Laminar

Laminar DeepScan™ Technology for AWS

Laminar DeepScan™ Technology for AWS

Introduction

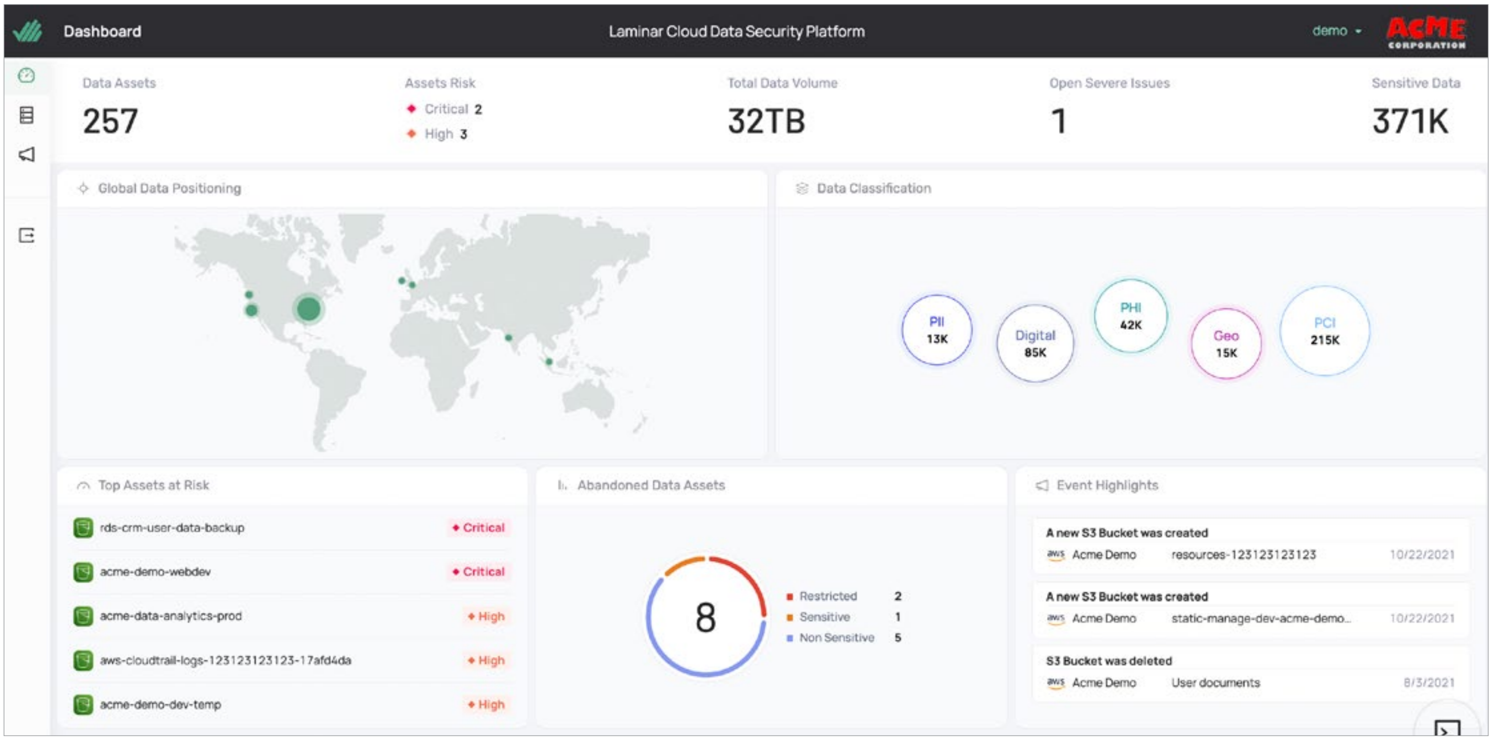
This whitepaper is designed for the technical audience to look a tad deeper into the **how** we do it. Laminar's Cloud Data Security Platform is based on our cloud-native DeepScan™ Technology that is both agentless, asynchronous, and with zero interruption to performance or data flow making it very simple and fast so customers can start to derive value from Laminar literally in a matter of minutes. This whitepaper will focus on the AWS environment and although DeepScan is very similar for Azure and GCP the exact cloud-native methods are native to each cloud service provider (CSP).

DEPLOYMENT

Installing Laminar is a simple install of a Cloud Formation stack. Laminar provides customers with a template and with just a few clicks in the AWS console is installed.

The Laminar deployment consists of three different parts:

1. **Cross-account Role** — allows Laminar service to query metadata about the client's environment, such as the names and sizes of S3 buckets. The Laminar service, running in Laminar's cloud account, assumes this role in order to understand what resources the client has and present them to the user. This role does not allow any access to the data itself, such that no data can ever leave the client's environment through the cross account role.
2. **Laminar Internal Role** — this is the role that Laminar's workload uses inside of the client environment to analyze and classify data. This role can only be assumed on compute nodes running inside of the client's environment — in particular lambda functions and EC2 instances. Laminar scans the data and sends only the metadata results of the classification (such as types and count of PII) to Laminar servers outside of the client network.
3. **Laminar Compute Environment** — Laminar creates a new private network inside of the client's cloud account. This network houses Laminar's compute resources and allows Laminar's service to access client data while keeping the data within the client's cloud environment.



Dashboard · Laminar Cloud Data Security Platform · demo

Data Assets	Assets Risk	Total Data Volume	Open Severe Issues	Sensitive Data
257	Critical 2 High 3	32TB	1	371K

Global Data Positioning

Data Classification: PII 13K · Digital 85K · PHI 42K · Geo 15K · PCI 215K

Top Assets at Risk

rds-crm-user-data-backup	Critical
acme-demo-webdev	Critical
acme-data-analytics-prod	High
aws-cloudtrail-logs-123123123123-17afd4da	High
acme-demo-dev-temp	High

Abandoned Data Assets: 8

Restricted	2
Sensitive	1
Non Sensitive	5

Event Highlights

A new S3 Bucket was created
aws Acme Demo resources-123123123123 10/22/2021

A new S3 Bucket was created
aws Acme Demo static-manage-dev-acme-demo... 10/22/2021

S3 Bucket was deleted
aws Acme Demo User documents 8/3/2021

The Laminar Dashboard

DISCOVERY

Laminar scans for and discovers all of the data assets in the client's cloud environment without the need for any user/owner credentials, just the cross-account role. Discovery is necessary because these data assets are unknown to data security teams as well as their corresponding access credentials.

Managed data assets are found by using AWS API calls through the cross-account role Laminar installs during deployment. Using the API, Laminar can list all of the S3 buckets, RDS databases, DynamoDB tables, and more.

Hosted data stores, such as EC2 instances running Postgres databases, can be more difficult to detect without Laminar. They can not be queried by using APIs and are often not documented at all.

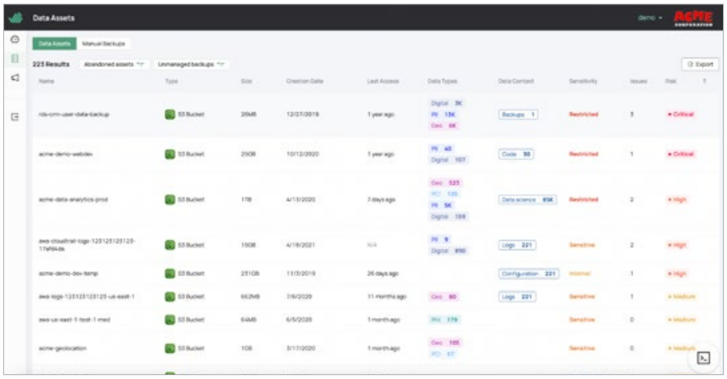
Using Laminar DeepScan™ Technology, Laminar scans the disk of all running EC2 instances in the client environment and discovers databases and other datastores. This operation happens on a clone of the disk and does not affect running systems at all. This takes place entirely within the client's cloud environment such that no potentially sensitive data escapes the client's cloud.

CLASSIFICATION

In order to read and classify client data, Laminar creates lambda functions that run within the client's cloud account. Laminar cannot access client's data from outside of this environment.

The way Laminar reads data is dependant on the type of data asset:

- **Managed Storage** — Services whose hosting and data access are fully controlled by a cloud provider, such as AWS S3. Laminar's internal role can simply use the AWS API to read this data
- **Managed DB** — Services such as RDS which are managed by AWS but not available through AWS API to read the data. Laminar creates an ephemeral clone of each RDS database. This clone is attached to the Laminar network inside of the client's cloud account so that it can be scanned. After classifying the data types, this clone is destroyed. Because Laminar only accesses clones of data, tasks performed by Laminar will never affect production systems.
- **Hosted** — Datastores running directly on virtual machines or other compute platforms. As described above, in order to discover these data stores, Laminar creates a clone of the disk and scans for datastores. Laminar uses these same clones to parse and classify the data from the datastore. Like with Managed data assets, this clone is attached to the Laminar network inside client environments and deleted after the data is classified.



Data Assets · demo

Data Assets · Manual Backups

223 Results · Abandoned assets · Unmanaged backups · Export

Name	Type	Size	Creation Date	Last Access	Data Types	Data Context	Sensitivity	Issues	Risk
rds-crm-user-data-backup	S3 Bucket	20MB	12/27/2019	1 year ago	Digital 3K PII 13K Geo 6K	Backups 1	Restricted	3	Critical
acme-demo-webdev	S3 Bucket	2GB	10/13/2020	1 year ago	PII 40 Digital 107	Code 30	Restricted	1	Critical
acme-data-analytics-prod	S3 Bucket	1TB	4/13/2020	7 days ago	Geo 325 PCI 120 PII 5K Digital 159	Data science 89K	Restricted	2	High
aws-cloudtrail-logs-123123123123-17afd4da	S3 Bucket	15GB	4/16/2021	N/A	PII 9 Digital 890	Logs 221	Sensitive	2	High
acme-demo-dev-temp	S3 Bucket	231GB	11/3/2019	26 days ago		Configuration 221	Internal	1	High
aws-logs-123123123123-us-east-1	S3 Bucket	662MB	7/6/2020	11 months ago	Geo 90	Logs 221	Sensitive	1	Medium
aws-us-east-1-test-1-med	S3 Bucket	64MB	6/5/2020	1 month ago	PHI 179		Sensitive	0	Medium
acme-geolocation	S3 Bucket	1GB	3/11/2020	1 month ago	Geo 105 PCI 67		Sensitive	0	Medium

Complete view of data assets and classified data in cloud account.

Once the data is read, Laminar classifies the data to search for sensitive data — PII, PCI, PHI, and more. Laminar also supports custom definitions of sensitive data to search for in cloud environments.

Laminar's classifier is based on three stages:

1. **Context Awareness** — Laminar uses contextual clues in order to ascertain the type of data. Most of the data found in cloud environments is structured data such as SQL databases and .CSV files or partially structured data such as .JSON or log files. Laminar uses the field names found in a data store, information about nearby fields, and a scan of nearby data above and below the current data to create contextual clues that help inform the classification of the data. Based on the contextual clues, Laminar creates a listing of possible data types and probabilities that are fed to the next stage
2. **Data Match** — Given the possible data types, Laminar determines if the data matches the expected format. This is usually done using regular expressions to match the data. Because of the contextual clues, Laminar is better to obtain a much lower error rate of both false positives and negatives than is usually possible with regular expressions alone.
3. **Verifier Function** — For many data types it is possible to further verify if the found data is actually of the suspected type. In these cases Laminar runs a verifier function to again check the data and ensure that no false positives are delivered. For example, the Luhn algorithm of credit card numbers is checked to confirm that the numbers found are actually credit card numbers.

Using these three stages, Laminar is able to correctly classify at a very high rate over many datastores. Laminar's algorithm is continuously improving, such that over time the types of data that Laminar recognizes and the accuracy with which it does so continues to increase.

SECURITY POSTURE

Laminar scans the access controls and security measures of each data asset. Together with the Discovery and Classification techniques discussed above, Laminar will generate a score for sensitivity and risk of each data asset that can give an understanding of the security posture of assets based on their contents as well as their permissions and security.

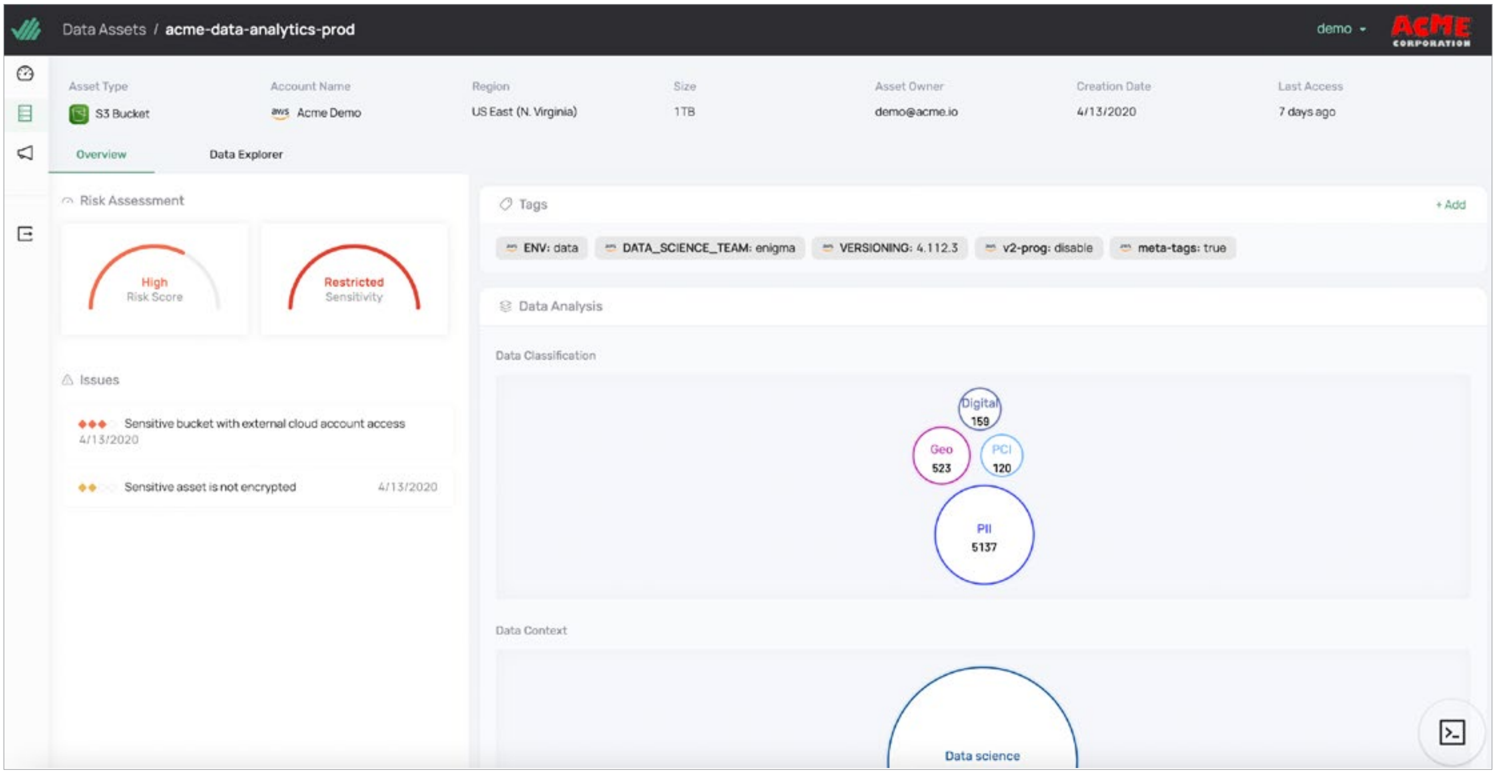
Laminar will generate issues based on either predefined or custom security rules. Predefined rules are continually updated by Laminar security researchers and represent issues such as:

- Database backup found in public S3 bucket
- PII not encrypted at rest
- Logging not enabled for database with PCI
- Abandoned data asset — data asset with sensitive data was not accessed in the past 90 days.

Custom rules can be defined by clients, such as:

- All PII should be in US
- Sensitive data should be tagged with "Owner"

Issues are identified and alerted based on prioritized risk exposure due to severity and/or data sensitivity.



Data Assets / acme-data-analytics-prod

Asset Type	Account Name	Region	Size	Asset Owner	Creation Date	Last Access
S3 Bucket	aws Acme Demo	US East (N. Virginia)	1TB	demo@acme.io	4/13/2020	7 days ago

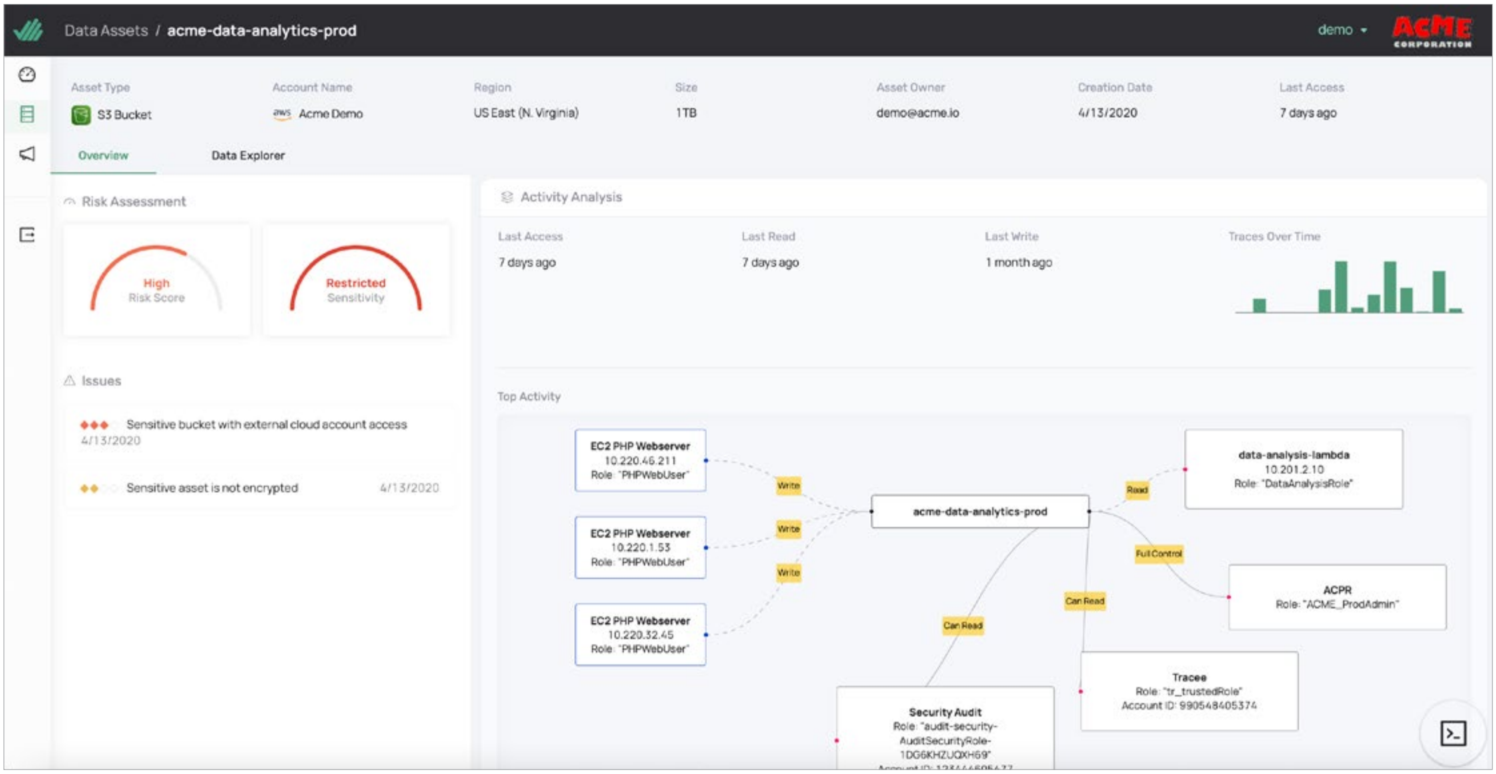
Tags: ENV: data · DATA_SCIENCE_TEAM: enigma · VERSIONING: 4.112.3 · v2-prog: disable · meta-tags: true

Issues

◆◆◆ Sensitive bucket with external cloud account access 4/13/2020

◆◆ Sensitive asset is not encrypted 4/13/2020

Immediatly see data sensitivity and data asset security posture.



Data Assets / acme-data-analytics-prod

Asset Type	Account Name	Region	Size	Asset Owner	Creation Date	Last Access
S3 Bucket	aws Acme Demo	US East (N. Virginia)	1TB	demo@acme.io	4/13/2020	7 days ago

Risk Assessment

High Risk Score

Restricted Sensitivity

Issues

◆◆◆ Sensitive bucket with external cloud account access 4/13/2020

◆◆ Sensitive asset is not encrypted 4/13/2020

Activity Analysis

Last Access	Last Read	Last Write
7 days ago	7 days ago	1 month ago

demo
ACME CORPORATION
Overview
Data Explorer
Traces Over Time
Top Activity
EC2 PHP Webserver 10.220.46.211 Role: "PHPWebUser"
EC2 PHP Webserver 10.220.1.53 Role: "PHPWebUser"
EC2 PHP Webserver 10.220.32.45 Role: "PHPWebUser"
Write
Write
Write
acme-data-analytics-prod
Read
data-analysis-lambda 10.201.2.10 Role: "DataAnalysisRole"
Full Control
ACPR Role: "ACME_ProdAdmin"
Can Read
Can Read
Tracee Role: "tr_trustedRole" Account ID: 990548405374
Security Audit Role: "audit-security-AuditSecurityRole-1DG6KHZUQXH69"

Continuous leak protection monitoring reads and writes.

LEAK PROTECTION

Access patterns of data assets can contain a great deal of information. Laminar analyzes the access logs of data assets to establish a baseline of normal accesses to the data. By detecting anomalies Laminar raises alerts when data is improperly accessed for an organization.

Oftentimes it is impractical to enable data access logging on all assets in an organization due to cloud hosting costs. However, everyone agrees that logging is a critical security measure for assets with your most sensitive data, but which ones are those? Laminar recognizes the client's most critical and sensitive assets and can make sure that access logging is enabled on those specific assets. This ensures that the most critical assets will be monitored.

ABOUT LAMINAR

Laminar delivers data leakage protection for everything you build and run in the cloud. Laminar's Cloud Data Security Platform is the first solution on the market delivering data-centric cloud security that allows you to Discover, Prioritize and Protect your sensitive cloud data. Continuously Discover and classify new datastores for complete visibility, Prioritize actions based on data sensitivity and risk posture, Protect data by improving security posture and actively Monitor for egress and access anomalies. Complete data observability for everything running in your public cloud accounts. Both agentless and asynchronous monitoring of datastores, compute as well as data egress channels allows sanctioned data movements and alerts when something's wrong. Data protection teams can reduce the attack surface, detect real-time data leaks and get back in control of their data. Founded in 2020 by a brilliant team of award winning Israeli red team experts, Laminar is proudly backed by Insight Partners, TLV Partners, and SentinelOne. To find out more please visit laminarsecurity.com.

Make your data flows Laminar!

Visit us: laminarsecurity.com