

Next Generation WAF (NG-WAF) With Amichai Shulman

Listen to Podcasts: <http://www.imperva.com/resources/podcasts.asp>

Brian Contos: Welcome to the Imperva Security Podcast. I'm Brian Contos, Chief Security Strategist for Imperva. Joining me today is Amichai Shulman, CTO and co-founder of Imperva.

Welcome to the show, Amichai.

Amichai Shulman: Thank you, Brian. It's always a pleasure to be here.

Brian Contos: Amichai, we've discussed dozens of different subjects on this podcast. One of the ones I think is really timely is WAF. And of course WAF is one of the products that Imperva makes, but there's been a lot of changes in WAF over the years.

What I'd like to do is look at a chronology of WAF. When did WAF really hit the market, and when it did what was it like?

Amichai Shulman: Sure. That's a great topic. It's time we mentioned the "W" word in our podcast series.

Brian Contos: [laughs] That's right. So, Amichai, when did WAF come on the scene? It's not really a new technology; it's been around for quite some time, hasn't it?

Amichai Shulman: I think I saw the first real WAF come out to the market somewhere around '98, maybe '99. It could be that some prototypes were available in '97.

Basically, the whole idea of web application attacks was very new, and people didn't really know what to expect from a web application firewall. So basically the first web application firewalls were trying to address the very basic SQL injection issue, very basic cross-site scripting issues, and then some of the emerging known vulnerabilities in web applications that started to become more popular at this time.

I think that the real deployment of web application firewalls started in the year 2000 and on, and by then web application firewalls were all of the proxy types. Again, because people were not really aware of what to expect and what are the real threats, there has been some strange variations on the WAF topic.

For example, I know that one of the products back then was very focused on defacement issues. And their solution was to really digitally sign everything that was supposed to out of the web application firewall, and then compare each response to that expected result. This is, of course, unthinkable in a world of real web application where you have changing data going out of the application each and every time. But that was the situation back then.

Brian Contos: Yeah. And I think if we're talking about 2000, web defacement was one of those historic, key issues that a lot of people were looking at. It was a lot of people trying to

get publicity. "Hey, I hacked this website," and get a little bit of notoriety. I think that pre-dates a lot of the industrialization of hacking that we've been talking about.

Out of curiosity, I guess it was 1997 or so when I started working with network-based firewalls. I think one of the first was CheckPoint that I was leveraging. Why has there always been this dichotomy between the web application firewall and the network firewall?

Do you have any insight into why perhaps CheckPoint, or later on Cisco with the Cisco Pix, didn't try to address this market and sort of left it as a standalone market even from the early days?

Amichai Shulman: I think that they were very busy trying to make their products work and provide more value to their customers based on their kind of offerings, to start chasing a totally different market and a problem.

The problems are very different. You can't really use the techniques and technologies you can find in network firewalls for web application firewalls. Again, you have to remember that back then no one really understood the problem well. Only after the year 2000, where e-commerce really got a boost, no one thought that it was interesting threat to organizations and to businesses.

Not until businesses came to the realization where there was no physical entity or physical presence for the business and everything was virtual, then the threat of application security became an important one.

By then those companies were very focused on the network-related programs, on network access controls, VPNs. Some of them tried to go into load balancing and so on. And they did not have the capacity to start developing knowledge and technologies in the domain of web application.

And that's where smart companies started to evolve. Some of them are in fact are much larger companies today, creating their own market.

Brian Contos: I guess thinking back, there was so much focus on things like state inspection and IP fragmentation attacks and evading firewalls. It was trying to get our hands around a very network-centric situation.

Amichai Shulman: Exactly. And until the point got their hands around those problems, you couldn't notice the application level problems.

Brian Contos: Exactly. That does make a lot of sense. So essentially to summarize, the first generation of WAF, there was a couple of different flavors, it sounds like, in terms of web defacement, focusing on very simple SQL injection and cross-site scripting, although all these issues, of course, still unfortunately are important today.

When would you say the second generation of WAF kicked in, and what were some of the new capabilities that that second generation began to offer?

Amichai Shulman: I think that really second generation WAF started when Imperva introduced the SecureSphere product into the market. There were, I think, a couple of very important changes.

One of them was trying to make the WAF a less-intrusive device in your network infrastructure. In the early days of the net generation, a web application firewall would have

Next Generation WAF (NG-WAF) With Amichai Shulman

been installed as a sniffer beside the web application server. Today, most of the deployments are as a network bridge in front of the web application server.

So these are very non-intrusive deployment methods. And that was one of the things that allowed web application firewalls to become more acceptable in production networks.

The other thing that characterizes the second generation of web application firewalls is that they became more of a tool for security administration, for IT operations, requiring less intervention and interaction with the application programmers.

This is a key factor for this kind of solution to become valued and valuable for an organization, because you cannot have your security operations rely on application programmers.

Brian Contos: So it sounds like the first generation very much leveraged by the application developers themselves and then the people running the web apps, is that fair to say?

Amichai Shulman: Well, it wasn't leveraged by them, but if you wanted to deploy a web application firewall, you'd need a very tight integration with those guys. Certainly you'd need their cooperation. Some of the devices required changes to the application in order to provide any security value. So this was clearly an organizational problem.

Brian Contos: So it wasn't a question of who's using it, it's the skill sets required to actually implement it and make use of it.

Amichai Shulman: Exactly.

Brian Contos: So here we are today, 2010, entering into what some might call the third generation of WAF or "WAF: the Next Generation." The evolution, what are some of the new whiz-bang things and capabilities and directions that WAF is going in now?

Amichai Shulman: I think that by now second generation WAF are coping pretty well with technical web application vulnerabilities, SQL injection, cross-site scripting. And most of them can handle pretty well as far as evasion techniques and so on.

It's now the time where the web application firewall takes a step further in protecting web application logic. As defense is improved, attackers are looking for yet another method to attack applications. And as technical vulnerabilities harder to exploit, attackers resort to business logic flaws.

In order to cope with those, you'll have to arm your web application firewall with capabilities such as business abstraction, where you are able to map the various technical application elements into business transactions and then apply some rules and security policies based on this mapping that would allow you to detect business logic attacks.

And business logic attacks rely on web traffic that looks very much like normal traffic in terms of structure, but it diverts from normal traffic by its rate or flow or other characteristics that are not part of the request itself. So they need a totally new set of capabilities in order to cope with it.

Another thing that we'll have to see from web application firewalls is an improved anti-automation capability. We're seeing more and more attacks today being executed not by an individual from a single laptop, but rather by an organized network of zombies

operated through a single command and control center, sending attack vectors in an automated way.

Brian Contos: Amichai, how would that differ from a preventative perspective from a single individual with a laptop, as opposed to an automated bot? What are the key differences there, and what are the key countermeasures?

Amichai Shulman: For instance, if you have a human being behind this attack and you start blocking requests, then probably after the second or third blocking incident, that individual would quit and go look somewhere else.

If you have an automated program keep sending you attack vectors - for example, if that program is now trying to exhaustively search for user names or passwords, or if that program is searching for resources that might be available through your application without authentication, it will not stop because of two or three blocking attempts.

One of the key things that should be added to a mitigation technique is the ability to slow down the rate of the attack, not necessarily just block the requests. So there is a new set of capabilities here that should be employed.

One of the things that we must remember when we're talking about this new breed of business logic-related automated attacks, is that it is very different to distinguish a good request from a bad request. So the mitigation technique should be such that even if a mistake is being made by the web application firewall, then that mistake will not affect an actual user but will have a dramatic effect on an automated device.

So for example, using CAPTCHAS is one technique that would allow a human being to go on with the process while disrupting an automated device.

Brian Contos: There's a couple really key takeaways I got from that. That's incredibly good stuff. I love the direction of it going from a solution that is still focused on the technical challenge, but also understands that there's risk that must be abstracted to a business layer to be properly explained to the analysts and the users looking at these things, writing rules and policies.

At the same time, the attacks are taking advantage of business logic, which isn't necessarily a technical flaw or a "hack, " but just taking advantage of the way the program was written and using it in a way that perhaps wasn't necessarily intended.

So in order to address business logic attacks, you have business-level abstraction. And then to take that into the real world, knowing that for that to be 100% is a ridiculous idea, so the notion that something that is very binary - allow A, block B - simply doesn't apply. So you use mitigation techniques that are more adaptive, if you will. As you mentioned, CAPTCHA, or I suppose, page redirection.

And there's probably a number of other things that can be done to slow down or prompt or ask for human interaction where, as you mentioned, an automated attack or some time of automated bot simply wouldn't be able to deal with that injection into the transaction.

Amichai Shulman: Exactly.

Brian Contos: I get the business abstraction stuff. I think that's great. I think that's certainly what we're hearing from organizations in the field. I'm always being asked, "How can this technology be up-leveled to more of a business-centric view, " for people who might

application developers and they simply don't speak SQL. They don't really understand this as well. I understand the mitigation changes.

Any other core things in this next generation of WAF that's really going to be important going forward to address these changes in the threatscape?

Amichai Shulman: There's actually one very important set of capabilities that I think web application firewalls must incorporate in the following decade. Looking back into the 2000s, attack traffic was very scarce. Because it was manual, it required a relatively high skill. If you look at the number of attack vectors being sent into an application at a given time, that number would be very low.

This has changed with the automation of hacking and the industrialization of hacking. Today you have much more attack traffic coming into an application at any given time. And even if you have a very accurate web application firewall, still the number of alerts that you will be getting is very high.

Most of those alerts, again, are originating in an automated device. And what I think is very important in order to make sure that operators and security response people are focused on the really dangerous attacks, on the real threats. Is to be able to very quickly identify those attacks that are coming from known hacking platforms, be those anonymous proxies or known compromised servers, or even desktops that have been known to have been part of an active bot-net.

And by incorporating the capability to have timely real-world information about hacking platforms and current attack vectors being used by those platforms, web application firewalls could quickly identify all this automatically-generated traffic and put it aside, leaving the operators with the real threats, the more sophisticated attacks, to look at and to handle.

Brian Contos: So it's a bit like knowing the reputation of the source of the traffic. And then being able to - as you mentioned, there's so many alerts, you have to prioritize them. If you can say, "Look, we know systems A, B, and C are part of this bot, or coming from a network that we know is attacking other sites..." what have you.

Based on that, we know how to prioritize it. Or essentially maybe just simply block it and not alert based on events from those systems. I would think that would do a couple of things. One, it would probably allow every individual to benefit from every other individual out there on the Internet, in terms of attacker information, knowing what's happening.

And also, it sound like locally it would help the individual administering the WAF, because now they're not dealing with 5,000 potential events a day, but now they're dealing with maybe five or six events a day that actually matter.

Amichai Shulman: Exactly. And I think that this capability of being proactive, looking for attack platforms and identifying them very quickly around the world, and then feeding this information into a web application firewall, is going to be a crucial capability for web application firewalls.

Brian Contos: So if I sort of take everything we've talked about, the first, the second, and we'll call this third generation of WAF. At the first level it was, one, understanding the attacks via blacklists and white lists and various more advanced mechanisms to detect things like cross-site scripting and SQL injection and known technical attacks.

Next Generation WAF (NG-WAF) With Amichai Shulman

Then it's also this introduction of not only understanding attacks, but users and profiling, and understanding how people are interacting. And I know something that we've talked about before in the past is this notion of vulnerability data and interacting with vulnerability assessment systems.

Application-level vulnerability assessment systems that can say, "Here's your machine, it has these vulnerabilities. We can virtually patch them or alert based on them, " whatever the case may be. But we know the vulnerability level.

Mix it with this attacker reputation we have, and I suppose at the bottom of all this, we'd be able to correlate all that data together to generate some type of adaptive response, which is a CAPTCHA or a block or an alert or something.

But that would be based on all these different fields and flavors, as opposed to just one single instance of an event coming through your firewall. You're basing it on this grand view of both local and external perspectives. That's something very unique. I don't think any solution has tried to incorporate those various flavors to date.

Amichai Shulman: It's very unique, and there's actually more to it. You mentioned the integration with vulnerability assessment tools. I think that, again, this is something that next-generation WAFs will have to do much better on, and that's integrating with systems in the enterprises. And vulnerability assessment tools are one of them.

But we have database security systems, we have data leakage preventive systems in our organization. And that, again, is a direction for web application firewalls to grow in.

I don't know about other vendors. I certainly do know that we are constantly working to incorporate these kinds of capabilities. And entering this second decade of our century, I think that our product at least is going to quickly incorporate much of the stuff that we need in order to give an answer to threats, to upcoming threats, according to what we just discussed.

Of course, I think that as we incorporate more capabilities, we're going to see attackers take their next step. Of course us, as well as other web application firewall vendors, will have to quickly adapt throughout this decade to the new threats and keep chasing attackers, hoping to be one step ahead of them.

Brian Contos: You mentioned a very important and, I think, very subtle point, that web application firewalls is probably a piece of the solution, in and of itself it's not - I tend to think of it as a piece of the data security story, where there's databases and data stores and other areas where you can get data-centric data.

And of course this interview is about WAF, and so we're very focused on that. But it is important to understand WAF is a part of the greater whole from a data security standpoint. Just like network-based firewalls, there's more to network security. We see the same thing on the data side.

One question, and I think probably our last question, to wrap things up is: how is this sort of movement or this adoption of cloud-based computing and this acceptance of people leveraging services like mass security service providers. Or maybe some hybrid of enabling technologies such as end-point WAF on your network, and you've also cloud and you've also got some MSSP.

Next Generation WAF (NG-WAF) With Amichai Shulman

How is that really going to impact the world of web application security? Clearly it seems like it will make it a little bit more difficult to manage. You don't have everything in your data center any longer. But what are some of perhaps the not-so-obvious things that that's going to change?

Amichai Shulman: I think that eventually organizations are going to have hybrid architecture in the ways you just mentioned. Some of the components are on-premise, some of them are externally managed, and some of them are actually in some type of cloud.

One way for organizations to go about is to look for individual solutions for each part of their architecture. Another option is for WAF vendors to have solutions that can be adapted to the different architectures. This is not only a technical issue, it's sometimes an organizational or a cultural issue.

Having the ability to either operate as a WAF in the cloud or have your WAF operated as a managed service, and then at the same time, have the same capabilities installed on an on-premise device. These are real challenges for WAF vendors, and I think that organizations would benefit from vendors who can provide solutions for all their architecture, rather than taking a separate vendor for each part of their virtual enterprise, let's call it that way.

Brian Contos: Clearly, if I have a WAF local to my enterprise, I also have a WAF that's maybe managed for some other service, and I'm using cloud computing for something, and they happen to have WAF there, I'm able to get a view or some type of aggregation of all that, it really allows me to my hands around my security posture, just as if everything was local.

I think ultimately, just because you're using managed services, just because you're using cloud computing, it doesn't mean you've transferred the risk to somebody else. It's still your sensitive data you're protecting, it's still your risk. So having visibility into that will remain key.

And to your point, certainly that's not just a technical challenge anymore, but now you probably have service-level agreements and contracts and SLAs that have to be approved by legal and this, that, and the other thing. It certainly adds some business complexity to things. It probably will take things a little bit longer to roll out, but hopefully the end result is something that's more robust.

Amichai Shulman: Yeah, definitely. Because I think long-term, organization cannot keep a vast plethora of different devices doing basically the same thing, handling the same risk. As you said, you'll eventually need a unified control, a way to set a unified set of policies for all part of the virtual enterprise.

Brian Contos: Amichai, I'd like to thank you.

Amichai Shulman: Thank you, Brian, for the opportunity.

Brian Contos: If you would like to learn more about this subject and Imperva, please visit <http://www.imperva.com> or send us an e-mail at blog@imperva.com.



Headquarters
3400 Bridge Parkway
Suite 101
Redwood Shores, CA 94065
Tel: +1-650-345-9000
Fax: +1-650-345-9004