Symantec by Broadcom

WHITE PAPER

Beyond the API

Security Events

WHITE PAPER

Beyond the API

Security Events

Symantec by Broadcom

TABLE OF CONTENTS

Security Events

Organizations implement security controls such as endpoint security, data loss prevention (DLP), and network security to protect themselves from threats such as advanced persistent threats, ransomware, and data breaches. Most of these controls serve dual purposes of both a source of protection (detect and prevent threats) and a point of visibility for triage, retrospective threat hunting, and forensic investigations. For both use cases, the event output of these technologies are important to any organization.

In organizations with dedicated security teams, these events are frequently consolidated with other event feeds such as authentication (Active Directory or identity provider), OS events (Windows Management Instrumentation [WMI] and other similar events), platform as a service (PaaS) and software as a service (SaaS) logs, firewall, DNS, and many others. For many organizations the destination of the events is a security information and event management system (SIEM). Increasingly, we see organizations with multiple tools such as user and entity behavior analytics and threat intelligence platforms, data lakes, and other tools that also consume this data. Many organizations do this work in the context of a security operation center (SOC). A SOC might utilize these data sets for a variety of different tasks, including the following processes:

- Threat detection
- Digital forensics and incident response
- Remediation activities
- Regulatory compliance

Security events are a critical component of these tasks, and ensuring that the events are of high value and are easily available is a critical task for any security vendor.

IN THE ON-PREMISES WORLD THERE IS A LACK OF WELL-ADOPTED STANDARDS.

Feeding the SOC

On-premises Events

Classically most security products have utilized a variety of different mechanisms for event export. Syslog, often utilizing the Common Event Format (CEF) is extremely common and supported by many products. However, both Syslog and CEF have numerous limitations around scale, network architecture, and data complexity. To overcome these limitations products have utilized other mechanisms including dedicated collectors (WMI and others), log push through protocols such as SCP/FTP, streaming mechanisms using custom or standard protocols, and so on.

In large systems, there are often issues with many devices sending data to a single location or single devices generating many events. This large amount of data can be problematic, because a single point of ingestion can cause issues with scaling and high availability. To solve this problem many organizations have created complex systems involving load balancers, multiple forwarders, and other custom workarounds just to deal with the scale of events.

In the on-premises world there is a lack of well-adopted standards. This deficit has led to many connectors or network configurations to ingest log and file data from a specific on-premises security product into a specific SIEM. Most SIEM vendors come with hundreds of different connectors. These vendors typically offer the ability to create custom ingesters that are both log and file based, or they offer network listeners (Syslog, HTTP, and so on) to deal with this problem. To avoid managing so many connectors, listeners, and other integrations between security products, customers are looking to consolidate the number of security tools they use. Consolidation is a trend that is driven by a need to reduce cost. The overhead of multiple ways to acquire data into the SOC, the SIEM product, or sometimes several SIEM products adds to the cost.

For organizations that are attempting to implement a data pipeline with multiple downstream tools, this problem only gets worse. Large custom solutions cause vendor lock in on both the producing and the ingesting side, making it difficult to migrate to a new solution. This environment even complicates trying new control points, analytics, and forensic tools. Even a proof of concept project often requires a solution for data exchange challenges.

Cloud Events

For cloud based controls, event data is typically provided through APIs. Event export typically comes in one of the following formats:

- File based downloading, often through a RESTful API
- Variants on RESTful APIs such as Microsoft Graph and so on
- Streaming APIs that are essentially a custom stateful API that streams data out as soon as it is created or the client is ready

Once again there are no standards, and every API implementation is unique. Interacting with the APIs requires a custom script or purpose built connector that not only has to interact with the unique API, but also ingest and map the data provided using the vendor's schema.

THE API MODEL HAS FUNDAMENTAL LIMITATIONS AROUND SCALE AND DATA PIPELINES.

APIs as an event export mechanism creates the following challenges:

- Since all APIs are different, they always require custom scripting or a vendor (either the producer or consumer) produced collector.
- For file and RESTful based APIs, they can potentially produce extremely large files that are difficult and failure prone to ingest.
- Streaming APIs almost by definition can only support a single client due to their stateful nature.
- Neither mechanism is conductive to load balancing or high availability, and at scale both configurations are challenging to support and maintain.
- For cloud to cloud, a pull mechanism is fundamentally inefficient because it requires a permanently running client that is either polling or constantly connecting.
- Given the lack of load balancing, APIs are essentially a single monolithic feed. When that feed is large enough, it can be difficult to ingest and failure recovery becomes difficult. Compared to the on-premises world where feeds tend to be broken out, moving to one gigantic feed can be problematic without load balancing and high availability.

Essentially APIs as an event export and ingestion mechanism are often inappropriate. APIs are offered by vendors as much through inertia as any other reason. The API model in general has fundamental limitations around scale and data pipelines. There is a reason why internally most organizations and infrastructure as a service (IaaS) providers do not use APIs for event streaming. Typically, they use protocols such as Pub/Sub, Kinesis, MQTT, Kafka, and other purpose built protocols for event streaming and data pipelines.

Fundamentally, protocols built for event streaming perform better, scale better, and offer features that are difficult to implement on both the API and client side of an API. APIs still have many use cases, but for event streaming they are the wrong tool for the job.

Data Standards

Historically, there has been no security event standard that has had broad adoption. Every vendor and technology has created data formats based on their needs, and the down stream tooling has had to adjust. Events might come across as delimited fields, possibly with a header or hardcoded field list, key pairs in CEF or JSON, and so on. This environment is challenging when attempting to correlate data between disparate data sets and is a key function of just about any tool that uses data from more than one source. A search based on a user could possibly fail if one source uses *user.name* and another source only uses *user*. The SIEM vendors in particular are aware of this problem and attempt to map many fields to their own data model on ingestion. This problem creates more upfront work for the downstream tool, and it also does not solve the problem for the data pipeline as a whole.

"SYMANTEC [A DIVISION OF BROADCOM] IS PROUD TO HAVE CONTRIBUTED OUR ICD SCHEMA AS THE FOUNDATION FOR THE OCSF PROJECT. THIS IS ANOTHER PROOF-POINT OF HOW WE SUPPORT OPEN STANDARDS ACROSS THE SECURITY INDUSTRY."

– ROB GREER, GENERAL MANAGER, SYMANTEC ENTERPRISE DIVISION AT BROADCOM

In the last few years there has been an initiative to develop global standards for vendors. The Symantec® product portfolio is in the vanguard of this initiative with the Symantec Integrated Cyber Defense (ICD) schema. Recently some new open source standards such as Open Cybersecurity Schema Framework (OCSF), which is descended from ICD, and Elastic Common Schema have been released. Vendor support for OCSF is rapidly expanding. This evolution is why Symantec product development is going to evaluate the support of multiple output schemas by premapping our outputs to an administrator selected mapping. The goal is to output different formats and different types of events to ensure the flexibility and easy adoption of the Symantec product portfolio, no matter what tools an organization adopts.

Modernizing Event Export in Cloud

Event export from SaaS applications at scale requires better solutions than the current APIs. Organizations need a solution that meets the following needs:

- Scales
- Utilizes standard mechanisms that are well supported across a variety of tooling
- Provides standardized data
- Is data pipeline friendly
- Supports both cloud to on-premises and cloud to cloud equally well

Symantec, a division of Broadcom, is offering two new unified options for event retrieval across all Symantec SaaS products.

As part of this unification, the Symantec product portfolio is standardizing the output schema of our products. All products will continue to generate their current output or Symantec ICD schema if desired. The software might also have new standardized output data types, including open source alternatives. Over time, Symantec product development will evaluate the support of additional outputs choices to support different use cases and tooling.

Customer Owned Bucket Push

The customer owned bucket push is a push mechanism where the Symantec product pushes files from the Symantec Enterprise Cloud platform into a customer owned bucket. This mechanism requires a public bucket that has *only write access* and supports Amazon S3, Google Cloud Storage, and Microsoft Azure Storage.

Benefits:

- Cloud buckets scale to any workload
- Native support by many tools as either a log location or a data store
- Cloud tooling available for data lifecycle and data pipelines
- Option for pure cloud to cloud and completely serverless
- Data assurance can be built into the system

The bucket push mechanism is a clean and fast way to receive security events in a purely cloud to cloud manner with no requirement for an on-premises component.

SUPPORT FOR BUCKET PUSH AND KAFKA STREAMING GIVE CUSTOMERS THE FLEXIBILITY THEY NEED FOR THE WIDE VARIETY OF ANALYTIC AND SIEM TOOLS THEY USE.

Customers frequently use analytics tools that support bucket push. In these instances, they can perform the following steps:

1. Configure the push from the Symantec product.
2. Configure pull from a tool.
3. Define the lifecycle (archive, delete, and so on) directly on the cloud provider and never touch the system again.

This solution is suitable for a wide range of organizations as long as they have the ability to utilize public buckets (write only) as a log intermediary.

Custom Kafka Topics

Kafka is an event streaming standard developed by the Apache foundation. Unlike an API, Kafka is a well known standard. It is supported by nearly every SIEM tool with libraries and tooling freely available in a variety of different form factors. Kafka supports Transport Layer Security encryption, and it can operate as a pull mechanism that is very similar to how the streaming APIs currently work. Essentially, clients would subscribe to a custom topic and retrieve events as they are produced.

Benefits:

- High throughput with built in load balancing and high availability
- Broad vendor support and technical experience with a mature ecosystem
- Extremely data pipeline friendly
- Lowest possible latency for receiving events

Kafka can support either cloud to on-premises or cloud to cloud easily, and it is supported and maintained by a large ecosystem of vendors and organizations. Kafka serves organizations or tools that do not support bucket based event retrieval. It is a pure pull mechanism that can work inside of firewalls either directly from the tools or as part of a data pipeline.

Client Support

Clients on the downstream side serve three primary purposes:

1. Provide the mechanism for either receiving or pulling the data. Examples include Syslog, HTTP/FTP/SCP, APIs, streaming protocols, and so on.
2. Parse the data and turn it into field names and values.
3. Map common fields to the internal schema to fulfill use cases such as correlation, threat hunting, and search.

In certain circumstances, a client might also be provided to create custom visualizations and dashboards.

ONE OF THE EXPLICIT GOALS OF THE NEW SYMANTEC LOG STREAMING MECHANISMS IS TO REMOVE THE REQUIREMENT FOR CUSTOM CLIENTS OR SCRIPTS ENTIRELY.

The following mechanisms are utilized by vendors (both originator and consumer) to ensure data exchanges:

- Output over standardized protocols and data formats. A common example of this output would be CEF over Syslog.
- A custom collector, either based on public documentation or through engineering collaboration is created by the downstream vendor. For specific vendors the generating vendor might create a technical add on or plugin to retrieve events.
- Documentation and public example scripts.

One of the explicit goals of the new Symantec log streaming mechanisms is to remove the requirement for custom clients or scripts entirely.

New Feeds

One other advantage for standardizing the Symantec product portfolio outputs is the ability to quickly publish new event feeds. If a product had to publish a new feed, it typically required a new API or a new update to an existing API. Either way, ingestion required either client side changes or a new custom script. This time and cost versus the benefit meant that new SaaS event feeds were rare and underutilized. Many products have additional features and feeds that can now quickly publish their events. In the past this change would have been a painful process. In the near future, Symantec product development will be evaluating the support of new event sources that are not currently available.

Conclusion

With these new event export mechanisms, the Symantec product portfolio is moving beyond the API. We are embracing better ways to ship security events from the Symantec Enterprise Cloud platform to where organizations need their data. At the same time, we are unifying mechanisms and data types across the entire portfolio to simplify management, removing the need for custom scripts and clients, and ultimately lowering the cost of ownership. By embracing standards and open schemas, Symantec Enterprise Cloud is easier to adopt and becomes a more effective part of the customer's security ecosystem. Together with our partners, we are eliminating the need for customized scripts, multiple connectors, and per-product data type mappings.

BROADCOM
connecting everything®

For more information, visit our website at: www.broadcom.com

BTA-SE-WP101 October 26, 2023