Perform a Query with Splunk

Scenario

You are a security analyst working at the e-commerce store Buttercup Games. You've been tasked with identifying whether there are any possible security issues with the mail server. To do so, you must explore any failed SSH logins for the root account.

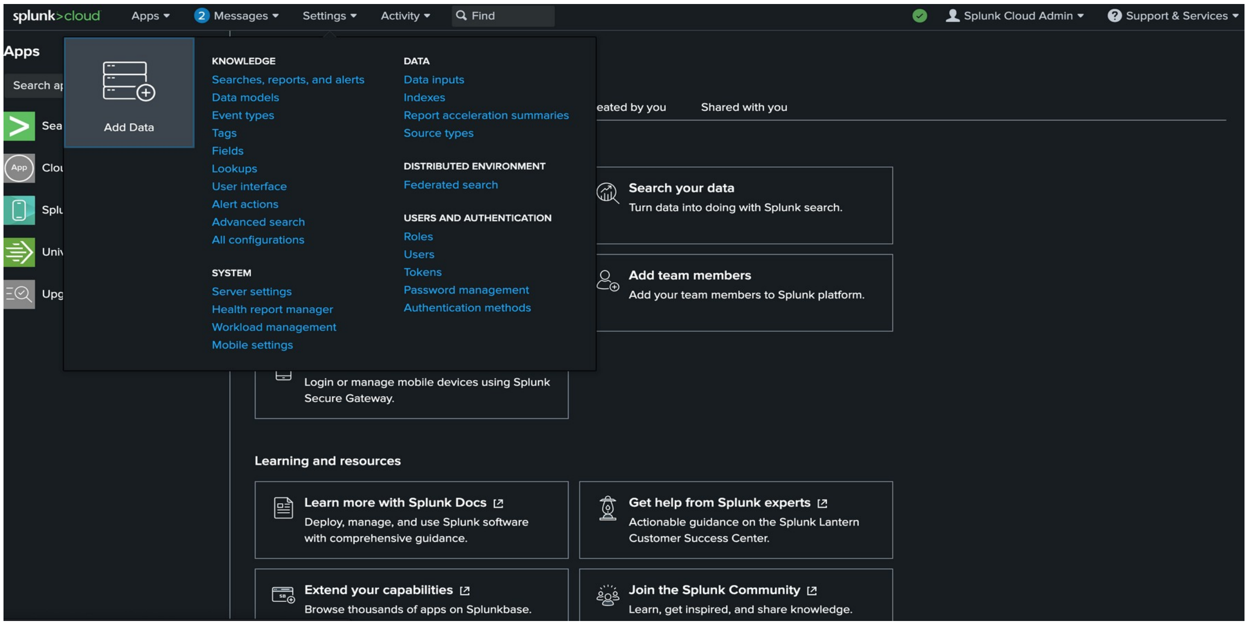
Project Description

Creating effective searches is an important skill because it enables you to quickly and accurately find the information you are looking for within a large amount of data. Quick and accurate searching is especially useful during incident response, because you might need to swiftly identify and address a security incident. Effective search techniques also help you efficiently identify patterns, trends, and anomalies within data.

Upload Data to Splunk:

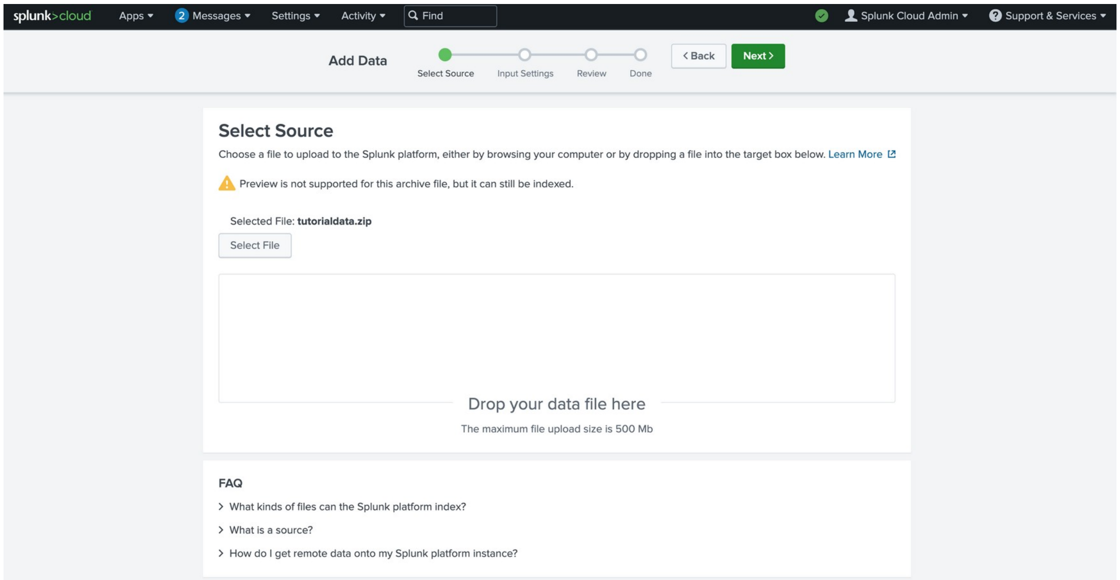
- Navigate to Splunk Home from your Splunk Cloud free trial instance. You might need to log in again using your credentials from Step 3.
- On the Splunk bar, click **Settings.** Then click the **Add Data** icon.

- Click **Upload**.



- Click the **Select File** button.

- Upload the *tutorialdata.zip* file, and click **Open**.



Select Source

Choose a file to upload to the Splunk platform, either by browsing your computer or by dropping a file into the target box below. Learn More

Preview is not supported for this archive file, but it can still be indexed.

Selected File: **tutorialdata.zip**

Select File

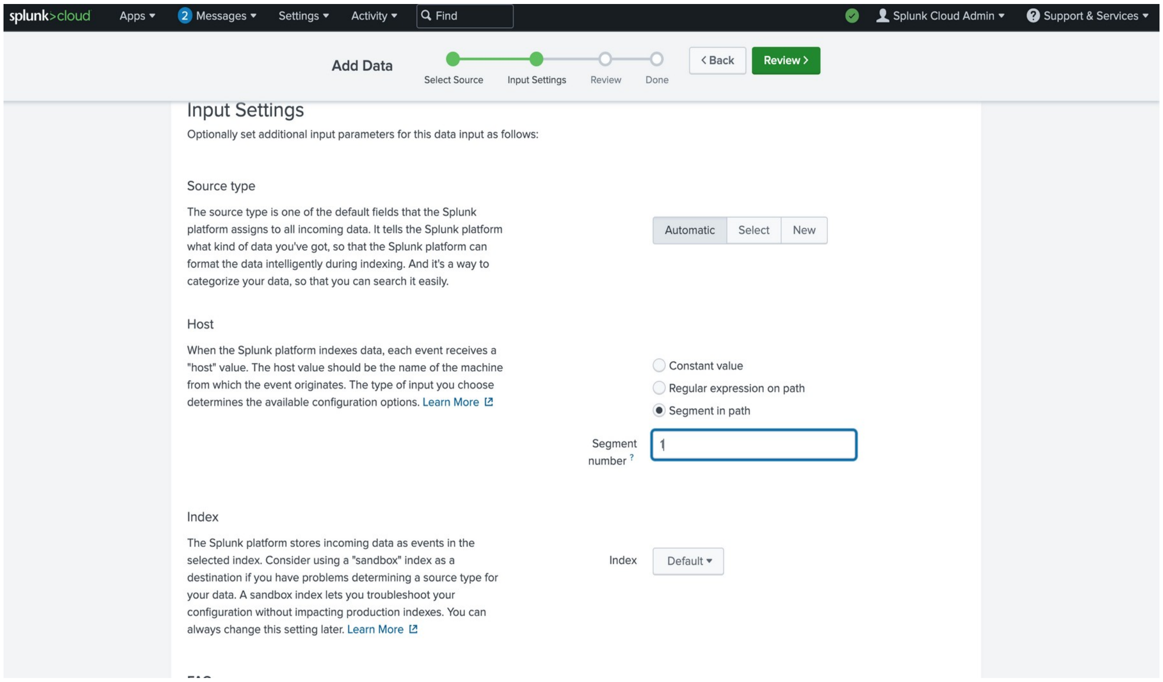
Drop your data file here

The maximum file upload size is 500 Mb

FAQ

- What kinds of files can the Splunk platform index?
- What is a source?
- How do I get remote data onto my Splunk platform instance?

- By the **Host** section, select **Segment in path** and enter **1** as the segment number.



Input Settings

Optionally set additional input parameters for this data input as follows:

Source type

The source type is one of the default fields that the Splunk platform assigns to all incoming data. It tells the Splunk platform what kind of data you've got, so that the Splunk platform can format the data intelligently during indexing. And it's a way to categorize your data, so that you can search it easily.

Automatic | Select | New

Host

When the Splunk platform indexes data, each event receives a "host" value. The host value should be the name of the machine from which the event originates. The type of input you choose determines the available configuration options. Learn More

○ Constant value
○ Regular expression on path
● Segment in path

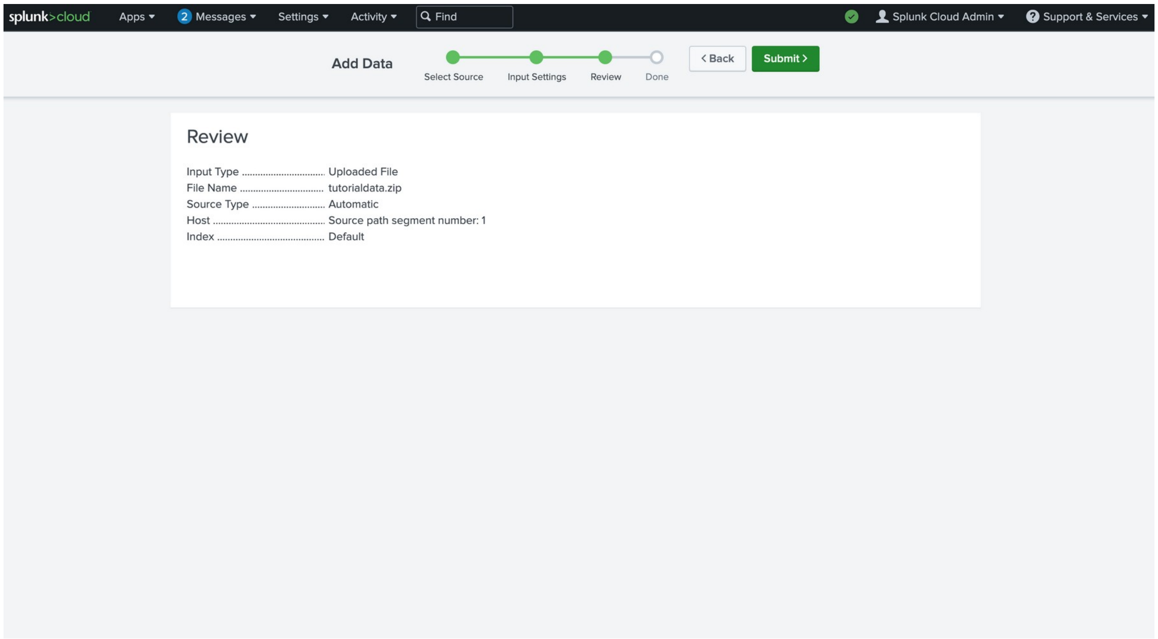
Segment number: 1

Index

The Splunk platform stores incoming data as events in the selected index. Consider using a "sandbox" index as a destination if you have problems determining a source type for your data. A sandbox index lets you troubleshoot your configuration without impacting production indexes. You can always change this setting later. Learn More

Index: Default ▾

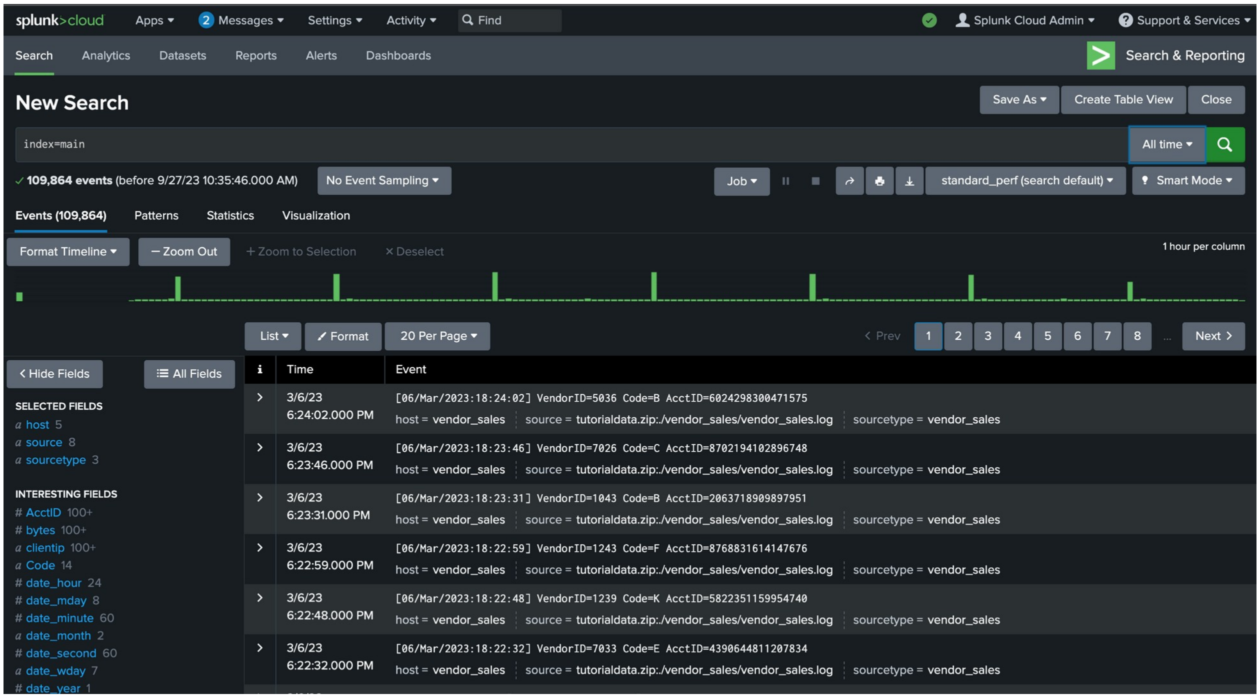
- Click the **Review** button and review the details of the upload before you submit. The details should be as follows: Input Type: Uploaded File File Name: tutorialdata.zip Source Type: Automatic Host: Source path segment number: 1 Index: Default.



Review

Input Type .... Uploaded File
File Name .... tutorialdata.zip
Source Type .... Automatic
Host .... Source path segment number: 1
Index .... Default

- Click **Submit**. Once Splunk has ingested the data, you will receive confirmation that the file was successfully uploaded.
- Navigate to Splunk Home. (To return to Splunk Home, click the Splunk Cloud logo on the Splunk Cloud page.)
- Click **Search & Reporting**. You may close any pop ups that appear.
- In the search bar, enter your search query: *index=main* This search term specifies the index. An **index** is a repository for data. Here, the index is a single dataset containing events from an index named main.
- Select **All Time** from the time range dropdown to search for all the events across all time.
- Click the search button. Note that the search button is represented by the magnifying glass icon. Your search should retrieve thousands of events.



```
index=main
```

✓ **109,864 events** (before 9/27/23 10:35:46.000 AM)

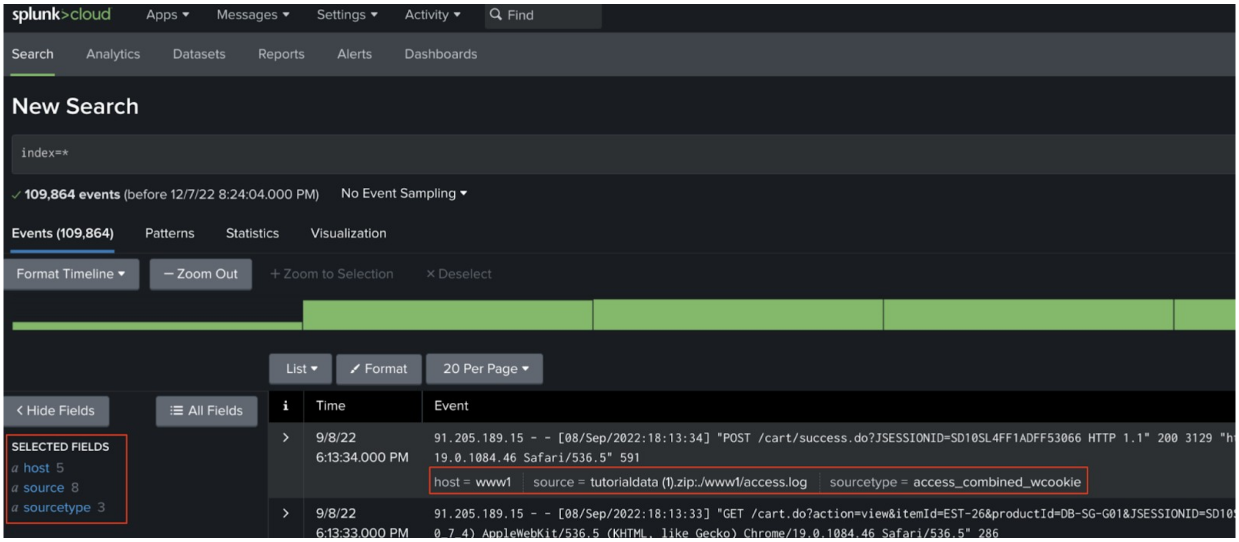
SELECTED FIELDS
- a host 5
- a source 8
- a sourcetype 3

INTERESTING FIELDS
- # AcctID 100+
- # bytes 100+
- a clientip 100+
- a Code 14
- # date_hour 24
- # date_mday 8
- # date_minute 60
- a date_month 2
- # date_second 60
- a date_wday 7
- # date_year 1

i	Time	Event
>	3/6/23 6:24:02.000 PM	[06/Mar/2023:18:24:02] VendorID=5036 Code=B AcctID=6024298300471575 host = vendor_sales source = tutorialdata.zip:./vendor_sales/vendor_sales.log sourcetype = vendor_sales
>	3/6/23 6:23:46.000 PM	[06/Mar/2023:18:23:46] VendorID=7026 Code=C AcctID=8702194102896748 host = vendor_sales source = tutorialdata.zip:./vendor_sales/vendor_sales.log sourcetype = vendor_sales
>	3/6/23 6:23:31.000 PM	[06/Mar/2023:18:23:31] VendorID=1043 Code=B AcctID=2063718909897951 host = vendor_sales source = tutorialdata.zip:./vendor_sales/vendor_sales.log sourcetype = vendor_sales
>	3/6/23 6:22:59.000 PM	[06/Mar/2023:18:22:59] VendorID=1243 Code=F AcctID=8768831614147676 host = vendor_sales source = tutorialdata.zip:./vendor_sales/vendor_sales.log sourcetype = vendor_sales
>	3/6/23 6:22:48.000 PM	[06/Mar/2023:18:22:48] VendorID=1239 Code=K AcctID=5822351159954740 host = vendor_sales source = tutorialdata.zip:./vendor_sales/vendor_sales.log sourcetype = vendor_sales
>	3/6/23 6:22:32.000 PM	[06/Mar/2023:18:22:32] VendorID=7033 Code=E AcctID=4390644811207834 host = vendor_sales source = tutorialdata.zip:./vendor_sales/vendor_sales.log sourcetype = vendor_sales

When Splunk indexes data, it attaches fields to each event. These fields become part of the searchable index event data. This helps security analysts easily search for and find the specific data they need. Now that you've run your first query, examine the search results and the fields.

For each event the fields are *host*, *source*, and *sourcetype*. Under **SELECTED FIELDS**, examine the same fields.



```
index=*
```

✓ **109,864 events** (before 12/7/22 8:24:04.000 PM)

SELECTED FIELDS
- a host 5
- a source 8
- a sourcetype 3

i	Time	Event
>	9/8/22 6:13:34.000 PM	91.205.189.15 - - [08/Sep/2022:18:13:34] "POST /cart/success.do?JSESSIONID=SD10SL4FF1ADFF53066 HTTP 1.1" 200 3129 "h 19.0.1084.46 Safari/536.5" 591 host = www1 source = tutorialdata (1).zip:./www1/access.log sourcetype = access_combined_wcookie
>	9/8/22 6:13:33.000 PM	91.205.189.15 - - [08/Sep/2022:18:13:33] "GET /cart.do?action=view&itemId=EST-26&productId=DB-SG-G01&JSESSIONID=SD10 0_7_4) AppleWebKit/536.5 (KHTML, like Gecko) Chrome/19.0.1084.46 Safari/536.5" 286

host: The host field specifies the name of the network host from which the event originated. In this search there are five hosts:

mailsv - Buttercup Games' mail server. Examine events generated from this host.

www1 - This is one of Buttercup Games' web applications.

www2 - This is one of Buttercup Games' web applications.

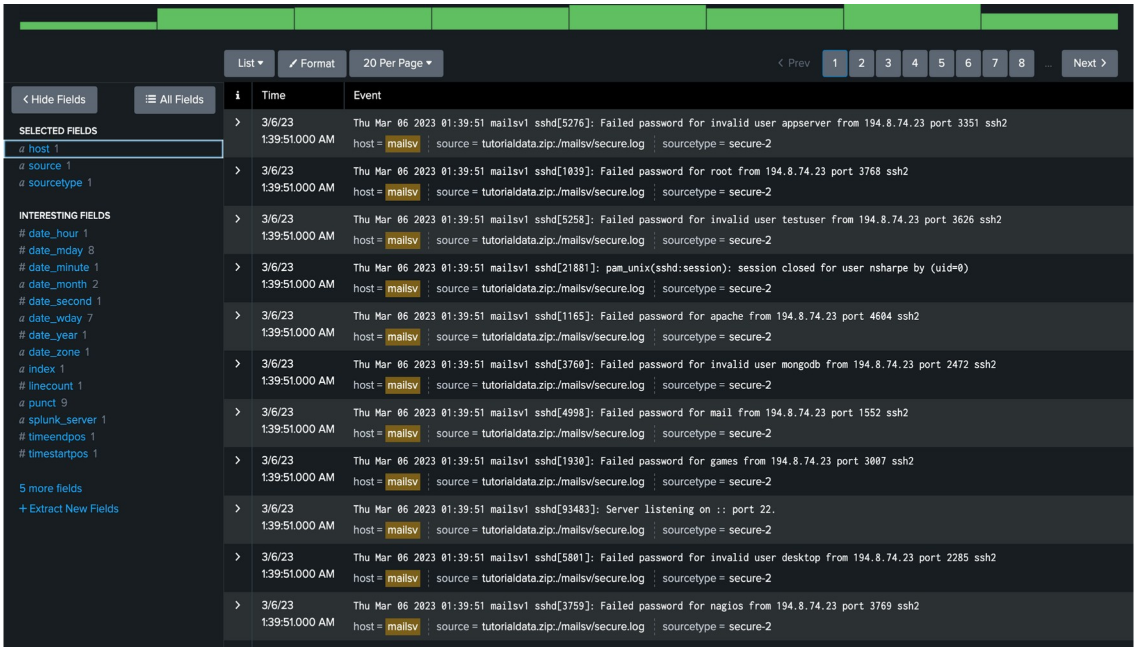
www3 - This is one of Buttercup Games' web applications.

vendor_sales - Information about Buttercup Games' retail sales.

source: The source field indicates the file name from which the event originates. You should identify eight sources. Notice */mailsv/secure.log*, which is a log file that contains information related to authentication and authorization attempts on the mail server.

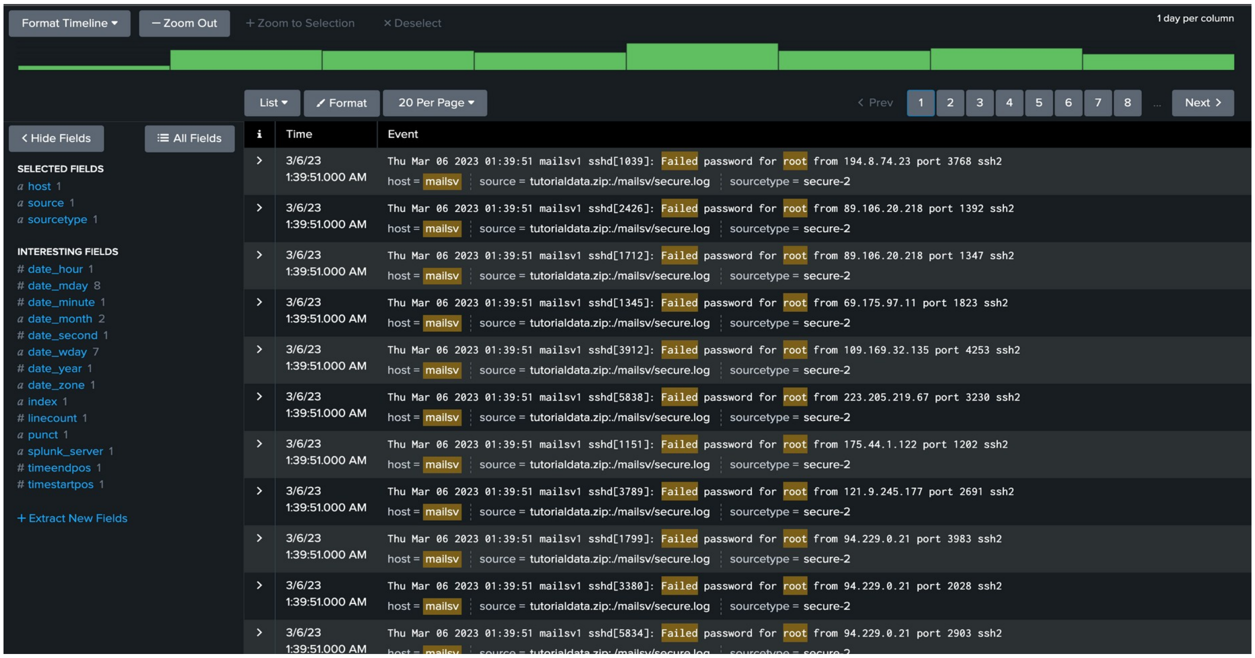
sourcetype: The sourcetype determines how data is formatted.

- Under **SELECTED FIELDS**, click **host** and click **mailsv**.



```
3/6/23 1:39:51.000 AM  Thu Mar 06 2023 01:39:51 mailsv1 sshd[5276]: Failed password for invalid user appserver from 194.8.74.23 port 3351 ssh2
                       host = mailsv   source = tutorialdata.zip:./mailsv/secure.log   sourcetype = secure-2
3/6/23 1:39:51.000 AM  Thu Mar 06 2023 01:39:51 mailsv1 sshd[1039]: Failed password for root from 194.8.74.23 port 3768 ssh2
                       host = mailsv   source = tutorialdata.zip:./mailsv/secure.log   sourcetype = secure-2
3/6/23 1:39:51.000 AM  Thu Mar 06 2023 01:39:51 mailsv1 sshd[5258]: Failed password for invalid user testuser from 194.8.74.23 port 3626 ssh2
                       host = mailsv   source = tutorialdata.zip:./mailsv/secure.log   sourcetype = secure-2
3/6/23 1:39:51.000 AM  Thu Mar 06 2023 01:39:51 mailsv1 sshd[21881]: pam_unix(sshd:session): session closed for user nsharpe by (uid=0)
                       host = mailsv   source = tutorialdata.zip:./mailsv/secure.log   sourcetype = secure-2
3/6/23 1:39:51.000 AM  Thu Mar 06 2023 01:39:51 mailsv1 sshd[1165]: Failed password for apache from 194.8.74.23 port 4604 ssh2
                       host = mailsv   source = tutorialdata.zip:./mailsv/secure.log   sourcetype = secure-2
3/6/23 1:39:51.000 AM  Thu Mar 06 2023 01:39:51 mailsv1 sshd[3760]: Failed password for invalid user mongodb from 194.8.74.23 port 2472 ssh2
                       host = mailsv   source = tutorialdata.zip:./mailsv/secure.log   sourcetype = secure-2
3/6/23 1:39:51.000 AM  Thu Mar 06 2023 01:39:51 mailsv1 sshd[4998]: Failed password for mail from 194.8.74.23 port 1552 ssh2
                       host = mailsv   source = tutorialdata.zip:./mailsv/secure.log   sourcetype = secure-2
3/6/23 1:39:51.000 AM  Thu Mar 06 2023 01:39:51 mailsv1 sshd[1930]: Failed password for games from 194.8.74.23 port 3007 ssh2
                       host = mailsv   source = tutorialdata.zip:./mailsv/secure.log   sourcetype = secure-2
3/6/23 1:39:51.000 AM  Thu Mar 06 2023 01:39:51 mailsv1 sshd[93483]: Server listening on :: port 22.
                       host = mailsv   source = tutorialdata.zip:./mailsv/secure.log   sourcetype = secure-2
3/6/23 1:39:51.000 AM  Thu Mar 06 2023 01:39:51 mailsv1 sshd[5801]: Failed password for invalid user desktop from 194.8.74.23 port 2285 ssh2
                       host = mailsv   source = tutorialdata.zip:./mailsv/secure.log   sourcetype = secure-2
3/6/23 1:39:51.000 AM  Thu Mar 06 2023 01:39:51 mailsv1 sshd[3759]: Failed password for nagios from 194.8.74.23 port 3769 ssh2
                       host = mailsv   source = tutorialdata.zip:./mailsv/secure.log   sourcetype = secure-2
```

- Enter index=main host=mailsv fail* root into the search bar.



SELECTED FIELDS
a host 1
a source 1
a sourcetype 1

INTERESTING FIELDS
# date_hour 1
# date_mday 8
# date_minute 1
a date_month 2
# date_second 1
a date_wday 7
# date_year 1
a date_zone 1
a index 1
# linecount 1
a punct 1
a splunk_server 1
# timeendpos 1
# timestartpos 1

```
i  Time                 Event
>  3/6/23               Thu Mar 06 2023 01:39:51 mailsv1 sshd[1039]: Failed password for root from 194.8.74.23 port 3768 ssh2
   1:39:51.000 AM       host = mailsv  source = tutorialdata.zip:./mailsv/secure.log  sourcetype = secure-2
>  3/6/23               Thu Mar 06 2023 01:39:51 mailsv1 sshd[2426]: Failed password for root from 89.106.20.218 port 1392 ssh2
   1:39:51.000 AM       host = mailsv  source = tutorialdata.zip:./mailsv/secure.log  sourcetype = secure-2
>  3/6/23               Thu Mar 06 2023 01:39:51 mailsv1 sshd[1712]: Failed password for root from 89.106.20.218 port 1347 ssh2
   1:39:51.000 AM       host = mailsv  source = tutorialdata.zip:./mailsv/secure.log  sourcetype = secure-2
>  3/6/23               Thu Mar 06 2023 01:39:51 mailsv1 sshd[1345]: Failed password for root from 69.175.97.11 port 1823 ssh2
   1:39:51.000 AM       host = mailsv  source = tutorialdata.zip:./mailsv/secure.log  sourcetype = secure-2
>  3/6/23               Thu Mar 06 2023 01:39:51 mailsv1 sshd[3912]: Failed password for root from 109.169.32.135 port 4253 ssh2
   1:39:51.000 AM       host = mailsv  source = tutorialdata.zip:./mailsv/secure.log  sourcetype = secure-2
>  3/6/23               Thu Mar 06 2023 01:39:51 mailsv1 sshd[5838]: Failed password for root from 223.205.219.67 port 3230 ssh2
   1:39:51.000 AM       host = mailsv  source = tutorialdata.zip:./mailsv/secure.log  sourcetype = secure-2
>  3/6/23               Thu Mar 06 2023 01:39:51 mailsv1 sshd[1151]: Failed password for root from 175.44.1.122 port 1202 ssh2
   1:39:51.000 AM       host = mailsv  source = tutorialdata.zip:./mailsv/secure.log  sourcetype = secure-2
>  3/6/23               Thu Mar 06 2023 01:39:51 mailsv1 sshd[3789]: Failed password for root from 121.9.245.177 port 2691 ssh2
   1:39:51.000 AM       host = mailsv  source = tutorialdata.zip:./mailsv/secure.log  sourcetype = secure-2
>  3/6/23               Thu Mar 06 2023 01:39:51 mailsv1 sshd[1799]: Failed password for root from 94.229.0.21 port 3983 ssh2
   1:39:51.000 AM       host = mailsv  source = tutorialdata.zip:./mailsv/secure.log  sourcetype = secure-2
>  3/6/23               Thu Mar 06 2023 01:39:51 mailsv1 sshd[3380]: Failed password for root from 94.229.0.21 port 2028 ssh2
   1:39:51.000 AM       host = mailsv  source = tutorialdata.zip:./mailsv/secure.log  sourcetype = secure-2
>  3/6/23               Thu Mar 06 2023 01:39:51 mailsv1 sshd[5834]: Failed password for root from 94.229.0.21 port 2903 ssh2
   1:39:51.000 AM       host = mailsv  source = tutorialdata.zip:./mailsv/secure.log  sourcetype = secure-2
```

This search expands on the search from the previous task and searches for the keyword fail*. The wildcard tells Splunk to expand the search term to find other terms that contain the word fail such as failure, failed, etc. Lastly, the keyword root searches for any event that contains the term root.