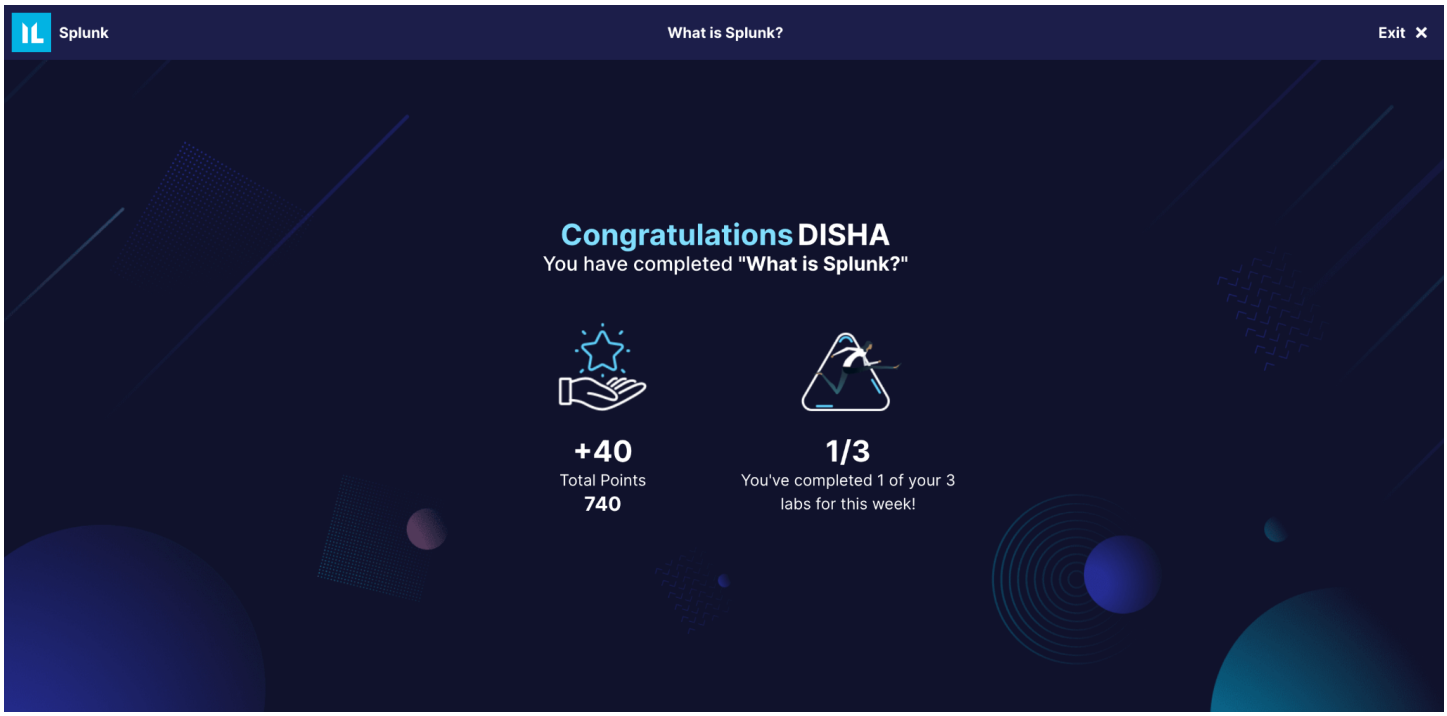
Network Monitoring- Lab 4

Name: Disha Sheshappa
Net ID: ds7615

Task 1: What is Splunk ?

Learnings:

- Data Source Variety:Understanding that Splunk can collect data from diverse sources such as files, network traffic, script outputs, Windows event logs, and more highlights the versatility of Splunk in handling data generated across an organization.
- Significance of Source Types: Recognizing the importance of source types in Splunk is crucial for efficient data handling. Source types determine how data is formatted and indexed, impacting the effectiveness of search queries and data analysis.
- Operational Insights:Learning about sources and source types in Splunk provides insights into efficient administration, filtering, and searching of data. This knowledge contributes to troubleshooting potential issues and identifying and preventing threats more quickly through a better understanding of data sources.

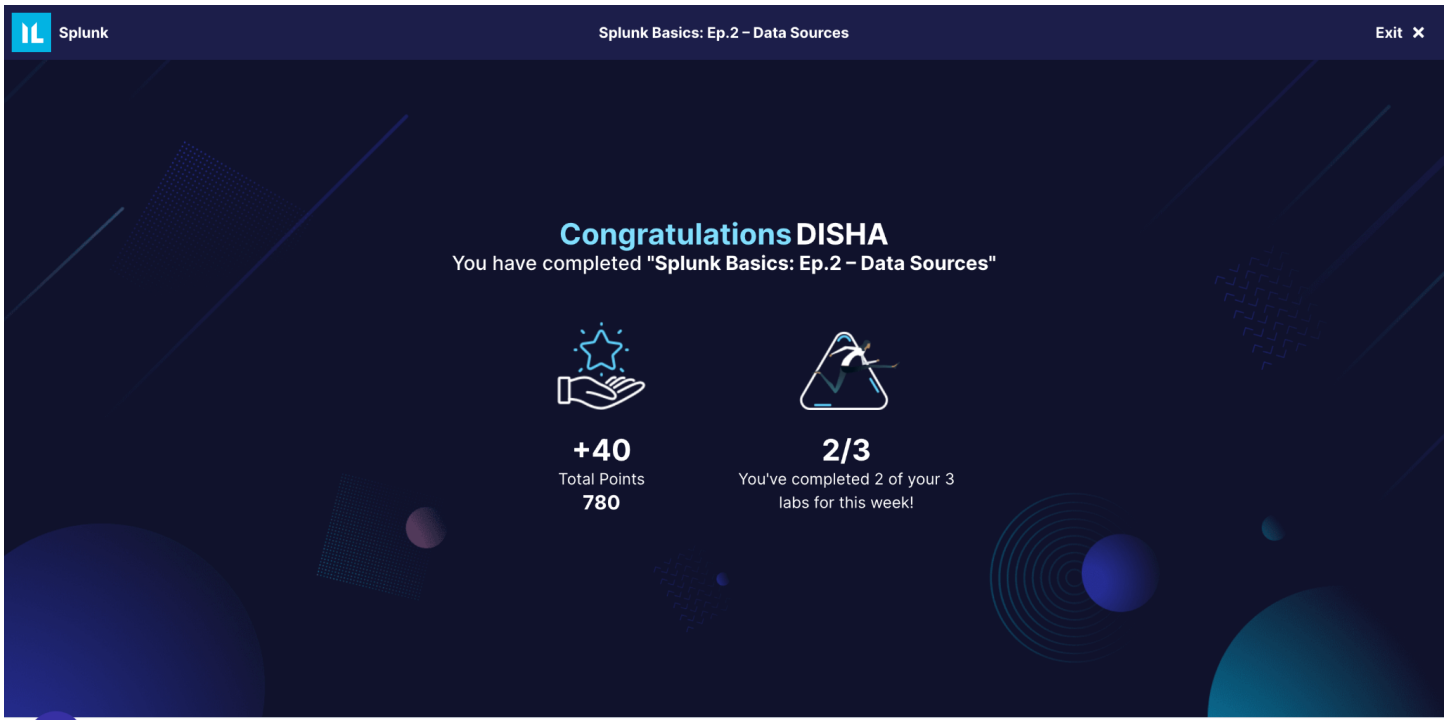


Splunk Basics: Ep.2 – Data Sources

Learnings:

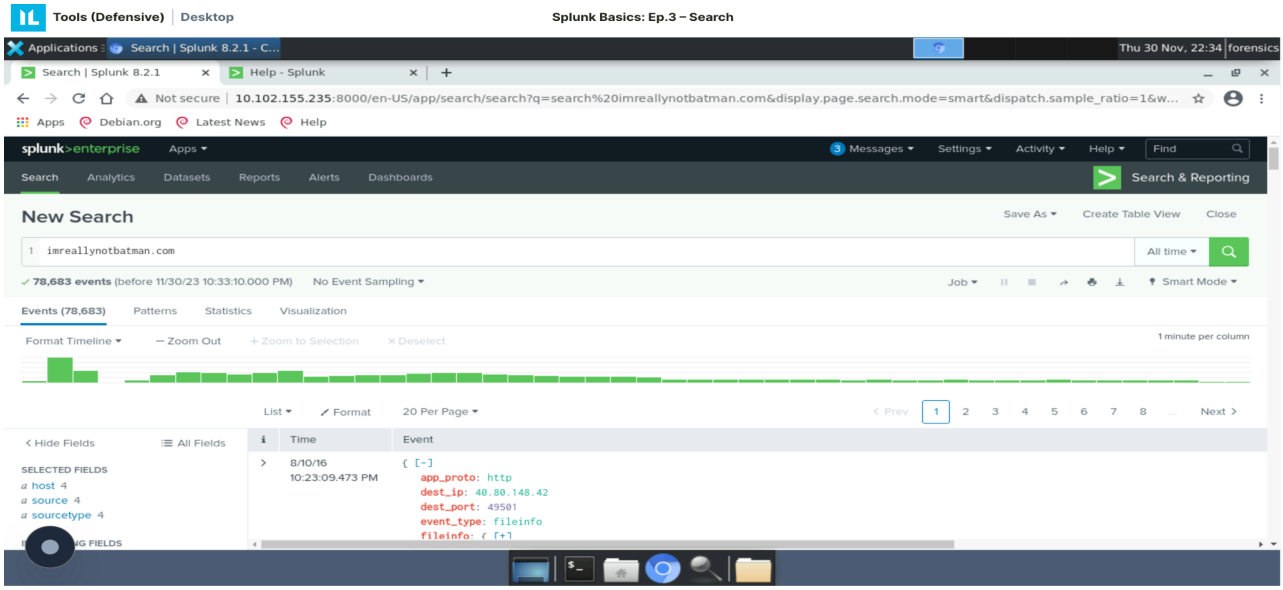
- Data Source Variety: Understanding that Splunk can collect data from diverse sources such as files, network traffic, script outputs, Windows event logs, and more highlights the versatility of Splunk in handling data generated across an organization.
- Significance of Source Types: Recognizing the importance of source types in Splunk is crucial for efficient data handling. Source types determine how data is formatted and indexed, impacting the effectiveness of search queries and data analysis.

- Operational Insights Learning about sources and source types in Splunk provides insights into efficient administration, filtering, and searching of data. This knowledge contributes to troubleshooting potential issues and identifying and preventing threats more quickly through a better understanding of data sources.



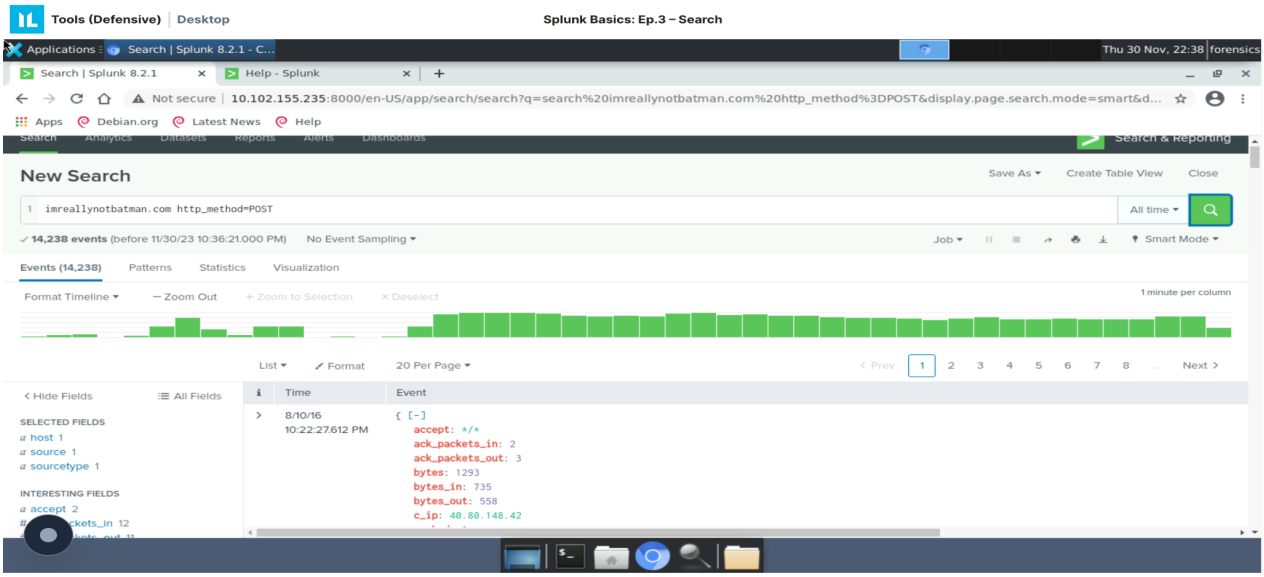
Splunk Basics: Ep.3 – Search

Q1 : Perform a search for the domain “imreallynotbatman.com”. How many events are returned?
Answer: The query returned 78,683 entries. When we directly give the domain name in the search bar it searches for the string in all the possible fields instead of limiting itself to one single field.



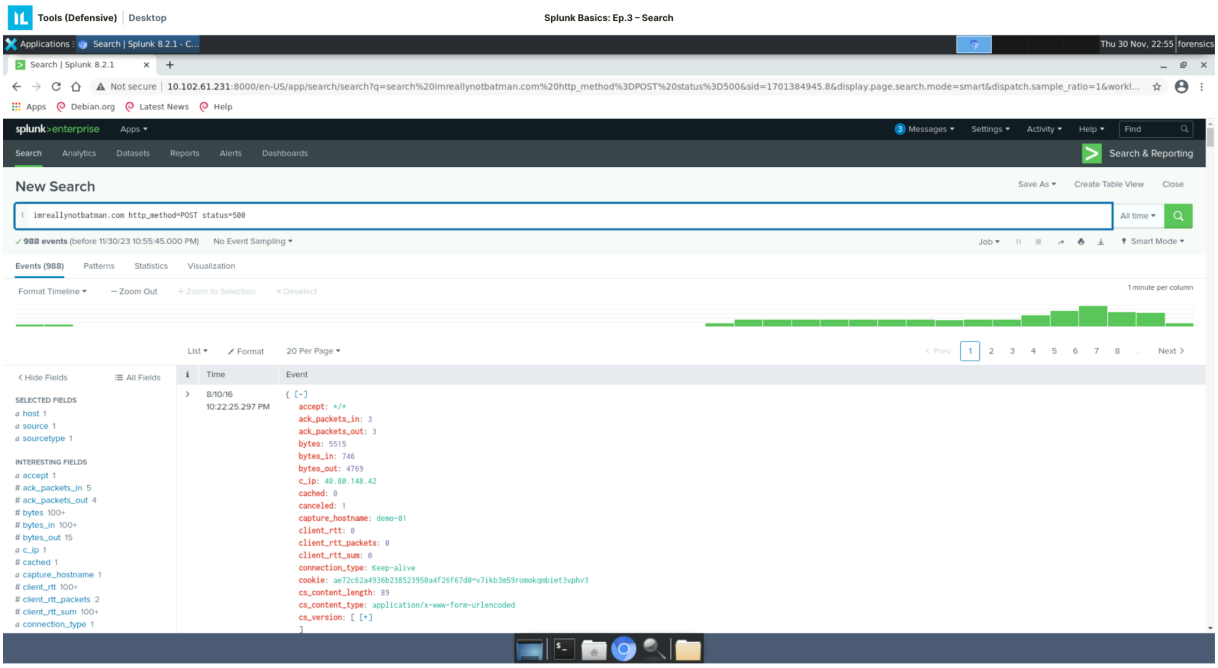
Q2: Perform a search for the domain “imreallynotbatman.com”, this time including the field "http_method=POST". How many events are returned?

Answer: the query returned 14,238 events. The search in the previous question only searched for the domain name. We added further refinement to it by mentioning http_method as POST. Hence the number of events filtered are less compared to previous one.



Q3: Perform a search for the domain “imreallynotbatman.com”, this time including the field "http_method=POST" and the field "status=500". How many events are returned?

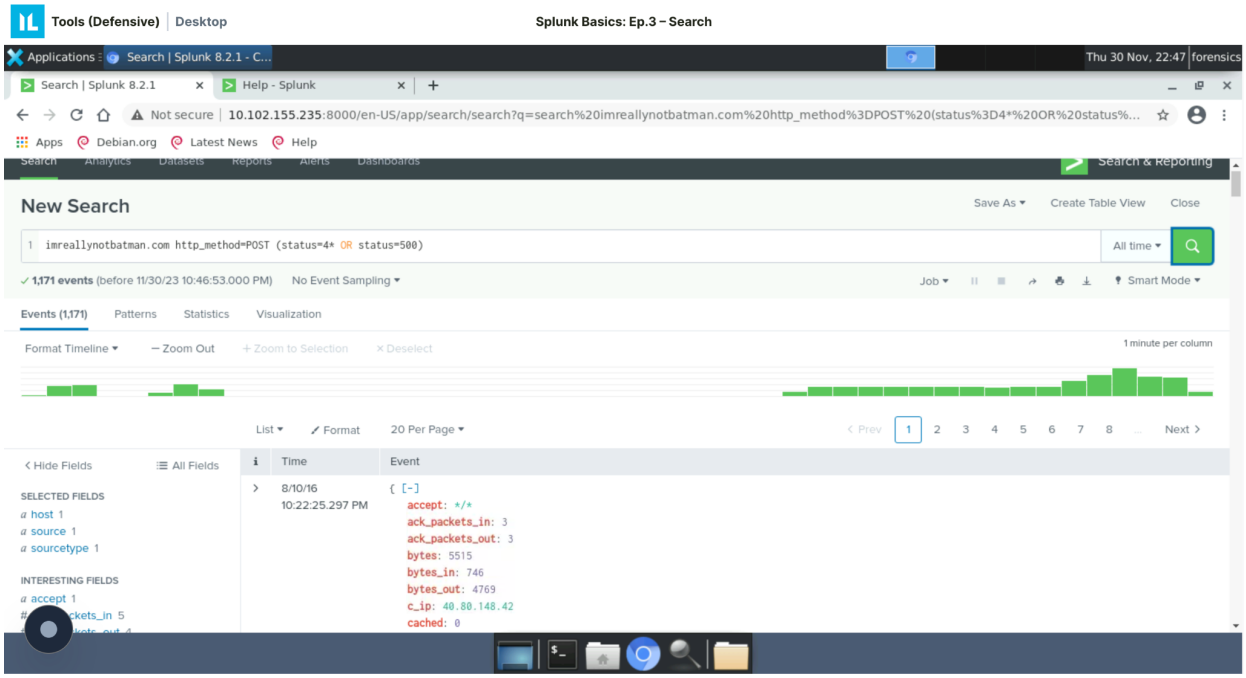
Answer: 988.
The search query explicitly mentions the status to be of value 500. All the other statuses are ignored. This kind of query can be used when we need to focus on only outcome and ignore rest of the noice.



Q4: Expand the search query from the previous question to also include all "status=4*" results. How many events are returned?

Answer: the number of entries returned is 1,171.

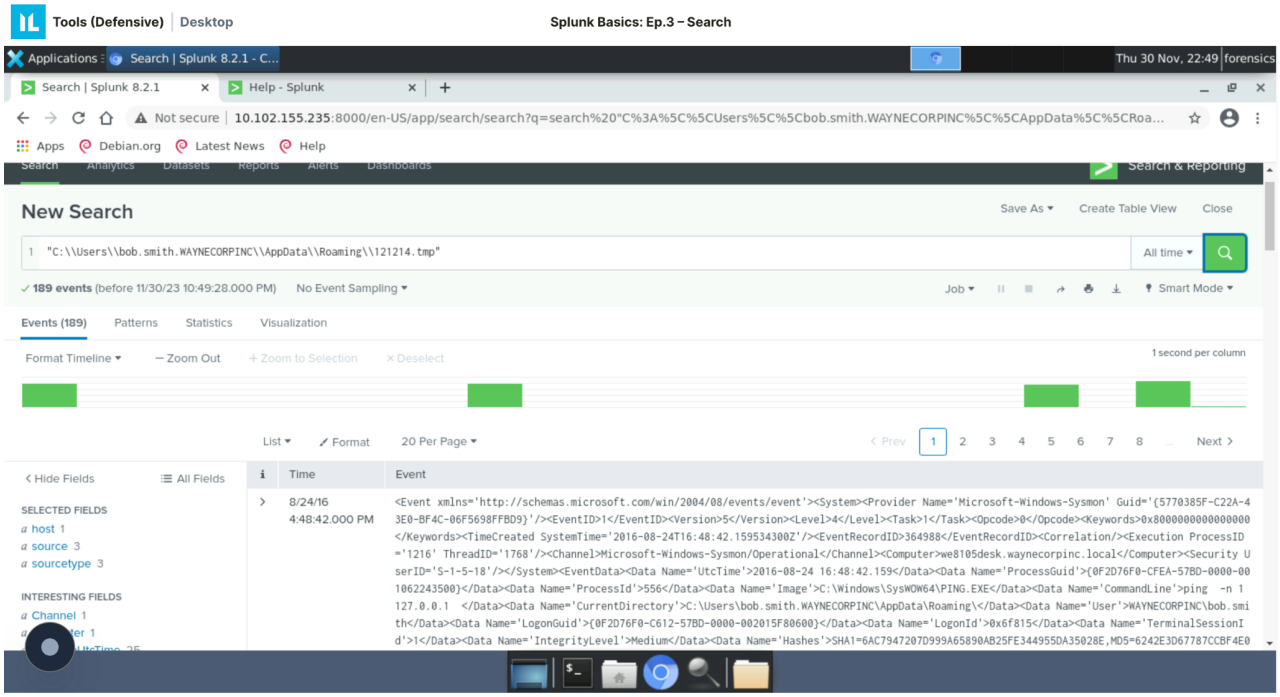
The search query used is imreallynotbatman.com http_method=POST (status=4* OR status=500). We are making use of OR operator here to filter entries which either have status 500 or any status starting with 4.



Q5: Perform a search for the filepath "C:\Users\bob.smith.WAYNECORPINC\AppData\Roaming\121214.tmp". How many events does it appear in?

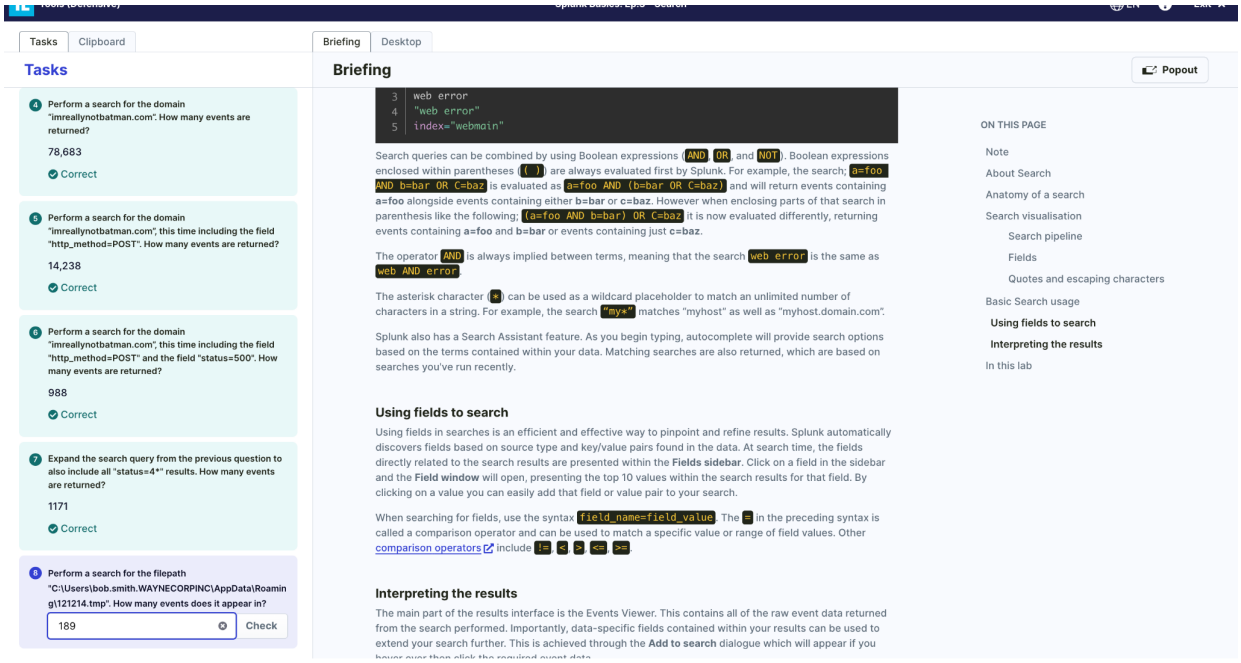
Answer: 189

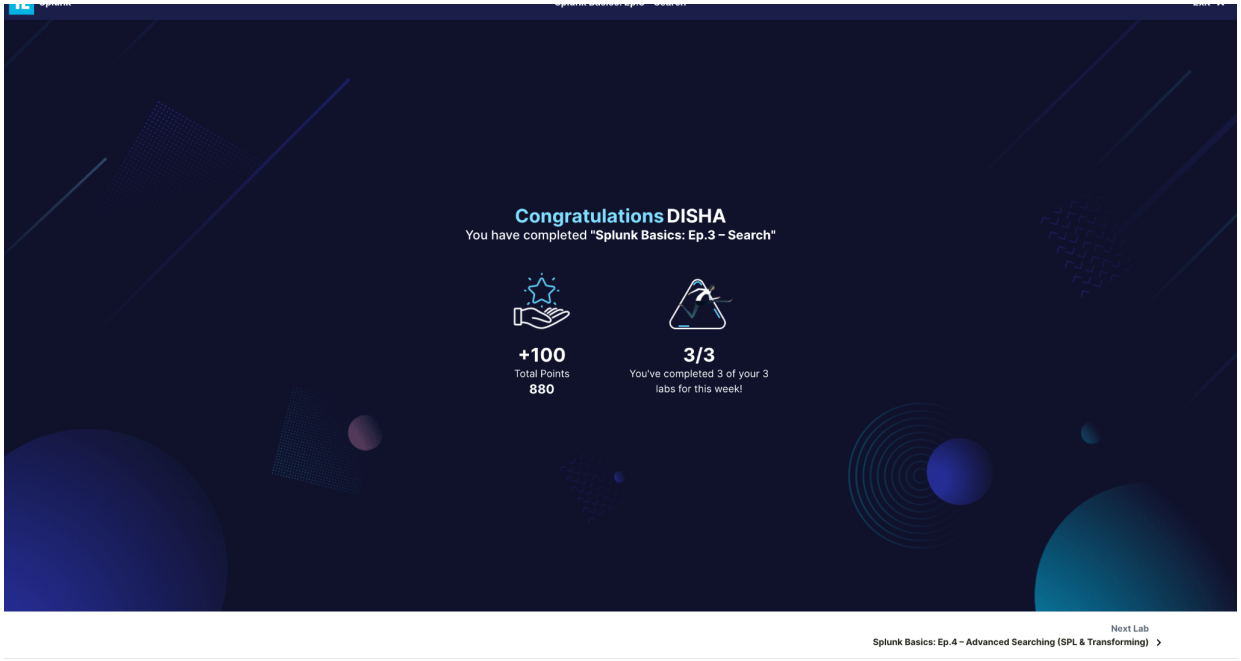
In this case if we directly search the string in the search bar we will not get any result. The key here is to enclose the string within double quotes. This is a necessary step for file paths containing spaces. As space is considered as start of new keyword in splunk.



Learnings:

- Basic Search Structure: Understood the core structure of Splunk searches, which consists of a series of instructions separated by the pipe symbol (|), allows for more effective data exploration.
- Search Pipeline: The search pipeline idea allows for step-by-step refining and augmentation of search results by chaining consecutive instructions together using the pipe character.
  Fields and Filtering: Understanding the role of fields in Splunk searches and their use as searchable key/value pairs aids in successful data filtering and refining.
- Boolean Expressions: Learned how to utilize Boolean expressions (AND, OR, NOT) and parenthesis in search queries to combine numerous search phrases for accurate results.
- Wildcard Usage and Escaping Characters: Utilizing wildcards, such as the asterisk (*) as a placeholder, and understanding the need to escape characters like quotes and backslashes enhances the flexibility and accuracy of search queries, especially in scenarios like searching for Windows file paths.
- Using wildcards as placeholders, such as the asterisk (*), and understanding the need to escape characters such as quotes and backslashes improves the flexibility and accuracy of search queries, especially when searching for Windows file paths.

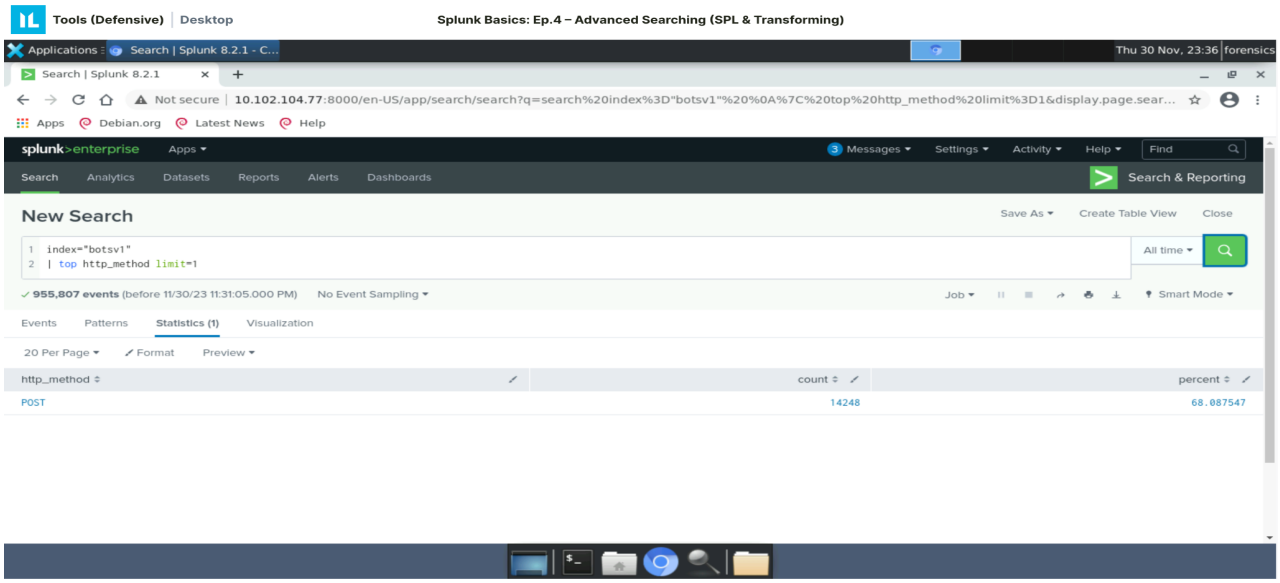




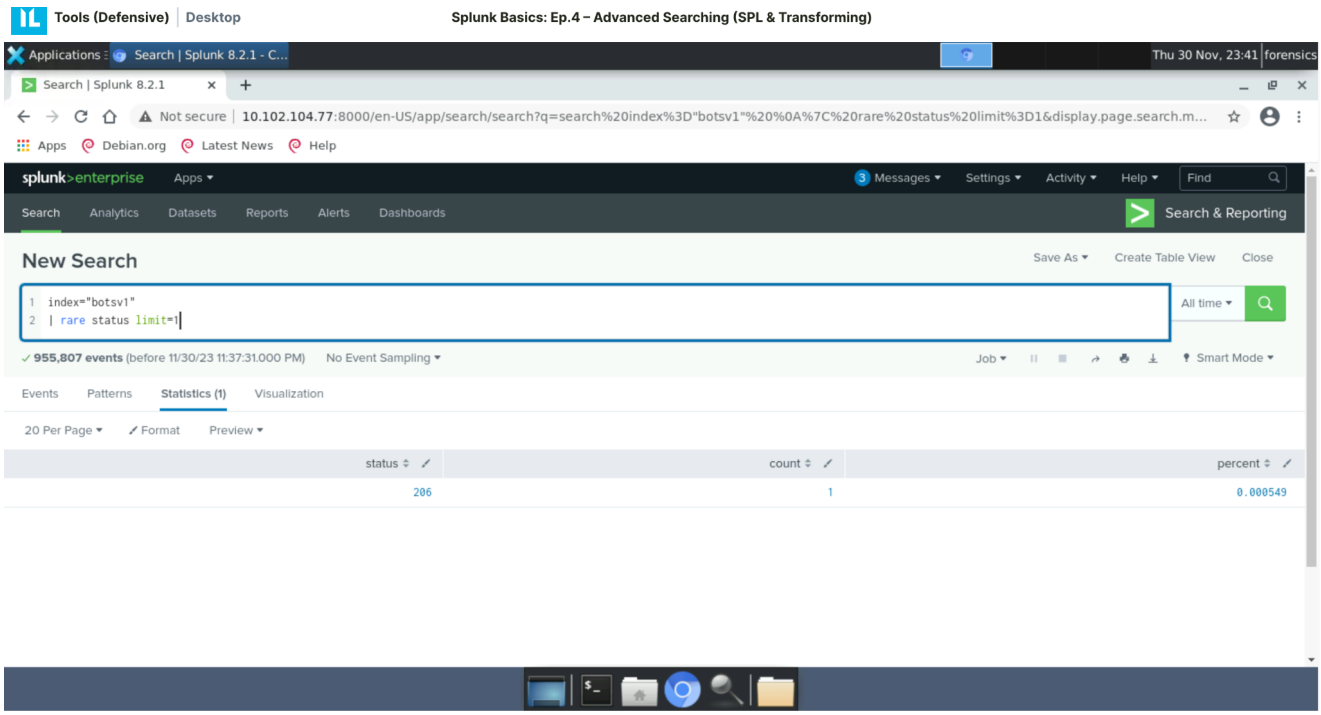
Splunk Basics: Ep.4 – Advanced Searching (SPL & Transforming)

Q1: Perform a search that lists the most common (top) “http_method” field values from the index “botsv1”. What percentage is given for the most common http_method present in the dataset?

Answer: 68
This task gives us an overview of how to make sense of the statistics tab in the splunk to get an aggregated view.

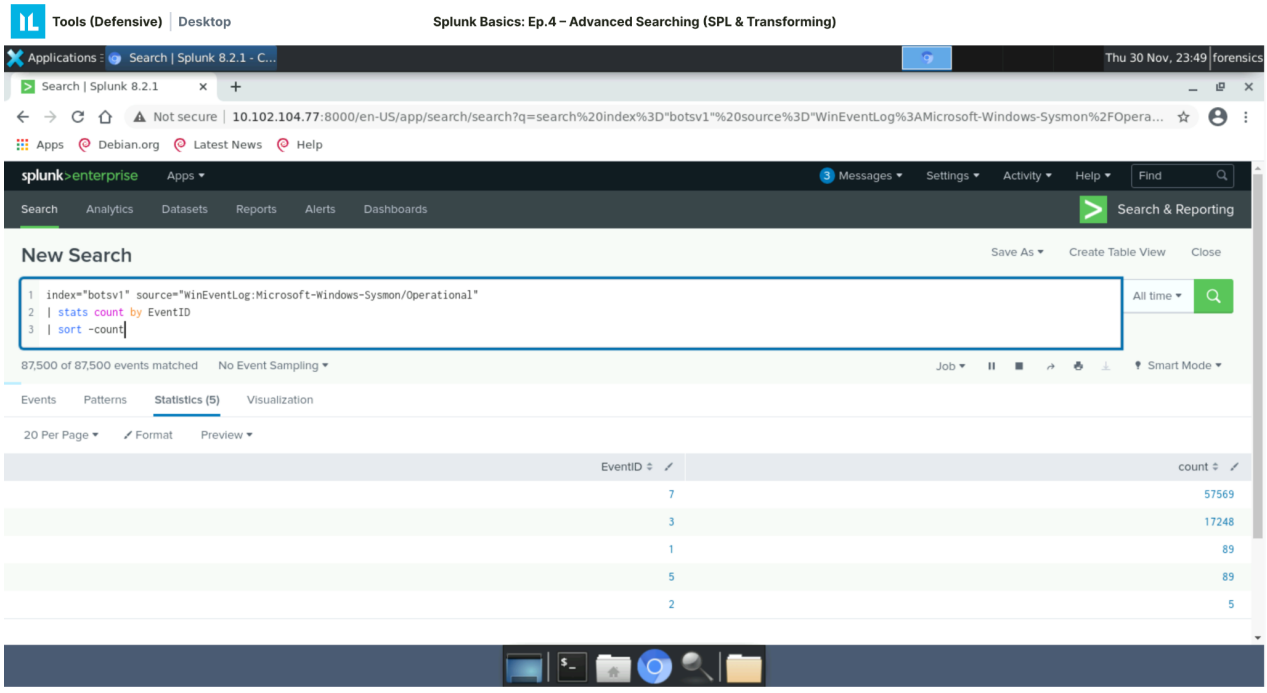


Q2: Perform a search that lists only the least common (rare) “status” field value from the index “botsv1”. What is the status code given?
Answer: 206



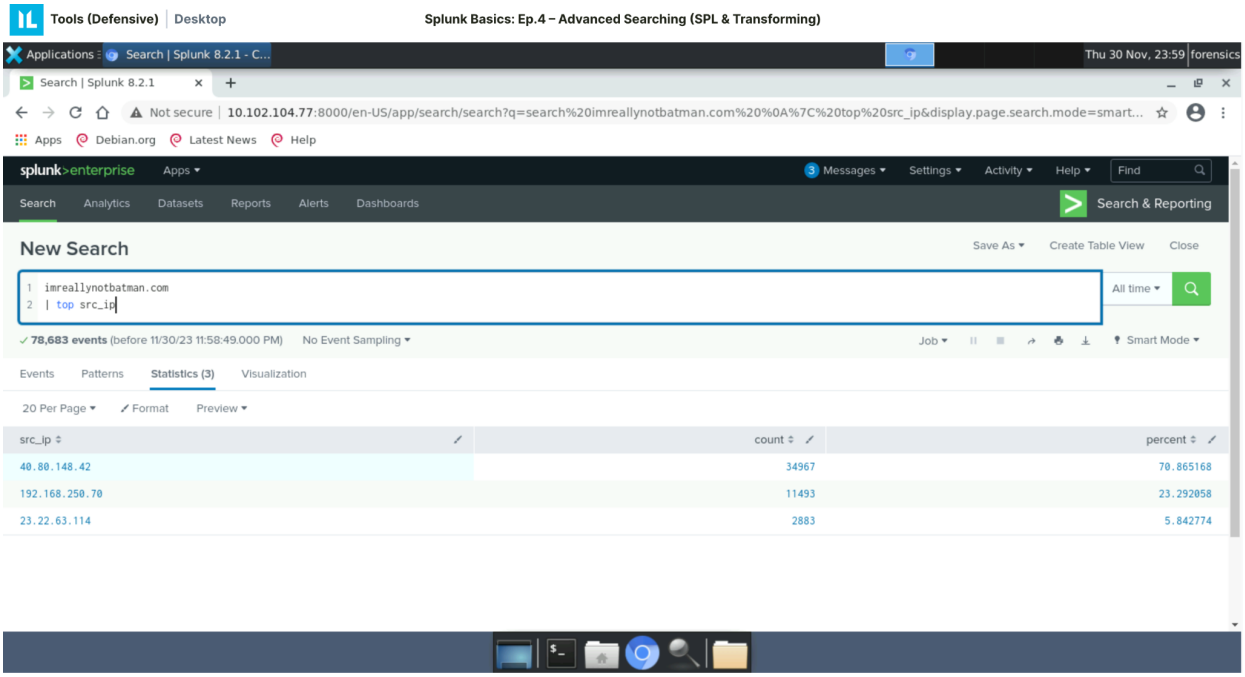
Q3: Perform a search using the stats command to count the number of events present by the field 'EventID' from the Source 'WinEventLog:Microsoft-Windows-Sysmon/Operational'. What is the EventID with the second most events?

Answer: 3

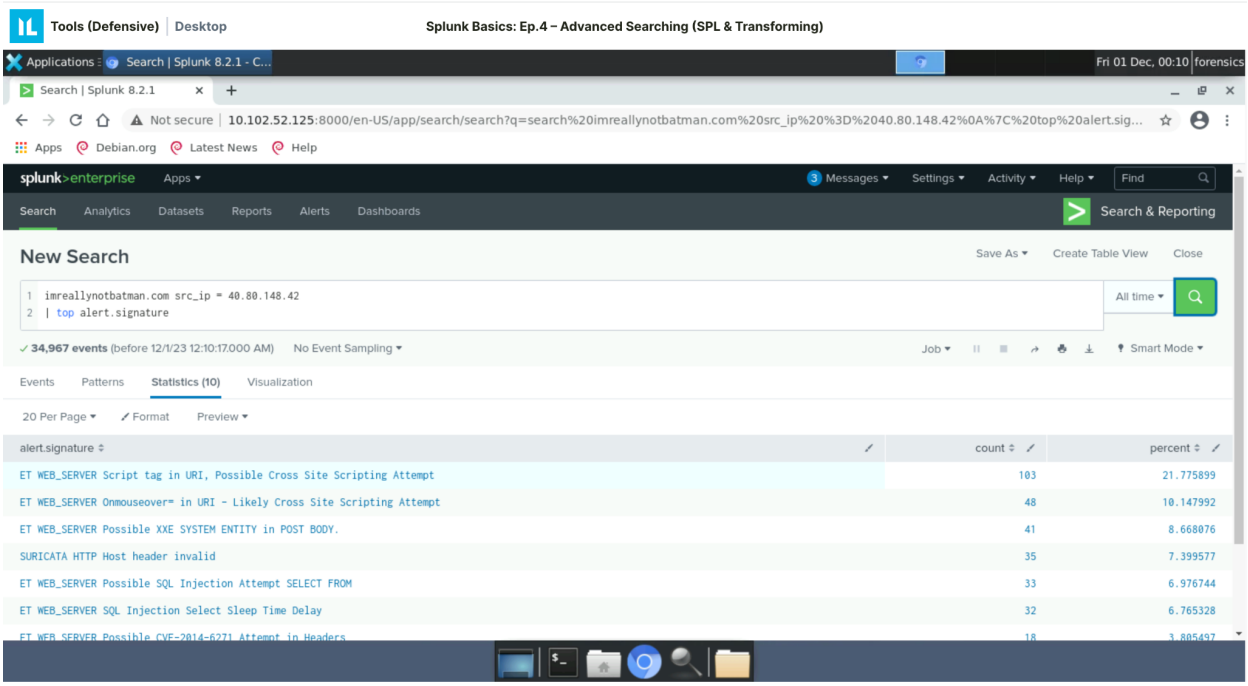


Q4: Perform a search for the domain “imreallynotbatman.com” and then use the 'top' command to determine the IP address of an attacker scanning the domain mentioned above for web app vulnerabilities (i.e., the 'top' 'src_ip').

Answer: 40.80.148.42



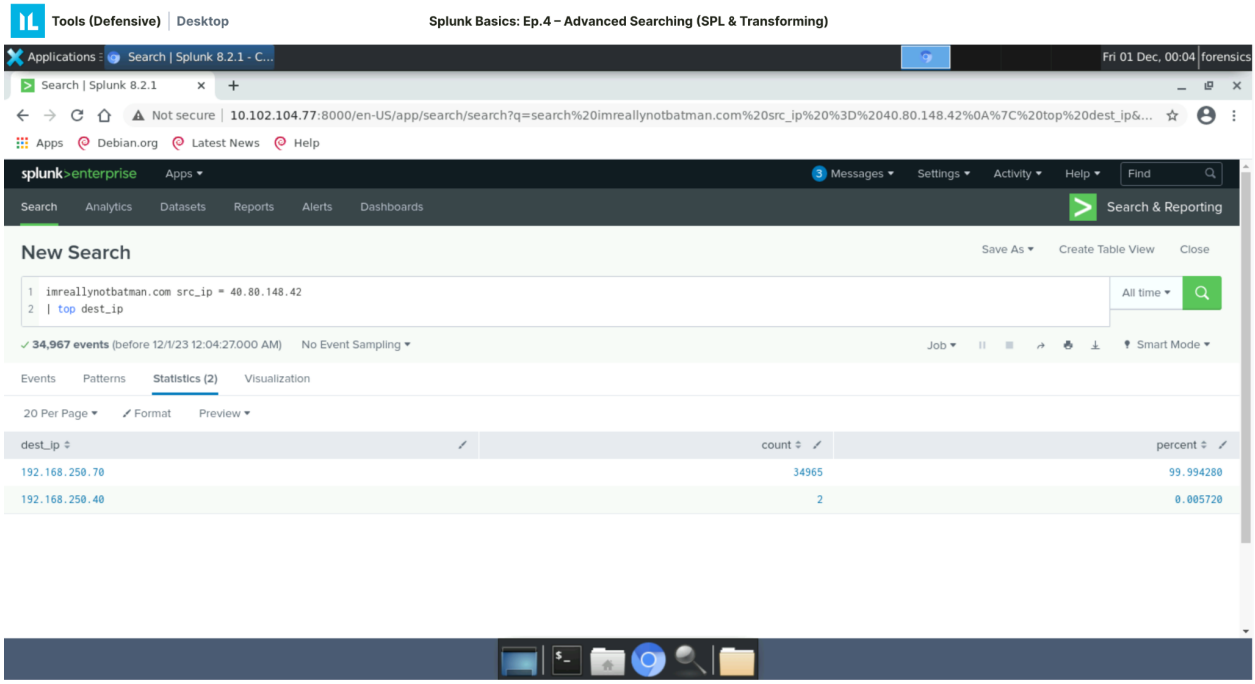
Q5: Perform a search using the domain and IP address from the previous question. What is the top 'alert.signature' field value reference?
Answer: ET WEB_SERVER Script tag in URI, Possible Cross Site Scripting Attempt
In this case our speculation from the previous answer that there might be an attack on the concerned IP is conformed.



Q6: Using the previously discovered ‘attacker IP’, determine the IP address of the web server being targeted by incoming attacker activity.

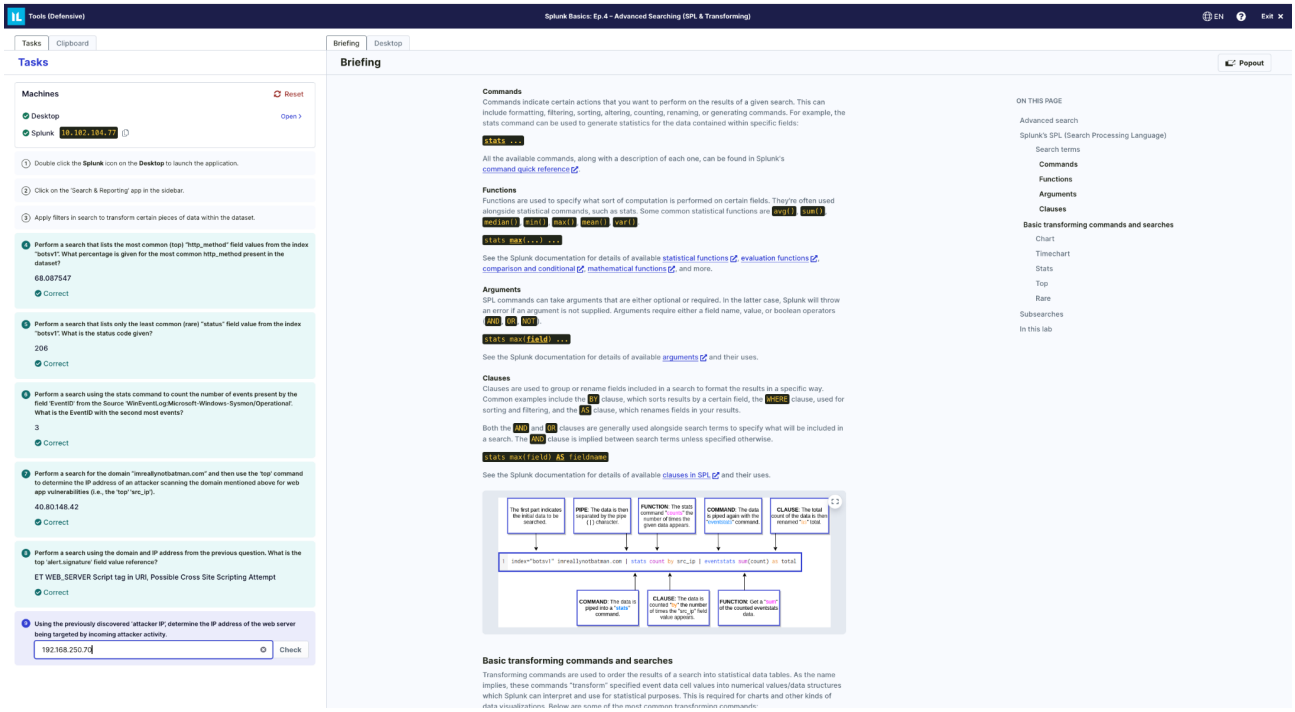
Answer: 192.168.250.70

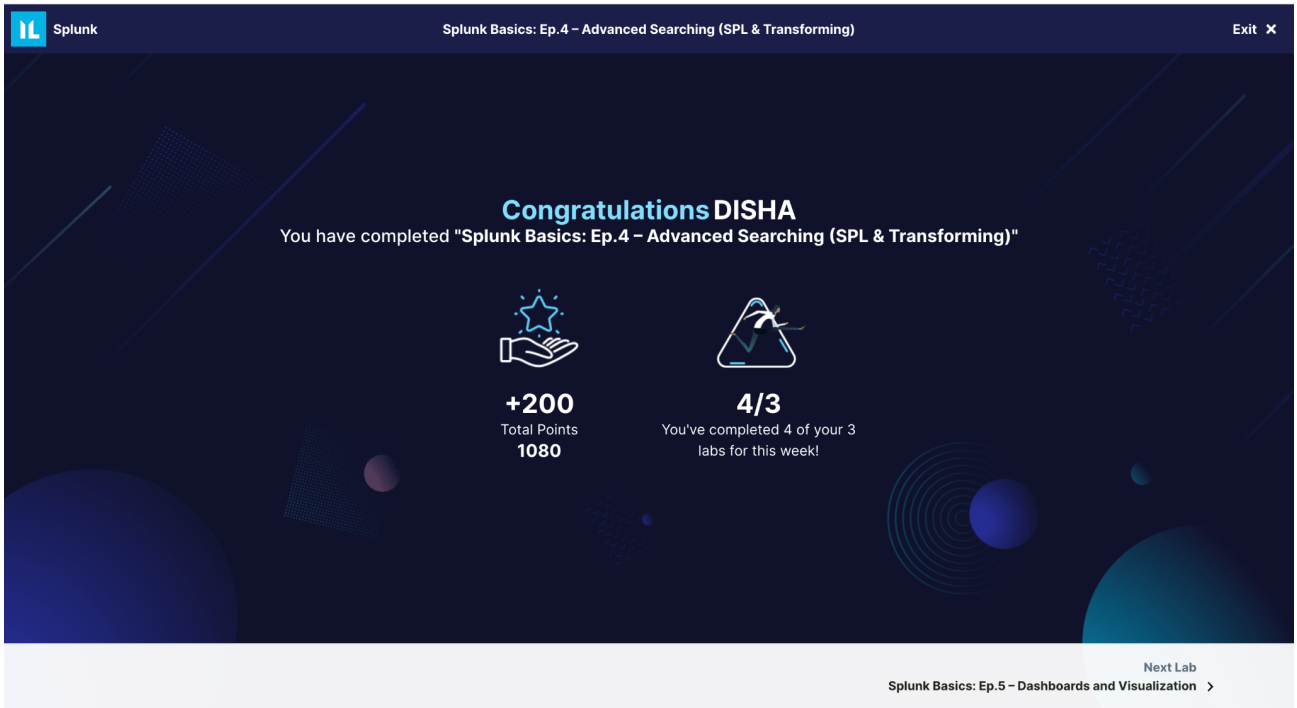
In this case we are able to find the source where attack is coming from. This can be further used to get more information of the attacker. Also to block the attacker.



Learnings:

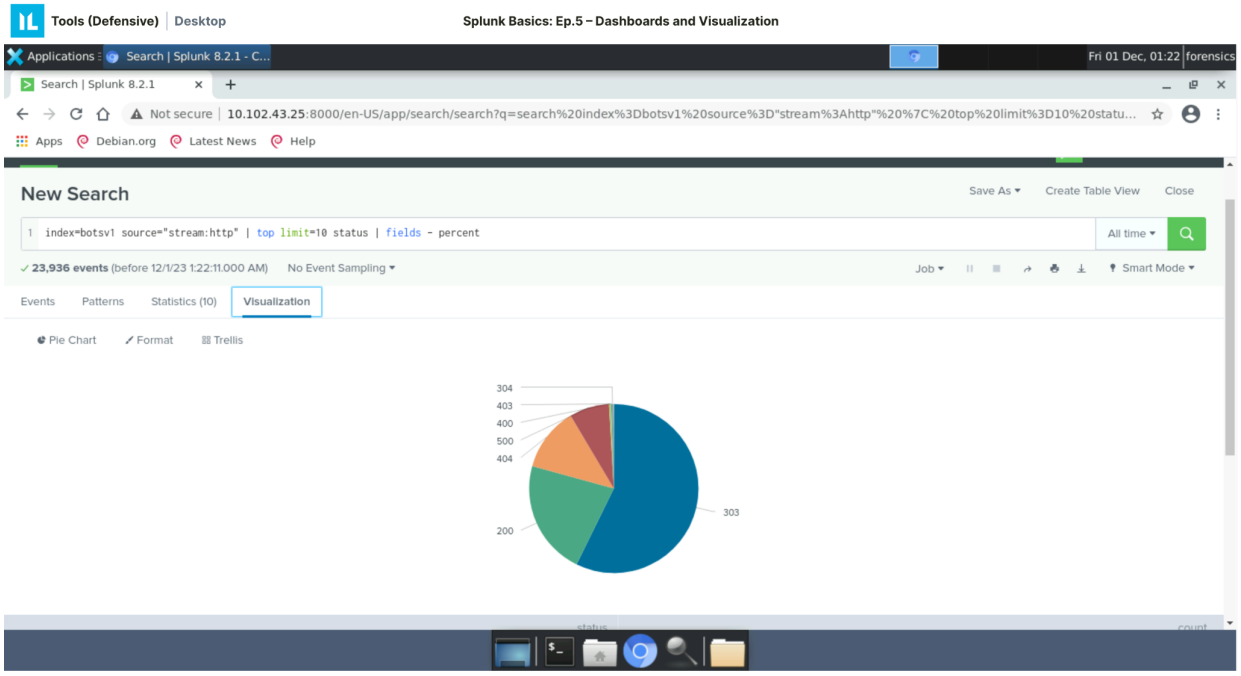
- Understanding Search Processing Language (SPL), Splunk's query language, is critical for dealing with event data, allowing for a wide variety of activities such as searching, filtering, and altering data.
- SPL Components: Understanding essential components like as search words, commands, functions, arguments, and clauses lays the groundwork for creating successful and exact search queries in Splunk.
- Learning converting commands such as Chart, Timechart, Stats, Top, and Rare enables the construction of visualizations and reports, assisting in the understanding of patterns and statistical summaries within the data.
- Subsearches: Recognizing subsearches as a strategy for narrowing down events by utilizing the results of one search as evidence for another improves the capacity to mix and evaluate diverse sets of data.
- Practical Application: The lab activities show how to use SPL commands and techniques in real-world situations, such as recognizing common and unusual field values, counting events based on certain criteria, and tracking probable attack scenarios using real-world data.


Tools (Defensive)
Splunk Basics: Ep.4 – Advanced Searching (SPL & Transforming)
EN
Exit
Tasks
Clipboard
Tasks
Machines
Reset
Desktop
Open
Splunk
10.102.184.77
1 Double click the Splunk icon on the Desktop to launch the application.
2 Click on the 'Search & Reporting' app in the sidebar.
3 Apply filters in search to transform certain pieces of data within the dataset.
4 Perform a search that lists the most common (top) "http_method" field values from the index "botsv1". What percentage is given for the most common http_method present in the dataset?
68.087547
Correct
5 Perform a search that lists only the least common (rare) "status" field value from the index "botsv1". What is the status code given?
206
Correct
6 Perform a search using the stats command to count the number of events present by the field 'EventID' from the Source 'WinEventLog:Microsoft-Windows-Sysmon/Operational'. What is the EventID with the second most events?
3
Correct
7 Perform a search for the domain "imreallynotbatman.com" and then use the 'top' command to determine the IP address of an attacker scanning the domain mentioned above for web app vulnerabilities (i.e., the 'top' 'src_ip').
40.80.148.42
Correct
8 Perform a search using the domain and IP address from the previous question. What is the top 'alert.signature' field value reference?
ET WEB_SERVER Script tag in URI, Possible Cross Site Scripting Attempt
Correct
9 Using the previously discovered 'attacker IP', determine the IP address of the web server being targeted by incoming attacker activity.
192.168.250.70
Check
Briefing
Desktop
Briefing
Popout
Commands
Commands indicate certain actions that you want to perform on the results of a given search. This can include formatting, filtering, sorting, altering, counting, renaming, or generating commands. For example, the stats command can be used to generate statistics for the data contained within specific fields:
stats ...
All the available commands, along with a description of each one, can be found in Splunk's command quick reference.
Functions
Functions are used to specify what sort of computation is performed on certain fields. They're often used alongside statistical commands, such as stats. Some common statistical functions are avg() sum() median() min() max() mean() var()
stats max(...) ...
See the Splunk documentation for details of available statistical functions, evaluation functions, comparison and conditional, mathematical functions, and more.
Arguments
SPL commands can take arguments that are either optional or required. In the latter case, Splunk will throw an error if an argument is not supplied. Arguments require either a field name, value, or boolean operators AND OR NOT.
stats max(field) ...
See the Splunk documentation for details of available arguments and their uses.
Clauses
Clauses are used to group or rename fields included in a search to format the results in a specific way. Common examples include the BY clause, which sorts results by a certain field, the CHECK clause, used for sorting and filtering, and the AS clause, which renames fields in your results.
Both the AND and OR clauses are generally used alongside search terms to specify what will be included in a search. The AND clause is implied between search terms unless specified otherwise.
stats max(field) AS fieldname
See the Splunk documentation for details of available clauses in SPL and their uses.
The first part indicates the initial data to be searched.
PIPE: The data is then separated by the pipe (|) character.
FUNCTION: The stats command "counts" the number of times the given data appears.
COMMAND: The data is piped again with the "eventstats" command.
CLAUSE: The total count of the data is then renamed "as" total.
1 index="botsv1" imreallynotbatman.com | stats count by src_ip | eventstats sum(count) as total
COMMAND: The data is piped into a "stats" command.
CLAUSE: The data is counted "by" the number of times the "src_ip" field value appears.
FUNCTION: Get a "sum" of the counted eventstats data.
Basic transforming commands and searches
Transforming commands are used to order the results of a search into statistical data tables. As the name implies, these commands "transform" specified event data cell values into numerical values/data structures which Splunk can interpret and use for statistical purposes. This is required for charts and other kinds of
ON THIS PAGE
Advanced search
Splunk's SPL (Search Processing Language)
Search terms
Commands
Functions
Arguments
Clauses
Basic transforming commands and searches
Chart
Timechart
Stats
Top
Rare
Subsearches
In this lab

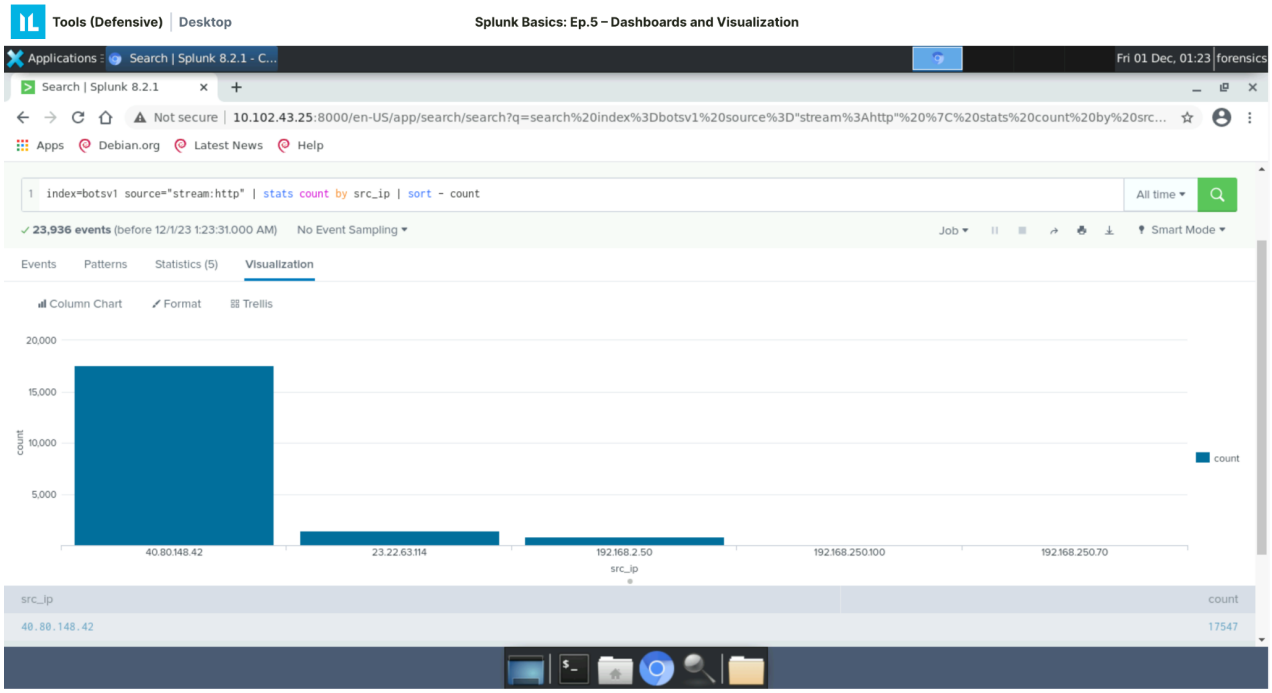

Splunk
Splunk Basics: Ep.4 – Advanced Searching (SPL & Transforming)
Exit
Congratulations DISHA
You have completed "Splunk Basics: Ep.4 – Advanced Searching (SPL & Transforming)"
+200
Total Points
1080
4/3
You've completed 4 of your 3 labs for this week!
Next Lab
Splunk Basics: Ep.5 – Dashboards and Visualization

Splunk Basics: Ep.5 - Dashboards and Visualization

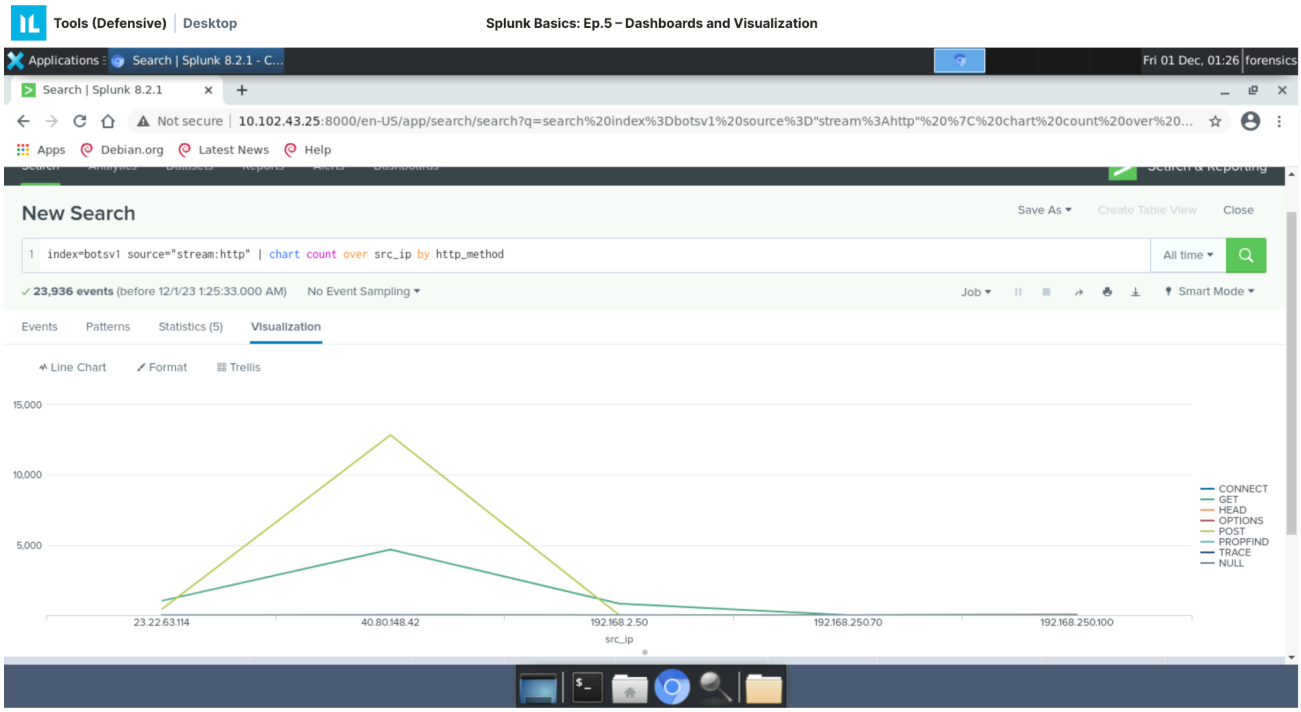
First chart

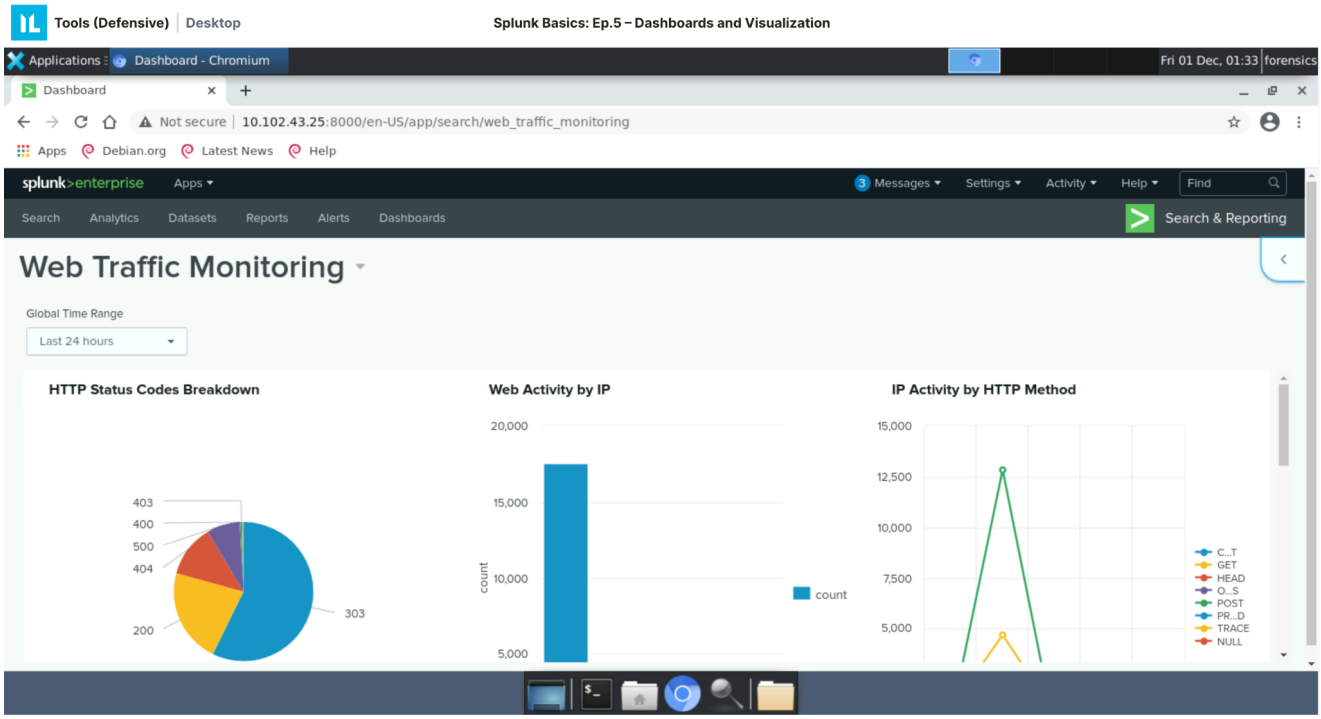


Second chart



Third Chart:



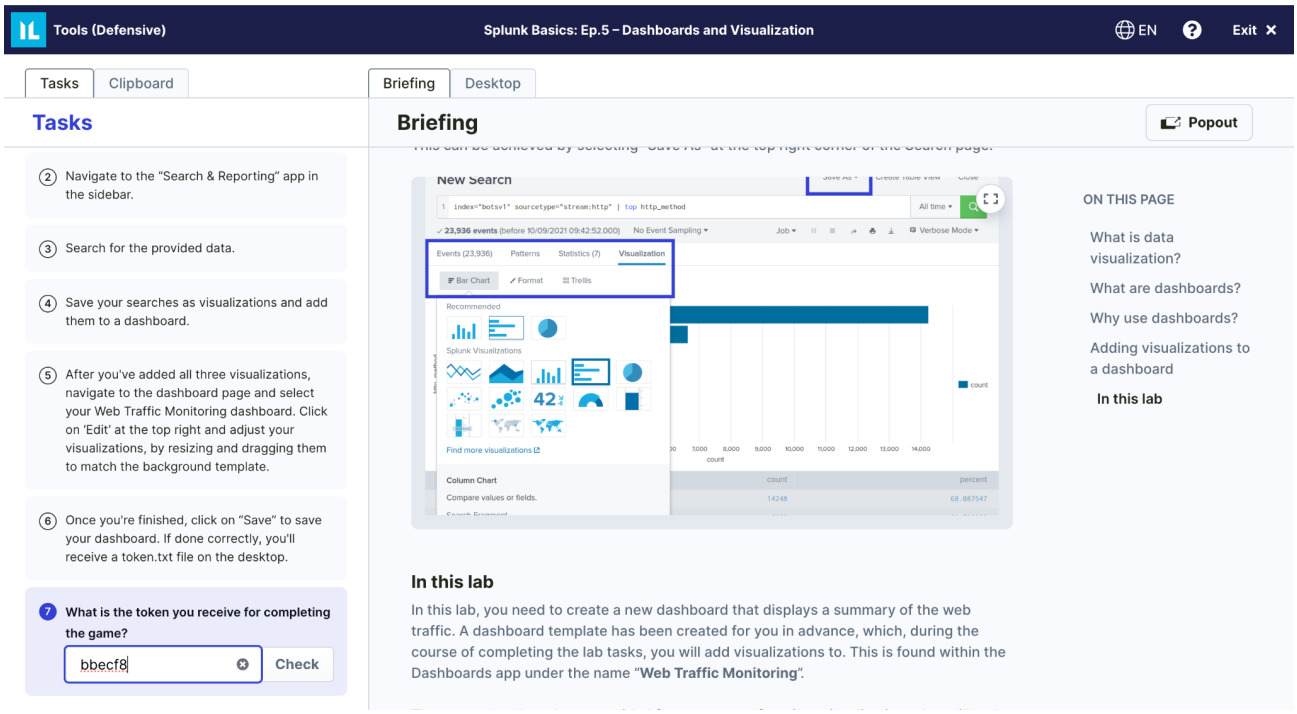


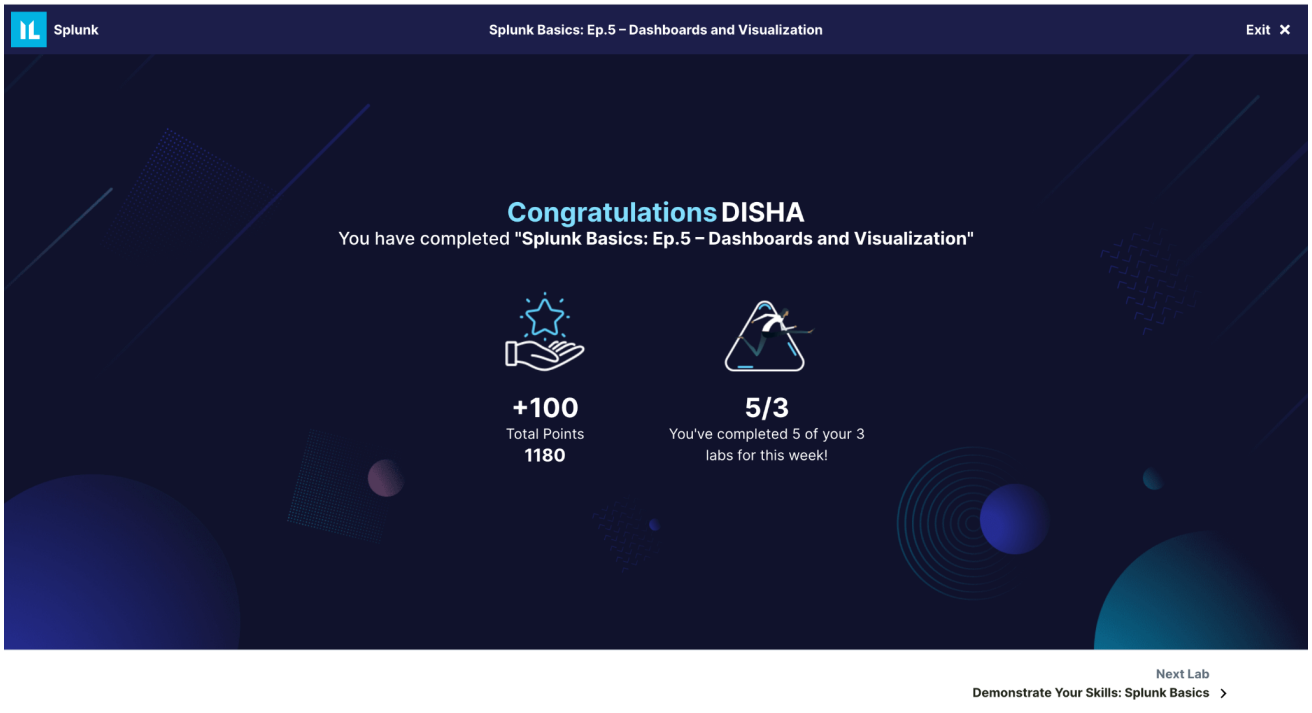
Learnings:

- Dashboard Creation: The lab gave hands-on experience in developing a dashboard in Splunk, highlighting the necessity of structuring and visualizing summary data.
- Visualization Types: Understanding how to translate search results into multiple visualization forms, such as Pie Charts, Bar Charts, and Line Charts, improves the capacity to effectively present data insights.

- Dashboard Customization: The process of adding, naming, and organizing visualizations inside a dashboard was explained, demonstrating how to modify dashboards to specific monitoring and data presentation needs.

Final Token:

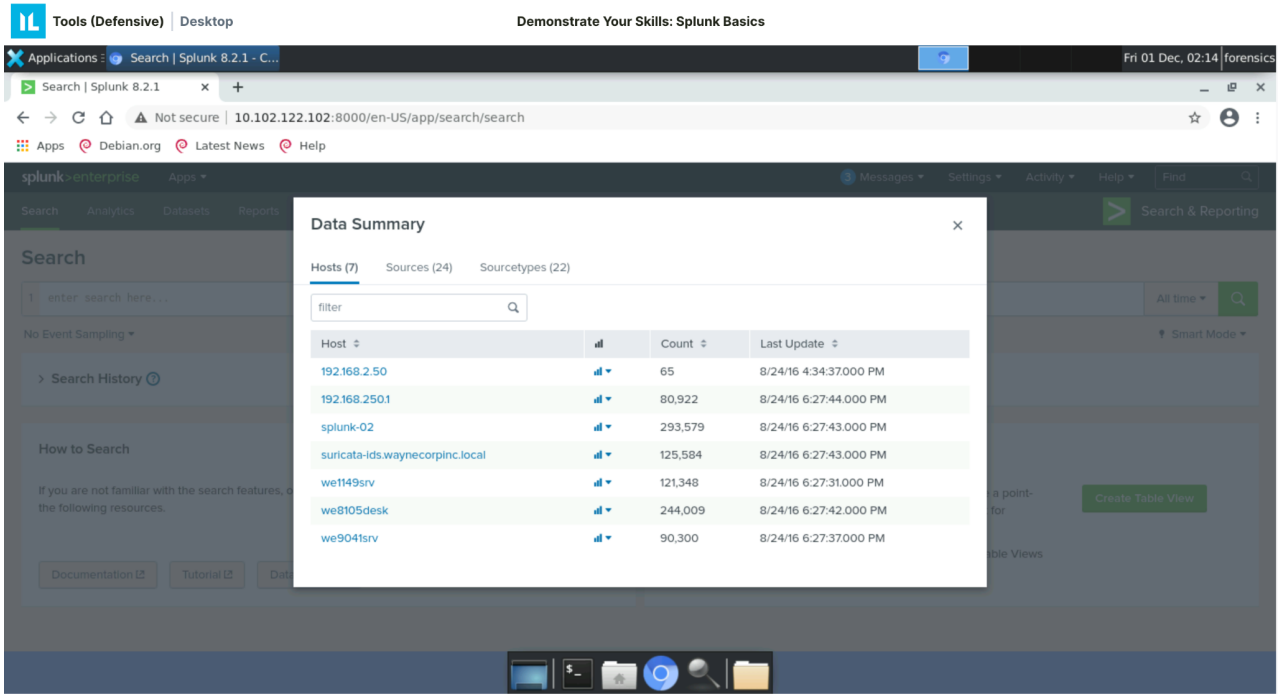




Demonstrate Your Skills: Splunk Basics

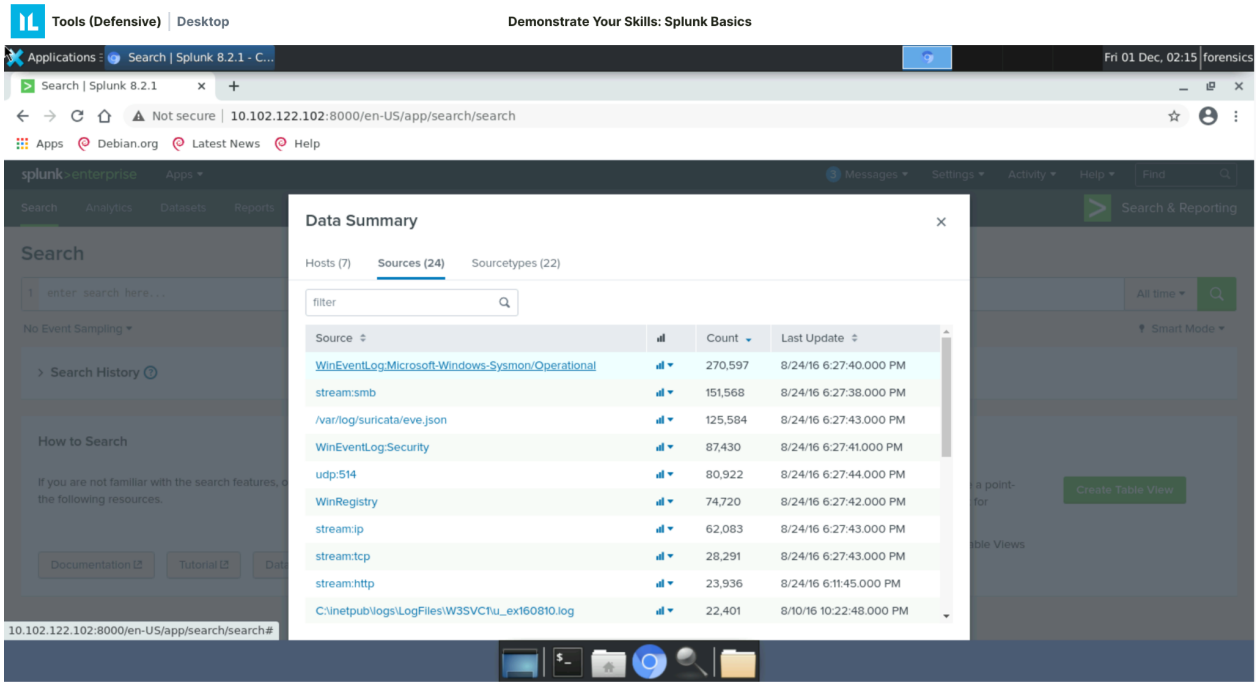
Q1:Select the Data Summary on the Search and Reporting App home page. How many hosts are there?

Answer: 7

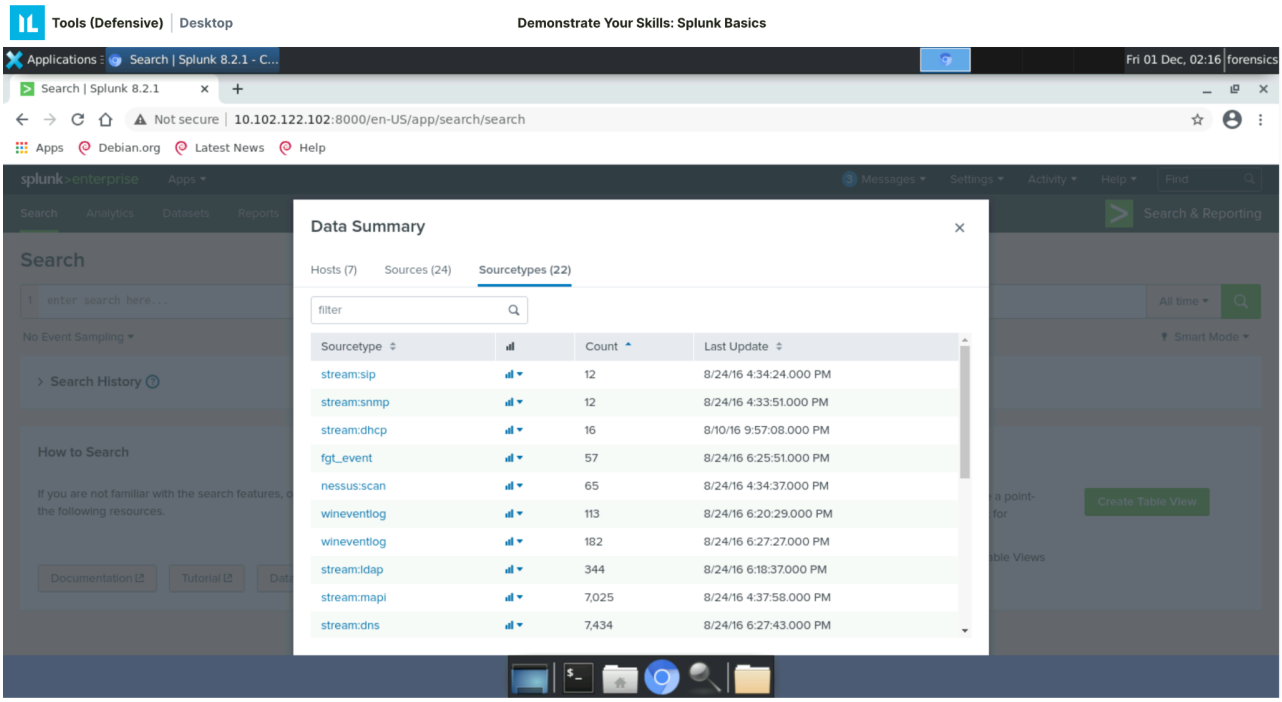


Q2:Looking at the data summary, which source has the highest count?

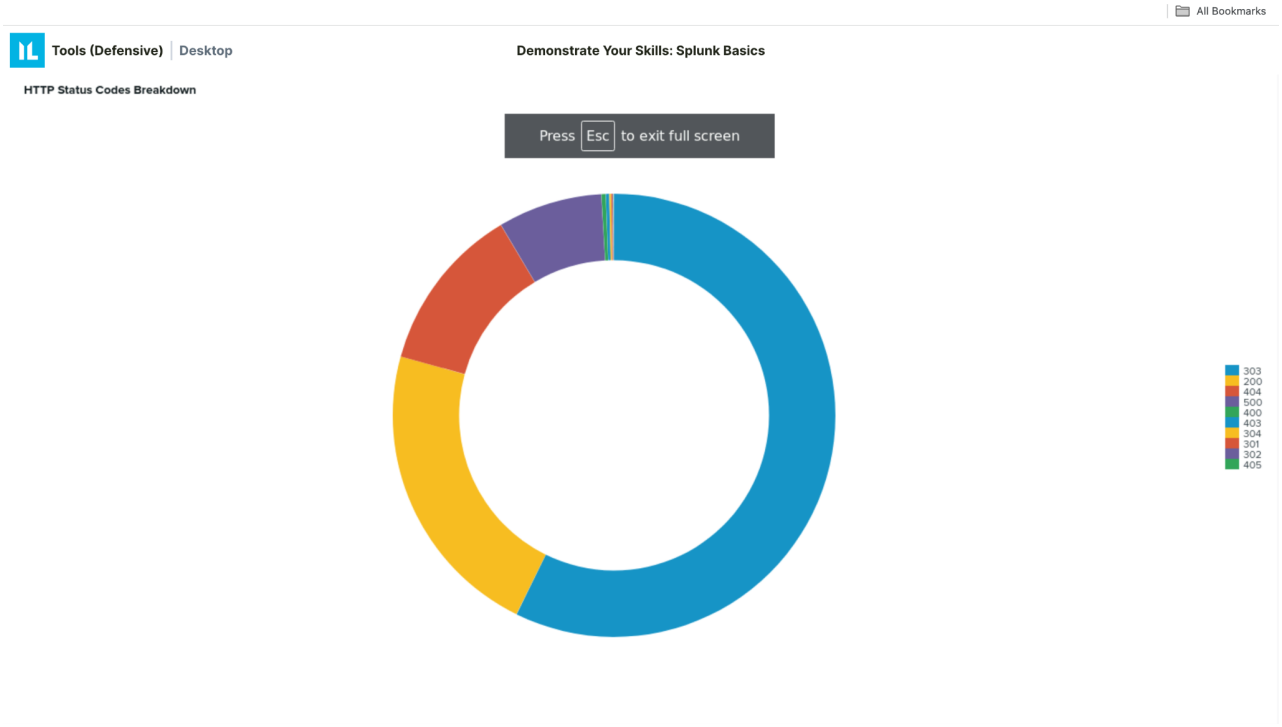
Answer: WinEventLog:Microsoft-Windows-Sysmon/Operational



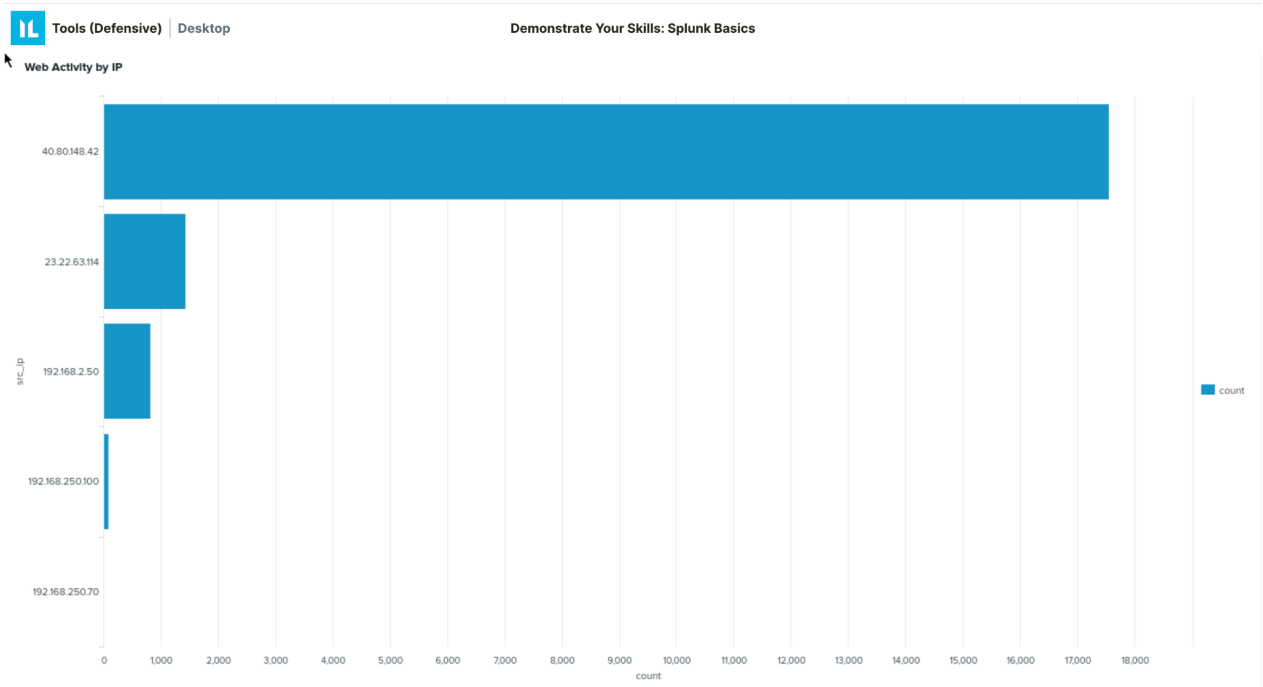
Q3: Looking at the data summary, provide one of the two sourcetypes with the lowest count.
Answer: stream:snmp



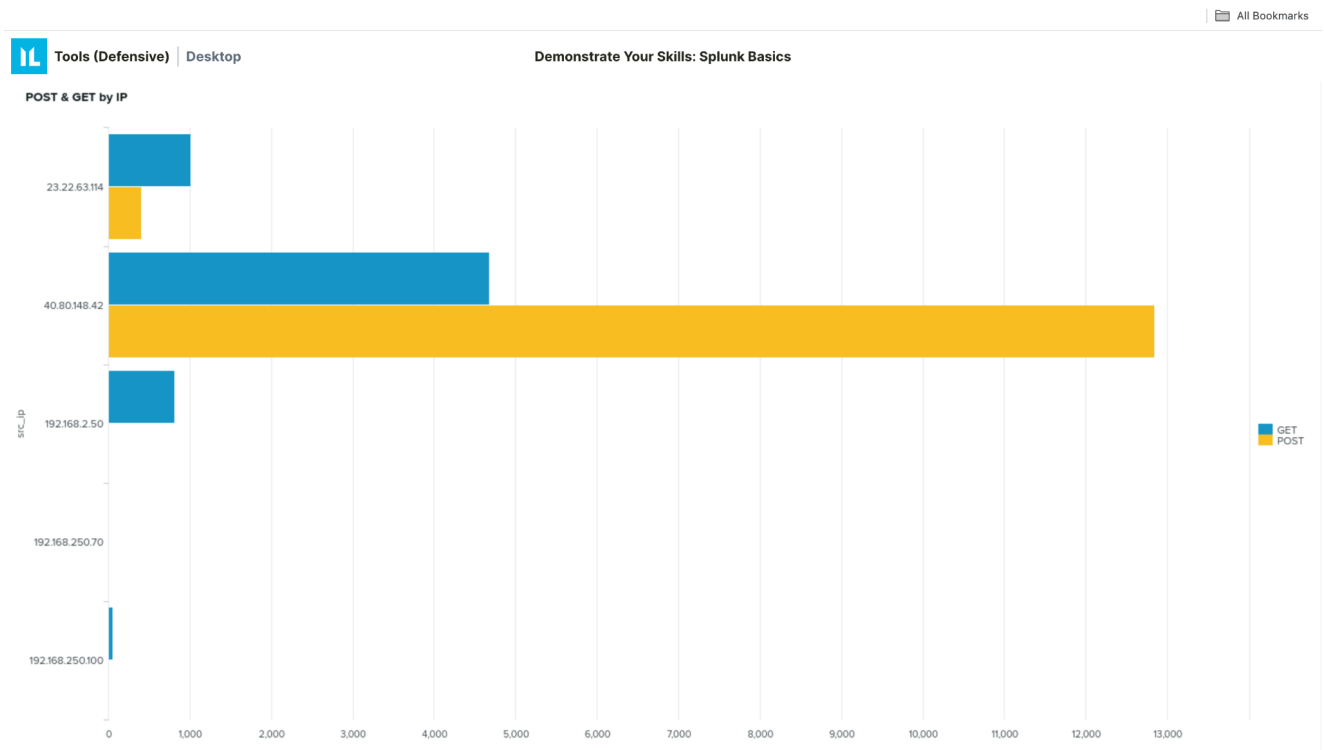
Q4:Navigate to the dashboard called Web Traffic Monitoring. Which status code appears the most?
Answer: 303



Q5:Which of the active IP addresses has the lowest number of requests?
Answer: 192.168.250.70

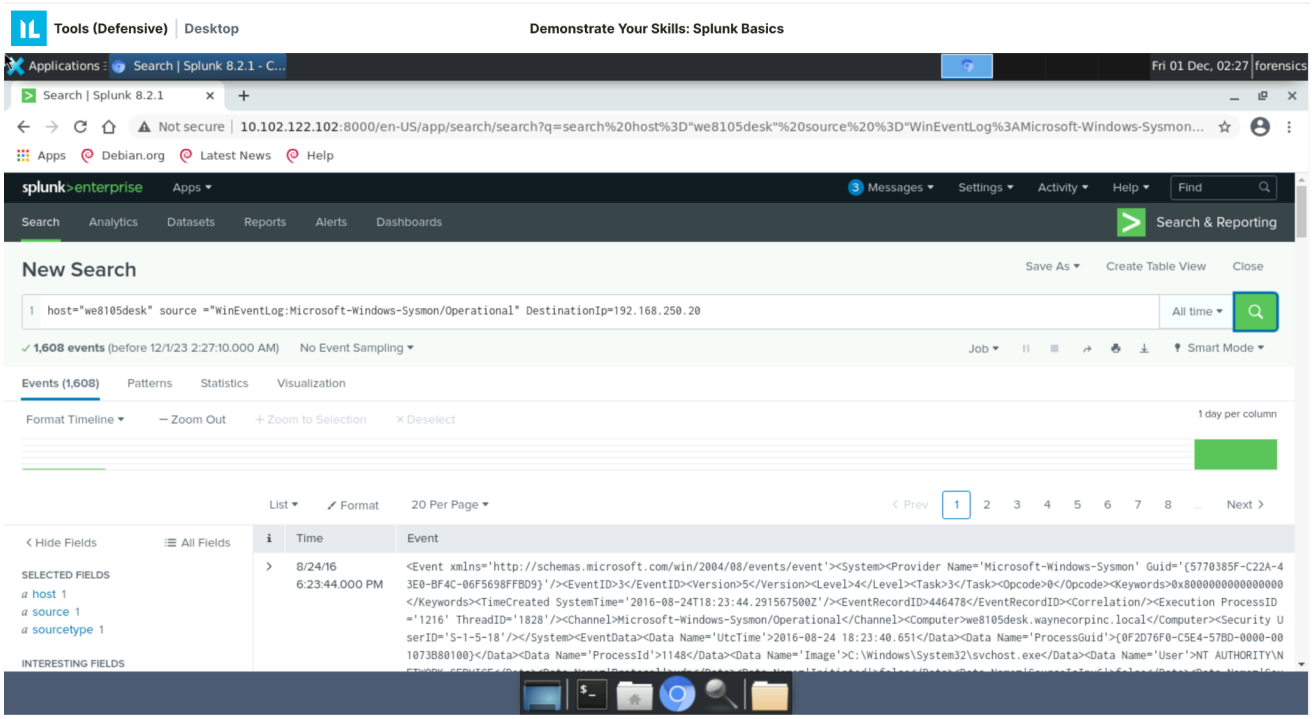


Q6:Which IP has the highest number of POST requests?
Answer: 40.80.148.42

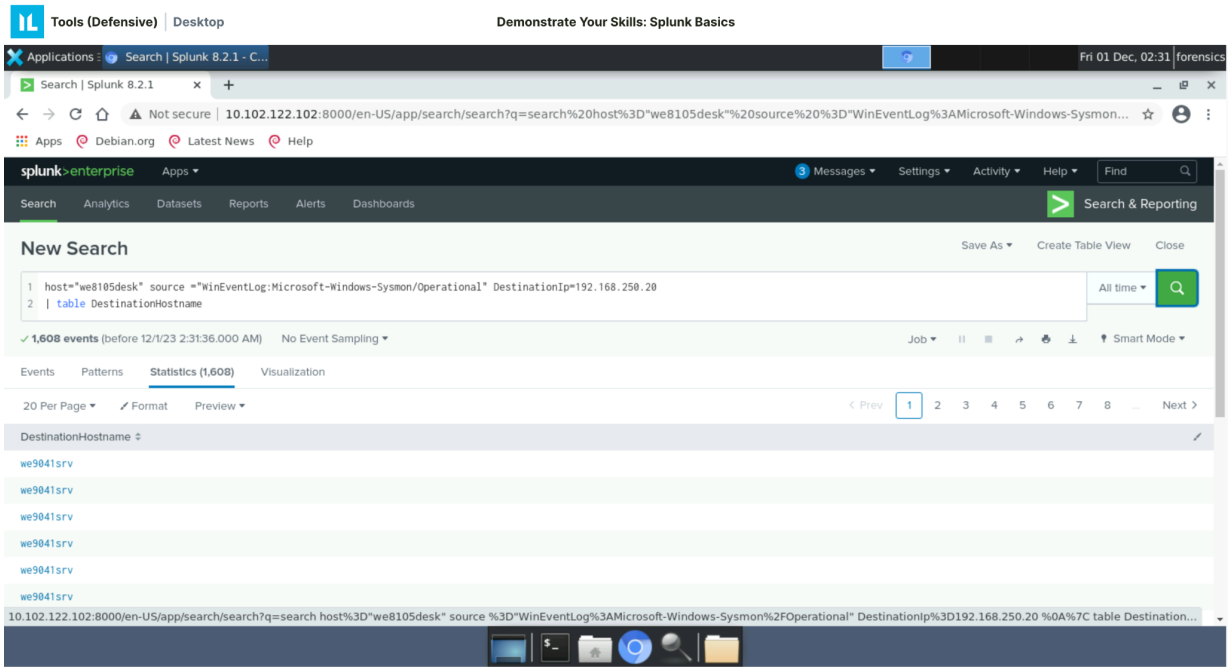


Q7:Search for the host we8105desk, source WinEventLog:Microsoft-Windows-Sysmon/Operational, and the 192.168.250.20 DestinationIp. How many events are returned?

Answer: 1608
This answer arrived when the search query Query = host=we8105desk source="WinEventLog:Microsoft-Windows-Sysmon/Operational" DestinationIp=192.168.250.20 was used.

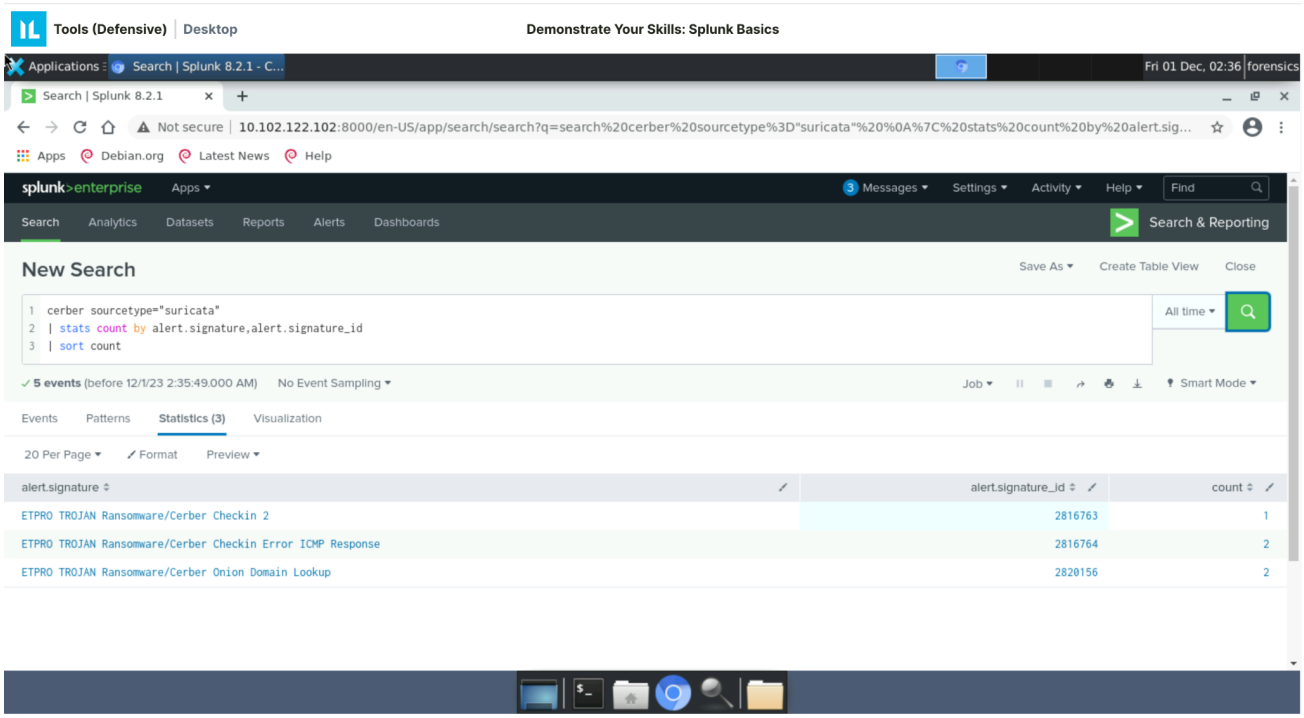


Q8: Looking at the results from the previous question, find the host name of the remote server. What is the DestinationHostname?
Answer: we9041srv



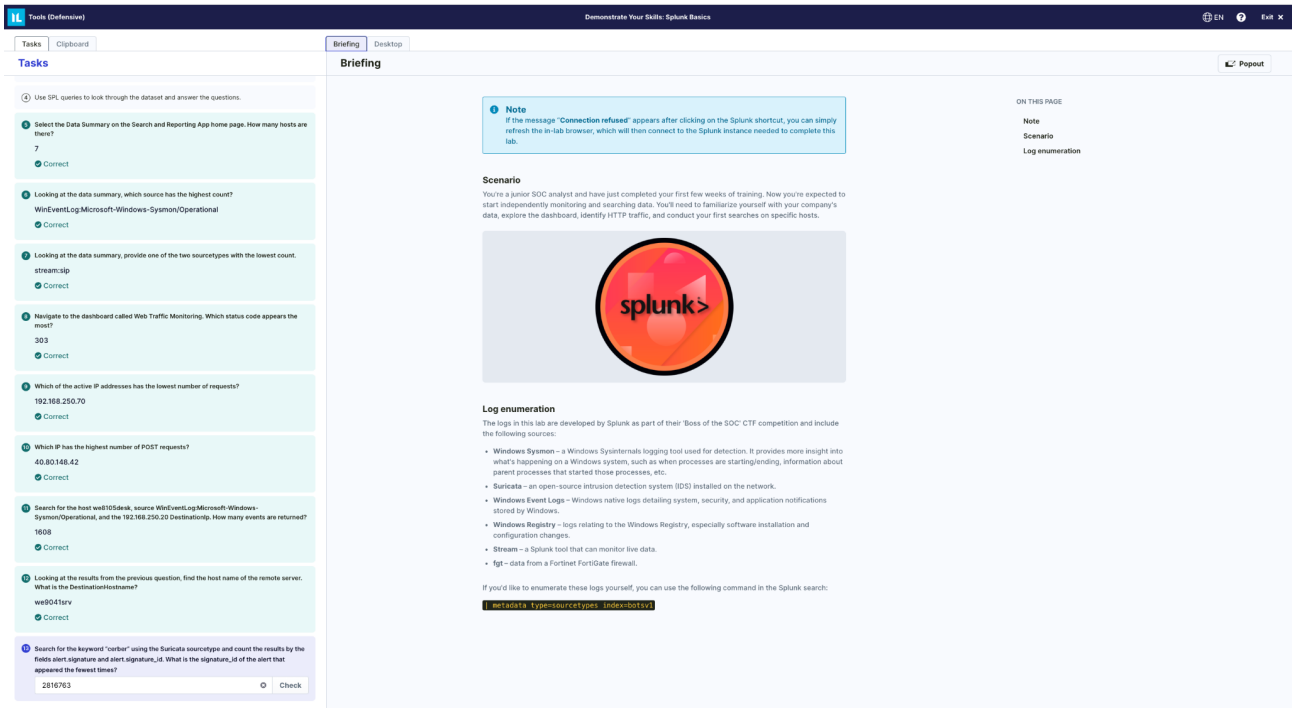
Q9:Search for the keyword “cerber” using the Suricata sourcetype and count the results by the fields alert.signature and alert.signature_id. What is the signature_id of the alert that appeared the fewest times?

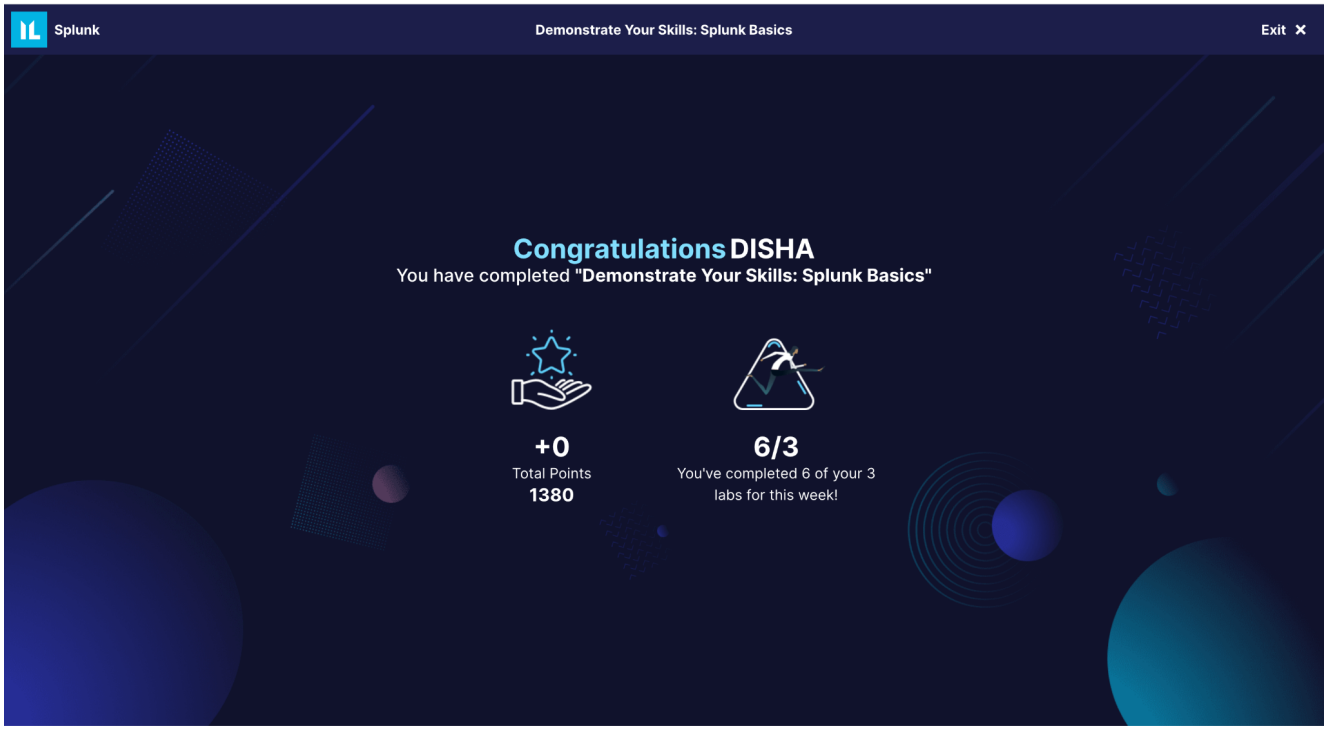
Answer: 2816763
To get this answer we had to construct search query as below.
Query= sourcetype=suricata cerber | stats count by alert.signature, alert.signature_id | sort count



Learnings:

- Log Enumeration:Understanding the sources of logs, such as Windows Sysmon, Suricata, Windows Event Logs, Windows Registry, Stream, and fgt, provides insight into the diverse data available for monitoring and analysis within the Splunk environment.
- Metadata Exploration: The use of metadata commands, like `| metadata type=sourcetypes index=botsv1`, facilitates the enumeration of logs, helping identify the types and sources of data available for analysis.
- Dashboard Navigation: Navigating and exploring dashboards, such as "Web Traffic Monitoring," allows for a more visual representation of data, aiding in quickly identifying patterns, anomalies, and critical information.
- Search Queries Performing targeted searches, such as searching for specific hosts, source types, and keywords like "cerber," enhances the ability to extract relevant information from the vast dataset and focus on specific security-related events.
- Data Analysis: Extracting meaningful insights from search results, including identifying the most common status codes, the IP with the lowest requests, and the host name of a remote server, demonstrates the practical application of data analysis skills for a SOC analyst.


Tools (Defensive)
Demonstrate Your Skills: Splunk Basics
EN
Exit
Tasks
Clipboard
Briefing
Desktop
Tasks
Briefing
Popout
Use SPL queries to look through the dataset and answer the questions.
5 Select the Data Summary on the Search and Reporting App home page. How many hosts are there?
7
Correct
6 Looking at the data summary, which source has the highest count?
WinEventLog:Microsoft-Windows-Sysmon/Operational
Correct
7 Looking at the data summary, provide one of the two sourcetypes with the lowest count.
stream:sip
Correct
8 Navigate to the dashboard called Web Traffic Monitoring. Which status code appears the most?
303
Correct
9 Which of the active IP addresses has the lowest number of requests?
192.168.250.70
Correct
10 Which IP has the highest number of POST requests?
40.80.148.42
Correct
11 Search for the host we8105desk, source WinEventLog:Microsoft-Windows-Sysmon/Operational, and the 192.168.250.20 DestinationIp. How many events are returned?
1608
Correct
12 Looking at the results from the previous question, find the host name of the remote server. What is the DestinationHostname?
we9041srv
Correct
13 Search for the keyword "cerber" using the Suricata sourcetype and count the results by the fields alert.signature and alert.signature_id. What is the signature_id of the alert that appeared the fewest times?
2816763
Check
Note
If the message "Connection refused" appears after clicking on the Splunk shortcut, you can simply refresh the in-lab browser, which will then connect to the Splunk instance needed to complete this lab.
Scenario
You're a junior SOC analyst and have just completed your first few weeks of training. Now you're expected to start independently monitoring and searching data. You'll need to familiarize yourself with your company's data, explore the dashboard, identify HTTP traffic, and conduct your first searches on specific hosts.
splunk>
Log enumeration
The logs in this lab are developed by Splunk as part of their 'Boss of the SOC' CTF competition and include the following sources:
Windows Sysmon – a Windows Sysinternals logging tool used for detection. It provides more insight into what's happening on a Windows system, such as when processes are starting/ending, information about parent processes that started those processes, etc.
Suricata – an open-source intrusion detection system (IDS) installed on the network.
Windows Event Logs – Windows native logs detailing system, security, and application notifications stored by Windows.
Windows Registry – logs relating to the Windows Registry, especially software installation and configuration changes.
Stream – a Splunk tool that can monitor live data.
fgt – data from a Fortinet FortiGate firewall.
If you'd like to enumerate these logs yourself, you can use the following command in the Splunk search:
| metadata type=sourcetypes index=botsv1
ON THIS PAGE
Note
Scenario
Log enumeration


Splunk
Demonstrate Your Skills: Splunk Basics
Exit
Congratulations DISHA
You have completed "Demonstrate Your Skills: Splunk Basics"
+0
Total Points
1380
6/3
You've completed 6 of your 3 labs for this week!

Writeup on Overall Learning:

The six Splunk lab activities provided a thorough grasp of the platform's capabilities and practical applications in network monitoring. The tour began with an introduction of Splunk fundamentals, stressing the tool's ability to collect different data from several sources inside an organization. The importance of source types became clear, emphasizing their function in data structuring and indexing, which is required for effective search queries and data analysis.

A detailed look into the search functionality evolved as the labs continued. Each phase gave insights into improving and supplementing search results, from simple searches to sophisticated queries utilizing the Search Processing Language (SPL). The ability to use Boolean expressions, wildcards, and escape characters improved the versatility and precision of search queries, which is critical for circumstances such as researching Windows file paths.

SPL components and transformational commands like Chart, Timechart, Stats, Top, and Rare were added during the research of sophisticated searching. This information was put to use in real-world scenarios such as finding common and uncommon field values, counting events, and tracking probable attack scenarios using actual data.

The laboratories then moved on to visualization, focusing the design and customisation of dashboards. This hands-on training emphasized the significance of arranging and presenting summary data, resulting in a significant skill set for efficient data presentation.

The last set of activities demonstrated the use of newly acquired abilities in the Splunk environment. The ability to navigate the Search and Reporting App, evaluate data summaries, and extract useful information from dashboards indicated a comprehensive awareness of Splunk's capabilities. The capacity to conduct focused searches and evaluate results for specific scenarios, such as identifying hosts, evaluating web traffic, and tracking security-related events, demonstrated the practical application of information gained.

In essence, the six lab exercises not only provided a thorough overview of Splunk's features, but also equipped participants with practical skills required for network monitoring, data analysis, and security incident response. The exploration of Splunk's capabilities has provided me with a versatile skill set applicable to real-world scenarios in the dynamic landscape of network and security operations.