White paper

Integrating dual WAF in the CI/CD process for better security

Dual WAF helps security teams test new rules and mitigate attacks with real-world production traffic.

Verizon's dual web application firewall (WAF) technology improves security by adding an audit WAF for monitoring and testing purposes to the standard production WAF. With two WAFs configured to operate in tandem, security operations centers (SOCs) are able to make data-driven decisions and update security policies with 100% predictability and minimal delay.

The problem

Security teams are constantly hunting out new vulnerabilities, but detecting issues is only one side of the equation. According to the Ponemon Institute's 2019 "Cost of a Data Breach" study, it took organizations an average of 206 days to identify security breaches and an additional 73 days to fix them.[1] That a security vulnerability remains exposed for more than two months on average highlights not just failures of communication but also of development practices when it comes to executing security fixes.

Common Vulnerabilities and Exposures (CVE)[2] notices and emergent online threats put security operations centers (SOCs) behind the curve, forcing development teams to play catch-up when it comes to security. Current practices outline a "test and patch" response model that can take up to three weeks[3] to implement even under ideal circumstances. SOCs can't repeatedly stop operations to implement changes without undermining their web presence. Even after patches are successfully applied, rule-set optimization and settings adjustments require tedious and costly attention and risk further network disruptions.

Organizations relying on prior-generation WAFs can be obstructed by their own continuous integration/continuous delivery (CI/CD) practices, which face a continuous bottleneck in rule-set changes. During time-critical zero-day events, SOCs have the unenviable choice of risking outages or remaining vulnerable to known threats. Without a live testing environment, changes both big and small can have unforeseen effects on production traffic. But not fixing problems immediately can result in costly downtime or data breaches.

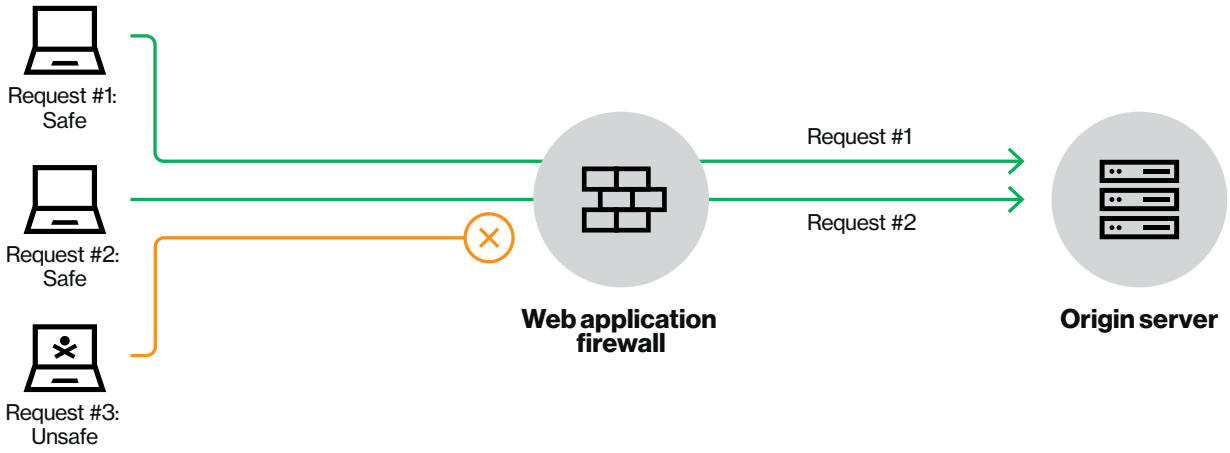
The solution

Dual WAF enables companies to run a production WAF and an audit WAF simultaneously on production traffic. Taking advantage of Verizon's waflz engine, dual WAF was designed from the outset to support multitenant integration and deliver high performance that's memory-constrained and deterministic.

A production WAF operates like an open-box WAF, with powerful and responsive tools for rule-set management and mitigation options for web traffic. Meanwhile, the audit WAF monitors the same production traffic and provides a testing ground for new rule sets. Changes to the audit WAF can be instantly promoted to the production WAF configurations, allowing for seamless updates without lengthy approval processes, disruptions or downtime.

In practical terms, that means that SOCs can effectively test patches in production and implement updates within minutes rather than weeks. This helps SOCs move toward a flexible and resilient "go live" approach.

Web applications firewall



Web applications and WAF

The modern internet relies on web applications as the means for customers and companies to reliably exchange information. A web application consists of servers running software remotely, accessed through a client interface such as a browser or a mobile application, and generally exposed to the public internet. These web applications are generally the entry point for individual users and encompass web services, business applications, online databases and communications.

While crucial to the online experience, the sheer range of functionality for web applications leaves them vulnerable to exploitation. Attacks at the Application Layer (Layer 7 of the OSI model) take many forms, including SQL injection (SQLi), cross-site scripting (XSS), and remote code execution (RCE). If successful, these attacks generally penetrate deeper into the exposed network, leading to data breaches and service interruptions. A common element of these attacks is malicious requests that exploit vulnerabilities in exposed application programming interface (API) endpoints.

A WAF serves as a gatekeeper between requests—both legitimate and malicious—and web applications. WAFs monitor incoming web traffic then filter that data according to assigned rules and policies. Legitimate requests are forwarded to their destination, while malicious requests are mitigated before they can cause harm.

How a WAF works

A WAF works by enforcing security policies between a user's requests and the web application.[4] The entire chain begins with an individual user, who accesses a web application through their client. Actions by the user, from mouse clicks to input text to passively created metadata, can generate HTTP requests for servers operating on the network edge or in a data center.

WAFs can be deployed at multiple locations: on content delivery network (CDN) edge servers, on a load balancer, on a cloud instance or on physical appliances that sit within a corporate network. The WAF is usually the first line of defense in front of the web or application server and is generally the first to receive an HTTP request sent by a client. If the request is determined to be legitimate, it will be forwarded to the web/application servers to generate a response. If the request is determined to be malicious, the WAF can perform various mitigation actions, such as denying the request by returning an HTTP 403 response. Questionable requests can also be logged for subsequent threat analysis.

Previous-generation WAFs

WAFs originated as physical hardware deployed directly within a corporate network. Because such appliance WAFs are network based and local, they have extremely low latency between verified requests and the destination server. Though some form of WAF is essential for security, those first-generation WAFs were costly to install, cumbersome to maintain and difficult to scale. Host-based WAFs were similarly constrained. Though generally requiring less upfront cost than network-based WAFs and offering more configuration options, host-based WAFs proved to be difficult to integrate and required significant operational resources to maintain.

Later, CDNs began offering cloud-based WAFs. Because a CDN's distributed edge servers are generally closer to traffic sources (i.e., users), the resulting performance was still competitive with appliance-based WAFs. These second-generation WAFs were regularly updated, easy to deploy and highly scalable, but their rule and policy systems failed to provide any form of transparency or configurability. Such "black box" WAFs left customers unable to customize their WAF to suit their applications.

Eventually, cloud service providers like Verizon recognized the limitations of black-box WAFs and began introducing "open-box" WAFs, designed with transparency and user accessibility as guiding principles. This new generation of WAFs empowered companies to tailor their WAF configurations to their applications and their security profiles and has been instrumental in containing the evolving threat of malicious traffic.

With full access to their entire event log, as well as comprehensive monitoring and reporting, companies deploying open-box WAFs are able to enjoy the full benefits of a perfectly tailored WAF that matches their unique requirements. By tuning rule configurations over time, security teams can reduce false positives where security flags legitimate traffic as malicious, adapt to changing application stacks and react better to the evolving threat landscape.

However, as companies adopted CI/CD practices, certain limitations of the new generation of WAFs became apparent. WAFs need to be integrated not just with a security information and event management (SIEM) solution, but also within a CI/CD process. A single WAF, even when fully optimized for existing applications, can prove to be a development obstacle, as new code is constantly deployed and changes need to be continually tested in multiple environments before they can be merged and deployed.

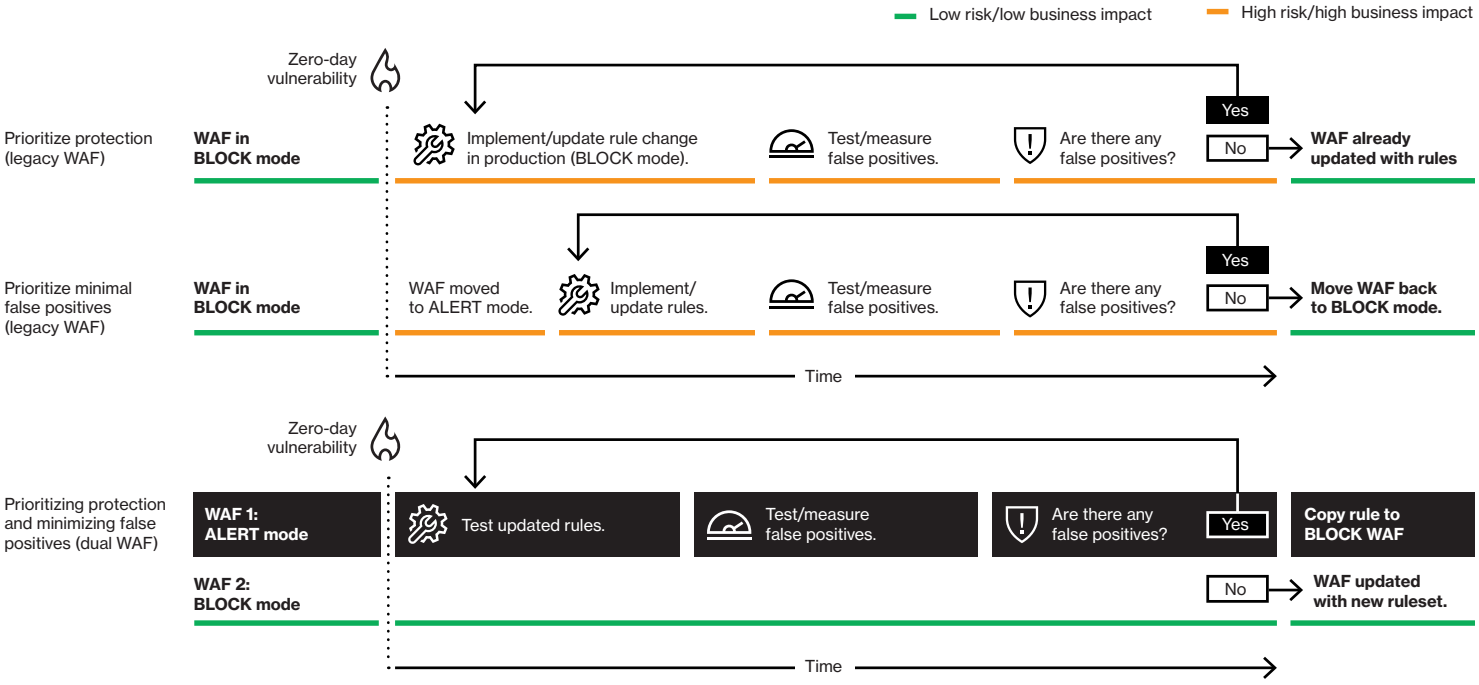
When a typical WAF serves as a gatekeeper for production traffic, changes to the WAF configuration or the underlying application can create high-risk situations. This creates a bottleneck in change validation and exposes operations to inaccurate assumptions, disrupting CI/CD practices that rely on consistent and automated testing.

The solution would be a WAF that integrates with the deployment pipeline, without compromising security or performance.

Introducing dual WAF technology

Verizon's dual WAF is built around the concept that two WAF configurations running on the same live production traffic can facilitate continuous, risk-free updates of security profiles, as well as application updates with complete predictability. To achieve this, dual WAFs were designed with the natively developed waflz engine to deliver the full range of open-box WAF capabilities while also enabling production testing with instant profile swapping. The end result is an always-on, always-accessible solution that takes full advantage of Verizon's extensive edge network.

Dual WAF differentiates from regular WAFs at the point when a user's request is received by the edge server. Here, the request is inspected simultaneously by both the production WAF profile (with full mitigation capability for malicious requests) and an audit WAF profile (running full sets of the proposed WAF config in monitoring mode). Instead of running a WAF in a staging network or a lower environment, which can only analyze nonproduction data, the audit WAF inspects the same traffic the production WAF is protecting and generates its own data based on independent configurations.



Continuous integration, delivery and deployment overview

Web application development often requires both rapid updates and integrated testing. Continuous integration (CI) improves on traditional software development by instituting a pipeline centered around frequent code merges in shared repositories and by implementing automated build and testing. Continuous delivery (CD) takes this a step further, bringing committed changes to production in a safe and sustainable manner. Continuous deployment might be seen as a further evolution in this progression, automating the release process with minimal lag between automated tests and production release.

Implementing dual WAF within the CI/CD process

Dual WAF configurations align with the CI/CD process by making application security an agile element of the development pipeline, rather than a roadblock that introduces tradeoffs. Where standard WAF requires preproduction test environments with either nonproduction or replay traffic, dual WAF allows for testing of security configurations with production traffic via the audit WAF, which shows real-time changes in behavior compared to the production WAF.

The following scenarios explore three of the biggest challenges faced by developers, as well as the solutions offered by employing a dual WAF configuration.

Scenario 1: Enhanced threat modeling with dual WAF support

CI/CD focuses on discovering issues throughout the development pipeline; a common understanding is that earlier discovery leads to easier fixes. However, for an application security team buried under a mountain of security tickets, CVEs and other requests, the operating model can quickly devolve into a frantic game of Whac-A-Mole: constantly executing high-priority fixes from a massive backlog without addressing the system that lets unattended issues grow into problems.

This failure to act can stem from numerous areas, a lack of direct accountability in policy development procedures or reliance on outdated systems among them. But whatever the cause, security vulnerabilities will inevitably cost a company money and damage its brand image. The Ponemon Institute reports that a data breach for a U.S. company will cost an average $8.19 million, with higher costs linked to longer exposure. With a quarter of the U.S. companies surveyed having reported a breach,[1] the best time to begin implementing serious changes is before they become an issue.

Organizational alignment is the most important step toward better security execution, and SecDevOps (also referred to as DevSecOps, DevOpsSec or Secure DevOps) has proven to be a robust approach to the CI/CD software development life cycle (SDLC). This ensures that the security team has a seat at the table throughout the product development process, as opposed to constantly playing catch-up with security patches. SecDevOps strives to integrate security with every aspect of the CI/CD pipeline, leading to better communication, clear lines of accountability and actionable response plans.

Integrating security as an active element of development requires recognizing that security must be more than just bolted-on control systems. Data security is an essential element of any web application, but a SecDevOps approach transcends the SDLC to include changes from an operational perspective. Verizon's dual WAF offers security teams powerful tools to achieve this goal, by enabling WAF testing at every stage of development.

The process begins with threat modeling from the very beginning of development; if Layer 7 vulnerability is inevitable, it's crucial that developers understand their operational environment and any inherent and potential threats. Verizon's dual WAF enables a "build everywhere, test everywhere" approach that gives developers effective tools for comprehensive testing, from using a waflz build in local and QA environments all the way through to production testing.

By enabling configuration adjustment using real-time production data throughout the SDLC, developers can refine their high-level threat model with increasingly granular information from the dual WAF logs. As attack vectors are exposed, dual WAF gives development teams accurate data on patch effectiveness and the tools to patch at low risk.

Scenario 2: Integrating and automating test-driven development while debugging

In a complex online business application, bugs aren't just possible, they're inevitable. Companies can expect a constant stream of bug reports, security tickets from internal scans or external bug bounties, CVEs, and even government notifications relating to security vulnerabilities. A product development team typically follows a "bag and tag" process, where newly reported bugs or feature requests are added to a backlog, to be groomed and resolved as developer time frees up.

This reactive approach puts developers on the back foot; as bug reports pile up, development teams can struggle to balance ongoing projects and product releases against making urgent fixes. The constant stop-and-go caused by efforts to fix bugs can lead to significant delays in code releases and can have a cascading effect, as project dependencies are affected by the weight of accumulated delays.

The best way to reduce the overall impact of bugs is to identify them early in the development process. Bugs emerge at all stages of the development cycle, but as a general rule, the later in the cycle they're uncovered, the more costly they

are to fix. Integrating test-driven development into a CI/CD framework catches bugs earlier and considerably reduces their overall impact.

While many categories of bugs are outside the purview of this white paper, Verizon's WAF provides a secure platform for test-driven development that is accessible at all stages of the CI/CD master process. From the earliest testable code until after deployment, the combination of waflz and the dual production/audit WAF configuration makes it easy to test changes under any conditions. This frees WAF from the security silo and gives development teams a clear view into how their patches will behave early in the development cycle, during QA and after production release.

Complexity breeds complications, and WAF is only one component of SecDevOps. By supporting automated unit testing throughout the pipeline via the waflz engine, issues can be resolved at any level without costly downtime. The dual WAF configuration enables testing of multiple production traffic profiles and supports instant go-live and rollback, so developers can be confident their fixes are working properly.

Scenario 3: Stepping outside the sandbox with virtual patching

Patches never happen in a vacuum. While development teams would preferably apply patches during low-traffic seasons and off-peak hours, reality often demands immediate action in nonoptimal conditions. High-traffic seasons and zero-day events place outside pressures on security updates, while infrastructure upgrades introduce large-scale and disruptive change management events. Development teams are often put in the unenviable position of either letting vulnerabilities go unpatched, or explaining to management why an exception to code freeze should be made.

SecDevOps adopts a broader perspective. It means not just bringing security awareness to all levels of the production process, but also bringing institutional awareness to security requirements. Dual WAF provides security teams with the ability to present solution plans to leadership with demonstrated certainty by using actual production data, while keeping those same teams apprised of traffic behavior. The audit WAF profile can provide predictive data on security changes and take the guesswork out of patching, ensuring reliable protection while eliminating downtime.

Staging environments, sometimes referred to as sandboxes, are a subset of servers that attempt to mimic production environments. Sandboxes are one of the newer testing platforms offered by some CDNs. While the ability to test configurations on lower environments is certainly preferable to no testing at all, a staging network has distinct disadvantages in replicating real-world conditions. Unit tests can be performed in a staging environment, and replay traffic can even be used to realistically reproduce some production conditions. No matter how good unit tests are, though, they're not a replacement for the sample size of the production traffic. The growing complexity of web applications with countless edge cases can diminish the potential for certainty and reproducibility of test results in a lower environment.

Unlike a sandbox environment, dual WAF is designed to have two security configurations run in tandem, monitoring live traffic data. The audit and production WAFs handle the same requests that the server receives, allowing companies to implement patches virtually without the need of simulated requests. This level of deployment maintains consistency and stronger security throughout the SDLC.

With dual WAF configurations, development teams can implement changes in the audit WAF with no disruption in traffic behavior, monitor the result with production traffic, and promote the effective changes to the production WAF. Real-time data from both production and audit WAFs can be superimposed to study the overall impact of the change and to support the decision-making process.

Practical applications: Dual WAF threat modeling

Threat modeling is a crucial component of network security practices, but implementing it within a CI/CD pipeline requires operational agility. The threat modeling process can be adapted on an individual basis to take advantage of a dual WAF configuration. Unlike previous SDLC approaches that implemented WAF technology as an end-stage coverall, Verizon's WAF was built to integrate testing and security at every level. Here are some standard steps in the threat modeling process, along with a brief description of how dual WAF can assist development teams:

Application assessment

A proper threat model begins with a comprehensive understanding of the assets at risk. That includes both the actual data the web application will be sending and receiving, the development platform to be employed, and less tangible assets like customer goodwill. From the outset, WAF should be configured to provide top-level protection to critical databases, as well as blanket coverage for any sensitive customer data.

Environmental threat assessment

Once a model of the assets in play is developed, the next step is to describe the likely threats to the web application, as well as their probable sources, including external hostile actors, internal errors and third-party complications. As the SDLC progresses, new threats will inevitably be uncovered, so flexibility is essential. Verizon WAF is highly configurable with various sensitivity settings, providing development teams with improved threat detection accuracy—even for previously unclassified traffic.

Vulnerability evaluation

Once a matrix of assets and threats is built, conduct a vulnerability evaluation. While it's not always possible to know where attacks will target, a company can map out anticipated vulnerabilities by connecting probable vectors to likely objectives. Dual WAF gives companies access to a robust and constantly updated rule set, ensuring that uncovered vulnerabilities are accounted for.

Priority modeling

Given the array of threats facing web applications, security triage is a fundamental element of threat modeling. By analyzing factors such as cost, likelihood of attack and scope of impact, a company can assign priority levels to its vulnerabilities and map the levels to action plans. The independent audit WAF profile enables early testing of rule-set updates, allowing for a more accurate assessment of threat impact.

Countermeasure modeling

Once a threat triggers an assigned priority plan, it's time to take action. That begins with an assessment of previously deployed countermeasures to avoid solving the same problem twice or disrupting related security. If existing measures are insufficient, develop targeted countermeasures to contain the risk. Dual WAF provides transparent and comprehensive policy overviews to ensure a clear understanding of your security profile, as well as tools to address emergent threats.

Deployment and verification

Having identified appropriate countermeasures, companies using dual WAF can swiftly push any updates to their audit WAF. Since no contingent systems need to be disrupted, a patch takes only minutes instead of weeks to start testing. Once results have been verified, the audit WAF profile can be live-swapped to the active WAF with no interruptions.

Refinement and optimization

After updates are in place, proper traffic monitoring is essential to manage unforeseen issues. Small-scale adjustments can be immediately tried on the audit WAF and transferred just as quickly to the production WAF configuration.

Conclusion

Verizon's dual WAF provides companies with the tools to integrate security at every level of their development process. An audit WAF operating on live traffic data allows teams to demonstrate solutions plans with complete certainty, making it easy to test changes before rolling them out. By taking the risk out of testing, dual WAF empowers developers with the SecDevOps principle of implementing security at every stage of the CI/CD pipeline. The end result is an organization aligned for better security execution.

verizon

1 "2019 Annual Study: Cost of a Data Breach," Ponemon Institute, 2019.

2 CVE is a constantly updated list of known cybersecurity vulnerabilities, consisting of standardized descriptions of common identifiers.

3 "Rewriting the Rules of Patch Management" IBM, 2017.

4 "Information Supplement: Application Reviews and Web Application Firewalls Clarified ver. 1.2," PCI Data Security Standards Council, Oct 2008.