splunk>

Troubleshooting Splunk Enterprise - Class Lab Exercises

Training Lab Environment

Throughout the course, you will access a standalone Splunk Enterprise instance in a private network environment. Your instructor will provide the following information for your Splunk instance:

Student ID {**x**}:	...................the last two digits of the internal ip address
SSH/RDC username {**user**}:	....
SSH/RDC password {**password**}:	....
Server address {**public_DNS**}:	.... **`<session>-<x>.class.splunk.com`**
splunk>enterprise Web access:	.... **`https://<public_DNS>:8000`**
splunk>enterprise Console access:	.... SSH or RDC
Splunk username:	....**`admin`**
Splunk user password:	................................same as SSH/RDC user password
`SPLUNK_HOME`	………………………refers to the **`/opt/splunk`** directory

You also have an unspecified Splunk deployment topology. As you progress through the lab activities, you will learn the complete topology of your Splunk lab environment.

To support the lab activities, download the lab-support apps from: https://splk.it/edu-tse-82

Check with your instructor if your class is using an alternate source to obtain the apps.

To edit the **`.conf`** files:

Splunk on Linux	Use **`vi`** or **`nano`**
Splunk on Windows	Use **Notepad++**

Typographical Convention

- **Blue** text in steps indicates text to be **added** or **replaced**.

Module 1 Lab Exercise – Know Your Environment

Description

In this exercise, you will:

- Confirm the lab access information provided by your instructor
- Perform basic Splunk configuration operations
- Capture a baseline condition of your Splunk instance with the diag UI
- Index and search the diag for system information

Steps

Task 1: Access your Splunk CLI terminal and change the basic Splunk settings.

1. Use SSH or RDC to connect to your dedicated Splunk server.

🐧 Splunk on Linux	Use SSH or PuTTY `ssh {user}@{public_DNS}`
⊞ Splunk on Windows	Use **Remote Desktop Connection**: a. Start Remote Desktop Connection and enter `{public_DNS}`. b. Enter `student` and the assigned password when prompted. c. In the remote Windows desktop, right-click the **Windows PowerShell**. d. Select **Run as administrator**.

2. Go to the Splunk `bin` directory.

🐧	`cd /opt/splunk/bin`
⊞	`cd "C:\Program Files\Splunk\bin"`

3. With Splunk CLI commands, confirm Splunk is running and retrieve the Splunk version.

```
./splunk status
Splunkd: Running (pid 1244)

./splunk version
Splunk 8.2.0 (build e053ef3c985f)
```

4. Using CLI, rename the Splunk server name and Default host name using the following convention You will be prompted for the Splunk administrator username and password:

5. Splunk server name: splunk## where ## is your {student-ID}

6. Default host name: splunk## where ## is your {student-ID}

```
./splunk set servername splunk##
  Splunk username: admin
  Password: {password}
  You need to restart the Splunk Server (splunkd) for your changes to take
  effect.

./splunk set default-hostname splunk##
  Default hostname set.
  You need to restart the Splunk Server (splunkd) for your changes to take
  effect.

./splunk restart
  ...
```

Task 2: Access Splunk Web and install the tse_lab01.spl app.

7. Direct your web browser to your **splunk>enterprise** instance: **`https://{public_DNS}:8000`**

   Log in as **`admin`** using your assigned password **`{password}`**.

> **NOTE:** When you get a prompt *Help us improve Splunk software*, click **Skip**.

8. Click **Apps** > **Manage Apps** and click **Install app from file**.

9. Browse for **`tse_lab01.spl`** and **Upload**.

   The **tse_lab01** app installs a few Splunk configuration files (**`indexes.conf`**, **`inputs.conf`**, etc.) that will be used throughout the lab exercises.

10. Click **Administrator** > **Account Settings** and change the **Full name** to *your name* and click **Save**.

11. Navigate to **Settings** > **Server settings** > **General settings** and confirm the changed values:

    Splunk server name: **`splunk##`**
    Default host name: **`splunk##`**

Task 3: Generate a baseline diag and index.

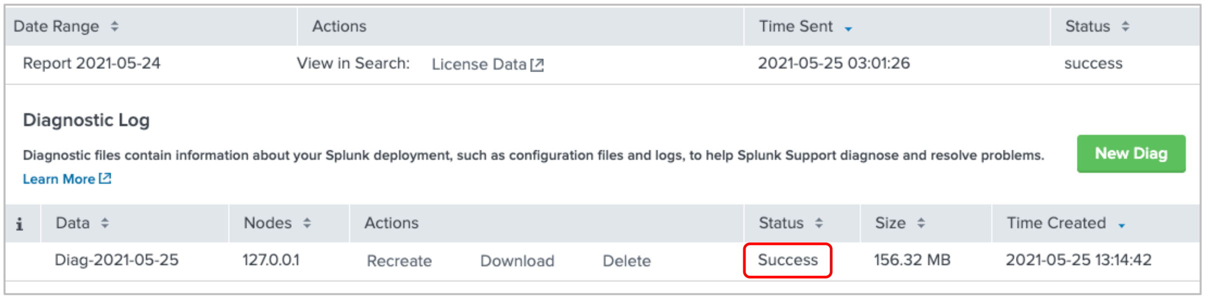
12. Navigate to **Settings** > **Instrumentation**.

13. Click **New Diag**.

14. Select **`127.0.0.1`** and click **Next**.

15. Click **Create**.

    Wait until the **Status** transitions to **`Success`**.



Date Range	Actions	Time Sent	Status
Report 2021-05-24	View in Search: License Data	2021-05-25 03:01:26	success

Diagnostic Log

Diagnostic files contain information about your Splunk deployment, such as configuration files and logs, to help Splunk Support diagnose and resolve problems.
Learn More

i	Data	Nodes	Actions	Status	Size	Time Created
	Diag-2021-05-25	127.0.0.1	Recreate Download Delete	Success	156.32 MB	2021-05-25 13:14:42

Check Your Work

Task 4: Index the baseline diag file for your record.

16. Navigate to **Settings** > **Add Data** > **Monitor**.

17. Index the diag file with **Index Once** option into the **`diag`** index:

> **NOTE:** Be sure to select the **`.tar.gz`** file, not the **`.json`** file. If your summary output looks different, do not submit. Go back and fix the step(s).

- The diag file **`<batchID>.tar.gz`** is saved in **`SPLUNK_HOME/var/run/diags`** (**`SPLUNK_HOME`** is the **`/opt/splunk`** directory.)
- Select the **Index Once** option
- Set **Source type** to **Automatic** and **App context** to **Search & Reporting**
- **Submit** when the summary of your input looks like this:

Input Type	File Monitor
Source Path	**`SPLUNK_HOME/var/run/diags/<batchID>.tar.gz`**
Continuously Monitor	No, index once
Whitelist	N/A
Blacklist	N/A
Sourcetype	Automatic
App Context	search
Host	**`splunk##`**
Index	**`diag`**

Task 5: Search the diag index for system information.

18. Determine the Splunk version and its system information by searching the **`diag`** index (All time):

```
index=diag source=*systeminfo.txt (version OR Uname OR process listing)
```

    a. Click **`Show all ### lines`** in the resulting event to expand.

b. Check the system environment values under `********** Uname **********`

19. Scroll down the expanded data and check the memory consumed by `splunkd` processes.

NOTE: Depending on the way events break, the information may span over multiple events.

Check the values under:

```
********** Process Listing (ps) **********
```

`ps aux` output lists process owner, process ID, CPU%, MEM%, total virtual memory used, non-swapped physical memory used, etc.

Check the values under:

```
********** Process Listing (tasklist) of splunkd.exe **********
```

`tasklist /V /FI IMAGENAME eq splunkd.exe` output lists name, PID, session name, session#, memory usage, status, user name, CPU time, etc.

20. To determine the system-wide resource usage, search (All time):

```
index=diag source=*resource_usage.log component=Hostwide | head 1 | table data.*
| transpose
```

NOTE: The indexing of your diag file may still be in progress. If you don't get a result, wait a couple of minutes and try the search again. If you still do not get results, you can continue to Step 19 and try again later.

21. To list the apps installed and determine their running status, search (All time):

```
index=diag source=*etc/apps/*app.conf
| rex field=source "etc/apps/(?<folderName>\w+)/"
| rex field=_raw "label.\=.(?<label>.+)"
| stats values(label) as Name, values(state) as State, values(is_visible) as
Visible by folderName
```

Module 2 Lab Exercise – Troubleshoot Indexing Issue

Description

In this exercise, you will:

- Configure Monitoring Console in standalone mode
- Install an app and enable a monitor input
- Investigate a search problem that is caused by blocked pipeline queues
- Resolve the missing data problem by analyzing the `splunkd.log`

Problem Setup Steps

Task 1: Configure the Monitoring Console to run in standalone mode.

1. From your Splunk Web, navigate to **Settings** > **Monitoring Console**.

2. To enable the app, click **Settings** > **General Setup** on the Monitoring Console menu.

3. Keep the default **Standalone** mode and click **Apply Changes** > **Refresh**.

4. Click **Indexing** > **Performance** > **Indexing Performance: Instance.**

5. Check the current data pipeline fill ratio of your Splunk instance.

   The **Status** in the **Indexing Overview** panel is ***normal*** and the fill ratio of each queue in the **Splunk Enterprise Data Pipeline** panel should be at **0%**. This indicates a healthy indexer.

Task 2: Install the tse_lab02.spl app and enable the pre-configured monitor input.

6. Click **Apps** > **Manage Apps** > **Install app from file**.

7. Browse for `tse_lab02.spl` and **Upload**.

8. Go to **Settings** > **Data inputs** > **Files & directories**.

9. Enable only the version of **trade_entries.log** input designated for your Splunk host OS.

```
/opt/log/tradelog/trade_entries.log
```

```
C:\opt\log\tradelog\trade_entries.log
```

10. To check if it is indexing, run the following search from the **Search & Reporting** app: (**All time**)

    ```
    index=* sourcetype=trade_entries
    ```

    Your search should return the ***No results found*** message.

Your Objectives

1. Investigate and identify the root cause of the search problem.

2. Fix the missing data problem so that all **`trade_entries`** events are available to search.

ROOT CAUSE:

....

....

....

REMEDY:

....

....

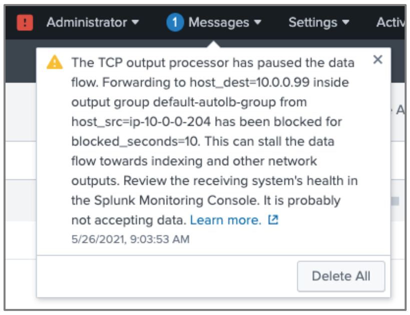
....

⛔ If you would like to run through the guided steps, continue on to **Task 3** starting on the next page.

Solution Steps

Task 3: Use Splunk tools to clarify the problems.

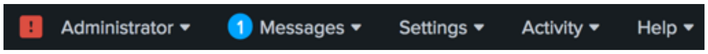
11. From Splunk Web, click **Messages** (you may have to refresh to see new messages) to view the message and then acknowledge it by deleting it.



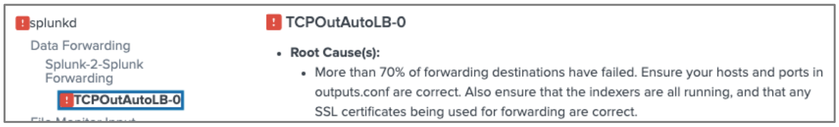
The TCP output processor has paused the data flow. Forwarding to host_dest=10.0.0.99 inside output group default-autolb-group from host_src=ip-10-0-0-204 has been blocked for blocked_seconds=10. This can stall the data flow towards indexing and other network outputs. Review the receiving system's health in the Splunk Monitoring Console. It is probably not accepting data. Learn more.

5/26/2021, 9:03:53 AM

12. Click the red **Health Status** icon.

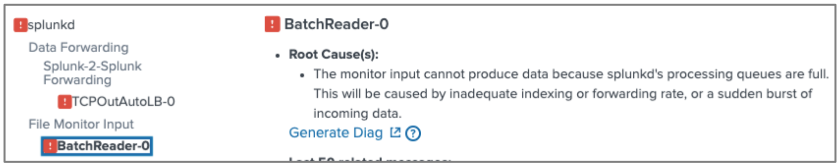


It should display a red-flagged component: `TCPOutAutoLB-0`, `BatchReader-0`, `Ingestion Latency`, and `TailReader-0`. Click the red-flagged components to get the triggered event logs. Note that it may take a few minutes for each alert to trigger.



TCPOutAutoLB-0

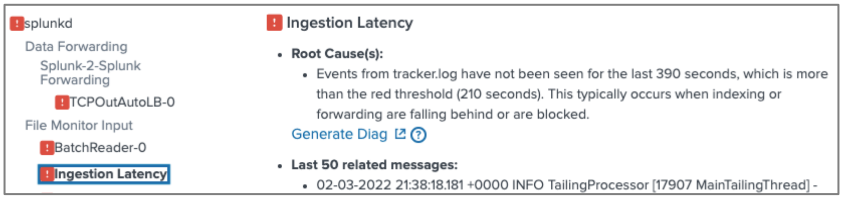
- **Root Cause(s):**
  - More than 70% of forwarding destinations have failed. Ensure your hosts and ports in outputs.conf are correct. Also ensure that the indexers are all running, and that any SSL certificates being used for forwarding are correct.



BatchReader-0

- **Root Cause(s):**
  - The monitor input cannot produce data because splunkd's processing queues are full. This will be caused by inadequate indexing or forwarding rate, or a sudden burst of incoming data.

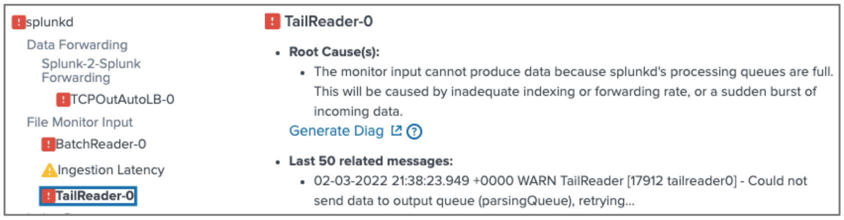
  Generate Diag



Ingestion Latency

- **Root Cause(s):**
  - Events from tracker.log have not been seen for the last 390 seconds, which is more than the red threshold (210 seconds). This typically occurs when indexing or forwarding are falling behind or are blocked.

  Generate Diag
- **Last 50 related messages:**
  - 02-03-2022 21:38:18.181 +0000 INFO TailingProcessor [17907 MainTailingThread] -



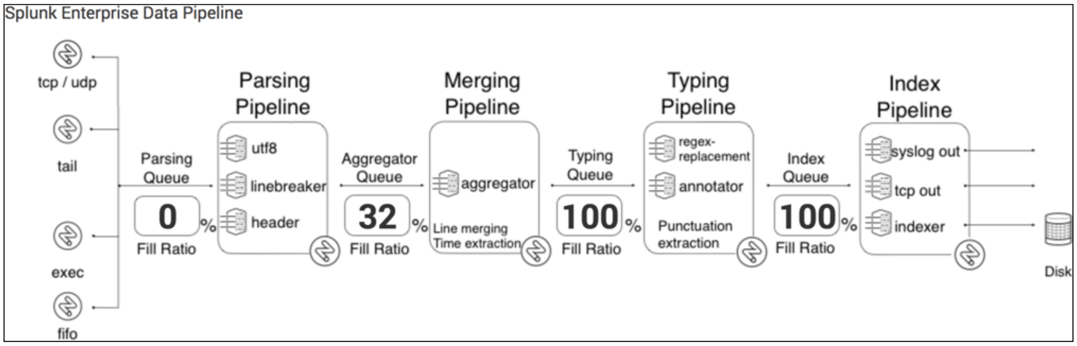
TailReader-0

- **Root Cause(s):**
  - The monitor input cannot produce data because splunkd's processing queues are full. This will be caused by inadequate indexing or forwarding rate, or a sudden burst of incoming data.

  Generate Diag
- **Last 50 related messages:**
  - 02-03-2022 21:38:23.949 +0000 WARN TailReader [17912 tailreader0] - Could not send data to output queue (parsingQueue), retrying...

13. Navigate to **Monitoring Console** > **Indexing** > **Performance** > **Indexing Performance: Instance**

    Notice the **Index Queue** is filled to 100% and spilling over to the prior queues. Based on the triggered message and the pipeline status shown in the MC dashboard, something in the forwarder settings has blocked the index pipeline.



Task 4: Clear the queue-blocking problems.

14. Go to **Settings** > **Forwarding and receiving** > **Configure forwarding**.

15. Try to `disable` and `delete` the two **Host** entries from the Web UI.

NOTE: Disabling has no effect. You will not be able to delete the listed hosts.

16. Locate the configured forwarding settings using the `btool`.

```
cd /opt/splunk/bin
./splunk btool outputs list --debug | grep -E 'apps|local'
```

```
cd 'C:\Program Files\Splunk\bin'
./splunk btool outputs list --debug | Select-String “apps|local”
```

NOTE: The output indicates that the forwarding settings were introduced from the `tse_lab02` are the result of Step 15.

17. To unblock the queues, disable the `outputs.conf` file in the `tse_lab02` app and the `system/local` directory.

    Restart Splunk.

```
cd /opt/splunk/etc/apps/tse_lab02/default
mv outputs.conf outputs.bad
rm /opt/splunk/etc/system/local/outputs.conf
cd /opt/splunk/bin
./splunk restart
```

```
cd 'C:\Program Files\Splunk\etc\apps\tse_lab02\default'
Rename-Item outputs.conf outputs.bad
del 'C:\Program Files\Splunk\etc\system\local\outputs.conf'
cd 'C:\Program Files\Splunk\bin'
./splunk restart
```

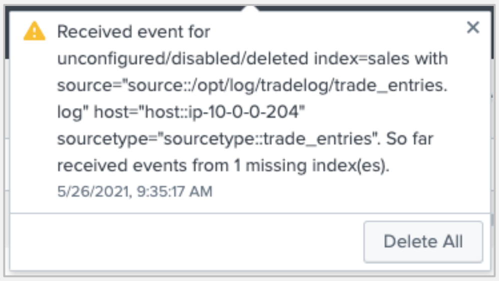
18. Confirm that the blocked queues are no longer an issue.

    a. Navigate to **Settings** > **Monitoring Console** > **Indexing** > **Indexing Performance**:**Instance Performance: Instance** and confirm that all queues are cleared.

    b. Click the health status icon and verify that all components are green.

    c. Click **Messages** and acknowledge the message(s) by clicking **Delete All**.

NOTE: You may also get another warning message for your Splunk instance:



⚠ Received event for unconfigured/disabled/deleted index=sales with source="source::/opt/log/tradelog/trade_entries.log" host="host::ip-10-0-0-204" sourcetype="sourcetype::trade_entries". So far received events from 1 missing index(es).
5/26/2021, 9:35:17 AM

19. Run the following search again from the **Search & Reporting** app: (All time)

```
index=* sourcetype=trade_entries
```

Your search should return the ***No results found*** message.

Task 5: Investigate the missing data problem and resolve.

20. Search the **metrics.log** to check if Splunk is processing the monitored file. (**Last 15 minutes**)

```
index=_internal sourcetype=splunkd metrics group=per_sourcetype_thruput
series=trade_entries | timechart sum(ev)

index=_internal sourcetype=splunkd metrics group=per_index_thruput | timechart
sum(ev) by series
```

NOTE: Checking the metrics.log indicates no apparent issues. The second search confirms that some events destined for `sales` were sent to the index queue.

21. Search the **splunkd.log** to check if Splunk has encountered any issues while indexing **trade_entries.log**. (Last 60 minutes)

    ```
    index=_internal sourcetype=splunkd log_level IN(warn, error) trade_entries |
    stats values(event_message) by component log_level _time
    ```

    It is the same message about a missing index you saw in step 19.

22. To fix the issue, navigate to **Settings > Indexes** and click **New Index.**

23. Enter `sales` as the **Index Name**, select `Troubleshooting Splunk Enterprise Indexes` for the **App** context, and **Save**.

NOTE: Splunk administrators can set the `lastChanceIndex` setting and an alert to capture the misconfigured indexing events.

24. Run the following search again from the **Search & Reporting** app: (**All time**)

    ```
    index=* sourcetype=trade_entries
    ```

    The health status should be green, and the search should produce results now.

Answers

ROOT CAUSE:

- The non-existent forwarding destinations in `outputs.conf` of `tse_lab02` app has caused a cascading queue blockage when the `trade_entries` input was enabled.

- The input setting in the app specified a missing index.

REMEDY:

- Disable the target servers until you can set the proper `outputs.conf` server list.

- Create the missing **sales** index.

Module 3 Lab Exercise – Troubleshoot Input Issue

Description

In this exercise, you will:

- Install an app and enable a monitor input
- Investigate a search problem that was caused by some input settings
- Fix the input-phase configurations so that the events are indexed properly

IMPORTANT: You MUST complete this lab exercise in its entirety before the start of Module 4 Lab Exercise. The solution steps in this lab exercise will no longer be valid once you begin the next lab exercise.

Problem Setup Steps

Task 1: Install the tse_lab03.spl app and enable the pre-configured monitor input.

1. From your Splunk Web, navigate to **Apps** > **Manage Apps** > **Install app from file**.

2. Browse for **`tse_lab03.spl`** and **Upload**.

3. Go to **Settings** > **Data inputs** > **Files & directories**.

4. Enable the directory monitor input **`/$SPLUNK_HOME/etc/apps/tse_lab03/data`**.

5. In the terminal, go to the **`data`** directory and review the content of a few log files.

```
cd /opt/splunk/etc/apps/tse_lab03/data                      #Linux
cd "\Program Files\Splunk\etc\apps\tse_lab03\data"         #Windows

ls -R

cat web/1.2.1.1.log
cat dmz/10.10.200.1.log
```

The directory contains a total of 20 .log files (7 from **`dmz`** and 13 from **`web`** sub-directories)

6. From the **Search & Reporting** app, search (**All time**):

   `index=cidr | stats count values(_time) by host, source`

7. Confirm the following artifacts from the search result:

   a. Only one source from the directory has been indexed
   b. The timestamps have been erroneously extracted from a numeric value in the events
   c. The **host** value transformation has failed

Your Objectives

1. Fix the timestamp extraction so that the current OS time is used for the `_time` value.
2. Investigate and remediate the **host** value transformation problem.

   This directory monitor input should generate 20 distinct **host** values from the 20 **source** names.

   The expected **host** values are in the form of `[dmz]:<ip_address>` *or* `[web]:<ip_address>`
3. Investigate the source count problem and address the root cause.

   This directory monitor input should generate 20 distinct **source** values.

ROOT CAUSE:

....

....

....

REMEDY:

....

....

....

If you would like to run through the guided steps, continue on to **Task 2** starting on the next page.

Solution Steps

Task 2: Review the index-time configuration settings specified in the tse_lab03 app.

8. In the command terminal, run the **`btool`** to check which input settings from **`tse_lab03`** are configured:

   Note from the output that the **host** is set to **`invalid`** and the **sourcetype** is set to **`cidrhost`**.

```
./splunk btool inputs list --app=tse_lab03
 [monitor:///opt/splunk/etc/apps/tse_lab03/data]                     #Linux
 [monitor:///C:\Program Files\Splunk\etc/apps/tse_lab03/data]     #Windows
 disabled = 0
 host = invalid
 index = cidr
 sourcetype = cidrhost
 whitelist = \.log
```

9. Run the **`btool`** to get the list of transformations configured from **`tse_lab03`**.

```
./splunk btool props list --app=tse_lab03
[cidrhost]
SHOULD_LINEMERGE = false
TRANSFORMS-subzero = subzero
[host::invalid]
TRANSFORMS-cidrhost = host_from_source

./splunk btool transforms list --app=tse_lab03
[host_from_source]
DEST_KEY = MetaData:Host
FORMAT = host::$1:$2
REGEX = data.(\w+).(\d{1,3}\.\d{1,3}\.\d{1,3}\.\d{1,3})
SOURCE_KEY = MetaData:Source
[subzero]
FORMAT = action::$1 reason::$2
REGEX = \d+ (ACCEPT|DENY)\s+(.*)$
```

Task 3: Fix the timestamp of the events to be the current time.

10. Search **All time** and get the name of the source:

    `index=cidr | stats count values(_time) by host, source`

    Note the source: ____

11. Navigate to **Settings** > **Add Data** > **Monitor**.

12. Index the source file from Step 10 with **Index Once** option.

13. In the **Set Source Type** step, search and select the **Source type**: **`cidrhost`** from the drop-down list.

    a. Expand the **Timestamp** option and click **Current time**.
    b. Confirm that the time changed in the preview panel and click **Next >**.
    c. Enter the following options:
       Name: **`cidrhost`**
       Category: **`custom`**
       App: **`system`**

       Note: Double-check that the **date** and the **Current Time** is properly set before saving.
    d. Click **Save**.
    e. Click **OK** to overwrite the existing sourcetype.

14. In the **Input Settings** step, set **App Context** to **`Troubleshooting M3 (tse_lab03)`** and **Index** to **`test`**.

    The summary of the input should be:

Input Type	File Monitor
Source Path	**`...tse_lab03/data/dmz/<source_from_step10>.log`** *or* **`...tse_lab03/data/web/<source_from_step10>.log`**
Continuously Monitor	No, index once
Sourcetype	**`cidrhost`**
App Context	**`tse_lab03`**
Host	**`splunk##`**
Index	**`test`**

15. **Submit** and **Start Searching**.

16. Proceed to the next task, if the **Time** of the events is current.

    If not, repeat Task 3.

Task 4: Investigate and fix the host name issue.

17. With the output from Step 9, get the full details of the host transformation **`host_from_source`**.

```
./splunk btool transforms list host_from_source --debug
.../etc/apps/tse_lab01/local/transforms.conf   [host_from_source]
.../etc/system/default/transforms.conf         CAN_OPTIMIZE = True
.../etc/system/default/transforms.conf         CLEAN_KEYS = True
.../etc/system/default/transforms.conf         DEFAULT_VALUE =
.../etc/system/default/transforms.conf         DEPTH_LIMIT = 1000
.../etc/apps/tse_lab01/default/transforms.conf DEST_KEY = MetaData:Host
.../etc/apps/tse_lab01/default/transforms.conf FORMAT = host::$1
.../etc/system/default/transforms.conf         KEEP_EMPTY_VALS = False
.../etc/system/default/transforms.conf         LOOKAHEAD = 4096
.../etc/system/default/transforms.conf         MATCH_LIMIT = 100000
.../etc/system/default/transforms.conf         MV_ADD = False
.../etc/apps/tse_lab01/default/transforms.conf REGEX = invalid
.../etc/apps/tse_lab01/default/transforms.conf SOURCE_KEY = MetaData:Host
.../etc/system/default/transforms.conf         WRITE_META = False
```

NOTE: The **`host_from_source`** invocation is handled by the stanza defined in **tse_lab01**. This stanza has already been taken by **tse_lab01**. To avoid any conflict with **tse_lab01**, change the **tse_lab03**'s stanza name.

18. Override the **`host_from_source`** transformation by copying the **`props.conf`** and **`transforms.conf`** files from the **`/tse_lab03/default/`** directory to the **`/tse_lab03/local`** directory. Once copied over, update the stanzas as indicated below:

    - **`/opt/splunk/etc/apps/tse_lab03/local/props.conf:`**

    ```
    [host::invalid]
    TRANSFORMS-cidrhost = cidrhost_from_source

    [cidrhost]
    #TIME_FORMAT = %s
    TRANSFORMS-subzero = subzero
    SHOULD_LINEMERGE = false
    ```

    - **`/opt/splunk/etc/apps/tse_lab03/local/transforms.conf:`**

    ```
    [cidrhost_from_source]
    SOURCE_KEY = MetaData:Source
    REGEX = data.(\w+).(\d{1,3}\.\d{1,3}\.\d{1,3}\.\d{1,3})
    FORMAT = host::$1:$2
    DEST_KEY = MetaData:Host

    [subzero]
    REGEX = \d+ (ACCEPT|DENY)\s+(.*)$
    FORMAT = action::$1 reason::$2
    ```

19. To reload the changes, restart Splunk.

20. Navigate to **Settings** > **Add Data** > **Monitor**.

21. Index the same source file from Step 10 with **Index Once** option.

    a. In the **Set Source Type** step, select the **Source type: `cidrhost`** from the drop-down list.

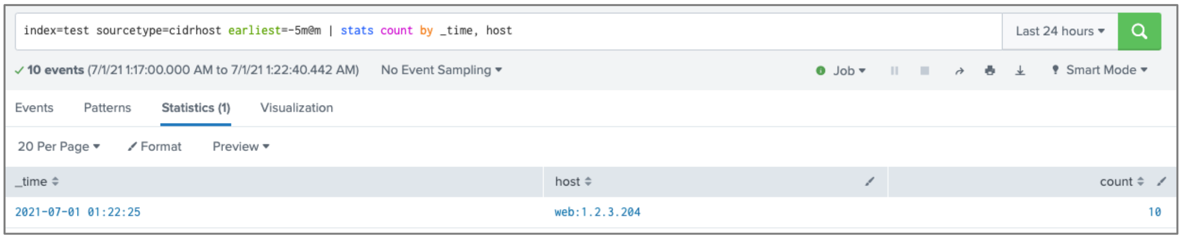
    b. In the **Input Settings** step, set **App Context** to **`Troubleshooting M3 (tse_lab03)`**, **Host field value** to **`invalid`**, and **Index** to **`test`**.

    c. Check the summary of the input and **Submit**:

Input Type	File Monitor
Source Path	**`...tse_lab03/data/dmz/<source_from_step10>.log`** *or* **`...tse_lab03/data/web/<source_from_step10>.log`**
Continuously Monitor	No, index once
Sourcetype	**`cidrhost`**
App Context	**`tse_lab03`**
Host	**`invalid`**
Index	**`test`**

22. Search: `index=test sourcetype=cidrhost earliest=-5m@m | stats count by _time, host`

23. If the **host** value of the latest event is in the form of `[dmz]:<IP>` or `[web]:<IP>`, proceed to the next task.



```
index=test sourcetype=cidrhost earliest=-5m@m | stats count by _time, host
```

✓ 10 events (7/1/21 1:17:00.000 AM to 7/1/21 1:22:40.442 AM) No Event Sampling ▾

_time	host	count
2021-07-01 01:22:25	web:1.2.3.204	10

If not, repeat Task 4.

Task 5: Investigate the source count problem and address the root cause.

24. To investigate monitor input activities, determine the installation time of the app `tse_lab03`.

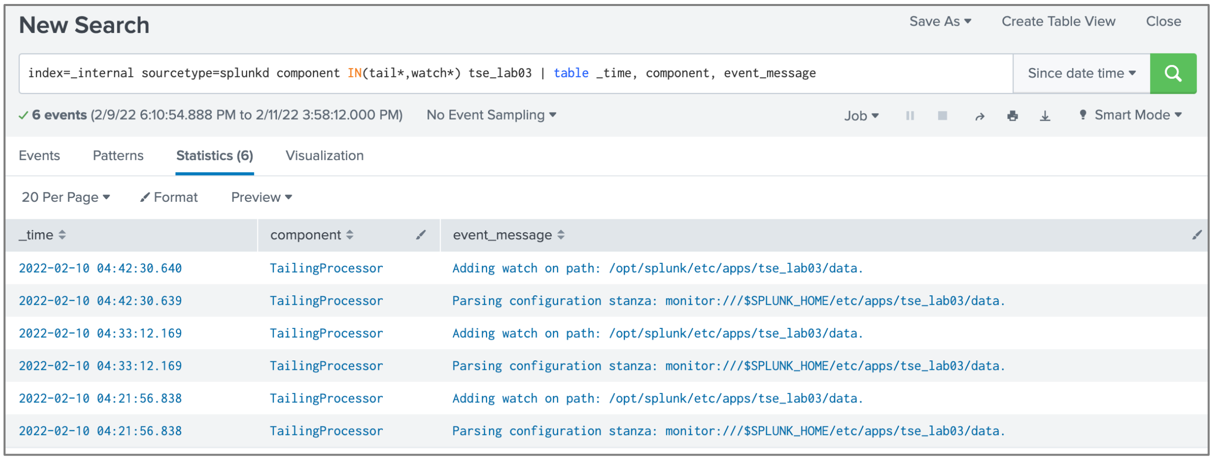
    a. Search (**All time**): `index=_internal sourcetype=splunkd ApplicationManager tse_lab03`

    b. Click the **Time** of the event and select **After this time** to define the search range for the subsequent searches.

25. To confirm that Splunk is monitoring the input from `tse_lab03`, search (**Since date time**):

    `index=_internal sourcetype=splunkd component IN(tail*,watch*) tse_lab03 | table _time, component, event_message`

> **NOTE:** You should get two types of `TailingProcessor` events:
> `Adding watch on path...` and `Parsing configuration stanza...`



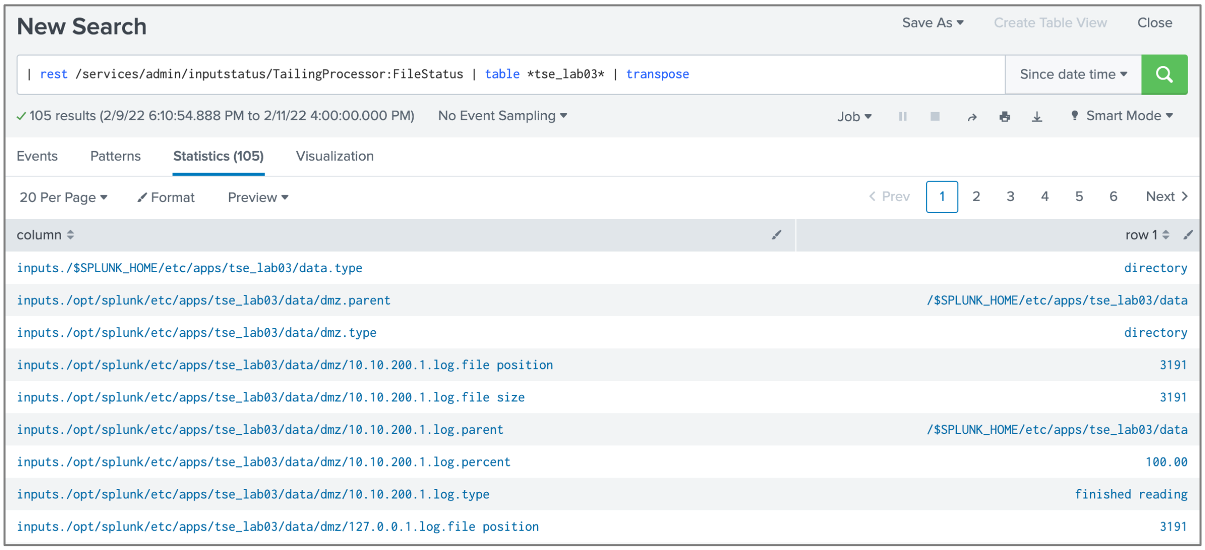
```
index=_internal sourcetype=splunkd component IN(tail*,watch*) tse_lab03 | table _time, component, event_message
```

✓ 6 events (2/9/22 6:10:54.888 PM to 2/11/22 3:58:12.000 PM) No Event Sampling ▾

_time	component	event_message
2022-02-10 04:42:30.640	TailingProcessor	Adding watch on path: /opt/splunk/etc/apps/tse_lab03/data.
2022-02-10 04:42:30.639	TailingProcessor	Parsing configuration stanza: monitor:///$SPLUNK_HOME/etc/apps/tse_lab03/data.
2022-02-10 04:33:12.169	TailingProcessor	Adding watch on path: /opt/splunk/etc/apps/tse_lab03/data.
2022-02-10 04:33:12.169	TailingProcessor	Parsing configuration stanza: monitor:///$SPLUNK_HOME/etc/apps/tse_lab03/data.
2022-02-10 04:21:56.838	TailingProcessor	Adding watch on path: /opt/splunk/etc/apps/tse_lab03/data.
2022-02-10 04:21:56.838	TailingProcessor	Parsing configuration stanza: monitor:///$SPLUNK_HOME/etc/apps/tse_lab03/data.

26. To check how many files from the directory have been processed by the monitor input, search:

```
| rest /services/admin/inputstatus/TailingProcessor:FileStatus | table
*tse_lab03* | transpose
```



```
| rest /services/admin/inputstatus/TailingProcessor:FileStatus | table *tse_lab03* | transpose
```

✓ 105 results (2/9/22 6:10:54.888 PM to 2/11/22 4:00:00.000 PM)

column	row 1
inputs./$SPLUNK_HOME/etc/apps/tse_lab03/data.type	directory
inputs./opt/splunk/etc/apps/tse_lab03/data/dmz.parent	/$SPLUNK_HOME/etc/apps/tse_lab03/data
inputs./opt/splunk/etc/apps/tse_lab03/data/dmz.type	directory
inputs./opt/splunk/etc/apps/tse_lab03/data/dmz/10.10.200.1.log.file position	3191
inputs./opt/splunk/etc/apps/tse_lab03/data/dmz/10.10.200.1.log.file size	3191
inputs./opt/splunk/etc/apps/tse_lab03/data/dmz/10.10.200.1.log.parent	/$SPLUNK_HOME/etc/apps/tse_lab03/data
inputs./opt/splunk/etc/apps/tse_lab03/data/dmz/10.10.200.1.log.percent	100.00
inputs./opt/splunk/etc/apps/tse_lab03/data/dmz/10.10.200.1.log.type	finished reading
inputs./opt/splunk/etc/apps/tse_lab03/data/dmz/127.0.0.1.log.file position	3191

Or, run the CLI command:

```
./splunk list inputstatus -input data
TailingProcessor:FileStatus :
        /$SPLUNK_HOME/etc/apps/tse_lab03/data
                type = directory

        /opt/splunk/etc/apps/tse_lab03/data/dmz
                parent = /$SPLUNK_HOME/etc/apps/tse_lab03/data
                type = directory

        /opt/splunk/etc/apps/tse_lab03/data/dmz/10.10.200.1.log
                file position = 3191
                file size = 3191
                parent = /$SPLUNK_HOME/etc/apps/tse_lab03/data
                percent = 100.00
                type = finished reading
...
```

NOTE: 20 files in the parent directory, `tse_lab03/data`, all report the same file position, size, percent, and type. This indicates some sort of CRC issue. Check the content of the log files and confirm that each file starts with the same content.

27. To force each source to have a unique CRC signature, add **`crcSalt = <SOURCE>`** to **inputs.conf**.

    `/opt/splunk/etc/apps/tse_lab03/local/inputs.conf:`

    ```
    [monitor:///$SPLUNK_HOME/etc/apps/tse_lab03/data]
    disabled = 0
    crcSalt = <SOURCE>
    ```

28. To reload the changes, restart Splunk.

29. Confirm that the directory monitor from **tse_lab03** is working as specified.

    Search: **`index=cidr earliest=-15m@m | stats count values(host) as hosts by source`**

NOTE: You should now get 20 sources returning 10 events each with the host name parsed from the source name.

Answers

ROOT CAUSE:

- The app did not provide any instruction to extract timestamps. Thus, Splunk recognized the numbers separating each event as epoch times and used them for the timestamps.
- Another app (**`tse_lab03)`** already has used the identical transformation stanza with a different definition.
- Splunk monitor process was only able to follow the first source due to the identical checksum. Because each source contains identical events, each ended up with the same checksum value.

REMEDY:

- Add **`DATETIME_CONFIG = CURRENT`** to the **`cidrhost`** sourcetype.
- Change the **`tse_lab03`**'s stanza name to avoid the invocation conflict.
- Use **`crcSalt = <SOURCE>`** to uniquely identify the checksum of each monitored file.

Module 4 Lab Exercise – Troubleshoot Deployment Client Issues

Description

In this exercise, you will:

- Investigate and resolve a distributed search problem
- Investigate and resolve a deployment client connection problem
- Search the distributed peers and piece together a picture of your lab environment

Problem Setup Steps

Task 1: Install the tse_lab04.spl app and run any search.

1. From your Splunk Web, navigate to **Apps** > **Manage Apps** > **Install app from file**.

2. Browse for **`tse_lab04.spl`** and **Upload**.

   This app configures your instance as a deployment client as well as a distributed search head.

3. Click **Restart Now** > **OK**, when prompted.

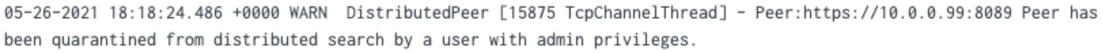
> **NOTE:** When Splunk starts, it runs a set of preliminary checks and logs the ***`My GUID is…`*** event to indicate the start of Splunk server daemon (**`splunkd`**) process. The following steps, 4 and 5, try to determine the latest splunkd start time based on this log entry.

4. Log back in when it is ready and search from the **Search & Reporting** app:
   `index=_internal sourcetype=splunkd ServerConfig "My GUID" | head 1`

5. Click the **Time** of the event, select **After this time** to set the search range, and search:
   `index=_internal host=splunk* sourcetype=splunkd log_level IN(ERROR, WARN)`

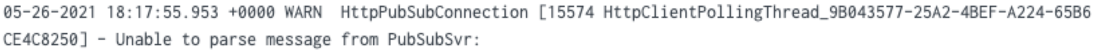
6. Observe the following component artifacts from the search result:

   - **`WARN DistributedPeer - Peer:https://10.0.0.99:8089`** indicates some sort of a search peer configuration issue.

     ```
     05-26-2021 18:18:24.486 +0000 WARN  DistributedPeer [15875 TcpChannelThread] - Peer:https://10.0.0.99:8089 Peer has
     been quarantined from distributed search by a user with admin privileges.
     ```

   - **`WARN HttpPubSubConnection - Unable to parse message from PubSubSvr:`** indicates some sort of a deployment client connection issue.

     ```
     05-26-2021 18:17:55.953 +0000 WARN  HttpPubSubConnection [15574 HttpClientPollingThread_9B043577-25A2-4BEF-A224-65B6
     CE4C8250] - Unable to parse message from PubSubSvr:
     ```

Your Objectives

NOTE: You are given the following information about the Splunk environment:

- Deployment server: **`10.0.0.200`**
- Search peers: **`10.0.0.88`** and **`10.0.0.99`**
- Remote username: **`ds_user`**
- Remote password: **`open.sesam3`**
- Remote splunkd port: *default*

1. Investigate the search peer connection issue and restore the distributed search capabilities.

2. Identify the root cause of your deployment client connection issue and resolve.

3. Determine the list of Splunk instances and their relationships within your lab environment and answer the following questions:

   a. How are the servers connected?

      ....

   b. Which servers are forwarding?

      ....

   c. Are the forwarders also deployment clients?

      ....

   d. Where is the deployment server?

      ....

   e. What apps have been deployed to the forwarders from the deployment server and when?

      ....

ROOT CAUSE:

....

....

REMEDY:

....

....

⛔ If you would like to run through the guided steps, continue on to **Task 2** starting on the next page.

Solution Steps

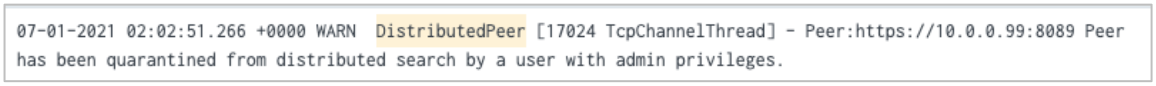
Task 2: Investigate the search peer connection issue and restore the distributed search function.

7. Search the **`splunkd.log`** for the latest search peer events (**Last 60 minutes**):

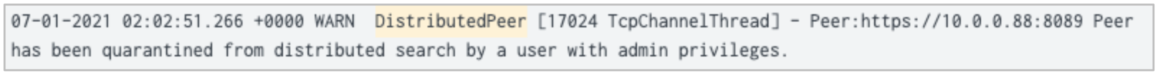
```
index=_internal host=splunk## sourcetype=splunkd DistributedPeer | head 5
```

NOTE: From this point on, you are running distributed searches. Replace **`splunk##`** with your Splunk host name.

8. Observer the following messages:



```
07-01-2021 02:02:51.266 +0000 WARN  DistributedPeer [17024 TcpChannelThread] - Peer:https://10.0.0.99:8089 Peer
has been quarantined from distributed search by a user with admin privileges.
```



```
07-01-2021 02:02:51.266 +0000 WARN  DistributedPeer [17024 TcpChannelThread] - Peer:https://10.0.0.88:8089 Peer
has been quarantined from distributed search by a user with admin privileges.
```

9. Navigate to **Settings** > **Distributed search** > **Search Peers** and check the status of the peers.

10. **Enable** and **Unquarantine** the peers.

    The **Health status** is **Sick** and shows the error messages.

11. Click the **Peer URI** and provide the distributed search authentication information:
    - Remote username: **`ds_user`**
    - Remote password: **`open.sesam3`**

12. If the **Health status** of both peers is ***Healthy***, search:

```
index=_internal sourcetype=splunkd | top splunk_server
```

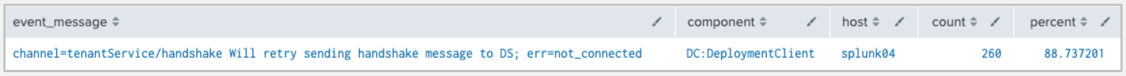
    If the results contain **`ip-10-0-0-88`** and **`ip-10-0-0-99`**, then the distributed search is working.

Task 3: Investigate the deployment client connection issue and resolve.

13. Search the **`splunkd.log`** for the latest deployment client events (**Last 60 minutes**):

```
index=_internal host=splunk## sourcetype=splunkd component=DC* | top
event_message component host
```

NOTE: Confirm that the messages are from your Splunk instance. One message indicates a deployment server connection issue and you should verify the configuration settings:



event_message	component	host	count	percent
channel=tenantService/handshake Will retry sending handshake message to DS; err=not_connected	DC:DeploymentClient	splunk04	260	88.737201

14. In the terminal, run **`btool`** to check the deployment client settings.

```
./splunk btool deploymentclient list --debug
apps/tse_lab04/default/deploymentclient.conf [target-
broker:deploymentServer]
etc/apps/tse_lab04/default/deploymentclient.conf  targetUri =
10.0.0.200:9997
```

NOTE: A wrong management port is configured for the **`targetUri`** attribute. The default is **`8089`**.

15. Override the **`targetUri`** value with the corrected port.
    a. Create the **`local`** directory in **`SPLUNK_HOME/etc/apps/tse_lab04`**.
    b. Copy **`default/deploymentclient.conf`** to **`local/deploymentclient.conf`**.
    c. Edit **`local/deploymentclient.conf:`**

```
[target-broker:deploymentServer]
targetUri = 10.0.0.200:8089
```

16. Restart Splunk.

    Take a short break. You will be logged out and asked to re-login shortly. Log back in and continue.

17. Search the **`splunkd.log`** again for the latest deployment client events:

```
index=_internal host=splunk## sourcetype=splunkd component=DC* "Handshake done"
earliest=-5m@m
```

Task 4: Determine the list of Splunk instances and their relationships in the environment.

18. Search **`_internal`** to answer the following questions (Last 7 days):

    a. How are the servers connected?

```
index=_internal sourcetype=splunkd connection* | stats count by sourceIp, host,
destPort
```

    b. Which servers are forwarding?

```
index=_internal sourcetype=splunkd tcpout_connections | stats count by host,
destIp, destPort | rename host as forwarder, destIp as indexer, destPort as
listening_port

index=_internal sourcetype=splunkd metrics group=tcpin_connections
connectionType=cooked* | stats sum(kb) by hostname, fwdType, lastIndexer
```

    c. Are the forwarders also deployment clients?

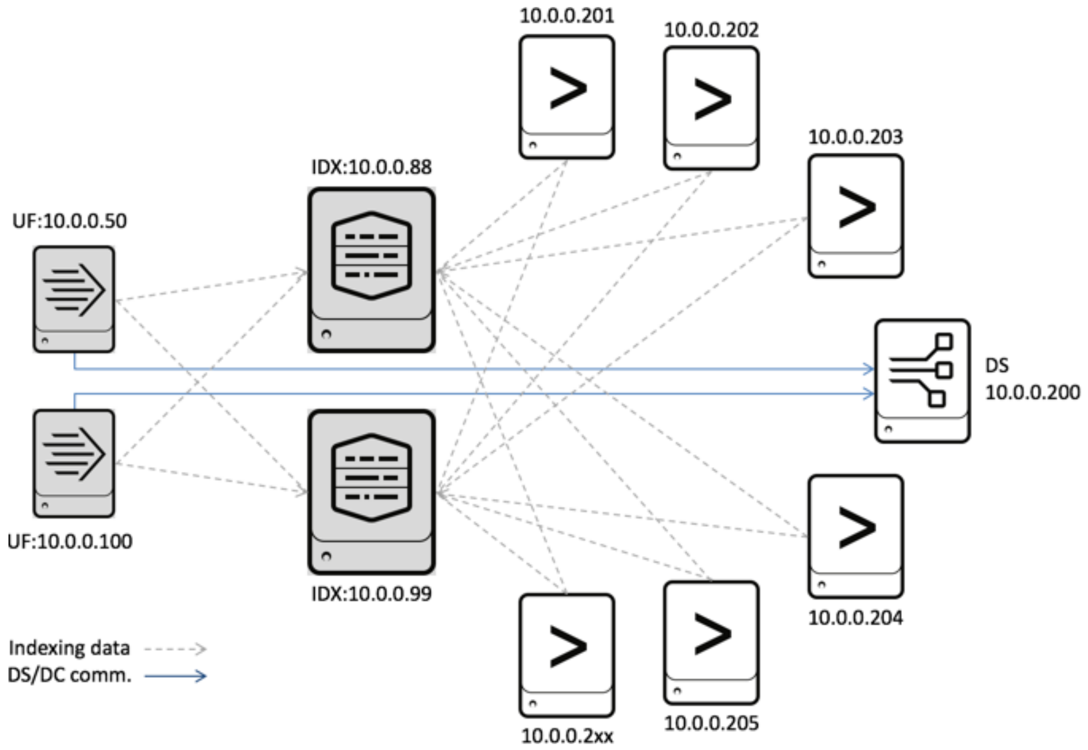
```
index=_internal sourcetype=splunkd component=DC* Handshake | stats count by host
```

d. Where is the deployment server?

```
index=_internal sourcetype=splunkd component=DeployedApplication url=* | table
host url
```

e. What apps have been deployed to the forwarders from the deployment server and when?

```
index=_internal sourcetype=splunkd component=DeployedApplication installing |
stats count latest(_time) AS latest_time by host app | convert ctime(latest_time)
```



Answers

ROOT CAUSE:

- The app was packaged with the distributed search functionality disabled and quarantined.
- The deployment client failed to connect because **`deploymentclient.conf`** in the app is set with the wrong Splunk management port.

REMEDY:

- Enable and un-quarantine the peers and provide the remote credential to distribute the search head's authentication key to the search peers.
- Change the port used in the **`targetUri`** attribute to **`8089`**.

Module 5 Lab Exercise – Troubleshoot LDAP Issues

Description

The problem in this lab exercise is in two parts:

Initially, the user **`dhale`** from the LDAP group **`splunkBizDev`** has reported that he cannot log into Splunk Web with his LDAP credentials. He is a member of the Splunk **`sales`** role and should be allowed to access Splunk Web. It turns out while all LDAP users do not have issues with other services, none of them are able to log into Splunk Web.

After addressing the initial issue, he has followed up with a second part of the issue. He is complaining that his searches frequently fail to run and return the message ***`Waiting for queued job to start`***.

In this exercise, you will:

- Investigate and resolve the LDAP user authentication issue
- Investigate and resolve **`dhale`**'s search problem

NOTE: In this lab environment, the account that binds the Splunk authentication service to the LDAP server is **`adsuser@buttercupgames.local`** and its password is **`open.sesam3`**.

The passwords for all LDAP users are set to **`open.sesam3`**.

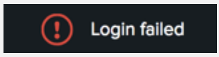
Problem Setup Steps

Task 1: Install the tse_lab05 app.

1. Navigate to **Apps** > **Manage Apps** page and install the **`tse_lab05.spl`** package.

2. Open a new private browser window and try to log into Splunk Web as the LDAP user **`dhale`**.

   Use **`open.sesam3`** for **dhale**'s password.

NOTE: You should get  Login failed. Do not close the private browser window.

3. After fixing the authentication issue, log in as **`dhale`** and run this search three times:

   ```
   index=* earliest=0
   ```

Your Objectives

1. Confirm the user authentication issue and restore the Splunk LDAP authentication service.
2. Investigate the user **`dhale`**'s search problem and address his issue.

ROOT CAUSE:

....

....

REMEDY:

....

....

⛔ If you would like to run through the guided steps, continue on to **Task 2** starting on the next page.

Solution Steps

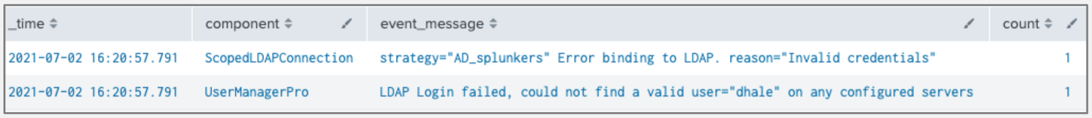
Task 2: Confirm the LDAP issue and restore the Splunk LDAP authentication service.

4. In the **`admin`**'s browser session, search internal logs for more clues.

```
index=_internal sourcetype=splunkd host=splunk## component IN(UserMan*,
ScopedLDAP*) earliest=-5m@m | stats count by _time component event_message
```

NOTE: You are running distributed searches. Replace **`splunk##`** with your Splunk host name.

The log messages should indicate some sort of LDAP service binding issues.



_time	component	event_message	count
2021-07-02 16:20:57.791	ScopedLDAPConnection	strategy="AD_splunkers" Error binding to LDAP. reason="Invalid credentials"	1
2021-07-02 16:20:57.791	UserManagerPro	LDAP Login failed, could not find a valid user="dhale" on any configured servers	1

5. Navigate to **Settings** > **Authentication methods** > **LDAP Settings** > **LDAP strategies** > **AD_splunkers**.

6. Re-enter and confirm **`open.sesam3`** for the **Bind DN Password**, and **Save**.

7. Run the following real-time search with **30 second window**:

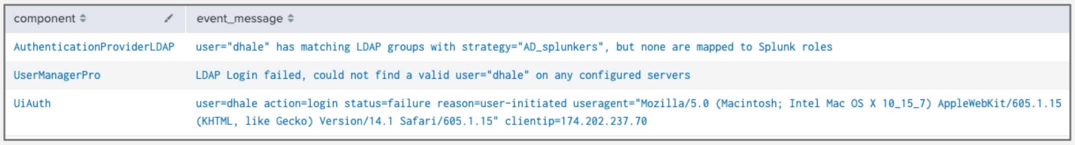
```
index=_internal sourcetype=splunkd host=splunk## dhale | table component
event_message
```

Stop the search the moment you get some results from the next action (Step 8).

8. Switch to **`dhale`**'s browser session and try to log into Splunk Web again.

NOTE: You should still not be able to log in with **`dhale`**'s account. However, the **`admin`**'s real-time search reveals additional information:

`user="dhale" has matching LDAP groups with strategy="AD_splunkers", but none are mapped to Splunk roles`



component	event_message
AuthenticationProviderLDAP	user="dhale" has matching LDAP groups with strategy="AD_splunkers", but none are mapped to Splunk roles
UserManagerPro	LDAP Login failed, could not find a valid user="dhale" on any configured servers
UiAuth	user=dhale action=login status=failure reason=user-initiated useragent="Mozilla/5.0 (Macintosh; Intel Mac OS X 10_15_7) AppleWebKit/605.1.15 (KHTML, like Gecko) Version/14.1 Safari/605.1.15" clientip=174.202.237.70

9. To fix the role mapping, stop the RT search if you haven't already.

10. Go to **Settings** > **Authentication methods** > **LDAP Settings** > **Map groups**.

11. Click **`splunkBizDev`**, add just the **`sales`** role to the **Selected Roles** column, and **Save**.

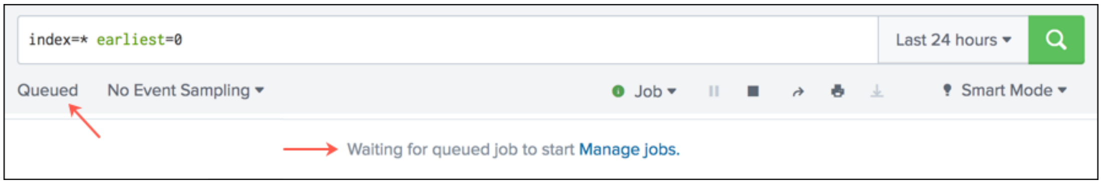
12. Switch to the **dhale**'s browser session and log into Splunk Web again.

    dhale is now in Splunk Web.

Task 3: Investigate the user dhale's search problem and address his issue.

13. In **dhale**'s browser session, search: **`index=* earliest=0`**

Repeat the same search until his search job returns the result ***`Waiting for queued job to start`***. It takes about 3~5 searches to see the following results:



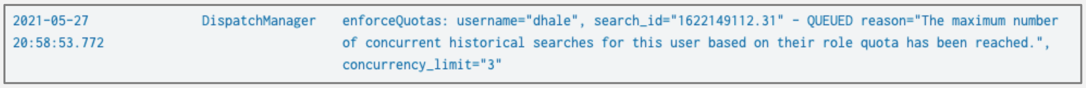
```
index=* earliest=0
```

Waiting for queued job to start Manage jobs.

14. Switch to the **admin**'s browser session and search:

```
index=_internal sourcetype=splunkd host=splunk## dhale NOT metrics earliest=-5m@m
| table _time component event_message
```

NOTE: To restrict excessive storage usage, a role quota enforcement is in effect according to the log message:



```
2021-05-27      DispatchManager   enforceQuotas: username="dhale", search_id="1622149112.31" - QUEUED reason="The maximum number
20:58:53.772                      of concurrent historical searches for this user based on their role quota has been reached.",
                                  concurrency_limit="3"
```

15. Go to **Activity** > **Jobs**.

16. Filter the job list with **App: All** and **Owner: Dwight Hale (dhale)**.

17. Stop the **Queued** jobs and delete the jobs that are **Done**, **Failed**, or **Finalized**.

NOTE: The users in this role should minimize the search range of stream searches. If they must use them over a large search window, then it is recommended to use the **Event Sampling** option. Splunk admins can also increase the role's **Limit total jobs disk quota**.

Answers

ROOT CAUSE:

- The LDAP configuration stanza provided in the **`tse_lab05`** app was incomplete and missing the LDAP group mapping to Splunk role.
- The **`sales`** role has the **Limit total jobs disk quota** set at **`40`** MB and the users of this role have exceeded this limit.

REMEDY:

- Provide the **Bind DN Password** and map the **`splunkBizDev`** LDAP group to the **`sales`** role.
- Either increase the size of the **Limit total jobs disk quota** and/or reduce the job artifact TTL.
- Tell users to limit unrestricted stream searches and advise them to use the event sampling option.

splunk>

Module 6 Lab Exercise – Troubleshoot Search Job Problems

Description

Splunk users have reported that their Splunk Web performance is poor, and their search jobs are inconsistent. In this exercise, you will:

- Identify the root cause of the poor performance
- Investigate and restore the search functionality

Problem Setup Steps

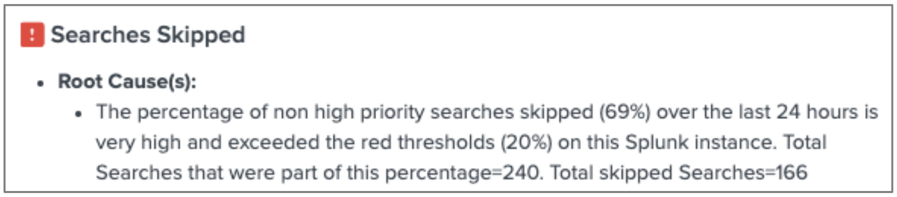
> **NOTE:** You may have noticed some performance issues. Or, Splunk Web UI might be unresponsive. This is intentional.

Task 1: Enable Monitoring Console in Standalone mode.

1. Go to **Settings** > **Monitoring Console** > **Settings** > **General Setup**.
2. Click **Edit** > **Edit Server Roles** and save only the **Search Head** role.

   Depending on the size of your display, you may need to scroll to the right to access the **Edit** link.
3. Click **Apply Changes** > **Go to Overview**.

   Notice the Health indicator message regarding skipped searches.



❗ Searches Skipped

- **Root Cause(s):**
  - The percentage of non high priority searches skipped (69%) over the last 24 hours is very high and exceeded the red thresholds (20%) on this Splunk instance. Total Searches that were part of this percentage=240. Total skipped Searches=166

Your Objectives

1. Describe the unusual symptoms you have noticed from Monitoring Console (MC).
2. Identify the root cause of the symptoms.
3. Recommend workarounds and remediate the symptoms.

ROOT CAUSE:

REMEDY:

⛔ If you would like to run through the guided steps, continue on to **Task 2** starting on the next page.

Solution Steps

Task 2: Stabilize Splunk.

4. Go to **Activity** > **Jobs**.

5. Filter the job list with **App: All** and **Owner: All**.

   Note the number of `Running` and `Done` jobs and which app they are from. You may use the search to get the similar list:

   ```
   | rest /services/search/jobs splunk_server=local | table label author eai:acl.app
   dispatchState diskUsage
   ```

6. Add **Status: Running** to the filter, select all running jobs, and delete them.

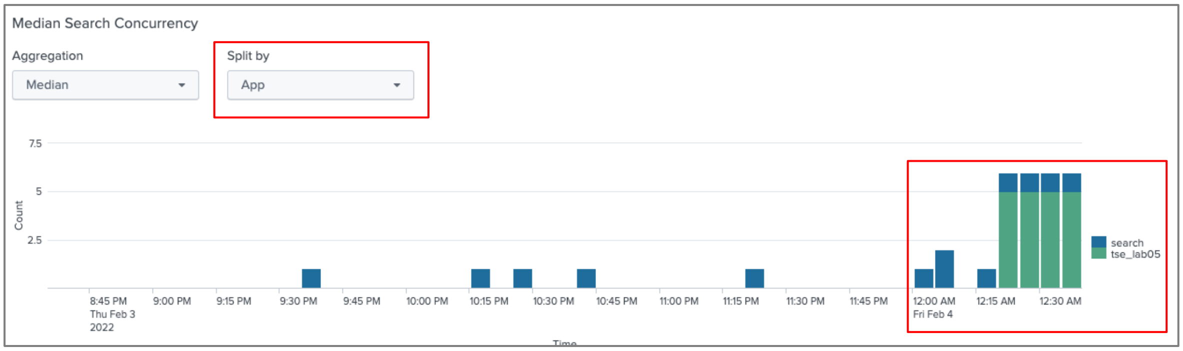
   As you delete the running jobs, new jobs may show up again.

Task 3: Use Monitoring Console to identify the root cause.

7. Go to the **Search** > **Search Activity: Instance** dashboard and note the search concurrency limit in the **Search Concurrency (Running/Limit)** panel.

8. Scroll down to the **Median Search Concurrency** panel and identify the type of search that is most common.

9. Change the **Split by** filter to `App` and identify the app that is monopolizing the search concurrency.



Median Search Concurrency

Aggregation: Median

Split by: App

10. Go to the **Search** > **Scheduler Activity: Instance** dashboard and check the **Count of Scheduler Executions** panel.

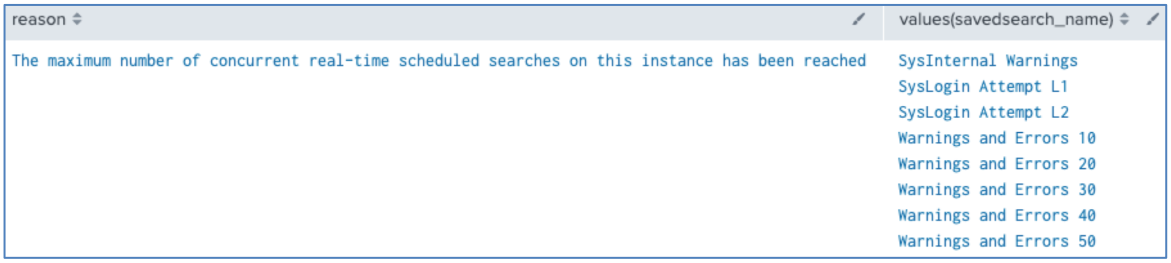
    A large portion of the jobs were skipped.

11. Scroll down to the **Count of Skipped Scheduled Reports** panel and check the **reason**.

12. To get the list of reports contributing to the scheduler problem, click the **Open in Search** icon in the bottom right corner of the **Count of Skipped Reports by Name and Reason** panel.

13. Click in the search box, press **Control-Shift-E** to expand the search, click **Open as new search**, and then edit:

```
host=splunk## index=_internal sourcetype=scheduler status="skipped" | stats
values(savedsearch_name) count by reason
```



reason	values(savedsearch_name)
The maximum number of concurrent real-time scheduled searches on this instance has been reached	SysInternal Warnings SysLogin Attempt L1 SysLogin Attempt L2 Warnings and Errors 10 Warnings and Errors 20 Warnings and Errors 30 Warnings and Errors 40 Warnings and Errors 50

14. Go to **Settings** > **Searches, reports, and alerts** and filter with **App: Troubleshooting M5 (tse_lab05)**.

15. From the **Owner** drop-down menu, select **All**.

16. Disable all real-time alerts found from the above search (**Edit** > **Disable**).
    - `SysInternal Warnings`
    - `SysLogin Attempt L1`
    - `SysLogin Attempt L2`
    - `Warnings and Errors 10`
    - `Warnings and Errors 20`
    - `Warnings and Errors 30`
    - `Warnings and Errors 40`
    - `Warnings and Errors 50`

Task 4: Confirm that the Search function is operational.

17. Go to the **Search > Scheduler Activity: Instance** dashboard.

18. Check the **Count of Scheduler Executions Over Time** panel again.

NOTE: The interval of the skipped schedules follow a pattern. You may need to change the time range to magnify the pattern (for example, select Last 15 minutes).

19. Scroll down to the **Runtime Statistics** panel.

    Note the **Average Runtime (sec)** of the report **delay Stats1** and the **Average Execution Latency (sec)** of the report **transaction search**.

20. Navigate to **Settings** > **Searches, reports, and alerts** and filter with **App: Troubleshooting M5 (tse_lab05)**.

    Note all reports are scheduled to run at the same time.

21. Disable the report **delay Stats1**.

22. Click **Edit** > **Edit Schedule** of the **`transaction search`** report and set the **Schedule Priority** to **`Highest`**.

23. Append the following to **`SPLUNK_HOME/etc/apps/tse_lab05/local/savedsearches.conf`**:

```
...

[default]
schedule_window = auto
```

24. Restart Splunk.

25. Take a 5 minute break.

26. Go to the **Scheduler Activity: Instance** dashboard.

27. Verify that **Skip Ratio** is trending down, more reports are completed, and the execution latency is reduced.

    a. Change the **Time Range** to **Last 15 minutes**.

    b. Check the **Count of Skipped Reports Over Time** and **Runtime Statistics** panels for confirmation.

Answers

ROOT CAUSE:

- Your Splunk instance can only support 12 scheduled searches concurrently -- 5 historcal, 5 real-time, and 2 summarization -- but the searches in the **tse_lab05** app trigger more alerts than your Splunk instance can handle.

  This has cascading consequences and it will eventually consume a large portion of CPU, memory, and disk space.

- A bunch of searches are scheduled with the same cron interval and they are overwhelming the scheduling resources every 5 minutes.

- Because of the number of concurrent searches and their TTLs, the job artifacts are accumulating faster than the reaping schedule, thus exhausting the available disk space on the server.

REMEDY:

- Upgrade the Splunk instance to new hardware that has more CPU cores, RAM, and disk space.

- Evaluate the need for real-time alerts and restrict usage of real-time searches.

  If an alert is absolutely necessary, then optimize it to run quickly with a suppression.

- For scheduled searches, distribute the schedules to run over different times. If they must run at the same time, use the schedule priority and schedule window. This will distribute the scheduling resources around the time.

Module 7 Lab Exercise – Troubleshoot Search Issues

Description

NOTE: If your Splunk instance is still extremely slow to respond, do this before you proceed.

1. To disable searches in **tse_lab05**, edit the **`default`** stanza in **`SPLUNK_HOME/etc/apps/tse_lab05/local/savedsearches.conf`**:

   ```
   [default]
   schedule_window = auto
   disabled = 1
   ```

2. Restart Splunk.

Splunk users have reported that suddenly their host events are missing from their searches. Another group has observed mysterious search behavior.

In this exercise, you will:

- Identify the root cause of the missing host issue and resolve
- Investigate and resolve the inconsistent search result problem

Problem Setup Steps

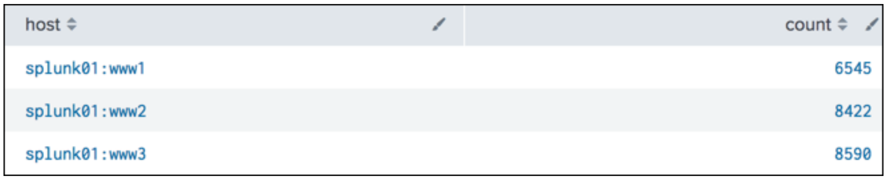
Task 1: Confirm the missing host search problem.

1. Navigate to **Apps** > **Manage Apps** page and install the **`tse_lab07.spl`** package.

2. From the search app, search (**Last 7 days**):

   ```
   index=itops sourcetype=ghostwww host=splunk##* | stats count by host
   ```

   You should get the event count of three hosts:



host	count
splunk01:www1	6545
splunk01:www2	8422
splunk01:www3	8590

3. Run another search (**Last 60 minutes**):

   ```
   index=itops sourcetype=ghostwww host=splunk##* | timechart count by host
   ```

   This search returns data from **`www2`** and **`www3`** but missing from **`www1`**.

Task 2: Confirm the mysterious result search problem.

4. Search (All time): `index=itops sourcetype=wocheese host=splunk## | top 20 country`

   The result returns **12** abbreviated country IDs including **BA** and **ZB**.

5. Search (All time): `index=itops sourcetype=wocheese host=splunk## country=BA`

   The result returns **121** events from **country=BA**.

6. Search (All time): `index=itops sourcetype=wocheese host=splunk## country=ZB`

   You get *No results found*.

Your Objectives

1. Determine when the data from **www1** went missing and why.
2. Resolve the issue so the new data from **www1** is indexed.
3. Determine the root cause of the task 2 search behaviors and provide more accurate search workaounds.

ROOT CAUSE:

....

....

REMEDY:

....

....

⛔ If you would like to run through the guided steps, continue on to **Task 3** starting on the next page.

Solution Steps

Task 3: Determine when the data from www1 went missing and why.

7. To determine the location of the input, run btool:

```
./splunk btool inputs list --debug | grep "access.log"
```

```
./splunk btool inputs list --debug | Select-String "access.log"
```

NOTE: The btool output indicates that the input setting was introduced from the **hf-student** app.

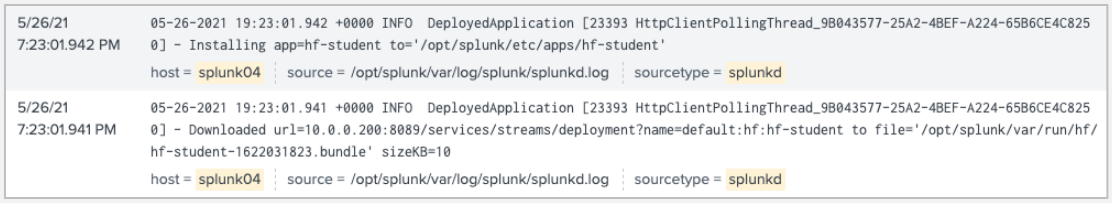
8. Investigate where this app came from (**Last 7 days**):

```
index=_internal sourcetype=splunkd host=splunk## component=ApplicationManager
hf-student
```

9. Click the **Time** of the event > **Before this time** and search:

```
index=_internal sourcetype=splunkd host=splunk## NOT metrics | head 10
```

NOTE: The messages indicate that it was pulled from the deployment server:



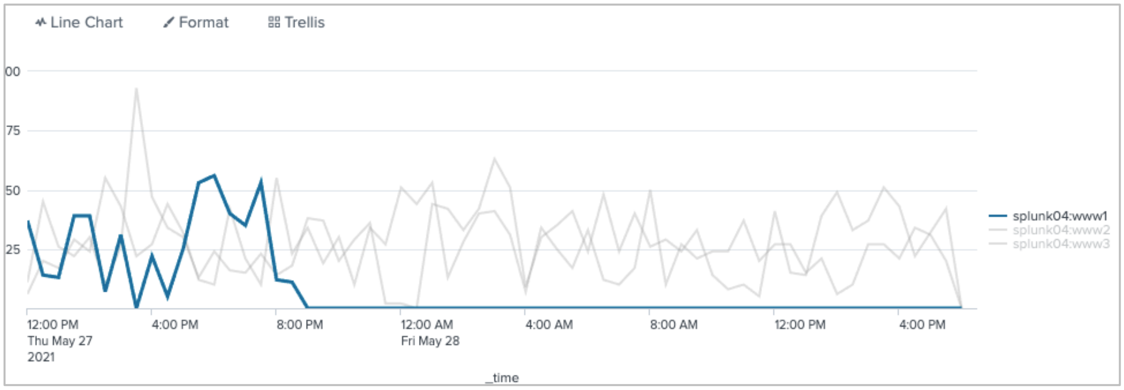
```
5/26/21           05-26-2021 19:23:01.942 +0000 INFO  DeployedApplication [23393 HttpClientPollingThread_9B043577-25A2-4BEF-A224-65B6CEC4C825
7:23:01.942 PM    0] - Installing app=hf-student to='/opt/splunk/etc/apps/hf-student'
                  host = splunk04   source = /opt/splunk/var/log/splunk/splunkd.log   sourcetype = splunkd

5/26/21           05-26-2021 19:23:01.941 +0000 INFO  DeployedApplication [23393 HttpClientPollingThread_9B043577-25A2-4BEF-A224-65B6CEC4C825
7:23:01.941 PM    0] - Downloaded url=10.0.0.200:8089/services/streams/deployment?name=default:hf:hf-student to file='/opt/splunk/var/run/hf/
                  hf-student-1622031823.bundle' sizeKB=10
                  host = splunk04   source = /opt/splunk/var/log/splunk/splunkd.log   sourcetype = splunkd
```

10. To confirm that the monitor input was working when it was first installed, change the time range scope to **since date time** and search:

```
index=itops sourcetype=ghostwww host=splunk##:* | timechart count by host
```

In the **Visualization** tab, select a **Line Chart** and mouse over the **www1** legend to confirm.



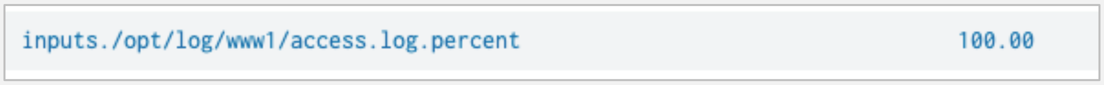
11. Get the timestamp of the last event in **/opt/log/www1/access.log**.

🐧	`tail -1 /opt/log/www1/access.log`
⊞	`Get-Content \opt\log\www1\access.log -Tail 1`

12. Check the input status of the file.

```
| rest /services/admin/inputstatus/TailingProcessor:FileStatus
splunk_server=local | transpose | search column=*www1*
```

NOTE: As far as Splunk file monitoring is concerned, all content from the file has been read. (**inputs...access.log.percent 100.00**)



```
inputs./opt/log/www1/access.log.percent                                    100.00
```

13. To get the last time Splunk processed the **www1** events for indexing, search **metrics.log** (**All time**):

```
index=_internal sourcetype=splunkd host=splunk## component=Metrics
group=per_host_thruput series=*www1 | timechart sum(ev)
```

14. Click the last point the event count was registered and click **Narrow to this time range** option to drill down.

    Repeat until you drill down to the last event count. This is the same timestamp if you search:

```
index=_internal sourcetype=splunkd host=splunk## component=ApplicationManager
```

NOTE: Compare the timestamp delta between the one from Step 11 and Step 14.

15. To check if there were any changes to Splunk, click the **Time** of the latest event, select **Before this time**, and search:

```
index=_internal sourcetype=splunkd host=splunk## component=ApplicationManager
```

    Change the search range to **Since date time** and run the same search again.

NOTE: The **tse_lab05** app was installed after **hf-student**.

16. Use **btool** to check if **tse_lab05** contains any transformations that may have affected the input:

```
./splunk btool props list --app=tse_lab05
[ghostwww]
TRANSFORMS-adios = adios

./splunk btool transforms list --app=tse_lab05
[adios]
DEST_KEY = queue
FORMAT = nullQueue
REGEX = (?msi).*www1.*
SOURCE_KEY = MetaData:Source
```

NOTE: The **tse_lab05** app is dropping the **www1** data based on the sourcetype.

Task 4: Disable the transformation of the ghostwww sourcetype.

17. Create a local version of **props.conf** to disable the invocation of the **adios** transformation:

    Add the following to **SPLUNK_HOME/etc/apps/tse_lab05/local/props.conf**:

    ```
    [ghostwww]
    TRANSFORMS-adios =
    ```

18. Reload the change by disabling and then enabling the **tse_lab05** app.

19. To confirm the fix, search (**Last 15 minutes**):

    index=itops sourcetype=ghostwww host=splunk##:www1 | timechart count by host

NOTE: In this environment, the **www1** event generation occurs at a random interval. If you don't get any events, wait a few minutes and try again. If you still don't get any results, just move on to the next task and search again at the end of the lab exercise.

Optional Task 5: Investigate the mysterious `wocheese` search behavior.

20. Use **`btool`** to get all configurations relating to the sourcetype **`wocheese`**:

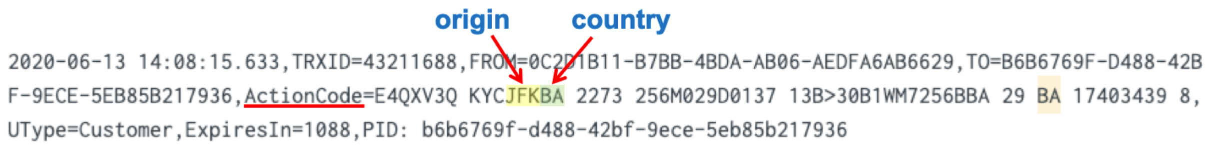
```
./splunk btool inputs list --debug | grep wocheese
.../apps/tse_lab07/default/inputs.conf                 sourcetype = wocheese

./splunk btool props list --app=tse_lab07
[wocheese]
EXTRACT-action_data =
  ActionCode=.+KYC(?P<origin>\w{3})(?P<country>\D{2})\s(?P<chizcode>\d{1,4})
EXTRACT-game_time =
  ^(?P<game_date>\d{4}-\d{2}-\d{2})\s(?P<game_time>\d{2}:\d{2}:\d{2}\.\d{3})
```

Replace the **`grep`** command and use the **`Select-String`** command instead.

21. Search (**All time**): **`index=itops sourcetype=wocheese host=splunk## country=BA`**

NOTE: **`country`** is a search-time field derived from the string following **`ActionCode`**.



```
2020-06-13 14:08:15.633,TRXID=43211688,FROM=0C2D1B11-B7BB-4BDA-AB06-AEDFA6AB6629,TO=B6B6769F-D488-42B
F-9ECE-5EB85B217936,ActionCode=E4QXV3Q KYCJFKBA 2273 256M029D0137 13B>30B1WM7256BBA 29 BA 17403439 8,
UType=Customer,ExpiresIn=1088,PID: b6b6769f-d488-42bf-9ece-5eb85b217936
```

22. Click **Job** > **Inspect Job** > **search.log**.

23. In the **`search.log`** browser window, press **Ctrl/Cmd+F** to open the browser Find tool.

    Search the log entry containing "**`base lispy`**".

NOTE: You get *`[ AND ba host::splunk## index::itops sourcetype::wocheese ]`*.

It represents the base query that Splunk uses to check the index buckets with. Note that it looks for the keyword **ba** in addition to the usual default metadata.

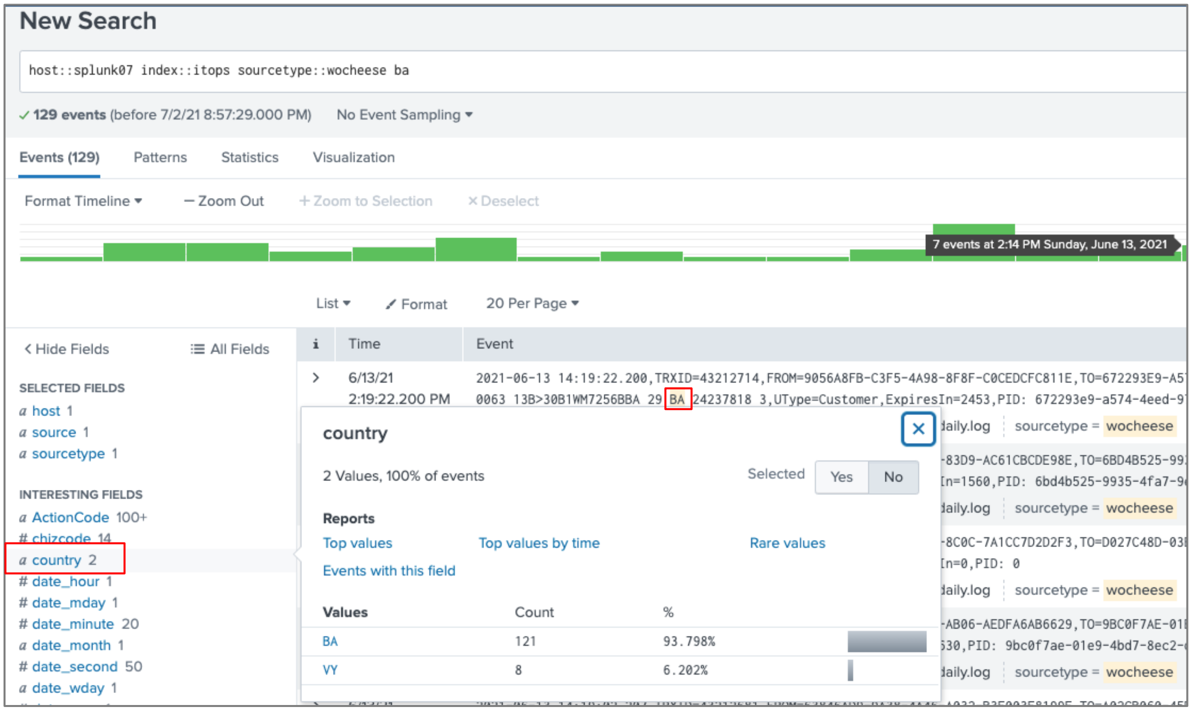
24. Search (**All time**): **`index=itops sourcetype=wocheese host=splunk## country=ZB`**

    Your search should return the ***No results found*** message.

25. Run the base lispy query in the search app.

    `host::splunk## index::itops sourcetype::wocheese ba`

    Note the string where the matched keyword is highlighted. The search, **`host::splunk## index::itops sourcetype::wocheese ba`**, is actually matching the wrong values.



```
host::splunk07 index::itops sourcetype::wocheese ba
```

✓ **129 events** (before 7/2/21 8:57:29.000 PM) No Event Sampling ▾

SELECTED FIELDS
- a host 1
- a source 1
- a sourcetype 1

INTERESTING FIELDS
- a ActionCode 100+
- # chizcode 14
- a country 2
- # date_hour 1
- # date_mday 1
- # date_minute 20
- a date_month 1
- # date_second 50
- a date_wday 1

i	Time	Event
>	6/13/21 2:19:22.200 PM	2021-06-13 14:19:22.200,TRXID=43212714,FROM=9056A8FB-C3F5-4A98-8F8F-C0CEDCFC811E,TO=672293E9-A5 0063 13B>30B1WM7256BBA 29 BA 24237818 3,UType=Customer,ExpiresIn=2453,PID: 672293e9-a574-4eed-9

```
daily.log  sourcetype = wocheese
-83D9-AC61CBCDE98E,TO=6BD4B525-99
In=1560,PID: 6bd4b525-9935-4fa7-9
daily.log  sourcetype = wocheese
-8C0C-7A1CC7D2D2F3,TO=D027C48D-03
In=0,PID: 0
daily.log  sourcetype = wocheese
-AB06-AEDFA6AB6629,TO=9BC0F7AE-01
530,PID: 9bc0f7ae-01e9-4bd7-8ec2-
daily.log  sourcetype = wocheese
```

country

2 Values, 100% of events

Selected: Yes [✓ No]

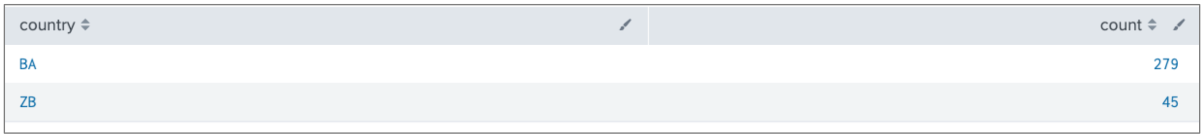
Reports

Values	Count	%
BA	121	93.798%
VY	8	6.202%

The proper searches should use **`country=*BA`** or **`country=*ZB`**.

26. To confirm, search (All time):

    `index=itops sourcetype=wocheese host=splunk## (country=*BA OR country=*ZB) | stats count by country`



country	count
BA	279
ZB	45

Answers

ROOT CAUSE:

- The missing data was caused by the transform stanza defined in the **`tseLlab05`** app. The stanza drops any events tagged with the **`ghostwww`** sourcetype and the **`host`** value containing **`www1`**.

- The mysterious search behavior was caused by a different expectation of how Splunk search works. The search-time field extraction works with the indexed keywords and they are based on the defined segmenters. By default, Splunk considers the **`<space>`** character as a keyword segmenter. If a tsidx bucket does not have the keyword that a search-time extraction is based on, then any subsequent extraction cannot be able to continue.

REMEDY:

- Edit the **`props.conf`** and disable the invocation of the event-dropping transformation.

- Force the search to match the format of keyword that the extraction is based on.

  A better but more complicated option is to use the **`fields.conf`**:

  ```
  [country]
  INDEXED_VALUE=*<VALUE>
  ```

  Place the **`fields.conf`** file in **`SPLUNK_HOME/etc/system/local/`** and restart.
  This definition generates a lispy term of "**`*zb`**" or "**`*ba`**" for searches such as **`country=ZB`**.
  Or, create an index-time field extraction.