Data Compliance Pipelines for FSI

Learn How to Filter, Mask, Route, and Monitor using Splunk Data Management pipelines to meet KYC, PCI, DORA, RMiT, and CPS requirements.

splunk>
a CISCO company

Forward-looking statements

This presentation may contain forward-looking statements that are subject to the safe harbors created under the Securities Act of 1933, as amended, and the Securities Exchange Act of 1934, as amended. All statements other than statements of historical facts are statements that could be deemed forward-looking statements. These statements are based on current expectations, estimates, forecasts, and projections about the industries in which we operate and the beliefs and assumptions of our management based on the information currently available to us. Words such as “expects,” “anticipates,” “targets,” “goals,” “projects,” “intends,” “plans,” “believes,” “momentum,” “seeks,” “estimates,” “continues,” “endeavors,” “strives,” “may,” variations of such words, and similar expressions are intended to identify such forward-looking statements. In addition, any statements that refer to (1) our goals, commitments, and programs; (2) our business plans, initiatives, and objectives; and (3) our assumptions and expectations, including our expectations regarding our financial performance, products, technology, strategy, customers, markets, acquisitions and investments are forward-looking statements. These forward-looking statements are not guarantees of future performance and involve significant risks, uncertainties and other factors that may cause our actual results, performance or achievements to be materially different from results, performance or achievements expressed or implied by the forward-looking statements contained in this presentation. Readers are cautioned that these forward-looking statements are only predictions and are subject to risks, uncertainties, and assumptions that are difficult to predict, including those identified in the “Risk Factors” section of Cisco’s most recent report on Form 10-Q filed on February 20, 2024 and its most recent report on Form 10-K filed on September 7, 2023, as well as the “Risk Factors” section of Splunk’s most recent report on Form 10-Q filed with the SEC on November 28, 2023. The forward-looking statements made in this presentation are made as of the time and date of this presentation. If reviewed after the initial presentation, even if made available by Cisco or Splunk, on Cisco or Splunk’s website or otherwise, it may not contain current or accurate information. Cisco and Splunk undertake no obligation to revise or update any forward-looking statements for any reason, except as required by law.

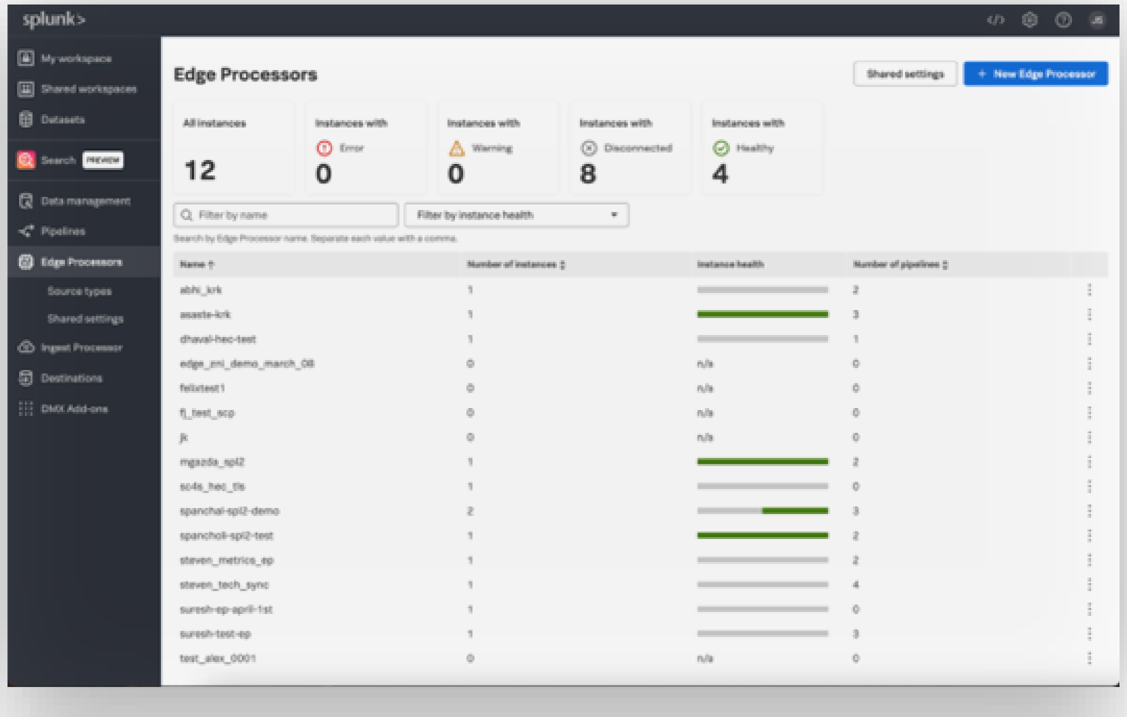
In addition, any information about new products, features, functionality or our roadmap outlines our general product direction and is subject to change at any time without notice. It is for informational purposes only and shall not be incorporated into any contract or other commitment or be relied upon in making a purchasing decision. We undertake no commitment, promise or obligation either to develop the features or functionalities described, in beta or in preview (used interchangeably), or to include any such feature or functionality in a future release. The development, release, and timing of any features or functionality described for our products remains at our sole discretion.

splunk>
a CISCO company

Why?

Efficiently Manage Compliance Data Pipelines

- Discover the specialized tools available for ensuring data compliance and understand how to use them effectively.
- Learn how to manage and monitor compliance-related data directly from your browser.
- Understand how to filter, mask, transform, enrich, and route sensitive data to multiple compliant destinations.
- See how you can receive rapid feedback to confirm that your compliance measures and data changes are functioning as intended.

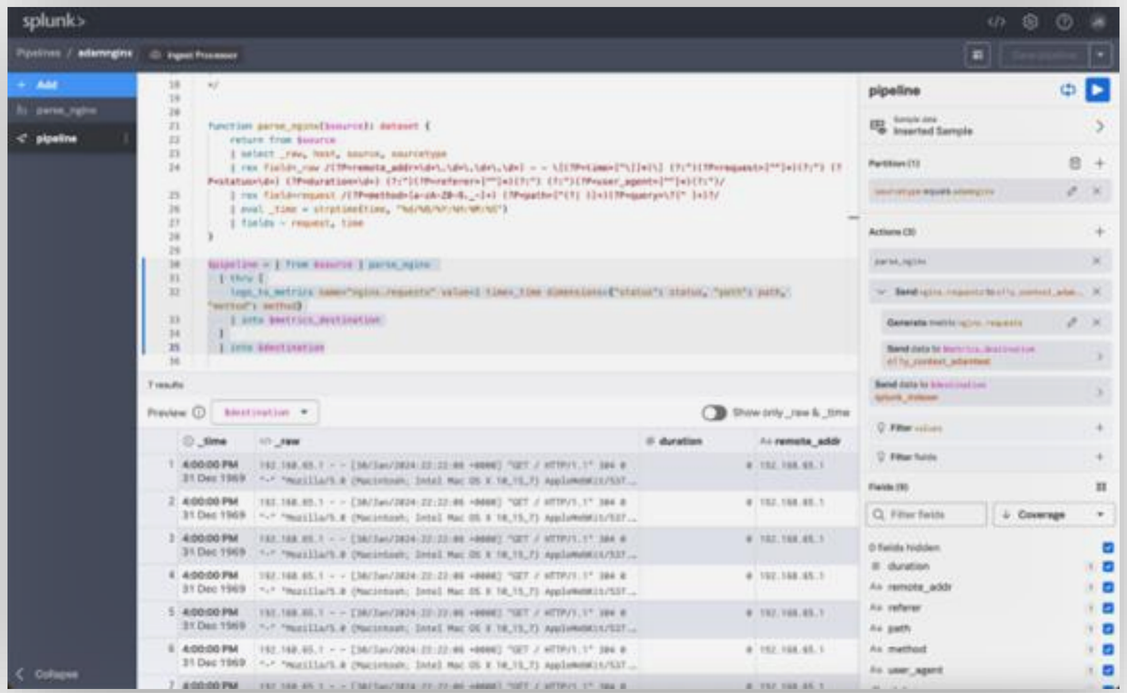


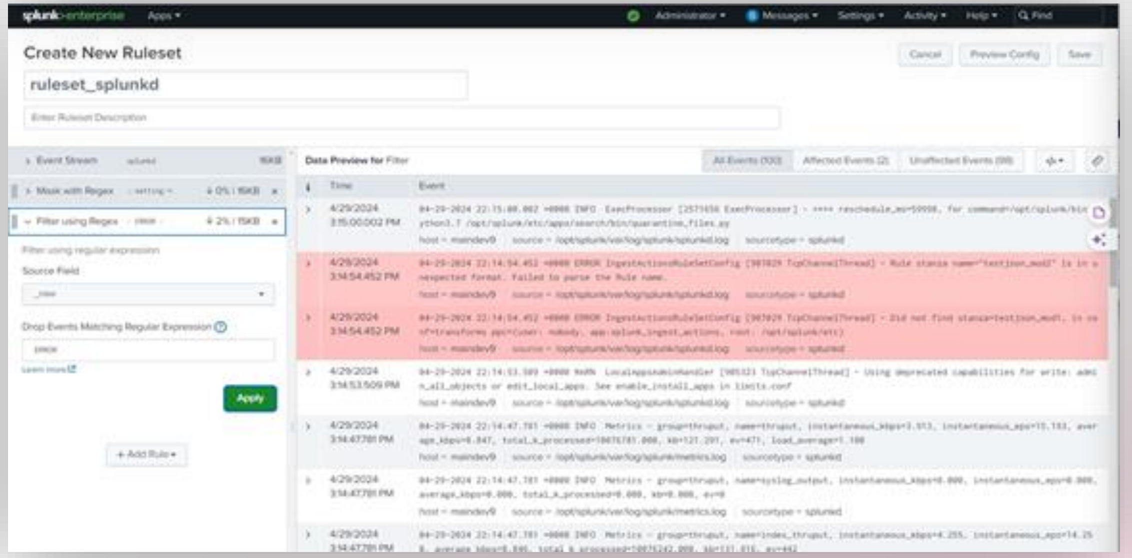
Edge Processors

All instances	Instances with Error	Instances with Warning	Instances with Disconnected	Instances with Healthy
12	0	0	8	4

Search by Edge Processor name. Separate each value with a comma.

Name ↑	Number of instances	Instance health	Number of pipelines
abhi_krk	1		2
asaste-krk	1		3
dhaval-hec-test	1		1
edge_zni_demo_march_08	0	n/a	0
felixtest1	0	n/a	0
fj_test_scp	0	n/a	0
jk	0	n/a	0
mgazda_spl2	1		2
sc4s_hec_tls	1		0
spanchal-spl2-demo	2		3
spanchol-spl2-test	1		2
steven_metrics_ep	1		2
steven_tech_sync	1		4
suresh-ep-april-1st	1		0
suresh-test-ep	1		3
test_alex_0001	0	n/a	0





Create New Ruleset

ruleset_splunkd

Workshop Agenda

- Building Data Compliance Pipelines
- Create an Edge Processor Node
- Create an Amazon S3 Destination
- Create a Data Pipeline for KYC
- Create a Data Pipeline for PCI
- Create a Data Pipeline for DORA
- Create a Data Pipeline for RMiT
- Create a Data Pipeline for CPS 230
- Splunk resources

REGISTRATION 5 MINS

Enroll in Today’s Workshop

Tasks

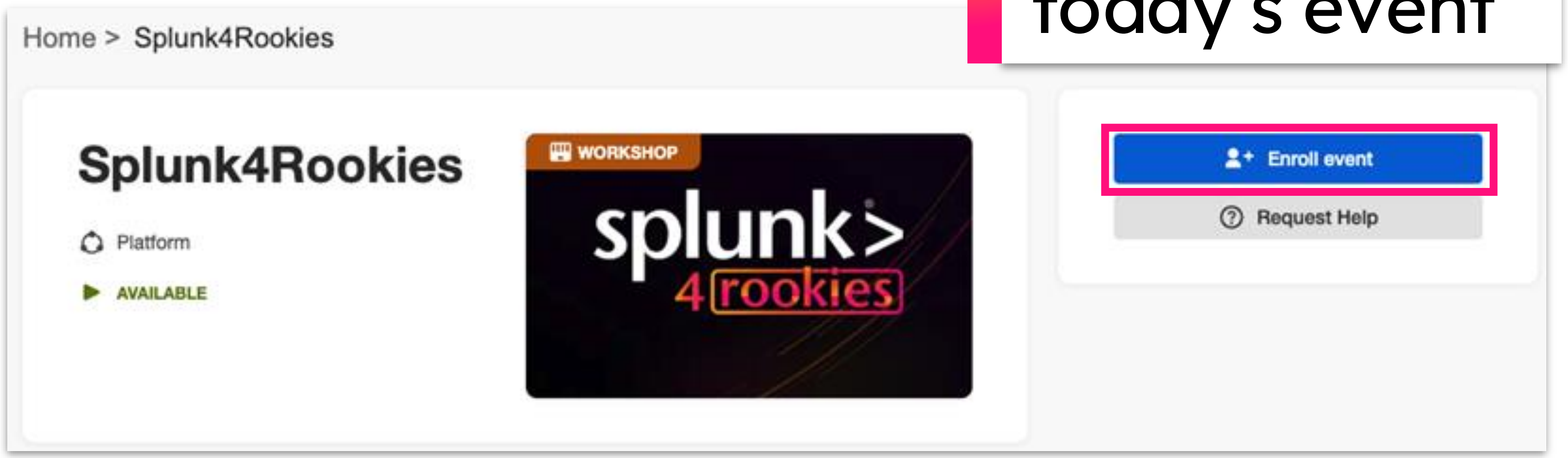
1. Get a splunk.com account if you don’t have one yet:
   https://splk.it/SignUp
2. Enroll in the Splunk Show workshop event:
   https://splunk.show/<uniqueID>
3. Download the hands-on lab guide:
   https://splk.it/Data-FSI-Lab-Guide

   Contains step-by-step instructions for all of today’s exercises!

4. Download a copy of today’s slide deck:
   https://splk.it/Data-FSI-Attendee

Goal

Enroll in today’s event



DOWNLOAD MATERIALS 5 MINS

Obtain the Materials for Today’s Workshop

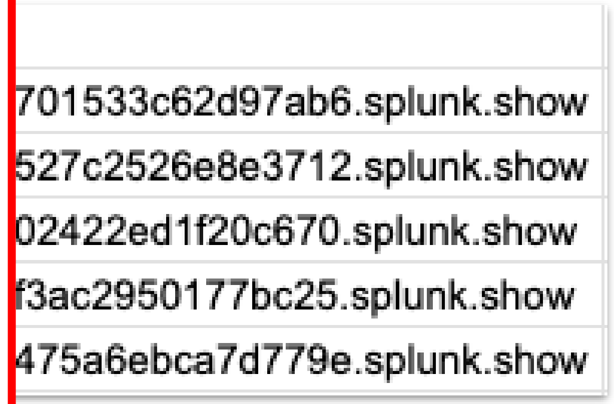
Tasks

1. Get your instan
<link to sheet w
2. Download the h
https://splk.it/D

Contains
for all of t

3. Download a cop
https://splk.it/D

Goal


701533c62d97ab6.splunk.show
527c2526e8e3712.splunk.show
02422ed1f20c670.splunk.show
f3ac2950177bc25.splunk.show
475a6ebca7d779e.splunk.show

name against one
ances in the list

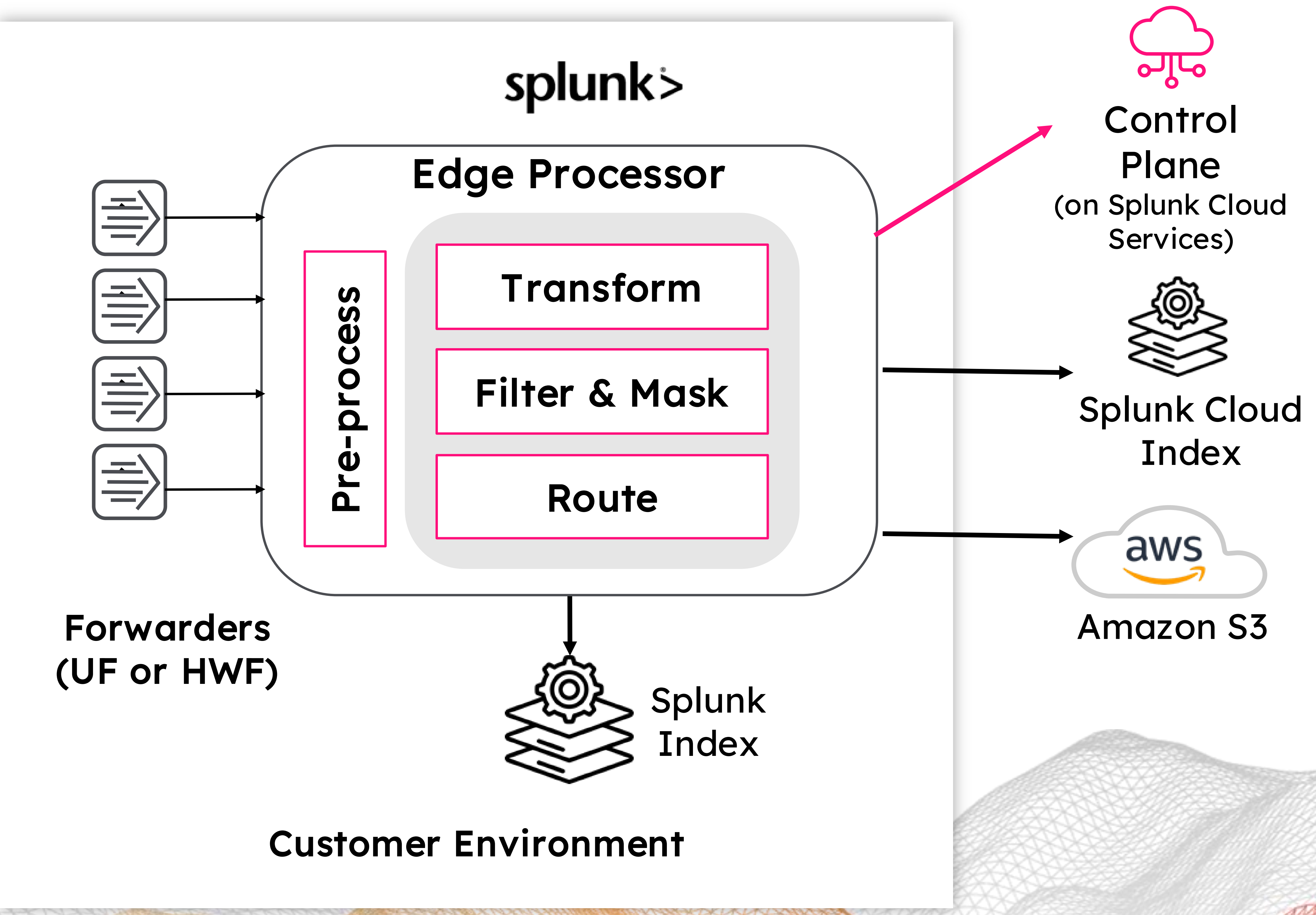
**** Optional slide for running a ‘normal’ workshop ****
(i.e. not an ‘event’) in Splunk Show

Presenter instructions:

1. **Create a ‘normal’ workshop in Splunk Show** selecting the required number of instances you require for your workshop (see the Splunk Show User Guide for Splunkers or for Partners)
2. Once the instances are all running, **export a CSV of the instances**
3. **Share the list of instances via a spreadsheet** (example here) or some other method that will allow attendees to obtain their own instance from the list.
4. **Copy the URL for your spreadsheet into step 1 of this slide** (underneath this text box!)
5. **Hide the previous 'Enroll in Today's Workshop' slide** and **unhide this slide**!
6. During the workshop **ask users to put their names against a free instance in the spreadsheet** as a way of tracking who is allocated to which instance
7. **Delete or move this text box off screen** before presenting!

What is Edge Processing?

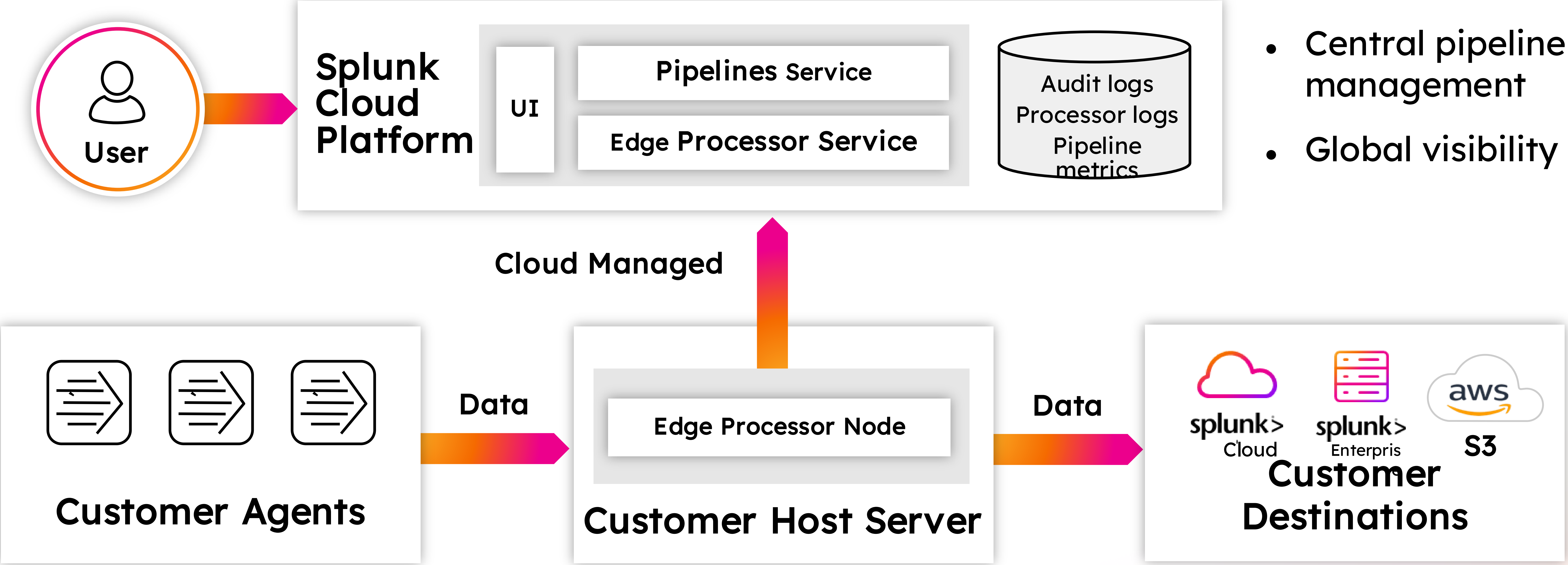
Double-click into the key use cases



- Filter verbose or low-value sources, like DEBUG logs or other noisy data
- Extract just the critical data
- Mask PII
- Route different “slices” of data to desired destinations

Edge Processor Overview

Use cases delivered through control plane in Splunk Cloud Platform



- Central pipeline management
- Global visibility

Edge Processor

Data Admin / SPL2 User

WHEN TO USE		HOW IT'S DONE
Reduce noise / volume	→	Filtering
Redact sensitive data	→	Masking
Add data via lookups / evals	→	**Enriching**
Add indexed fields	→	**Transforming**
Send to indexes / S3	→	Routing

WHAT YOU NEED

- Edge Processing nodes in your own network
- **Cloud stack for pipeline authoring and management of Edge Processing nodes**
- Defined sourcetypes

Today’s Scenario

Your Company

- Buttercup Bank is a leading international financial institution, serving customers across multiple countries with a comprehensive range of banking services through its global online platform.

- Buttercup Bank has recently invested in Data Compliance Pipelines to strengthen its management, monitoring, and protection of sensitive financial data, ensuring adherence to global regulatory standards such as PCI, KYC, DORA, RMiT, and CPS 230.

Your Role

- You are one of the chosen few: a Splunk power user!

- Your responsibility is to ensure proper data handling, regulatory compliance, and risk management across the bank's international operations.

- The teams you work with include:
  - IT Operations
  - DevOps
  - Business Analytics
  - Security and Fraud

BUTTERCUP
ENTERPRISES

Know Your Customer (KYC) - Global

Requirement 1: Archive Raw Data

- Send an unmodified copy of the raw data to Amazon S3 for archival.

Requirement 2: Sanitize PII

- Redact, mask, or hash Personally Identifiable Information (PII) such as full names, dates of birth, emails, phone numbers, street addresses, and document IDs, before the data reaches Splunk Cloud for analysis.

Requirement 3: Filter Relevant Events

- Keep only KYC events that have been "approved”.
- e.g., event_type == "kyc_approved" and event_status == "decision_approved".

Requirement 4: Route to a Specific Index

- Send the processed and sanitized data to a KYC specific index in Splunk Cloud.

Payment Card Industry Data Security Standard (PCI) | AMER

Requirement 1: Sanitize Sensitive Fields

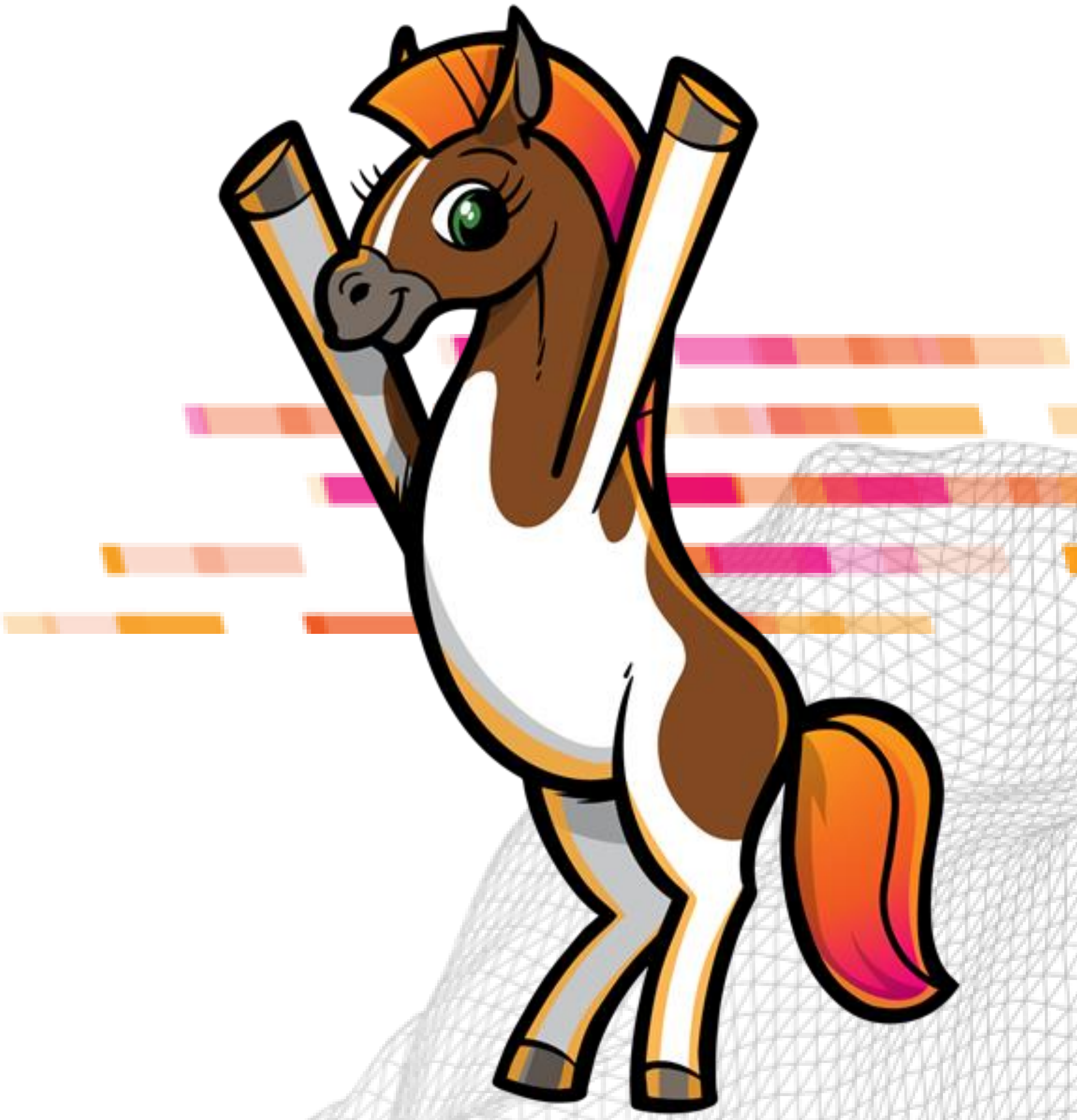
- Redact cardholder names, and mask Primary Account Numbers (PANs) and account numbers according to PCI DSS requirements.

Requirement 2: Drop Non-Storable Fields

- Remove fields that are explicitly forbidden from being stored (e.g., CVV, track data, PIN).

Requirement 3: Route to Specific Index

- Send the processed and sanitized data to a PCI specific index in Splunk Cloud.



BUTTERCUP
ENTERPRISES

Digital Operational Resilience Act (DORA) | EMEA

Requirement 1: Archive Raw Data

- Send an unmodified copy of the raw data to Amazon S3 for archival.

Requirement 2: Optimize Event Size

- Drop the _raw field after archival.

Requirement 3: Pseudonymize/Anonymize Data

- Mask IP addresses and pseudonymize User IDs using a salted hash for enhanced privacy while allowing correlation if needed.

Requirement 4: Route Based on Severity

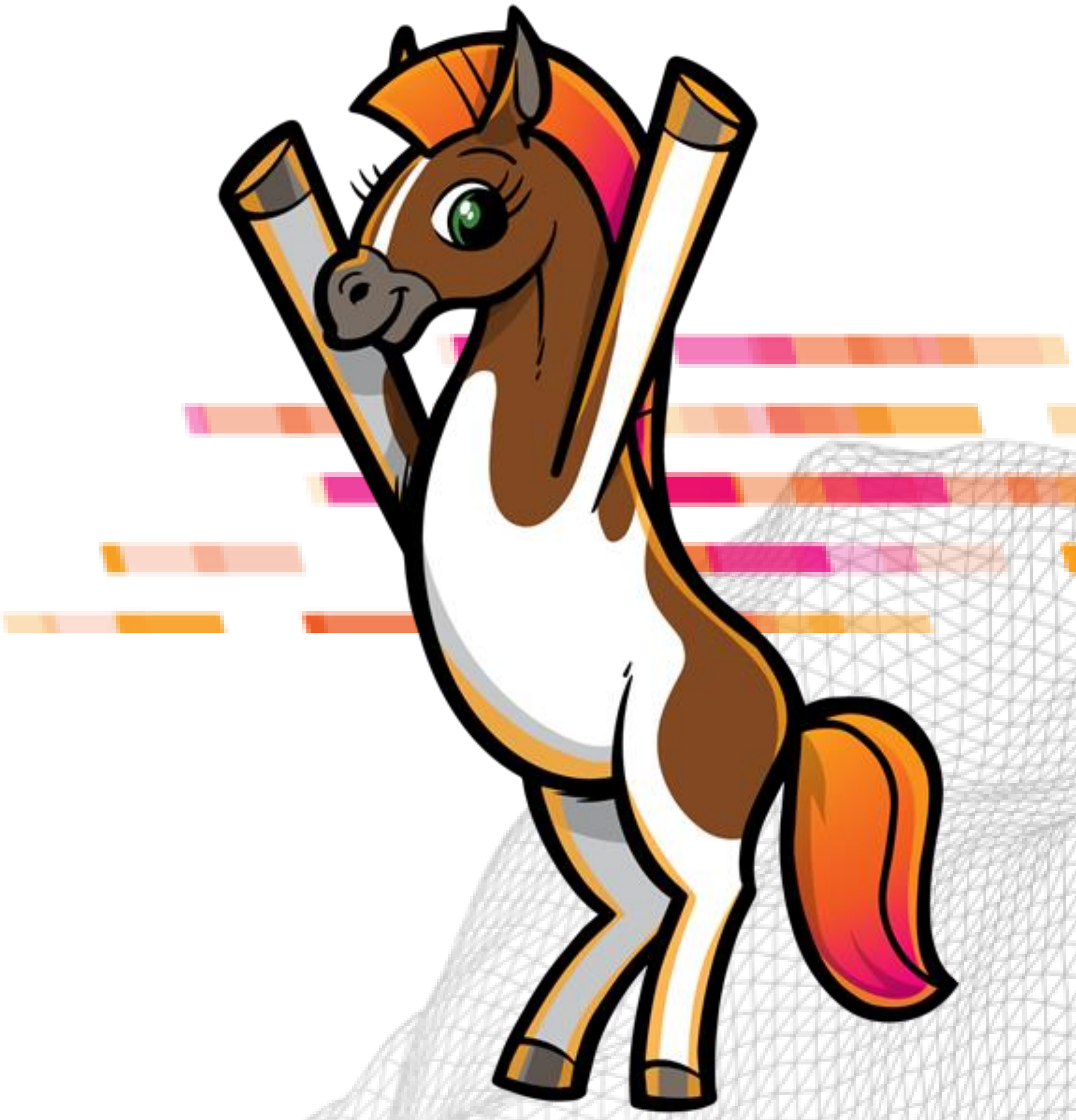
- Direct events with "ERROR" or "CRITICAL" severity to a specific index (dora_errors_events) and "WARNING" severity events to another (dora_warning_events).

Requirement 5: Minimize Data

- Remove unnecessary fields (incident_description, description, transaction_id, session_id) to reduce data volume for the default stream.

Requirement 6: Route to Specific Index

- Send the processed and sanitized data to a DORA specific index in Splunk Cloud.



BUTTERCUP
ENTERPRISES

Risk Management in Technology (RMiT) | APJC

Requirement 1: Enrich Data with Lookups

- Enhance events with asset inventory information (criticality, location, service impact) using a lookup file.

Requirement 2: Define Risk Levels

- Assign a risk_level (High, Medium, Low) based on the criticality from the lookup.

Requirement 3: Mask Sensitive Operational Details

- Mask IP addresses, pseudonymize user_id, mask reporting_user_or_process, and redact event_description if it contains PII.

Requirement 4: Filter Non-Actionable Logs

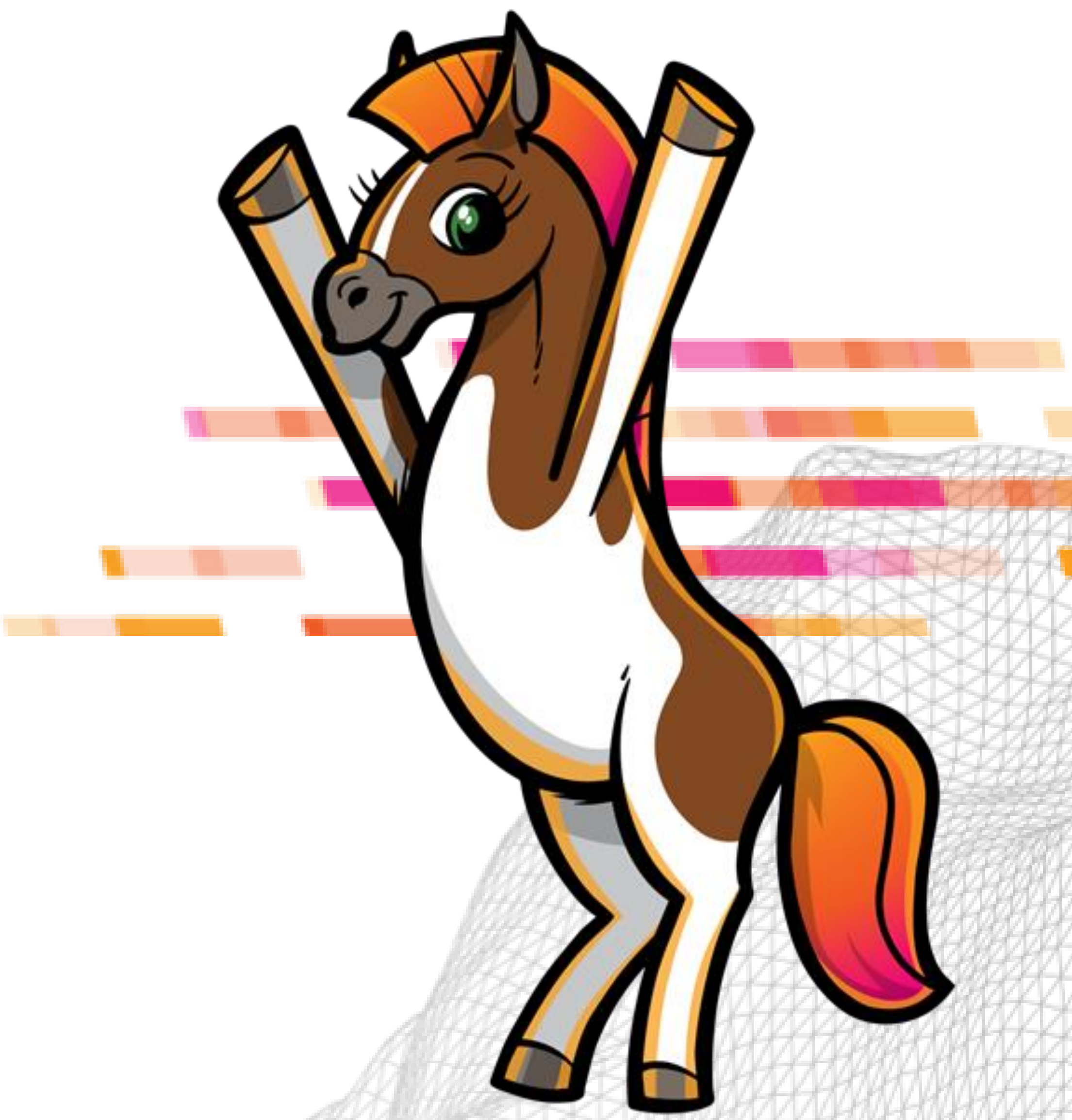
- Retain only relevant events based on status and criticality.

Requirement 5: Minimize Data

- Remove low-value or empty fields from the default stream.

Requirement 6: Route Based on Event Type

- Direct events to different Splunk indexes (rmit_cyber_events, rmit_change_events, rmit_tprm_events) based on their event_type.



BUTTERCUP
ENTERPRISES

CPS 230 | A/NZ

Requirement 1: Enrich Data with Lookups

- Enhance events with asset inventory information (tolerance level, critical operation name, and business unit) using a lookup file.

Requirement 2: Define Risk Levels

- Assign a risk_level (High, Medium, Low) based on the criticality from the lookup.

Requirement 3: Mask Sensitive Operational Details

- Mask IP addresses, pseudonymize user IDs, mask reporting users/processes, and redact event descriptions if they contain PII.

Requirement 4: Filter Non-Actionable Logs

- Retain only relevant events based on status and criticality.

Requirement 5: Minimize Data

- Remove low-value or empty fields from the default stream.

Requirement 6: Route Based on APRA Notification Candidate Status

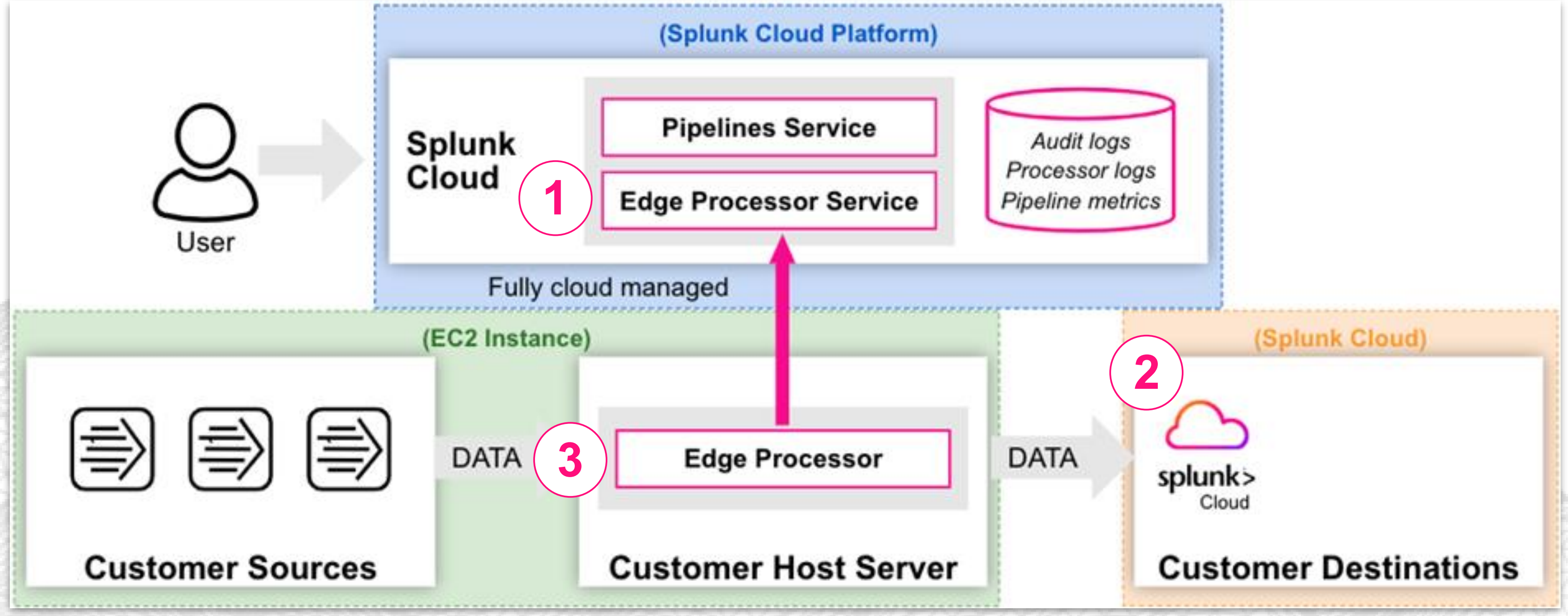
- Direct events to different Splunk indexes (cps_apra_notify_events, cps_significant_events) based on whether they are APRA notification candidates.

LAB EXERCISE 1 5 MINS

Access Your Lab Environment

Tasks

1. Log in to Splunk Show
2. Follow the lab guide steps to access:
   1. Edge Processor Service
   2. Splunk Cloud
   3. EC2 instance (‘EP – Edge Node’)



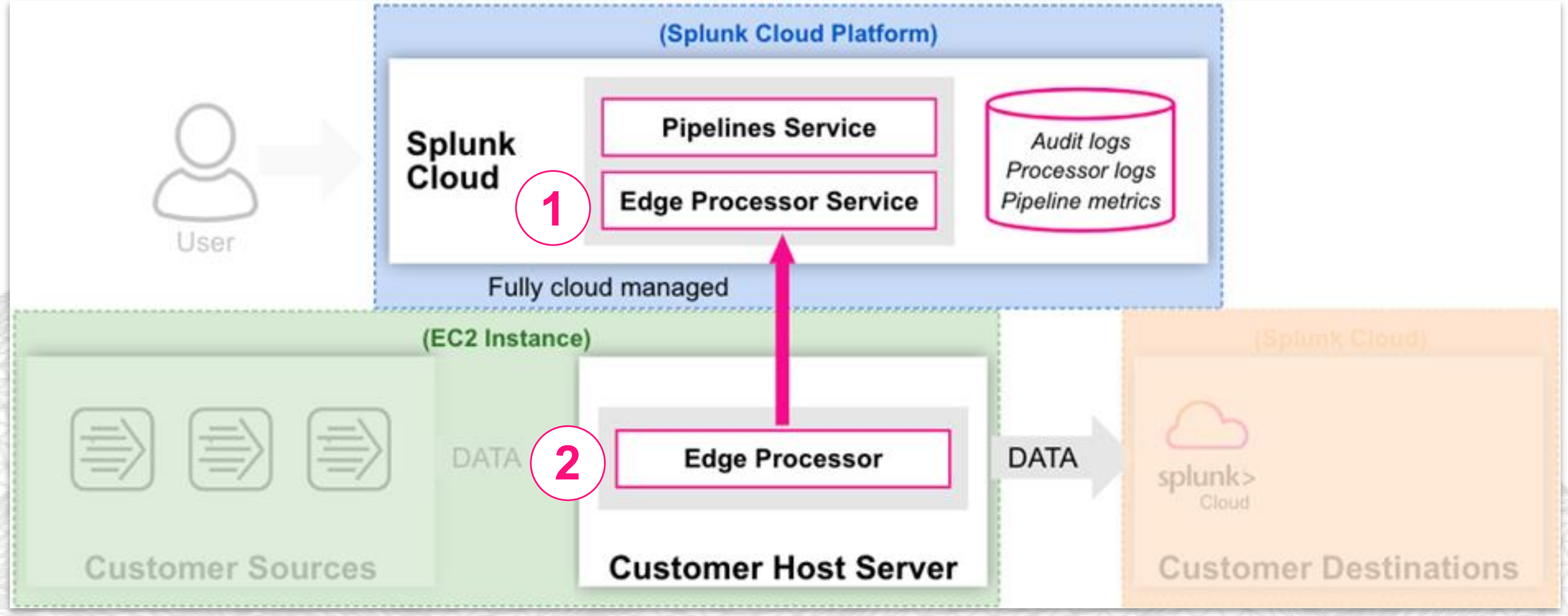
Reminder
Download the lab guide for step-by-step instructions!

LAB EXERCISE 2 5 - 10 MINS

Create an Edge Processor

Tasks

1. Create a new Edge Processor via Edge Processor Service
2. Install a new Edge Processor instance on your EC2 server



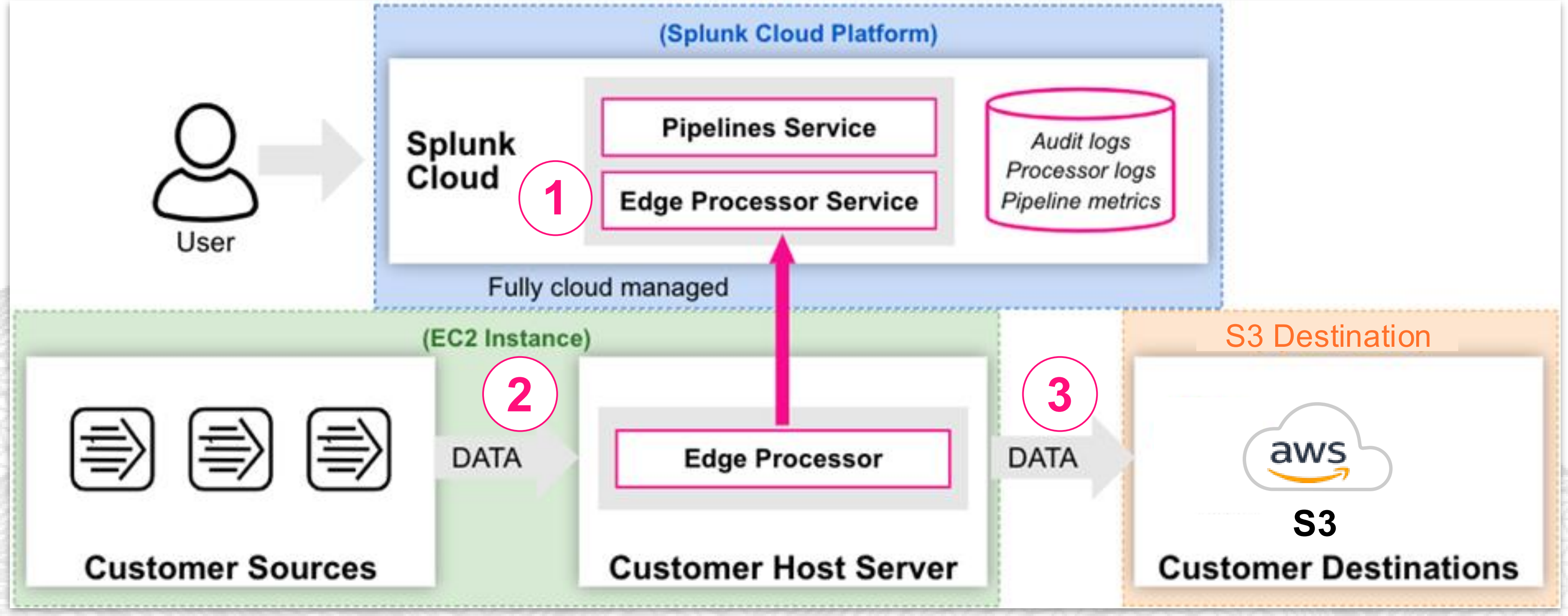
LAB EXERCISE 3 5 – 10 MINS

Create an Amazon S3 Destination

Tasks

1. Create a new Amazon S3 Destination via Edge Processor Service

2. Configure Edge Processor Service to support Amazon S3 Destination



Know Your Customer (KYC) - Universal

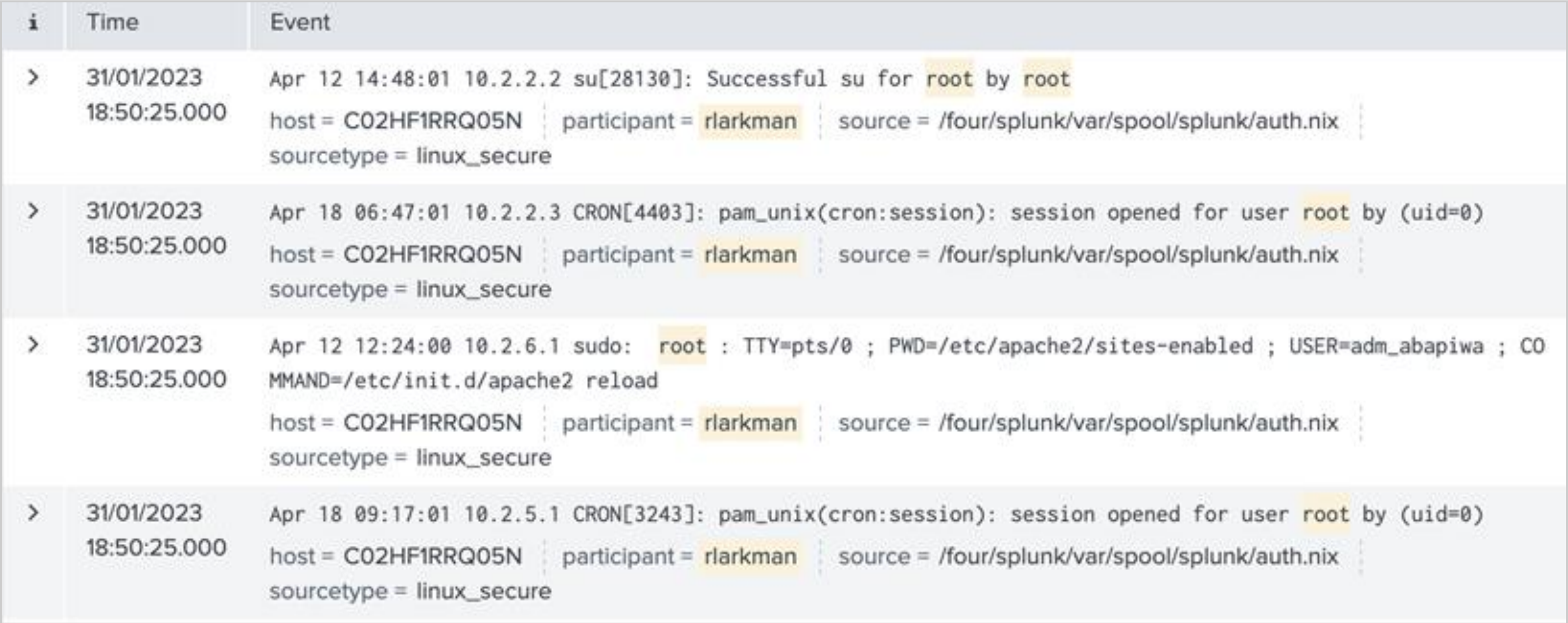
Tasks

Create an Edge Processor pipeline to:

- Send data to Archive Location (S3)
- Sanitize PII Events & Filter/Mask events relating to KYC events
- Send the “kyc” events to the ‘kyc_events’ index in Splunk Cloud
- Add a custom field to your data so you can identify it in Splunk Cloud

Goal

```
index=kyc_events participant="<YourName>"
```



i	Time	Event

>	31/01/2023
	18:50:25.000
		Apr 12 14:48:01 10.2.2.2 su[28130]: Successful su for root by root
		host = C02HF1RRQ05N participant = rlarkman source = /four/splunk/var/spool/splunk/auth.nix
		sourcetype = linux_secure

>	31/01/2023
	18:50:25.000
		Apr 18 06:47:01 10.2.2.3 CRON[4403]: pam_unix(cron:session): session opened for user root by (uid=0)
		host = C02HF1RRQ05N participant = rlarkman source = /four/splunk/var/spool/splunk/auth.nix
		sourcetype = linux_secure

>	31/01/2023
	18:50:25.000
		Apr 12 12:24:00 10.2.6.1 sudo: root : TTY=pts/0 ; PWD=/etc/apache2/sites-enabled ; USER=adm_abapiwa ; COMMAND=/etc/init.d/apache2 reload
		host = C02HF1RRQ05N participant = rlarkman source = /four/splunk/var/spool/splunk/auth.nix
		sourcetype = linux_secure

>	31/01/2023
	18:50:25.000
		Apr 18 09:17:01 10.2.5.1 CRON[3243]: pam_unix(cron:session): session opened for user root by (uid=0)
		host = C02HF1RRQ05N participant = rlarkman source = /four/splunk/var/spool/splunk/auth.nix
		sourcetype = linux_secure

LAB EXERCISE 5 10 MINS

PCI-DSS - AMER

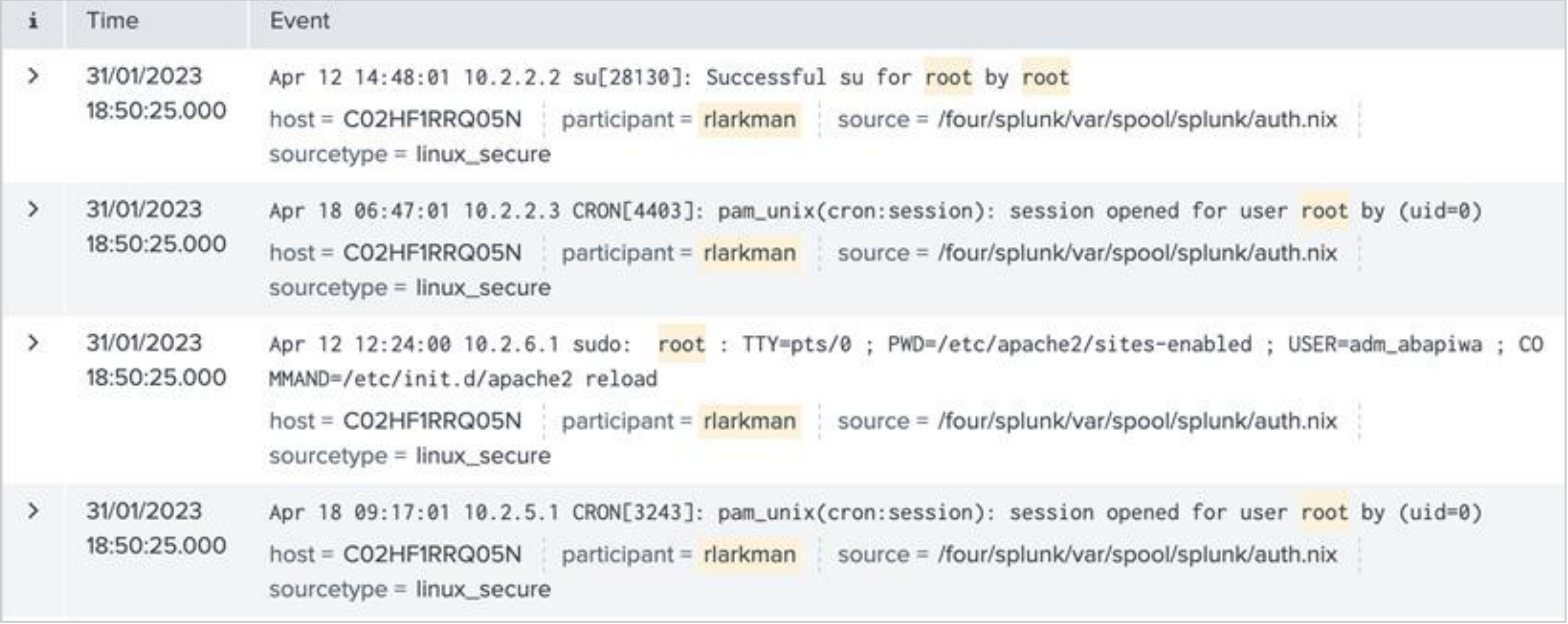
Tasks

Create an Edge Processor pipeline to:

- Sanitize Sensitive Fields
- Drop Non-Storable Fields
- Send the “pci” events to the ‘pci_events’ index in Splunk Cloud
- Add a custom field to your data so you can identify it in Splunk Cloud

Goal

```
index=pci_events participant="<YourName>"
```



i	Time	Event

>	31/01/2023
	18:50:25.000
		Apr 12 14:48:01 10.2.2.2 su[28130]: Successful su for root by root
		host = C02HF1RRQ05N participant = rlarkman source = /four/splunk/var/spool/splunk/auth.nix
		sourcetype = linux_secure

>	31/01/2023
	18:50:25.000
		Apr 18 06:47:01 10.2.2.3 CRON[4403]: pam_unix(cron:session): session opened for user root by (uid=0)
		host = C02HF1RRQ05N participant = rlarkman source = /four/splunk/var/spool/splunk/auth.nix
		sourcetype = linux_secure

>	31/01/2023
	18:50:25.000
		Apr 12 12:24:00 10.2.6.1 sudo: root : TTY=pts/0 ; PWD=/etc/apache2/sites-enabled ; USER=adm_abapiwa ; COMMAND=/etc/init.d/apache2 reload
		host = C02HF1RRQ05N participant = rlarkman source = /four/splunk/var/spool/splunk/auth.nix
		sourcetype = linux_secure

>	31/01/2023
	18:50:25.000
		Apr 18 09:17:01 10.2.5.1 CRON[3243]: pam_unix(cron:session): session opened for user root by (uid=0)
		host = C02HF1RRQ05N participant = rlarkman source = /four/splunk/var/spool/splunk/auth.nix
		sourcetype = linux_secure

LAB EXERCISE 6 10 MINS

Digital Operational Resilience Act - (DORA)

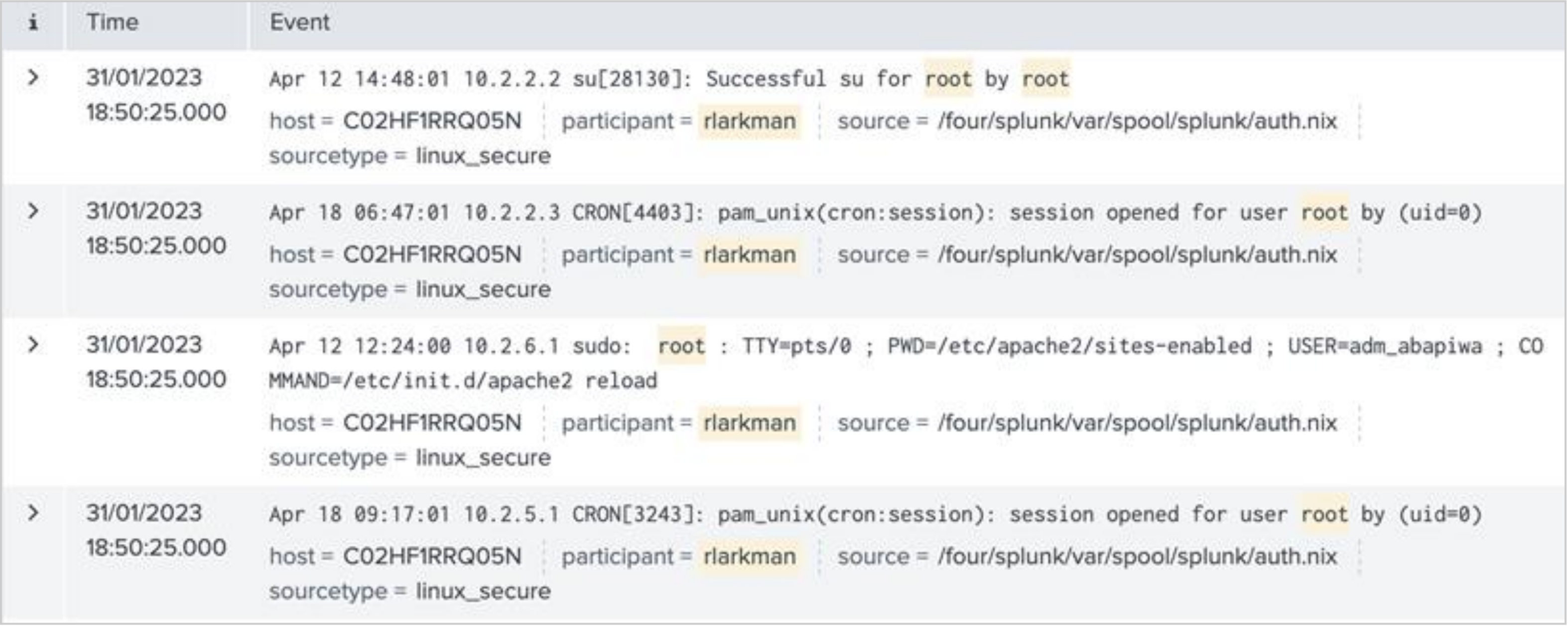
Tasks

Create an Edge Processor pipeline to:

- Send data to Archive Location (S3)
- Filter/Mask events relating to KYC events
- Route “dora” events based on severity to specific index in Splunk Cloud
- Add a custom field to your data so you can identify it in Splunk Cloud

Goal

```
index=dora_events participant="<YourName>"
```



i	Time	Event

>	31/01/2023
	18:50:25.000
		Apr 12 14:48:01 10.2.2.2 su[28130]: Successful su for root by root
		host = C02HF1RRQ05N participant = rlarkman source = /four/splunk/var/spool/splunk/auth.nix
		sourcetype = linux_secure

>	31/01/2023
	18:50:25.000
		Apr 18 06:47:01 10.2.2.3 CRON[4403]: pam_unix(cron:session): session opened for user root by (uid=0)
		host = C02HF1RRQ05N participant = rlarkman source = /four/splunk/var/spool/splunk/auth.nix
		sourcetype = linux_secure

>	31/01/2023
	18:50:25.000
		Apr 12 12:24:00 10.2.6.1 sudo: root : TTY=pts/0 ; PWD=/etc/apache2/sites-enabled ; USER=adm_abapiwa ; COMMAND=/etc/init.d/apache2 reload
		host = C02HF1RRQ05N participant = rlarkman source = /four/splunk/var/spool/splunk/auth.nix
		sourcetype = linux_secure

>	31/01/2023
	18:50:25.000
		Apr 18 09:17:01 10.2.5.1 CRON[3243]: pam_unix(cron:session): session opened for user root by (uid=0)
		host = C02HF1RRQ05N participant = rlarkman source = /four/splunk/var/spool/splunk/auth.nix
		sourcetype = linux_secure

LAB EXERCISE 7 10 MINS

Risk Management in Technology (RMiT) - APJC

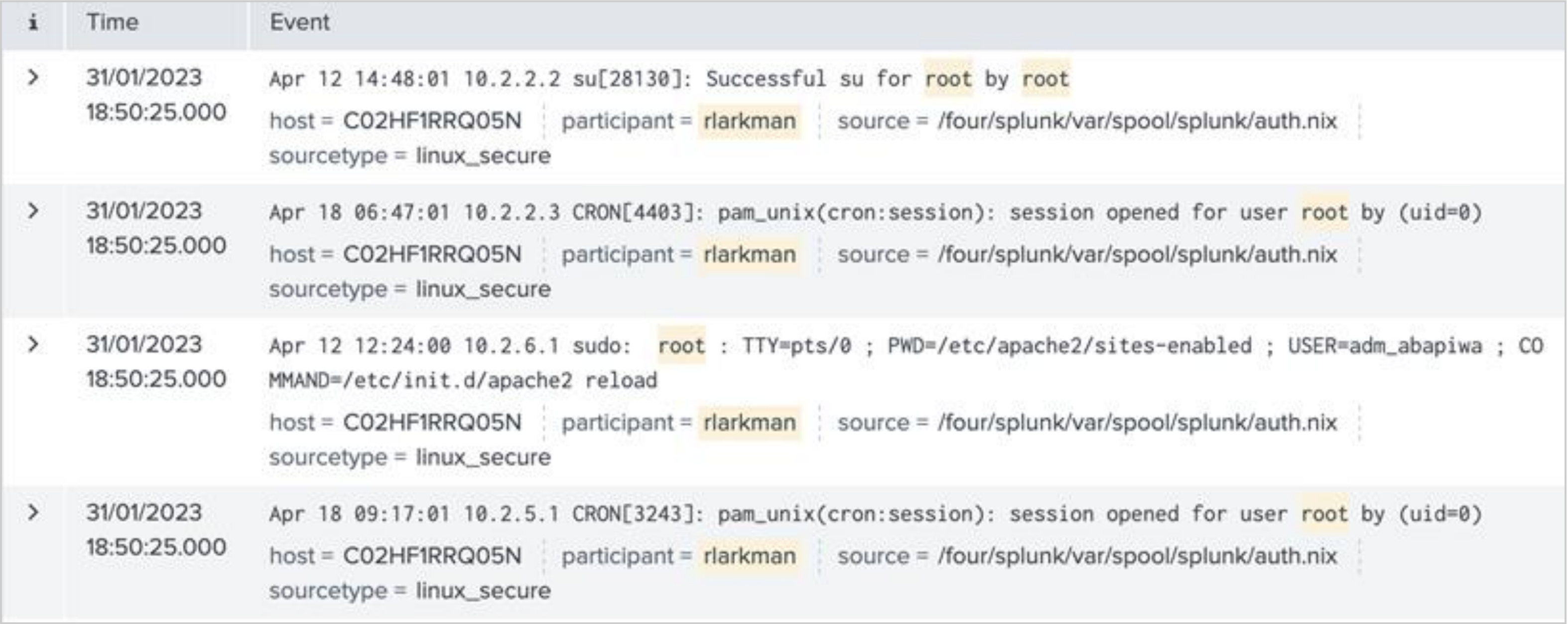
Tasks

Create an Edge Processor pipeline to:

- Enrich events with lookups
- Mask sensitive operational details
- Filter Non-Actionable Logs
- Send the “rmit” events to the ‘rmit_events’ index in Splunk Cloud
- Add a custom field to your data so you can identify it in Splunk Cloud

Goal

```
index=rmit_events participant="<YourName>"
```



```
i  Time                      Event
>  31/01/2023 18:50:25.000   Apr 12 14:48:01 10.2.2.2 su[28130]: Successful su for root by root
                             host = C02HF1RRQ05N   participant = rlarkman   source = /four/splunk/var/spool/splunk/auth.nix
                             sourcetype = linux_secure
>  31/01/2023 18:50:25.000   Apr 18 06:47:01 10.2.2.3 CRON[4403]: pam_unix(cron:session): session opened for user root by (uid=0)
                             host = C02HF1RRQ05N   participant = rlarkman   source = /four/splunk/var/spool/splunk/auth.nix
                             sourcetype = linux_secure
>  31/01/2023 18:50:25.000   Apr 12 12:24:00 10.2.6.1 sudo:  root : TTY=pts/0 ; PWD=/etc/apache2/sites-enabled ; USER=adm_abapiwa ; CO
                             MMAND=/etc/init.d/apache2 reload
                             host = C02HF1RRQ05N   participant = rlarkman   source = /four/splunk/var/spool/splunk/auth.nix
                             sourcetype = linux_secure
>  31/01/2023 18:50:25.000   Apr 18 09:17:01 10.2.5.1 CRON[3243]: pam_unix(cron:session): session opened for user root by (uid=0)
                             host = C02HF1RRQ05N   participant = rlarkman   source = /four/splunk/var/spool/splunk/auth.nix
                             sourcetype = linux_secure
```

LAB EXERCISE 8 10 MINS

CPS 230 – AN/Z

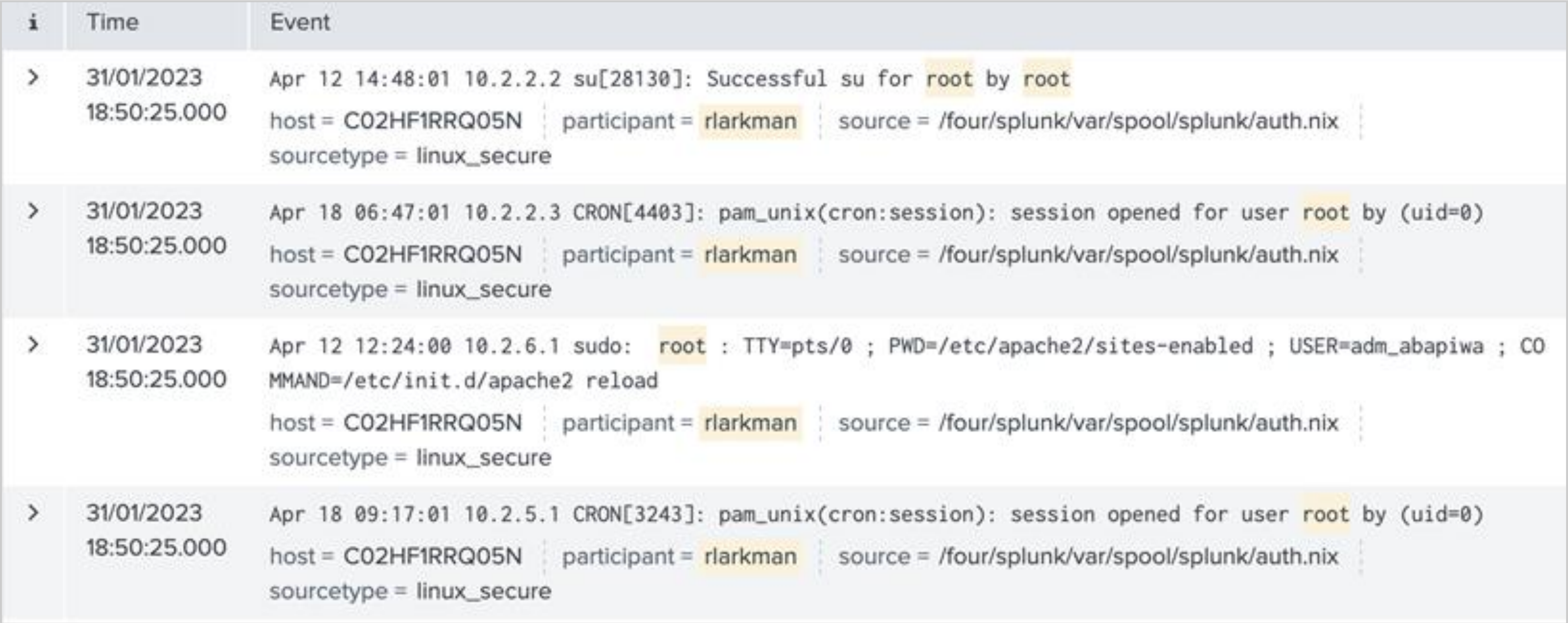
Tasks

Create an Edge Processor pipeline to:

- Enrich events with lookups
- Define Risk Levels
- Mask sensitive operational details
- Filter Non-Actionable Logs
- Send the “cps” events to the ‘cps_events’ index in Splunk Cloud
- Add a custom field to your data so you can identify it in Splunk Cloud

Goal

```
index=cps_events participant="<YourName>"
```



i	Time	Event

>	31/01/2023
	18:50:25.000
		Apr 12 14:48:01 10.2.2.2 su[28130]: Successful su for root by root
		host = C02HF1RRQ05N participant = rlarkman source = /four/splunk/var/spool/splunk/auth.nix
		sourcetype = linux_secure

>	31/01/2023
	18:50:25.000
		Apr 18 06:47:01 10.2.2.3 CRON[4403]: pam_unix(cron:session): session opened for user root by (uid=0)
		host = C02HF1RRQ05N participant = rlarkman source = /four/splunk/var/spool/splunk/auth.nix
		sourcetype = linux_secure

>	31/01/2023
	18:50:25.000
		Apr 12 12:24:00 10.2.6.1 sudo: root : TTY=pts/0 ; PWD=/etc/apache2/sites-enabled ; USER=adm_abapiwa ; COMMAND=/etc/init.d/apache2 reload
		host = C02HF1RRQ05N participant = rlarkman source = /four/splunk/var/spool/splunk/auth.nix
		sourcetype = linux_secure

>	31/01/2023
	18:50:25.000
		Apr 18 09:17:01 10.2.5.1 CRON[3243]: pam_unix(cron:session): session opened for user root by (uid=0)
		host = C02HF1RRQ05N participant = rlarkman source = /four/splunk/var/spool/splunk/auth.nix
		sourcetype = linux_secure

Workshop Summary

- Today you’ve used Edge Processor to create Data Compliance Pipelines that:
  - ✓ Route data
  - ✓ Filter and mask data
  - ✓ Transform data
  - ✓ Enrich Data
  - ✓ Archive Data

- Splunk documentation: https://docs.splunk.com/Documentation/SplunkCloud/latest/EdgeProcessor

Workshop Questions

Thank you

splunk> a CISCO company