

# Starting Your Splunk Journey – Get Your Data In

PLA1906C

**Ben Marcus**

Sr. Staff IT Engineer | Qualcomm



# Forward-Looking Statements



During the course of this presentation, we may make forward-looking statements regarding future events or plans of the company. We caution you that such statements reflect our current expectations and estimates based on factors currently known to us and that actual events or results may differ materially. The forward-looking statements made in the this presentation are being made as of the time and date of its live presentation. If reviewed after its live presentation, it may not contain current or accurate information. We do not assume any obligation to update any forward-looking statements made herein.

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. Splunk undertakes no obligation either to develop the features or functionalities described or to include any such feature or functionality in a future release.

Splunk, Splunk>, Data-to-Everything, D2E and Turn Data Into Doing are trademarks and registered trademarks of Splunk Inc. in the United States and other countries. All other brand names, product names or trademarks belong to their respective owners. © 2020 Splunk Inc. All rights reserved



# Ben Marcus

Sr. Staff IT Engineer | Qualcomm

<https://www.linkedin.com/in/heybigben>



# Agenda

## Overview

Will discuss common ways of getting data into Splunk and give ideas for example data sources and why different types of sources are useful.

- 1. General Data Onboarding**
- 2. Splunk Universal Forwarder**
- 3. Remote Syslog**
- 4. Splunk HTTP Event Collector (HEC)**
- 5. Scripted Inputs via Universal Forwarder (UF)**
- 6. Modular Inputs**
- 7. Database (DBX) Connect**
- 8. Other transform, routing, onboarding tools**

# General Data Onboarding

 How to interface with data - push vs pull

 What are you going to do with the data

 Sizing


 Timestamp extraction

 Compatible apps

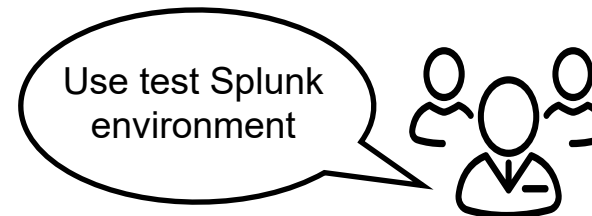
[Splunk Docs - Getting Data in](#)

 Event breaking

 Max length (truncation)

 Host, source, sourcetype, index

 Common Information Model (CIM)  
compatible fields



# Splunk Universal Forwarder (UF)

## Monitor Log Files

### Linux syslogs



- /var/log/syslog, /var/log/authlog, /var/log/sudo.log, /var/log/syslogs/local1-7
- /root/.bash\_history

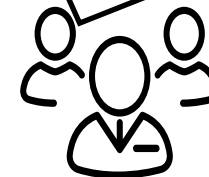
### Application Logs



- Webservers (Apache/Nginx)
- Revision Control
- Database (mysql, postgres)
- DNS
- LDAP
- OS Query (/var/log/osquery/osqueryd.snapshots.log, /var/log/osquery/osqueryd.results.log)

```

/opt/splunkforwarder/etc/apps
  /linuxsyslogs/default/inputs.conf
  /clearcase/default/inputs.conf
  /lsf/default/inputs.conf
  /webserver/default/inputs.conf
  /dns/default/inputs.conf
  /osquery/default/inputs.conf
  /database/default/inputs.conf
  
```



Package "apps"

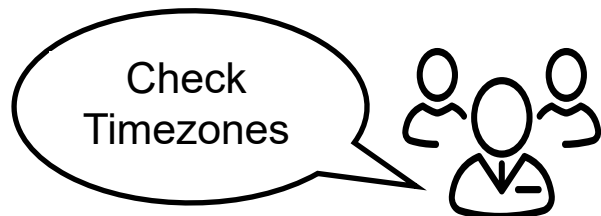
[Splunk Docs - Universal Forwarder Manual](#)

# Remote Syslog

Easiest for appliances and devices where you can't run the Universal Forwarder directly

Appliances send remote syslog via UDP/TCP 514 port

- Firewalls (Fortinet, PaloAlto)
- Routers (Cisco, Arista)
- VMware
- VPN's
- Web proxies, web security gateways
- Load balancers (Citrix Netscaler, F5 BigIP)
- Servers (HP ILO, Dell IDRAC, Supermicro BMC)



Management - Remote Syslog

SNMP Settings AlertMail Remote Syslog

Remote Syslog Settings

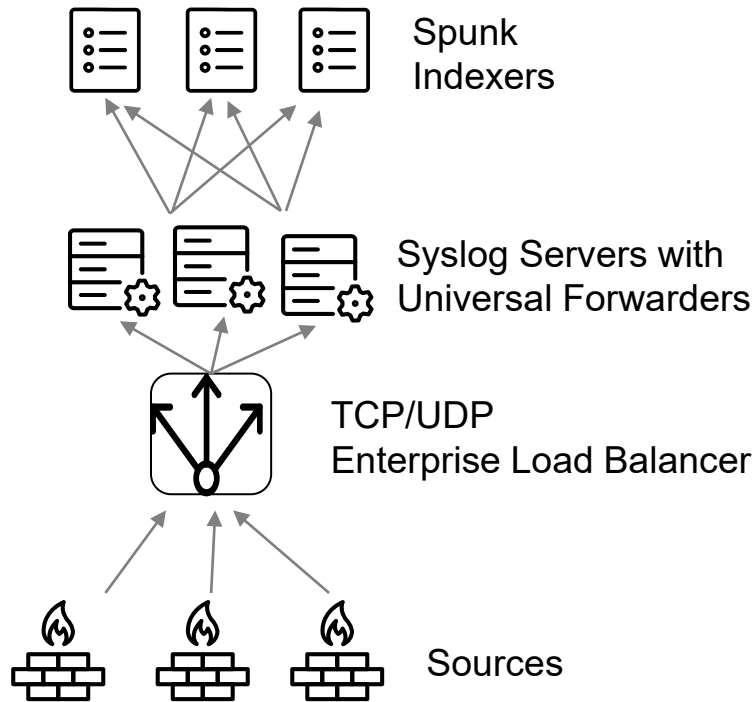
Enable iLO Remote Syslog

Remote Syslog Port

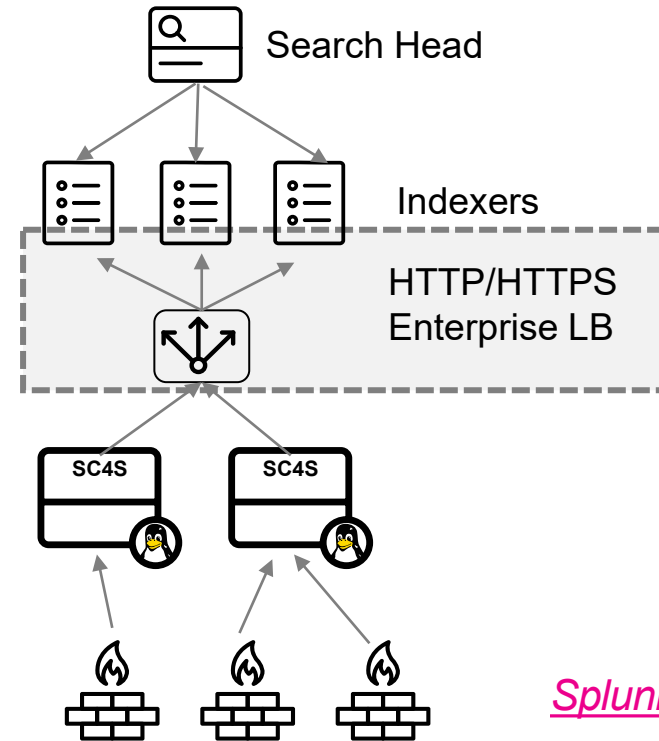
Remote Syslog Server

# Remote Syslog Architectures

## Classic



## Splunk Connect 4 Syslog (SC4S)



[Splunk docs - SC4S](#)

Performant and Scalable Syslog Data Ingest



# Splunk HTTP Event Collector (HEC)

Great method for application developers

Easy for scripts and custom applications to post data to Splunk

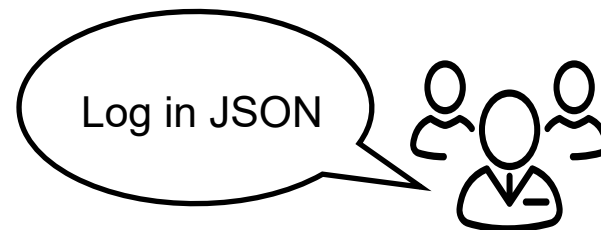
- collectd, telegraf – linux agents send metrics to Splunk HEC.  
Splunk App Infra
- Webhooks (Zoom, Github, Plex)
- Custom - Data center power stats

Collectd.conf

```
<Plugin write_splunk>  
  server hec-prod-splunk.mydomain.com  
  port "443"  
  token "<redacted>"  
  ssl true  
  verifyssl false  
  Dimension "department:corp"  
</Plugin>
```

```
curl -k https://hecserver.yourdomain.com:443/services/collector/event -H "Authorization: Splunk  
secrettoken" -d '{"event": "hello world via Splunk HEC"}'
```

[Splunk Docs - HEC](#)



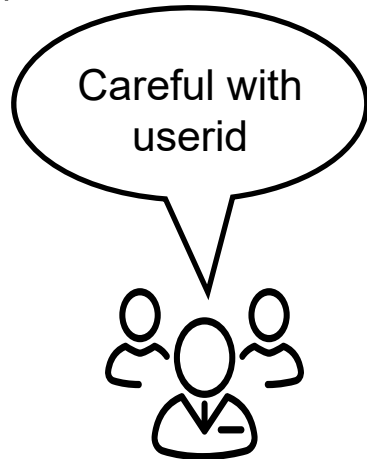
# Scripted Inputs via Splunk UF

Linux commands  
(Splunk \*nix TA)

- top
- ps
- netstat
- uptime
- df
- lsof

Windows commands

- Powershell



```
[script://./bin/nfsstatjson.sh]
interval = 1800
sourcetype = json_nfsmountstat
source = json_nfsmountstat
index = nfsmounts
disabled = 0
```

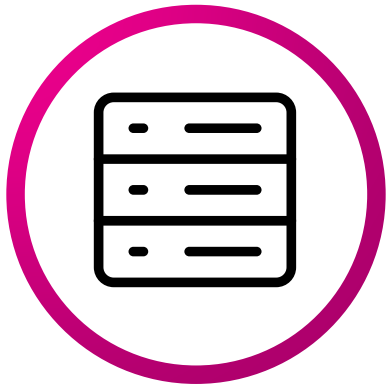
```
[powershell://conninfo]
script="$SplunkHome\etc\apps\myap\bin\Conn.ps1"
schedule = */20 * 9-16 * 1-5
```

# Windows Data

Splunk UF on Windows

[Splunk Docs - Windows UF](#)

## WinEventLog Logs



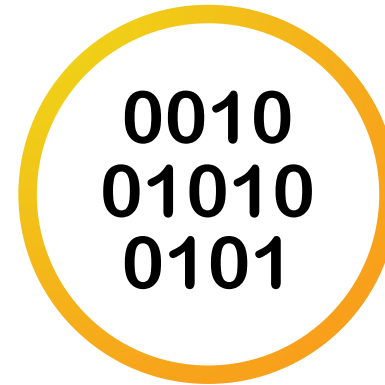
System  
Security  
Application  
Apps and Services

## WinPerfMon Performance Counters



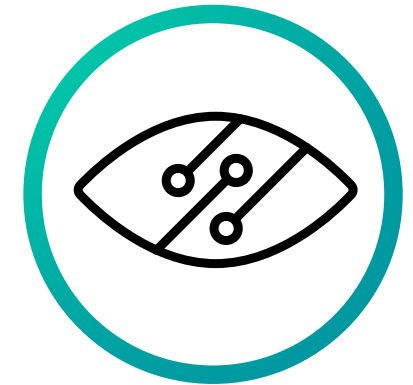
CPU  
Memory  
Network  
LogicalDisk

## WinHostMon Host Properties



OperatingSystem  
Processor  
NetworkAdapter  
Service  
Process  
Disk

## ADMon Active Directory



AD baseline  
AD changes

# Modular Inputs (Splunk Heavy Forwarder)

App or add on extending Splunk Enterprise framework to define a custom input as if it were a native input.

The logo for ServiceNow, featuring the word "servicenow" in a lowercase, sans-serif font.The logo for Carbon Black, with "CARBON" in a smaller, uppercase font above "BLACK" in a larger, bold, uppercase font.The logo for Netskope, featuring a stylized icon of three interconnected nodes above the word "netskope" in a lowercase, sans-serif font.The logo for Digital Shadows, consisting of the words "digital shadows" in a lowercase, sans-serif font inside a dark rectangular box.

App has special input to pull/obtain data from cloud or via app API

REST API modular input – obtain data via remote application REST endpoint.  Phantom



[Splunk Apps - Splunkbase](#)



# Cloud Connectors

## AWS

**Splunk app AWS**  
**Kinesis Firehose**  
(sends directly to Splunk HEC)  
**Splunk Universal Forwarder**



CloudTrail, CloudWatch  
(metrics/logs/VPC flow),  
Config, Billing, ALB/NLB  
(load balancers),  
Cloudfront, S3

## Azure

**Splunk Add-on for Microsoft**  
**Cloud Services,**  
**Splunk Add-on for Office 365**  
**Splunk HEC**



Azure AD User Audit,  
Storage Accounts, Virtual  
Machines, Subscriptions,  
Billing

## GCP

**Splunk Add on for Google**  
**Cloud**  
**Splunk HEC**



G Suite Admin Console,  
Stackdriver Logging, Cloud  
Security, Command Center,  
GCP, GKE and Anthos,  
Stackdriver Metrics

# Splunk Database Connect

- Database connect–dbx v3 type inputs
  - Most JDBC/ODBC drivers will work
- SQL input – periodically run query and index results
- Tail table with unique key
- Enrichment via Splunk database lookups

```
| dbxlookup lookup="userdept" | table host, user, title, dept
```

The Oracle logo, consisting of the word "ORACLE" in a bold, red, sans-serif font.

PostgreSQL

The Vertica logo, consisting of the word "VERTICA" in a bold, black, sans-serif font.

[Splunk Docs - DBX](#)

# Other Transform, Routing, Onboarding Tools

- Splunk Data Stream Processor – DSP
- Fluentd (fluentd.org) – Open source data collector/router
- Kafka (kafka.apache.org) – Distributed streaming platform
  - Event Streaming Platform for Kafka – confluent.io
- Cribl (cribl.io) – LogStream processor, Ingest at scale
  - Universal receiver: Receive data from multiple sources, send to Splunk
  - Pull data at high scale from Kafka, S3, Azure Event hubs, and other pull sources
  - Receive data on standard protocols such as Syslog and SNMP Traps
  - Batch ingest of historic event logs, on file systems or in S3



[Know More – Data Stream Processor](#)

# Other Data Sources

- Splunk Stream (packet data)
- Netflow (summary metadata on connections)







# Thank You

Please provide feedback via the

**SESSION SURVEY**

