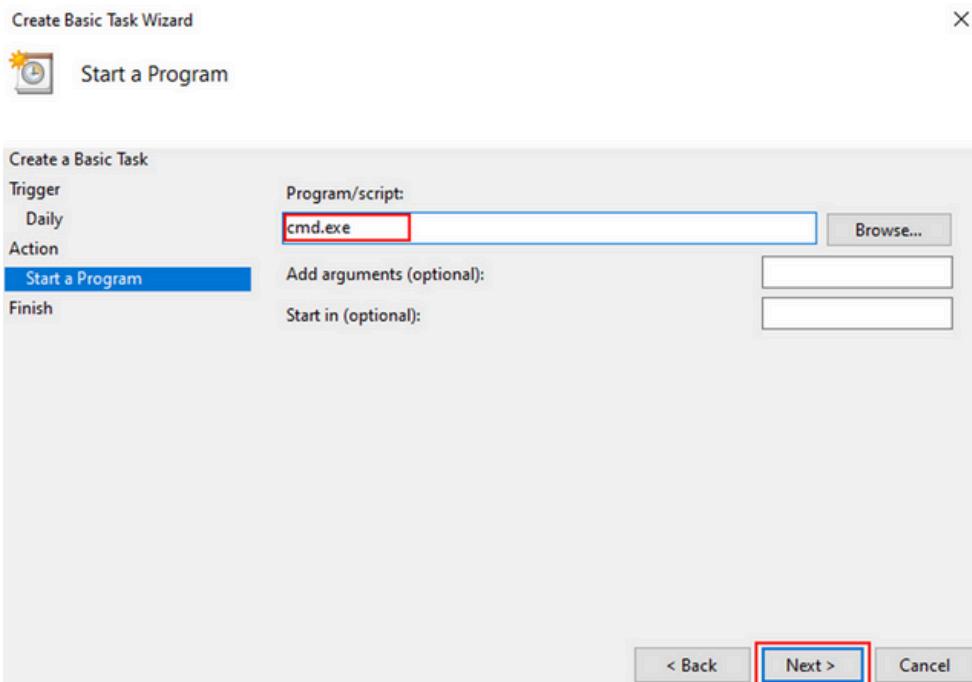


- **Action:** Start a program



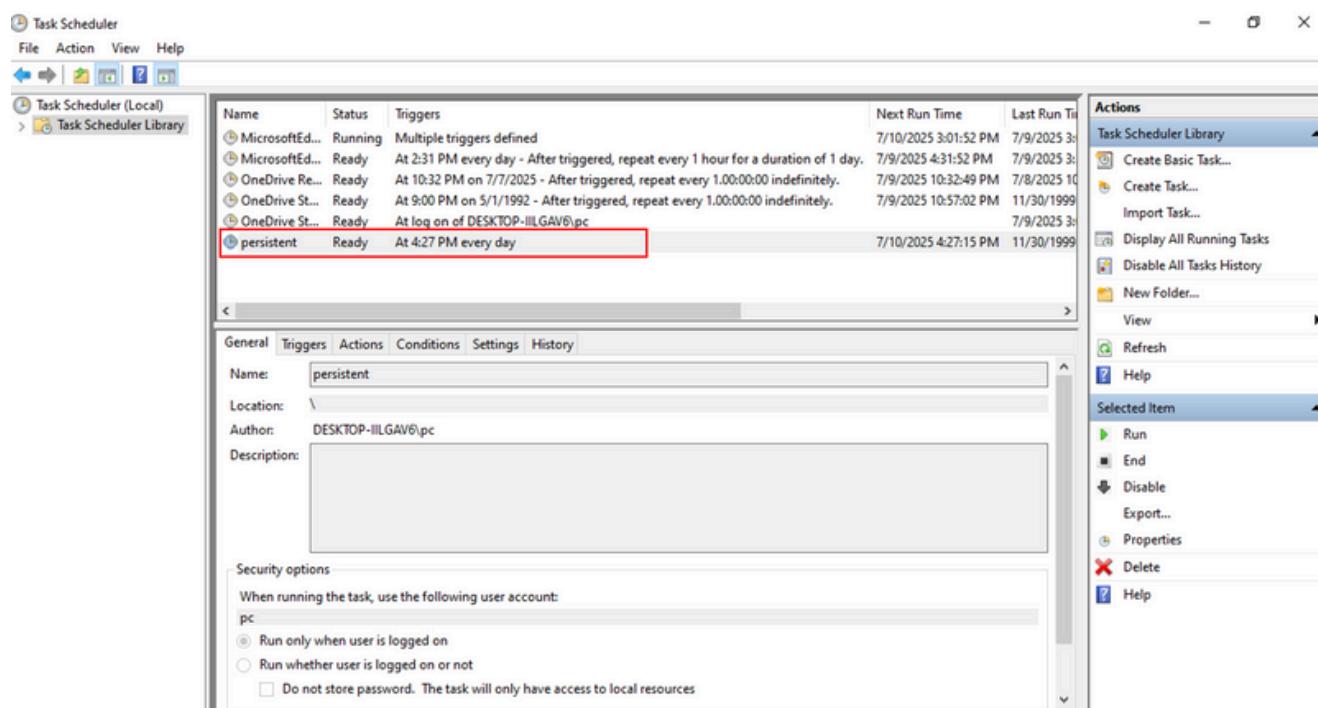
- **Program/script:** Type cmd.exe

Getting External Data into Splunk



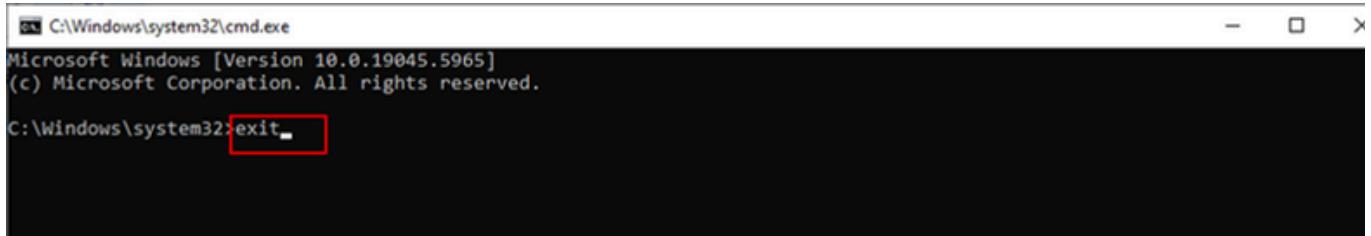
4. Click **Finish**.

5. You'll see the task listed as **Ready**.



6. To test it, right-click the task and choose **Run**, a command prompt should open.

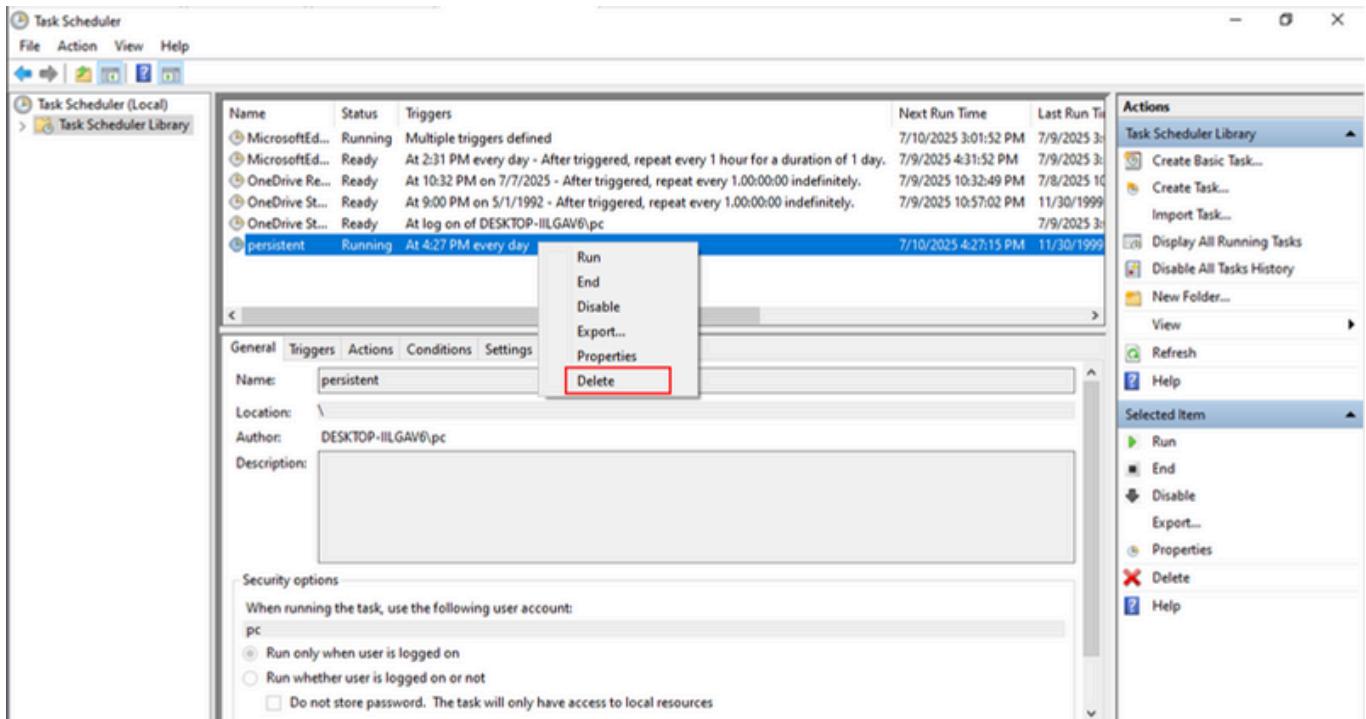
Getting External Data into Splunk



```
C:\Windows\system32\cmd.exe
Microsoft Windows [Version 10.0.19045.5965]
(c) Microsoft Corporation. All rights reserved.

C:\Windows\system32\exit
```

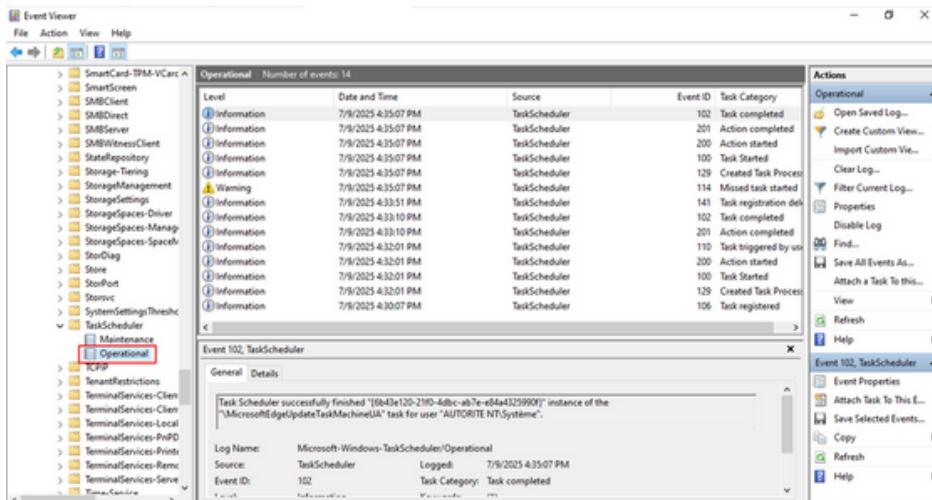
7. After testing, you can delete the task.



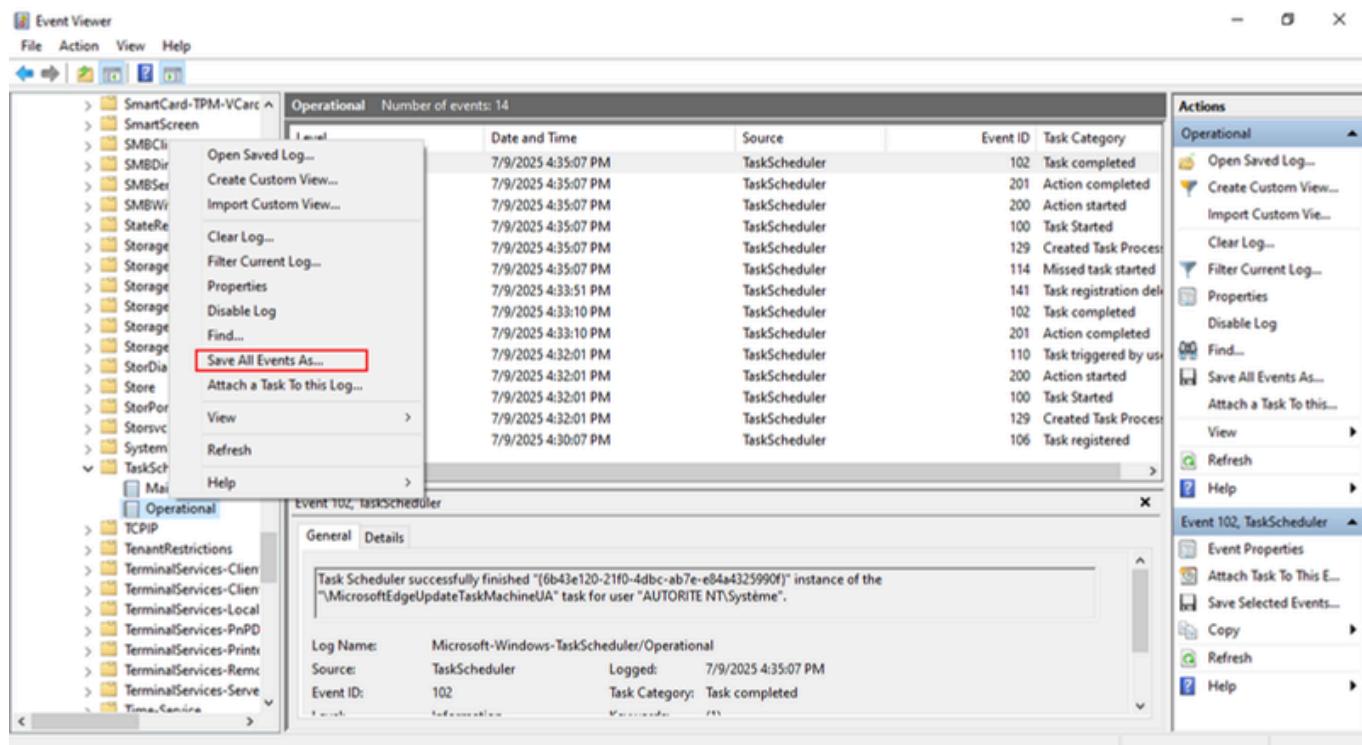
3. Save Task Scheduler Logs

1. Go back to Event Viewer →

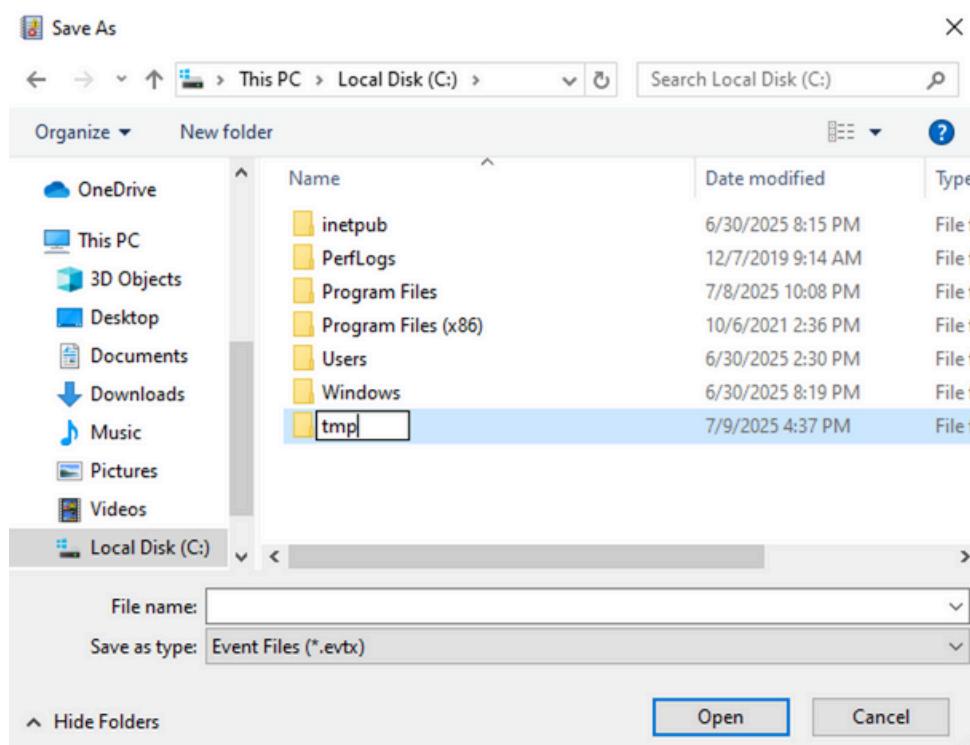
- Applications and Services Logs > Microsoft > Windows > TaskScheduler > Operational



2. Right-click Operational → Save All Events As...

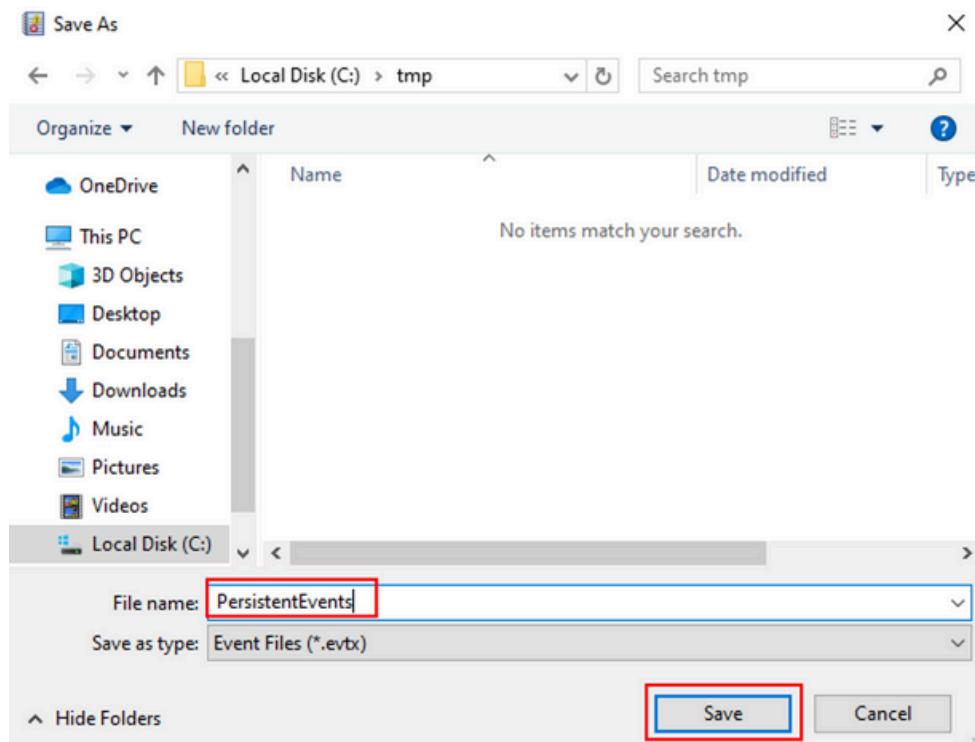


3. Choose a folder like C:\tmp

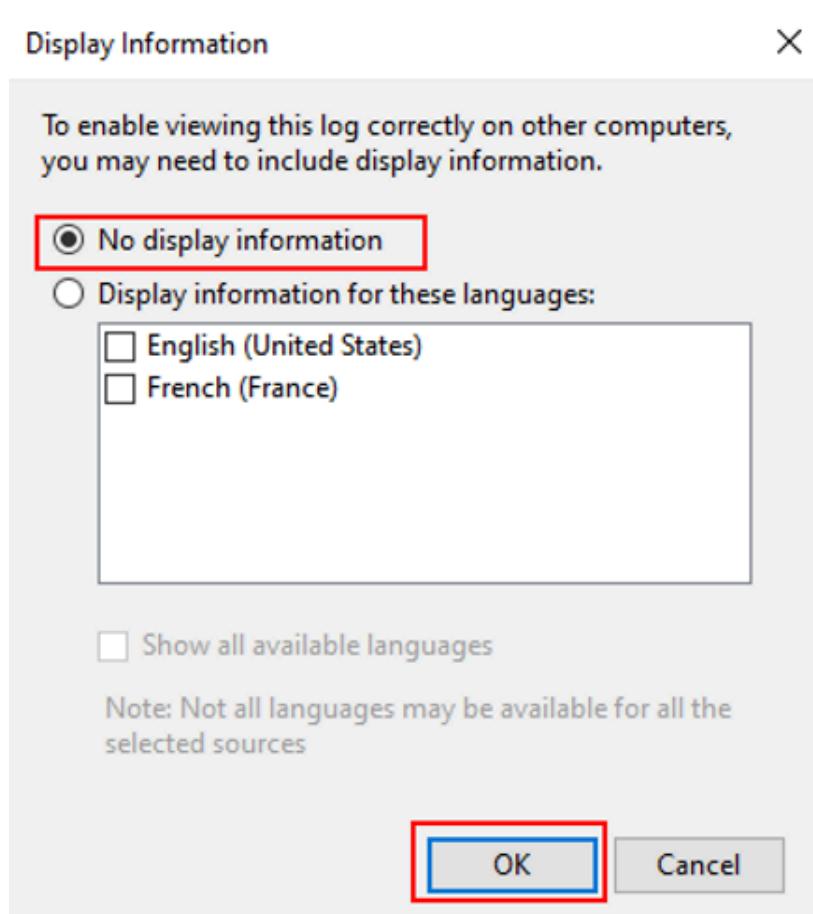


- Name the file: PersistentEvents.evtx
- Choose **Event Files (*.evtx)** as the file type
- Click **Save**

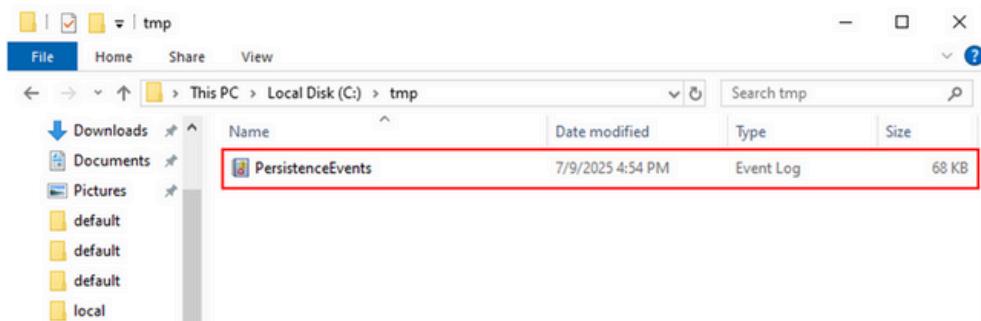
Getting External Data into Splunk



6. When asked about display info, choose **No display information**



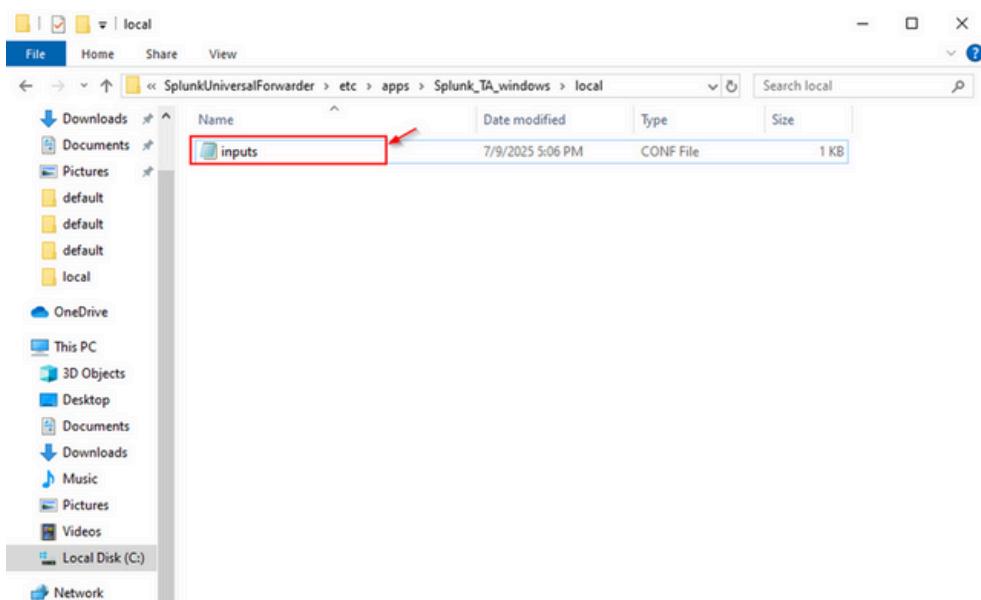
7. Confirm the file is saved in C:\tmp



4. Configure Splunk Universal Forwarder to Monitor the Folder

1. Open the file:

- C:\Program Files\SplunkUniversalForwarder\etc\apps\Splunk_TA_windows\local\inputs.conf



2. Add this configuration:

A screenshot of a Notepad window titled '*inputs - Notepad'. The menu bar includes File, Edit, Format, View, Help. The content of the file is as follows:

```
[WinEventLog://Security]
disabled = 0

[monitor://C:\tmp]
disabled = 0
index = test
```

The '[monitor://C:\tmp]' section is highlighted with a red box.

- disabled = 0 means monitoring is active
- index = test sends the data to the index named test

Getting External Data into Splunk

3. Restart the Universal Forwarder:

- Open **Admin Command Prompt**
- Go to: C:\Program Files\SplunkUniversalForwarder\bin
- Run: splunk restart

```
Administrator: Command Prompt
Microsoft Windows [Version 10.0.19045.5965]
(c) Microsoft Corporation. All rights reserved.

C:\Windows\system32>cd "\Program Files\SplunkUniversalForwarder\bin"
C:\Program Files\SplunkUniversalForwarder\bin>splunk restart
SplunkForwarder: Stopped

Splunk> The Notorious B.I.G. D.A.T.A.

Checking prerequisites...
    Checking mgmt port [8089]: open
    Checking conf files for problems...
    Done
    Checking default conf files for edits...
    Validating installed files against hashes from 'C:\Program Files\SplunkUniversalForwarder\splunkforwarder-9.4.3-237ebbd22314-windows-x64-manifest'
        All installed files intact.
    Done
All preliminary checks passed.

Starting splunk server daemon (splunkd)... .

SplunkForwarder: Starting (pid 9616)
Done
```

5. Verify Logs in Splunk Web

To check if logs are coming in:

1. Open Splunk Web in your browser.
2. In the search bar, type: index=test
3. You should see events from the folder C:\tmp, like: WinEventLog:Microsoft-Windows-TaskScheduler/Operational

The screenshot shows the Splunk Web interface with a search bar containing "index=test". Below the search bar, it says "14 events (7/9/25 4:00:00.000 PM to 7/10/25 4:57:04.000 PM)" and "No Event Sampling". The main area displays 14 event entries in a table format. Each event row includes a timestamp, a detailed log message, and several filter buttons at the bottom. On the left side, there are two panels: "SELECTED FIELDS" and "INTERESTING FIELDS", both listing various event properties like host, index, source, etc. The top navigation bar includes links for Search, Analytics, Datasets, Reports, Alerts, Dashboards, and a "Search & Reporting" button.

Onboarding Windows Registry Data

Monitoring the Windows Registry provides valuable insights for numerous use cases, aiding in security investigations and system health monitoring.

Why Monitor the Windows Registry?

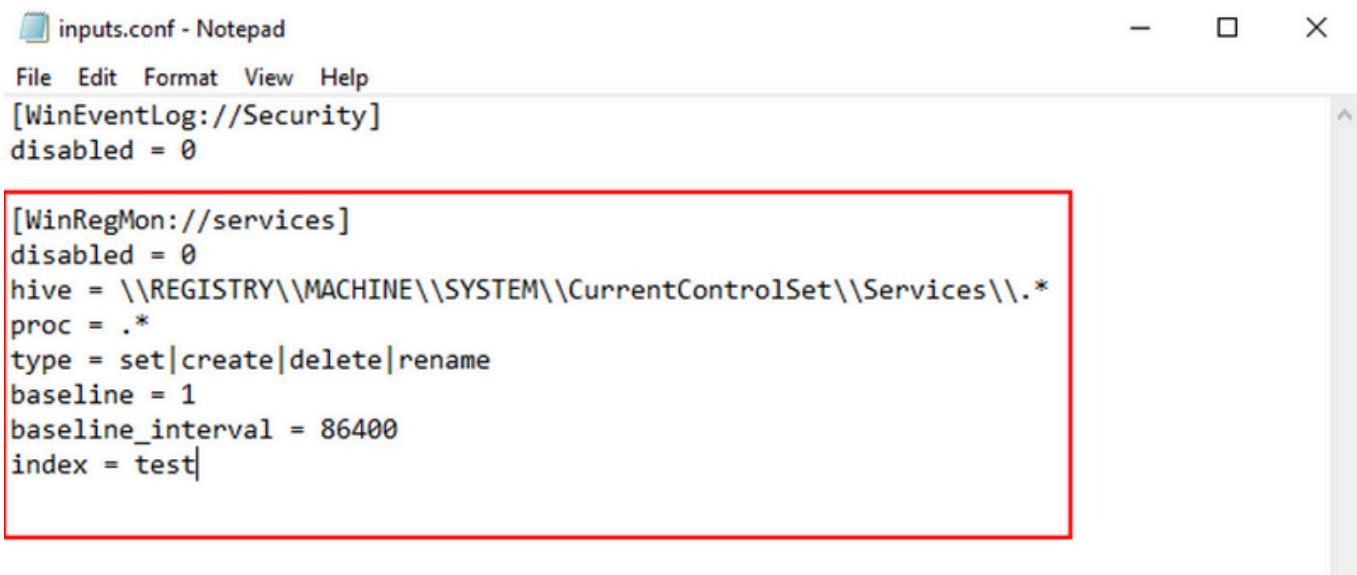
The Windows Registry is like a big settings database for Windows and its apps. Monitoring it helps you:

- **Detect new services:** Can show if malware or unknown software was installed.
- **Track autoruns:** Malware often sets itself to start with Windows.
- **Spot disabled security features:** Like WDigest being turned off — a sign of attack.
- **Help investigations:** Shows info about installed software, COM objects, etc.
- **Find weak settings:** Bad registry configs can be risky.
- **Detect unknown devices:** Registry logs can show if someone plugged in unauthorized hardware.

Configure Splunk to Monitor Registry

To collect registry data, you need to edit the inputs.conf file on the Windows machine where Splunk Universal Forwarder is installed.

• Configuration



```
inputs.conf - Notepad
File Edit Format View Help
[WinEventLog://Security]
disabled = 0

[WinRegMon://services]
disabled = 0
hive = \\REGISTRY\\MACHINE\\SYSTEM\\CurrentControlSet\\Services\\.*
proc = .*
type = set|create|delete|rename
baseline = 1
baseline_interval = 86400
index = test|
```

What Each Line Means:

- WinRegMon://services: Just a name for this config.
- hive: Which part of the registry to watch, here it's services.
- proc: Watch all processes making changes.
- type: What kind of changes to track:
 - set: Change a value
 - create: Make a new key/value
 - delete: Remove a key/value
 - rename: Rename a key
- baseline = 1: Helps detect changes from a clean state.
- baseline_interval = 86400: Re-check every 24 hours.
- index = test: Send the data to the test index in Splunk.

Where to Put This Config

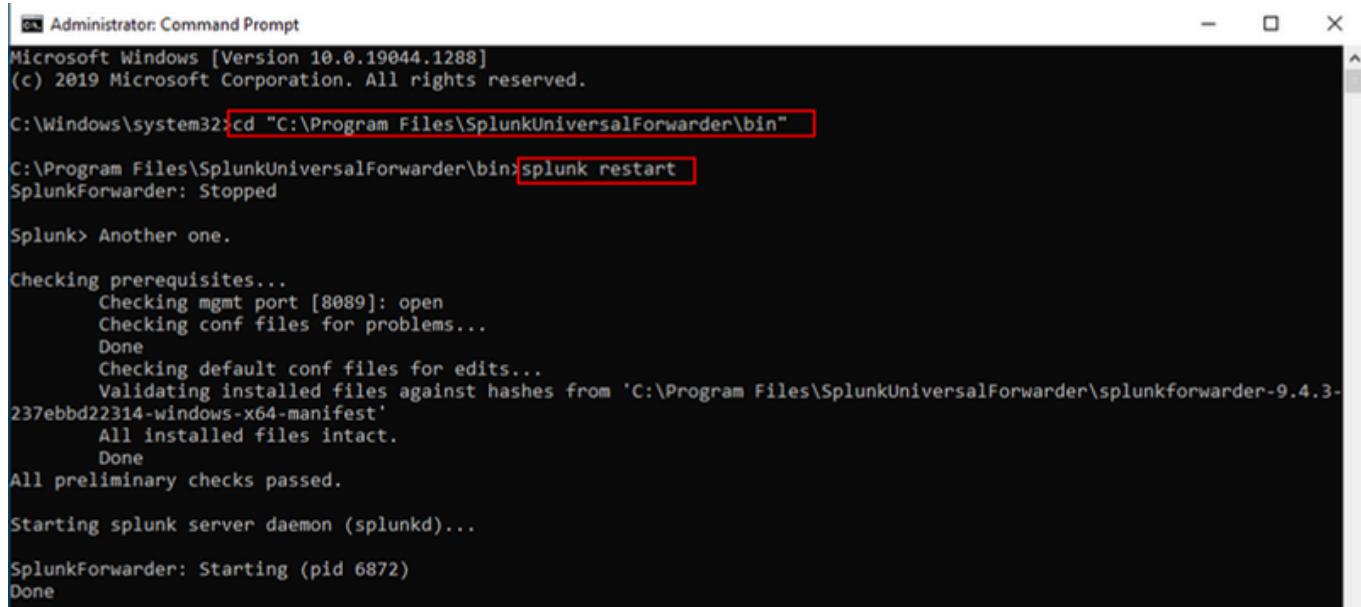
- File path:

C:\Program

Files\SplunkUniversalForwarder\etc\apps\Splunk_TA_windows\local\inputs.conf

Restart

After editing, restart Splunk Universal Forwarder:



```
Administrator: Command Prompt
Microsoft Windows [Version 10.0.19044.1288]
(c) 2019 Microsoft Corporation. All rights reserved.

C:\Windows\system32>cd "C:\Program Files\SplunkUniversalForwarder\bin"
C:\Program Files\SplunkUniversalForwarder\bin>splunk restart
SplunkForwarder: Stopped

Splunk> Another one.

Checking prerequisites...
    Checking mgmt port [8089]: open
    Checking conf files for problems...
    Done
    Checking default conf files for edits...
    Validating installed files against hashes from 'C:\Program Files\SplunkUniversalForwarder\splunkforwarder-9.4.3-237ebbd22314-windows-x64-manifest'
        All installed files intact.
        Done
All preliminary checks passed.

Starting splunk server daemon (splunkd)...

SplunkForwarder: Starting (pid 6872)
Done
```

Getting External Data into Splunk

Search and Analyze Registry Data in Splunk

Once data is coming in, you can search it in Splunk Web.

Basic Search:

The screenshot shows the Splunk Web interface with a search bar containing "index=test". Below the search bar, it says "1,896 events (7/10/25 3:00:00.000 PM to 7/11/25 3:03:03.000 PM) No Event Sampling". The main area displays two event entries in a table format. Both events are from "host = DESKTOP-SR75NGF" and "source = WinRegistry". The first event has "sourceType = WinRegistry" and the second has "sourceType = WinEventLog". Other fields like "event_status", "pid", "process_image", and "registry_type" are also visible.

Filter for Registry Events:

This screenshot shows a more advanced search where the user has added "sourceType=WinRegistry" to the search bar. A modal window titled "sourceType" is open, showing a table of values: "WinRegistry" (Count: 1,882, %: 99.26%) and "WinEventLog" (Count: 14, %: 0.73%). The user has selected "WinRegistry". The search results table below shows two events matching the criteria, both from "host = DESKTOP-SR75NGF" and "source = WinRegistry".

If you don't see results, try changing the time range or check if sampling is enabled.

Getting External Data into Splunk

What You'll See in Results

Each registry event will show:

- time: When it happened
 - event_status: Was it successful?
 - process_image: Which program made the change
 - registry_type: Type of change (e.g., createKey, SetValue)
 - host: Which machine
 - source: Where the data came from
 - sourcetype: Should be WinRegistry
 - action: What happened (e.g., modified)
 - key_path: Full path to the registry key
 - registry_key_name: Name of the key
 - registry_path: Path to the key
 - registry_value_name: Name of the value
 - registry_value_type: Type of data (e.g., REG_BINARY)

Ingesting Dynamic Windows Registry Data

Splunk's powerful data ingestion features allow organizations to collect and analyze machine data for security, operations, and performance. Going beyond basic log collection, a strategic approach focuses on targeted data to reduce noise and improve insights.

Strategic Data Collection Using inputs.conf

The inputs.conf file on Splunk Universal Forwarders is where you define what data to collect. It gives you fine control to:

- Filter specific events
- Reduce resource usage
- Focus on relevant security signals

Example: Targeted Windows Security Event Log Ingestion

Instead of collecting all logs, you can focus on critical security events, like changes to security-enabled groups (e.g., local Administrators group).

Key Event IDs:

- 4732: A user was added to a security-enabled local group
- 4733: A user was removed from a security-enabled local group

Sample Configuration in inputs.conf:



```

inputs.conf - Notepad
File Edit Format View Help
[WinEventLog://Security]
disabled = 0
whitelist = EventCode="(4732|4733)"
index = test

[WinRegMon://services]
disabled = 0
hive = \\REGISTRY\\MACHINE\\SYSTEM\\CurrentControlSet\\Services\\.*
proc = .*
type = set|create|delete|rename
baseline = 1
baseline_interval = 86400
index = test

```

- whitelist: Only collects events with ID 4732 or 4733
- index: Sends data to the security_monitoring index
- After modifying inputs.conf, a **restart** of the Splunk Universal Forwarder is mandatory for the changes to take effect:

Getting External Data into Splunk

```
C:\Program Files\SplunkUniversalForwarder\bin\splunk restart
SplunkForwarder: Stopped

Splunk> Another one.

Checking prerequisites...
    Checking mgmt port [8089]: open
    Checking conf files for problems...
        Done
    Checking default conf files for edits...
    Validating installed files against hashes from 'C:\Program Files\SplunkUniversalForwarder\splunkforwarder-9.4.3-237ebbd22314-windows-x64-manifest'
        All installed files intact.
        Done
All preliminary checks passed.

Starting splunk server daemon (splunkd)...

SplunkForwarder: Starting (pid 4812)
Done
```

This method is called source-side filtering, it reduces bandwidth and indexing costs by filtering logs before they reach Splunk.

Verifying Ingestion

To test if it works:

1. Add/remove a user from the local Administrators group
2. Search in Splunk:

The screenshot shows a Splunk search interface with the following details:

- Search Bar:** index=test sourcetype=XmIWinEventLog EventCode=4732 OR EventCode=4733
- Results Summary:** 4 events (7/11/25 12:00:00.000 AM to 7/11/25 6:12:11.000 PM) | No Event Sampling
- Time Range:** 0 events at 11 AM on Friday, July 11, 2025
- Events View:** Shows 4 XML event logs. One event is fully expanded, showing details like EventID, Version, Level, Task, Opcode, Keywords, TimeCreated, ProcessID, ThreadID, Channel, Computer, Security, MemberName, MemberSid, TargetUserName, Utilisateurs, TargetDomainName, BuiltIn, TargetSid, SubjectUserSid, DESKTOP-SR7SNGF\pc, SubjectUserName, SubjectDomainName, SubjectLogonId, and PrivilegeList. The event also includes attributes for EventCode (4732), host (DESKTOP-SR7SNGF), index (test), source (XmIWinEventLog:Security), and sourcetype (XmIWinEventLog).

Beyond EventCode: Using whitelist Flexibly

The whitelist (and its opposite, blacklist or denylist) can filter any field Splunk parses, not just EventCode.

This means you can:

- Target specific usernames
- Filter by process name
- Focus on registry paths, etc.

For advanced use, check Splunk's official documentation on allow/deny list formats.

Onboarding Generic Linux System Logs

To collect Linux logs (like /var/log/auth.log) and send them to Splunk, you need the Splunk Add-on for Unix and Linux installed on your Universal Forwarder (UF).

Step 1: Download the Add-on

- Go to Splunkbase and download **Splunk Add-on for Unix and Linux** (.tgz file).

Step 2: Transfer to the Universal Forwarder

- Use SCP to copy the .tgz file to your UF:

```
PS C:\WINDOWS\system32> scp C:\Users\pc\Downloads\splunk-add-on-for-unix-and-linux_1010.tgz ubuntu@192.168.1.15:/home/ubuntu/
ubuntu@192.168.1.15's password:
splunk-add-on-for-unix-and-linux_1010.tgz                                         100% 132KB 21.4MB/s  00:00
```

- Verify with:

```
ubuntu@ubuntu:~$ ls -l
total 87100
-rw-rw-r--  1 ubuntu  ubuntu   134865 Jul 12 16:26 splunk-add-on-for-unix-and-linux_1010.tgz
drwxr-x--- 11 ubuntu  ubuntu     4096 Jul 12 13:43 splunkforwarder
-rw-rw-r--  1 ubuntu  ubuntu  89045711 Jun  4 18:47 splunkforwarder-9.4.3-237ebbd22314-linux-amd64.tgz
```

You should see the .tgz file.

Step 3: Install the Add-on

- Move the file to the apps directory:

```
ubuntu@ubuntu:~$ mv splunk-add-on-for-unix-and-linux_1010.tgz /home/ubuntu/splunkforwarder/etc/apps/
```

- Navigate to the directory and extract:

```
ubuntu@ubuntu:~$ cd ./splunkforwarder/etc/apps/
ubuntu@ubuntu:~/splunkforwarder/etc/apps$ ls
audit_trail           learned
introspection_generator_addon    search
journald_input          splunk-add-on-for-unix-and-linux_1010.tgz  SplunkUniversalForwarder
splunk_httpinput
splunk_internal_metrics
```

```
ubuntu@ubuntu:~/splunkforwarder/etc/apps$ tar -xf splunk-add-on-for-unix-and-linux_1010.tgz
ubuntu@ubuntu:~/splunkforwarder/etc/apps$ ls
audit_trail           search
introspection_generator_addon  splunk-add-on-for-unix-and-linux_1010.tgz  SplunkUniversalForwarder
journald_input         splunk_httpinput
learned               splunk_internal_metrics
```

You should see `Splunk_TA_nix`.

- **Restart UF:**

```
ubuntu@ubuntu:~/splunkforwarder/bin$ sudo ./splunk restart
[sudo] password for ubuntu:
Warning: Attempting to revert the SPLUNK_HOME ownership
Warning: Executing "chown -R ubuntu:ubuntu /home/ubuntu/splunkforwarder"
splunkd is not running.
splunkd.pid doesn't exist...
Splunk> The IT Search Engine.

Checking prerequisites...
    Checking mgmt port [8089]: open
    Checking conf files for problems...
        Done
    Checking default conf files for edits...
    Validating installed files against hashes from '/home/ubuntu/splunkforwarder/splunkforwarder-9.4.3-237ebbd22314-linux-amd64-manifest'
        All installed files intact.
        Done
All preliminary checks passed.

Starting splunk server daemon (splunkd)...
Done
```

4. Alternative Method (Splunk Web)

- Go to **Apps (Manage) → Browse More Apps** in Splunk Enterprise.

The screenshot shows the Splunk Enterprise web interface. The top navigation bar includes links for 'Administrator', 'Messages', 'Settings', 'Activity', 'Help', and a 'Find' search bar. On the left, there's a sidebar titled 'Apps' with a 'Manage' button highlighted by a red box. Below it is a search bar and a list of available apps: 'Search & Reporting', 'Audit Trail', 'Splunk Secure Gateway', and 'Upgrade Readiness App'. The main content area is titled 'Hello, Administrator' and shows sections for 'Bookmarks', 'Dashboard', 'Search history', 'Recently viewed', 'Created by you', and 'Shared with you'. Under 'Bookmarks', there are sections for 'My bookmarks (0)' and 'Shared with my organization (0)', each with an 'Add bookmark' button.

Getting External Data into Splunk

The screenshot shows the Splunk Enterprise Apps interface. At the top, there's a navigation bar with 'splunk>enterprise' and various dropdown menus like 'Administrator', 'Messages', 'Settings', 'Activity', 'Help', and 'Find'. Below the navigation is a search bar with 'filter' and a magnifying glass icon. To the right of the search bar are buttons for 'Browse more apps', 'Install app from file', and 'Create app'. A dropdown menu for '25 per page' is open, showing options 1, 2, and Next. The main area is a table titled 'Apps' with columns: Name, Folder name, Version, Update checking, Visible, Sharing, Status, and Actions. The table lists several Splunk add-ons, including 'SplunkDeploymentServerConfig', 'SplunkForwarder', 'SplunkLightForwarder', 'Splunk Add-on for Microsoft Windows', 'Log Event Alert Action', 'Webhook Alert Action', 'Apps Browser', 'Audit Trail', 'indexes', and 'introspection_generator_addon'. Each row shows details like version (e.g., 9.0.1), status (e.g., Enabled), and actions like 'Edit properties' or 'View objects'.

Name	Folder name	Version	Update checking	Visible	Sharing	Status	Actions
SplunkDeploymentServerConfig	SplunkDeploymentServerConfig		Yes	No	App Permissions	Enabled Disable	Edit properties View objects
SplunkForwarder	SplunkForwarder		Yes	No	App Permissions	Disabled Enable	
SplunkLightForwarder	SplunkLightForwarder		Yes	No	App Permissions	Disabled Enable	
Splunk Add-on for Microsoft Windows	Splunk_TA_windows	9.0.1	Yes	No	Global Permissions	Enabled Disable	Edit properties View objects View details on Splunkbase
Log Event Alert Action	alert_logevent	9.4.2	Yes	No	App Permissions	Enabled	Edit properties View objects
Webhook Alert Action	alert_webhook	9.4.2	Yes	No	App Permissions	Enabled	Edit properties View objects
Apps Browser	appsbrowser	9.4.2	Yes	No	App Permissions	Enabled	Edit properties View objects
Audit Trail	audit_trail	1.0.0	Yes	Yes	App Permissions	Enabled	Launch app Edit properties View objects
indexes	indexes	1.0.0	Yes	No	App Permissions	Enabled Disable	Edit properties View objects
introspection_generator_addon	introspection_generator_addon	9.4.2	Yes	No	App Permissions	Enabled Disable	Edit properties View objects

- Search for **Splunk Add-on for Unix and Linux** and click Install.

The screenshot shows the 'Browse More Apps' interface in Splunk Enterprise. At the top, there's a search bar with 'linux' and a magnifying glass icon. Below the search bar are buttons for 'Best Match', 'Newest', and 'Popular'. The main area shows a grid of app cards. One card for 'Splunk Add-on for Unix and Linux' is highlighted with a red box around its 'Install' button. Another card for 'Sandfly Agentless Security for Linux' is also highlighted with a red box around its 'Install' button. Other cards visible include 'BeyondTrust Privilege Management for Unix and Linux' and 'Splunk Add-on for Linux'. On the left, there's a sidebar with a 'Category' filter section containing checkboxes for various IT operations and compliance categories. The search results show 43 apps found.

Step 4: Configure Data Input

- Create a local directory inside Splunk_TA_nix if it doesn't exist.

```
ubuntu@ubuntu:~/splunkforwarder$ cd ./etc/apps/
ubuntu@ubuntu:~/splunkforwarder/etc/apps$ ls
audit_trail           search
introspection_generator_addon  splunk-add-on-for-unix-and-linux_1010.tgz  SplunkUniversalForwarder
journald_input         splunk_httpproto
learned               splunk_internal_metrics
ubuntu@ubuntu:~/splunkforwarder/etc/apps$ cd Splunk_TA_nix/
ubuntu@ubuntu:~/splunkforwarder/etc/apps/Splunk_TA_nix$ mkdir local
```

- Create inputs.conf:

```
ubuntu@ubuntu:~/splunkforwarder/etc/apps/Splunk_TA_nix$ cd local/
ubuntu@ubuntu:~/splunkforwarder/etc/apps/Splunk_TA_nix/local$ nano inputs.conf
```

Getting External Data into Splunk

```
GNU nano 6.2
```

```
[monitor:///var/log/auth.log]
sourcetype = linux_secure
index = test
```

```
inputs.conf *
```

- Save and exit.

- **Restart UF:**

```
ubuntu@ubuntu:~/splunkforwarder/bin$ sudo ./splunk restart
[sudo] password for ubuntu:
Warning: Attempting to revert the SPLUNK_HOME ownership
Warning: Executing "chown -R ubuntu:ubuntu /home/ubuntu/splunkforwarder"
Stopping splunkd...
Shutting down. Please wait, as this may take a few minutes.

Stopping splunk helpers...

Done.
splunkd.pid doesn't exist...

Splunk> The IT Search Engine.
```

6. Verifying Data Ingestion in Splunk Enterprise

After restarting the Splunk Universal Forwarder, you can confirm that authentication logs from your Linux system are being successfully sent to Splunk.

Step-by-Step Verification

1. Perform a Search

- Go to the Search & Reporting app
- In the search bar, enter:

The screenshot shows the Splunk 'New Search' interface. The search bar contains 'index=test'. Below the search bar, it says '20 events (7/12/25 7:34:00.000 PM to 7/12/25 8:34:22.000 PM) No Event Sampling'. The 'Events (20)' tab is selected. The results table has columns for Time and Event. The first two events are:

Time	Event
Jul 12 20:16:11	splunk sudo: pam_unix(sudo:session): session opened for user root(uid=0) by splunk(uid=1000)
Jul 12 20:16:11	host = splunk index = test source = /var/log/auth.log sourcetype = linux_secure
Jul 12 20:16:11	splunk sudo: splunk : TTY=tty1 ; PWD=/home/splunk/splunk/bin ; USER=root ; COMMAND=../splunk restart
Jul 12 20:16:11	host = splunk index = test source = /var/log/auth.log sourcetype = linux_secure

2. Review Search Results

You should see events coming from /var/log/auth.log. These events confirm that your Universal Forwarder is working correctly.

Getting External Data into Splunk

The screenshot shows the Splunk Event Editor interface. At the top, there's a header with 'Time' and 'Event'. Below that, a message says 'Jul 12 20:16:11 splunk sudo: pam_unix(sudo:session): session opened for user root(uid=0) by splunk(uid=1000)'. A 'Event Actions' button is visible. The main area is a table where users can select fields from a list. The table has columns for 'Type', 'Field', 'Value', and 'Actions'. The 'Selected' section contains fields like host (splunk), index (test), source (/var/log/auth.log), and sourcetype (linux_secure). The 'Event' section lists various event-related fields such as action (success), app (sudo), authentication_service (pam_unix), and eventtype (nix-all-logs). The 'Time' section includes _time (2025-07-12T20:16:11.000+00:00). The 'Default' section includes linecount (1), punct (_"), and splunk_server (splunk). The 'Actions' column contains dropdown menus for each selected field.

Example Event Fields

Here are some of the fields you'll see in the search results:

- host: The name of your Linux server
- index: test (as configured)
- source: /var/log/auth.log
- sourcetype: linux_secure
- action: e.g., success for successful logins
- user: The username involved in the login
- _time: When the event happened

Onboarding Authentication Logs from Ubuntu UF

You can collect logs from /var/log/auth.log using two methods:

1. Using the full Splunk TA-linux_secure add-on
2. Using only inputs.conf for a lightweight setup

Prerequisites

- Splunk Universal Forwarder installed on Ubuntu
- SSH or console access to the server
- A running Splunk indexer
- A working outputs.conf file

1. Set Up outputs.conf

This file tells the Universal Forwarder where to send data.

Steps:

- Copy your existing outputs.conf to:

/home/ubuntu/splunkforwarder/etc/system/local/

- Use scp from your local machine:

```
:\Windows\system32>scp "C:\Program Files\SplunkUniversalForwarder\etc\system\local\outputs.conf" ubuntu@192.168.1.15:/home/ubuntu/splunkforwarder/etc/system/local/
ubuntu@192.168.1.15's password:                                         100% 152     0.2KB/s  00:00
outputs.conf
```

```
ubuntu@ubuntu:~/splunkforwarder/etc/system/local$ cat outputs.conf
[tcpout]
defaultGroup = default-autolb-group

[tcpout:default-autolb-group]
server = 192.168.1.13:9997

[tcpout-server://192.168.1.13:9997]
```

- Make sure the IP and port in outputs.conf match your Splunk indexer (default port is 9997)

2. Ensure Splunk User Can Read Log Files

The Splunk Forwarder runs under a specific user (often splunk). This user needs read access to /var/log/auth.log.

Add Read Permission

```
ubuntu@ubuntu:~$ sudo chmod o+r /var/log/auth.log  
[sudo] password for ubuntu:
```

- This is less secure if the file contains sensitive data.

Check

```
ubuntu@ubuntu:~$ ls -l /var/log/auth.log  
-rw-r--r-- 1 syslog adm 3594 Jul 13 15:49 /var/log/auth.log
```

- This is safer and more standard for log access.

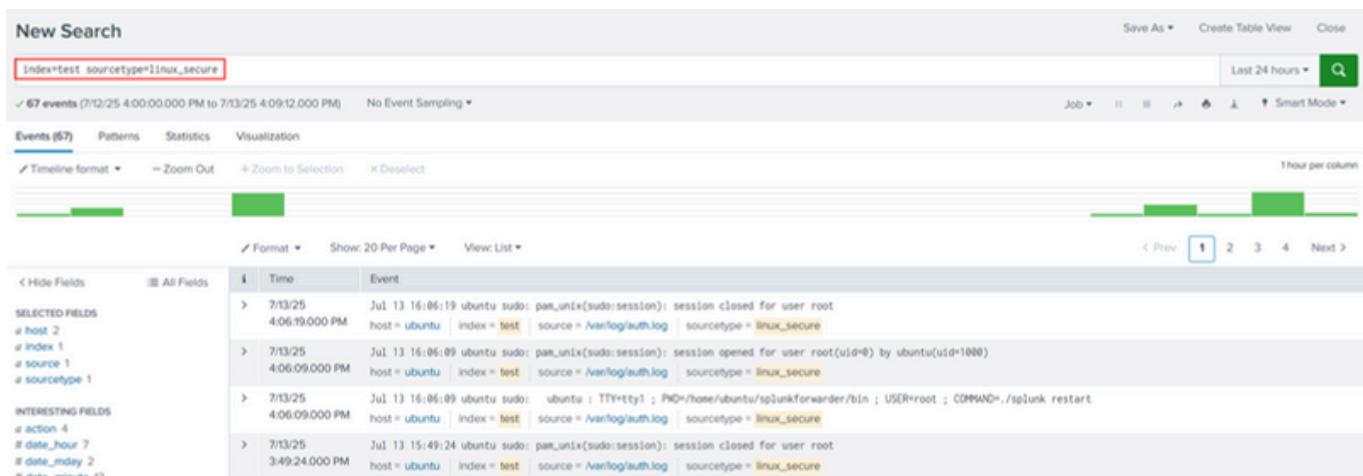
3. Restarting the Universal Forwarder

After making changes (permissions, config files), restart the service:

```
ubuntu@ubuntu:~$ cd /home/ubuntu/splunkforwarder/bin  
ubuntu@ubuntu:~/splunkforwarder/bin$ sudo ./splunk restart  
[sudo] password for ubuntu:  
Warning: Attempting to revert the SPLUNK_HOME ownership  
Warning: Executing "chown -R ubuntu:ubuntu /home/ubuntu/splunkforwarder"  
Stopping splunkd...  
Shutting down. Please wait, as this may take a few minutes.  
  
Stopping splunk helpers...  
  
Done.  
splunkd.pid doesn't exist...  
  
Splunk> Be an IT superhero. Go home early.
```

4. Final Verification in Splunk

- Go to Splunk Web
- Search:



The screenshot shows the Splunk Web interface with a search bar containing "index:test sourcetype=linux_secure". The search results table displays 67 events from July 13, 2013, between 4:06 PM and 4:09 PM. The table includes columns for Time, Event, and several other fields like host, index, source, and sourcetype. The sourcetype is consistently listed as "linux_secure". Some events show "host = ubuntu" and "index = test". One event at 4:06:19 PM shows a session closed for user root. Another at 4:06:09 PM shows a session opened for user root. A third at 4:06:24 PM shows a session closed again. The interface also shows timeline navigation, event sampling, and various search and visualization options.

- You should see parsed events from /var/log/auth.log, confirming successful onboarding!

Configuring Ingestion for Apache Web Server Access Logs

If you're running an Apache web server and want to send access logs to Splunk in a clean, structured way, here's a simple guide to help you do that using JSON format and the Splunk Universal Forwarder.

1. Configure Apache to Log in JSON Format

Splunk works best with structured data like JSON. So first, we'll set up Apache to log access data in JSON.

1.1 Define Custom Log Format

- Go to Apache config folder:

```
ubuntu@apache:~$ cd /etc/apache2/
ubuntu@apache:/etc/apache2$ ls -al
total 88
drwxr-xr-x  8 root root  4096 Jul 13 20:45 .
drwxr-xr-x 97 root root  4096 Jul 13 20:46 ..
-rw-r--r--  1 root root  7224 Apr  3 09:05 apache2.conf
drwxr-xr-x  2 root root  4096 Jul 13 20:45 conf-available
drwxr-xr-x  2 root root  4096 Jul 13 20:45 conf-enabled
-rw-r--r--  1 root root 1782 Mar 18 2024 envvars
-rw-r--r--  1 root root 31063 Mar 18 2024 magic
drwxr-xr-x  2 root root 12288 Jul 13 20:45 mods-available
drwxr-xr-x  2 root root  4096 Jul 13 20:45 mods-enabled
-rw-r--r--  1 root root   320 Mar 18 2024 ports.conf
drwxr-xr-x  2 root root  4096 Jul 13 20:45 sites-available
drwxr-xr-x  2 root root  4096 Jul 13 20:45 sites-enabled
```

- Create a new config file (e.g. log-splunk.conf) inside conf-available:

```
ubuntu@apache:/etc/apache2$ sudo nano conf-available/log-splunk.conf
```

- Add this custom log format:

```
GNU nano 6.2                                     conf-available/log-splunk.conf *
<IfModule log_config_module>
#
# The following directives define some format nicknames for use with
# a CustomLog directive (see below).
#
LogFormat "%h %l %u %t \"%r\" %>s %b \"%{Referer}i\" \"%{User-Agent}i\"" combined
LogFormat "%h %l %u %t \"%r\" %>s %b" common

<IfModule logio_module>
    # You need to enable mod_logio.c to use %I and %
<client=%h, status=%>s, uri_path=\"%U\", uri_query=\"%q\", user=\"%u\"> splunk_json

    #LogFormat "{\"time\": \"\${%s}t.\${usec_frac}t\", \"bytes_in\": \"%I\", \"bytes_out\": \"%O\", \"cookie\": \"\${%s}t.\${usec_frac}t\", \"status\": \"%>s\", \"method\": \"\${%m}\", \"uri\": \"\${uri}\", \"user\": \"\${user}\", \"host\": \"\${%h}\", \"port\": \"\${%p}\", \"proto\": \"\${%P}\", \"referer\": \"\${%{Referer}i}\", \"user_agent\": \"\${%{User-Agent}i}\", \"splunk_json\": true}" combined
    #LogFormat "%h %l %u %t \"%r\" %>s %b \"%{Referer}i\" \"%{User-Agent}i\" %I %O" combinedio

</IfModule>
#
# The location and format of the access logfile (Common Logfile Format).
# If you do not define any access logfiles within a <VirtualHost>
# container, they will be logged here. Contrariwise, if you *do*
# define per-<VirtualHost> access logfiles, transactions will be
# logged therein and *not* in this file.
#
# CustomLog "logs/access_log" common
#
# If you prefer a logfile with access, agent, and referer information
# (Combined Logfile Format) you can use the following directive.
#
CustomLog "/var/log/apache2/access.log" splunk_json
#CustomLog "logs/access_log" splunk_json
#CustomLog "logs/access_log" combined
</IfModule>
```

- Save and exit (Ctrl+X, then Y, then Enter).

1.2 Apply the JSON Format to Your Virtual Host

- Edit your default virtual host file:

```
ubuntu@apache:/etc/apache2$ sudo nano sites-available/000-default.conf
```

- Inside <VirtualHost *:80>, find or add this line:

Getting External Data into Splunk

```
GNU nano 6.2                               sites-available/000-default.conf *
<VirtualHost *:80>
    # The ServerName directive sets the request scheme, hostname and port that
    # the server uses to identify itself. This is used when creating
    # redirection URLs. In the context of virtual hosts, the ServerName
    # specifies what hostname must appear in the request's Host: header to
    # match this virtual host. For the default virtual host (this file) this
    # value is not decisive as it is used as a last resort host regardless.
    # However, you must set it for any further virtual host explicitly.
    #ServerName www.example.com

    ServerAdmin webmaster@localhost
    DocumentRoot /var/www/html

    # Available loglevels: trace8, ..., trace1, debug, info, notice, warn,
    # error, crit, alert, emerg.
    # It is also possible to configure the loglevel for particular
    # modules, e.g.
    #LogLevel info ssl:warn

    ErrorLog ${APACHE_LOG_DIR}/error.log
    CustomLog ${APACHE_LOG_DIR}/access.log splunk_json

    # For most configuration files from conf-available/, which are
    # enabled or disabled at a global level, it is possible to
    # include a line for only one particular virtual host. For example the
    # following line enables the CGI configuration for this host only
    # after it has been globally disabled with "a2disconf".
    #Include conf-available/serve-cgi-bin.conf
</VirtualHost>

# vim: syntax=apache ts=4 sw=4 sts=4 sr noet
```

- Save and exit.
- Enable the config:

```
ubuntu@apache:/etc/apache2$ sudo a2enconf log-splunk
Enabling conf log-splunk.
To activate the new configuration, you need to run:
  systemctl reload apache2
```

- Apply Apache Changes

```
ubuntu@apache:/etc/apache2/conf-enabled$ sudo systemctl reload apache2
```

1.3 Verify Apache is Logging in JSON

- Generate some traffic (e.g., open your website).

Getting External Data into Splunk

```
ubuntu@apache:/etc/apache2/conf-enabled$ curl http://127.0.0.1
<!DOCTYPE html PUBLIC "-//W3C//DTD XHTML 1.0 Transitional//EN" "http://www.w3.org/TR/xhtml1/DTD/xhtml1-transitional.dtd">
<html xmlns="http://www.w3.org/1999/xhtml">
  <!--
    Modified from the Debian original for Ubuntu
    Last updated: 2022-03-22
    See: https://launchpad.net/bugs/1966004
  -->
  <head>
    <meta http-equiv="Content-Type" content="text/html; charset=UTF-8" />
    <title>Apache2 Ubuntu Default Page: It works</title>
    <style type="text/css" media="screen">
      *
      margin: 0px 0px 0px 0px;
      padding: 0px 0px 0px 0px;
    }

    body, html {
      padding: 3px 3px 3px 3px;

      background-color: #D8DBE2;

      font-family: Ubuntu, Verdana, sans-serif;
      font-size: 11pt;
      text-align: center;
    }

    div.main_page {
      position: relative;
      display: table;
    }
  
```

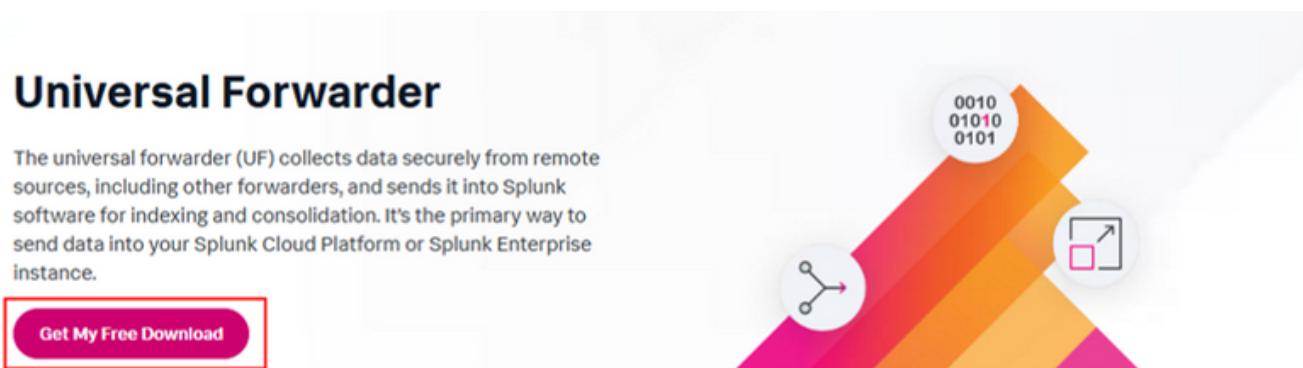
- Check the log:

```
ubuntu@apache:/var/log/apache2$ tail access.log
time=1752507052.017982, bytes_in=73, bytes_out=10926, cookie="-", server=127.0.1.1, dest_port=80, http_content_type="-", http_method="GET", http_referrer="-", http_user_agent="curl/7.81.0", ident="-", response_time_microseconds=2580, client=127.0.0.1, status=200, uri_path="/index.html", uri_query="", user="-"
```

You should see JSON-formatted entries.

2. Install Splunk Universal Forwarder

- Go to Splunk Downloads and get the .tgz for Linux.



Getting External Data into Splunk

Windows	Linux	Mac OS	Free BSD	Solaris	AIX		
PPCLE	4.x+, or 5.x+ kernel Linux distributions	.rpm	32.47 MB	Download Now	Copy wget link	More	
		.tgz	32.6 MB	Download Now	Copy wget link	More	
ARM	4.14+, 5.4+ kernel Linux distributions with libc v2.21+, 6.x+ kernel, Graviton+ Servers 64-bit	.deb	52.17 MB	Download Now	Copy wget link	More	
		.rpm	84.76 MB	Download Now	Copy wget link	More	
		.tgz	75.01 MB	Download Now	Copy wget link	More	
64-bit	4.x+, 5.x+, 6.x+ kernel Linux distributions	.rpm	97.21 MB	Download Now	Copy wget link	More	
		.deb	64.56 MB	Download Now	Copy wget link	More	
		.tgz	84.92 MB	Download Now	Copy wget link	More	
s390x	4.x+, or 5.x+ kernel Linux distributions	.tgz	31.0 MB	Download Now	Copy wget link	More	
		.rpm	30.71 MB	Download Now	Copy wget link	More	

- Use wget to download it on your server.

```
ubuntu@apache:~$ wget -O splunkforwarder-9.4.3-237ebbd22314-linux-amd64.tgz "https://download.splunk.com/products/universalforwarder/releases/9.4.3/linux/splunkforwarder-9.4.3-237ebbd22314-linux-amd64.tgz"
--2025-07-14 16:01:23-- https://download.splunk.com/products/universalforwarder/releases/9.4.3/linux/splunkforwarder-9.4.3-237ebbd22314-linux-amd64.tgz
Resolving download.splunk.com (download.splunk.com)... 52.84.66.28, 52.84.66.112, 52.84.66.10, ...
Connecting to download.splunk.com (download.splunk.com)|52.84.66.28|:443... connected.
HTTP request sent, awaiting response... 200 OK
Length: 89045711 (85M) [binary/octet-stream]
Saving to: 'splunkforwarder-9.4.3-237ebbd22314-linux-amd64.tgz'

splunkforwarder-9.4.3-2 100%[=====] 84.92M 1.95MB/s in 43s

2025-07-14 16:02:11 (1.98 MB/s) - 'splunkforwarder-9.4.3-237ebbd22314-linux-amd64.tgz' saved [89045711/89045711]
```

- Extract it:

```
ubuntu@apache:~$ tar -xzvf ./splunkforwarder-9.4.3-237ebbd22314-linux-amd64.tgz
splunkforwarder/
splunkforwarder/swidtag/
splunkforwarder/swidtag/splunk-UniversalForwarder-primary.swidtag
splunkforwarder/opt/
splunkforwarder/opt/jemalloc-4k-stats/
splunkforwarder/opt/jemalloc-4k-stats/include/
splunkforwarder/opt/jemalloc-4k-stats/include/jemalloc/
splunkforwarder/opt/jemalloc-4k-stats/include/jemalloc/jemalloc.h
splunkforwarder/opt/jemalloc-4k-stats/lib/
```

Getting External Data into Splunk

- Enable the forwarder:

```
ubuntu@apache:~$ sudo /home/ubuntu/splunkforwarder/bin/splunk enable boot-start -user ubuntu -systemd-managed 1
```

- Start the forwarder:

```
ubuntu@apache:~$ ./splunkforwarder/bin/splunk start
Warning: Attempting to revert the SPLUNK_HOME ownership
Warning: Executing "chown -R ubuntu:ubuntu /home/ubuntu/splunkforwarder"
Couldn't change ownership for /home/ubuntu/splunkforwarder/etc: Operation not permitted

Splunk> Needle. Haystack. Found.

Checking prerequisites...
    Checking mgmt port [8089]: open
    Checking conf files for problems...
        Done
    Checking default conf files for edits...
    Validating installed files against hashes from '/home/ubuntu/splunkforwarder/splunkforwarder-9.4.3-237ebb
d22314-linux-amd64-manifest'
        All installed files intact.
        Done
All preliminary checks passed.

Starting splunk server daemon (splunkd)...
==== AUTHENTICATING FOR org.freedesktop.systemd1.manage-units ===
Authentication is required to start 'SplunkForwarder.service'.
Authenticating as: ubuntu
Password:
==== AUTHENTICATION COMPLETE ===
Done
```

3. Install Splunk Add-on for Apache

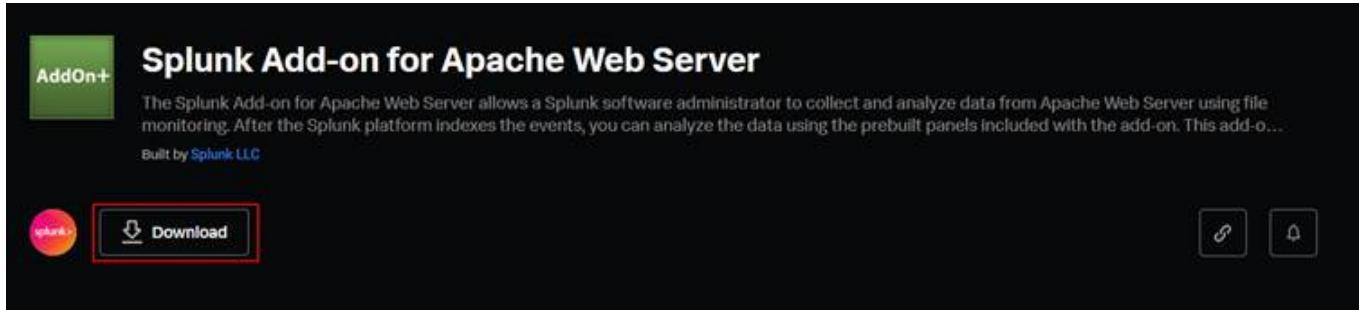
Search for the appropriate Apache-related apps or add-ons on Splunkbase to enable log collection and analysis from Apache Web Servers within Splunk.

The screenshot shows the Splunkbase interface with a search bar containing 'apache'. Below the search bar, there are three tabs: 'Best Match', 'Newest', and 'Popular'. A sidebar on the left lists various categories such as IT Operations, Security, Fraud & Compliance, Business Analytics, Utilities, Artificial Intelligence, IoT & Industrial Data, DevOps, Directory Service, Email, Endpoint, Firewall, Generic, Identity Management, Information, Investigative, Network Access Control, Network Device, Network Security, Reputation, Sandbox, and SIEM. The main content area displays two search results:

- Atlas ITSI Content Pack for Apache Web Server**: This entry includes a green 'Install' button. The description states: "The ITSI Content Pack for Apache from Presidio Splunk Solutions is specifically designed to monitor the performance and health of Apache web servers. It leverages Splunk ITSI to provide in-depth analysis and visualization of logs for Apache, ensuring critical systems are operating optimally. This content pack is an essential tool for IT professionals... [More](#)". It was released 2 months ago and has 153 downloads.
- Splunk Add-on for Apache Web Server**: This entry also includes a green 'Install' button. The description states: "The Splunk Add-on for Apache Web Server allows a Splunk software administrator to collect and analyze data from Apache Web Server using file monitoring. After the Splunk platform indexes the events, you can analyze the data using the prebuilt panels included with the add-on". It was released 3 years ago and has 32439 downloads.

4. Download Splunk Add-on for Apache

- Go to Splunkbase and download the Splunk Add-on for Apache Web Server.



- Move it to the forwarder's etc/apps folder:

scp from local machine:

```
PS C:\WINDOWS\system32> scp C:\Users\pc\Downloads\splunk-add-on-for-apache-web-server_210.tgz ubuntu@192.168.1.17:/home/ubuntu  
ubuntu@192.168.1.17's password:  
splunk-add-on-for-apache-web-server_210.tgz
```

100% 37KB 12.0MB/s 00:00

mv on the server:

```
ubuntu@apache:~$ mv splunk-add-on-for-apache-web-server_210.tgz /home/ubuntu/splunkforwarder/etc/apps/
```

- Extract the add-on:

```
ubuntu@apache:~/splunkforwarder/etc/apps$ tar -xf splunk-add-on-for-apache-web-server_210.tgz
```

This will create a directory like Splunk_TA_apache.

- Restart the forwarder:

```
ubuntu@apache:~/splunkforwarder/bin$ sudo ./splunk restart  
Warning: Attempting to revert the SPLUNK_HOME ownership  
Warning: Executing "chown -R ubuntu:ubuntu /home/ubuntu/splunkforwarder"  
Stopping splunkd...  
Shutting down. Please wait, as this may take a few minutes.  
  
Stopping splunk helpers...  
  
Done.  
splunkd.pid doesn't exist...  
  
Splunk> Needle. Haystack. Found.  
  
Checking prerequisites...  
  Checking mgmt port [8089]: open  
  Checking conf files for problems...  
  Done  
  Checking default conf files for edits...  
  Validating installed files against hashes from '/home/ubuntu/splunkforwarder/splunkforwarder-9.4.3-237ebb  
d22314-linux-amd64-manifest'  
    All installed files intact.  
    Done  
All preliminary checks passed.  
  
Starting splunk server daemon (splunkd)...  
Done
```

5. Configure Log Monitoring

- Create the local folder if it doesn't exist:

```
ubuntu@apache:~/splunkforwarder/etc/apps/Splunk_TA_apache$ mkdir local
```

- Create inputs.conf:

```
ubuntu@apache:~/splunkforwarder/etc/apps/Splunk_TA_apache/local$ sudo nano inputs.conf
```

- Add this config:

```
GNU nano 6.2                                inputs.conf
[monitor:///var/log/apache2/access.log]
sourcetype = apache:access:json
index = test
disabled = 0
```

- Save and restart:

```
ubuntu@apache:~/splunkforwarder/bin$ sudo ./splunk restart
[sudo] password for ubuntu:
Warning: Attempting to revert the SPLUNK_HOME ownership
Warning: Executing "chown -R ubuntu:ubuntu /home/ubuntu/splunkforwarder"
Stopping splunkd...
Shutting down. Please wait, as this may take a few minutes.

Stopping splunk helpers...

Done.
splunkd.pid doesn't exist...

Splunk> Needle. Haystack. Found.

Checking prerequisites...
    Checking mgmt port [8089]: open
    Checking conf files for problems...
        Done
    Checking default conf files for edits...
    Validating installed files against hashes from '/home/ubuntu/splunkforwarder/splunkforwarder-9.4.3-237ebb
d22314-linux-amd64-manifest'
        All installed files intact.
        Done
All preliminary checks passed.

Starting splunk server daemon (splunkd)...
Done
```

6. Verify Logs in Splunk

- Go to Splunk Web.
- Open Search & Reporting.
- Run:

Getting External Data into Splunk

The screenshot shows the Splunk web interface. At the top, there is a search bar with the query "Index:test host:apache sourcetype=apache.access json". Below the search bar, it says "1 event (7/16/25 2:00:00.000 PM to 7/16/25 2:42:54.000 PM) No Event Sampling". On the right side of the header, there are buttons for "Last 24 hours", "Job", "Smart Mode", and a search icon. Below the header, there are tabs for "Events (0)", "Patterns", "Statistics", and "Visualization". Under "Events (0)", there are sub-options: "Timeline format", "Zoom Out", "Zoom to Selection", and "Deselect". To the right of these, there is a "1 hour per column" button and a green progress bar.

In the main content area, there is a table titled "Event". The table has columns: "Time" and "Event". There is one event listed:

Time	Event
7/16/25 2:42:49.670 PM	{ "bytes_in": 73 "bytes_out": 18926 "client": "127.0.0.1" "cookie": " "dest_port": 80 "http_content_type": " "http_method": "GET" "http_referer": " "http_user_agent": "curl/7.81.0" "ident": " "response_time_microseconds": 324 "server": "127.0.0.1" "status": 200 "time": 1752676369.679181 "url_path": "/index.html" "url_query": " "user": " } Show as raw text

Below the table, there are buttons for "host = apache", "index = test", "source = /var/log/apache2/access.log", and "sourcetype = apache.access.json".

You should see your Apache logs in JSON format

Comprehensive Guide to Onboarding CSV Files

CSV files are everywhere simple, structured, and easy to work with. But when it comes to Splunk, you need to make sure the data is ingested properly so you can search and analyze it easily. Here's a step-by-step guide to help you do that using Indexed Extractions.

What Are Indexed Extractions?

Normally, Splunk reads each line of a file as a raw event. That means even the header (like rank, domain, tld) shows up as an event.

With Indexed Extractions, Splunk reads the header and uses it to extract fields from each line. So instead of raw lines, you get structured events like:

rank=1, domain=google.com, tld=com

```
ubuntu@ubuntu:~$ ls -la /var/log/top*
-rw-rw-r-- 1 ubuntu ubuntu 28941 Jul 17 13:00 /var/log/top-1000.csv
-rw-rw-r-- 1 ubuntu ubuntu 28941 Jul 18 10:45 /var/log/top-1000.txt
ubuntu@ubuntu:~$ head /var/log/top-1000.csv
# These are the top 1000 domains and their top level domains
rank, domain, tld
1, google.com, com
2, microsoft.com, com
3, www.google.com, com
4, netflix.com, com
5, cloud.netflix.com, com
6, prod.cloud.netflix.com, com
7, data.microsoft.com, com
8, ftl.netflix.com, com
```

- **Important:** All parsing and field extraction must happen on the Universal Forwarder. Settings on the Indexer won't work here.

Monitor Your CSV File

To tell Splunk to watch your CSV file for new data:

- Create top1000_csv_inputs and default folders
- Go to your app's inputs directory
- Create inputs.conf:

Getting External Data into Splunk

```
ubuntu@ubuntu:~$ cd ./splunkforwarder/etc/apps/
ubuntu@ubuntu:~/splunkforwarder/etc/apps$ mkdir top1000_csv_inputs
ubuntu@ubuntu:~/splunkforwarder/etc/apps$ cd top1000_csv_inputs/
ubuntu@ubuntu:~/splunkforwarder/etc/apps/top1000_csv_inputs$ mkdir default
ubuntu@ubuntu:~/splunkforwarder/etc/apps/top1000_csv_inputs$ cd default/
ubuntu@ubuntu:~/splunkforwarder/etc/apps/top1000_csv_inputs/default$ nano inputs.conf
```

- Add this stanza:

```
GNU nano 6.2                                         inputs.conf *
[monitor:///var/log/top-1000.csv]
index = test
disabled = 0
```

Verify Ingestion in Splunk

- Go to Splunk Web and search:

Time	Event
7/17/25 100:19:00 PM	1000,pull-hls-f77-sg01.tiktokcdn.com,com
7/17/25 100:19:00 PM	host = ubuntu index = test source = /var/log/top-1000.csv sourcetype = csv
7/17/25 100:19:00 PM	999,syndication.twitter.com,com
7/17/25 100:19:00 PM	host = ubuntu index = test source = /var/log/top-1000.csv sourcetype = csv
7/17/25 100:19:00 PM	998,pull-ad-sg01.tiktokcdn.com,com
7/17/25 100:19:00 PM	host = ubuntu index = test source = /var/log/top-1000.csv sourcetype = csv
7/17/25 100:19:00 PM	997,rtb.mafsrver.com,com
7/17/25 100:19:00 PM	host = ubuntu index = test source = /var/log/top-1000.csv sourcetype = csv
7/17/25 100:19:00 PM	996,pull-hls-f58-sg01.tiktokcdn.com,com
7/17/25 100:19:00 PM	host = ubuntu index = test source = /var/log/top-1000.csv sourcetype = csv
7/17/25 100:19:00 PM	995,pull-oxaf-f16-tt02.tiktokcdn.com,com
7/17/25 100:19:00 PM	host = ubuntu index = test source = /var/log/top-1000.csv sourcetype = csv

- Click on an event and check the Selected Fields. You should see fields like rank, domain, and tld.

Top 10 Values	Count	%
com	765	76.65%
net	175	17.53%
io	19	1.904%
org	16	1.603%
ms	5	0.501%
apple	3	0.301%
fi	2	0.2%
goog	2	0.2%
internal	2	0.2%
me	2	0.2%

Getting External Data into Splunk

The screenshot shows a Splunk search interface with the following details:

- Search Bar:** Index=top source="/var/log/top-1000.csv" tld=10
- Event Count:** 19 events (7/17/25 12:00:00.000 AM to 7/17/25 12:48:00.000 PM)
- Event List:** A table with 19 rows, each containing a timestamp, event details, and sourcetype (csv). Some fields are highlighted in yellow.
- Left Panel:** Shows selected fields (host, index, source, sourcetype) and interesting fields (domain, eventtype, linecount, rank, splunk_server, timestamp, tld).

Improve Field Extraction (props.conf)

If Splunk doesn't extract fields correctly, you can force it to treat the file as CSV:

- Go to: `~/splunkforwarder/etc/system/default/`
- Edit `props.conf`:

```
ubuntu@ubuntu:~/splunkforwarder/etc/system/default$ nano props.conf
```

- Add this:

```
[csv]
SHOULD_LINEMERGE = False
puddown_type = true
INDEXED_EXTRACTIONS = csv
KV_MODE = none
category = Structured
description = Comma-separated value format. Set header and other settings in "Delimited Settings"
```

- Check:

The screenshot shows a Splunk search interface with the following details:

- Search Bar:** Index=top source="/var/log/top-1000.csv" tld=10
- Event List:** A table with 19 rows, each containing a timestamp, event details, and sourcetype (csv). Some fields are highlighted in yellow.
- Left Panel:** Shows selected fields (host, index, source, sourcetype) and interesting fields (domain, eventtype, linecount, rank, splunk_server, timestamp, tld).
- Context Menu:** An open context menu over a row where sourcetype is set to CSV. It includes options like "Selected" (which is highlighted), "Reports", "Top values", "Rare values", and "Events with this field".

Handle Multiple Files with Different Sourcetypes

If you have more than one file (e.g., .txt or other formats), give each one a unique sourcetype:

```
ubuntu@ubuntu:~/splunkforwarder/etc/system/default$ head /var/log/top-1000.txt
# These are the top 1000 domains and their top level domains
rank, domain, tld
1, google.com, com
2, microsoft.com, com
3, www.google.com, com
4, netflix.com, com
5, cloud.netflix.com, com
6, prod.cloud.netflix.com, com
7, data.microsoft.com, com
8, ftl.netflix.com, com
```

- Edit inputs.conf again:

```
ubuntu@ubuntu:~/splunkforwarder/etc/apps/top1000_csv_inputs/default$ nano ./inputs.conf
```



```
GNU nano 6.2                               ./inputs.conf *
[monitor:///var/log/top-1000.csv]
index = test
disabled = 0

[monitor:///var/log/top-1000.txt]
sourcetype = top_domains
index = test
crcSalt = splunk4analysts
disabled = 0
```

This helps you search and manage data more easily.

Restart Splunk and Confirm

- Restart the forwarder:

```
ubuntu@ubuntu:~/splunkforwarder$ sudo ./bin/splunk restart
[sudo] password for ubuntu:
Warning: Attempting to revert the SPLUNK_HOME ownership
Warning: Executing "chown -R ubuntu:ubuntu /home/ubuntu/splunkforwarder"
Stopping splunkd...
Shutting down. Please wait, as this may take a few minutes.

Stopping splunk helpers...

Done.
```

Getting External Data into Splunk

- Go back to Splunk Web and search:

New Search

Index= test

2,071 events (7/17/25 100:00:00 PM to 7/18/25 107:56:00 PM) No Event Sampling ▾

Events (2,071) Patterns Statistics Visualization

Timeline format ▾ Zoom Out ▾ Zoom to Selection X Deselected

Format ▾ Show 20 Per Page ▾ View: List ▾

sourcetype

3 Values, 100% of events Selected Yes No

Reports

Top values Top values by time Rare values

Events with this field

Values	Count	%
CSV	1,000	48.28%
top_domains	1,000	48.28%
linux_secure	71	3.42%

12:33:00:00 PM host = ubuntu index = test source = /var/log/auth.log sourcetype = linux_secure
 7/18/25 12:33:00 ubuntu ssh[1478]: Accepted password for ubuntu from 192.168.1.8 port 49938 ssh2
 12:33:00:00 PM host = ubuntu index = test source = /var/log/auth.log sourcetype = linux_secure
 7/18/25 12:32:07:00 ubuntu systemd-logind[885]: Removed session 3.
 12:32:07:00 PM host = ubuntu index = test source = /var/log/auth.log sourcetype = linux_secure

New Search

Index= test sourcetype=top_domains

1,000 events (7/17/25 100:00:00 PM to 7/18/25 110:22:00 PM) No Event Sampling ▾

Events (1,000) Patterns Statistics Visualization

Timeline format ▾ Zoom Out ▾ Zoom to Selection X Deselected

Format ▾ Show 20 Per Page ▾ View: List ▾

Time Event

7/18/25 10:45:21:00 AM	1000,pull-hls-f77-sg01.tiktokcdn.com.com host = ubuntu index = test source = /var/log/top-1000.txt sourcetype = top_domains
7/18/25 10:45:21:00 AM	999,syndication.twitter.com.com host = ubuntu index = test source = /var/log/top-1000.txt sourcetype = top_domains
7/18/25 10:45:21:00 AM	998,pull-as1-sg01.tiktokcdn.com.com host = ubuntu index = test source = /var/log/top-1000.txt sourcetype = top_domains
7/18/25 10:45:21:00 AM	997,rtb.mafisrvr.com.com host = ubuntu index = test source = /var/log/top-1000.txt sourcetype = top_domains
7/18/25 10:45:21:00 AM	996,pull-hls-f58-sg01.tiktokcdn.com.com host = ubuntu index = test source = /var/log/top-1000.txt sourcetype = top_domains
7/18/25 10:45:21:00 AM	995,pull-ccaf-f16-tt82.tiktokcdn.com.com host = ubuntu index = test source = /var/log/top-1000.txt sourcetype = top_domains
7/18/25 10:45:21:00 AM	994,pull-ccaf-f18-sg01.tiktokcdn.com.com host = ubuntu index = test source = /var/log/top-1000.txt sourcetype = top_domains

- Check if fields like tld are extracted correctly. You can even explore top values to confirm data quality.

New Search

Index= test

1,000 events (7/17/25 100:00:00 PM to 7/18/25 110:22:00 PM) No Event Sampling ▾

Events (1,000) Patterns Statistics Visualization

Timeline format ▾ Zoom Out ▾ Zoom to Selection X Deselected

Format ▾ Show 20 Per Page ▾ View: List ▾

tld

15 Values, 99.8% of events Selected Yes No

Reports

Top values Top values by time Rare values

Events with this field

Top 10 Values	Count	%
com	765	76.653%
net	175	17.535%
io	19	1.904%
org	16	1.603%
ms	5	0.501%
apple	3	0.301%
fi	2	0.2%
goog	2	0.2%
internal	2	0.2%
se	2	0.2%

Custom Data Sources and the "Great 8" Best Practices

When you're working with custom logs, like niche software, in-house apps, or unique formats, Splunk won't have a ready-made add-on. That's where custom sourcetypes and Technical Add-ons (TAs) come in.

To make sure Splunk parses your data correctly, you need to configure how events are broken, timestamped, and indexed. This is done using the "Great 8" parsing settings in props.conf.

The "Great 8" Parsing Settings Explained

Here are the 8 key settings you should always define when creating a custom sourcetype:

1. **TIME_PREFIX**

- Tells Splunk where the timestamp starts.
- Example: timestamp=" if your log looks like timestamp="2023-10-26..."

2. **TIME_FORMAT**

- Defines the format of the timestamp.
- Example: %Y-%m-%d %H:%M:%S for 2023-10-26 14:30:00

3. **MAX_TIMESTAMP_LOOKAHEAD**

- How many characters Splunk should scan after TIME_PREFIX to find the timestamp.

4. **SHOULD_LINEMERGE**

- Set to false if each log entry is one line. Helps performance.

5. **LINE_BREAKER**

- Regex that defines where one event ends and the next begins.
- Example: (\\r\\n)+ for new lines.

6. **TRUNCATE**

- Max size of a single event (in bytes). Prevents huge events from being indexed.

7. **EVENT_BREAKER**

- Lets Splunk break events at the Universal Forwarder level (not just indexer).

8. **EVENT_BREAKER_ENABLE**

- Enables the above setting. Set to true.

Example: Onboarding Ubuntu dpkg.log

Let's say you want to ingest /var/log/dpkg.log from an Ubuntu system.

1. Create App Directory

```
ubuntu@ubuntu:~/splunkforwarder/etc/apps$ mkdir -p dpkg/default  
ubuntu@ubuntu:~/splunkforwarder/etc/apps$ cd dpkg/default/
```

2. Define Parsing Rules in props.conf

```
ubuntu@ubuntu:~/splunkforwarder/etc/apps/dpkg/default$ nano props.conf
```

- **props.conf Content for dpkg Logs**

```
GNU nano 6.2                                         props.conf *
```

```
[ubuntu_dpkg_log]  
TIME_PREFIX = ^  
TIME_FORMAT = %Y-%m-%d %H:%M:%S  
MAX_TIMESTAMP_LOOKAHEAD = 20  
SHOULD_LINEMERGE = false  
LINE_BREAKER = (\r\n)+  
TRUNCATE = 2000  
EVENT_BREAKER = (\r\n)+  
EVENT_BREAKER_ENABLE = true
```

This configuration explicitly sets several key parsing parameters for the ubuntu_dpkg_log sourcetype:

- **TIME_PREFIX = ^**: Indicates the timestamp is expected at the beginning of each event.
- **TIME_FORMAT = %Y-%m-%d %H:%M:%S**: Specifies the timestamp format (e.g., 2023-10-26 14:30:00).
- **MAX_TIMESTAMP_LOOKAHEAD = 20**: Sets the maximum characters Splunk will scan for the timestamp after the TIME_PREFIX.
- **SHOULD_LINEMERGE = false**: Prevents multiple lines from being combined into a single event, indicating each log entry is a single line.
- **LINE_BREAKER = (\r\n)+**: Defines that a new line character signifies a new event, consistent with SHOULD_LINEMERGE = false.
- **TRUNCATE = 2000**: Sets the maximum event size to 2000 bytes, acting as a safeguard against overly large events.
- **EVENT_BREAKER = (\r\n)+**: Enables event breaking at the Universal Forwarder level based on new lines.
- **EVENT_BREAKER_ENABLE = true**: Activates the EVENT_BREAKER functionality on the forwarder.

3. Configure Input in inputs.conf

```
ubuntu@ubuntu:~/splunkforwarder/etc/apps/dpkg/default$ nano inputs.conf
```

```
GNU nano 6.2                                         inputs.conf *
[monitor:///var/log/dpkg.log]
sourcetype = dpkg
index = test
disabled = 0
```

This configuration specifies:

- **[monitor:///var/log/dpkg.log]:** Instructs Splunk to monitor the file /var/log/dpkg.log.
- **sourcetype = dpkg:** Assigns the dpkg sourcetype to the data from this log file, ensuring it uses the parsing rules defined in props.conf.
- **index = test:** Directs the ingested data to the test index in Splunk.
- **disabled = 0:** Ensures this input is active (0 means enabled).

4. Restart the Forwarder

```
ubuntu@ubuntu:~/splunkforwarder$ sudo ./bin/splunk restart
[sudo] password for ubuntu:
Warning: Attempting to revert the SPLUNK_HOME ownership
Warning: Executing "chown -R ubuntu:ubuntu /home/ubuntu/splunkforwarder"
Stopping splunkd...
Shutting down. Please wait, as this may take a few minutes.

Stopping splunk helpers...

Done.
splunkd.pid doesn't exist...

Splunk> 4TW

Checking prerequisites...
    Checking mgmt port [8089]: open
    Checking conf files for problems...
        Done
    Checking default conf files for edits...
        Validating installed files against hashes from '/home/ubuntu/splunkforwarder/splunkforwarder-9.4.3-237ebbd22314-linux-amd64-manifest'
            All installed files intact.
            Done
All preliminary checks passed.

Starting splunk server daemon (splunkd)...
Done
```

Getting External Data into Splunk

5. Verify in Splunk Web

- Search:

The screenshot shows the Splunk Web interface with the search bar set to "Index:test sourcetype=dpkg". The results table displays 145 events from July 20, 2025, at 9:08:39 AM. The table has columns for Time, Event, and several other fields like host, index, source, and sourcetype. The sourcetype is consistently listed as "dpkg". The interface includes navigation buttons for previous/next pages and a "Format" dropdown.

Time	Event
2025-07-20 09:08:22	status installed man-db:amd64 2.1.0-2+1 host = ubuntu index = TEST source = /var/log/dpkg.log sourcetype = dpkg
2025-07-20 09:08:22	status half-configured man-db:amd64 2.1.0-2+1 host = ubuntu index = TEST source = /var/log/dpkg.log sourcetype = dpkg
2025-07-20 09:08:22	triggered man-db:amd64 2.1.0-2+1 <none> host = ubuntu index = TEST source = /var/log/dpkg.log sourcetype = dpkg
2025-07-20 09:08:22	status installed net-tools:amd64 1.60-120181103.0eebece-1ubuntu5.4 host = ubuntu index = TEST source = /var/log/dpkg.log sourcetype = dpkg
2025-07-20 09:08:22	status unpacked net-tools:amd64 1.60-120181103.0eebece-1ubuntu5.4 host = ubuntu index = TEST source = /var/log/dpkg.log sourcetype = dpkg
2025-07-20 09:08:22	configure net-tools:amd64 1.60-120181103.0eebece-1ubuntu5.4 <none> host = ubuntu index = TEST source = /var/log/dpkg.log sourcetype = dpkg
2025-07-20 09:08:22	startup packages configure host = ubuntu index = TEST source = /var/log/dpkg.log sourcetype = dpkg
2025-07-20 09:08:22	status unpacked net-tools:amd64 1.60-120181103.0eebece-1ubuntu5.4 host = ubuntu index = TEST source = /var/log/dpkg.log sourcetype = dpkg

You should see clean, timestamped events from dpkg.log, parsed using your custom sourcetype.

Extracting Data Fields using the EXTRACT Command

Splunk is powerful because it doesn't just store logs, it makes data searchable. One of the key ways it does this is through field extraction, which turns raw log lines into structured, searchable fields.

What Is Search-Time Field Extraction?

Unlike index-time extraction (which happens when data first enters Splunk), search-time extraction happens when you run a search. This means:

- You can define new fields anytime, even after data is already indexed.
- No need to re-ingest or re-index your data.
- Super flexible, perfect for evolving log formats or changing analysis needs.

Things to Keep in Mind

- Search-time extraction uses CPU on the Search Head every time you run a query.
- In large environments, make sure field extraction configs are consistent across all Search Heads, especially in clusters.

How to Extract Fields in Splunk

Splunk offers three main methods to extract fields:

1. EXTRACT (Inline in props.conf)

- Define regex directly in props.conf under your sourcetype.
- Best for simple to medium complexity extractions.

2. REPORT (Using transforms.conf)

- Use props.conf + transforms.conf for more complex logic or reusable regex.
- Great for shared rules across multiple sourcetypes or when using lookups.

3. Automatic Field Discovery

Splunk can auto-extract fields from structured formats like:

- Key-Value pairs: user=jsmith status=success
- JSON: { "user": "jsmith", "status": "success" }
- XML: <user>jsmith</user>

No config needed, Splunk does the work for you!

Step-by-Step: Extracting Fields via UI

```
splunk@splunk:~$ tail /var/log/fail2ban.log
2025-07-21 11:52:08,288 fail2ban.jail      [69865]: INFO  Creating new jail 'sshd'
2025-07-21 11:52:08,305 fail2ban.jail      [69865]: INFO  Jail 'sshd' uses pyinotify {}
2025-07-21 11:52:08,311 fail2ban.jail      [69865]: INFO  Initiated 'pyinotify' backend
2025-07-21 11:52:08,313 fail2ban.filter    [69865]: INFO  maxLines: 1
2025-07-21 11:52:08,340 fail2ban.filter    [69865]: INFO  maxRetry: 5
2025-07-21 11:52:08,340 fail2ban.filter    [69865]: INFO  findtime: 600
2025-07-21 11:52:08,340 fail2ban.actions   [69865]: INFO  banTime: 600
2025-07-21 11:52:08,341 fail2ban.filter    [69865]: INFO  encoding: UTF-8
2025-07-21 11:52:08,342 fail2ban.filter    [69865]: INFO  Added logfile: '/var/log/auth.log' (pos = 0, has
h = 4f91f2b22f109ae4b6c9210830dc14ed35ffffa9f)
2025-07-21 11:52:08,372 fail2ban.jail      [69865]: INFO  Jail 'sshd' started
splunk@splunk:~$ cd ./splunk/etc/apps/fail2ban/default/
splunk@splunk:~/splunk/etc/apps/fail2ban/default$ cat props.conf
[fail2ban_logs]
TIME_PREFIX = ^
TIME_FORMAT = %Y-%m-%d %H:%M:%S,%3N
MAX_TIMESTAMP_LOOKAHEAD = 24
SHOULD_LINEMERGE = false
LINE_BREAKER = ([\r\n]+)
TRUNCATE = 10000
EVENT_BREAKER = ([\r\n]+)
EVENT_BREAKER_ENABLE = true
```

1. Search your data

Example:

	Time	Event
SELECTED FIELDS	7/21/25 11:52:08.372 AM	host = splunk index = test source = /var/log/fail2ban.log sourcetype = fail2ban
INTERESTING FIELDS	7/21/25 11:52:08.340 AM	host = splunk index = test source = /var/log/fail2ban.log sourcetype = fail2ban
	7/21/25 11:52:08.341 AM	host = splunk index = test source = /var/log/fail2ban.log sourcetype = fail2ban
	7/21/25 11:52:08.340 AM	host = splunk index = test source = /var/log/fail2ban.log sourcetype = fail2ban
	7/21/25 11:52:08.340 AM	host = splunk index = test source = /var/log/fail2ban.log sourcetype = fail2ban
	7/21/25 11:52:08.340 AM	host = splunk index = test source = /var/log/fail2ban.log sourcetype = fail2ban
	7/21/25 11:52:08.340 AM	host = splunk index = test source = /var/log/fail2ban.log sourcetype = fail2ban
	7/21/25 11:52:08.340 AM	host = splunk index = test source = /var/log/fail2ban.log sourcetype = fail2ban

2. Click "Event Actions" → "Extract Fields"

Build Event Type	Value	Actions
Extract Fields	host	<input type="button" value="▼"/>
Show Source	source	<input checked="" type="checkbox"/> <input type="button" value="▼"/>
	/var/log/fail2ban.log	<input type="button" value="▼"/>
Event	eventtype	<input checked="" type="checkbox"/> <input type="button" value="▼"/>
	nix-all-logs	<input type="button" value="▼"/>
Time	_time	<input type="button" value="▼"/>
Default	linecount	<input checked="" type="checkbox"/> <input type="button" value="▼"/>
	1	<input type="button" value="▼"/>
	punct	<input checked="" type="checkbox"/> <input type="button" value="▼"/>
	" "	<input type="button" value="▼"/>
	splunk_server	<input checked="" type="checkbox"/> <input type="button" value="▼"/>
	splunk	<input type="button" value="▼"/>

3. Choose “Regular Expression” method

Extract Fields

Select Method Select Fields Validate Save **Next >**

Select Method

Indicate the method you want to use to extract your field(s). Learn more [?](#)

I prefer to write the regular expression myself >

Source type
fail2ban

2025-07-21 11:52:08,372 fail2ban.jail [69865]: INFO Jail 'sshd' started

Regular Expression

(.*?)

Splunk Enterprise will extract fields using a Regular Expression.

Delimiters

x|y|z

Splunk Enterprise will extract fields using a delimiter (such as commas, spaces, or characters). Use this method for delimited data like comma separated values (CSV files).

4. Highlight the value you want to extract

Example: highlight INFO to create a field called log_level.

5. Name the field and click “Add Extraction”

Extract Fields

Select Method **Select Fields** Validate Save < Back **Next >**

Select Fields

Highlight one or more values in the sample event to create fields. You can indicate one value is required, meaning it must exist in an event for the regular expression to match. Click on highlighted values in the sample event to modify them. To highlight text that is already part of an existing extraction, first turn off the existing extractions. Learn more [?](#)

2025-07-21 11:52:08,372 fail2ban.jail [69865]: **INFO** Jail 'sshd' started

Extract	Require
Field Name	log_level
Sample Value	INFO
Add Extraction	

Getting External Data into Splunk

6. Preview and refine the regex if needed

Preview
If you see incorrect results below, click an additional event to add it to the set of sample events. Highlight its values to improve the extraction. You can remove incorrect values in the next step.

Events log_level

✓ 15 events (4/22/25 12:00:00.000 AM to 7/21/25 10:06:00.000 PM) 20 per page *

filter Apply Sample: 1,000 events ▾ All events ▾ All Events Matches Non-Matches

_raw	log_level
✓ 2025-07-21 11:52:08,372 fail2ban.jail [69865]: INFO Jail 'sshd' started	INFO
✓ 2025-07-21 11:52:08,342 fail2ban.filter [69865]: INFO Added logfile: '/var/log/auth.log' (pos = 0, hash = 4f91f2b22f109ae4b6c9210830dc14ed35ffffa9f)	INFO
✓ 2025-07-21 11:52:08,341 fail2ban.filter [69865]: INFO encoding: UTF-8	INFO
✓ 2025-07-21 11:52:08,340 fail2ban.actions [69865]: INFO banTime: 600	INFO
✓ 2025-07-21 11:52:08,340 fail2ban.filter [69865]: INFO findtime: 600	INFO
✓ 2025-07-21 11:52:08,340 fail2ban.filter [69865]: INFO maxRetry: 5	INFO
✓ 2025-07-21 11:52:08,313 fail2ban.filter [69865]: INFO maxLines: 1	INFO
✓ 2025-07-21 11:52:08,311 fail2ban.jail [69865]: INFO Initiated 'pyinotify' backend	INFO
✓ 2025-07-21 11:52:08,305 fail2ban.jail [69865]: INFO Jail 'sshd' uses pyinotify ()	INFO
✓ 2025-07-21 11:52:08,288 fail2ban.jail [69865]: INFO Creating new jail 'sshd'	INFO
✓ 2025-07-21 11:52:08,288 fail2ban.database [69865]: WARNING New database created. Version '4'	WARNING
✓ 2025-07-21 11:52:08,280 fail2ban.database [69865]: INFO Connected to fail2ban persistent database '/var/lib/fail2ban/fail2ban.sqlite3'	INFO
✓ 2025-07-21 11:52:08,267 fail2ban.observer [69865]: INFO Observer start...	INFO
✓ 2025-07-21 11:52:08,266 fail2ban.server [69865]: INFO Starting Fail2ban v0.11.2	INFO
✓ 2025-07-21 11:52:08,266 fail2ban.server [69865]: -----	INFO

Select Fields
Highlight one or more values in the sample event to create fields. You can indicate one value is required, meaning it must exist in an event for the regular expression to match. Click on highlighted values in the sample event to modify them. To highlight text that is already part of an existing extraction, first turn off the existing extractions. Learn more ↗

2025-07-21 11:52:08,372 fail2ban.jail [69865]: INFO Jail 'sshd' started

Hide Regular Expression ▾ View in Search ↗

[\n]*\n*\n*([Pp]log_level)\n

Edit the Regular Expression

Preview
If you see incorrect results below, click an additional event to add it to the set of sample events. Highlight its values to improve the extraction. You can remove incorrect values in the next step.

Events log_level

✓ 15 events (4/22/25 12:00:00.000 AM to 7/21/25 10:06:00.000 PM) 20 per page *

7. Save the extraction and set permissions

Extract Fields < Back Existing fields ▾

⚠️ If you manually edit and then preview the regular expression below, you cannot return to the automatic field extraction workflow.

Use the event listing below to validate the field extractions produced by your regular expression.

Regular Expression [Regular Expression Reference](#) [View in Search](#)

[\n]*\n*\n*([Pp]log_level)\n

Events log_level **Save**

✓ 15 events (4/22/25 12:00:00.000 AM to 7/21/25 10:06:00.000 PM) 20 per page *

filter Apply Sample: 1,000 events ▾ All events ▾ All Events Matches Non-Matches

Getting External Data into Splunk

8. Explore your new field in search results

Success!

You have extracted additional fields from your data (sourcetype=fail2ban).

Edit your field extractions at any time by going to [Field Extractions](#).

What would you like to do next?

- [Explore the fields I just created in Search](#) 
- [Extract more fields](#)

New Search

Save As ▾ Create Table View Close

Index=_* OR index** sourceType=fail2ban

Last 24 hours ▾

17 events (7/20/25 100:00:00 PM to 7/21/25 11:59:00 PM) No Event Sampling ▾

Job ▾ II III A Smart Mode ▾

Events (17) Patterns Statistics Visualization

Timeline format ▾ Zoom Out ▾ Zoom to Selection ▾ Deselect ▾

1 hour per column

Format ▾ Show: 20 Per Page ▾ View List ▾

Hide Fields All Fields

	Time	Event
Selected Fields	host 1 index 1 source 1 sourcetype 1	> 7/21/25 12:51:59.312 fail2ban.filter [69865]: WARNING [sshd] Please check jail has possibly a timezone issue. Line with odd timestamp: Jul 21 12:17:26 splunk mongod: looking for plugins in '/home/build/build-home/opt/mongo/lib/sasl2'. failed to open directory, error: No such file or directory host = splunk index = test source = /var/log/fail2ban.log sourcetype = fail2ban
Interesting Fields	#date_hour 2 #date_minute 2 #date_second 1 #date_second_2 #date_wday 1 #date_year 1	> 7/21/25 12:51:59.307 PM fail2ban.filter [69865]: WARNING [sshd] Simulate NOW in operation since found time has too large deviation 1753100246.0 - 1753102319.3058023 +/- 68 host = splunk index = test source = /var/log/fail2ban.log sourcetype = fail2ban
	> 7/21/25 11:52:08.372 fail2ban.jail [69865]: INFO jail 'sshd' started host = splunk index = test source = /var/log/fail2ban.log sourcetype = fail2ban	
	> 7/21/25 11:52:08.372 fail2ban.filter [69865]: INFO Added logfile: '/var/log/auth.log' (pos = 0, hash = 4f91f2b22f189ae4b6c921883dc14ed3fffffa1f) host = splunk index = test source = /var/log/fail2ban.log sourcetype = fail2ban	
	> 7/21/25 11:52:08.342 AM fail2ban.filter [69865]: INFO encoding: UTF-8 host = splunk index = test source = /var/log/fail2ban.log sourcetype = fail2ban	

7/21/25 12:51:59.312 PM 2025-07-21 12:51:59.312 fail2ban.filter [69865]: WARNING [sshd] Please check jail has possibly a timezone issue. Line with odd timestamp: Jul 21 12:17:26 splunk mongod: looking for plugins in '/home/build/build-home/opt/mongo/lib/sasl2'. failed to open directory, error: No such file or directory

Event Actions ▾

Type	Field	Value	Actions
Selected	host	splunk	▼
	index	test	▼
	source	/var/log/fail2ban.log	▼
	sourcetype	fail2ban	▼
Event	eventtype	nix-all-logs	▼
		nix_errors {error}	▼
		nix_to_data	▼
	log_level	WARNING	▼
	tag	error	▼
Time	_time	2025-07-21T12:51:59.312+00:00	
Default	linecount	1	▼
	punct	-.-.---- --- --- --- ---	▼
	splunk_server	splunk	▼

Extracting Data Fields using the REPORT Command

Why Use REPORT for Field Extraction in Splunk?

When working with custom log formats in Splunk, like **fail2ban logs** or other non-standard sources, you often need to extract fields using regular expressions (regex). Splunk gives you two main ways to do this: EXTRACT and **REPORT**.

While EXTRACT is great for quick setups, **REPORT** is the better choice when you want clean, reusable, and scalable configurations.

Benefits of Using REPORT

1. Regex Reusability

Define your regex once in transforms.conf and reuse it across multiple sourcetypes in props.conf.

2. Modular Configuration

Keeps your regex logic separate from sourcetype definitions — easier to manage and troubleshoot.

3. Advanced Field Mapping

Use FORMAT to assign names to regex capture groups, making your field extraction more readable and precise.

Example: Extracting Fields from fail2ban Logs

1. Define the Extraction in transforms.conf

```
splunk@splunk:~/splunk/etc/apps/fail2ban/default$ nano transforms.conf
```

```
GNU nano 6.2                                     transforms.conf
[fail2ban_fields]
REGEX = ^(?:<timestamp>\d{4}-\d{2}-\d{2}\s+\d{2}:\d{2}:\d{2},\d{3})+
FORMAT = timestamp::$1 component::$2 pid::$3 level::$4 jail::$5 message::$6
```

- This regex captures key parts of the log and maps them to field names like timestamp.

2. Apply the Extraction in props.conf

```
splunk@splunk:~/splunk/etc/apps/fail2ban/default$ nano props.conf
```

```
GNU nano 6.2                                         props.conf
[fail2ban]
SHOULD_LINEMERGE = false
REPORT-fail2ban = fail2ban_fields
TIME_PREFIX = ^
TIME_FORMAT = %Y-%m-%d %H:%M:%S,%3N
MAX_TIMESTAMP_LOOKAHEAD = 24
```

- This tells Splunk to apply the fail2ban_fields extraction to all events with sourcetype=fail2ban.

3. Restart Splunk to Apply Changes

```
splunk@splunk:~/splunk/etc/apps/fail2ban/default$ sudo ../../../../../../bin/splunk restart
```

- Note: The relative path is used because you're deep inside the app directory.

4. Verify in Splunk Search

Search:

The screenshot shows a Splunk search interface with the following details:

- Search Bar:** Index=test sourcetype=fail2ban
- Results Summary:** 79 events (before 11/3/25 6:26:08.000 PM) - No Event Sampling
- Event List:** One event is visible in the main pane:


```
2025-07-28 19:46:33.925 fail2ban.filter [875]: WARNING [sshd] Please check jail has possibly a timezone issue. Line with odd timestamp: Jul 28 19:40:29 splunk mongo
d: looking for plugins in '/home/build/build-home/opt/mongo/lib/sasl2'.
Failed to open directory, error: No such file or directory
```
- Event Actions Panel:** Shows the event's timestamp (2025-07-28 19:46:33.925) highlighted with a green box.
- Selected Fields Sidebar:** Shows fields like host, source, sourcetype, and timestamp.
- Interesting Fields Sidebar:** Shows various date-related fields.

- You should now see fields like timestamp in the left sidebar under Selected Fields, confirming that your extraction is working!

