SIEM the Skyscanner Way

Integrating Splunk into our security practice

Marc Santamaria Ortega
Senior Security Engineer | Skyscanner
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 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
Key Takeaways

1. Splunk versatility allows to customize it to your needs

2. Infrastructure as Code can prove very useful on a Splunk environment

3. Automation is very helpful in alerting and monitoring
1) Splunk analytics as code
   Infrastructure as Code for Splunk Cloud searches

2) Automating alert processing
   Automating processing of Splunk alerts via AWS Lambda

3) How we monitor our alert pipeline
   TTD/TTR dashboards, error monitoring, sources and alert health

4) KVStore and log enriching
   Correlating company information with Splunk resources

5) Favorite controls
Splunk Analytics as Code

Why?

In Skyscanner we use Infrastructure as Code (IaC) to manage our Cloud resources

• Doing the same in Splunk allows us to integrate our best practices into it
Splunk Analytics as Code

Why?

In Skyscanner we use Infrastructure as Code (IaC) to manage our Cloud resources
• Doing the same in Splunk allows us to integrate our best practices into it

<table>
<thead>
<tr>
<th>Benefit</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Version control</td>
<td>Keep historic of changes with descriptions or tickets that drove the change</td>
</tr>
</tbody>
</table>
Splunk Analytics as Code

Why?

In Skyscanner we use Infrastructure as Code (IaC) to manage our Cloud resources

• Doing the same in Splunk allows us to integrate our best practices into it

<table>
<thead>
<tr>
<th>Benefit</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Version control</td>
<td>Keep historic of changes with descriptions or tickets that drove the change</td>
</tr>
<tr>
<td>Change management</td>
<td>Have peer reviews on any change and getting automatically notified of changes to review</td>
</tr>
</tbody>
</table>
Splunk Analytics as Code

Why?

In Skyscanner we use Infrastructure as Code (IaC) to manage our Cloud resources
• Doing the same in Splunk allows us to integrate our best practices into it

<table>
<thead>
<tr>
<th>Benefit</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Version control</td>
<td>Keep historic of changes with descriptions or tickets that drove the change</td>
</tr>
<tr>
<td>Change management</td>
<td>Have peer reviews on any change and getting automatically notified of changes to review</td>
</tr>
<tr>
<td>Resilience</td>
<td>Being able to easily reproduce/rollback any change</td>
</tr>
</tbody>
</table>
Splunk Analytics as Code

Why?

In Skyscanner we use Infrastructure as Code (IaC) to manage our Cloud resources
• Doing the same in Splunk allows us to integrate our best practices into it

<table>
<thead>
<tr>
<th>Benefit</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Version control</td>
<td>Keep historic of changes with descriptions or tickets that drove the change</td>
</tr>
<tr>
<td>Change management</td>
<td>Have peer reviews on any change and getting automatically notified of changes to review</td>
</tr>
<tr>
<td>Resilience</td>
<td>Being able to easily reproduce/rollback any change</td>
</tr>
<tr>
<td>Auditability</td>
<td>Have a detailed audit trace of all changes</td>
</tr>
</tbody>
</table>
Splunk Analytics as Code

Why?

In Skyscanner we use Infrastructure as Code (IaC) to manage our Cloud resources
- Doing the same in Splunk allows us to integrate our best practices into it

<table>
<thead>
<tr>
<th>Benefit</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Version control</td>
<td>Keep historic of changes with descriptions or tickets that drove the change</td>
</tr>
<tr>
<td>Change management</td>
<td>Have peer reviews on any change and getting automatically notified of changes to review</td>
</tr>
<tr>
<td>Resilience</td>
<td>Being able to easily reproduce/rollback any change</td>
</tr>
<tr>
<td>Auditability</td>
<td>Have a detailed audit trace of all changes</td>
</tr>
<tr>
<td>Stability</td>
<td>Having a source of truth for our alerts which replaces any manual change</td>
</tr>
</tbody>
</table>
Splunk Analytics as Code

Why?

In Skyscanner we use Infrastructure as Code (IaC) to manage our Cloud resources

- Doing the same in Splunk allows us to integrate our best practices into it

<table>
<thead>
<tr>
<th>Benefit</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Version control</td>
<td>Keep historic of changes with descriptions or tickets that drove the change</td>
</tr>
<tr>
<td>Change management</td>
<td>Have peer reviews on any change and getting automatically notified of changes to review</td>
</tr>
<tr>
<td>Resilience</td>
<td>Being able to easily reproduce/rollback any change</td>
</tr>
<tr>
<td>Auditability</td>
<td>Have a detailed audit trace of all changes</td>
</tr>
<tr>
<td>Stability</td>
<td>Having a source of truth for our alerts which replaces any manual change</td>
</tr>
<tr>
<td>Contextuality</td>
<td>Include info that cannot be easily added like tool related, links to internal docs of the log source or its procedure, …</td>
</tr>
</tbody>
</table>
Splunk Analytics as Code

Why?

In Skyscanner we use Infrastructure as Code (IaC) to manage our Cloud resources
• Doing the same in Splunk allows us to integrate our best practices into it

<table>
<thead>
<tr>
<th>Benefit</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Version control</td>
<td>Keep historic of changes with descriptions or tickets that drove the change</td>
</tr>
<tr>
<td>Change management</td>
<td>Have peer reviews on any change and getting automatically notified of changes to review</td>
</tr>
<tr>
<td>Resilience</td>
<td>Being able to easily reproduce/rollback any change</td>
</tr>
<tr>
<td>Auditability</td>
<td>Have a detailed audit trace of all changes</td>
</tr>
<tr>
<td>Stability</td>
<td>Having a source of truth for our alerts which replaces any manual change</td>
</tr>
<tr>
<td>Contextuality</td>
<td>Include info that cannot be easily added like tool related, links to internal docs of the log source or its procedure, …</td>
</tr>
<tr>
<td>Usability</td>
<td>Have more control over our 225+ Splunk scheduled searches, allowing for an easier review of them</td>
</tr>
</tbody>
</table>
Splunk Analytics as Code

How?

Splunk API is very powerful and allows to check/modify many configurations
- [https://docs.splunk.com/Documentation/Splunk/latest/RESTREF/RESTsearch#saved.2Fsearches](https://docs.splunk.com/Documentation/Splunk/latest/RESTREF/RESTsearch#saved.2Fsearches)

We use GitHub as our code repository and store our scheduled searches in separate files with a folder per each Project/Source

We have defined a script as part of our Splunk IaC that for each Splunk scheduled search
- Parses the template defined
- Encodes any needed field
- Create/updates the scheduled search in Splunk Cloud
- Updates permissions of the scheduled search
Our alert template allows to define any Splunk parameter we need and also to include custom ones.
Versions over time of a search

Requests to make changes on searches

Classification of searches
Modification of a scheduled search

Slack notification of repository change
Automating Alert Processing

Why?

Slack is our standard way of managing requests

We have some needs that are difficult to cover with Splunk by default Slack integrations

• Show as many results as Slack message limit allows
• Capability to process alerts from Slack and reduce size without losing access to results
• Create a JIRA ticket from a Slack message
• Include documentation to internal procedures, tool that generates the source log and other relevant info
• Do pre-processing in specific alerts to extract information from other systems or automate analysis
Automating Alert Processing

How?

Slack allows to create apps that bring interaction to messages
  • https://api.slack.com/messaging/interactivity

AWS Lambda allows us to create a serverless function with an API
  • https://docs.aws.amazon.com/lambda/latest/dg/welcome.html

We have defined a Python Lambda that does the following:
  • Processes requests from both Splunk and Slack
  • For each Splunk alert, retrieves all results from Splunk API and depending on each case does automated checks, creates a ticket, updates info on another tool, sends a Slack message, …
  • We classify alerts for the main source that drives them and add context info to Slack messages
Automating Alert Processing

Diagram

- Splunk
  - Splunk Cloud
  - Splunk ES
  - Heavy Forwarder
- AWS
  - API Gateway
  - SNS
  - Lambda
  - AWS API
- JIRA
- Slack

Scheduled search with action webhook

Retrieve results from Splunk API or modify KV Stores

Create JIRA ticket or check ticket status

Send alerts to Slack

Click Slack button

Boto3 API calls
JIRA tickets created by the Lambda

Example of different functions/integration s of the Lambda
How we Monitor our Alert Pipeline

Systems always can fail and with alerts depending on different systems we need to ensure that issues in all steps are detected (log source system, Splunk, Lambda, …)

- Detect if logs are arriving as expected into Splunk (both by host or sourcetype)
- Detect issues with Splunk (processing or integration errors, unexpected changes, …)
- Detect issues with the Lambda (timeouts, errors/exceptions caught, alerts not processed, …)

We have also built a Splunk dashboard that correlates Splunk audit information and Lambda logs

- Offers visibility on TTD and TTR of alerts (Time to detect and time to response)
- Allows us to do quarterly alert reviews knowing which alerts raise more often, which ones are ignored or false positives, which ones take longer to process, …
Monitoring searches in Splunk

- Splunk - Alert not processed by Lambda
- Splunk - Alert not processed by Splunk
- Splunk - Apps pending to be updated
- Splunk - DBConnect error
- Splunk - DBConnect query failed
- Splunk - Delete permission changes
- Splunk - Error in Python scripts
- Splunk - Error in Sck Lambda
- Splunk - Error in alert execution
- Splunk - Host not sending logs in last 2h
- Splunk - License of 100GB surpassed
- Splunk - Multiple failed login attempt
- Splunk - Multiple failed login attempt followed by successful one
- Splunk - New host sending logs
- Splunk - New sourcetype sending logs
- Splunk - Possible logs deleted in Splunk Cloud
- Splunk - SecOps Lambda has timeout
- Splunk - Skycraper fewer in logs
- Splunk - Sourcetype logs not arriving in last 2h
- Splunk - Summary of alerts not managed by GF yesterday
- Splunk Info - Alert changes
- Splunk Info - App changes
- Splunk Info - Log storage Buckets deleted
- Splunk Info - Object changes
- Splunk Info - Server restarted

Alert review dashboard

TTD/TTR dashboard
Alert review template

Attendees
- Type attendees here using "@"

Discussion items
Review:
- Alerts with summary dashboard
  - [link]
- Suggestions on Confluence page
  - Alert Review Suggestions
- Review pending tasks from previous Alert Reviews
- Review top hosts/sourcetypes not sending logs
  - [link]
- Review current temporary exceptions
  - [link]

Gather feedback on:
- Which alerts create more noise
- Reason to not action some alerts
- Alerts which management/procedures is not clear
- Comments on Splunk alerts

Decide
- Alerts that should be disabled/changed
- Alerts that we should work on
- Alerts that will be migrated first into Hashi for automation
  - ...

Action items
General tasks
- [Task template]

Alerts to be created
- [Task template]

Alerts to be modified
- [Task template]

Alerts to be deleted
- [Task template]

New VictorOps alerts
- [Task template]

Alerts to be automated
- [Task template]

Exceptions
- [Task template]

Procedures to be modified
- [Task template]
KVStore and Log Enriching

We ingest into Splunk data from multiple sources but some of it is specially useful to include in KV Stores to enrich Splunk alerts and dashboards on search time

- Skyscanner office IP ranges
- Skyscanner Cloud accounts info and IP ranges
- Ownership of Projects by Skyscanner teams
- Status of tickets opened
- Old employees
- Temporary exceptions in alerts

We automatically manage our KVStores via Splunk API and keep them up to date

- https://dev.splunk.com/enterprise/docs/developapps/manageknowledge/kvstore/usetherestapitomanagekv/
KVStore and Log Enriching

Diagram

- **GitHub**
  - API
  - Obtain info of Office IPs and internal ranges

- **AWS**
  - API
  - Obtain list of public IPs
  - `describe_network_interfaces`
  - Obtain list of accounts
  - `list_accounts`

- **Internal tools**
  - API
  - Obtain project ownership

- **Azure**
  - API
  - Obtain employee information

- **JIRA**
  - API
  - Obtain status of tickets

- **Scheduled Scripts**
  - Python
  - Update KVStore entries
  - `storage/collections/data/{collection}/{id}`

- **Splunk**
  - API
  - `splunk>`
Favorite Controls

VictorOps

For Out of Hours (OOH) notifications in Skyscanner we use VictorOps

Splunk and VictorOps can be integrated in several ways

- Via Splunk app for VictorOps
- Via VictorOps API
- **Via mails to VictorOps**
  - `{Tenant}`+`{Routing Key}`@alert.victorops.com

Allows to notify OOH support in an easy way with all the data needed to handle the alert
Email: Splunk Alert: CloudTrail - EMR get-shell YARN detected

SUMMARY


View results in Splunk: https://splunk.com/443/app/search/app/splunksplit

96 drw:who YARN eu-west-1 get-shell
j-GGTASWBoD5G

If you believe you've received this email in error, please see your Splunk administrator. splunk > the engine for machine data

state_start_time: 5:18 PM
subject: Splunk Alert: CloudTrail - EMR get-shell YARN detected
timestamp: 5:18 PM
VO_ALERTRCV_TIME: 5:18 PM
VO_ALERTTYPE: SERVICE
VO_MONITOR_TYPE: 8
VO_ORGANIZATION_ID: globalops
VO_UUID: 0a851a1e5-9c76-458b-b5a5-c9140b286cab

Email Splunk Alert: CloudTrail - EMR get-shell YARN detected


Observed by: gongkaplanzai
Favorite Controls

DMARC

DMARC allows to protect your email domains and prevent unauthorized usage for phishing, spoofing and other scams.

Managing and monitoring multiple email domains usually is a hassle or too expensive.

Splunk allows you to easily ingest the data and prepare your own dashboards and controls.

- TA-dmarc add-on for Splunk - [https://splunkbase.splunk.com/app/3752/](https://splunkbase.splunk.com/app/3752/)
<table>
<thead>
<tr>
<th>DMARC - All domains</th>
<th>SPF policy evaluated</th>
<th>DKIM evaluated</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>127</strong></td>
<td><strong>29,292,175</strong></td>
<td></td>
</tr>
</tbody>
</table>

**DMARC policy applied**

- Rejected
- Pass

**SPF results**

- Unknown
- Temporarily failed
- Permanent

**DKIM results**

- Unknown
- Transient
- Permanent
What’s Next with Splunk in Skyscanner?

Improve our Operational Threat Intelligence with Splunk Enterprise Security

Migrate more Splunk configurations to Infrastructure as Code

Replace the Lambda that processes Splunk Alerts with a microservice shared between all the Security members to unify automations and reuse integrations

Correlate information from our Vulnerability DBs and scanners with our inventories of both servers and workstations to revamp our Vulnerability Management
## Links Related with the Presentation

<table>
<thead>
<tr>
<th>Application</th>
<th>Links</th>
</tr>
</thead>
</table>
| Splunk      | https://docs.splunk.com/Documentation/Splunk/latest/RESTREF/RESTsearch#saved.2Fsearches  
             | https://docs.splunk.com/Documentation/Splunk/latest/RESTUM/RESTusing#Access_Control_List  
             | https://dev.splunk.com/enterprise/docs/developapps/manageknowledge/kvstore/useherestapitomanagekv/  
             | https://splunkbase.splunk.com/app/3752/ |
| AWS         | https://docs.aws.amazon.com/apigateway/latest/developerguide/welcome.html  
             | https://docs.aws.amazon.com/sns/latest/dg/welcome.html  
             | https://docs.aws.amazon.com/lambda/latest/dg/welcome.html |
| Slack        | https://api.slack.com/start  
             | https://api.slack.com/messaging/interactivity |
| JIRA         | https://developer.atlassian.com/server/jira/platform/rest-apis/ |
| GitHub       | https://developer.github.com/v3/ |
                     | https://docs.microsoft.com/en-us/powershell/module/addsadministration/get-adobject |
Thank You

Please provide feedback via the SESSION SURVEY