ThreatML: Machine Learning Done Right

Our detection-in-depth approach leverages predictions to surface real, actionable threats without tuning

If you Google “supervised learning,” you’ll learn that it’s a subcategory of machine learning (ML) and artificial intelligence (AI).

What powers the model is the use of labeled datasets to train algorithms in order to predict outcomes.

In the world of cloud security, we call this predictive threat modeling. High-level takeaways in this ebook include:

1. Why ThreatML™ + Rules is ML done right for cloud security
2. How the deep telemetry collection and rules of F5® Distributed Cloud App Infrastructure Protection (AIP), formerly known as Threat Stack, enabled us to develop supervised learning technology for high-efficacy threat detection
3. Why the future of cloud security is supervised learning

The Modernization of Cloud Infrastructure

Organizations are placing more emphasis than ever on the modernization of applications and architectures to reduce costs and promote increased productivity and flexibility. In fact, according to a recent survey by 451 Research, 87% of organizations operate both modern and traditional architectures, with modernization deemed necessary when legacy systems are too rigid to adapt to rapidly changing business conditions.[1] But as organizations work to keep pace with the rate of change and digitization of products and services, new challenges arise, especially when it comes to cybersecurity.

Most of the innovation taking place today is in the cloud. In a recent IDC Futurescape report, the research company estimated that over 90% of enterprises worldwide will rely on a mix of on-premises/dedicated private clouds, multiple public clouds, and legacy platforms to meet their infrastructure needs.[2] Therefore, security teams must efficiently identify and respond to security and compliance risk across their entire infrastructure and application stack—irrespective of where those workloads run.

This is easier said than done given the heterogeneous nature of cloud infrastructure, even for advanced cloud-native organizations.

Evaluating the Challenges in Securing the Entire Cloud Infrastructure

Most organizations care about running their security operations as best as they can, given their constraints. Constraints might mean being short-staffed, having a small budget, or doing other jobs in addition to security. Doing the best job means having the right balance of coverage, context, and cost. Security teams need to have visibility to all threats in their environment, the proper context for relevance into the more significant needs of the organization, and most importantly, to reduce key security metrics like mean-time-to-know (MTTK) and mean-time-to-respond (MTTR).

But we all know that too many alerts can be overwhelming, and can obscure attacker behavior. Teams can become ineffective if they aren’t focused on the highest priority threats. Other vendors attempt to solve this with just anomaly detection—but simply finding anomalies that depart from the baseline with little context is not the answer. Furthermore, reducing alert fatigue through silencing alerts creates a false sense of security and is not a strategy that reduces or manages risk. Thus, maintaining complete visibility over the environment with highly prioritized alerts in context becomes imperative.

"Time and time again, security ops teams are intrigued by platforms claiming to drastically reduce noisy alert environments. This is equivalent to going to your car mechanic and having them turn off the check engine light appearing on your dashboard so that you can simply pass an inspection. There's a way to have the best of both worlds, in other words, more context into what the most pressing problems are with your car. This same logic can be applied to cloud security."

Chris Ford, RVP of Product and Engineering, F5

Overcome Known and Unknown Threats with Full Stack Observability and Machine Learning Done Right

In today's digital era, security organizations need to help drive business transformations. However, striking the right balance between technology enablement and keeping up with the growing number of security and compliance requirements can feel like an uphill battle.

ThreatML considers the growing complexity of today's IT environment and incorporates the value of our rich security telemetry and alerting rules to help address this challenge.

Through our rules + ML approach, security teams can optimize their investigation and remediation efforts by improving MTTK and MTTR. As a result, security teams can empower their organizations to embrace strategic and transformational technologies, such as containers and Kubernetes, while also ensuring appropriate security and compliance measures are in place to protect their critical systems and sensitive information.

The goal is for customers to take advantage of predictive machine learning via ThreatML that deliver high-efficacy alerts that represent real, actionable threats to the environment.

Distributed Cloud AIP's Continued Innovation

ThreatML is the industry's first and only solution that addresses a modern problem in cloud security: Providing actionable, relevant findings without sacrificing depth and breadth of detection. In other words, you'll only get the highest priority alerts surfaced without sacrificing full observability. And because it's automated, it requires very little tuning and maintenance.

One of our core differentiators is that our platform analyzes 60 Billion+ events daily and applies both rules and machine learning, to uncover known and unknown threats. With our new predictive modeling, we can use supervised learning to automatically prioritize those findings based on what's relevant to your organization—that is, what behavior is not predictable. This ultimately lessens the amount of investigating our customers' internal security teams will be required to do.

Supervised learning aims to massively lower the burden on already resource-strapped security teams without introducing gaps in coverage or timeliness of detection.

Understanding our Vision for ThreatML with Supervised Learning for Predictive Risk

Our supervised learning approach is the next natural progression for ThreatML and offers customers full coverage with even more highly relevant context for their security operations. The challenge with supervised learning is that it takes copious amounts of labeled data to train the models. Since we already have a real-time rules engine, ThreatML provides superior insights compared to other vendors, making it easier for us to label data very efficiently.

Thanks to these capabilities, we're able to pioneer supervised learning to bring the industry high-efficacy threat detection. By enhancing ThreatML, our detection-in-depth approach surfaces the alerts that need immediate remediation, complete with the necessary context to take action.

Supervised learning presents the added benefit of filtering out behaviors that don't represent threats. They're predictable, so don't require an analyst's direct attention, allowing them to focus on fewer but higher priority alerts while having confidence that complete detection is not compromised.

With this next step in our machine learning evolution, Distributed Cloud AIP offers customers the industry's most relevant security alerts on both the known and unknown threats without sacrificing deep observability collection.

Here's how we do it:

Rules are built-in and configurable, meaning out-of-the-box you get detection on known security risks and compliance triggers. This is the first method of detection to keep your organization secure.

Supervised learning is the prioritization piece of ThreatML. This looks at an event AND all relevant surrounding data to make predictions about the event. The goal is to answer the question: "Given the historical behavior in this workload, was this event predictable or not?" Our approach follows the "Observe, Detect, Know, Respond" pattern.

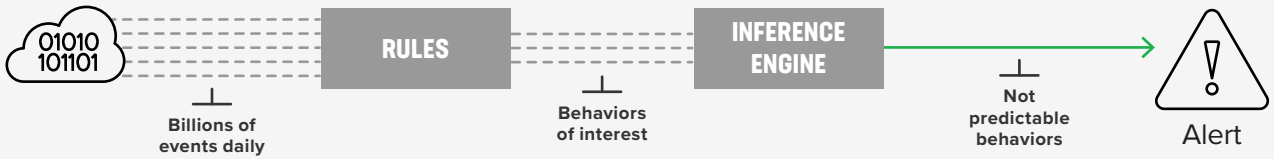

Figure 1: ThreatML's path to high-efficacy alerts in context.

Observe

Distributed Cloud AIP collects the richest telemetry across the full cloud stack, bringing in data from containers, Kubernetes, host machines, and the cloud management console.

We process over 60 billion events each day across our entire customer base. This extensive telemetry collection is the foundation for our comprehensive approach to cloud security monitoring.

Customers benefit from both the **size** and continuous **collection of telemetry**.

Customers benefit from the size and continuous collection of this telemetry through having full coverage of all behavior within their cloud environment. This is a significant differentiator for Distributed Cloud AIP that fuels our multiple detection methods.

Detect

Surfacing meaningful security alerts from massive amounts of data requires a detection-in-depth approach. We achieve this through both behavior-based rules and ML-driven detection. Many cloud security providers force customers to choose one or the other, but you need both to accurately detect risk in your environment. And with Distributed Cloud AIP and ThreatML, you get both.

Our rules engine lets customers monitor the known threats in their environment. Rules are beneficial in that they capture risk within well-known behavioral patterns that customers define on their systems. Consistent monitoring is essential when watching for insider threats or providing a complete view of security posture during a compliance audit.

We also detect through supervised machine learning. Using these types of algorithms in cloud security requires labeling billions of events daily to train them. ThreatML takes a novel approach to supervised learning by taking all the data we collect and using the rules engine to label it real time. This allows ThreatML to look at and make predictions about all the behaviors being detected.

Rules plus machine learning allows you to detect both the known and unknown threats—and eliminate false negatives, or normal behavior that is actually malicious.

Know

ThreatML with supervised learning is our approach to improving customer Mean Time To Know (MTTK) and Mean Time To Respond (MTTR) by enabling them to take the right actions, faster. The goal is to lower the operational burden on already stressed security teams by surfacing only the behavior that is most important to the organization.

ThreatML is our inference engine that uses predictive modeling with labeled data taken from rule matches. Supervised learning takes this data and make determinations about whether or not the surrounding events signal predictable behavior. If the engine discovers any unpredictable behavior, ThreatML surfaces it as a priority alert with the relevant context that needs to be addressed.

Because of our industry-leading telemetry collection and proprietary in-house data platform, we are uniquely positioned to make behavior-based security and risk predictions about customers' environments—and in real-time, meaning customers can be much more proactive in addressing the most pressing alerts to their business. This is a key advantage compared to other vendors that often leverage a third-party data warehouse to store and analyze data, hindering real-time capabilities. Additionally, with supervised learning, complete security coverage, and the backing of a deep telemetry collection and ruleset, customers can have confidence that they are addressing the most prominent security needs to their business.

Respond

Even the best set of rules and ML technology can't replace human intuition, reasoning, and decision-making skills. Our approach combines rules and supervised learning to reduce the alert volume so that customers can hone in on only the most critical behavioral threats. Our platform does the hard work of full stack coverage and sorting out the relevant behavior from the noise to only surfacing the highest priority alerts that need attention—ultimately driving down key metrics like MTTK and MTTR.

We also provide security expertise through our 24/7/365 SOC team that can help support running the daily security operations. Additionally, Distributed Cloud AIP Managed Security Services can take over the detection and triaging of the high-severity issues and investigate alerts on the customer's behalf while also using our extensive telemetry collection and supervised learning.

The more our experts work with the customer's team, the more action taken on the customer end will inform ThreatML's models. Over time, we will evolve to understand which parts of the infrastructure are most important to the particular business. As a result, customers can focus on the highest priority alerts in their environment before those alerts become threats.

The Future of Cloud Security is Supervised Learning

Distributed Cloud AIP, especially ThreatML, continues to prove its commitment to addressing customer pain points and values. We're working to make our cybersecurity solution as proactive and efficient as possible to support the security professionals in ultimately managing the risk in their environment. We are building on top of the crucial foundational pieces such as our telemetry collection and data platform, rules, and evolving ML to continue to deliver the most relevant and actionable alerts without sacrificing detection.

Building on top of crucial foundational pieces such as **telemetry collection** and **data platform**, **rules**, and **evolving ML**, delivers relevant and actionable alerts.

We promise to constantly challenge the status quo, innovate to solve problems and help you to drive greater value within your organization. ThreatML with supervised learning is a major step forward in cloud security. While others try to catch up, we will continue to innovate. So stay tuned as we get closer to further revolutionizing cloud security.

Appendix

[1] "Access to Talent Driving Managed Services Opportunity," 451 Research, https://go.451research.com/2020-mi-access-to-talent-driving-managed-service-opportunity.html

[2] "IDC FutureScape: Worldwide Cloud 2020 Predictions," IDC, https://www.idc.com/research/viewtoc.jsp?containerId=US44640719

THREAT STACK: NOW PART OF F5

Threat Stack is now F5 Distributed Cloud App Infrastructure Protection (AIP). If you'd like to learn more about this solution, the company's Security Operations Center (including Distributed Cloud AIP Managed Security Services and Distributed Cloud AIP Insights), and more, feel free to contact our cloud security and compliance experts.

Let our experts take your cloud security worries off your shoulders, so you can get down to business. To learn more or to schedule a demo, visit our website today.