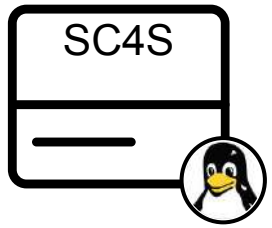
Splunk4Rookies - Splunk Connect for Syslog (SC4S)



#Splunk4Rookies

splunk>

Forward-Looking Statements

This presentation may contain forward-looking statements regarding future events, plans or the expected financial performance of our company, including our expectations regarding our products, technology, strategy, customers, markets, acquisitions and investments. These statements reflect management's current expectations, estimates and assumptions based on the information currently available to us. 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.

For additional information about factors that could cause actual results to differ materially from those described in the forward-looking statements made in this presentation, please refer to our periodic reports and other filings with the SEC, including the risk factors identified in our most recent quarterly reports on Form 10-Q and annual reports on Form 10-K, copies of which may be obtained by visiting the Splunk Investor Relations website at www.investors.splunk.com or the SEC's website at www.sec.gov. 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 us, on our website or otherwise, it may not contain current or accurate information. We disclaim any obligation to update or revise any forward-looking statement based on new information, future events or otherwise, except as required by applicable law.

In addition, any information about 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. We undertake no 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.

splunk>

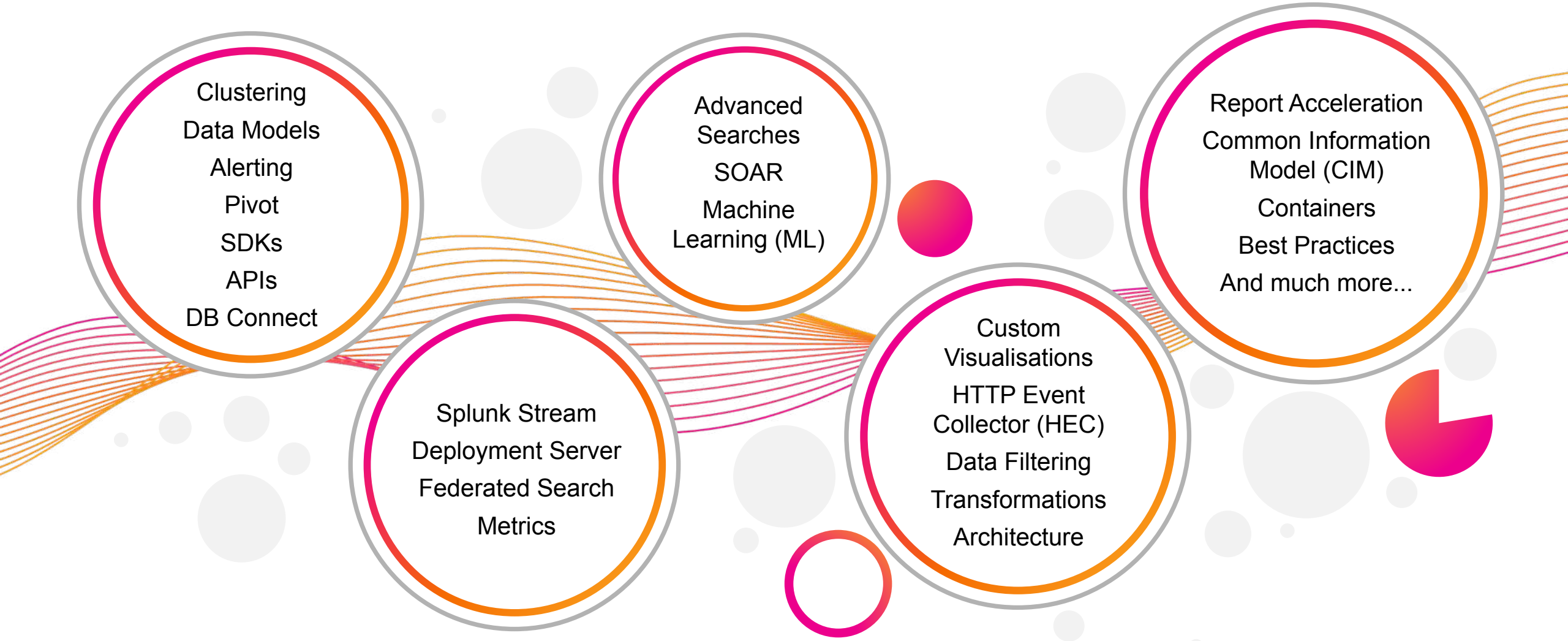
Please introduce yourself!

- Name
- Company/organisation
- Role
- Are you currently using Splunk?
- What are you interested in using Splunk for?

Workshop Agenda

- The value of data
- Splunk’s investigative approach to data
- SC4S overview
- Why use it?
- SC4S 101
- INFOSEC FTW!
- SC4S 201

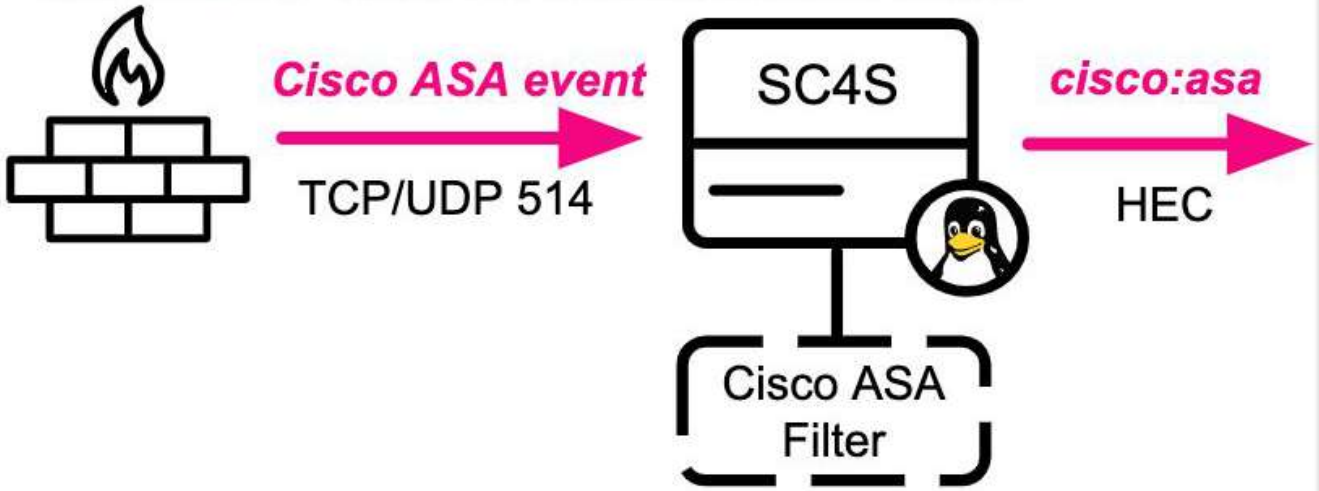
There's a Lot More to Splunk



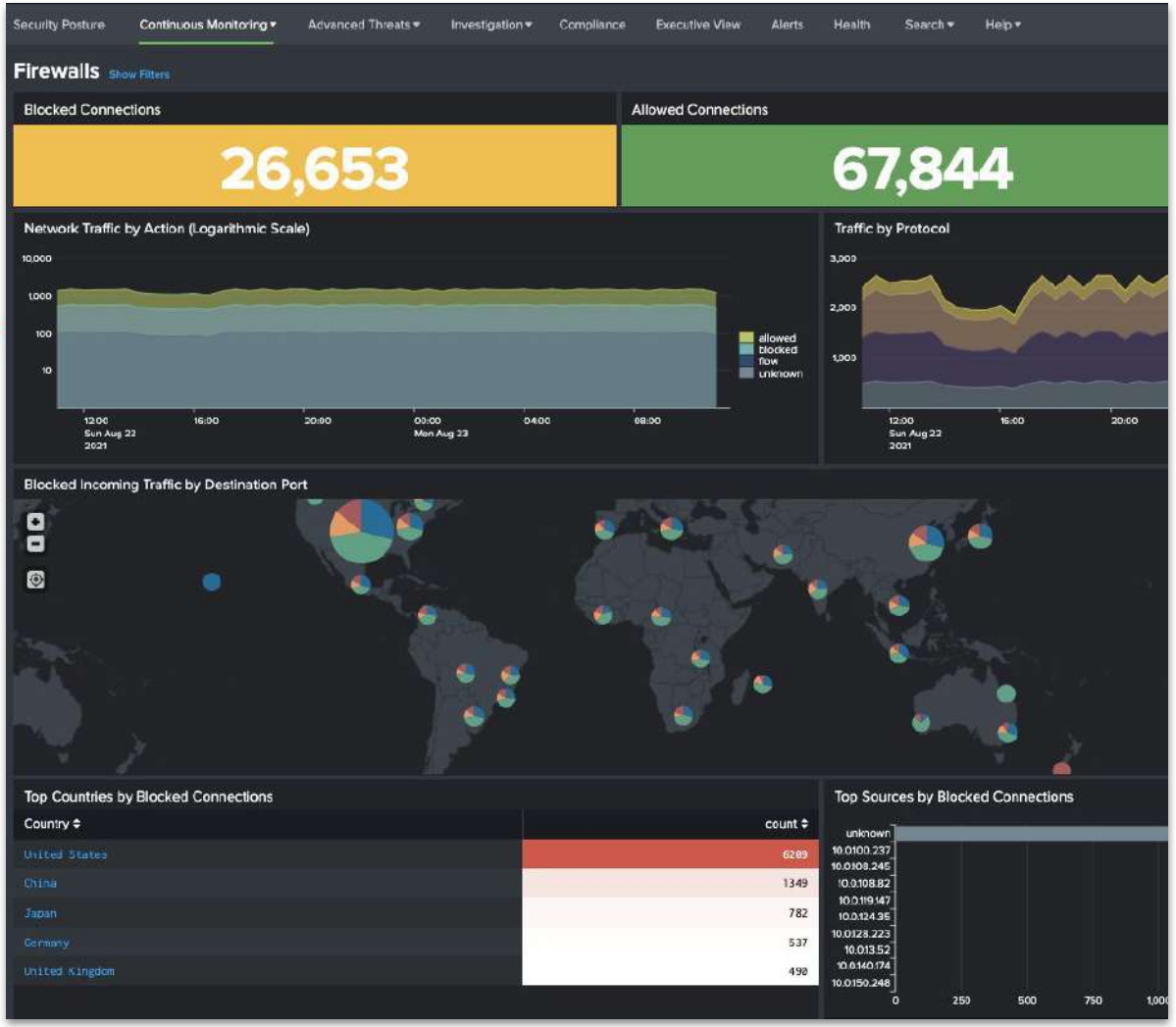
Visit https://splunk.com/training to learn more!

splunk>

Objective: SIEM in ~3 hours



Identify > Parse > Format



splunk>

REGISTRATION 5 MINS

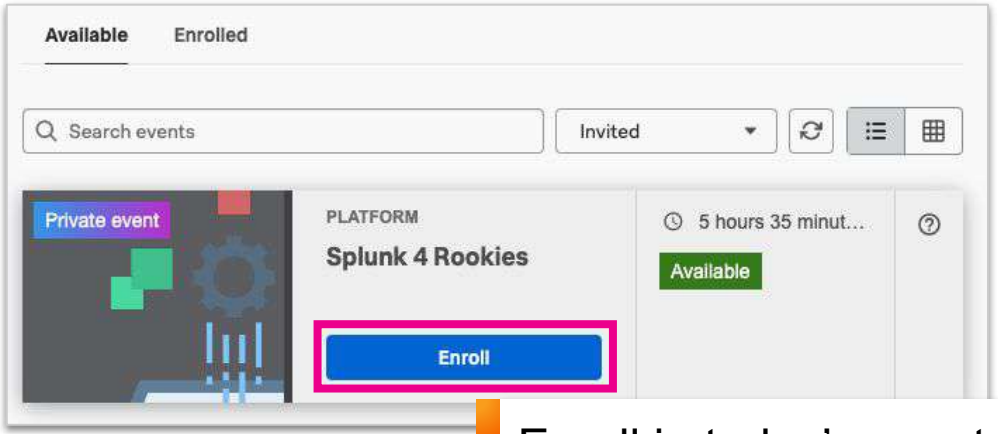
Enroll in Today's Workshop

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://show.splunk.com/event/<eventID>
3. Download the hands-on lab guide: https://splk.it/S4R-SC4S-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/S4R-SC4S-Attendee

Goal



Enroll in today's event

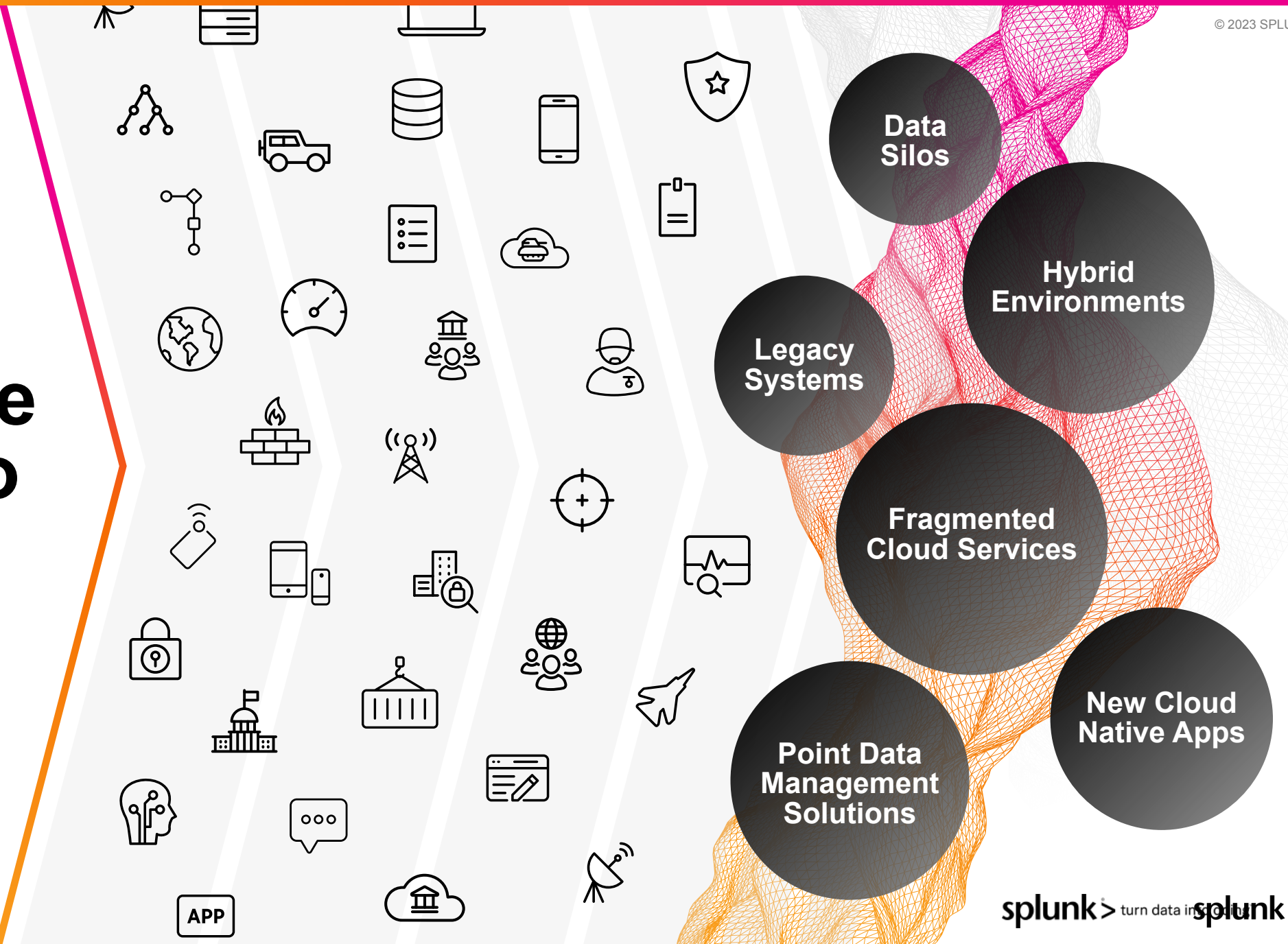
splunk>

Data is your Competitive Advantage

Data isn't just a record.
Data makes things happen.
Splunk makes it possible.

splunk> turn data into doing

Turning Real-time Data Into Action is Hard



The Power of Splunk

Delivering Unified Security and Observability

See

End-to-end visibility

No sampling or blind spots

Act

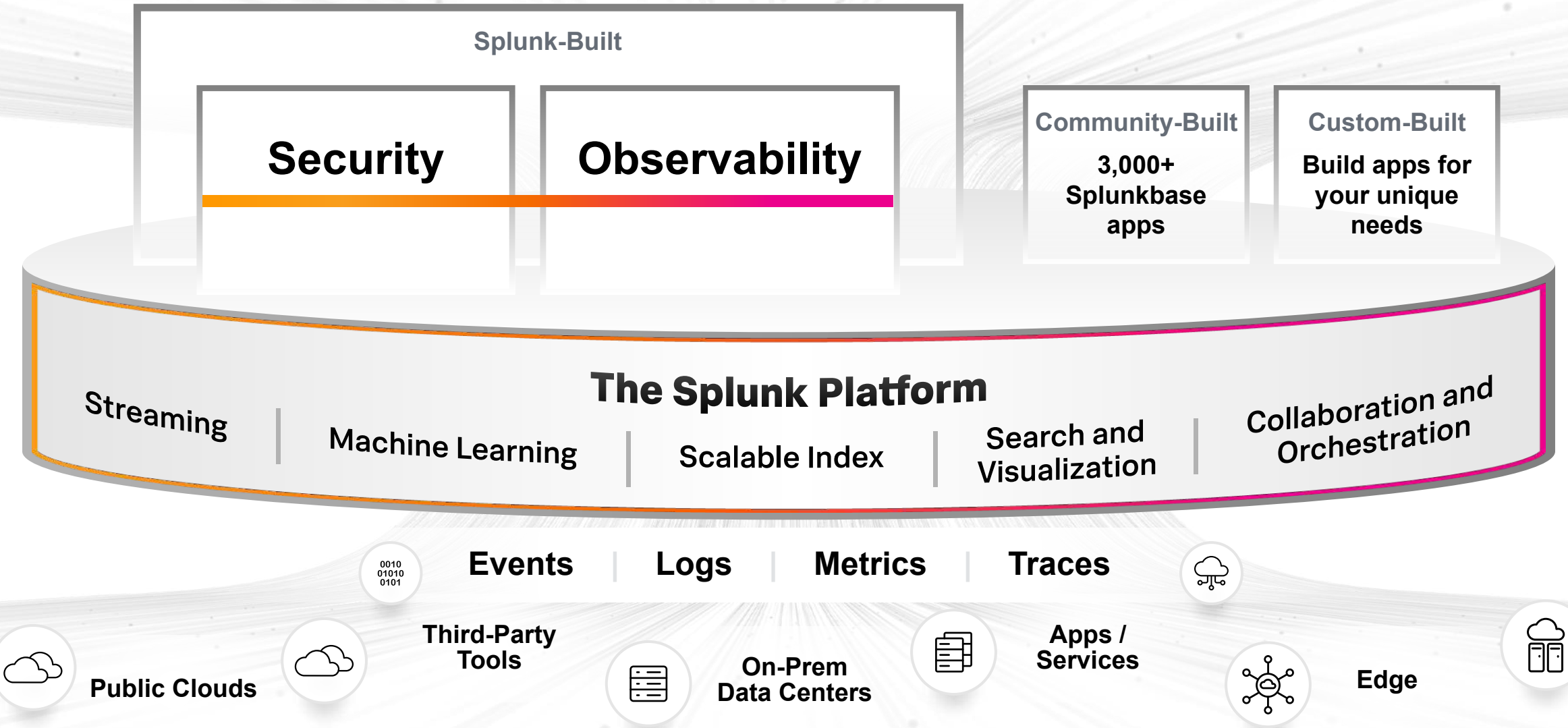
Investigate across massive data sets and take **action** fast

Extend

Extend the **platform** to use data to solve problems across the business

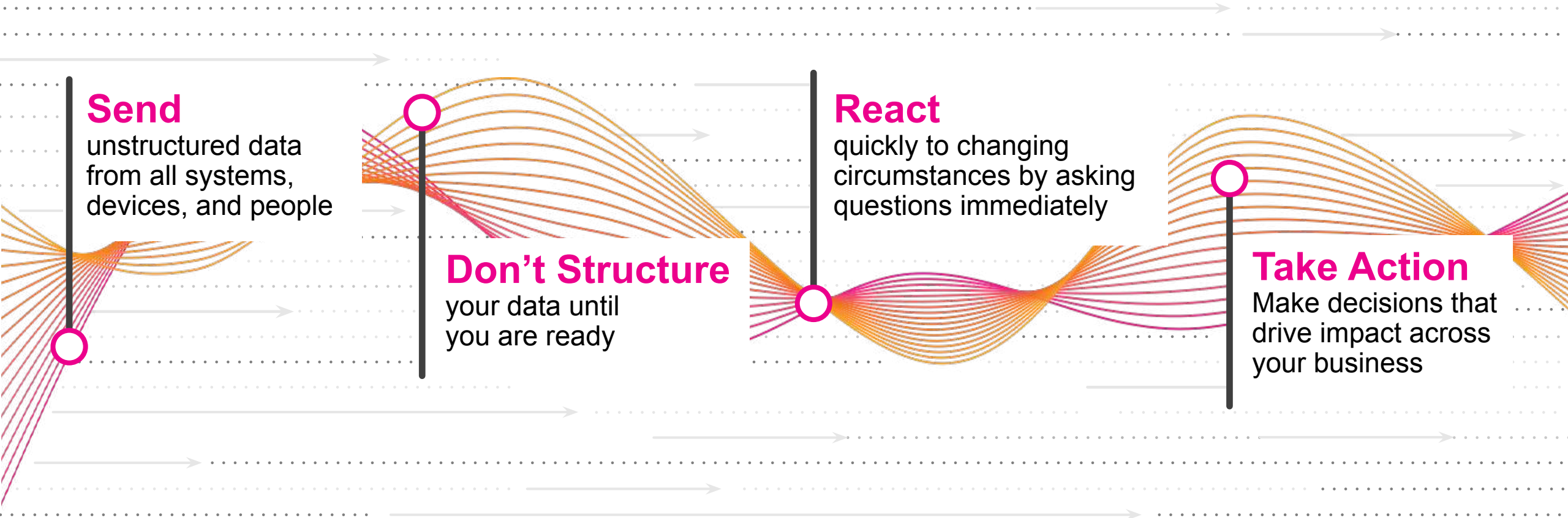
splunk>

The Unified Security and Observability Platform



Spot Risk and Opportunity in Your Complex Data

We solved one of the biggest challenges in data with our investigative approach



splunk>

Splunk as a Service

Fastest time to value | Minimum Infrastructure | Maximum Value

3 Simple Steps

1. Onboard data
2. Onboard users
3. Get value from your data

splunk>cloud™

- **Fastest time to value**
- **Software as a service** - AWS or GCP
- **Secure** - ISO 27001, SOC 2 Type II, PCI DSS, HIPAA, FedRAMP Moderate, DoD IL5, IRAP
- **Encryption-in-transit** - plus optional encryption-at-rest
- **Resilient infrastructure**
- **100% uptime guarantee**
- **24/7 NOC/SOC support team**

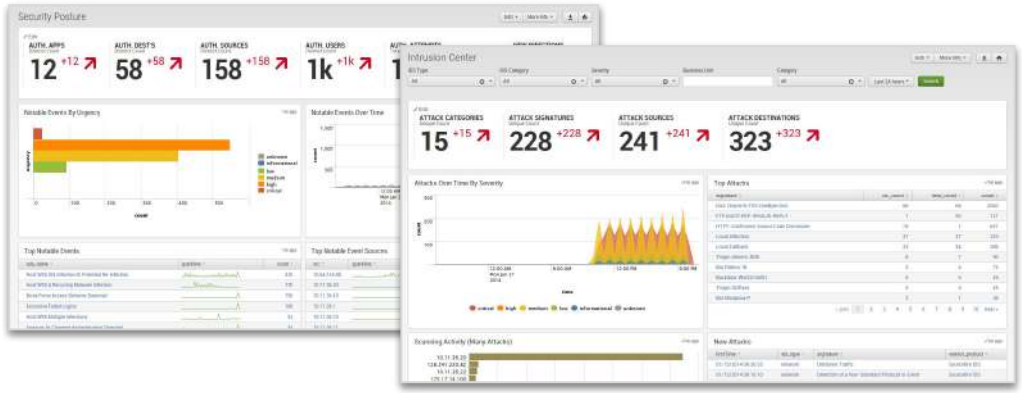
Flexible options for data collection and forwarding

Splunk Cloud Service Description: https://splk.it/SplunkCloudServDesc

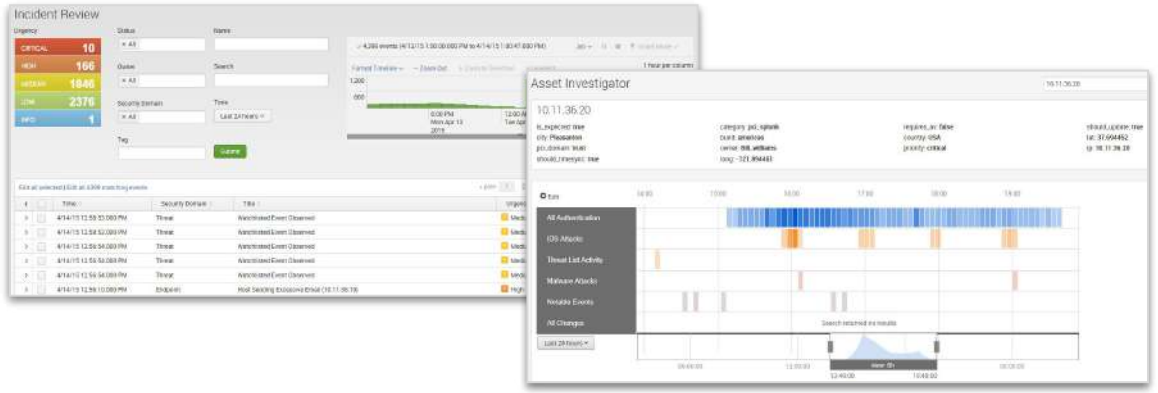
splunk>

Splunk Enterprise Security - The Next Level

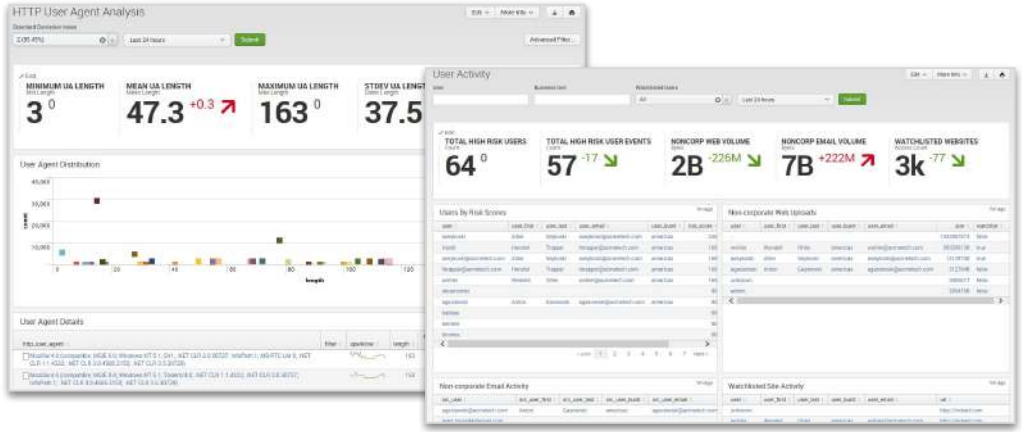
Pre-built searches, alerts, reports, dashboards, incident workflow, and threat intelligence feeds



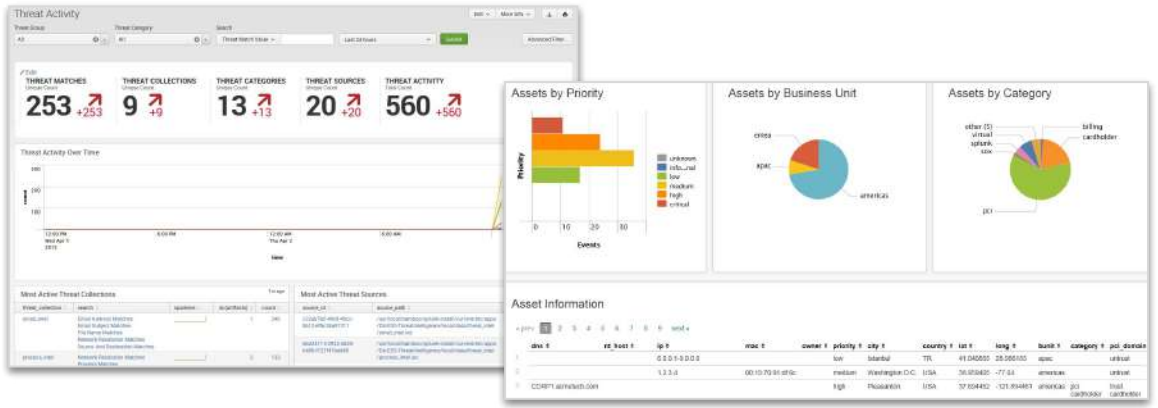
Alerts & Dashboards & Reports



Incident Investigations & Management



Statistical Outliers & Risk Scoring & User Activity



Threat Intel & Asset & Identity Integration

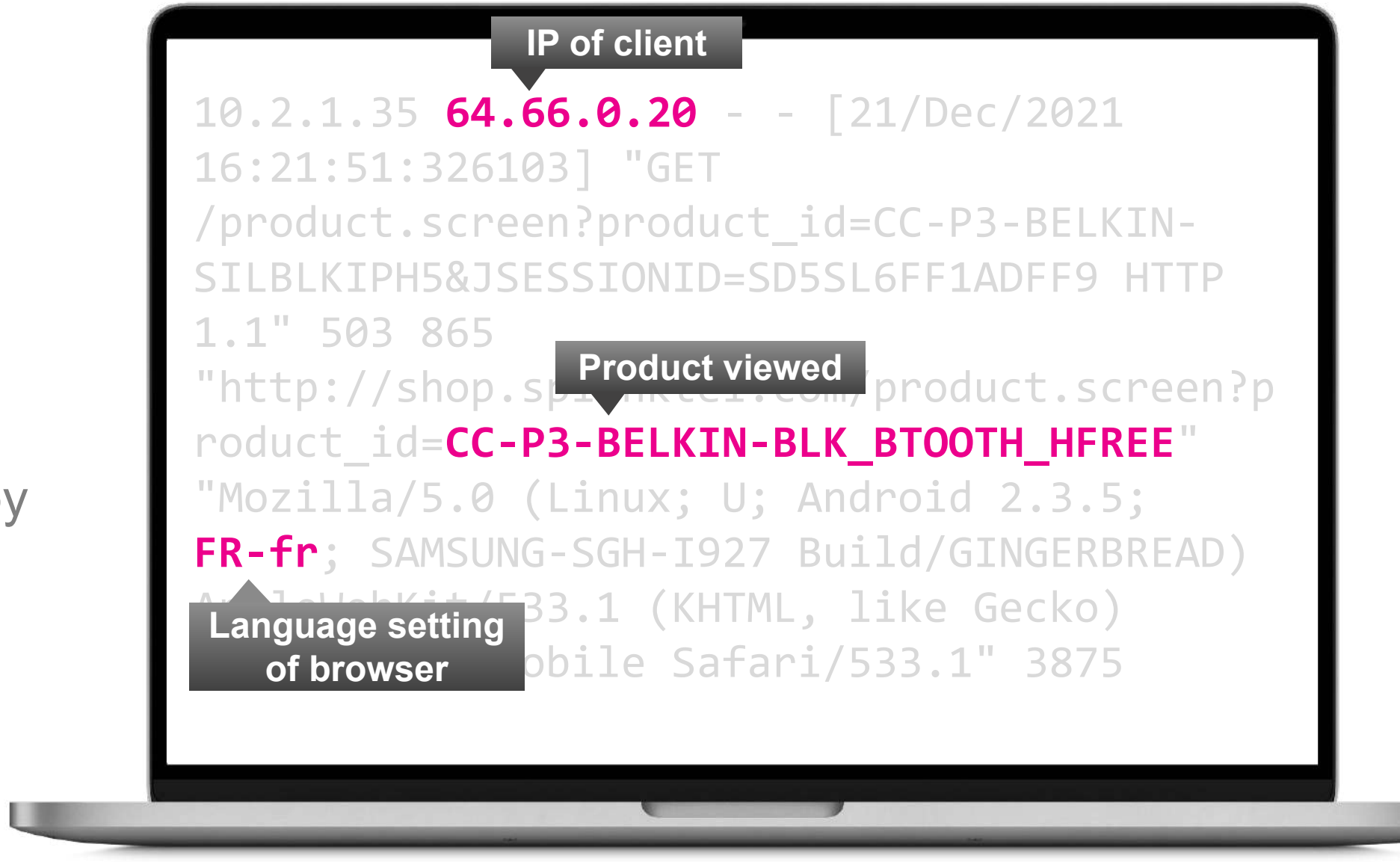
Learn more: https://www.splunk.com/en_us/products/enterprise-security.html

splunk>

Machine data is valuable not complex!

Marketing Use Case

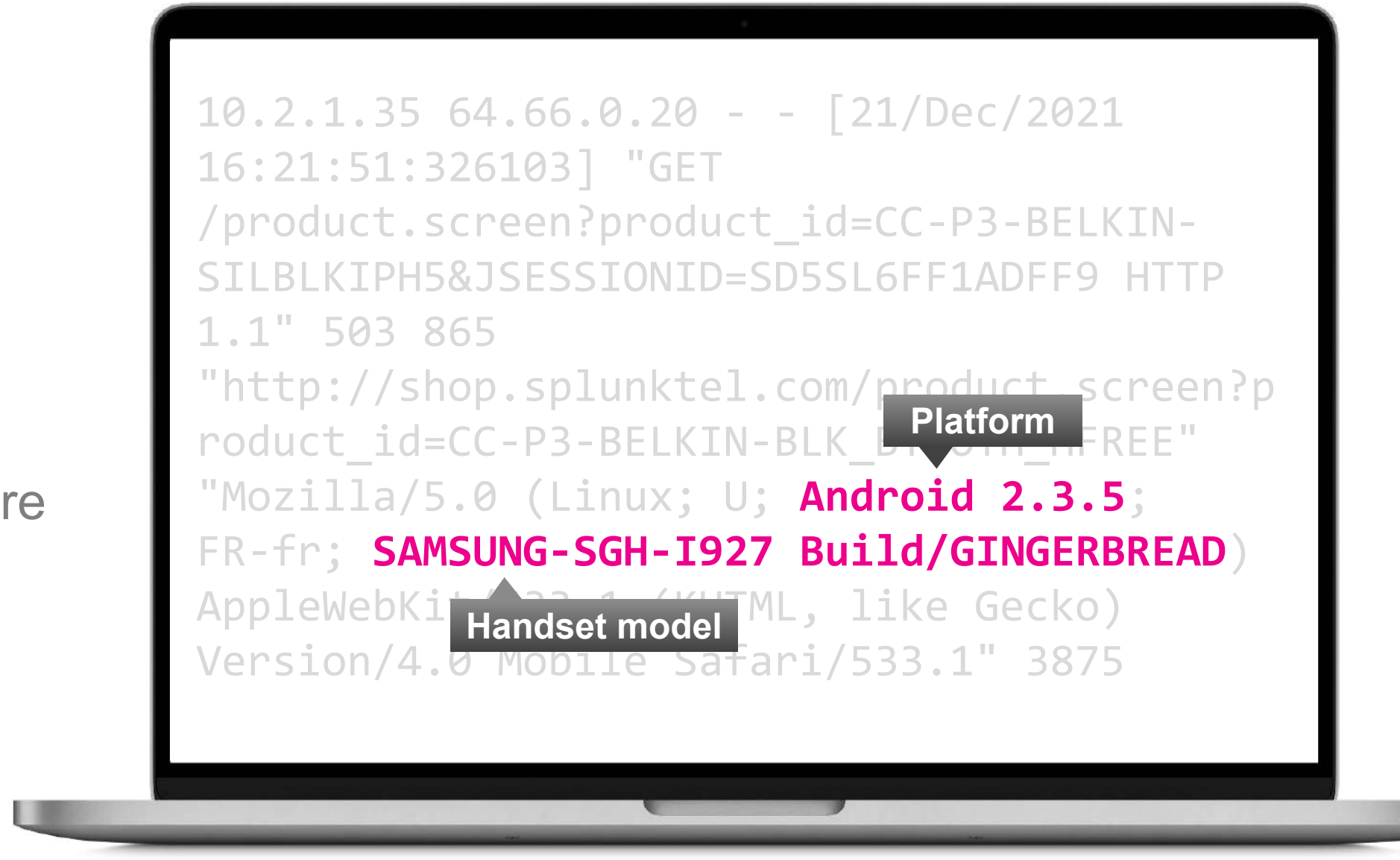
Show the top products viewed by language



splunk>

DevOps Use Case

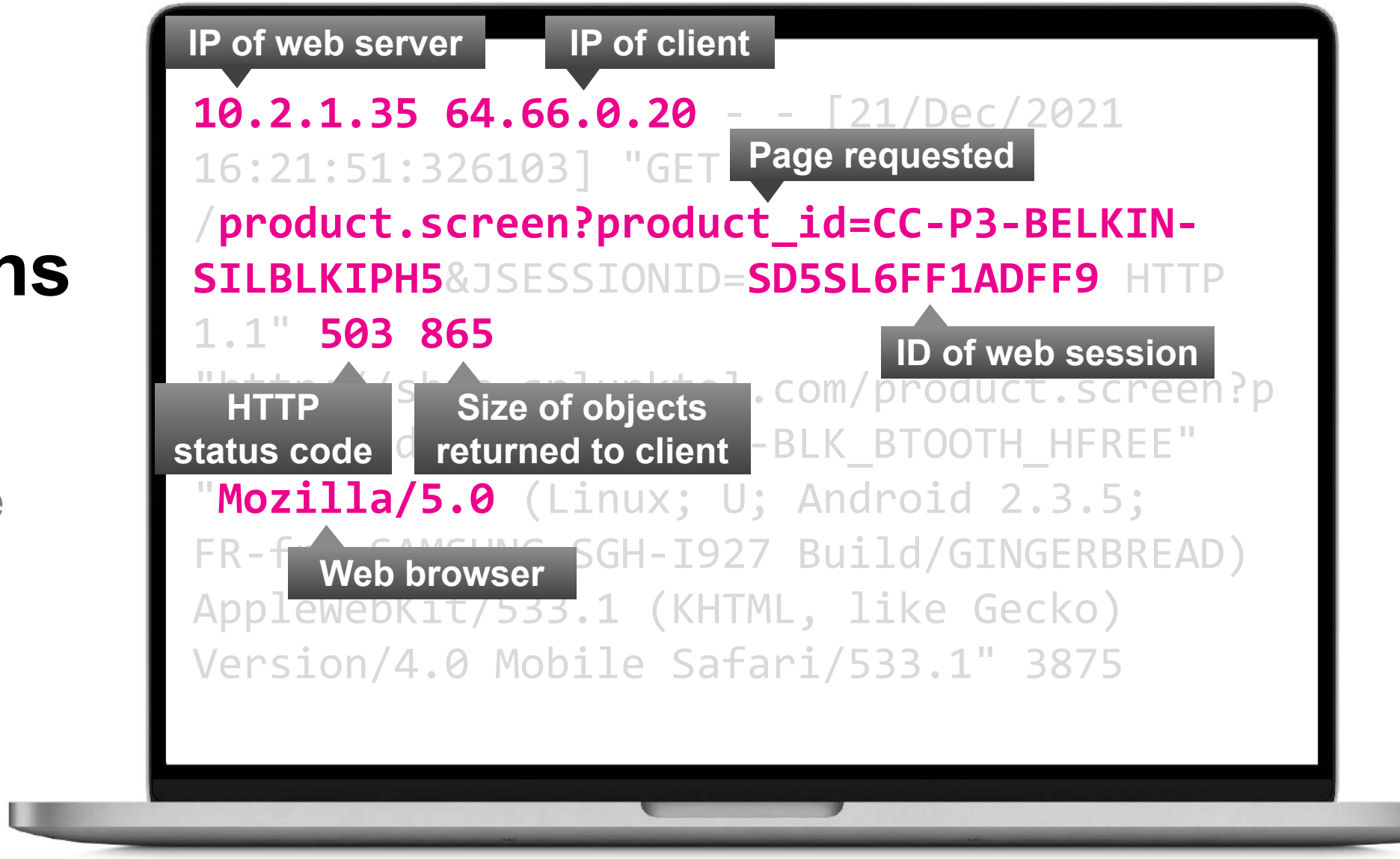
Which mobile handsets should I test the most before releasing my new app?



splunk>

IT Operations Use Case

Which web pages are generating the most errors?



splunk>

Splunk Connect for Syslog (SC4S) Overview

splunk> turn data into doing™

“Look, No ~~Hands~~ Splunk!”

Splunk Connect for Syslog (SC4S) is a wrapped version of syslog-ng in a container, preconfigured with parsers and meta for many data sources out of the box.

Key point: Reduced management surface - no Splunk add-ons or deployment server required!

Apps and Add-ons

- 3000+ free apps and add-ons available from https://splunkbase.splunk.com/
- Built either by Splunk, our technology partners or members of our user community
- Prebuilt packages that help to enhance and extend the Splunk platform
- Provide content and capabilities – such as reports, dashboards and integrations – for a specific technology, purpose or use case, with the flexibility to customise for your own needs

Apps

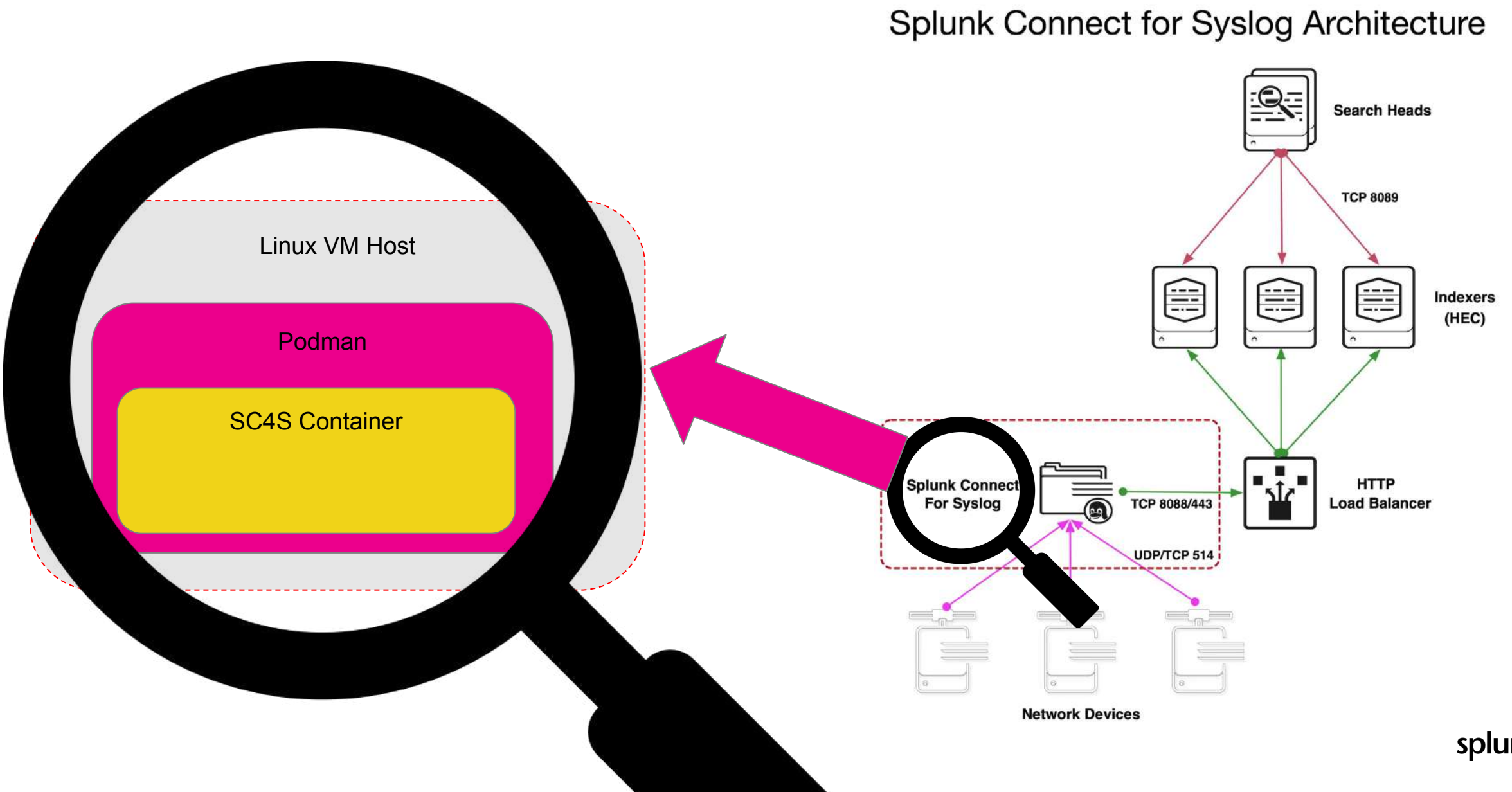
Content designed to bring fast time-to-value from your data in Splunk, including pre-built dashboards, reports, alerts, visualisations and workflows

Add-ons

Provide specific capabilities to Splunk, such as getting data in, mapping data, or providing saved searches and macros

splunk>

Architecture Example



splunk>

Syslog Data Collection at Scale

In concert with attempts to centralize syslog, many admins will co-locate several syslog-ng servers for horizontal scale, and load balance to them with a front-side load balancer. For many reasons (that go beyond this short discussion) this is not a best practice.

Briefly, the attempt to load balance for scale (and HA – see below) will actually cause more data loss.
This is due to normal device operations and higher buffer loss than would be the case if a simple, robust single server (or shared-IP cluster) were used.

Front-side load balancing will also cause inadequate data distribution on the upstream side, leading to data unevenness on the Splunk indexers.

HA Considerations and Challenges

- In addition to scale, many opt to load balance for **high availability**. While a sound approach for stateful, application-level protocols such as http, it **does not work well for stateless, unacknowledged syslog traffic.**
- Again, in the attempt to design for HA, more data ends up being lost vs. **more simple designs such as vMotioned VMs.**
- With syslog, always remember that the protocol itself is lossy, and there will be data loss (think CD-quality (lossless) vs. MP3). Syslog data collection can be made, at best, “Mostly Available”.

UDP vs. TCP

- Paradoxically, UDP for syslog actually ends up being a better choice for resilience for syslog. For an excellent discussion on this topic (as well as the “myth” of load balancers for HA), see Performant AND Reliable Syslog: UDP is best.

Learn more: https://splunk-connect-for-syslog.readthedocs.io/en/latest/architecture/

splunk>

SC4S Architectural Considerations

There are some key architectural considerations and recommendations that will yield extremely performant and reliable syslog data collection while minimizing the “over-engineering” that is common in many syslog data collection designs.

These recommendations are not specific to Splunk Connect for Syslog, but rather stem from the syslog protocol itself – and its age.

- **The Syslog Protocol**
  - The syslog protocol was **designed in the mid 1980s** to offer very high-speed, network-based logging for network and security devices that were (especially at the time) starved for CPU and I/O resources.
  - For this reason, the protocol was designed for speed and efficiency at the expense of resilience/reliability. UDP was chosen due to its ability to **“send and forget”** the events over the network without regard (or acknowledgment) of receipt.
  - In later years, TCP was added as a transport, as well as TLS/SSL. In spite of these additions, **UDP still retains favor** as a syslog transport for most data centers, and for the same reasons as originally designed.
  - Because of these tradeoffs selected by the original designers (and retained to this day), **traditional methods used to provide scale and resiliency do not necessarily transfer to the syslog world.**

- **Collector Location**
  - Due to syslog being a “send and forget” protocol, it does not perform well when routed through substantial (and especially WAN) network infrastructure. This includes front-side load balancers. **The most reliable way to collect syslog traffic is to provide for edge collection** (not centralized). **Resist the urge to centrally locate any syslog server (sc4s included)** and expect the UDP and (stateless) TCP traffic to “make it”. Data loss will undoubtedly occur.

Learn more: https://splunk-connect-for-syslog.readthedocs.io/en/latest/architecture/

splunk>

What's a Podman, Anyway?

Podman is a **daemonless, open source, Linux native tool** designed to make it easy to find, run, build, share and deploy applications using **Open Containers Initiative (OCI)** containers and container Images. Podman provides a command line interface (CLI) familiar to anyone who has used the Docker Container Engine.

Most users can simply alias Docker to Podman (alias docker=podman) without any problems. Similar to other common Container Engines (Docker, CRI-O, containerd), Podman relies on an OCI compliant Container Runtime (runc, crun, runv, etc) to interface with the operating system and create the running containers. This makes the running containers created by Podman nearly indistinguishable from those created by any other common container engine.

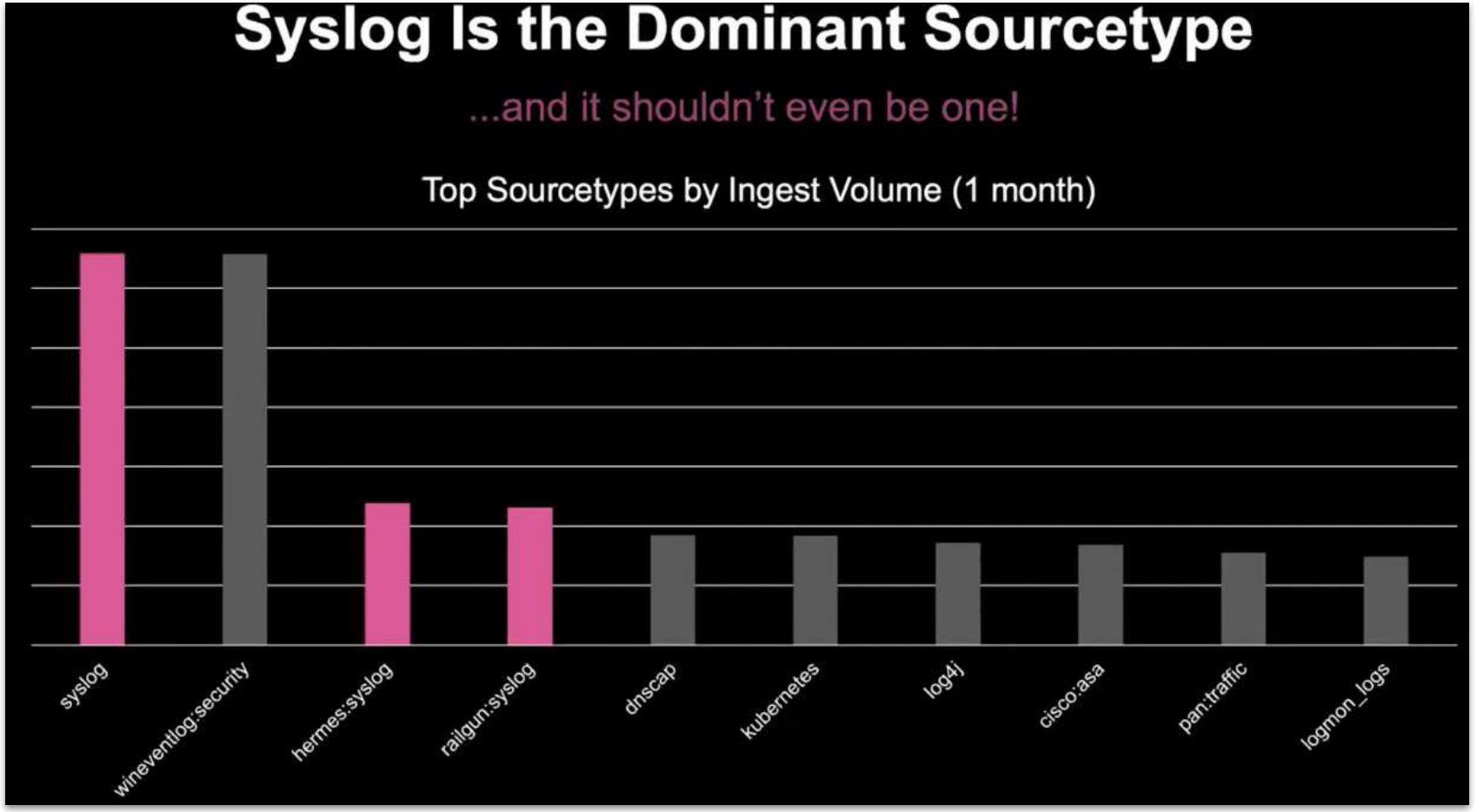
Containers under the control of Podman can either be run by root or by a non-privileged user. Podman manages the entire container ecosystem which includes pods, containers, container images, and container volumes using the libpod library. Podman specializes in all of the commands and functions that help you to maintain and modify OCI container images, such as pulling and tagging. It allows you to create, run, and maintain those containers and container images in a production environment.

SC4S:
Why Use It?

splunk>

Why Did Splunk Build SC4S?

"Syslog" is not a sourcetype! I'll repeat that: **"Syslog" is not a sourcetype!**



splunk>

Why Did Splunk Build SC4S?

It was clear from our research that the “status quo” of syslog ingest into Splunk could not continue due to complexity and scale issues, and customer feedback showed continued challenges with the problem.

Thus Splunk Connect for Syslog was born, and as the formal product gelled, two overarching questions were continually asked during the development process:

- **What solution would significantly improve on current practices for syslog?**
- **How can we meet the needs of 80+% of our customers out of the box?**

Specifically, we have seen the following key indicators of a need for a solution:

- **Syslog has NOT been “deprecated” or become “legacy”** for most device vendors**.** In fact, a major security device vendor has *returned* to syslog after a long departure.
- **Syslog data volume has increased significantly** due to enterprise growth and device throughput, requiring formalization of the techniques explored in the previous blog to cope with the volume.
- **Syslog continues to be a majority data type**, by volume, for nearly 100% of Splunk’s customer base.
- **Customer demand for a turnkey, scalable solution** to the problem has increased significantly.

Learn more: https://www.splunk.com/en_us/blog/tips-and-tricks/splunk-connect-for-syslog-turnkey-and-scalable-syslog-gdi.html

splunk>

SC4S Design Goals

We felt that if we brought the benefits below to the Splunk user community, then the solution would be viable.

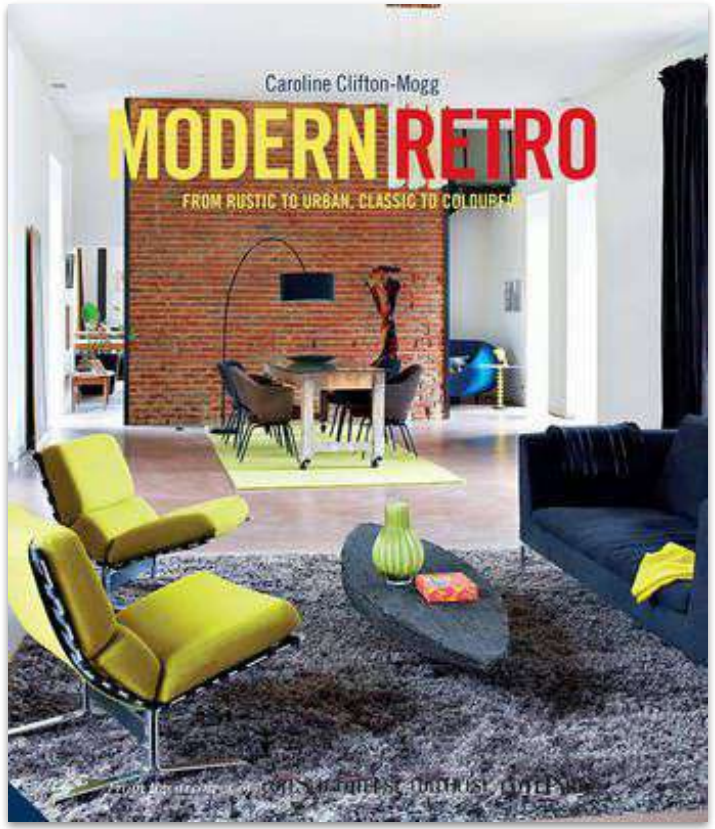
These benefits include:

- ✓ **Lowering the burden, both on customers and Splunkers, of getting syslog data into the Splunk platform**
- ✓ **Providing a consistent, documented, and repeatable syslog collection infrastructure**
- ✓ **Providing turnkey data ingestion for 15 top sourcetypes at first release**
- ✓ **Improving the “data hygiene” of incoming syslog data with proper sourcetyping and enriched metadata**
- ✓ **Reducing Splunk overhead in processing syslog data**
- ✓ **Significantly enhancing scale and data distribution**

What's in it For Me?

Why move from what you have, or start off with SC4S?

- ✓ **Modern platforms to handle an old skool protocol**
- ✓ **Reproducible**
- ✓ **Scalable**
- ✓ **Predictable**
- ✓ **Fits into modern CI/CD tool chains**
- ✓ **No Splunk TAs to manage/update**



splunk>

Today's Scenario: Buttercup Cyber

Your Company

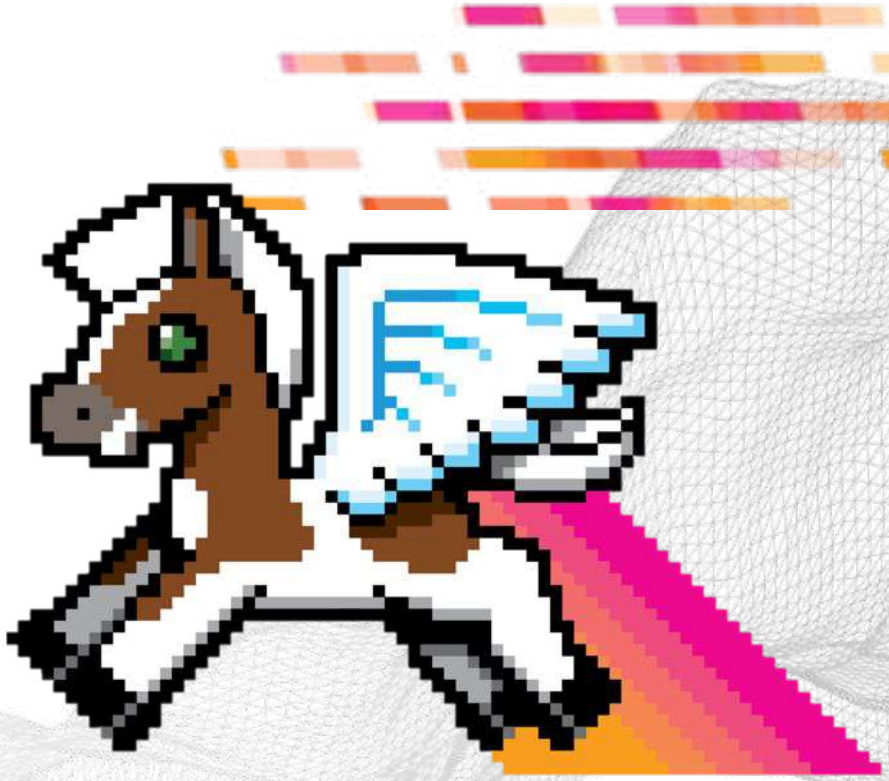
- Buttercup Cyber is a large MSSP looking to rapidly deploy managed SIEM services across its customers

Your Role

- You are one of the chosen few: An MSSP Engineer
- Your responsibility is to deploy SC4S and the Splunk INFOSEC app to a new customer - initially for cisco events

Thanks to some amazing engineering you already have

- Splunk, with the required add-ons and apps installed
- SC4S ready as a systemd service, managed by podman



splunk>

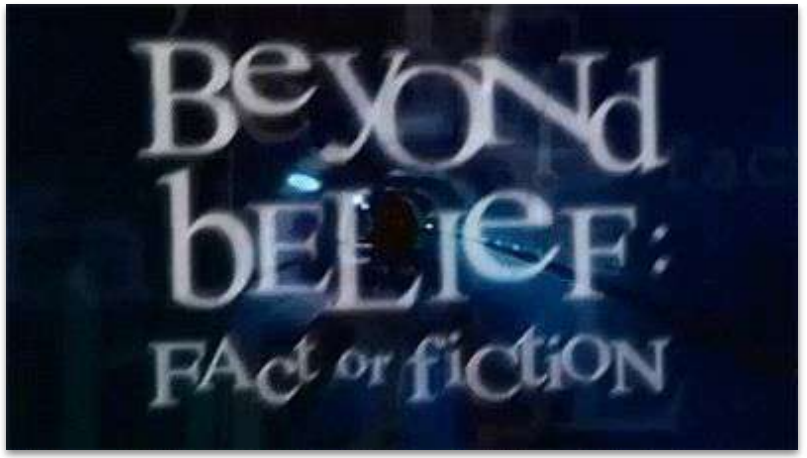
Suspend Belief

Datagens

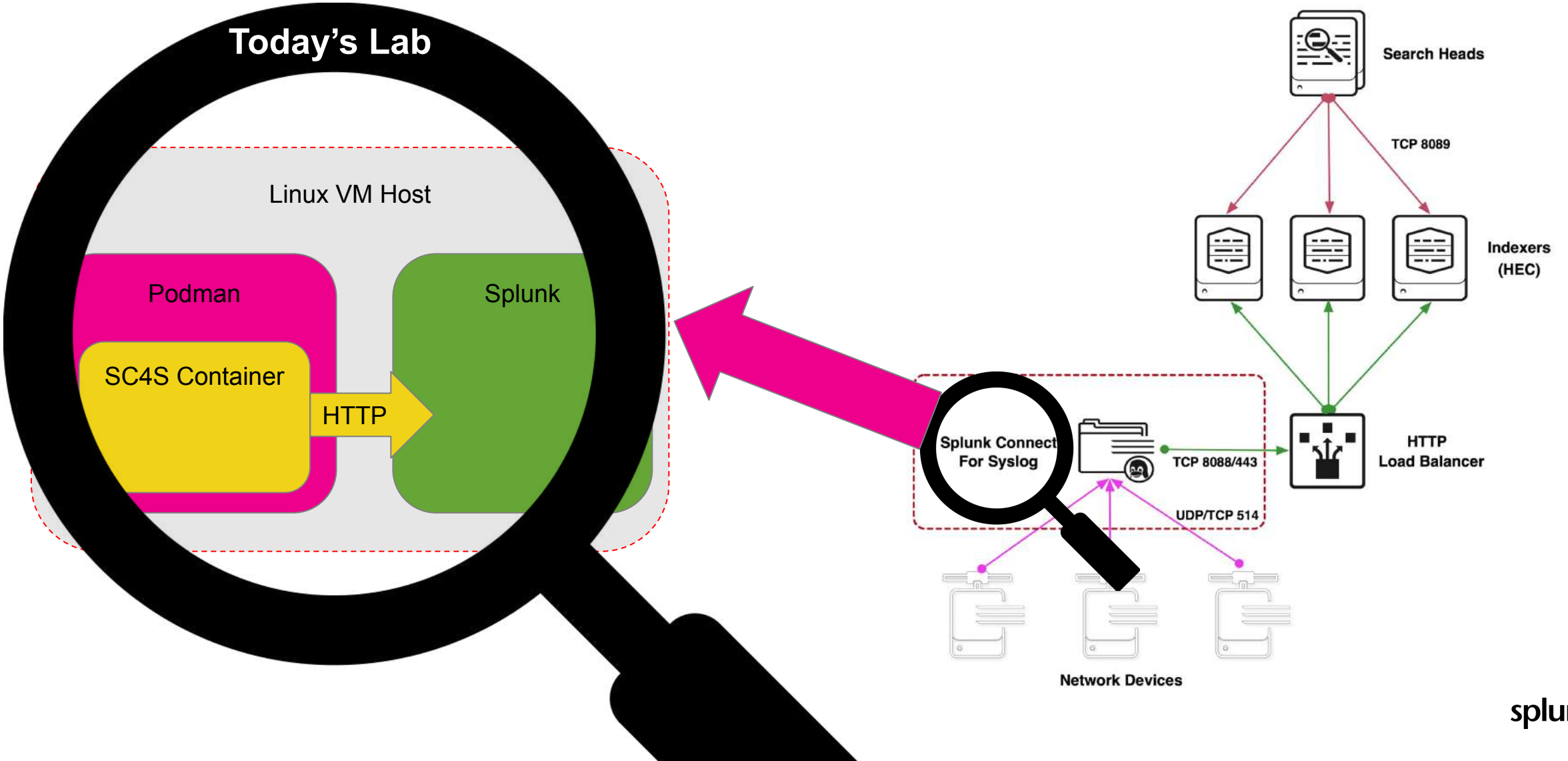
- This lab is based on data generators - with that there will be quirks, bumps and things that just don't make real world sense; Please look past this and focus on the overall goal!
- If you see something quirky - it’s not a problem with Splunk, it's the datagen

Where's my SSH?

- Interaction with SC4S and systemd/podman will be via the Splunk UI
  - This is not its natural flow but due to most corporate internet access policies it’s a must
  - We don't want to spend half the workshop teaching how to use PuTTY!



Todays Lab Architecture



splunk>

SC4S systemd Service

Set to the latest image of version 1 every time

podman pull (pull latest image as set above)

Note the **env_file** location

Note the name of the service

Note the location of the **sc4s.service** file

```
echo "
## Created with JB Splunk Install script by magic
[Unit]
Description=SC4S Container
Wants=NetworkManager.service network-online.target
After=NetworkManager.service network-online.target
[Install]
WantedBy=multi-user.target
[Service]
Environment=\"SC4S_IMAGE=ghcr.io/splunk/splunk-connect-for-syslog/container:1\"
# Required mount point for syslog-ng persist data (including disk buffer)
Environment=\"SC4S_PERSIST_MOUNT=splunk-sc4s-var:/var/lib/syslog-ng\"
# Optional mount point for local overrides and configurations; see notes in docs
Environment=\"SC4S_LOCAL_MOUNT=/opt/splunk/sc4s/local:/etc/syslog-ng/conf.d/local:z\"
# Optional mount point for local disk archive (EWMM output) files
Environment=\"SC4S_ARCHIVE_MOUNT=/opt/splunk/sc4s/archive:/var/lib/syslog-ng/archive:z\"
# Uncomment the following line if custom TLS certs are provided
Environment=\"SC4S_TLS_MOUNT=/opt/splunk/sc4s/tls:/etc/syslog-ng/tls:z\"
TimeoutStartSec=0
ExecStartPre=/usr/bin/podman pull \$SC4S_IMAGE
ExecStartPre=/usr/bin/bash -c \"/usr/bin/systemctl set-environment SC4SHOST=$(hostname -s)\"
ExecStart=/usr/bin/podman run \\
        -e \"SC4S_CONTAINER_HOST=\${SC4SHOST}\" \\
        -v \$SC4S_PERSIST_MOUNT \\
        -v \$SC4S_LOCAL_MOUNT \\
        -v \$SC4S_ARCHIVE_MOUNT \\
        -v \$SC4S_TLS_MOUNT \\
        --env-file=/opt/splunk/sc4s/env_file \\
        --network host \\
        --name SC4S \\
        --rm \$SC4S_IMAGE
Restart=on-abnormal
" > /lib/systemd/system/sc4s.service
```

Learn more: https://splunk-connect-for-syslog.readthedocs.io/en/latest/gettingstarted/podman-systemd-general/

splunk>

The Basics

The env_file ….. file…..

There are only a few items that SC4S needs to start up out of the box:

- HEC URL (either a list of endpoints or load balancer VIP)
- HEC Token

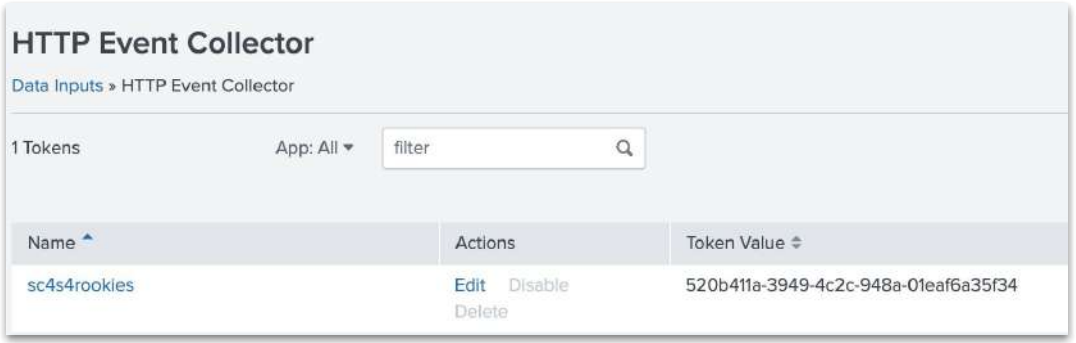
Best practice: Use FQDN and TLS verification, also use a HTTP load balancer in larger environments.

```
# Output config
SC4S_DEST_SPLUNK_HEC_DEFAULT_URL=https://idx1.example.com:8088,https://idx2.example.com:8088
SC4S_DEST_SPLUNK_HEC_DEFAULT_TOKEN=e342a3c6-5d59-40ed-9ffe-9f2b75786802

#Uncomment the following line if using untrusted SSL certificates
SC4S_DEST_SPLUNK_HEC_TLS_VERIFY=yes
```

splunk>

What the HEC?

- SC4S will send data to Splunk via the **HTTP Event Collector (HEC)**
  - We have already created one for you today
  - **NOTE:** When filling out the form for the HEC token, it is recommended that the “Selected Indexes” pane be left blank and that a **lastChanceIndex** be created so that all data received by SC4S will land somewhere in Splunk



- HEC is a fast and efficient way to send data to Splunk Enterprise and Splunk Cloud Platform
  - Notably, HEC enables you to send data over HTTP (or HTTPS) directly to Splunk Enterprise or Splunk Cloud Platform from your application. HEC was created with application developers in mind, so that all it takes is a few lines of code added to an app for the app to send data.
  - HEC is token-based, so you never need to hard-code your Splunk Enterprise or Splunk Cloud Platform credentials in your app or supporting files.
  - HEC provides a new way for developers to send application logging and metrics directly to Splunk Cloud Platform and Splunk Enterprise via HTTP in a highly efficient and secure manner.

Learn more: https://docs.splunk.com/Documentation/SplunkCloud/latest/Data/UsetheHTTPEventCollector

splunk>

Splunk Setup

Indexes

The following default indexes are used by SC4S, but can be customised (we have already created them for you today):

- email
- epav
- netauth
- netdlp
- netdns
- netfw
- netids
- netops
- netwaf
- netproxy
- netipam
- oswinsec
- osnix
- em_metrics

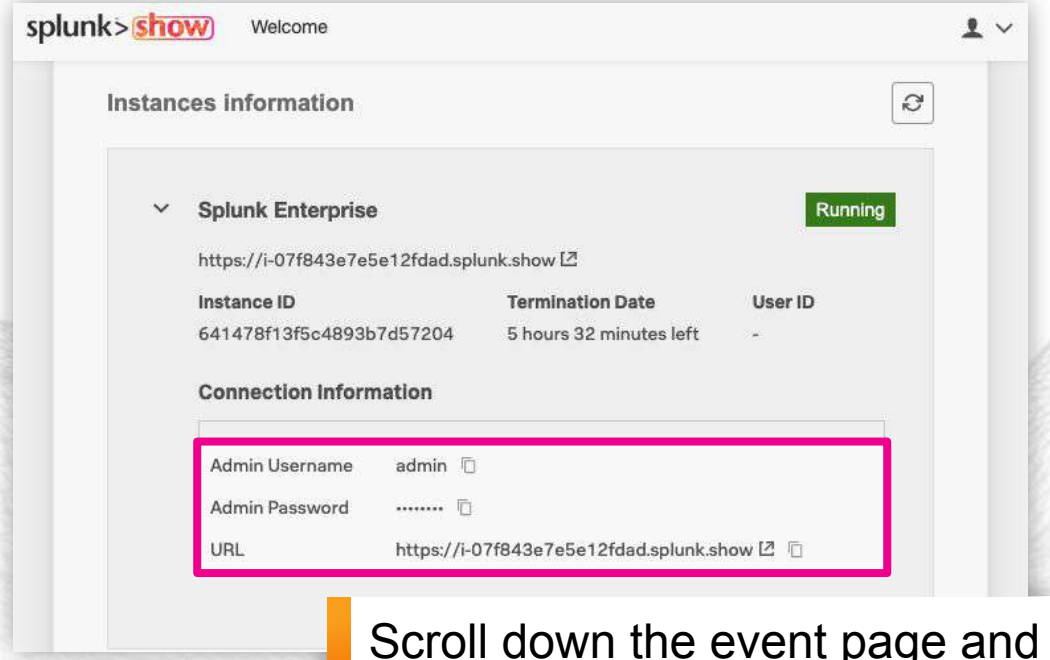
Optional opt-in for SC4S operational metrics; ensure this is created as a metrics index)

Learn more: https://splunk-connect-for-syslog.readthedocs.io/en/latest/gettingstarted/quickstart_guide/#splunk-setup

SC4S 101

Login to Splunk

Locate your instance URL and credentials in the Splunk Show event https://show.splunk.com



Scroll down the event page and expand the **Splunk Enterprise** section to view your login details

Log in to your Splunk instance

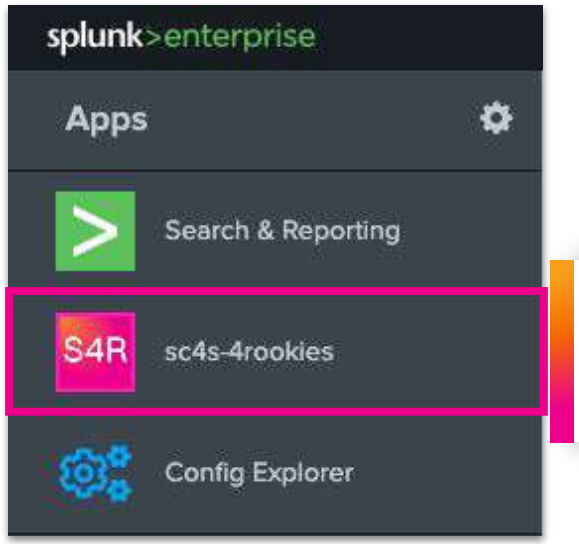
Login using the credentials from Splunk Show

splunk>

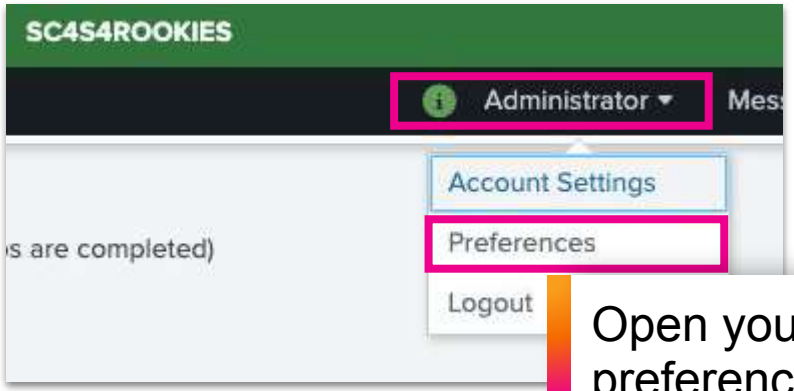
Lab 1: Setup Your Environment

splunk>

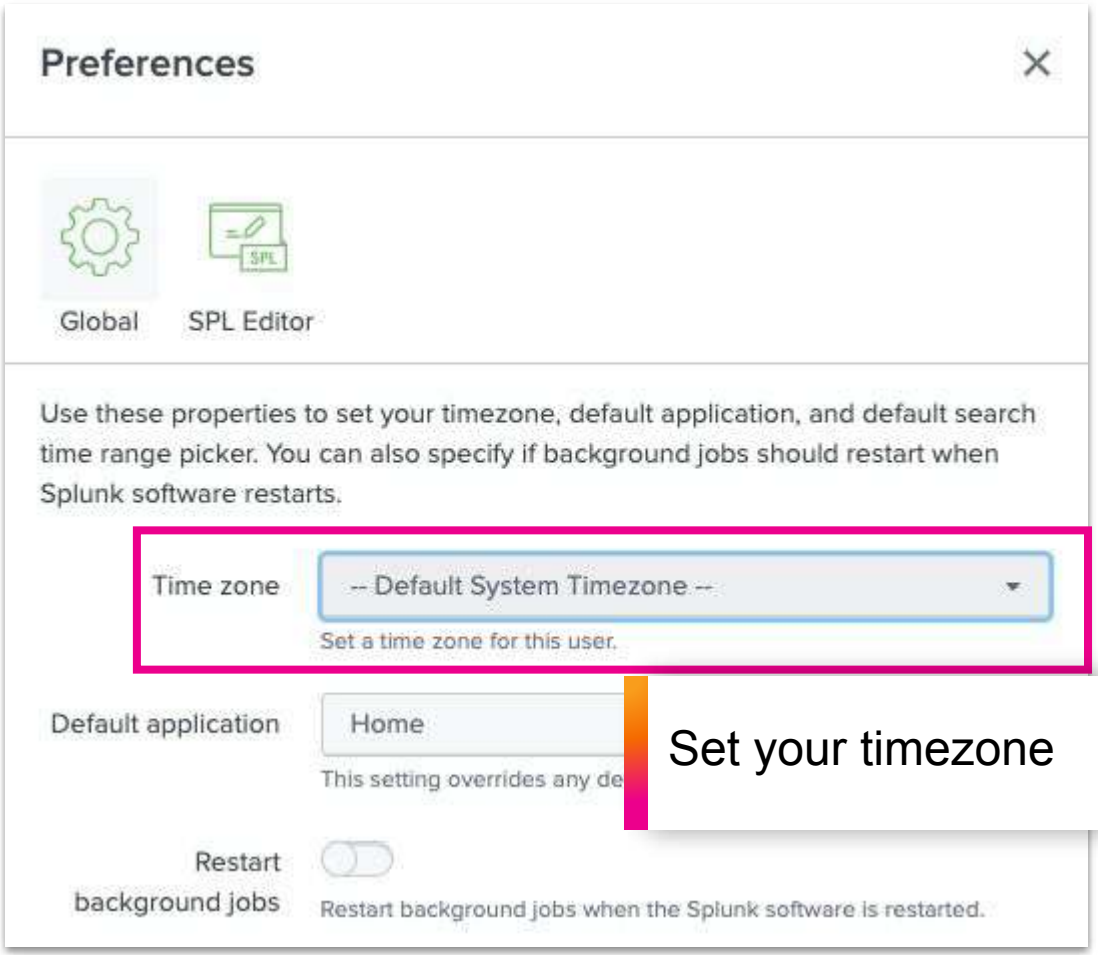
Login and Set Your Timezone



Click on the **sc4s-4rookies** app



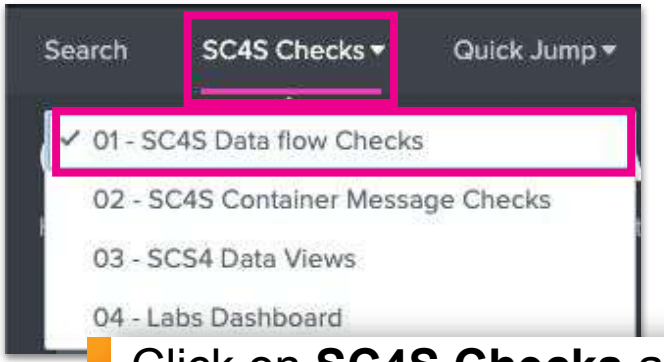
Open your user preferences



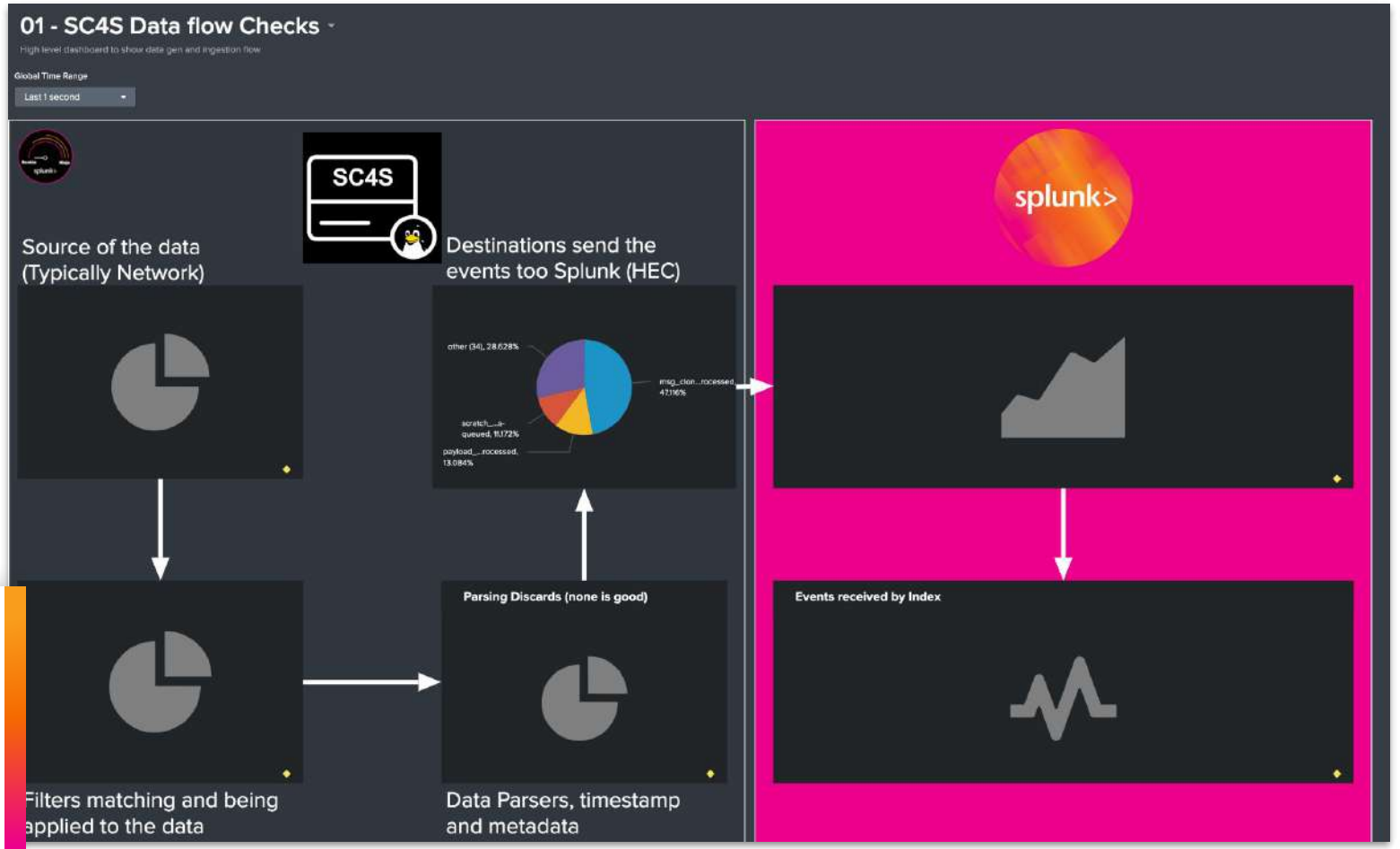
Set your timezone

splunk>

Check the Data Flow



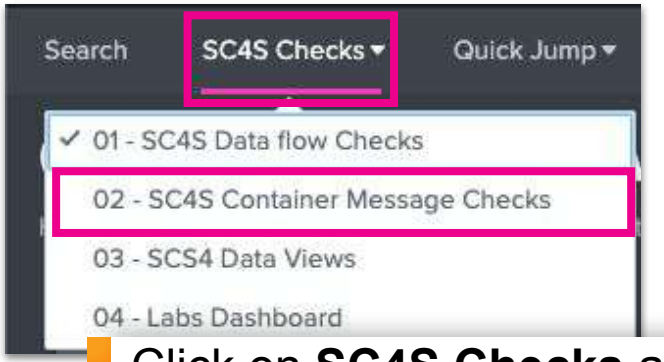
Click on **SC4S Checks** and then on **01 - SC4S Data flow Checks**



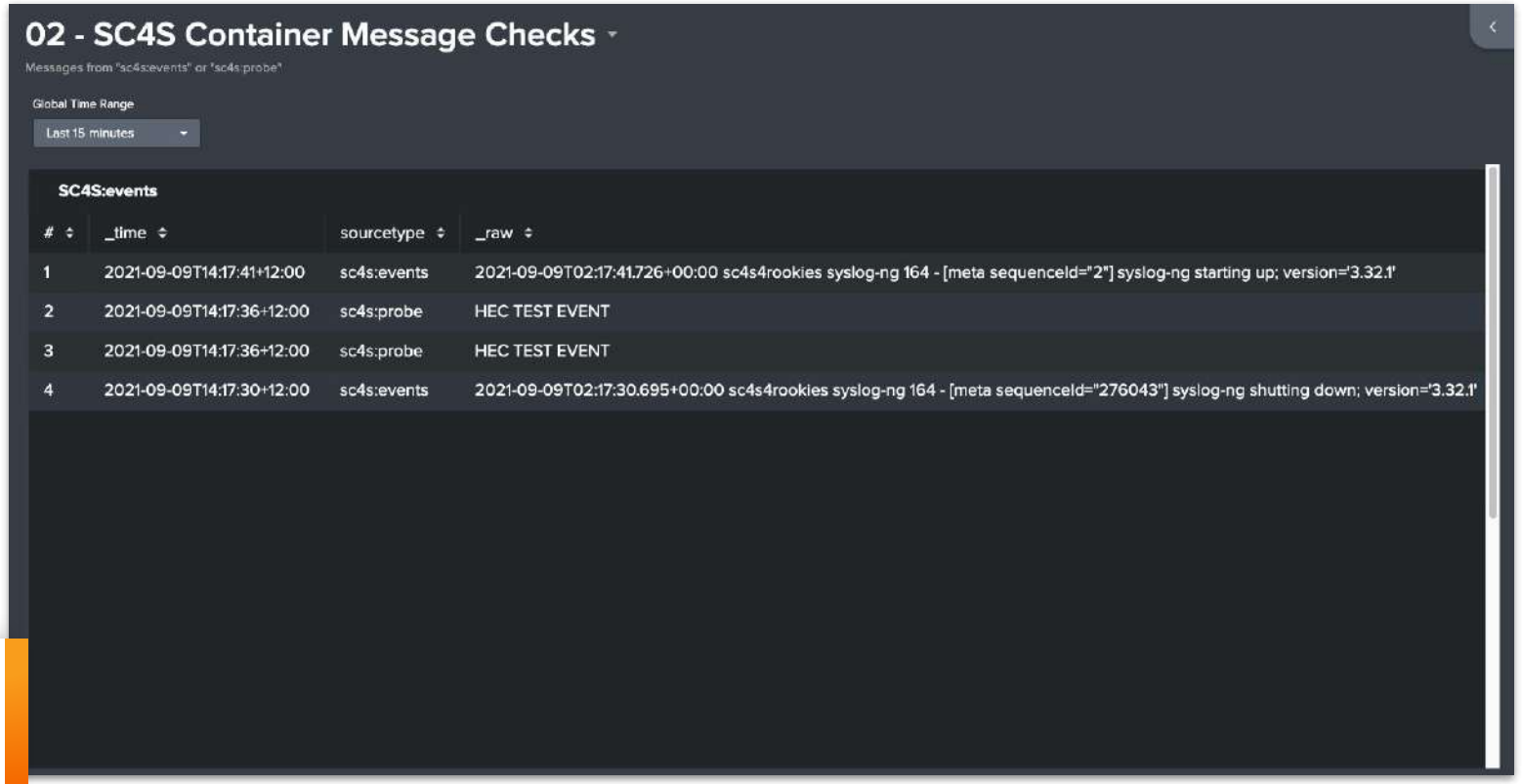
All inputs are turned off and the sc4s service is stopped after the initial image creation

You may see some historical data or blank graphs - this is normal

Check the Data Flow



Click on **SC4S Checks** and then on **02 - SC4S Container Message Checks**



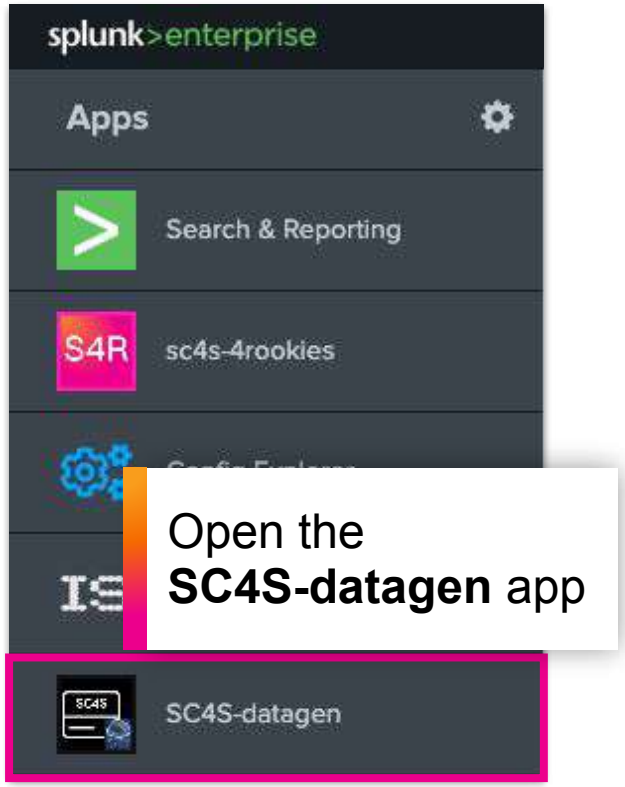
All inputs are turned off and the sc4s service is stopped after the initial image creation

You may see some historical data or blank tables - this is normal

splunk>

Restart the SC4S systemd Service

Restart the systemd service for podman and SC4S by enabling the input, waiting a few seconds, then disabling it again:



Open the **SC4S-datagen** app

1. Set the toggle to Enabled

2. Wait for **10 seconds…**

1. Set the toggle to Disabled

Note: In environments with CLI access you can also restart the service using the following command:

```
sudo systemctl restart sc4s
```

splunk>

Enable the Cisco Datagen

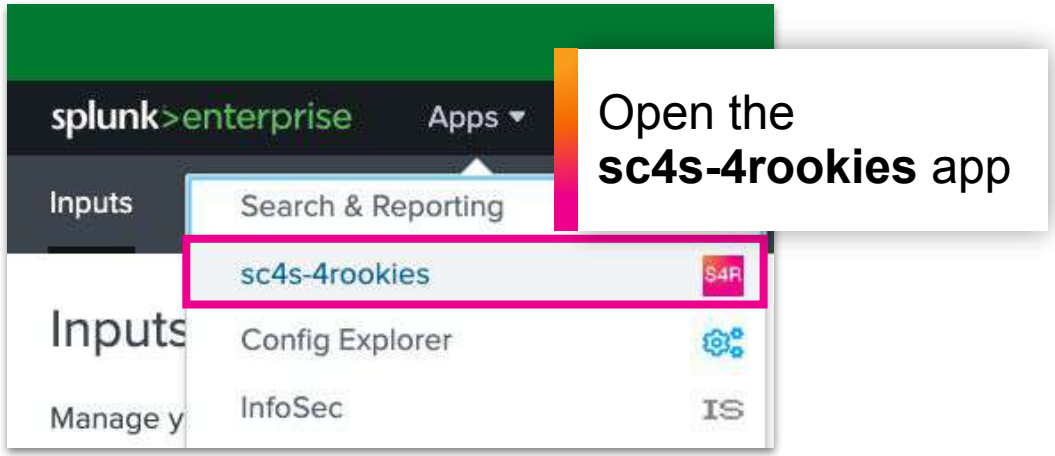
1. Set the toggle to Enabled

This input will start the cisco datagen - leave this one enabled

Now let's check the data flow dashboard again

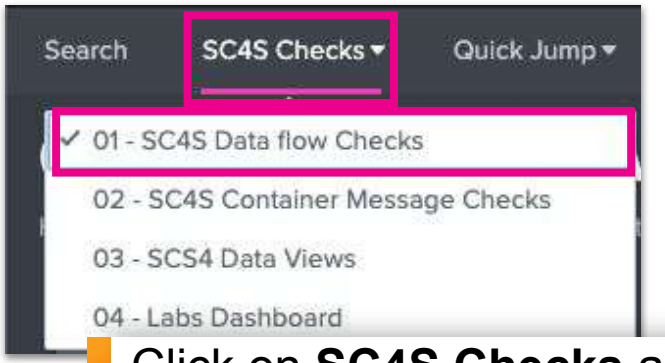
splunk>

Go Back to the sc4s-4rookies App

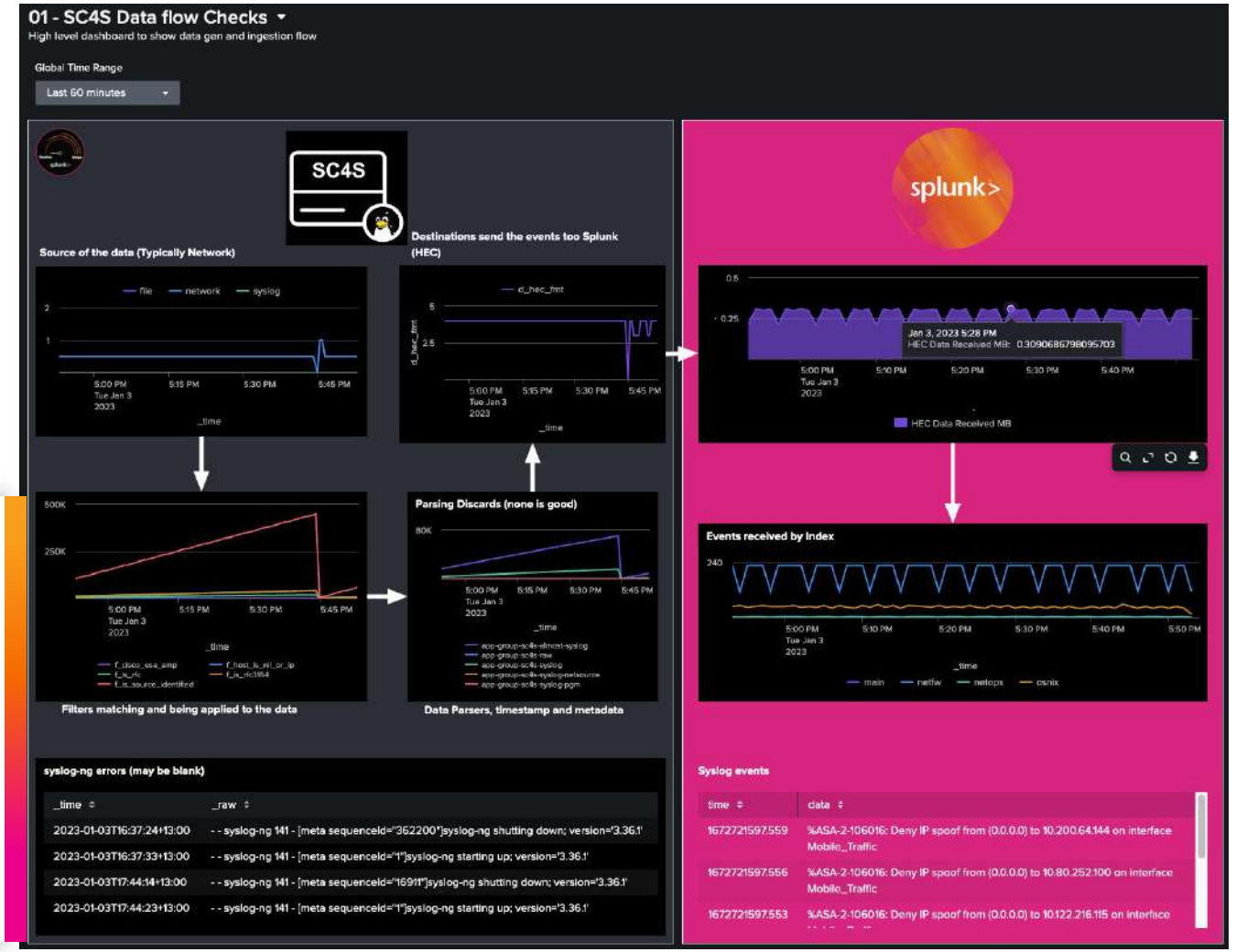


Open the **sc4s-4rookies** app

Check the Data Flow



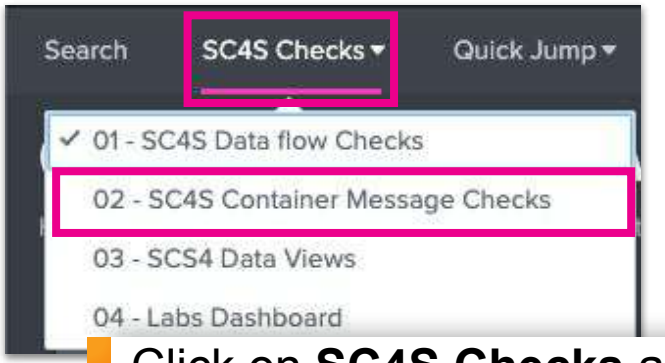
Click on **SC4S Checks** and then on **01 - SC4S Data flow Checks**



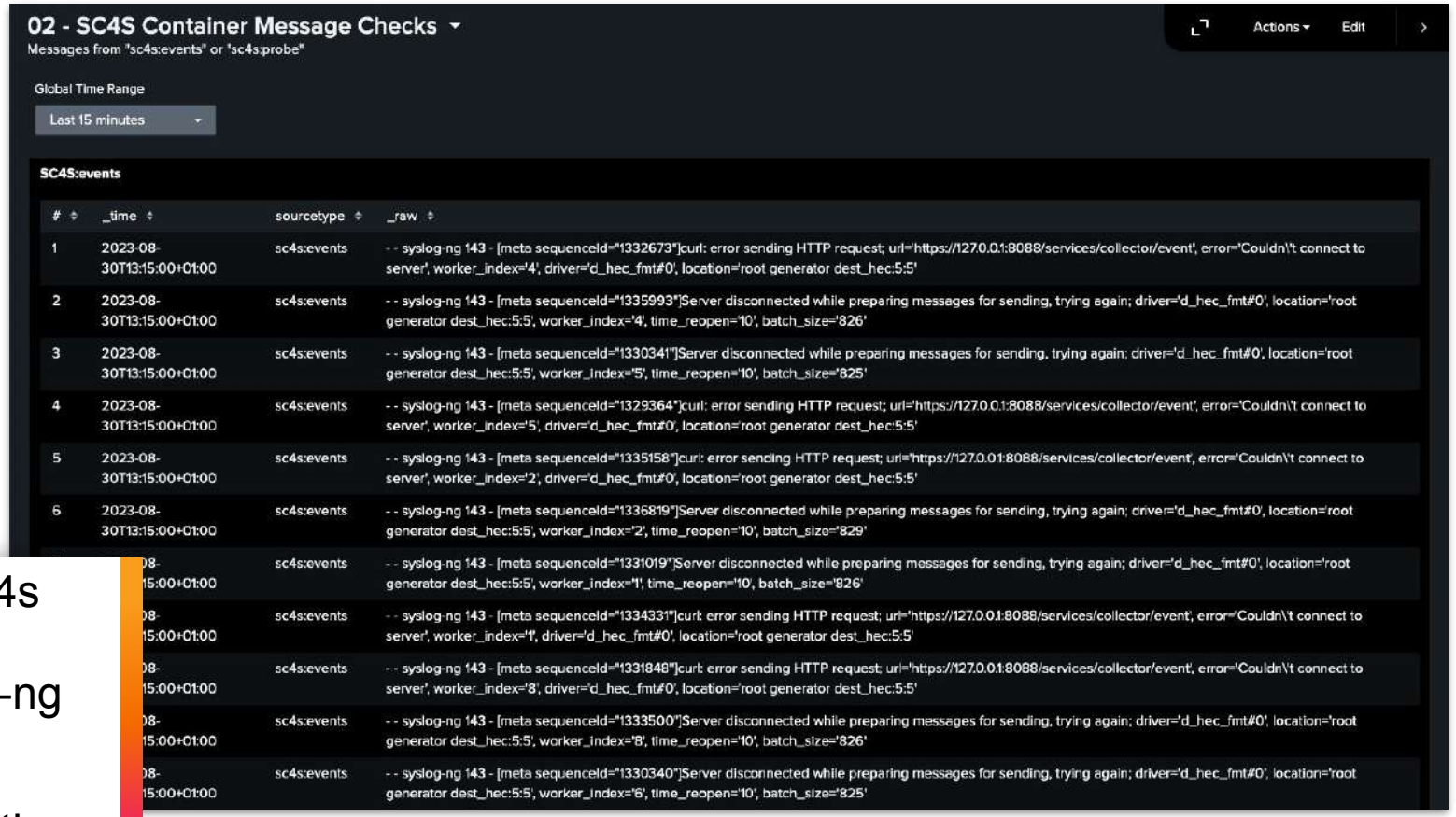
Now we have started the Cisco datagen you should see a steady flow of data start to populate your dashboard

The arrows indicate the flow of the events through your log paths and out to Splunk over HEC to the indexes

Check the Data Flow



Click on **SC4S Checks** and then on **02 - SC4S Container Message Checks**



Now we have restarted the sc4s service you should see some startup messages from syslog-ng from inside the container

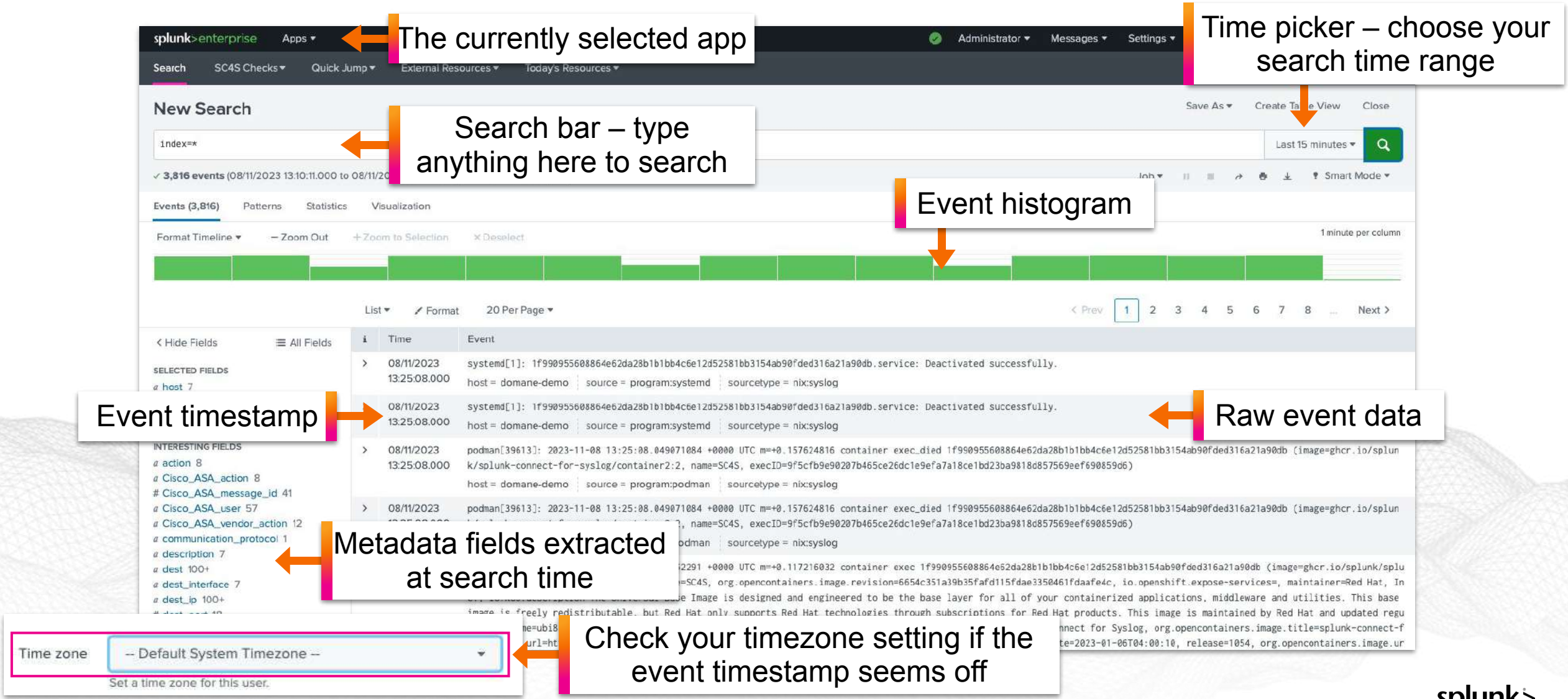
(If not you may have disabled the sc4s restart input too quickly a few slides back - try that again!)

splunk>

Lab 2: Introduction to Search

splunk>

Open your app and have a play!



splunk>

Start Exploring Your Data

Example searches

`dns`	Find all events containing “dns”
`dns *.com`	Find all events containing “dns” and words ending with “*.com”
`dns (*.com OR *.org)`	Boolean operators (AND/OR/NOT) – must be UPPERCASE!
`sourcetype="cisco:asa"`	Use *fieldname* = *value* to ensure accurate search results

How would you find dns events that are NOT .google.com?

```
dns NOT *.google.com
dns query!=*.google.com
```

Splunk Quick Reference Guide: https://splk.it/SplunkQuickRef

splunk>

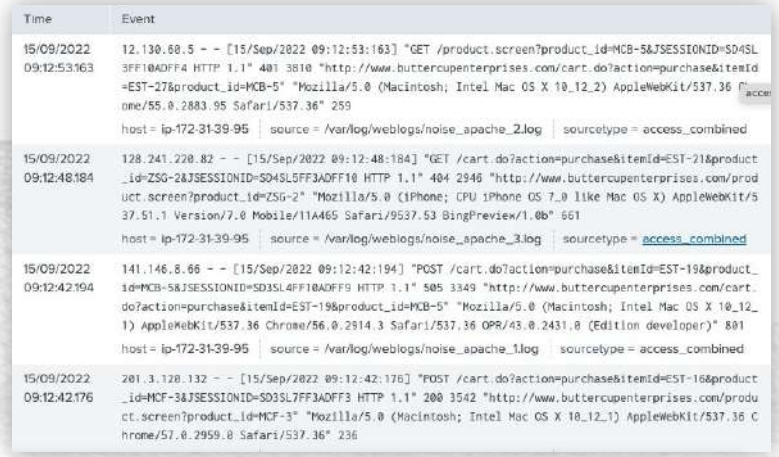
Splunk's Search Processing Language (SPL)

Search Terms (constraints) **Commands**

```
sourcetype="cisco:asa" | stats count by query | rename count as "number of events"
```

Pipe character: Output of left is input to right **Functions**

e.g. `sourcetype="cisco:asa"` `| stats count by query` `| rename count as "number of events"`



query	count
005.tv	1
0575360.com	2
100pstatic.nl	1

query	number of events
005.tv	1
0575360.com	2
100pstatic.nl	1

Want to know more? Check out:
Splunk Quick Reference Guide: https://splk.it/SplunkQuickRef
Search manual: https://splk.it/SplunkSearchManual

splunk>

Lab 3: Infosec App

Introducing the InfoSec App for Splunk

Download: https://splunkbase.splunk.com/app/4240

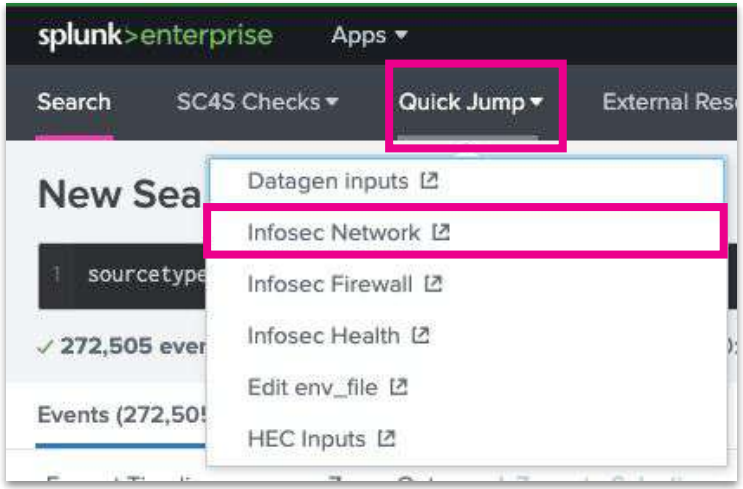
InfoSec App for Splunk is your starter security pack.

The InfoSec app is free (as in beer) to install and use and is designed to address the most common security use cases, including continuous monitoring and security investigations, as well as a number of advanced threat detection use cases.

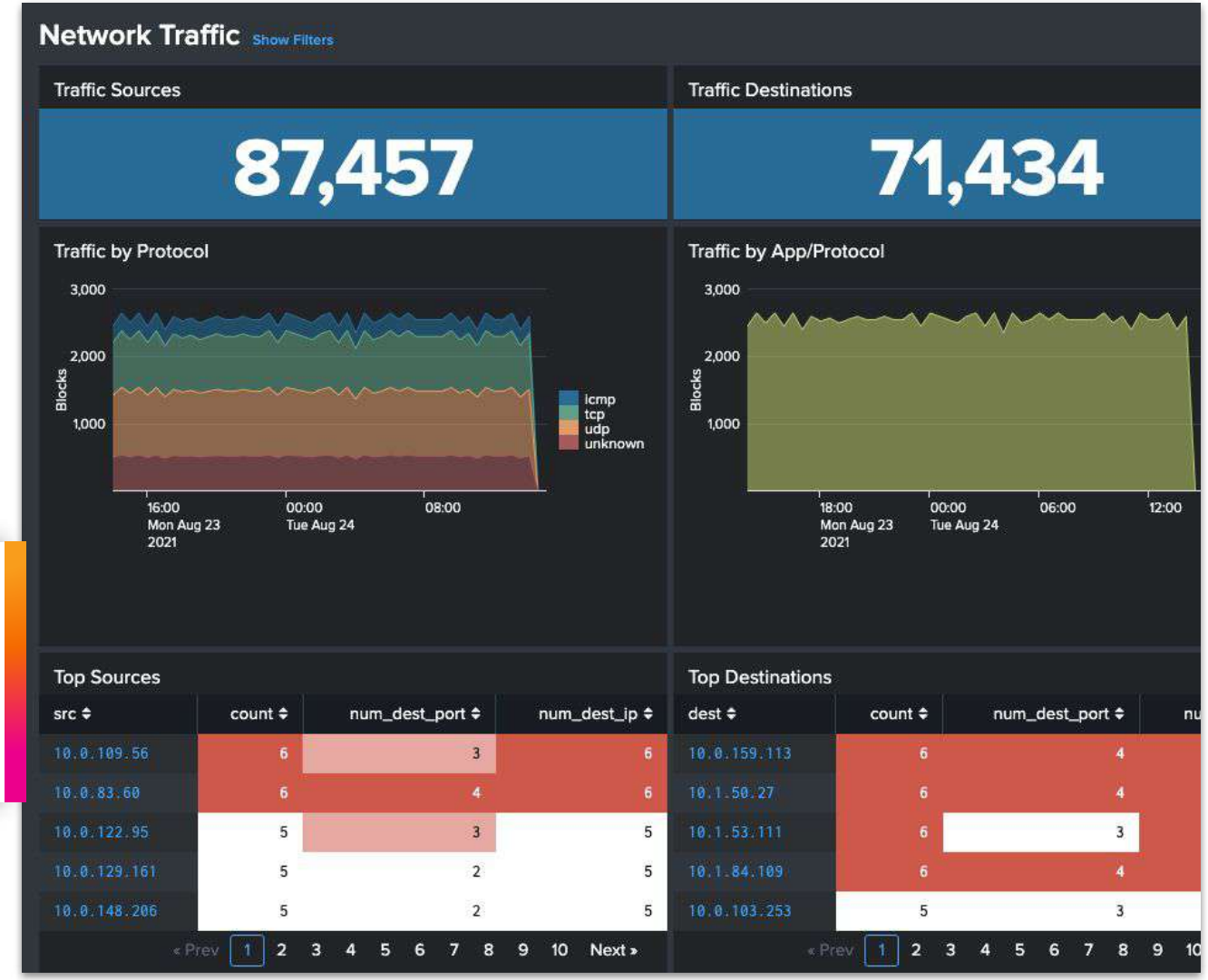
All of the components of the InfoSec app can be easily expanded using free security resources available for Splunk like Splunk Security Essentials

splunk>

InfoSec App for Splunk: Examples

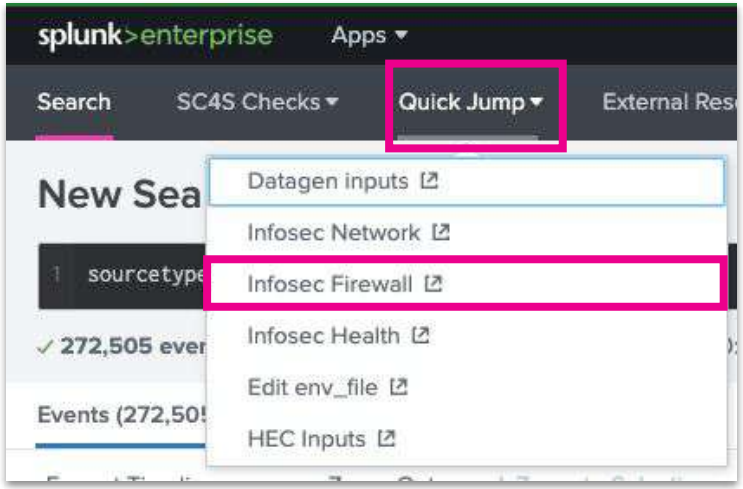


Network Traffic
This dashboard provides visibility into your network traffic



splunk>

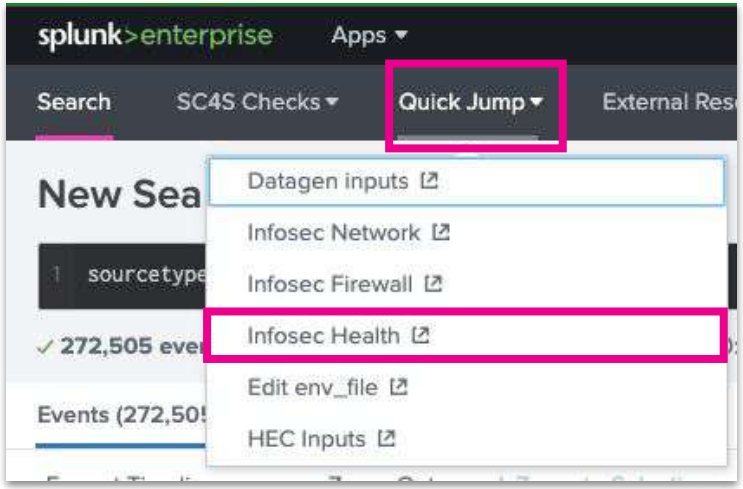
InfoSec App for Splunk: Examples

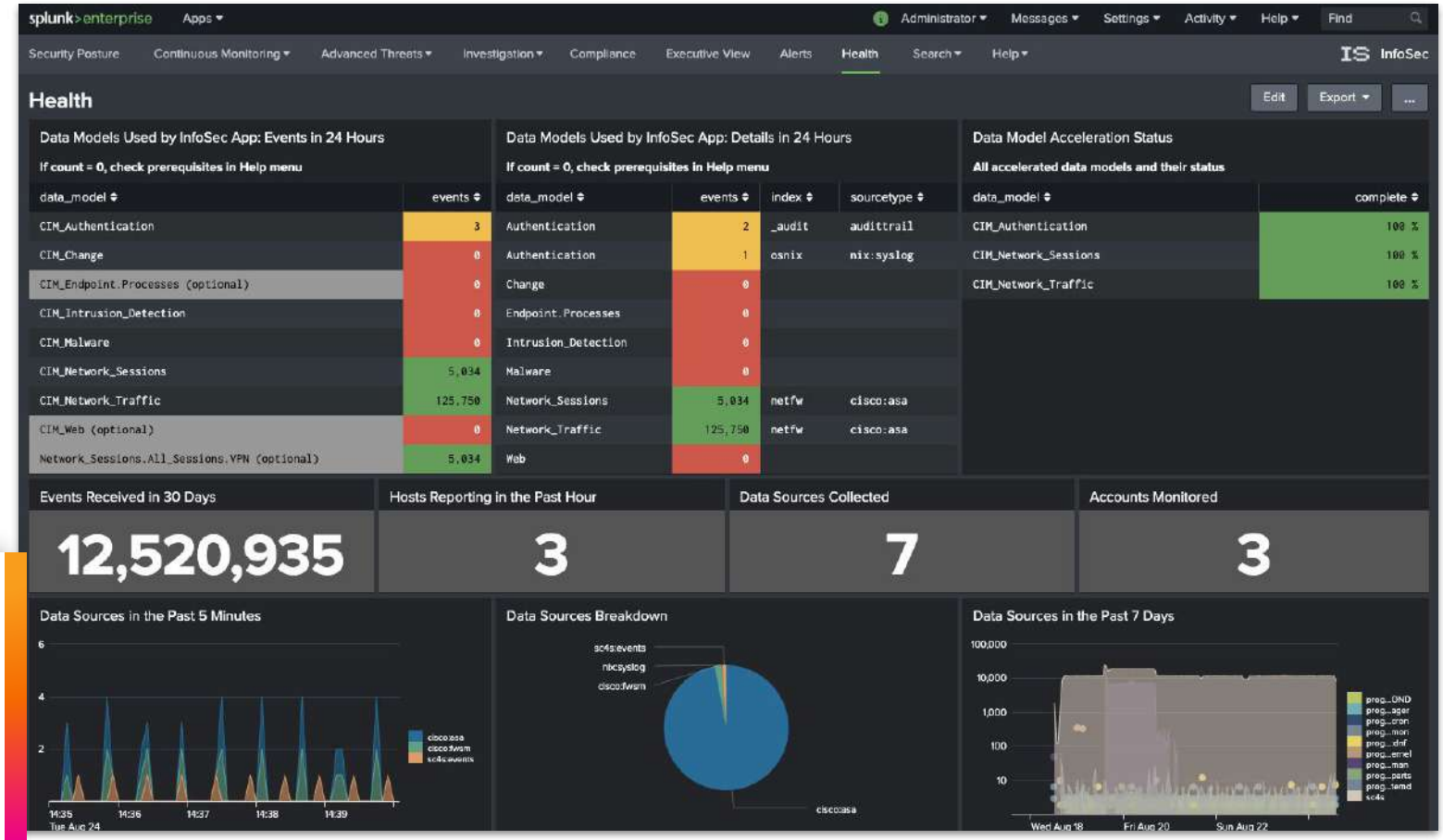


Firewalls
Provides visibility into your firewalls, regardless of vendor or the presence of data from multiple vendors in the same environment.

splunk>

InfoSec App for Splunk: Examples



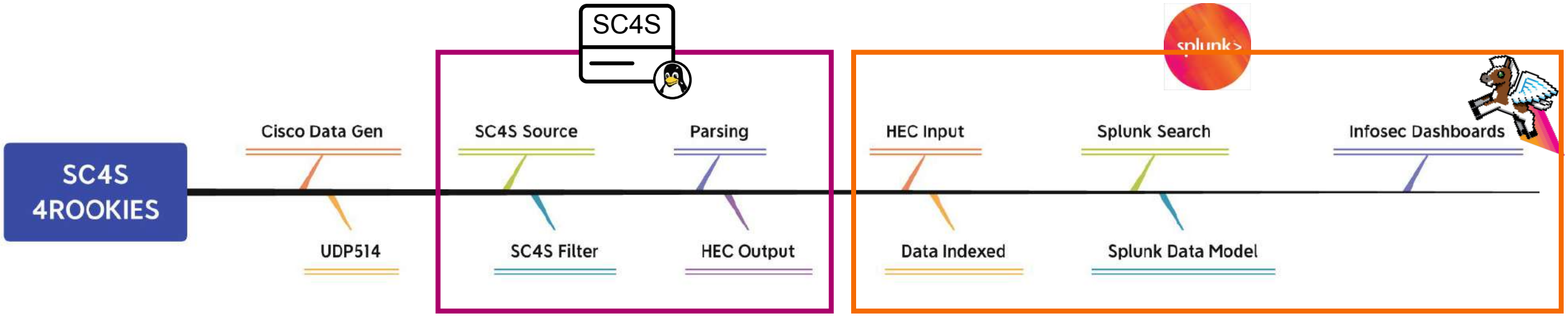


Health
Provides visibility into your InfoSec Health.

InfoSec is built on top of Splunk's amazing CIM data models.

splunk>

How is This All Working?



In SC4S, most of the log path is abstracted via the configuration mechanisms, which frees the administrator from understanding the nuances of the syslog-ng syntax.

But the structure of a log path will become immediately apparent when developing a new one, and an understanding of how the parts fit together is crucial.

- Global Options
- Log Paths (Source, Dest, Filter)
  - Sources
  - Filters
  - Message Parsing
  - Destinations

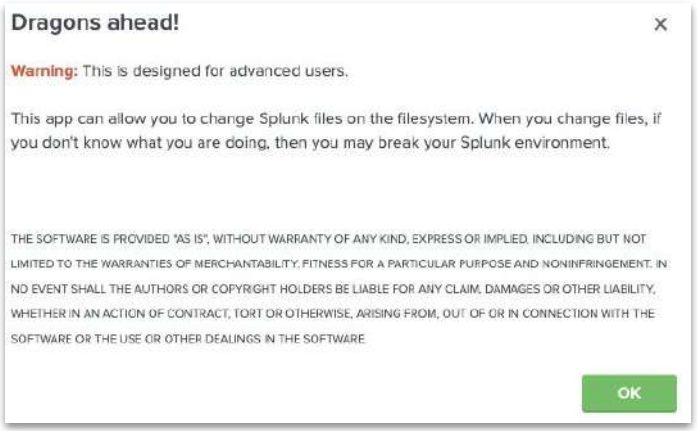
In our case today…

- **= UDP 514**
- **= `cisco:asa` matches**
- **= Timestamps, host, etc. extraction**
- **= Splunk HEC (127.0.0.1:8088)**

Lab 4:
Config Explorer

splunk>

Warning: You Can Break Things!

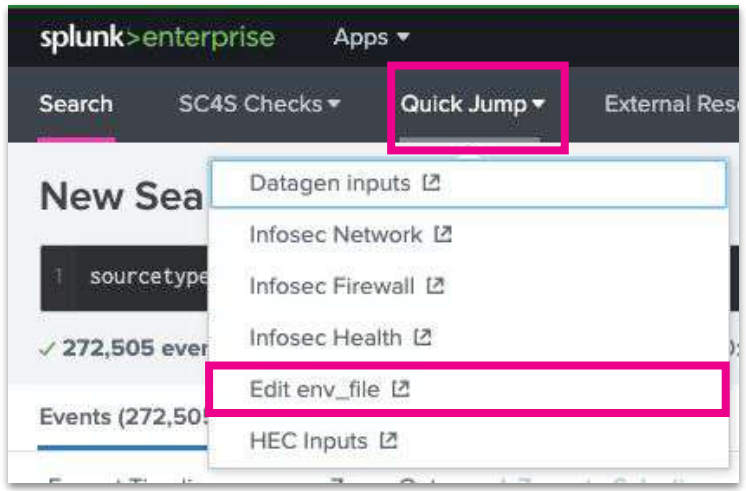


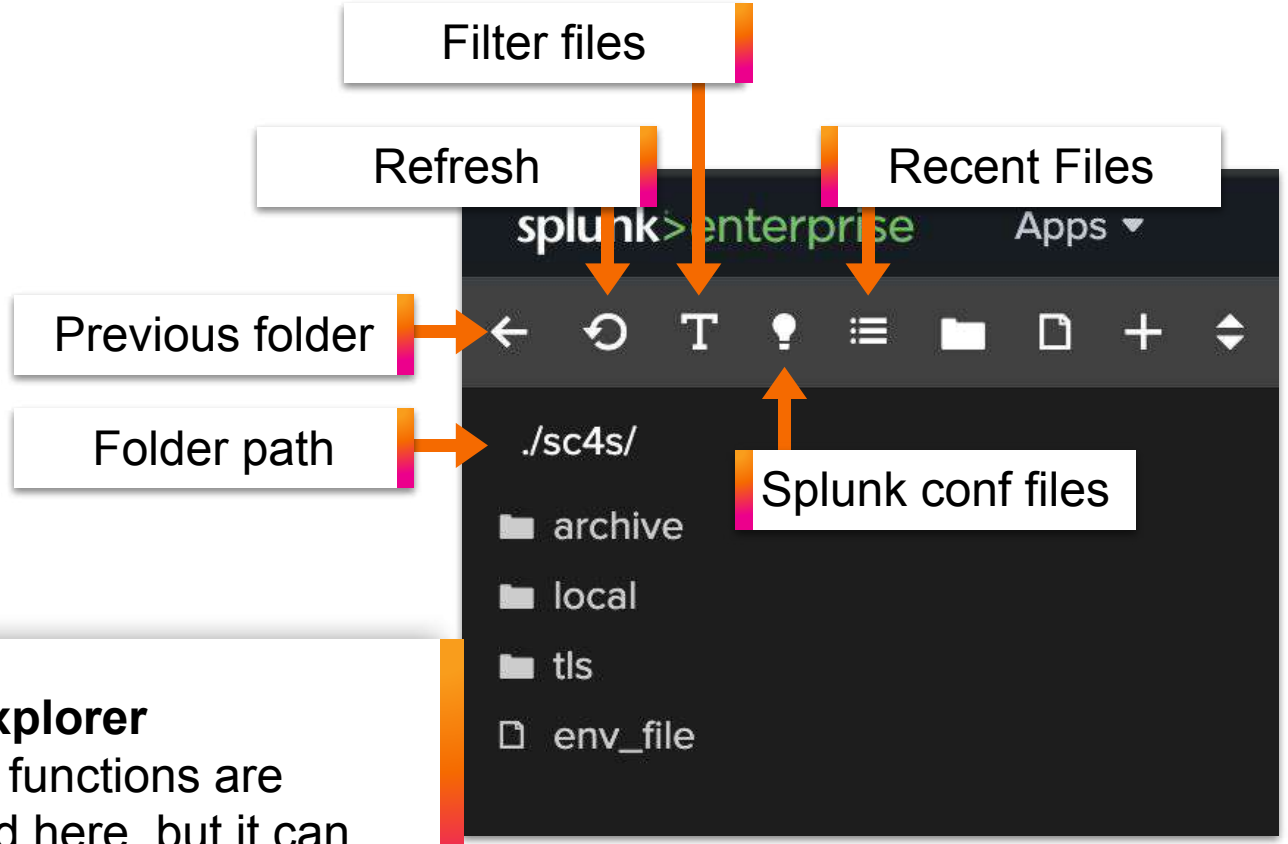
- Today you are an **admin**: don't let the power go to your head!
- You can edit all the files you can see in the **Config Explorer** app - this does NOT mean you should!
- Follow the path and guidance of the labs to the best of your ability
- As the old Unix adage goes:

> *"Unix was not designed to stop you from doing stupid things, because that would also stop you from doing clever things" - Doug Gwyn*

splunk>

Config Explorer 101





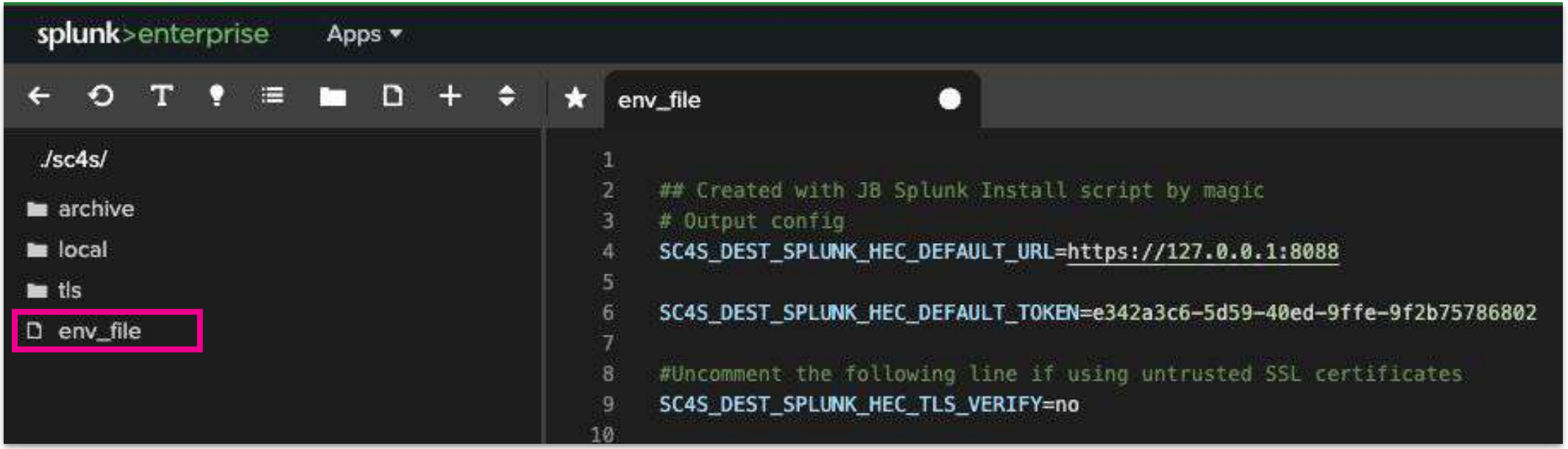
Config Explorer
The basic functions are highlighted here, but it can also do much more than this!

splunk>

Config Explorer 102

It's like Visual Studio Code, but in Splunk!

Click on the **`env_file`** to see the contents:



Learn more: https://splunk.github.io/splunk-connect-for-syslog/main/configuration/#global-configuration

splunk>

Introducing SC4S Configuration Files

Files	Configuration Level	Typical Use
env_file	Global and Required Settings	• Required SC4S configuration • Unique listening ports • Alternate destinations • Kernel settings • Debug/Development settings
splunk_metadata.csv	Splunk Metadata	• Set/override Splunk metadata • Set event output format
compliance_meta_by_source.*	• Compliance Overrides • Event sub-filters	• Override Splunk metadata for subset of events • Add indexed fields to subset of events
vendor_product_by_source.*	Vendor/Product Event Categorization	• Set/override Splunk metadata • Set event output format
• lp-<vendor>.conf.tmpl d_<destination>.conf • s_<source>.conf	Platform Extension (Log Paths)	• Add new data source to SC4S • Add new destination to SC4S • Add new collection method (source) to SC4S

Covered in "SC4S 101"

Covered in "SC4S 201"

For another day…

Learn more:
https://www.splunk.com/en_us/blog/tips-and-tricks/splunk-connect-for-syslog-turnkey-and-scalable-syslog-gdi-part-3.html

splunk>

Lab 5: Set a Custom Index

splunk>

Introducing splunk_metadata.csv

splunk_metadata.csv is located **/sc4s/local/context/** in your lab instance and is a CSV file containing a “key” that is referenced in the log path for each data source. We will use the **Config Explorer** app to change the metadata and re-route some events.

The columns in this file are key, metadata, and value:

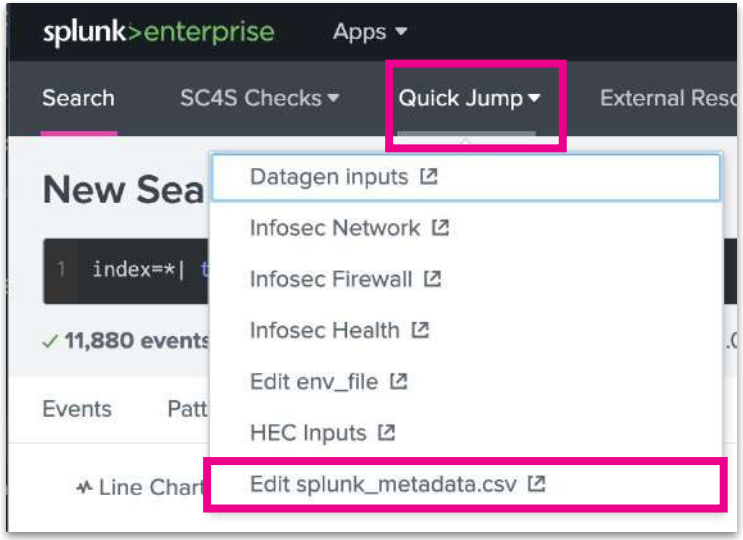
- **index** to specify an alternate value for index
- **source** to specify an alternate value for source
- **host** to specify an alternate value for host
- **sourcetype** to specify an alternate value for sourcetype
  - Be very careful when changing this - only change it if an upstream TA is not being used or a custom TA (built by you) is being used
- **sc4s_template** to specify an alternate value for the syslog-ng template that will be used to format the event that will be indexed by Splunk.
  - Changing this carries the same warning as the sourcetype above; this will affect the upstream TA. The template choices are documented elsewhere in this Configuration section.

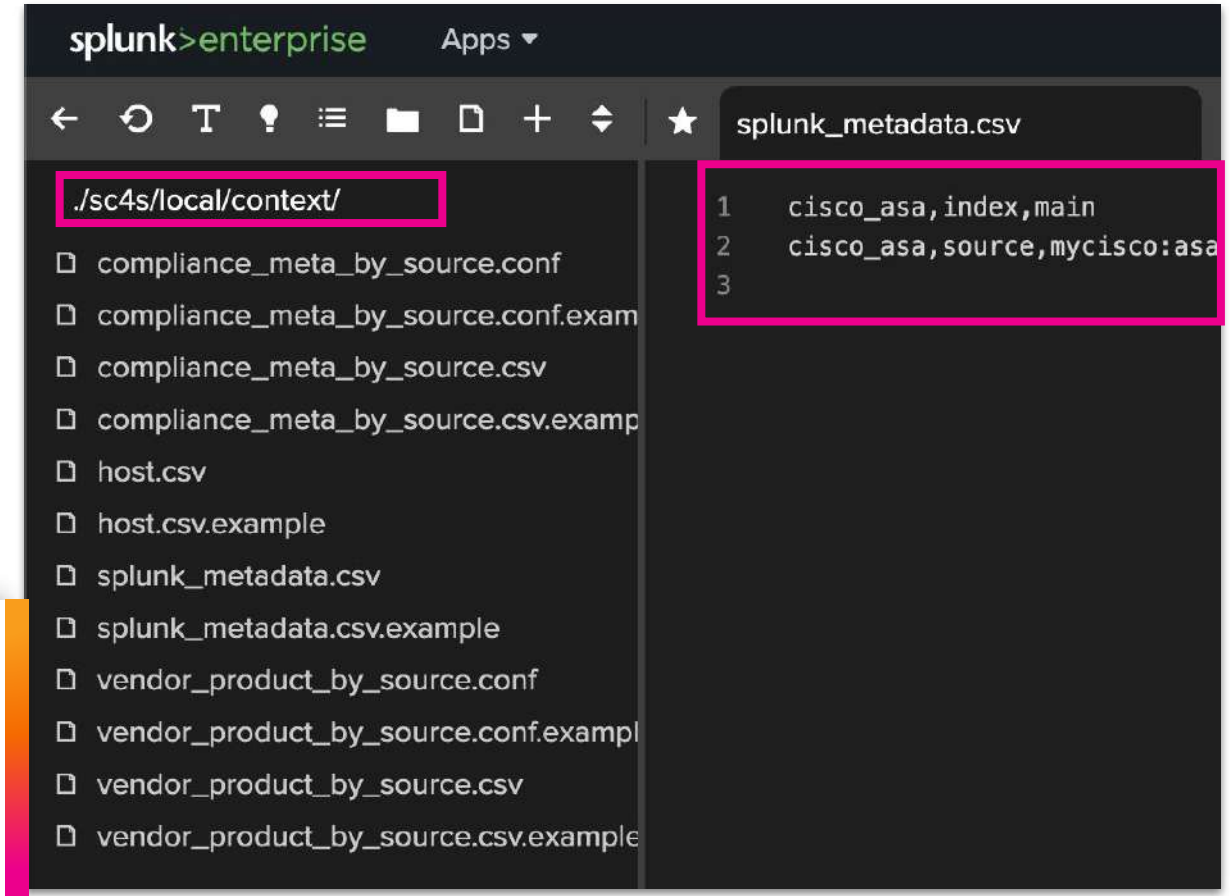
We will set:

index to “**main**”
source to “**mycisco:asa**”

Learn more: https://splunk.github.io/splunk-connect-for-syslog/main/configuration/#sc4s-metadata-configuration

Edit splunk_metadata.csv





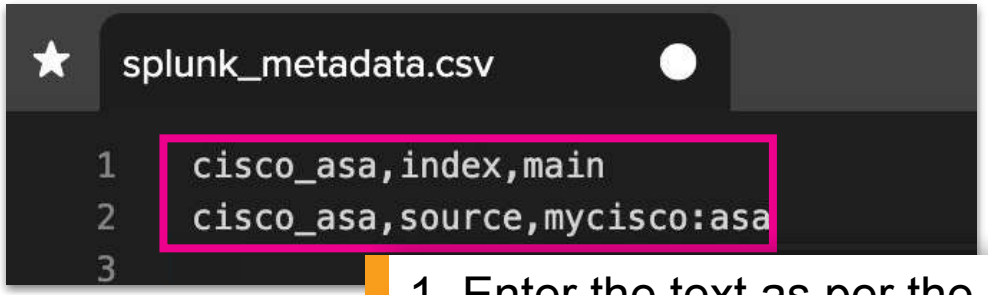
As you can see the file is currently empty.

We are going to set some custom items, restart SC4S and enjoy the show!

Learn more: https://splunk.github.io/splunk-connect-for-syslog/main/configuration/#sc4s-metadata-configuration

splunk>

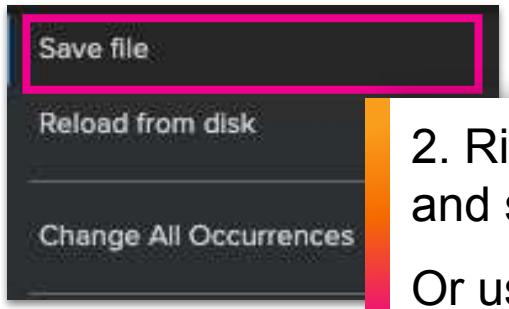
Edit splunk_metadata.csv



1. Enter the text as per the screenshot to the left

change:

- **index** to **main**
- **source** to **mycisco:asa**



2. Right click in the window and select **Save file**

Or use **cmd+S** (Mac) or **ctrl+S** (Windows)

This configuration will take precedence over the default for the **cisco_asa** filter.

The default values of:

key = cisco_asa	key = cisco_asa
metadata = index	metadata = source
value = **netfw**	value = **cisco:asa**

will become:

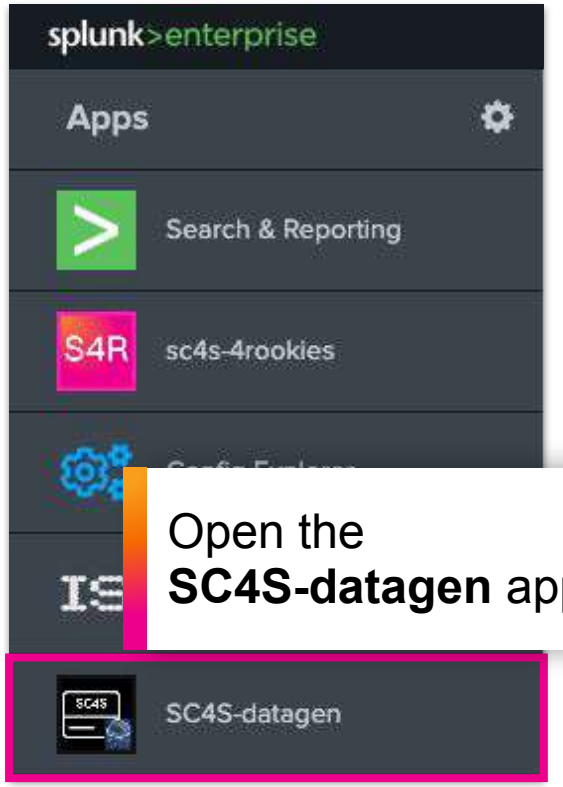
key = cisco_asa	key = cisco_asa
metadata = index	metadata = source
value = **main**	value = **mycisco:asa**

Learn more: https://splunk.github.io/splunk-connect-for-syslog/main/sources/Cisco/#product-asa-and-ftd-firepower

splunk>

Restart SC4S systemd Service

Restart the systemd service for podman and SC4S by enabling the input, waiting a few seconds, then disabling it again:



Open the **SC4S-datagen** app

1. Set the toggle to Enabled

2. Wait for **10 seconds…**

1. Set the toggle to Disabled

Note: In environments with CLI access you can also restart the service using the following command:

```
sudo systemctl restart sc4s
```

splunk>

Check the Changes Have Been Applied

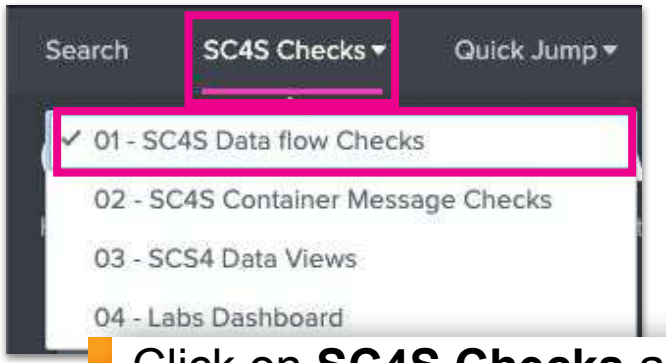
Run the following Splunk search:

```
index=main sourcetype=cisco:asa
```

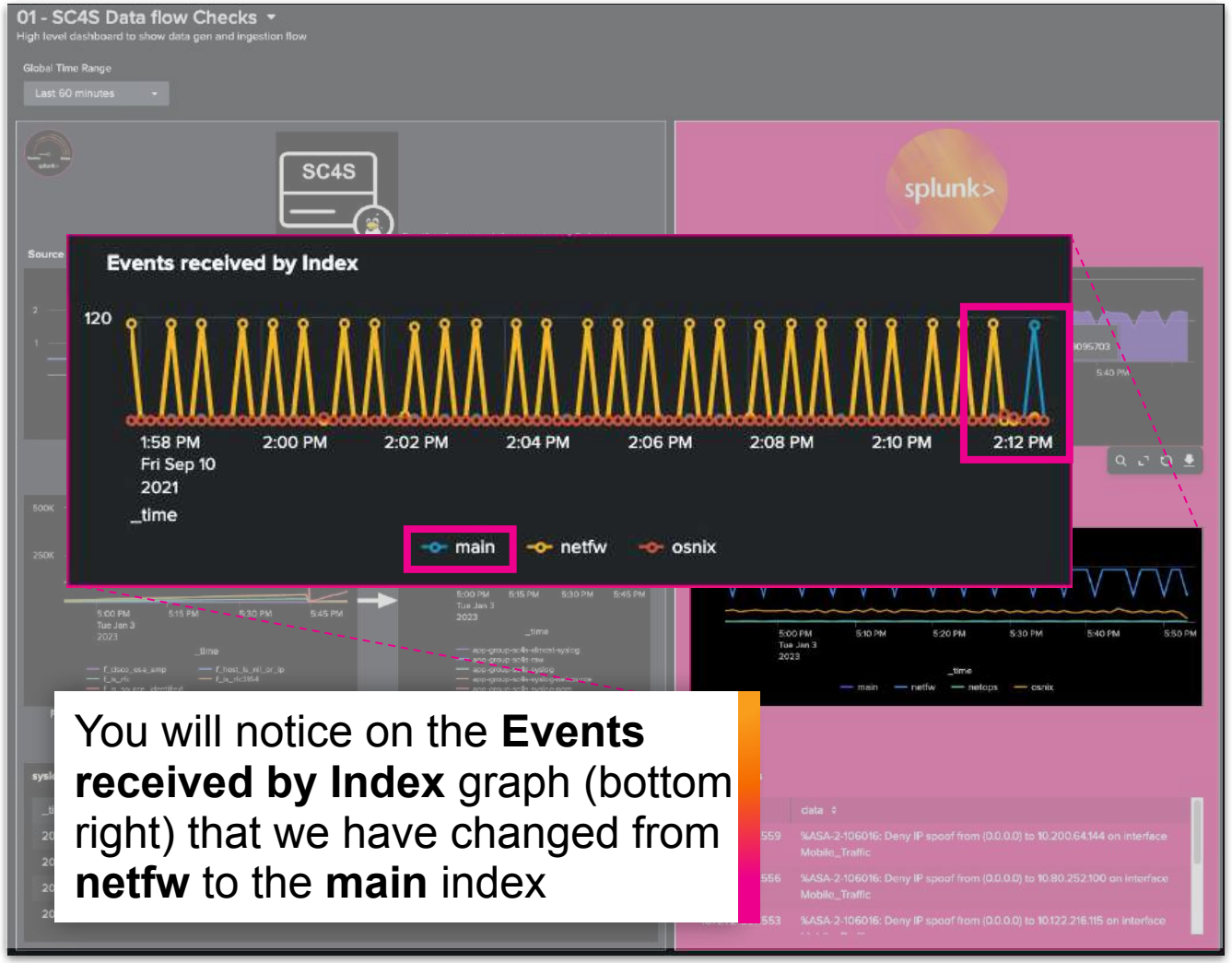
Learn more: https://splunk.github.io/splunk-connect-for-syslog/main/configuration/#sc4s-metadata-configuration

splunk>

Check the Index Change



Click on **SC4S Checks** and then on **01 - SC4S Data flow Checks**



You will notice on the **Events received by Index** graph (bottom right) that we have changed from **netfw** to the **main** index

splunk>

Check the Index Change

Run the following Splunk search:

```
index=* | timechart count by index
```

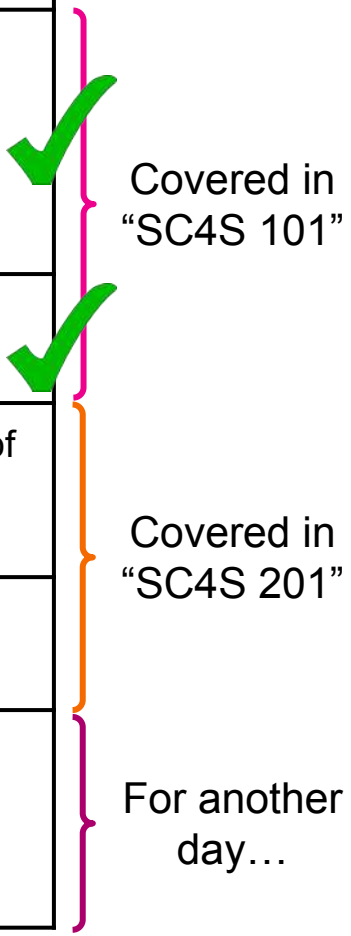
Note the index change here when we restarted SC4S

Learn more: https://splunk.github.io/splunk-connect-for-syslog/main/configuration/#sc4s-metadata-configuration

splunk>

Introducing SC4S Configuration Files

Files	Configuration Level	Typical Use
env_file	Global and Required Settings	• Required SC4S configuration • Unique listening ports • Alternate destinations • Kernel settings • Debug/Development settings
splunk_metadata.csv	Splunk Metadata	• Set/override Splunk metadata • Set event output format
compliance_meta_by_source.*	• Compliance Overrides • Event sub-filters	• Override Splunk metadata for subset of events • Add indexed fields to subset of events
vendor_product_by_source.*	Vendor/Product Event Categorization	• Set/override Splunk metadata • Set event output format
• lp-<vendor>.conf.tmpl d_<destination>.conf • s_<source>.conf	Platform Extension (Log Paths)	• Add new data source to SC4S • Add new destination to SC4S • Add new collection method (source) to SC4S

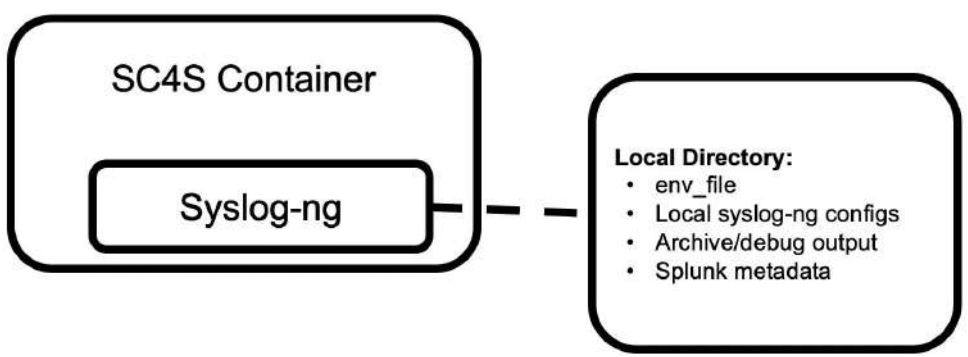


Learn more:
https://www.splunk.com/en_us/blog/tips-and-tricks/splunk-connect-for-syslog-turnkey-and-scalable-syslog-gdi-part-3.html

splunk>

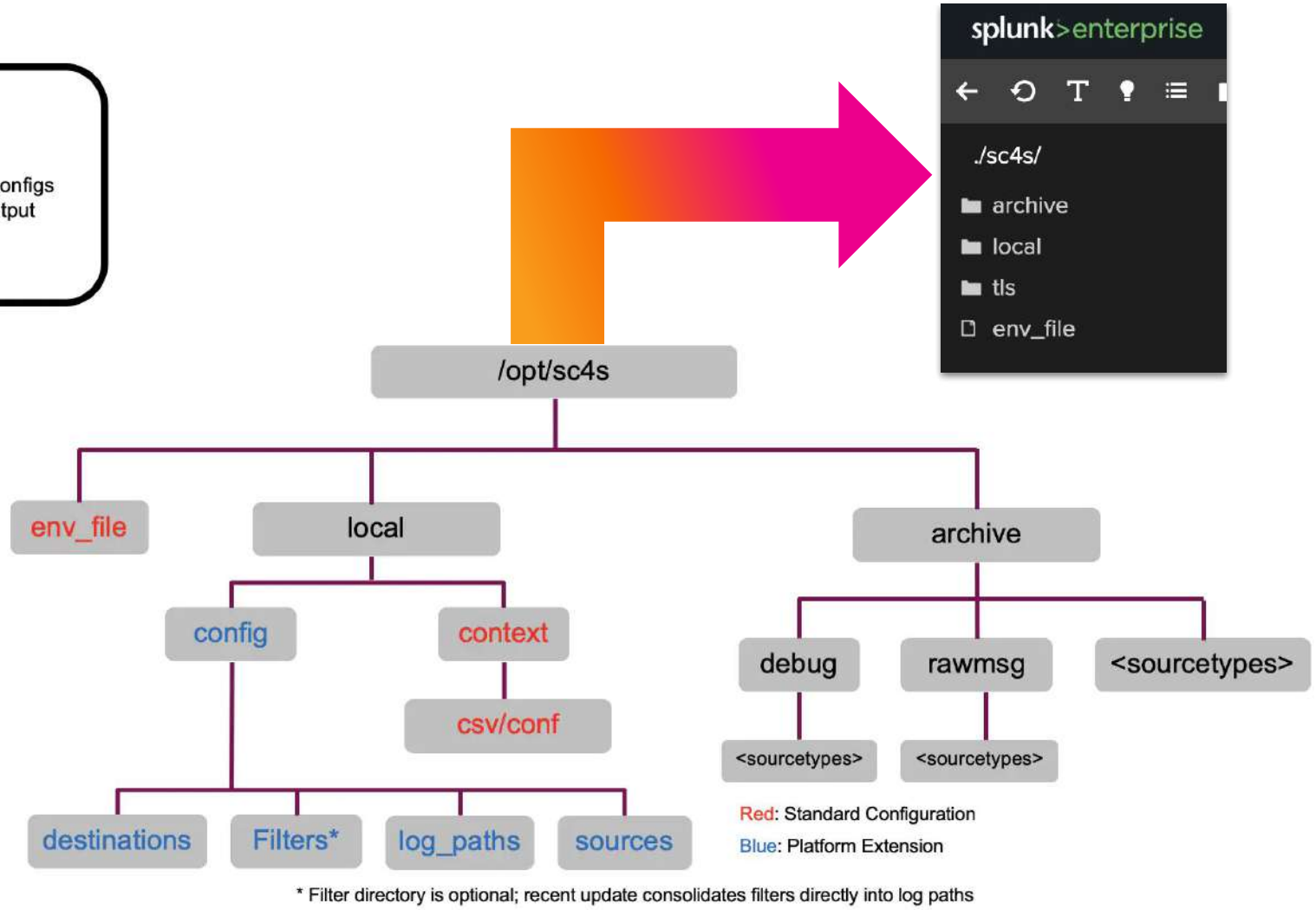
SC4S 201

SC4S File System



Directories and files shown in red are part of a standard SC4S configuration and cover the required components such as the URL and token for the Splunk HEC endpoint (which is the default destination for SC4S traffic) as well as configuration for Splunk metadata (index, etc.)

Files contained in the blue directories contain configurations for locally configured collection methods (sources), output destinations, and - most commonly - log paths (filters) for data sources that SC4S does not support out of the box



Learn more: https://www.splunk.com/en_us/blog/tips-and-tricks/splunk-connect-for-syslog-turnkey-and-scalable-syslog-gdi-part-3.html

splunk>

Lab 6: Set a Custom Filter

splunk>

Introducing Compliance Overrides

In many cases it is appropriate to provide the same index/metadata overrides but based on PCI scope, geography, or other criterion rather than globally.

This is accomplished by the use of a file that uniquely identifies these source exceptions via **syslog-ng filters**, which maps to an associated lookup of alternate **indexes**, **sources**, or **other metadata**.

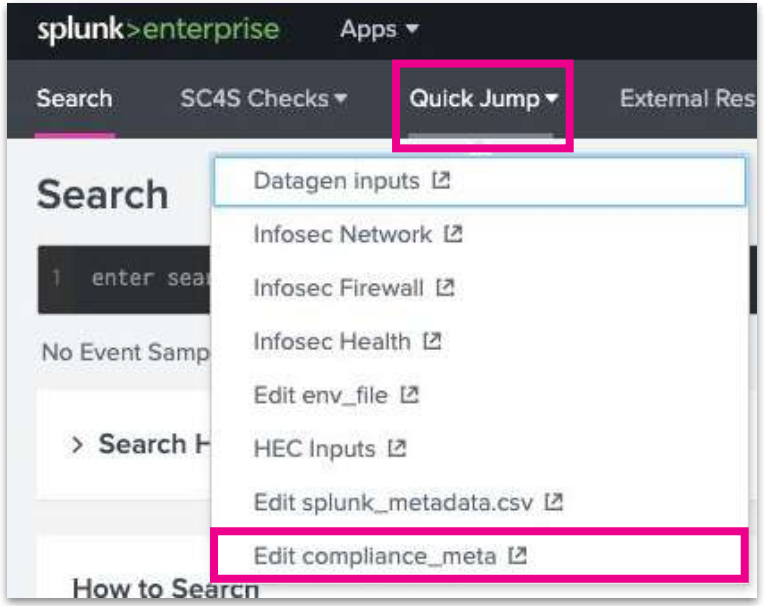
In addition, (indexed) fields can also be added to further classify the data.

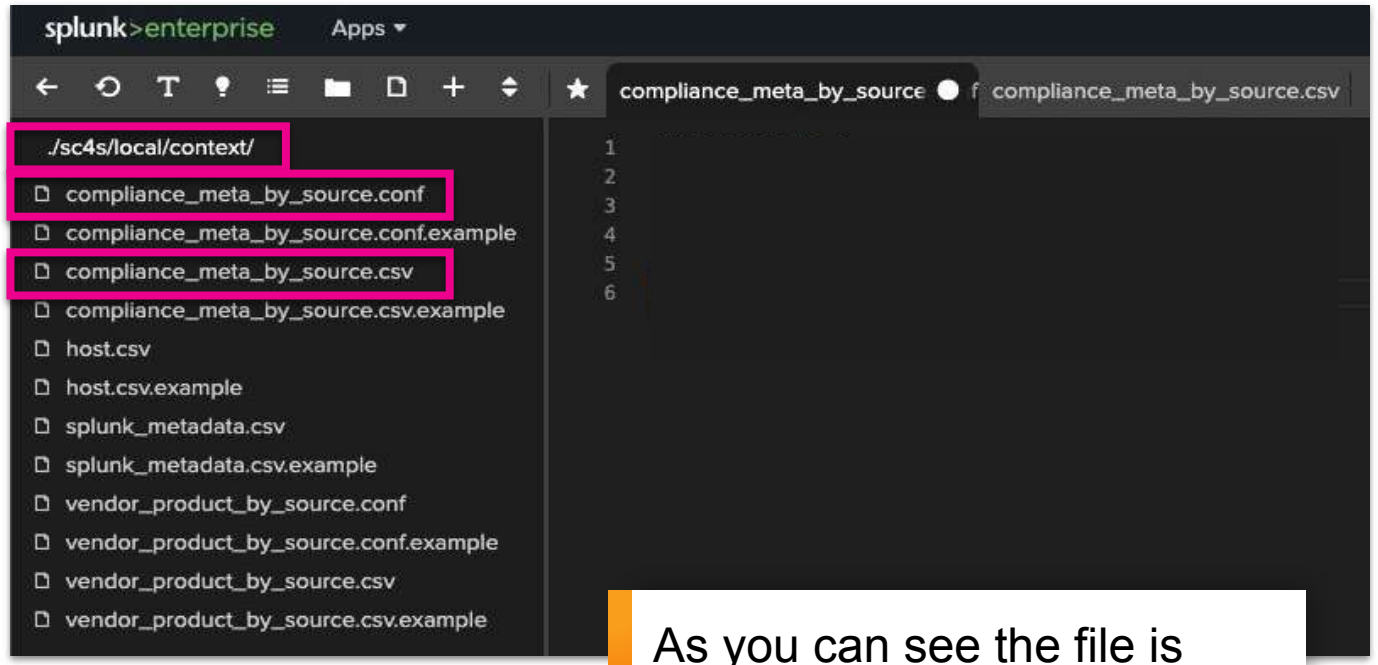
compliance_meta_by_source.csv and **compliance_meta_by_source.conf** are located in **/sc4s/local/context/** in your lab instance.

We will use the Config Explorer app to change the metadata and add some enrichment to the events - such as **datacentre**, **city**, **country** and some other fun things.

Learn more:
https://splunk.github.io/splunk-connect-for-syslog/main/configuration/#override-index-or-metadata-based-on-host-ip-or-subnet-compliance-overrides

Filter and Enrich





As you can see the file is currently empty

We are going to set some custom items, restart SC4S and enjoy the show

Learn more:
https://splunk.github.io/splunk-connect-for-syslog/main/configuration/#override-index-or-metadata-based-on-host-ip-or-subnet-compliance-overrides

splunk>

Filter and Enrich

compliance_meta_by_source.conf

Create **filters** in json format to match the events you want to manipulate

In this case we are choosing our Cisco device

compliance_meta_by_source.csv

Create **csv** data - this will be added to each event as it flows.

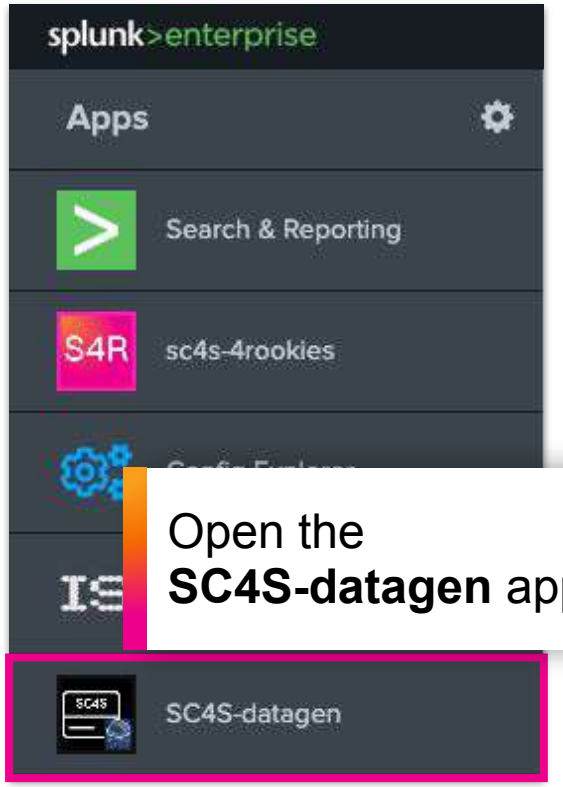
This is a fun example but there are production consequences such as **more disk space used** as fields are indexed. Use a **lookup** where possible as it's more efficient!

Learn more:
https://splunk.github.io/splunk-connect-for-syslog/main/configuration/#override-index-or-metadata-based-on-host-ip-or-subnet-compliance-overrides

splunk>

Restart SC4S systemd Service

Restart the systemd service for podman and SC4S by enabling the input, waiting a few seconds, then disabling it again:



Open the **SC4S-datagen** app

1. Set the toggle to Enabled

2. Wait for **10 seconds…**

1. Set the toggle to Disabled

Note: In environments with CLI access you can also restart the service using the following command:

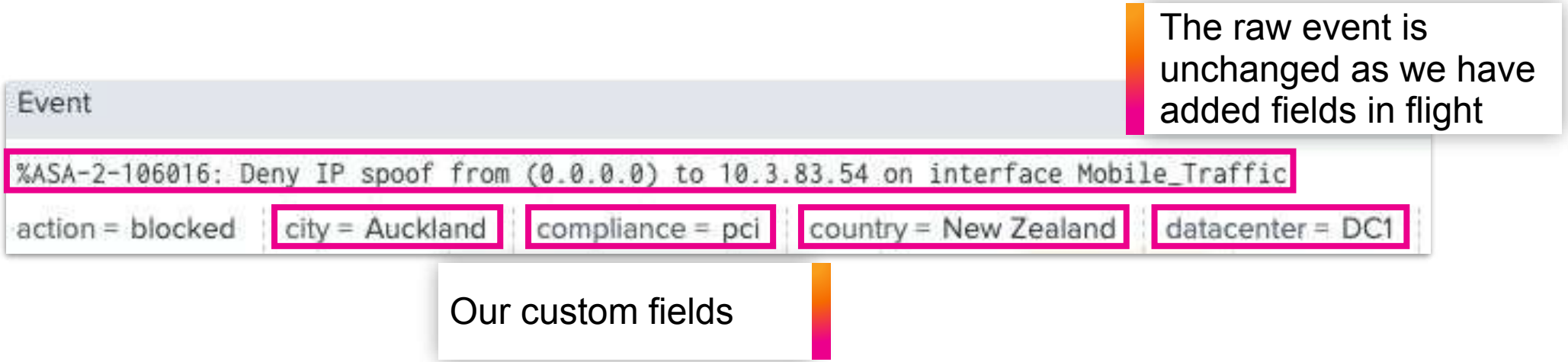
```
sudo systemctl restart sc4s
```

splunk>

Check the Changes Have Been Applied

Run the following Splunk search:

```
index=*
```



Learn more:
https://splunk.github.io/splunk-connect-for-syslog/main/configuration/#override-index-or-metadata-based-on-host-ip-or-subnet-compliance-overrides

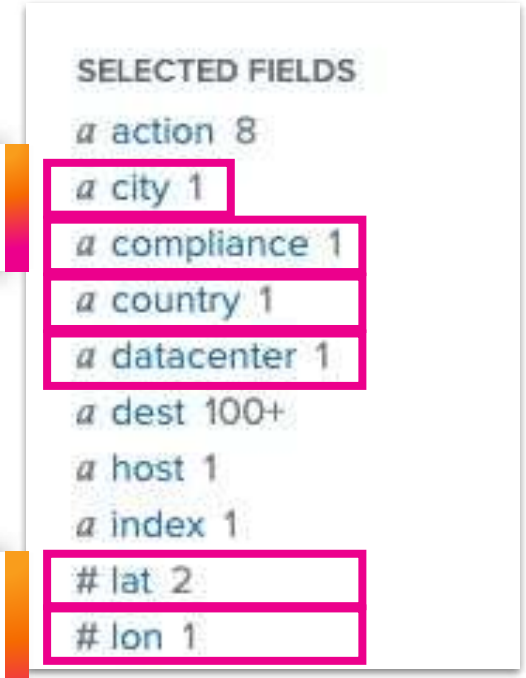
splunk>

Check the Changes Have Been Applied

```
index=main
```

Our custom fields

These ones look fun… what can we do with these?

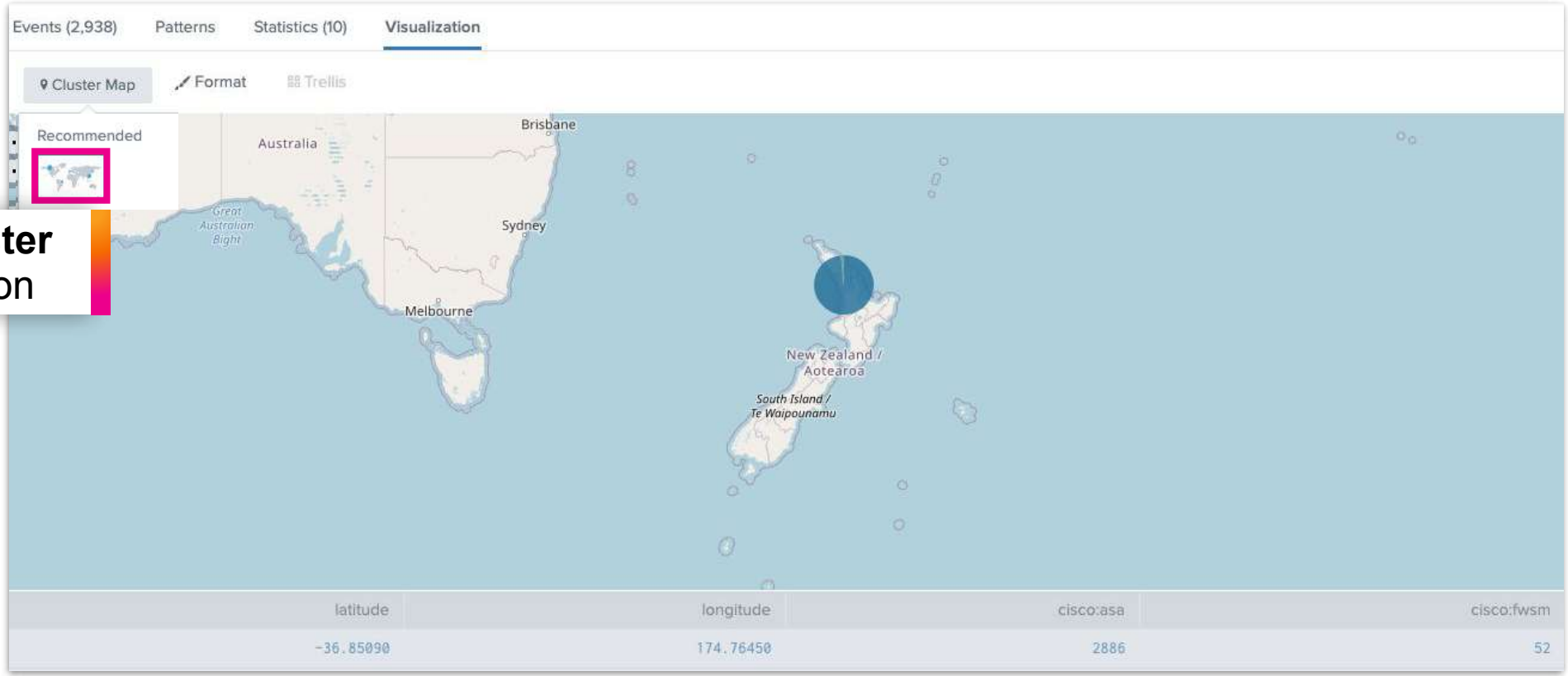


splunk>

Map It!

```
index=main sourcetype=*cisco*
| geostats count by sourcetype latfield=lat longfield=lon
```

Select the **Cluster Map** visualisation

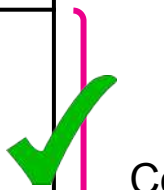


Learn more:
https://splunk.github.io/splunk-connect-for-syslog/main/configuration/#override-index-or-metadata-based-on-host-ip-or-subnet-compliance-overrides

splunk>

Introducing SC4S Configuration Files

Files	Configuration Level	Typical Use
env_file	Global and Required Settings	• Required SC4S configuration • Unique listening ports • Alternate destinations • Kernel settings • Debug/Development settings
splunk_metadata.csv	Splunk Metadata	• Set/override Splunk metadata • Set event output format
compliance_meta_by_source.*	• Compliance Overrides • Event sub-filters	• Override Splunk metadata for subset of events • Add indexed fields to subset of events
vendor_product_by_source.*	Vendor/Product Event Categorization	• Set/override Splunk metadata • Set event output format
• lp-<vendor>.conf.tmpl d_<destination>.conf • s_<source>.conf	Platform Extension (Log Paths)	• Add new data source to SC4S • Add new destination to SC4S • Add new collection method (source) to SC4S



Covered in "SC4S 101"

Covered in "SC4S 201"

For another day…

Learn more:
https://www.splunk.com/en_us/blog/tips-and-tricks/splunk-connect-for-syslog-turnkey-and-scalable-syslog-gdi-part-3.html

splunk>

Lab 7: null_queue

splunk>

Introducing dest-rewrite-fwsm_drop.conf

dest-rewrite-fwsm_drop.conf is located in **/sc4s/local/config/app_parsers/** in your lab instance. We will use the Config Explorer app to change the settings as follows:

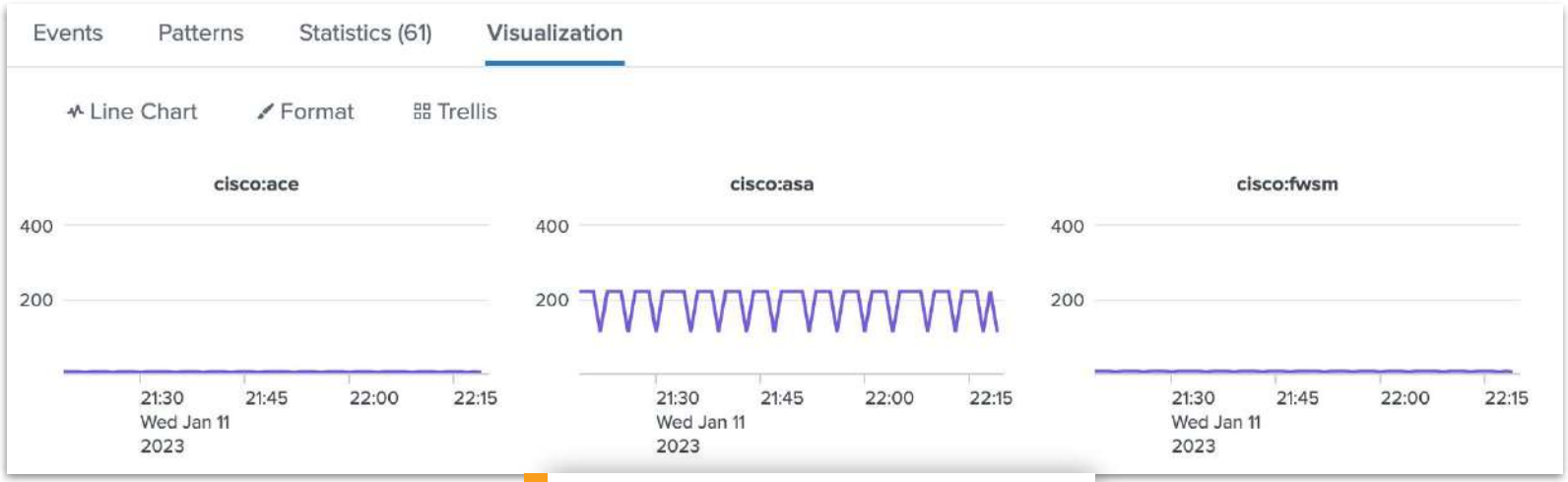
- We will use the **vendor** and **product** fields to **drop all syslog from our main firewall** (just for fun)
- File name is arbitrary, but needs to end in **.conf** and be located in the **/sc4s/local/config directory**

Learn more: https://splunk.github.io/splunk-connect-for-syslog/main/sources/#filtering-events-from-output

null_queue

```
index=* host=192* sourcetype=*cisco*
| timechart count by sourcetype
```

Enable the trellis layout and split by **sourcetype**



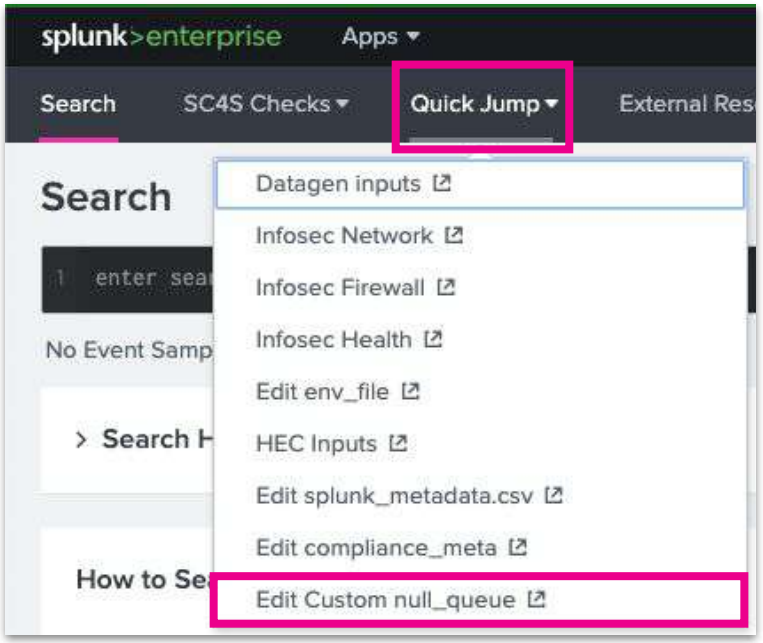
We are currently getting data from:

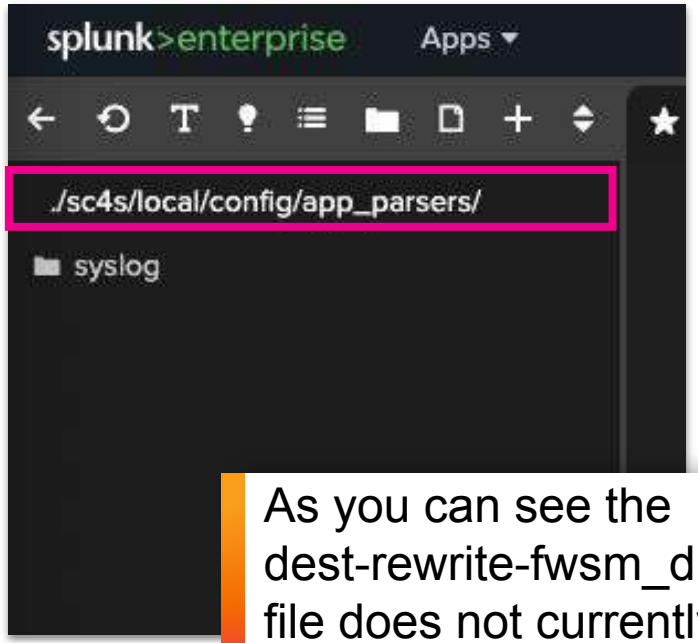
- **ace**
- **asa**
- **fwsm**

Learn more: https://splunk.github.io/splunk-connect-for-syslog/main/configuration/#dropping-all-data-by-ip-or-subnet

splunk>

null_queue





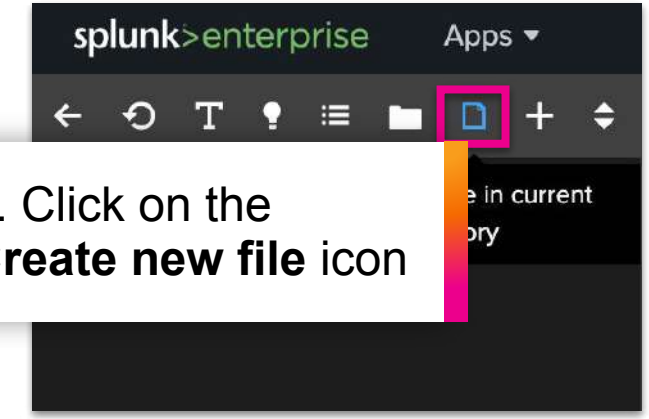
As you can see the dest-rewrite-fwsm_drop.conf file does not currently exist

We are going to create the file and its contents

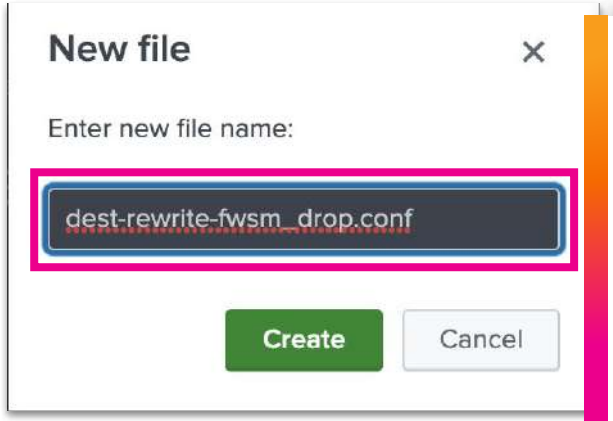
Learn more: https://splunk.github.io/splunk-connect-for-syslog/main/configuration/#dropping-all-data-by-ip-or-subnet

splunk>

null_queue



1. Click on the **Create new file** icon

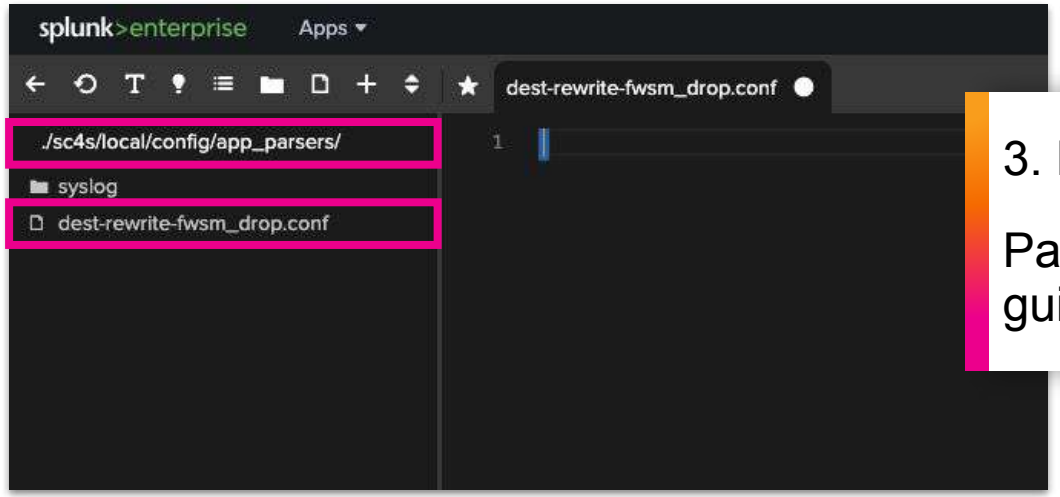


2. Enter the file name:

dest-rewrite-fwsm_drop.conf

then click on **Create**.

This file name is arbitrary but make sure it makes sense to you!



3. Double check the file path

Paste the file contents from the lab guide

Learn more: https://splunk.github.io/splunk-connect-for-syslog/main/sources/#filtering-events-from-output

splunk>

null_queue

dest-rewrite-fwsm_drop.conf

dest-rewrite-fwsm_drop.conf

```
1   # /opt/splunk/sc4s/local/config/app_parsers/dest-rewrite-fwsm_drop.conf
2
3   block parser fwsm_drop-postfilter() {
4       channel {
5   # In this case the outcome is drop the event
6           rewrite(r_set_dest_splunk_null_queue);
7       };
8   };
9   application fwsm_drop-postfilter[sc4s-postfilter] {
10      filter {
11          "${fields.sc4s_vendor}" eq "cisco" and
12          "${fields.sc4s_product}" eq "fwsm"
13  # Note regex reads as start from first position
14  # Any atleast 1 char that is not a `-`
15  # constant '-7-'
16  # and message('^%[^\-]+-7-');
17      };
18      parser { fwsm_drop-postfilter(); };
19  };
```

Create a **filter** in json format to match the events you want to send to **null_queue**, i.e. drop.

In this case we are choosing our **Cisco** vendor and **FWSM** product

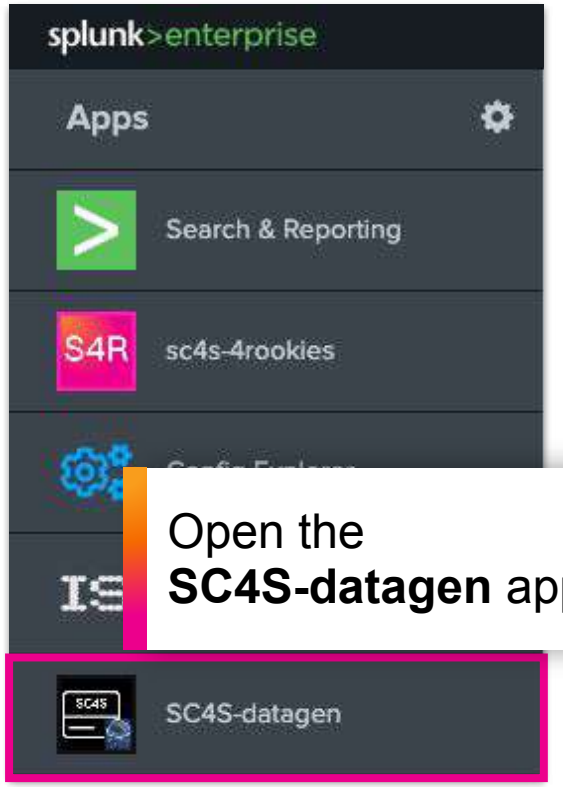
See the workshop lab guide for the full file contents to copy and paste!

Learn more: https://splunk.github.io/splunk-connect-for-syslog/main/sources/#filtering-events-from-output

splunk>

Restart SC4S systemd Service

Restart the systemd service for podman and SC4S by enabling the input, waiting a few seconds, then disabling it again:



Open the **SC4S-datagen** app

1. Set the toggle to Enabled

2. Wait for **10 seconds…**

1. Set the toggle to Disabled

Note: In environments with CLI access you can also restart the service using the following command:

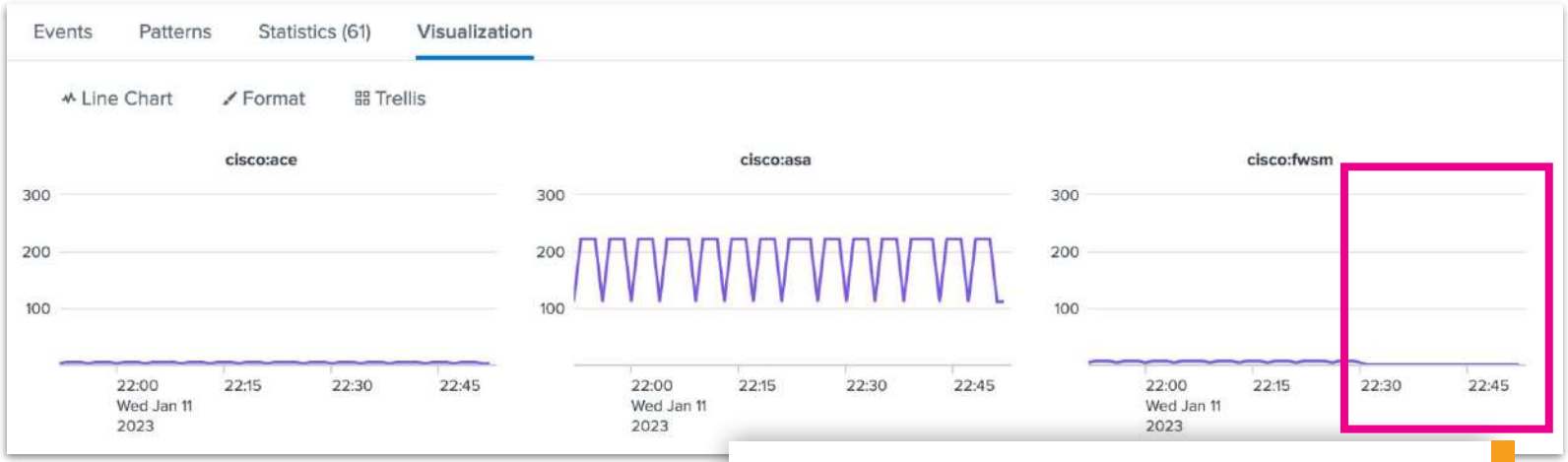
```
sudo systemctl restart sc4s
```

splunk>

null_queue

```
index=main host=192* sourcetype=*cisco*
| timechart count by sourcetype
```

Enable the trellis layout and split by **sourcetype**



We are currently getting data from:

- **ace**
- **asa**

fwsm has dropped away thanks to our filter

Learn more: https://splunk.github.io/splunk-connect-for-syslog/main/configuration/#dropping-all-data-by-ip-or-subnet

splunk>

Introducing SC4S Configuration Files

Files	Configuration Level	Typical Use
env_file	Global and Required Settings	• Required SC4S configuration • Unique listening ports • Alternate destinations • Kernel settings • Debug/Development settings
splunk_metadata.csv	Splunk Metadata	• Set/override Splunk metadata • Set event output format
compliance_meta_by_source.*	• Compliance Overrides • Event sub-filters	• Override Splunk metadata for subset of events • Add indexed fields to subset of events
vendor_product_by_source.*	Vendor/Product Event Categorization	• Set/override Splunk metadata • Set event output format
• lp-<vendor>.conf.tmpl d_<destination>.conf • s_<source>.conf	Platform Extension (Log Paths)	• Add new data source to SC4S • Add new destination to SC4S • Add new collection method (source) to SC4S

Covered in "SC4S 101"

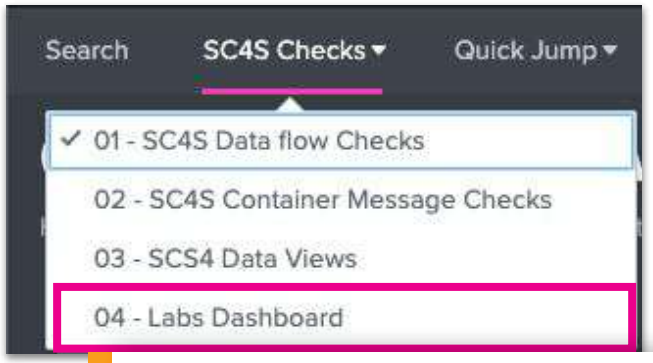
Covered in "SC4S 201"

For another day…

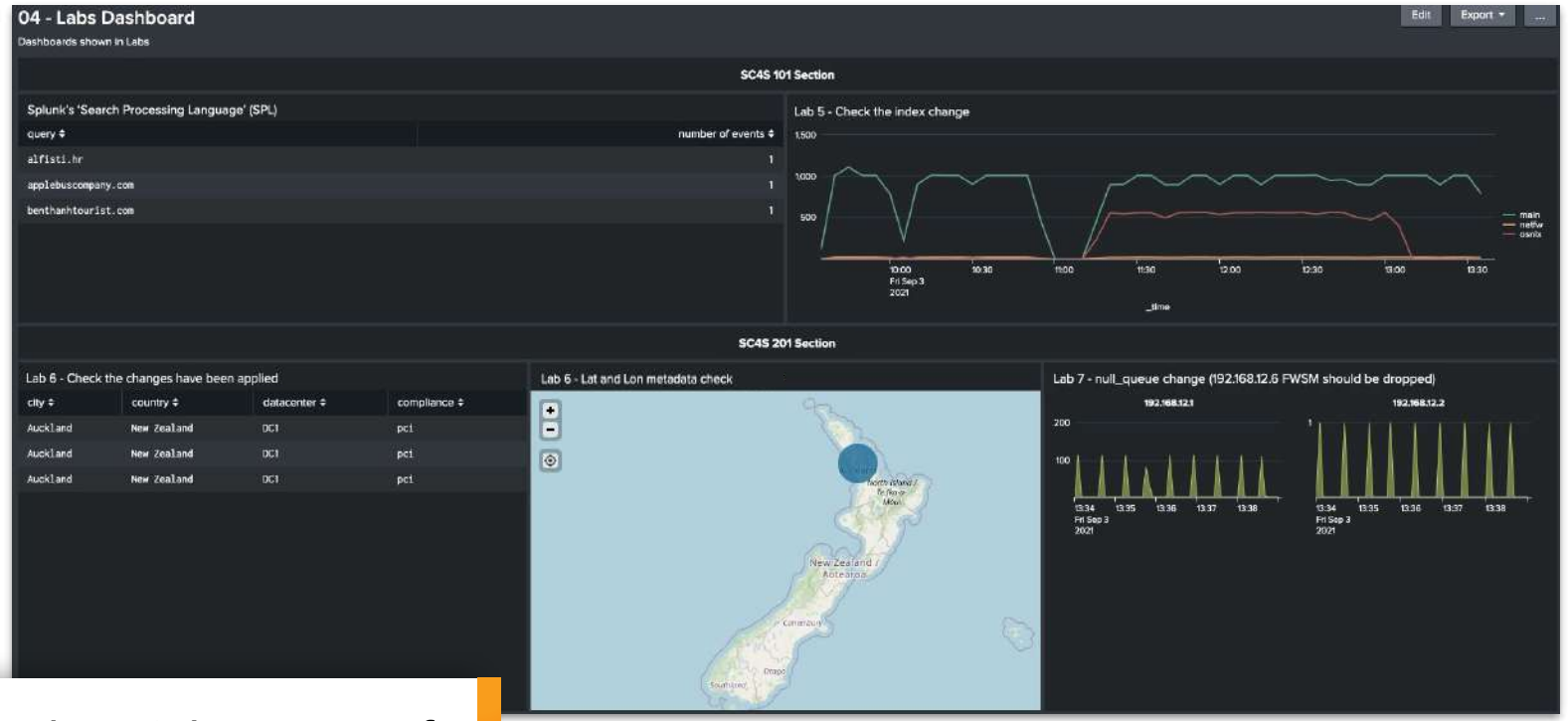
Learn more:
https://www.splunk.com/en_us/blog/tips-and-tricks/splunk-connect-for-syslog-turnkey-and-scalable-syslog-gdi-part-3.html

splunk>

Check Labs Dashboard



Click on **SC4S Checks** and then on **04 - Labs Dashboard**



This Dashboard contains some of the panels from the labs.

Yours may not match exactly, but this is a good checkpoint.

splunk>

SC4S
Notes for the Future

Multiple HEC Targets

splunk>

HTTP Load Balancer Setup

Set up the Splunk HEC endpoints **behind a load balancer (VIP) configured for HTTPS round robin WITHOUT sticky session**.

Alternatively, a list of HEC endpoint URLs can be configured in SC4S (native syslog-ng load balancing) if no load balancer is in place.

In most scenarios the recommendation is to use an external load balancer, as that makes longer term maintenance simpler by eliminating the need to manually keep the list of HEC URLs specified in SC4S current.

However, if a LB is not available, **syslog-ng native load balancing can be used with 10 or fewer Indexers** where HEC is used exclusively for syslog.

Also adjust the number of **workers to be a multiple of the receivers**, so if there are **3 indexers** set workers to **12** to ensure even event loading. Using **SC4S_DEST_SPLUNK_HEC_WORKERS** in the **env_file**

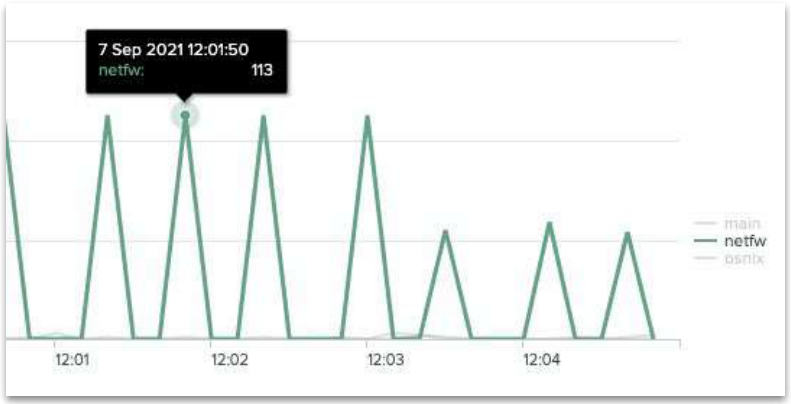
```
# Output config
SC4S_DEST_SPLUNK_HEC_DEFAULT_URL=https://idx1.example.com:8088,https://idx2.example.com:8088
SC4S_DEST_SPLUNK_HEC_DEFAULT_TOKEN=e342a3c6-5d59-40ed-9ffe-9f2b75786802

#Uncomment the following line if using untrusted SSL certificates
SC4S_DEST_SPLUNK_HEC_TLS_VERIFY=yes
```

Learn more: https://splunk.github.io/splunk-connect-for-syslog/main/configuration/#splunk-hec-destination-configuration

splunk>

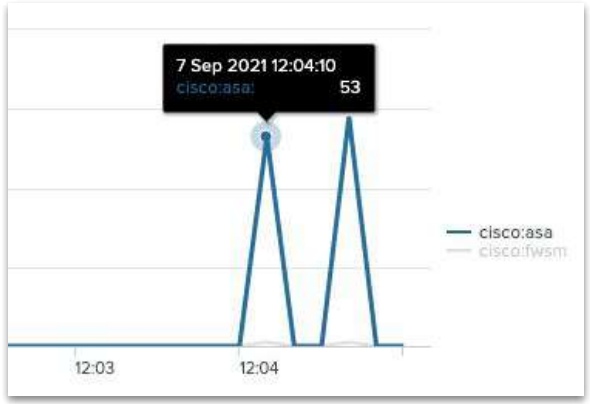
Load Balancing

HTTP Target 1



There are **113** events per burst. When a second HEC target is added, you can see that's roughly halved as the load is split

HTTP Target 2



Here they are - on our new target!

Syslog-ng native load balancing can be used with 10 or fewer Indexers

You can see in the example above that the split of data is almost evenly distributed between the indexers.

splunk>

systemd Service Template

splunk>

SC4S systemd Service

Key point - check your systemd file matches the systemd file service template for the container you are running.

Options:

- You could lock the version in by using **`Environment="SC4S_IMAGE=ghcr.io/splunk/splunk-connect-for-syslog/container:2.1.1"`** for example
- Comment out the "**`ExecStartPre=/usr/bin/podman pull \$SC4S_IMAGE`**" line to prevent the pull on every service restart
- Run a local docker repo rather than pull from **ghcr.io** directly to manage versions (see offline setup on next slide)

Offline Setup

splunk>

Offline Container Installation

Follow the instructions in the link below to “stage” SC4S by downloading the container so that it can be loaded “out of band” on a host machine, such as an air gapped system, without internet connectivity.

High Level steps:

1. Download container image “**oci_container.tgz**” from our GitHub page (see below)
2. Distribute the container to the air gapped host machine using an appropriate file transfer utility
3. Load the container image using the relevant Docker or podman command
4. Note the container ID of the resultant load
5. Use the container ID to create a local label
6. Use the local label **sc4slocal:latest** in the relevant unit or yaml file to launch SC4S (see the runtime options above) by setting the **SC4S_IMAGE** environment variable in the unit file (example below), or the relevant **image:** tag if using Docker compose/swarm
7. Using this label will cause the runtime to select the locally loaded image, and will not attempt to obtain the container image via the internet

Learn more: https://splunk.github.io/splunk-connect-for-syslog/main/gettingstarted/#offline-container-installation

splunk>

You Finished the Hands-on Exercises!

You made it!

Splunk Resources

Where to go after today's workshop

splunk>

Splunk Connect for Syslog Resources

There are many resources available to enable your success with SC4S!

In addition to the main repo and documentation, there are many other resources available:

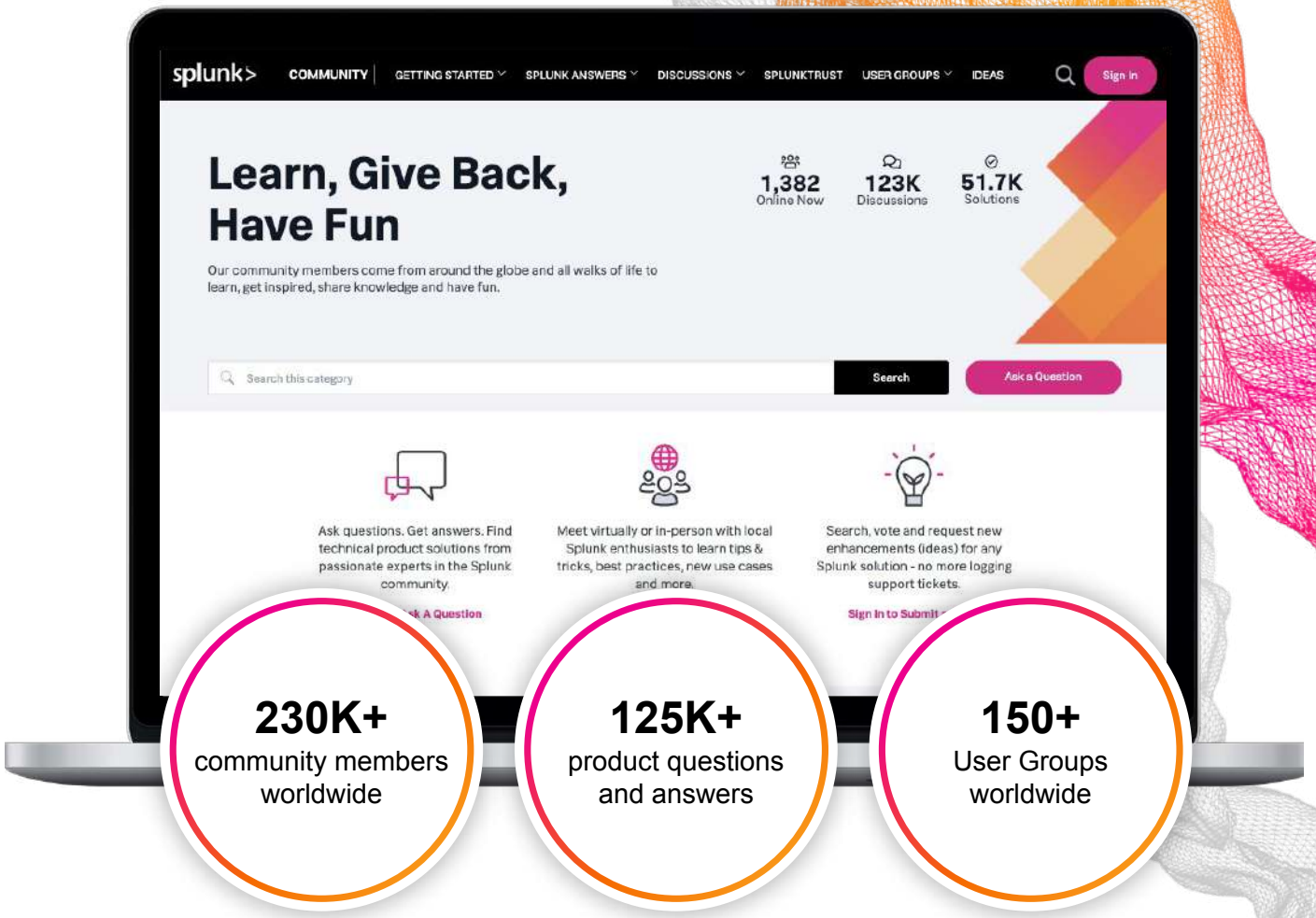
- **Part 1:** History, Design Goals, and Architecture
- **Part 2:** Splunk Connect for Syslog Configuration Overview
- **Part 3:** Splunk Connect for Syslog Configuration In Depth
- **Part 4:** Splunk Connect for Syslog: Extending the Platform
- **Support & Community Discussions:** https://splunk-usergroups.slack.com, #splunk-connect-for-syslog
- **Issues or Enhancements:** https://github.com/splunk/splunk-connect-for-syslog/issues

We wish you the best of success with SC4S. Get involved, try it out, ask questions, contribute new data sources, and make new friends!

Splunk Community

https://community.splunk.com

- A free way to connect, learn, have fun, and find success with Splunk
- Ask questions, get answers, and find solutions from passionate experts in the community
- Meet in-person or virtually with like-minded enthusiasts, in your area or by interest
- Search for, vote on, or submit your own ideas for new enhancements for any product or solution



splunk>

Splunk Events

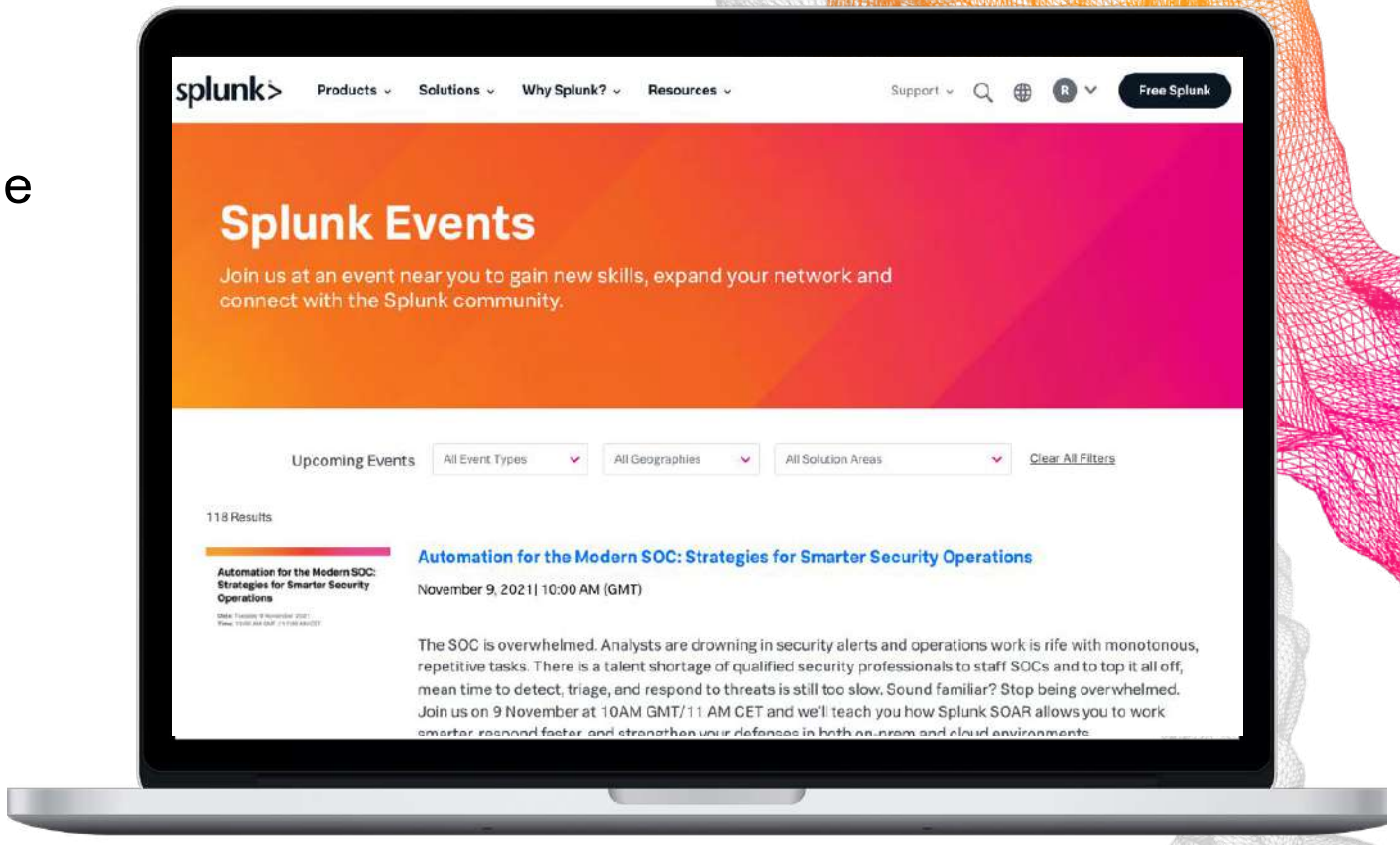
https://events.splunk.com

- Expand your network and connect with the global and local Splunk community

.conf

https://conf.splunk.com

- Join us at .conf24!
- Hundreds of on-demand sessions from product updates to learning new Splunk skills!

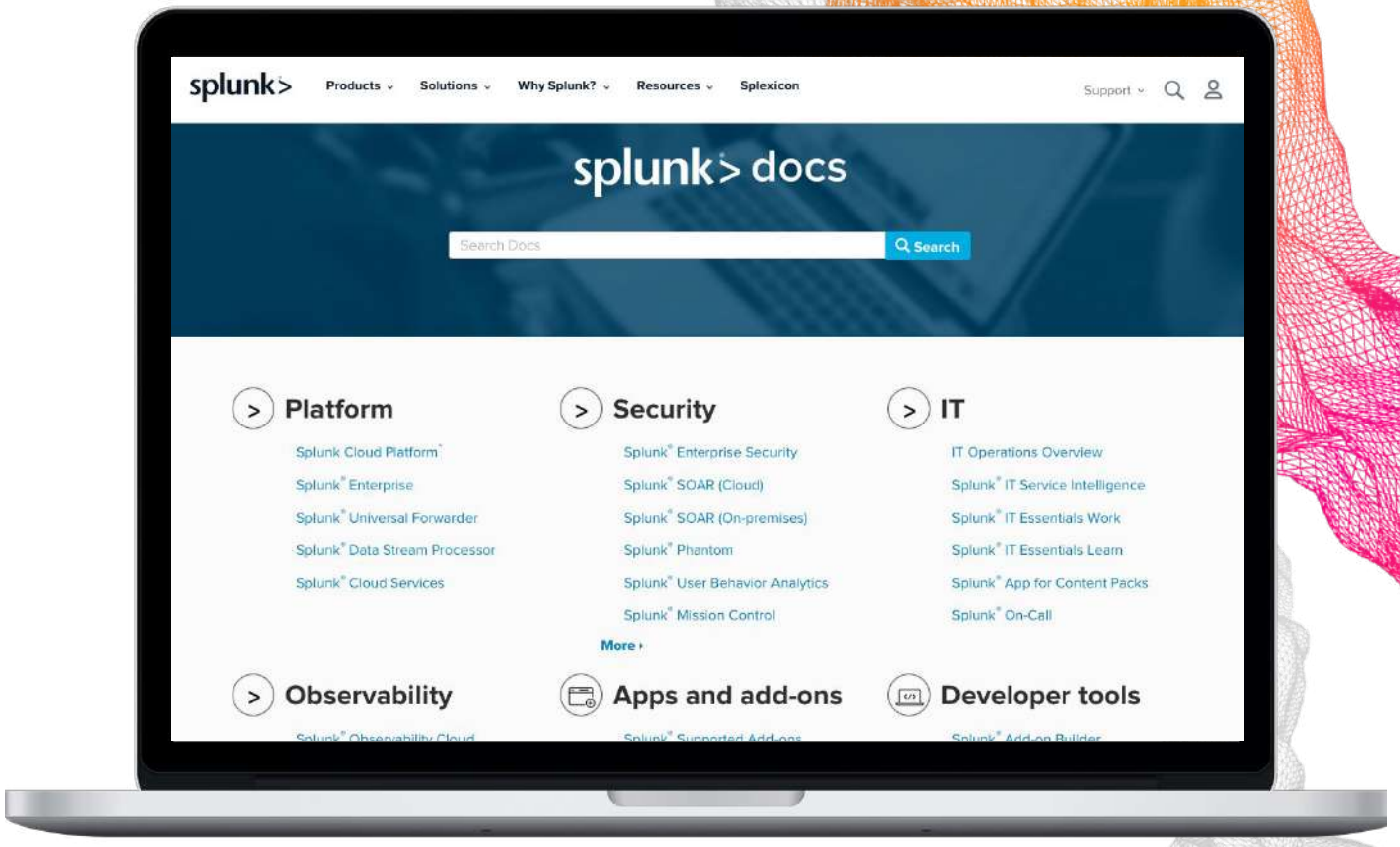


splunk>

Documentation

https://docs.splunk.com

- Search reference for SPL
- Step-by-step tutorials
  Search
  https://splk.it/SplunkSearchTutorial
  Dashboard Studio
  https://splk.it/SplunkDashStudioTutorial
- Product references
- Procedures/guides
- And more!

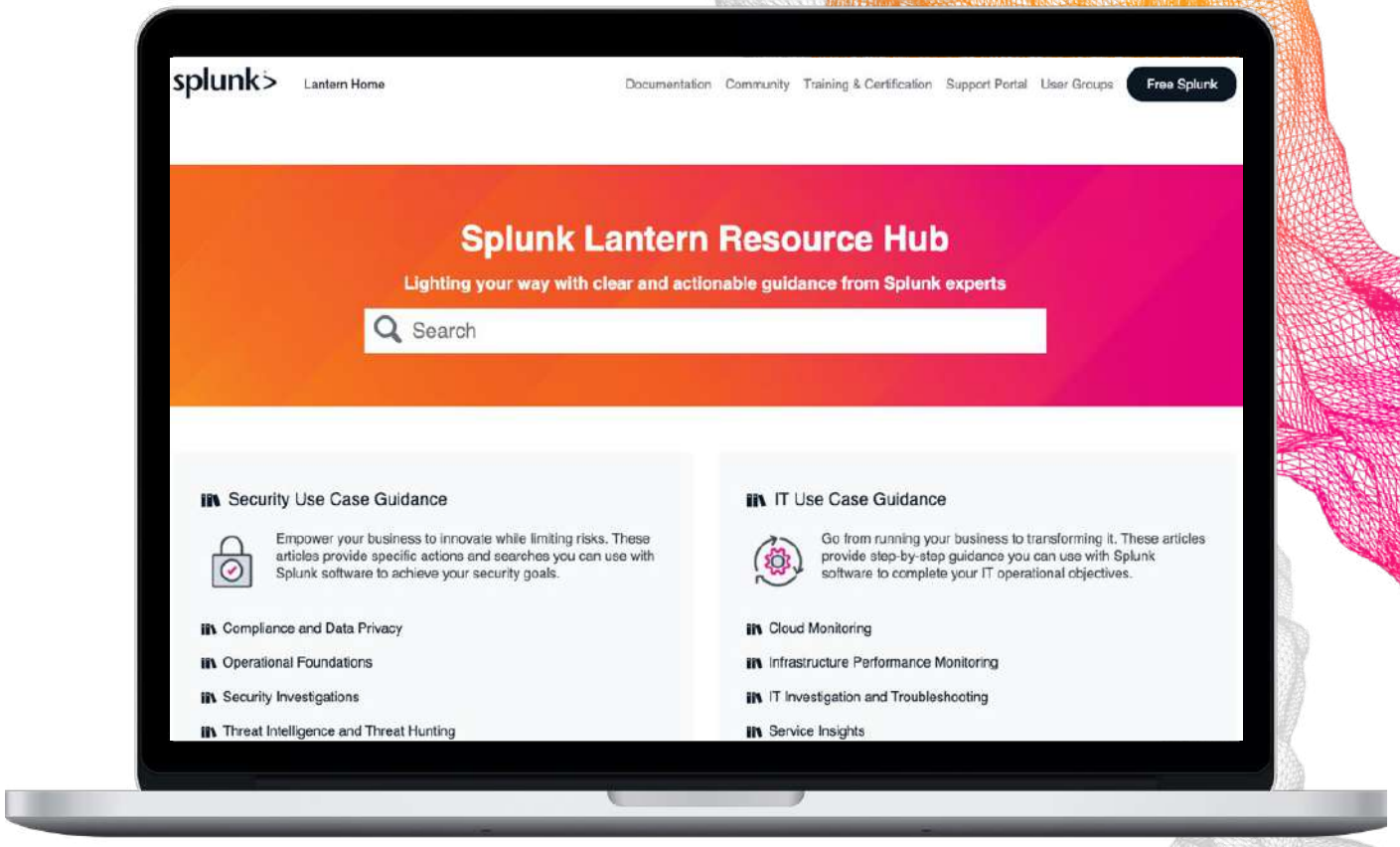


splunk>

Splunk Lantern

https://lantern.splunk.com

- Use case library
- Step-by-step procedures
- Map use cases to data sources
- Splunk Success Framework to realize value across your organisation

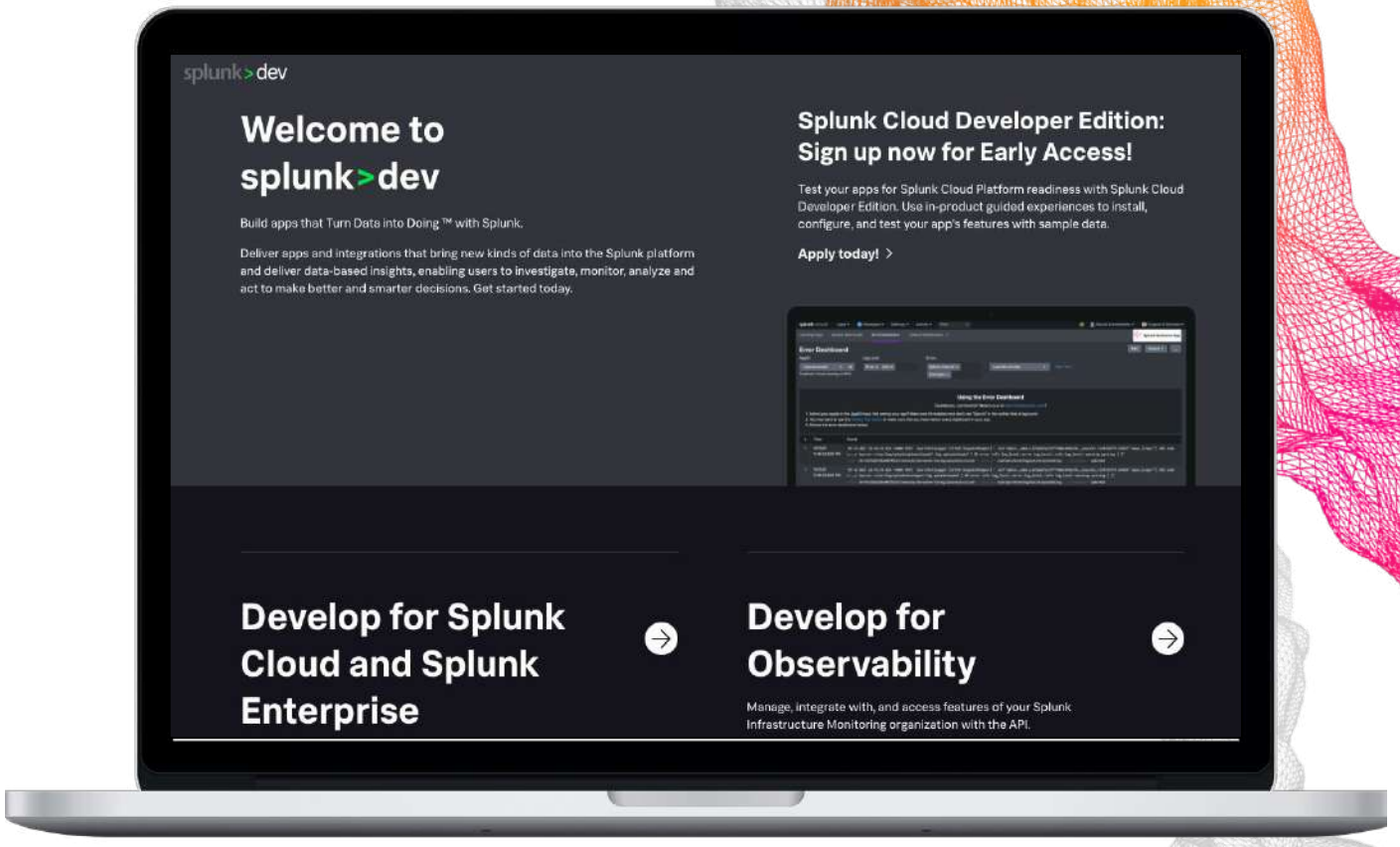


splunk>

Developer Resources

https://dev.splunk.com

- Developer Guide
- API Reference
- Tutorials
- Downloads
  APIs, libraries, tools
- Code examples
- Free Developer licence
- Splunk Cloud Developer Edition
  Test your apps for Splunk Cloud readiness

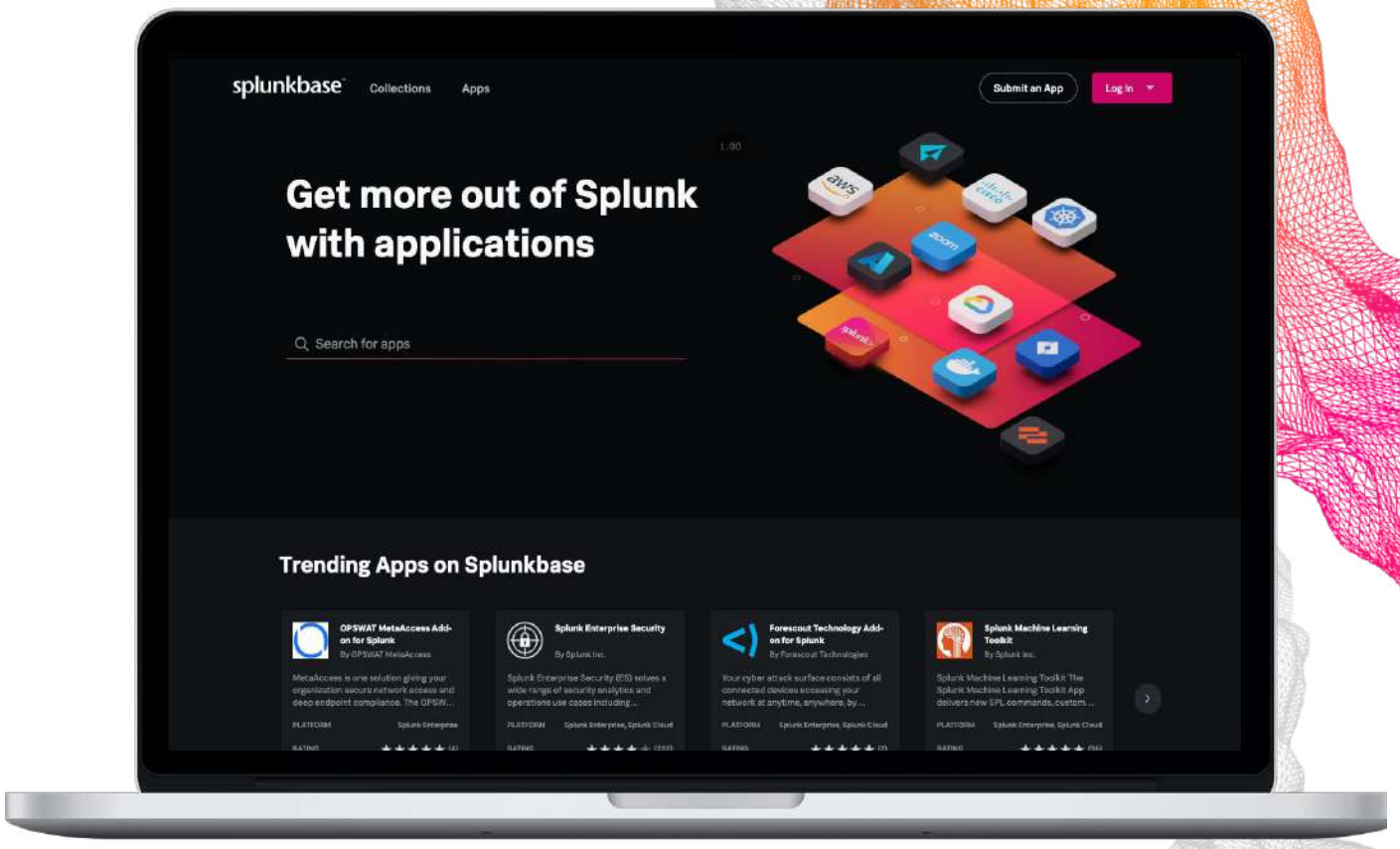


splunk>

Splunk Apps & Add-ons

https://splunkbase.splunk.com/

- 3000+ apps and add-ons
- Pre-built searches, reports, visualisations and integrations for specific use cases and technologies
- Download apps and customise them based on your requirements
- Fast time to value from your data
- Build and contribute your own apps!



Training & Certification

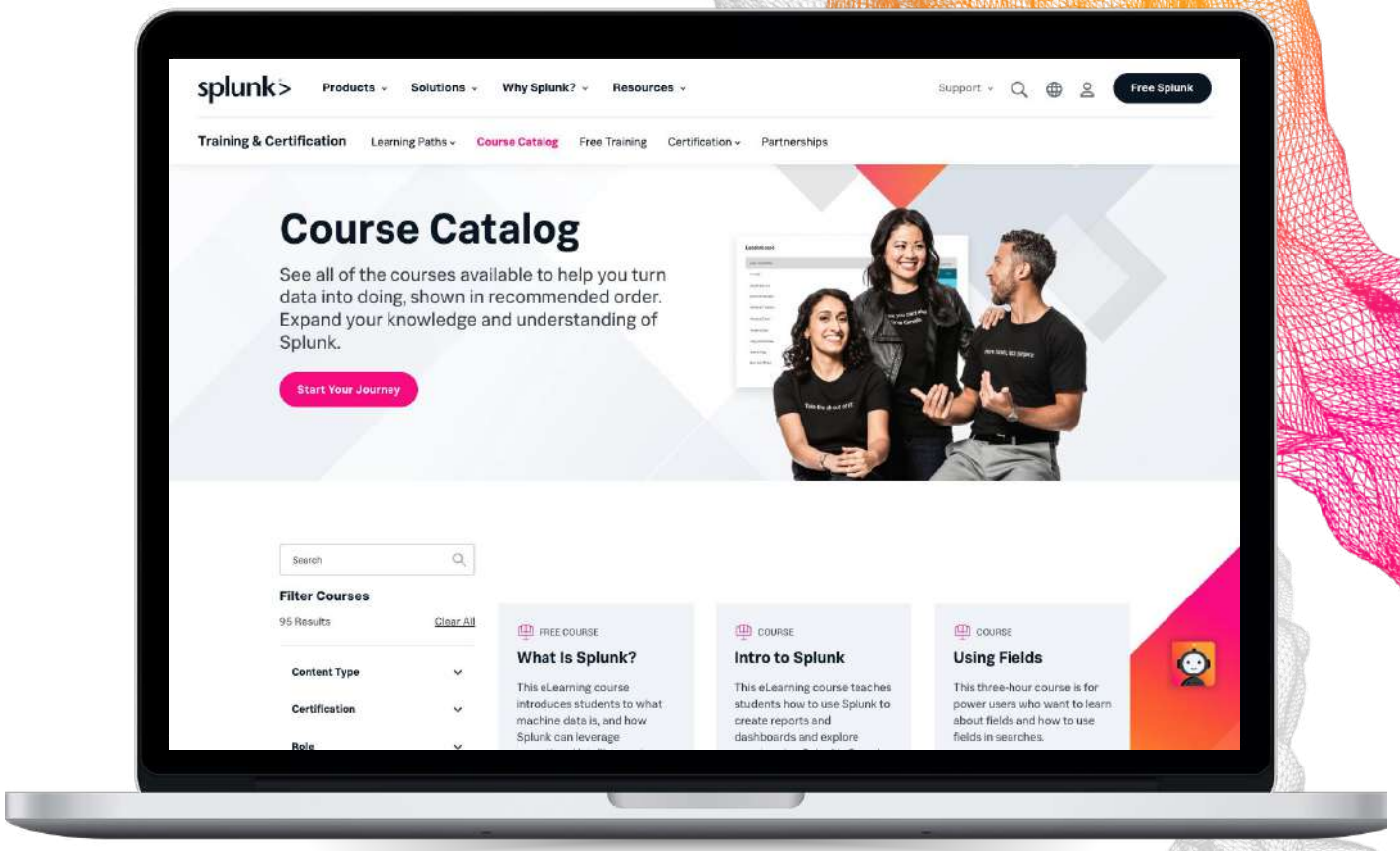
https://splunk.com/training

Online education classes
Instructor-led and self-paced eLearning

Certification tracks for different roles
User, Power User, Admin, Architect and Developer

Splunk Education Rewards
Complete training and receive points that you can redeem for Splunk swag!

Free education!
Free single-subject eLearning courses to kick start your Splunk learning



splunk>

Thank You!