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.

Getting the Most Out of Splunk® Enterprise, OCSF, and Amazon Security Lake

SEC1744B

Tom Smit

Principal Security Strategist | Splunk

James Brodsky

GVP, Global Security Strategists | Splunk

splunk> .conf23

Tom Smit

Principal Security Strategist | Splunk

James Brodsky

GVP, Global Security Strategists | Splunk

splunk> .conf23

Agenda

1) What is OCSF?
   – Why is it important?
2) What is Amazon Security Lake
3) Data Lakes
4) Splunk with Amazon Security Lake
5) Other Things

What is OCSF?

Hint: It's not a product.

Framework used to define common data schema for data

Vehicle to drive an open, vendor agnostic standard, like STIX

Initial 18 enterprises -> 60+ orgs contributing

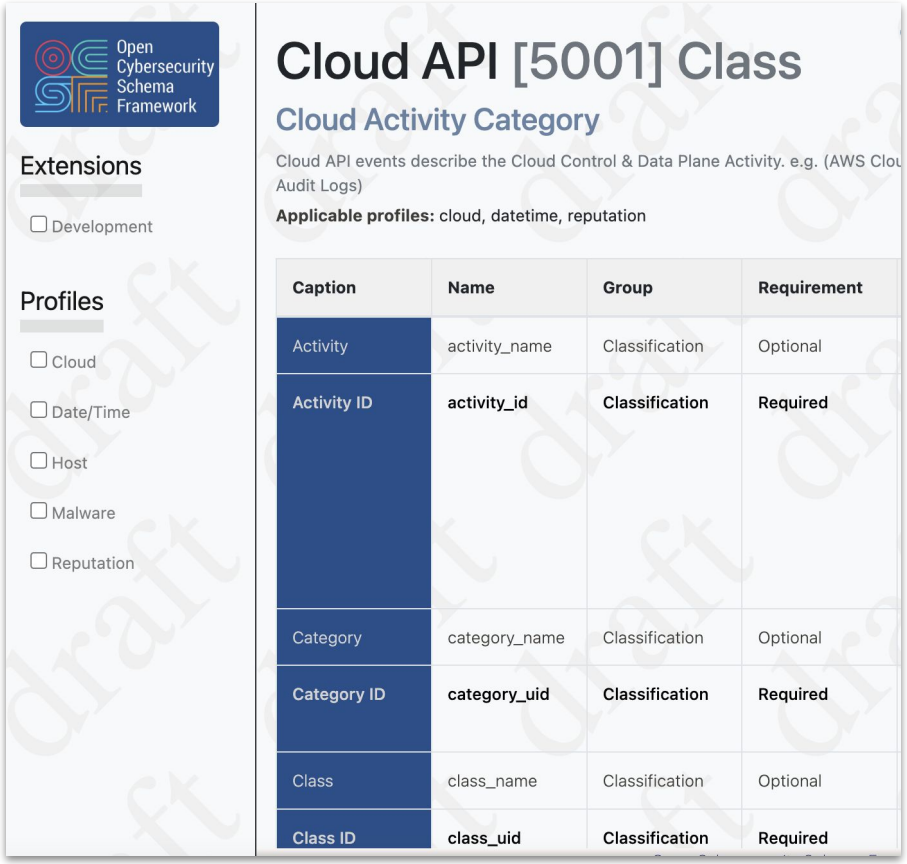
Open Cybersecurity Schema Framework

OCSF is an open source project on GitHub, where individuals and organizations can collaborate and contribute to framework development.

Disclaimer: all logos displayed are the intellectual property of their respective owners

What are some cool things about OCSF?

- Not “owned” by anyone like CEF or CIM were.
  - Agnostic to product, ETL, storage, ingest methods
- “Core” and “Extended” schemas allow for customization
  - Even past security?
  - Adding in MITRE ATT&CK mapping?
  - “Profiles” add capability to share data across classes
    - “Malware” profile might add ATT&CK info and process info to “System Activity” class



https://ocsf.io

What are some cool things about OCSF?

- Lots of support
  - Splunk, AWS, Symantec, Trend, Sumo, Okta, Zscaler, Crowdstrike, PANW, Dtex, Jupiter One, Tanium, IronNet, IBM, Securonix, Rapid7, SFDC
- Less possibility for “abuse” (Data Typing, UID for event classes)
- It seems like folks are super interested (analyst whitepapers)

Open Cybersecurity Schema Framework

Extensions

- Development

Profiles

- Cloud
- Date/Time
- Host
- Malware
- Reputation

Cloud API [5001] Class

Cloud Activity Category

Cloud API events describe the Cloud Control & Data Plane Activity. e.g. (AWS Clou Audit Logs)

Applicable profiles: cloud, datetime, reputation

Caption	Name	Group	Requirement
Activity	activity_name	Classification	Optional
Activity ID	activity_id	Classification	Required
Category	category_name	Classification	Optional
Category ID	category_uid	Classification	Required
Class	class_name	Classification	Optional
Class ID	class_uid	Classification	Required

https://ocsf.io

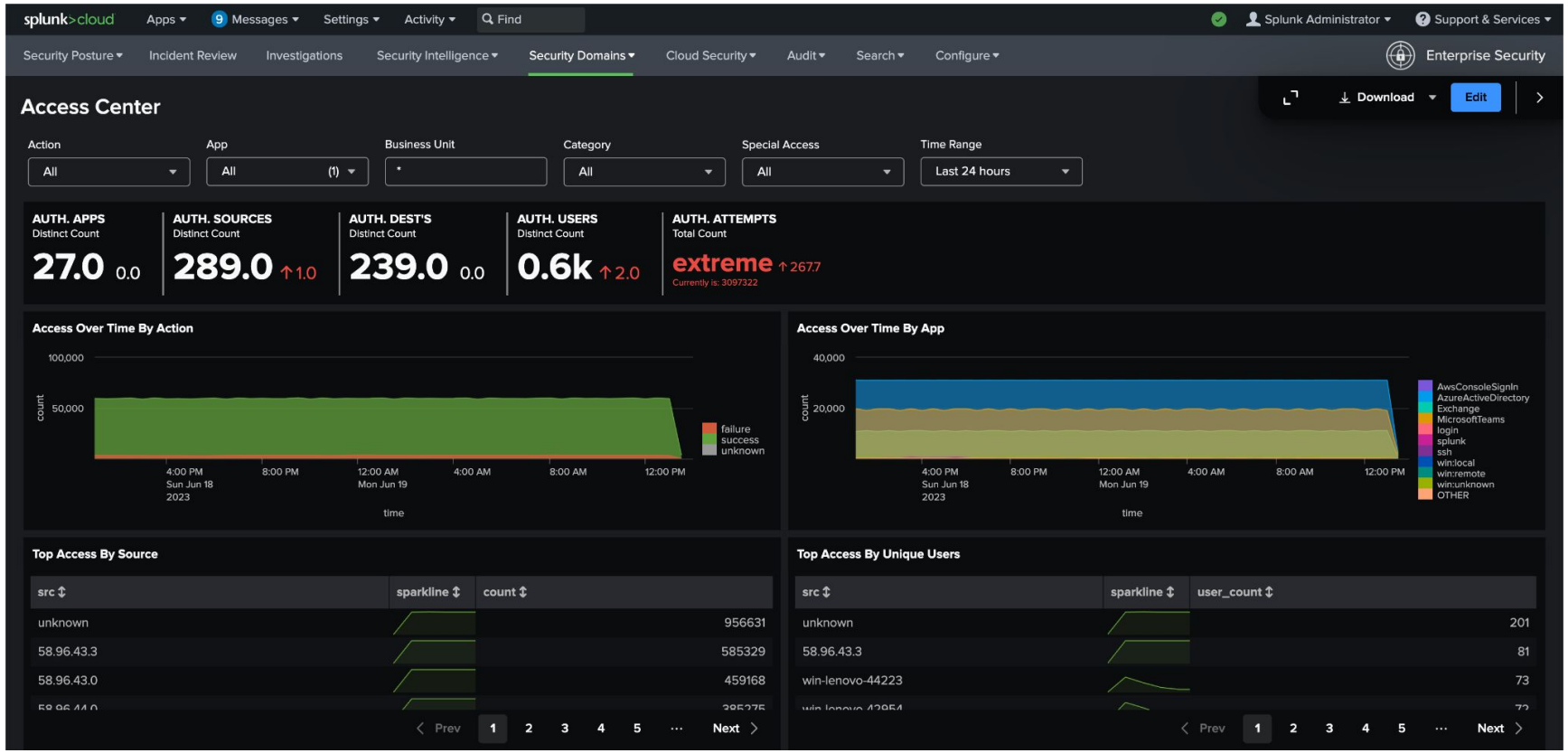
And some unknowns?

- Will vendors jump on the bandwagon and release OCSF-compatible feeds?
- What happens to all the data in the native feed?
- Where IS ATT&CK mapping anyway?
- Will practitioners do their own mappings, and then what happens when official mappings occur?
- What happens when the schema of OCSF changes?
- Will processing shift to the edge and how heavy is that?
- Storing schema along with every event could increase ingest?

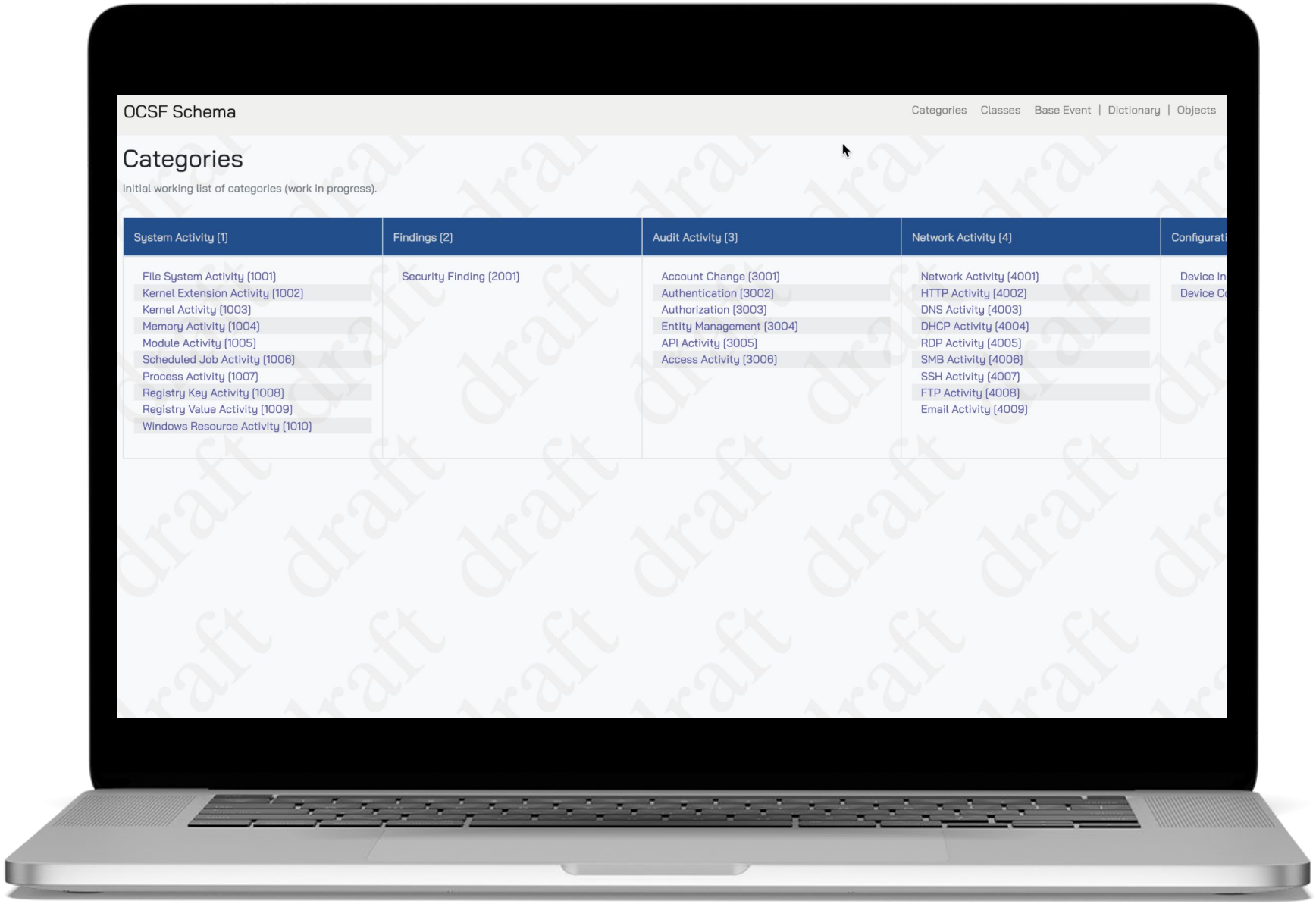
And the big one….

Will Splunk get rid of CIM??

Currently in beta, the OCSF -> CIM translator

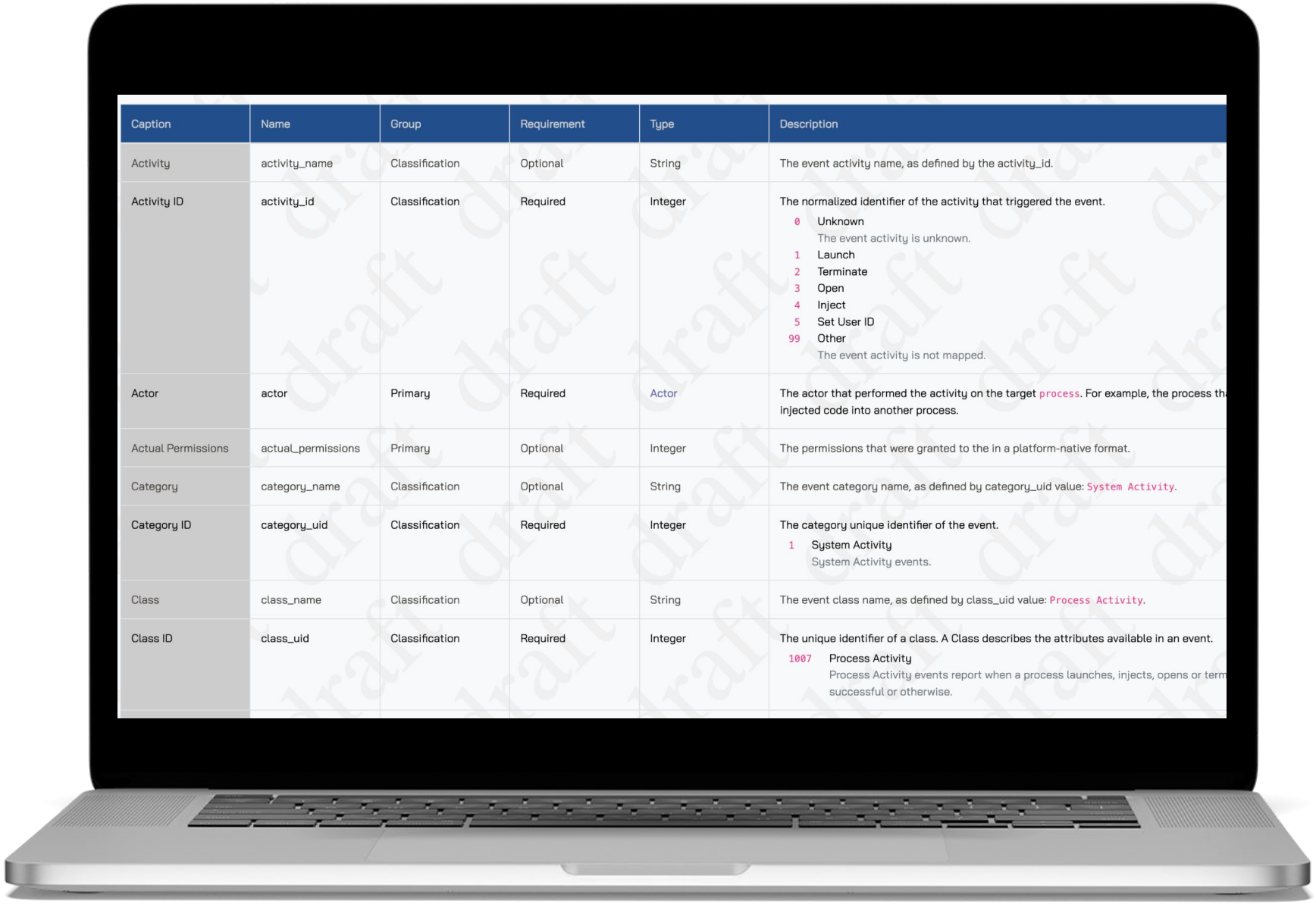


OCSF Browser



splunk> .conf23

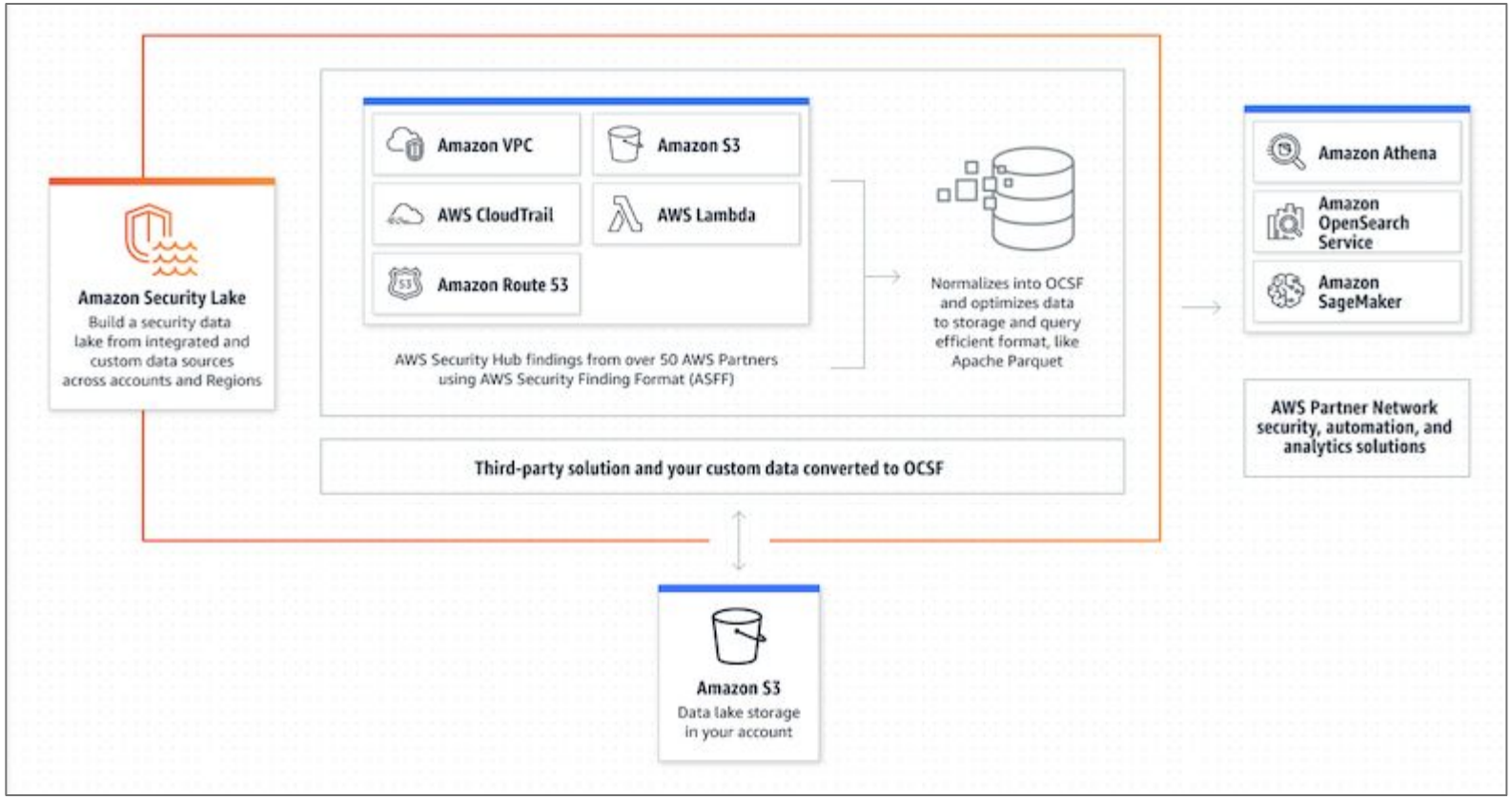
OCSF Browser



Caption	Name	Group	Requirement	Type	Description
Activity	activity_name	Classification	Optional	String	The event activity name, as defined by the activity_id.
Activity ID	activity_id	Classification	Required	Integer	The normalized identifier of the activity that triggered the event. 0 Unknown The event activity is unknown. 1 Launch 2 Terminate 3 Open 4 Inject 5 Set User ID 99 Other The event activity is not mapped.
Actor	actor	Primary	Required	Actor	The actor that performed the activity on the target process. For example, the process th injected code into another process.
Actual Permissions	actual_permissions	Primary	Optional	Integer	The permissions that were granted to the in a platform-native format.
Category	category_name	Classification	Optional	String	The event category name, as defined by category_uid value: System Activity.
Category ID	category_uid	Classification	Required	Integer	The category unique identifier of the event. 1 System Activity System Activity events.
Class	class_name	Classification	Optional	String	The event class name, as defined by class_uid value: Process Activity.
Class ID	class_uid	Classification	Required	Integer	The unique identifier of a class. A Class describes the attributes available in an event. 1007 Process Activity Process Activity events report when a process launches, injects, opens or term successful or otherwise.

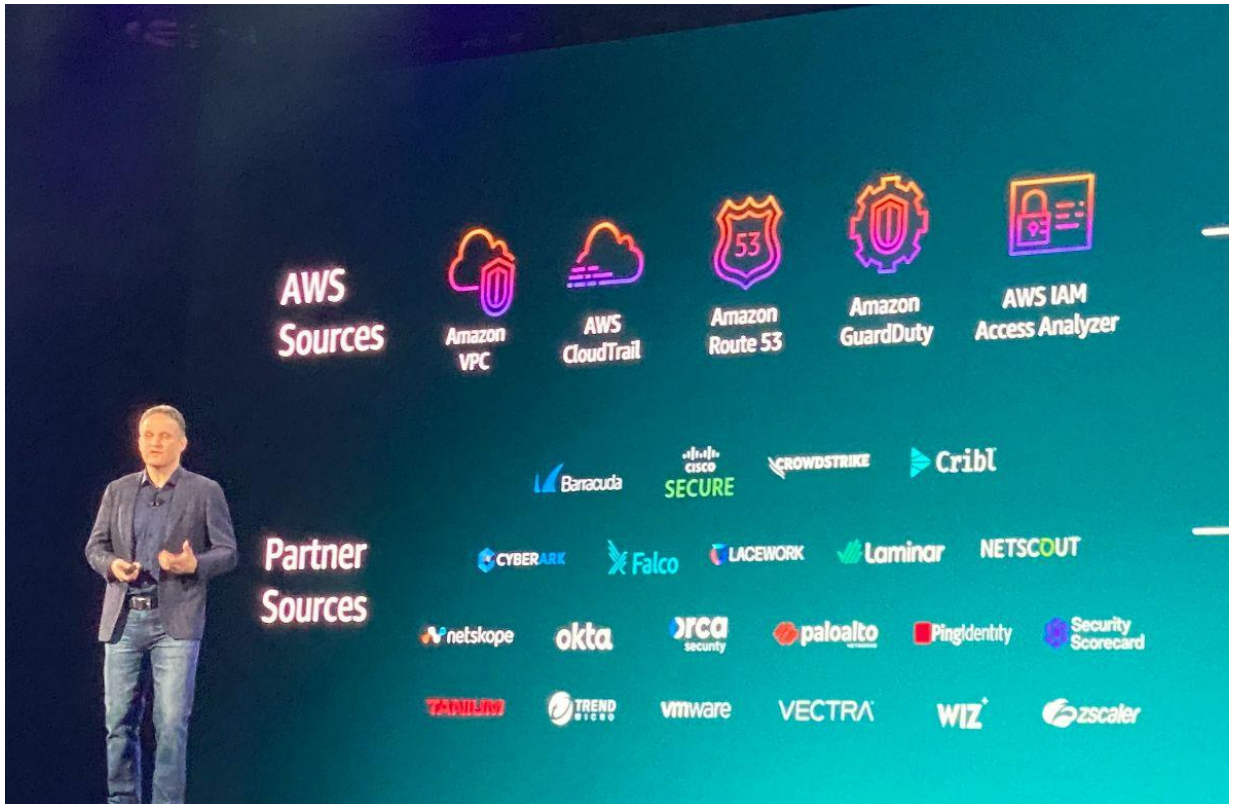
What is Amazon Security Lake?

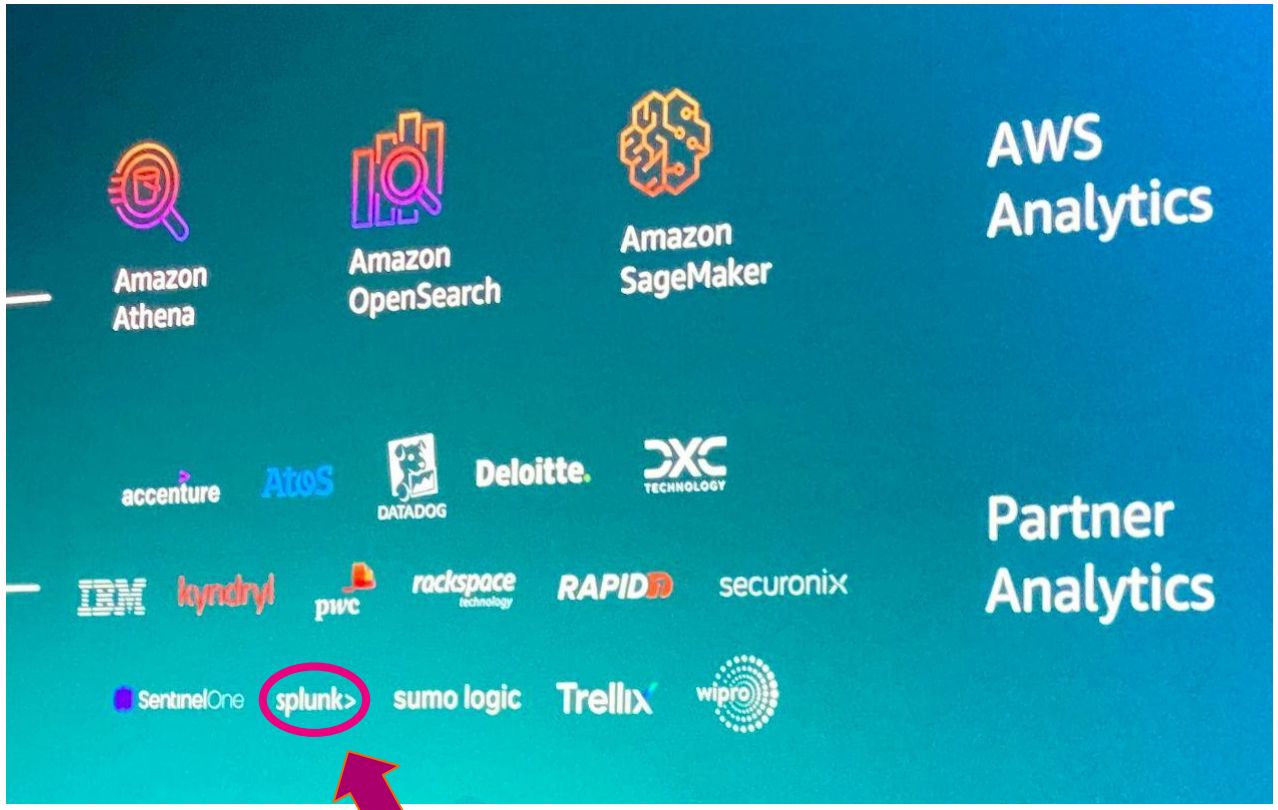
New AWS service that stores security data in OCSF format for storage and analysis



Amazon Security Lake is a service that centralizes data from AWS and 3rd party services in OCSF format and makes it available to native and 3rd party tools for search and analysis.

What is Amazon Security Lake?





Disclaimer: all logos displayed are the intellectual property of their respective owners

Why, exactly?

Why would Splunk partner?

- Reduce effort and spend on data storage and normalization of security data
- Today, Splunk is a leader in detection, response and investigation, but in the future, the focus should shift from storing and manipulating data to analysis and outcomes
- Bottom line, our customers don’t want to move data from public to private clouds and worry about ETL and parsing. They want to focus on gaining insights from their data to deliver outcomes.

Is Security Lake a Splunk replacement?

Image Credit: Pixabay

- Security Lake is a data storage layer service only
  - Requires Athena or OpenSearch to search
  - Requires other tools for analysis, dashboards
- For near-real time security monitoring and TDIR, you'll want your data searchable from Splunk
- Splunk allows AWS customers advanced query performance and capabilities, such as scheduling searches, running reports and creating dashboards that security teams rely on in the SOC today.

Security Lake Pricing

Pricing after a brief free trial period can be complicated - there is also an OCSF charge

US East (N. Virginia) | US East (Ohio) | US West (Oregon) | Europe (Ireland) | Europe (Frankfurt) | Asia Pacific (Tokyo) | Asia Pacific (Sydney)

Region			US East (N. Virginia)
CloudTrail logs			$0.75 per GB
Other logs			
First	10	TB	$0.25 per GB
Next	20	TB	$0.15 per GB
Next	20	TB	$0.075 per GB
Over	50	TB	$0.05 per GB
Normalization			$0.035 per GB

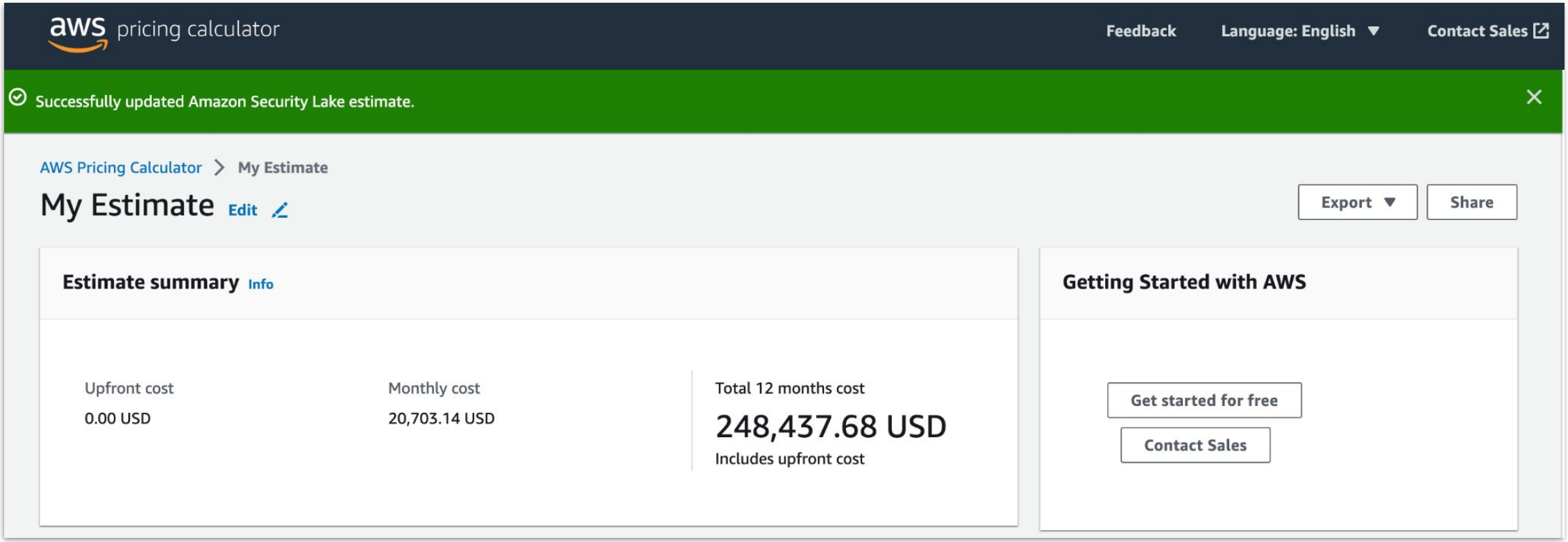
You are required to have a CloudTrail organization trail configured to collect CloudTrail management events into your security data lake.

Your data is stored in Amazon S3 and standard S3 charges apply.

Security Lake also orchestrates other AWS services on your behalf. You will incur separate charges for AWS services used and resources set up as part of your security data lake - separate pricing for AWS Glue, Amazon EventBridge, AWS Lambda, Amazon SQS, and Amazon SNS.

Pricing Example!

Your mileage, and budget, may vary….



For illustration only. Individual estimates and pricing will vary.

Pricing Example!

Your mileage, and budget, may vary….

Edit Amazon Security Lake Info

▼ Show calculations

6,200 GB of CloudTrail events x 0.75 USD per GB = 4,650.00 USD (cost for CloudTrail events ingested)
Total CloudTrail events cost: 4,650.00 USD
Tiered price for: 24800 GB of other AWS logs ingested
10240 GB of other AWS logs ingested x 0.2500000000 USD = 2560.00 USD
14560 GB of other AWS logs ingested x 0.1500000000 USD = 2184.00 USD
Total tier cost: 2560.00 USD + 2184.00 USD = 4744.0000 USD (cost of other AWS logs)
Total cloudtrail other events cost: 4,744 USD
6,200 GB of CloudTrail events + 24,800 GB of other AWS logs ingested = 31,000.00 GB (Total GB of data normalization)
31,000.00 GB x 0.035 USD = 1,085.00 USD (cost for data normalization cost)
Total data normalization cost: 1,085.00 USD
4,650.00 USD (CloudTrail events ingested) + 4,744 USD (other AWS logs) + 1,085.00 USD (Data normalization cost) = 10,479.00
Security Lake pricing (monthly): 10,479.00 USD

Edit Amazon Simple Storage Service (S3) Info

3100000

▼ Show calculations

Tiered price for: 31000 GB
31000 GB x 0.0230000000 USD = 713.00 USD
Total tier cost = 713.0000 USD (S3 Standard storage cost)
100,000,000 PUT requests for S3 Standard Storage x 0.000005 USD per request =
100,000,000 GET requests in a month x 0.0000004 USD per request = 40.00 USD (
31,000 GB x 0.0007 USD = 21.70 USD (S3 select returned cost)
3,100,000 GB x 0.002 USD = 6,200.00 USD (S3 select scanned cost)
713 USD + 40.00 USD + 500.00 USD + 21.70 USD + 6,200.00 USD = 7,474.70 USD
S3 Standard cost (monthly): 7,474.70 USD

S3 Standard cost (upfront): 0.00 USD

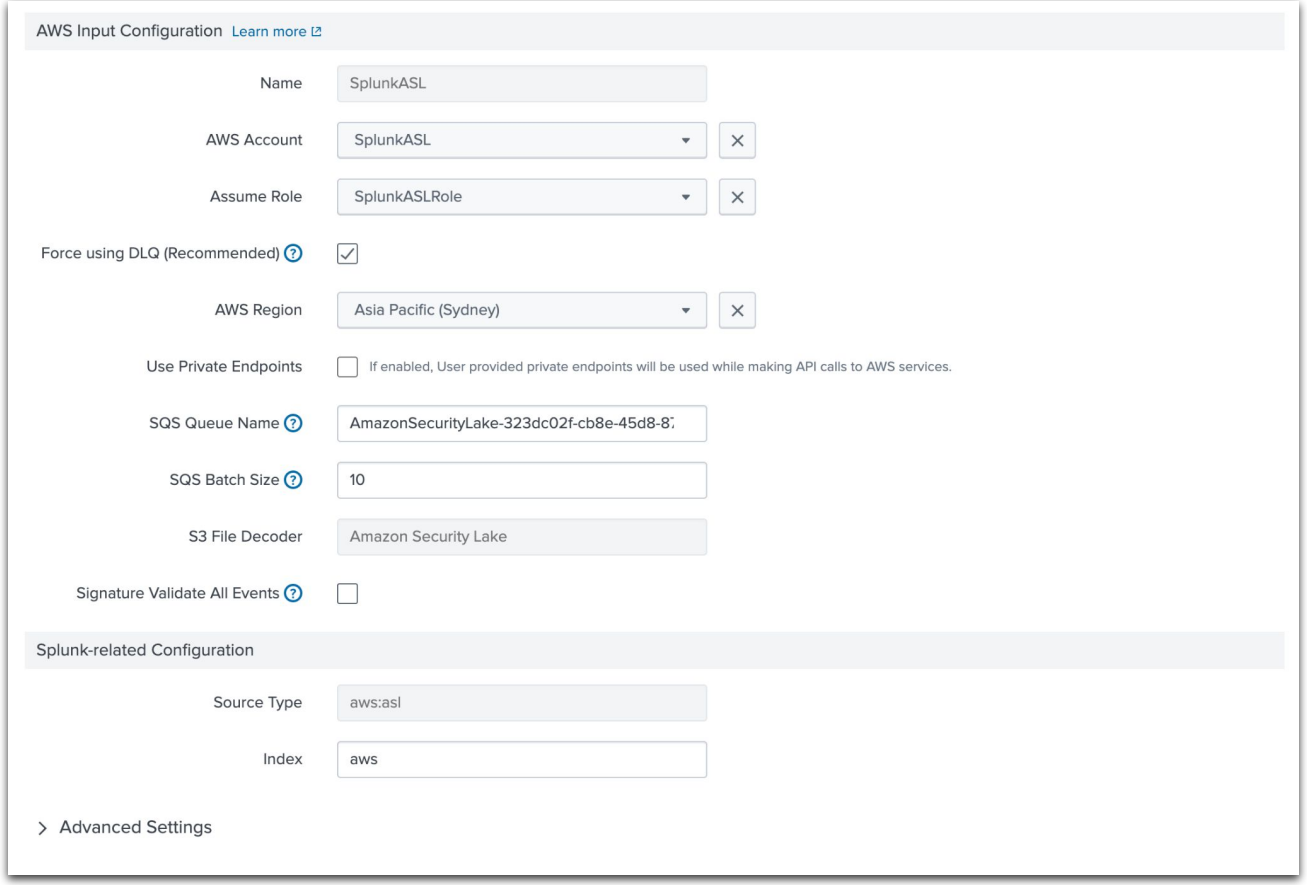
Total Upfront cost: 0.00 USD
Total Monthly cost: 10,224.14 USD

Show Details ▼

- 1TB daily (200GB Cloudtrail, 800GB other)
- Transfer outbound 10% of data in a month
- Scan 100% of the data 100 times

splunk> .conf23

Splunk® Add-On for AWS



- Bring Amazon Security Lake Data into Splunk so you can search it
- Security Lake data won't work with ES capabilities out of the box
- Product plans to improve ES to work with OCSF data in the future
- With the level of maturity currently, it doesn't quite make sense to ingest data into Security Lake before ingesting it into Splunk - except AWS data

Data Lakes

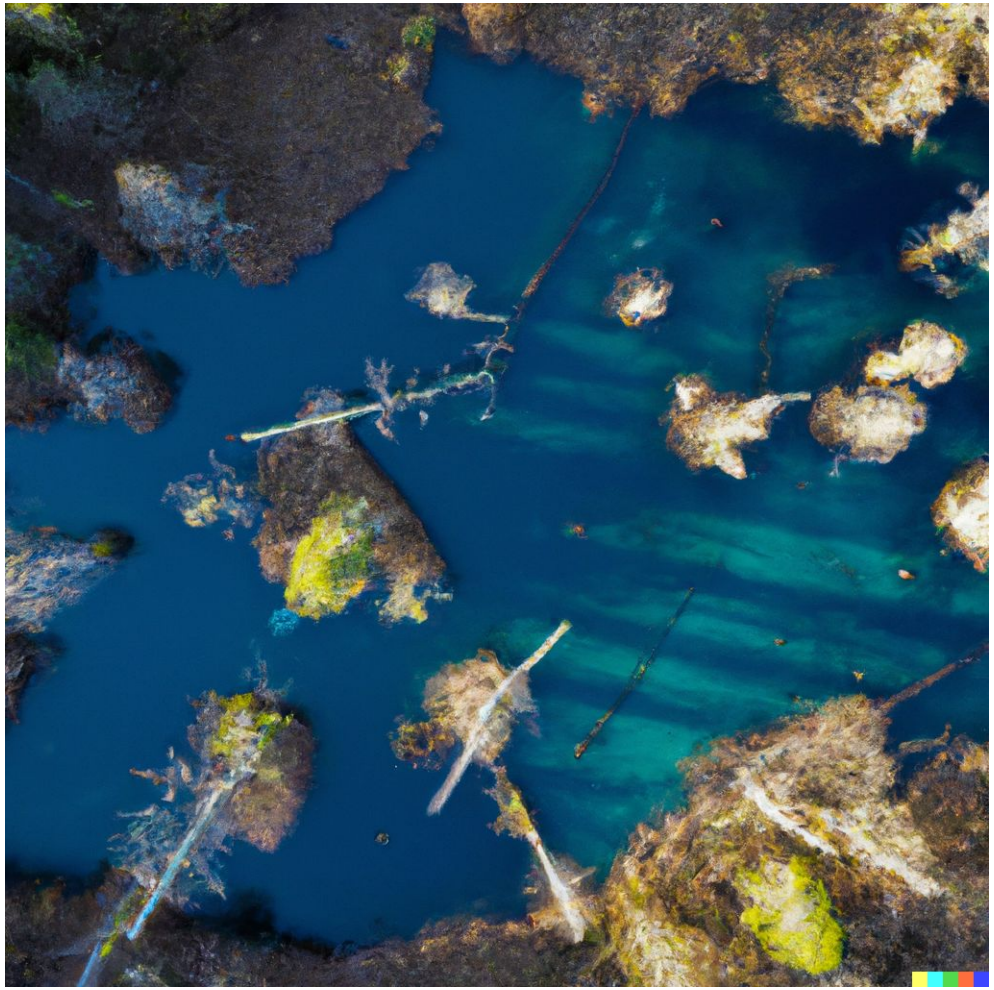
Data Lakes and Security are not new

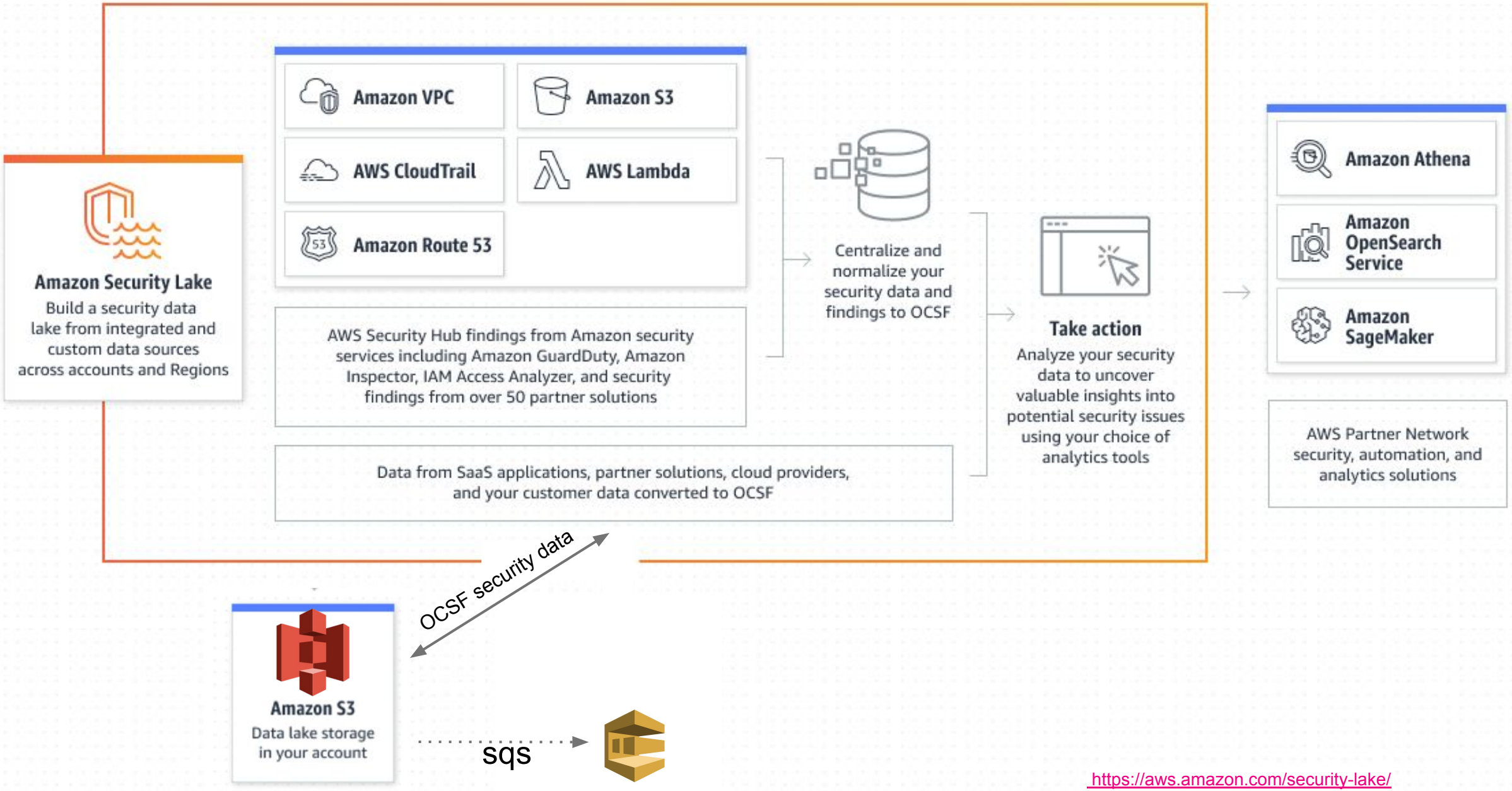
Search data where it resides instead of moving it around

Analytics are now more advanced than what we had with hadoop

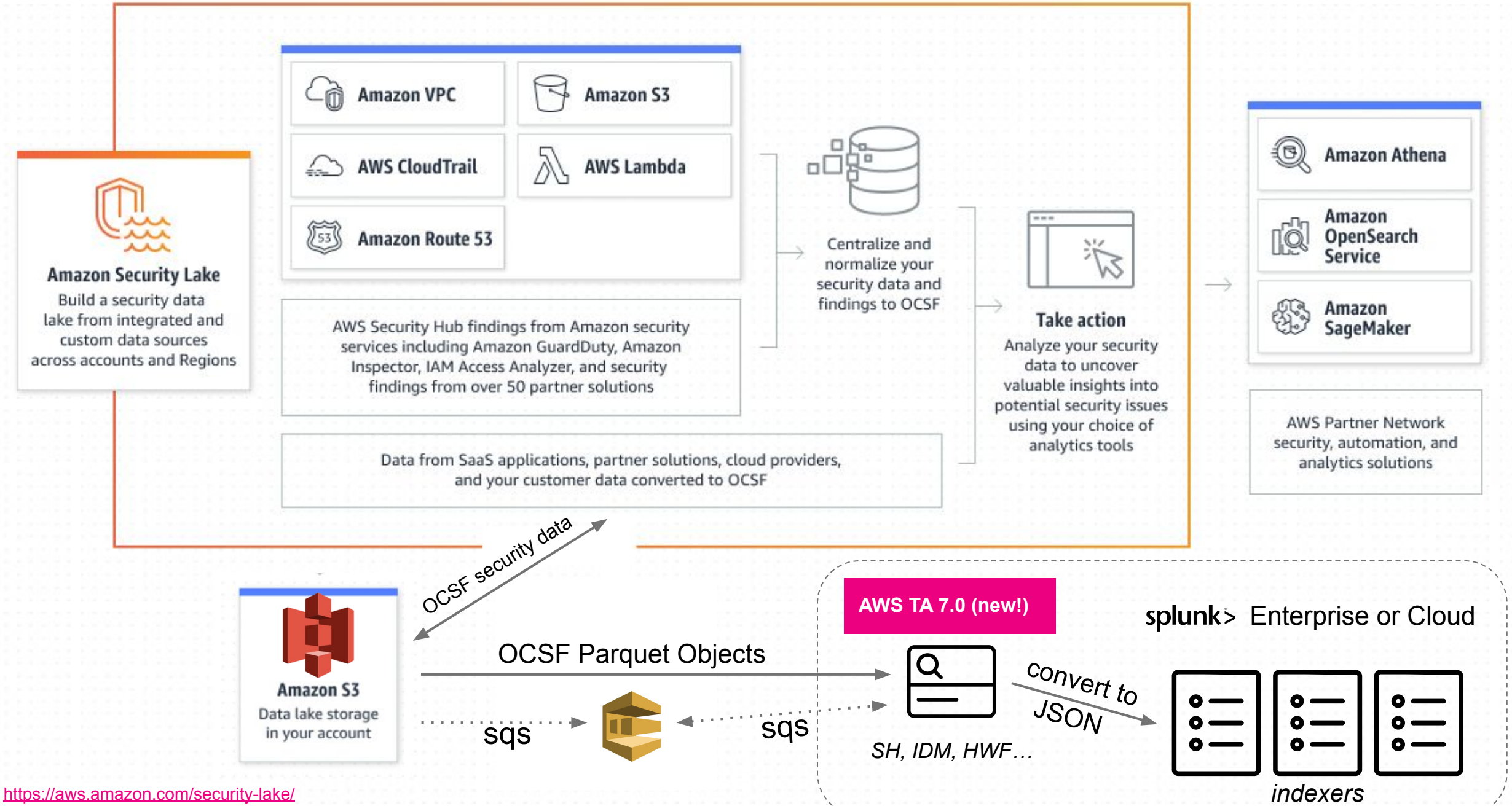
Open standards (OCSF) start to make Data Lakes make more sense

Search where it is vs. search in Splunk





https://aws.amazon.com/security-lake/



https://aws.amazon.com/security-lake/

```
index=*
```

✓ **46,204 events** (6/5/23 6:00:00.000 PM to 6/6/23 6:25:12.000 PM) No Event Sampling ▾ Job ▾

Events (46,204) Patterns Statistics Visualization

Format Timeline ▾ — Zoom Out + Zoom to Selection × Deselect

List ▾ Format 20 Per Page ▾ < Prev 1 2 3 4 5

< Hide Fields ☰ All Fields

SELECTED FIELDS
- a host 2
- a source 100+
- a sourcetype 2

INTERESTING FIELDS
- a action 1
- # activity_id 6
- a activity_name 7
- a actor.idp.name 1
- a actor.invoked_by 5
- a actor.session.created_time 100+
- a actor.session.issuer 4
- a actor.session.mfa 2
- # actor.user.account_uid 2
- a actor.user.credential_uid 100+
- a actor.user.name 2
- a actor.user.type 3
- a actor.user.uid 5
- a actor.user.uuid 5

sourcetype ×

2 Values, 100% of events Selected Yes No

Reports

Top values Top values by time Rare values

Events with this field

Values	Count	%
aws:asl	46,056	99.68%
stash	148	0.32%

```
cloud: { [+]
}
dst_endpoint: null
http_request: { [+]
}
metadata: { [+]
}
mfa: null
resources: [ [+]
```

splunk> .conf23

SELECTED FIELDS
- a host 2
- a source 100+
- a sourcetype 2

INTERESTING FIELDS
- a action 1
- # activity_id 6
- a activity_name 7
- a actor.idp.name 1
- a actor.invoked_by 5
- a actor.session.created_time 100+
- a actor.session.issuer 4
- a actor.session.mfa 2
- # actor.user.account_uid 2
- a actor.user.credential_uid 100+
- a actor.user.name 2
- a actor.user.type 3
- a actor.user.uid 5
- a actor.user.uuid 5
- a api.operation 6
- a api.request.uid 100+
- a api.response.error 1
- a api.response.message 1
- a api.service.name 2
- a api.version 1
- a app 4
- a category_name 3
- # category_uid 3
- a class_name 4
- # class_uid 4
- # cloud.account_uid 1
- a cloud.provider 1
- a cloud.region 7
- a cloud.zone 1
- a connection_info.boundary 2

6/6/23
6:20:29.000 PM

```
{ [-]
  activity_id: 3
  activity_name: Update
  actor: { [+]
  }
  api: { [+]
  }
  category_name: Audit Activity
  category_uid: 3
  class_name: API Activity
  class_uid: 3005
  cloud: { [+]
  }
  dst_endpoint: null
  http_request: { [+]
  }
  metadata: { [+]
  }
```

app

4 Values, 99.68% of events

Selected: Yes | No

Reports

Top values | Top values by time | Rare values

Events with this field

Values	Count	%
Amazon VPC	25,152	54.612%
CloudTrail	20,817	45.199%
Security Hub	60	0.13%
Route 53	27	0.059%

host = $decideOnStartup | source = s3://aws-security-data-lake-ap-southeast-2-i9d9o9wfggi1xp94cf61rixgb/aws/S3_... | sourcetype =

Possibilities Virtually Unlimited

Read out of a Lake, build meta-index in Splunk, query that index instead

Federated Search into lake

Federated Search of S3

Searching/Querying other Data Lakes with dbconnect

- Snowflake and Databricks prove that this can be done and the apps are on splunkbase

Custom search commands in Splunk that do what you can and inject it into the search pipeline

- Sentinel One does this currently

An example combined use case

Assumes future Federated Search capabilities

- Store Cloudtrail data in Splunk and VPC flow logs in Amazon Security Lake
- ES runs regular detections against recent Cloudtrail data
  - Access and account activity type of detections.
- Enterprise Security detects an unusual login activity from a EC2 instance to a S3 bucket based on Cloudtrail data in Splunk
- The analyst wants to see all of the network traffic that came to and from the EC2 that connected to the S3 bucket instances
- Analyst uses federated search to query VPC flow log in Lake
  - Depending on how far back you are pulling VPC flow log data that is tied to the EC2 instance will depend on how much data you are searching.
- Federated Search (via Athena) will take multiple queries for the analyst to complete their investigation, or…
- Federated Search (Indexer) data is already indexed to query against so no need to launch multiple queries

Wrap Up!

Please don’t leave yet, you’ll give us a complex - but 5 stars!

- If you’ve turned on Amazon Security Lake, talk with your account team to get the new AWS technology addon up and running
  - If you’re also using Enterprise Security (or need CIM), check out the OCSF to CIM beta conversion tool (chat with your account team)

Watch this space for Federated Search

Sessions to check out:

PLA1422A - Federated Search across S3

THE2070 - AWS Security with Splunk

SEC1969A - Automation with Splunk to support SOC2 in AWS

Thank You

splunk> .conf23