TRACEABLE

SOLUTIONS BRIEF

API Data Lake

The Basis for Context-Aware Security, Intelligence and Scalability

API Data Lake - Full Context for Complete API Security

Data lakes can provide significant value to security practitioners in various ways. By centralizing and storing massive amounts of raw, structured, and unstructured data from diverse sources, data lakes enable security professionals to gain deeper insights, improve threat detection and forensic analysis, and streamline security operations.

Data Lakes and APIs

API security requires full context. This means learning behavioral patterns – both normal and abnormal – researching reconnaissance tactics, and identifying any unknown attack in the environment that did not have known signatures to protect APIs.

The entire point of a data lake is to power the full context of APIs, and support the use cases of the "unknown unknowns". Since unknown threats, including zero-day attacks, make up the vast majority of attacks against APIs, a data lake is essential to constantly look for anomalous or malicious behavior.

Once the detections have been identified with that context, API protection becomes the next logical step. This can be done with signature or machine learning algorithms to protect APIs from specific business logic attacks or generic exploits in language or frameworks.

Without an API data lake, it's impossible to address unknown attacks and identify anomalies. You simply must have context-aware security through a data lake for complete API security.

SOLUTION HIGHLIGHTS

- **Forensic Analysis:** understand root cause of the incident, identify contributing factors, and develop recommendations for future threat prevention.
- **Threat Modeling**: provides the necessary data to build confident models that finds more true positives, not false positives.
- **Confirm known findings** concerning an intrusion and discover new information about it.
- **Analyze API traffic** related to malicious activities
- **Analyze logs and API access information** from request/responses to identify the perpetrator(s) of a network breach

Why an API Data Lake is Essential for Complete API Security

An API data lake provides the basis for robust security, intelligence and scalability in API security. It provides the full context necessary to practice proactive security.

Without it, you become a canned checkpoint to satisfy a checkbox. And unfortunately, that's how many think about security – compliance checkboxes.

There are multiple ways that an API data lake provides value to security practitioners.

Forensic Analysis: An API Data Lake powers in-depth post-mortem analysis. This is also known as a post-incident analysis or after-action review, and is a process conducted by security practitioners to investigate and analyze a security incident after it has occurred.

Forensic analysis, powered by an API Data Lake, allows you to understand the root cause of the incident, identify the factors that contributed to its occurrence, and develop recommendations for preventing similar incidents in the future.

This capability is essential for organizations to continuously improve their security posture, assess their attack surface risk, and learn from past incidents.

Threat Modeling: Machine Learning is only as powerful as the data it has to analyze, and the security of APIs comes down to the quality of the data being analyzed. If you do not have a data lake, or if a solution is only looking at edge traffic, then what data are you using to power your algorithms?

Without a large set of data to work with, you can't build confident models that will provide more true positives than false positives. False positives waste everybody's time. The data lake provides you a large enough data set to build confident models that will be finding more true positives than false positives.

Low and Slow Attacks: an API Data Lake provides the wide view of data necessary to identify attacks that happen over a long period of time. These are becoming so common that most attackers are no longer siphoning through millions of records in one shot – they know that there are solutions in place that monitor those activities. Instead, they target just a few records, then a few more several hours later, setting it up for a "low drip" of data taken per day.

Zero-Day Attacks: if a zero-day comes out tomorrow, can you be sure that your system was able to identify that zero day yesterday? How would you know where the apps and APIs, which are impacted by them, reside? What critical information could be compromised if those APIs were attacked using that vulnerability?

Without that knowledge, no platform could know that you were compromised. With a data lake - you can analyze the data traces to see if you are vulnerable, or if it has already happened in your environment.

Scalability: Scalability in data lakes for security professionals is essential, as organizations need to handle growing volumes of security data, such as logs, network traffic, and security events, to protect their IT infrastructure and maintain a strong security posture.

The ability to scale up as data volumes increase helps security professionals maintain performance, conduct advanced analytics, and respond to threats effectively.

The Bottom Line:

The API Data Lake provides security context -- a function that no other solution can claim, without a data lake as a vital component of their solution architecture.

Traceable:
The Industry's Only API Data Lake for Complete API Security

The Traceable API data Lake captures all of your API calls (not limited to just malicious requests) and information about actors accessing the APIs, infrastructure where the APIs are hosted, and dependencies on other APIs including 3rd party APIs.

With the API data lake, Traceable provides the most complete security context of your APIs by building API data flows, API usage patterns (normal and abnormal), and user activity sessions covering an extended period of time.

Traceable provides the option to vary the retention of the data lake depending on the customer's use case.

Longer retention in the API Data Lake enables:

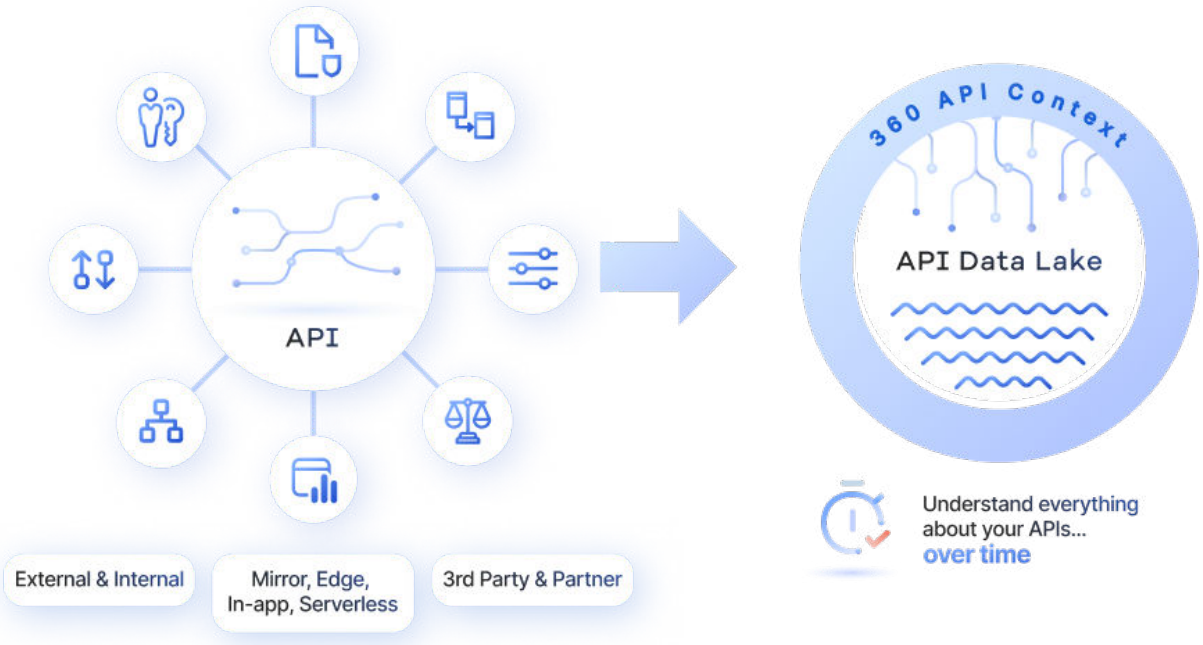
Security Posture Management and Discovery via API Catalog: provides audit teams with complete and verifiable details of all API activity (malicious and non-malicious) by leveraging the API data lake. This is especially useful where logging is found to be insufficient.

Compliance and Security Testing - For APIs that are rarely used, we are able to look at historical data and replay that for vulnerability scanning. Without the historical replay capability, only those APIs that are frequently exercised may be scanned.

Threat Detection - In addition to all the legacy attacks, Traceable is able to detect long-lived complex attacks starting from the recon stage. We are also able to detect and protect against slow and latent attacks whose impact culminates over a period of time by leveraging the data lake.

API DLP - Traceable detects and tracks all activity around sensitive data and data sets, including which users and APIs have accessed it, when, from what service, and if it has been exfiltrated. All this is possible to see over long periods of time for even slow leaks using the data captured in the Traceable Data Lake.

Threat Management - Forensics such as root cause analysis of attacks and other ad-hoc security analytics is made possible by the use of the data lake which provides advanced query capabilities over historical data. This also includes the ability to detect advanced API fraud and API abuse use cases.



TRACEABLE

Traceable is the industry's leading API Security company that helps organizations achieve API protection in a cloud-first, API-driven world. With an API Data Lake at the core of the platform, Traceable is the only intelligent and context-aware solution that powers complete API security – security posture management, threat protection and threat management across the entire Software Development Lifecycle – enabling organizations to minimize risk and maximize the value that APIs bring to their customers. To learn more about how API security can help your business, book a demo with a security expert.

www.traceable.ai/request-a-demo