AWS Native SIEM and OCSF

- by Ashish Ghodake

About me

- Cloud Security Enthusiast
- Opensource Lover
- Biker & Trekker

- arghodaketech@gmail.com
- https://www.linkedin.com/in/arghodaketech
- https://mrpool404.medium.com/
- https://www.instagram.com/mr.pool.404/

What’s in for today ?

- OCSF
- AWS Security Lake
- Building Complete SIEM on AWS

Open Cybersecurity Schema Framework (OCSF)

- open-source project, delivering an extensible framework for developing schemas, along with a vendor-agnostic core security schema
- developed by AWS along with its partners
- has set of data types, an attribute dictionary, and the taxonomy.
- agnostic to storage format, data collection and ETL processes

OCSF Taxonomy - Presonas

- basically users of the schema
- Types
  - *Author* : creates or extends schema
  - *Producer* : generates events natively into the schema
  - *Mapper* : translates events from different schema to OCSF
  - *Analyst* : end user who searches the data

OCSF Taxonomy - Data Types

- *Scalar data types* - like strings, integers, floating point, numbers and Boolean, IP Address, MAC address, Timestamp, etc.
- *Complex data types* - are termed as objects and each object itself can be a data type
- *Arrays* - support any data types
- *Attribute* - unique identifier for a specific Validatable data type
- *Event Class* – structured set of attributes that describe a particular type of event
- *Categories* – organize event classes that represent particular domain
- *Profiles* – Overlay additional attributes to event class for better co-relation
- *Extensions* – allow schema to be extended without manipulating the core schema

OCSF Event Class

- Event classes have the semantic for the events
- *Base Event* class has the required, recommended and optional attributes that apply to all the event classes
- Special base class attributes
  - *Unmapped* : all attributes in the original event that are not mapped
  - *Raw_data* : optional attribute that holds the source raw data as whole
  - *Type_uid* : combination of class_uid and activity_id which is unique across the schema
- Event class can have constraints which can be more versatile than simple required constraints
- Associations are relations between event classes

OCSF Event Categories

System Activity	Network Activity	IAM	Findings	Discovery	Application Activity
File System Activity	Network Activity	Account Change	Security Finding	Device Inventory Info	Web Resource Activity
Kernel Extension Activity	HTTP Activity	Authentication		Device Config State	Application Lifecycle
Kernel Activity	DNS Activity	Authorize Session			API Activity
Memory Activity	DHCP Activity	Entity Management			Web Resource Access Activity
Module Activity	RDP Activity	User Access management			
Scheduled Job Activity	SMB Activity	Group Management			
Process Activity	SSH Activity				
	Email Activity				
	Network File Activity				
	Email File Activity				
	Email URL Activity				

Pros

- Opensource !!!
- Unified standardized security data
- Avoid vendor lock in
- Improve query efficiency

Cons

- One more schema to learn (hopefully the last ??)
- Still in early phases of adoption

AWS Security Lake

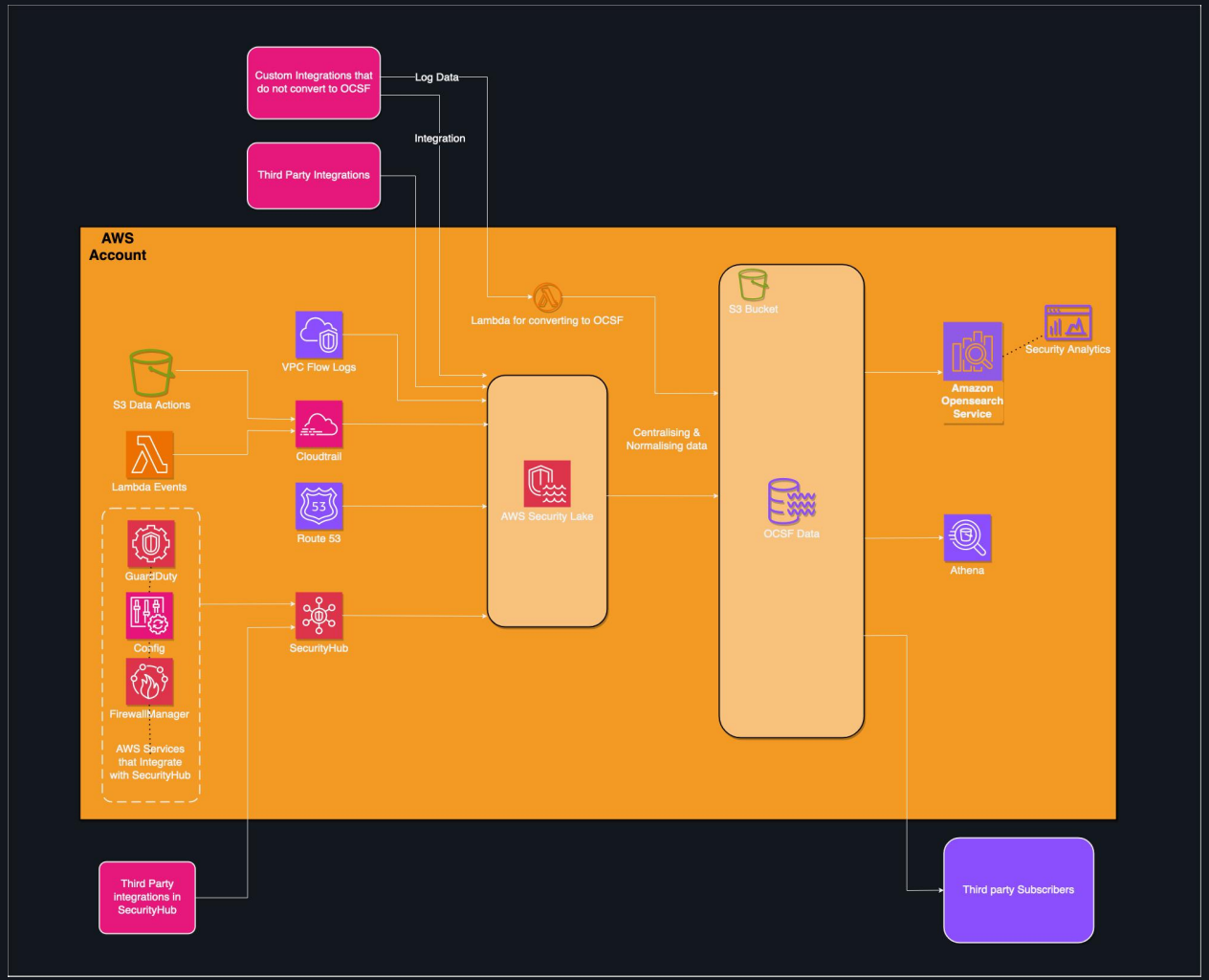
- Managed service by AWS to centralize security data from AWS as well as Non-AWS sources into purpose-built data lake
- It transforms the security data into OCSF format and stores it into the data lake built on S3
- It partitions data to optimize querying directly onto S3
- Optimizes lifecycle management

Sources

- Sources are the services or applications that generate security logs
- Can be AWS services like:
  - Cloudtrail
  - Route53 Resolvers
  - SecurityHub
  - VPC Flow
- Can be custom sources
- Custom Sources need to be transformed into OCSF and partitioned by the customer

Subscribers

- Subscribers are the services or applications that consume the logs to give out security insights
- Subscribers can be of two types:
- Subscriber Data access – subscribers that get notified of all the new objects created
- Subscriber Query access – subscribers that have access to querying Security Lake data
- AWS Athena can also directly query the security lake
- Third Party tools which can query S3 can also be used to query the security lake


Custom Integrations that do not convert to OCSF
Log Data
Integration
Third Party Integrations
AWS Account
Lambda for converting to OCSF
S3 Bucket
Security Analytics
VPC Flow Logs
S3 Data Actions
Amazon Opensearch Service
Cloudtrail
Lambda Events
Centralising & Normalising data
AWS Security Lake
Route 53
OCSF Data
Athena
GuardDuty
Config
SecurityHub
FirewallManager
AWS Services that Integrate with SecurityHub
Third Party integrations in SecurityHub
Third party Subscribers

Pros

- Fully Managed Services
- AWS Support available
- Opensource components
- Large number of integrations from third party available

Cons

- Cost
- Integrations may require custom code

References

- https://github.com/ocsf/
- https://github.com/ocsf/.github/blob/main/profile/Contributors.md
- https://schema.ocsf.io/
- https://github.com/ocsf/ocsf-docs/blob/main/Understanding%20OCSF.pdf
- https://aws.amazon.com/security-lake/
- https://docs.aws.amazon.com/security-lake/latest/userguide/what-is-security-lake.html