Splunk: Basics

Task 1 Introduction

Splunk is one of the leading SIEM solutions in the market that provides the ability to collect, analyze and correlate the network and machine logs in real-time. In this room, we will explore the basics of Splunk and its functionalities and how it provides better visibility of network activities and help in speeding up the detection.

Learning Objective and Pre-requisites

If you are new to SIEM, please complete the [Introduction to SIEM]. This room covers the following learning objectives:

- Splunk overview
- Splunk components and how they work
- Different ways to ingest logs
- Normalization of logs

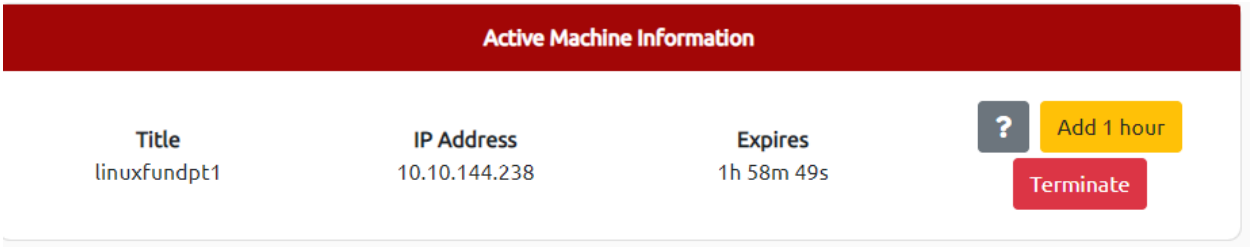
Task 2 Connect with the Lab

Room Machine

Before moving forward , simply press the green button on the top-right of this

task indicated by the arrow on the right:

Once deployed, a card will appear at the top of the room, showing the IP address assigned to the Machine.



Active Machine Information

Title	IP Address	Expires
linuxfundpt1	10.10.144.238	1h 58m 49s

Splunk Instance can be accessed by copy and pasting the MACHINE_IP into the web browser on the AttackBox, or via the VPN at . The machine will take up to 3-5 minutes to start.

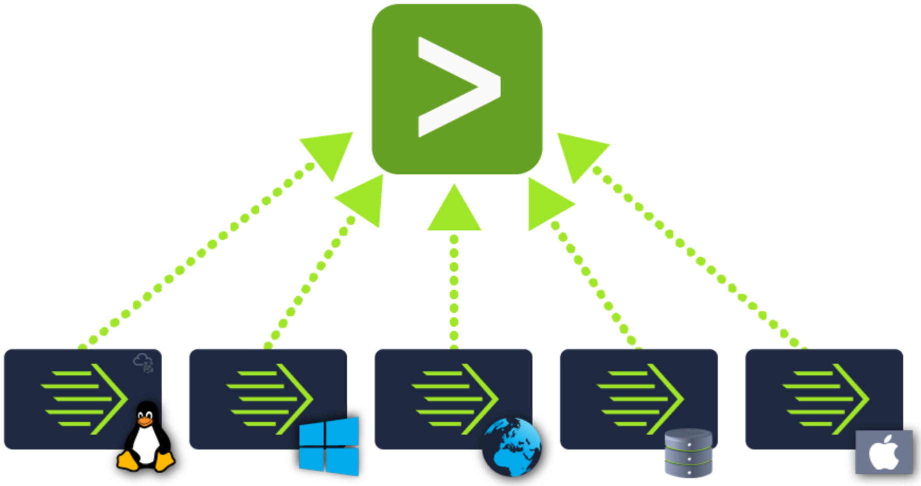
Task 3 Splunk Components

Splunk has three main components, namely Forwarder, Indexer, and Search Head. These components are explained below:

Splunk Forwarder

Splunk Forwarder is a lightweight agent installed on the endpoint intended to be monitored, and its main task is to collect the data and send it to the Splunk instance. It does not affect the endpoint's performance as it takes very few resources to process. Some of the key data sources are:

- Web server generating web traffic.
- Windows machine generating Windows Event Logs, PowerShell, and Sysmon data.
- Linux host generating host-centric logs.
- Database generating DB connection requests, responses, and errors.

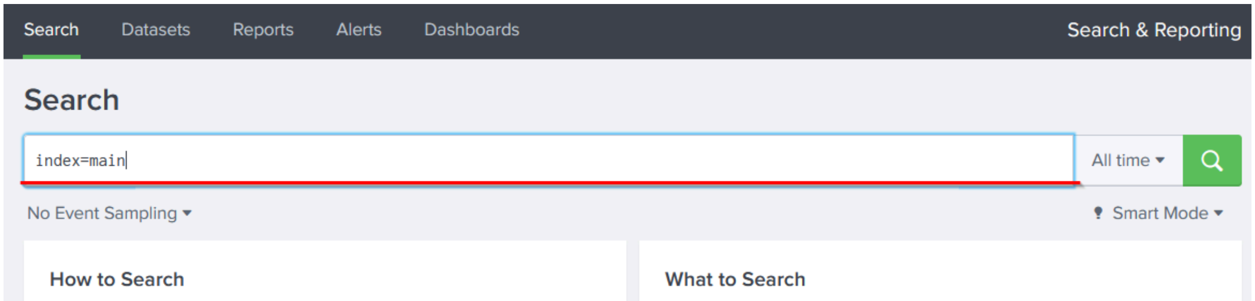


Splunk Indexer

Splunk Indexer plays the main role in processing the data it receives from forwarders. It takes the data, normalizes it into field-value pairs, determines the datatype of the data, and stores them as events. Processed data is easy to search and analyze.

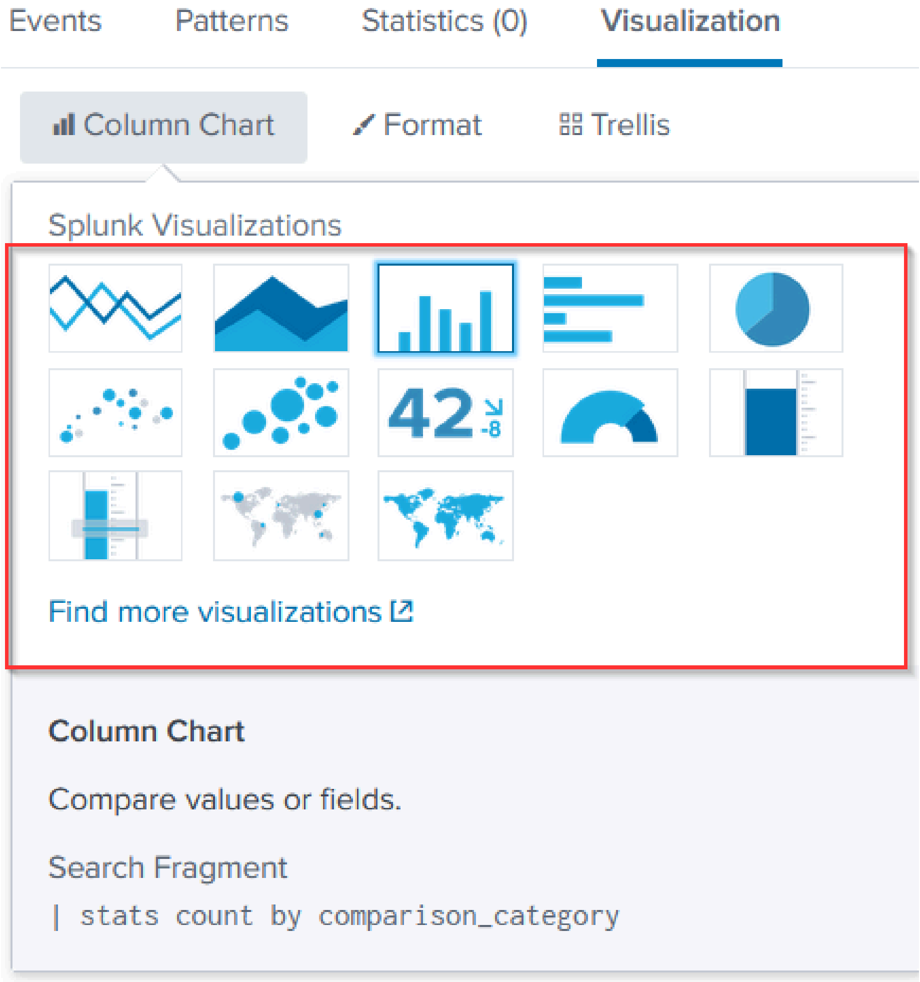
Search Head

Splunk Search Head is the place within the Search & Reporting App where users can search the indexed logs as shown below. When the user searches for a term or uses a Search language known as Splunk Search Processing Language, the request is sent to the indexer and the relevant events are returned in the form of field-value pairs.



```
index=main
```

Search Head also provides the ability to transform the results into presentable tables, visualizations like pie-chart, bar-chart and column-chart, as shown below:



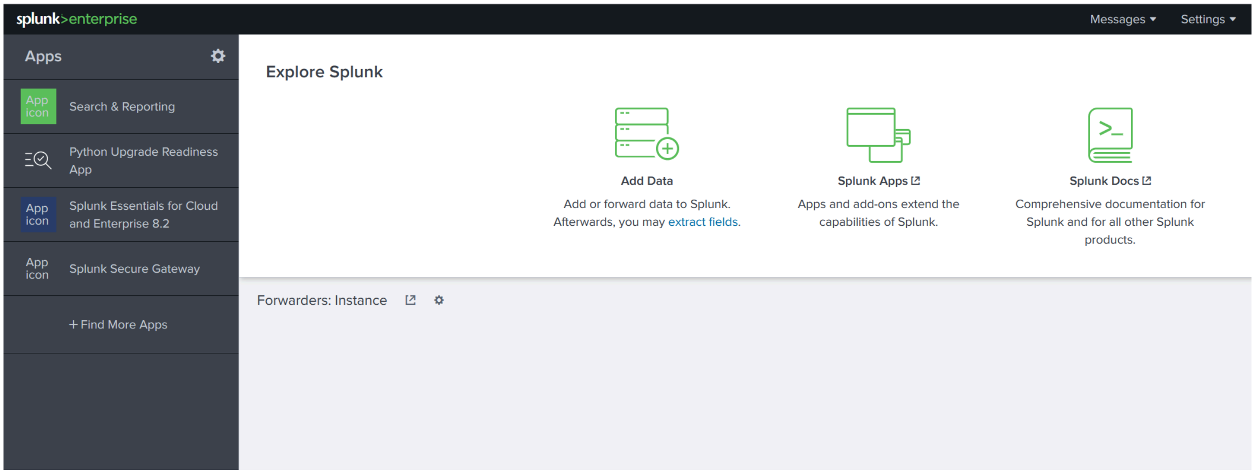
```
| stats count by comparison_category
```

Answer the questions below
Which component is used to collect and send data over the Splunk instance?
Forwarder

Task 4 **Navigating Splunk**

Splunk Bar

When you access Splunk, you will see the default home screen identical to the screenshot below.

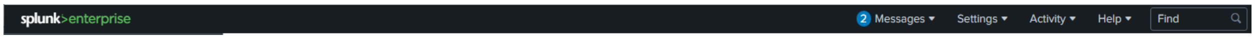


Add or forward data to Splunk. Afterwards, you may extract fields.

Apps and add-ons extend the capabilities of Splunk.

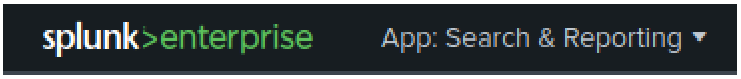
Comprehensive documentation for Splunk and for all other Splunk products.

Let's look at each section, or panel, that makes up the home screen. The top panel is the **Splunk Bar** (below image).



In the Splunk Bar, you can see system-level messages (**Messages**), configure the Splunk instance (**Settings**), review the progress of jobs (**Activity**), miscellaneous information such as tutorials (**Help**), and a search feature (**Find**).

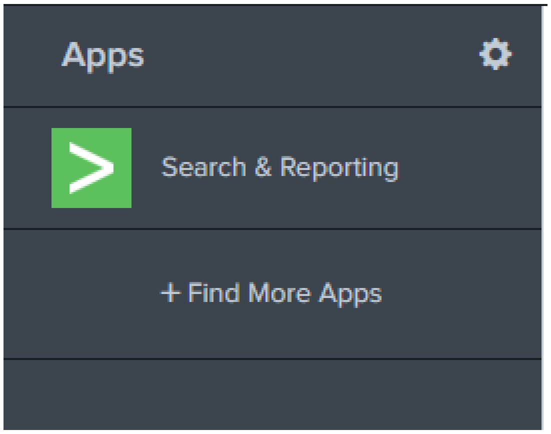
The ability to switch between installed Splunk apps instead of using the **Apps panel** can be achieved from the Splunk Bar, like in the image below.



Apps Panel

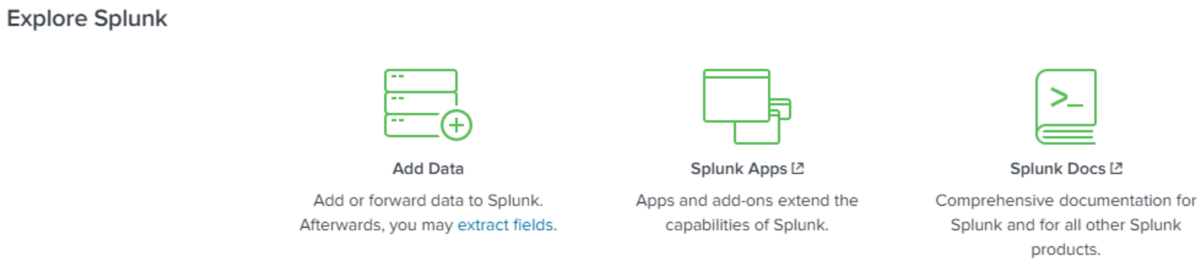
Next is the **Apps Panel** . In this panel, you can see the apps installed for the Splunk instance.

The default app for every Splunk installation is **Search & Reporting** .



Explore Splunk

The next section is **Explore Splunk** . This panel contains quick links to add data to the Splunk instance, add new Splunk apps, and access the Splunk documentation.



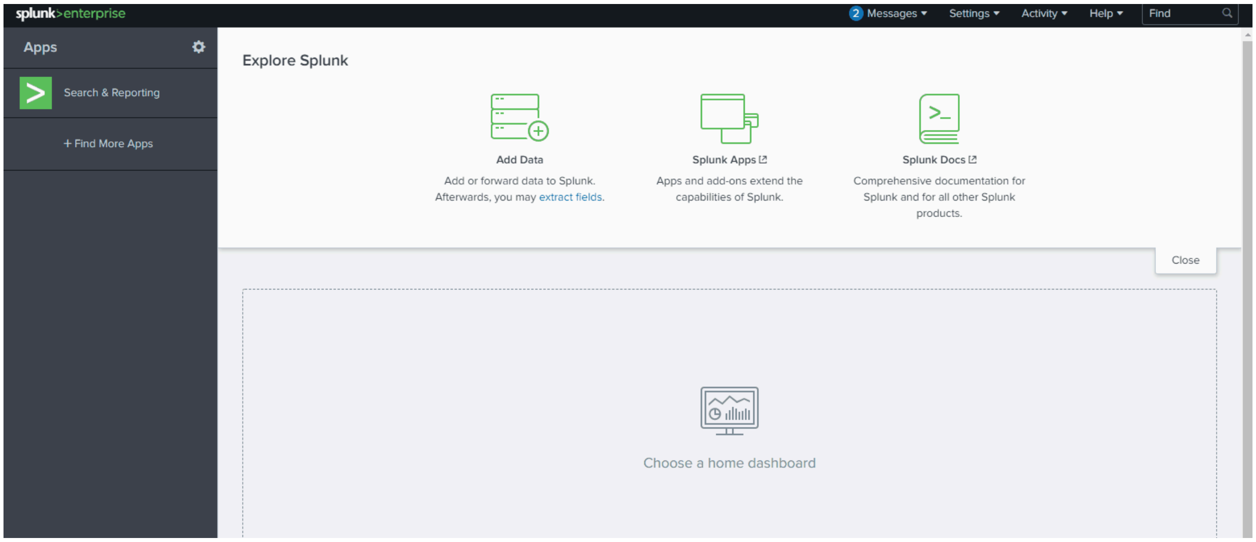
Add or forward data to Splunk. Afterwards, you may extract fields.

Apps and add-ons extend the capabilities of Splunk.

Comprehensive documentation for Splunk and for all other Splunk products.

Splunk Dashboard

The last section is the **Home Dashboard** . By default, no dashboards are displayed. You can choose from a range of dashboards readily available within your Splunk instance. You can select a dashboard from the dropdown menu or by visiting the **dashboards listing page** .



Add or forward data to Splunk. Afterwards, you may extract fields.

Apps and add-ons extend the capabilities of Splunk.

Comprehensive documentation for Splunk and for all other Splunk products.

Choose a home dashboard

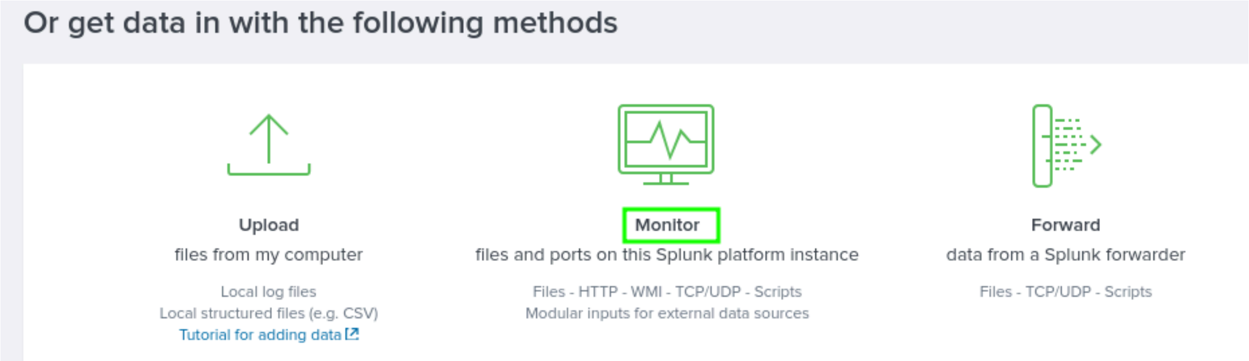
You can also create dashboards and add them to the Home Dashboard. The dashboards you create can be viewed isolated from the other dashboards by clicking on the **Yours** tab.

Please review the Splunk documentation on Navigating Splunk [here](#).

Answer the questions below

In the Add Data tab, which option is used to collect data from files and ports?

Go the home page -> add data
Here we can see the Monitor option



Or get data in with the following methods

Upload
files from my computer
Local log files
Local structured files (e.g. CSV)
Tutorial for adding data

Monitor
files and ports on this Splunk platform instance
Files - HTTP - WMI - TCP/UDP - Scripts
Modular inputs for external data sources

Forward
data from a Splunk forwarder
Files - TCP/UDP - Scripts

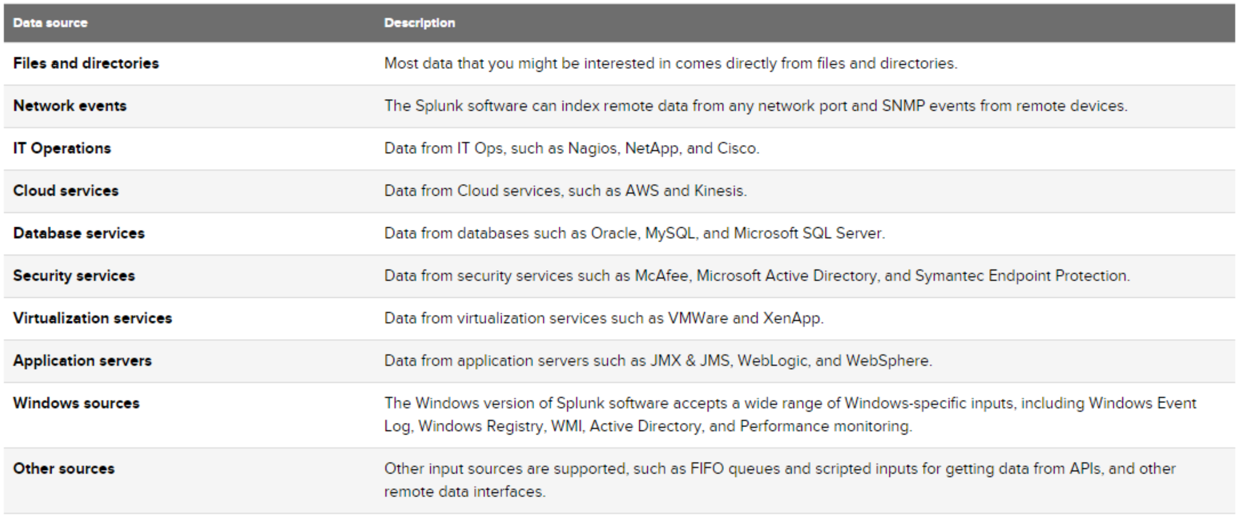
Monitor

Task 5 Adding Data

Splunk can ingest any data. As per the Splunk documentation, when data is added to Splunk, the data is processed and transformed into a series of individual events.

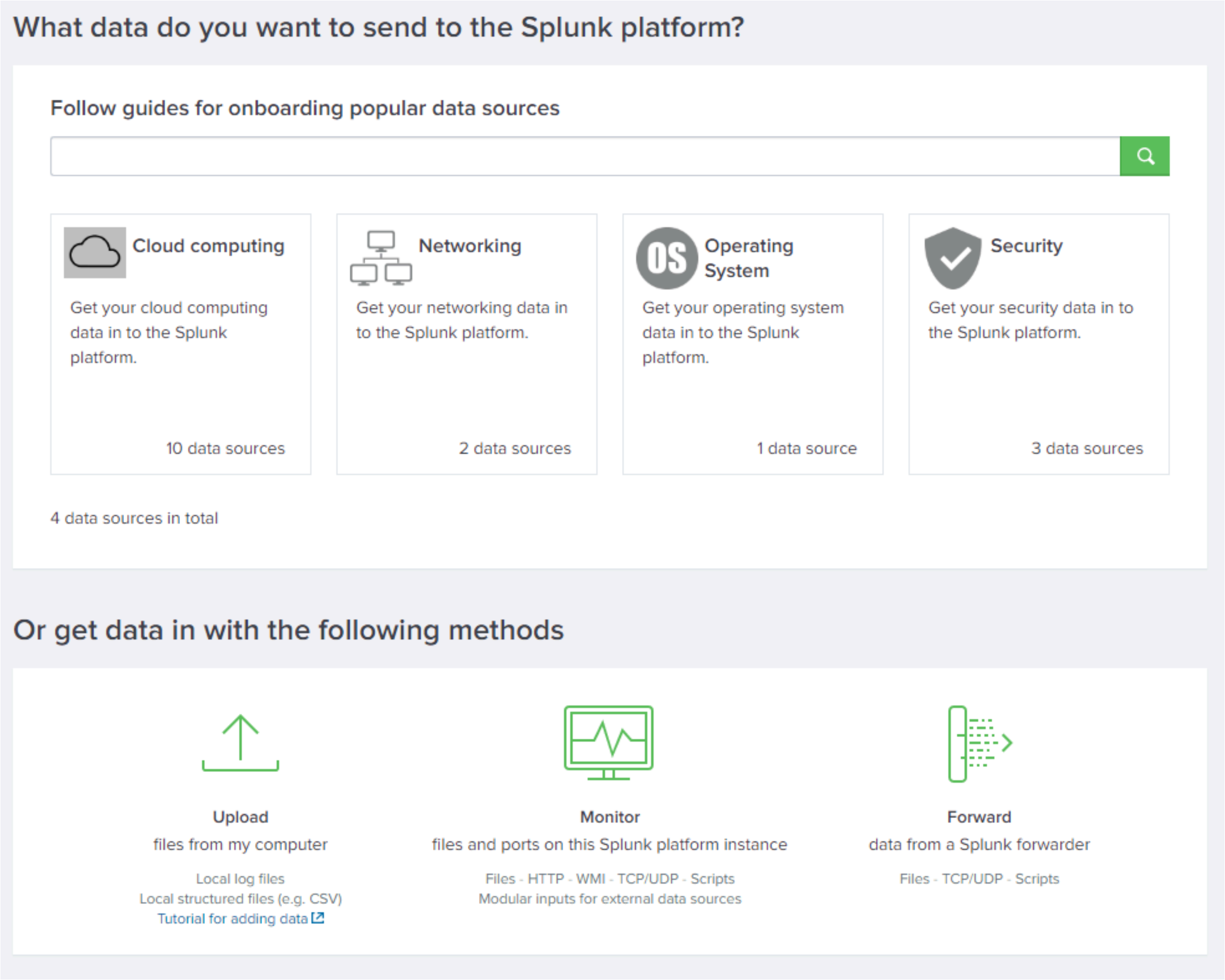
The data sources can be event logs, website logs, firewall logs, etc.

Data sources are grouped into categories. Below is a chart listing from the Splunk documentation detailing each data source category.



Data source	Description
Files and directories	Most data that you might be interested in comes directly from files and directories.
Network events	The Splunk software can index remote data from any network port and SNMP events from remote devices.
IT Operations	Data from IT Ops, such as Nagios, NetApp, and Cisco.
Cloud services	Data from Cloud services, such as AWS and Kinesis.
Database services	Data from databases such as Oracle, MySQL, and Microsoft SQL Server.
Security services	Data from security services such as McAfee, Microsoft Active Directory, and Symantec Endpoint Protection.
Virtualization services	Data from virtualization services such as VMWare and XenApp.
Application servers	Data from application servers such as JMX & JMS, WebLogic, and WebSphere.
Windows sources	The Windows version of Splunk software accepts a wide range of Windows-specific inputs, including Windows Event Log, Windows Registry, WMI, Active Directory, and Performance monitoring.
Other sources	Other input sources are supported, such as FIFO queues and scripted inputs for getting data from APIs, and other remote data interfaces.

In this room, we're going to focus on **VPN logs**. When we click on the link (from the Splunk home screen), we're presented with the following screen.



What data do you want to send to the Splunk platform?

Follow guides for onboarding popular data sources

Cloud computing

Get your cloud computing data in to the Splunk platform.

10 data sources

Networking

Get your networking data in to the Splunk platform.

2 data sources

Operating System

Get your operating system data in to the Splunk platform.

1 data source

Security

Get your security data in to the Splunk platform.

3 data sources

4 data sources in total

Or get data in with the following methods

Upload

files from my computer

Local log files
Local structured files (e.g. CSV)
Tutorial for adding data

Monitor

files and ports on this Splunk platform instance

Files - HTTP - WMI - TCP/UDP - Scripts
Modular inputs for external data sources

Forward

data from a Splunk forwarder

Files - TCP/UDP - Scripts

We will use the Upload Option to upload the data from our local machine. Download the attached log file and upload it on Splunk.

As shown above, it has a total of 5 steps to successfully upload the data.

1. **Select Source** -> Where we select the Log source.
2. **Select Source Type** -> Select what type of logs are being ingested.
3. **Input Settings** ->Select the index where these logs will be dumped and hostName to be associated with the logs.
4. **Review** -> Review all the gif
5. **Done** -> Final step, where the data is uploaded successfully and ready to be analyzed.

As you can see, there are **A LOT** more logs we can add to the Splunk instance, and Splunk supports various source types.

Download the attached log file "VPN_logs" and upload this file into the Splunk instance with the right source type.

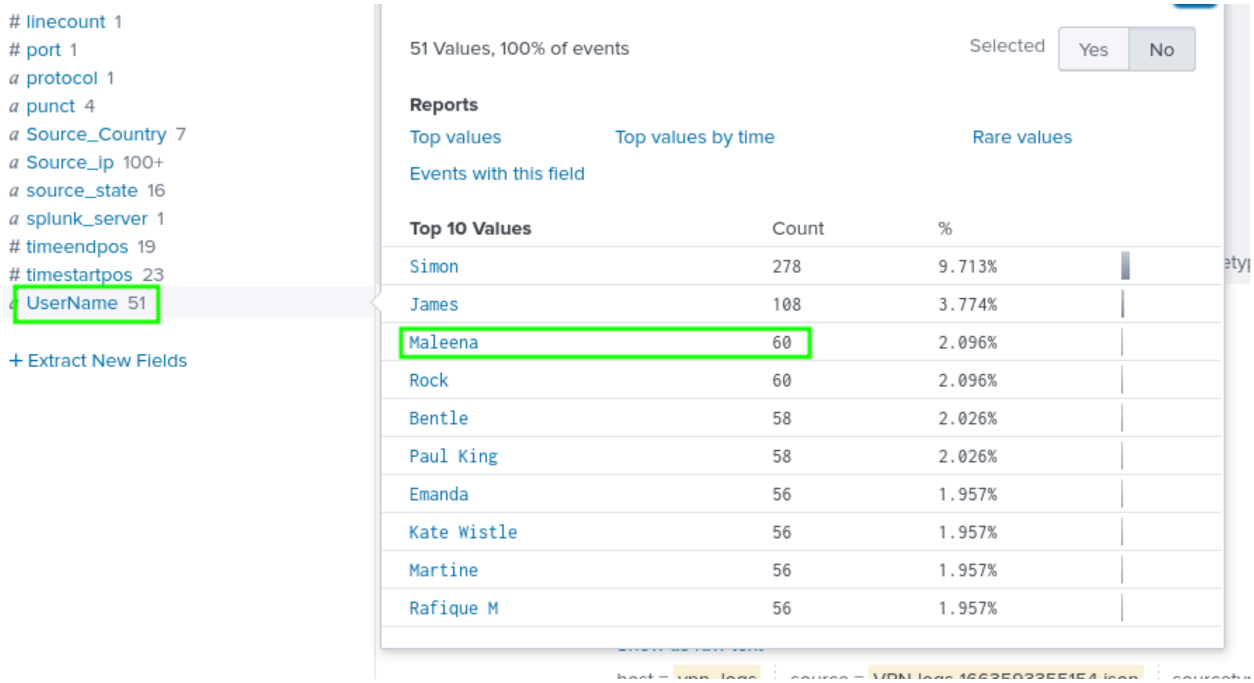
Answer the questions below

Upload the data attached to this task and create an index "VPN_Logs". How many events are present in the log file?

2862

How many log events by the user Maleena are captured?

Use the username option in the left side and then look for the user Maleena



linecount 1
port 1
a protocol 1
a punct 4
a Source_Country 7
a Source_ip 100+
a source_state 16
a splunk_server 1
# timeendpos 19
# timestartpos 23
a UserName 51

+ Extract New Fields

51 Values, 100% of events

Selected Yes No

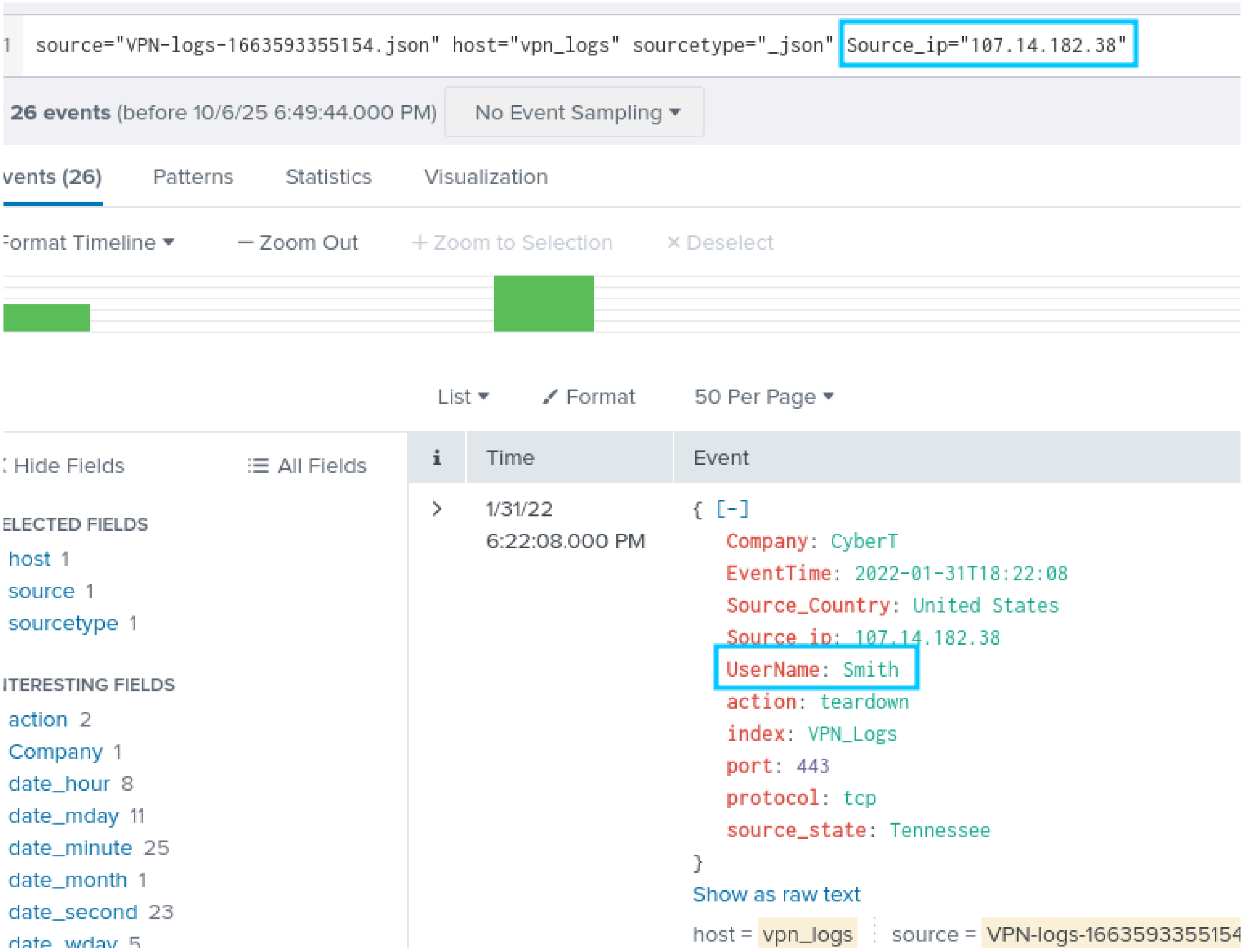
Reports
Top values   Top values by time   Rare values
Events with this field

Top 10 Values	Count	%
Simon	278	9.713%
James	108	3.774%
Maleena	60	2.096%
Rock	60	2.096%
Bentle	58	2.026%
Paul King	58	2.026%
Emanda	56	1.957%
Kate Wistle	56	1.957%
Martine	56	1.957%
Rafique M	56	1.957%

60

What is the name associated with IP 107.14.182.38?

Add this part of query Source_ip=”**107.14.182.38**” to the search field , and find the username in the filtered logs



```
source="VPN-logs-1663593355154.json" host="vpn_logs" sourcetype="_json" Source_ip="107.14.182.38"
```

26 events (before 10/6/25 6:49:44.000 PM) No Event Sampling ▾

vents (26)

ELECTED FIELDS
host 1
source 1
sourcetype 1

ITERESTING FIELDS
action 2
Company 1
date_hour 8
date_mday 11
date_minute 25
date_month 1
date_second 23
date_wday 5

1/31/22 6:22:08.000 PM

```
{ [-]
   Company: CyberT
   EventTime: 2022-01-31T18:22:08
   Source_Country: United States
   Source_ip: 107.14.182.38
   UserName: Smith
   action: teardown
   index: VPN_Logs
   port: 443
   protocol: tcp
   source_state: Tennessee
}
```

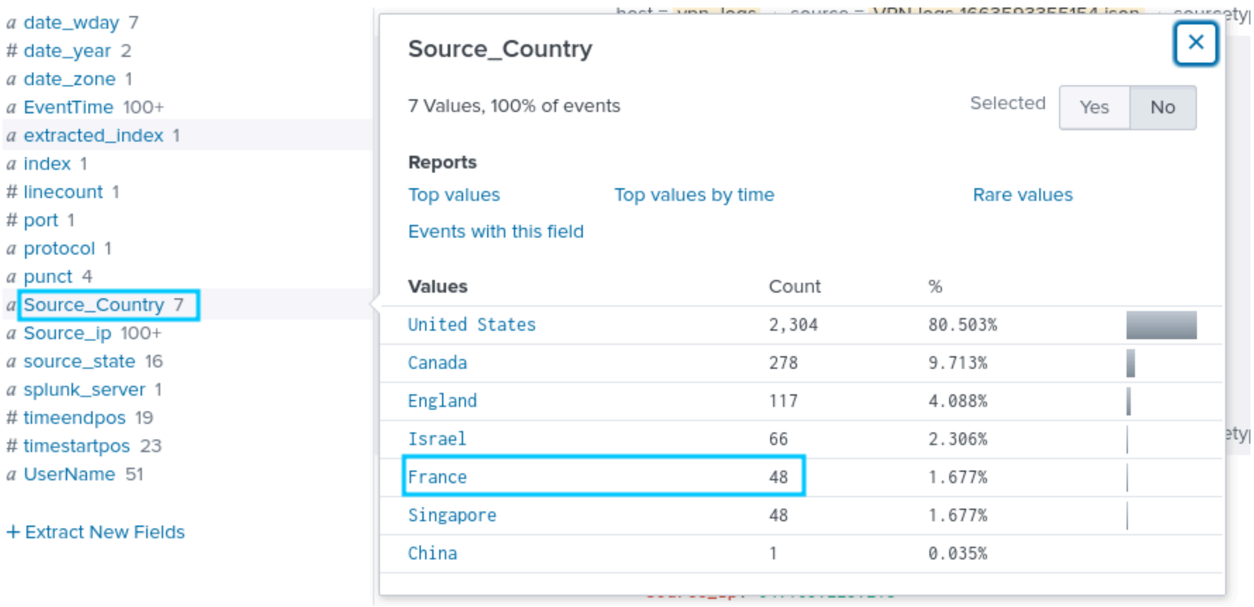
host = vpn_logs source = VPN-logs-1663593355154

smith

What is the number of events that originated from all countries except France?

To find the total except France remove the france from the total (2862)

To find france click on the source_Country on the left side



a date_wday 7
# date_year 2
a date_zone 1
a EventTime 100+
a extracted_index 1
a index 1
# linecount 1
# port 1
a protocol 1
a punct 4
a Source_Country 7
a Source_ip 100+
a source_state 16
a splunk_server 1
# timeendpos 19
# timestartpos 23
a UserName 51

+ Extract New Fields

Source_Country

7 Values, 100% of events

Selected Yes [✓ No]

Reports

Values	Count	%
United States	2,304	80.503%
Canada	278	9.713%
England	117	4.088%
Israel	66	2.306%
France	48	1.677%
Singapore	48	1.677%
China	1	0.035%

2814

How many VPN Events were observed by the IP 107.3.206.58?

Add the Source_ip="107.3.206.58" part of query to the existing one to filter by the ip address

14