

Discovery, Assessment & Classification – an Interview with Dana Tamir, Senior Product Marketing Manager with Imperva

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. If you would like to learn more about this subject and Imperva, visit <http://www.imperva.com>, check out our blog at <http://blog.imperva.com>, follow us on Twitter at <http://www.twitter.com/imperva>, or send us an e-mail at blog@imperva.com.

Joining me today is Dana Tamir, Senior Product Marketing Manager with Imperva. Welcome to the show, Dana.

Dana Tamir: Hi, Brian. Thank you for having me.

Brian Contos: Dana, before we begin, could you tell us a little bit about what it is that you do at Imperva?

Dana Tamir: As a product marketing manager at Imperva I'm leading the launch activities for our new SecureSphere discovery and assessment solution, this solution helps organizations map out databases on the network, finding any rogue servers that might be out there, classifying data hosted on these servers, and assessing any vulnerabilities that may put that data at risk. We combine analysis of these scans, provide organizations with a unique ability to manage risk to data with a focus on data itself not just the platform that hosts that data. By understanding risk to data, organizations can make better decisions and better plan their remediation efforts.

Brian Contos: Dana, what is DAS, why should organizations care about this?

Dana Tamir: DAS is "Discovery and Assessment Server," and there are a few different use cases that we find in organizations out there. The first one is discovery and classification. When you don't know which databases are on your network or what data they contain, it's very hard to manage and protect that data, so for many organizations this is a huge challenge. The dynamic nature of IT environments means that these environments always change. They have business needs that dictate changes to the applications, which effects the database requirements on the back end, so they either have changes in the database platform itself, or the data that it contains.

We also see cases in which databases are cloned or created for test environments. Sometimes they're cloned with production environments and may contain sensitive data. And then these servers are just left out there. They're forgotten. They're not protected. If you don't know about these servers, and you're not aware of these servers, there's a higher chance that these servers are going to be attacked, and you'll experience a data breach.

So with database discovery and classification, we're insuring that you're fully aware of these new systems and any changes to existing systems. Now there are various privacy and data

protection laws that require protection of information collected and maintained on individuals from protecting the data from disclosure or misuse. You can't really protect it, if you don't know where it is.

Brian Contos: Yes, that makes a lot of sense. I know for example, even on the network side, when you're talking about network firewalls and switchers and routers, over time things get forgotten in data closets and people leave and contractors move around, and things get easily forgotten. And certainly during merging or acquisition activity -- Bank A buys Bank B -- sometimes it can be very difficult to determine those. Is the solution such that it takes a snapshot and most organizations would run this on an ongoing basis, either daily, weekly, monthly, something like that? And if that is the case, how often would an organization want to reassess or rescan for these databases?

Dana Tamir: That's a very good question. It depends on the organization. In most organizations, a periodic scan of their databases makes a lot of sense. In environments that change less frequently, maybe once a quarter, if it's more frequent changes, maybe once a week, or once a month. Anytime you schedule this run, you want to compare it to previous runs, and you want to sure up any changes that occurred since the last run, and that's one of the important things that would help you understand how the environment changes and where the data is hosted and where you should focus your efforts.

Brian Contos: So, it seems pretty intuitive when we talk about A) Let's simply discover where the database is. Here's a database I knew about, and here's three I didn't -- replicated or rogue system or what have you. It also seems rather intuitive to B) Then look inside that database and determine what sensitive data resides in that database and help to classify that. You mentioned another piece, more of a risk-based piece looking at vulnerabilities. Is this similar to a network vulnerability scan?

Dana Tamir: Absolutely. The second step is, after you know where that data is, is to understand the risk it might be exposed to, and so there's a need for configuration audits and vulnerability assessments, which insures that databases are configured in a secure manner and according to industry best practices. These are practice scans that discover vulnerabilities and misconfigurations and enable organizations to be proactive in their remediation efforts taking care of these known vulnerabilities and reduce the risk of an exploit.

Brian Contos: OK, so in addition to scanning for vulnerabilities, I can also do audit checks, like you said, best practices or certain types of configuration standards.

Dada: That's correct.

Brian Contos: Are there any sort of leading industry configuration audits that the data solution will check for?

Dana Tamir: Absolutely. We have predefined policies for assessing compliance with PCI. We have SOX policies. We have specific policies for compliance with NIST standards, and CIS benchmarks. So all the commonly used industry best practices are existing in the product. That includes the STIG, almost forgot to mention that. They are available in the product out of the box, but a nice-to-have option that you have in the process is also to create custom policies. So if your organization has built its own policy and its own preferred configuration of database servers, and you want to evaluate your database against your internal policy, you can build the policy to match that policy and find any server that doesn't comply with that policy.

Brian Contos: And for those of you out there that heard Dana mention STIG, that is Department of Defense guideline that's commonly used by groups like DISA for baselines to systems. Dana, so we know a little bit about the theory and the technology. What are some of the real life use cases that we've seen in the field with people actually leveraging the solution?

Dana Tamir: So, I mentioned before, there's a use for discovering classification, so either for compliance with privacy laws or data protections laws, we see a use of the product for scanning databases to understand where sensitive data resides. That's the first step for preventing any data leakage and insuring that data is protected. There is also a use case related to PCI. PCI DSS requirement 3.2 states specifically that you should not store sensitive authentication data after authorization. This is regarding the magnetic stripe information. Merchants are not allowed to keep that information anywhere within their infrastructure.

But if they need to confirm that the data isn't stored there, they need a way to scan their databases to ensure that that data does not exist, and if it does exist, remove it from the database. This is an example where you not only scan the database to know about the existence of data, but you actually take action, and remove that data from the database.

Brian Contos: So, again, it's not just a tool for reporting. You actually take corrective action.

Dana Tamir: Absolutely. Another example of corrective action is based on vulnerability assessments. If you run a vulnerability assessment, you discover that there is a vulnerability on your server that maybe a patch exists for, or maybe the server is just misconfigured. You can remediate that. SecureSphere specifically gives you an advantage, because when SecureSphere Discovery and Assessment Server is integrated with our database firewall, we can also provide a virtual patching solution. Virtual patching allows you to block attempts to exploit known vulnerabilities on the network, without deploying the patch. That way you remediate, or you have a mitigation for that vulnerability, without actually changing anything on your database or application.

Brian Contos: Yeah, and we all know, as we get closer to the holiday season, and so many organizations are moving into a code freeze, if they have discovered a vulnerability, in most cases they're not allowed to make any changes, or patches, or configuration adjustments. Virtual patching certainly can assist in that, especially since it can be done in a pure blocking mode, or in a monitoring mode, so at least you're aware that someone's trying to exploit that vulnerability. I think that's a great real-life solution to a problem that many organizations face around their applications and databases.

Dana, anybody familiar with Imperva is probably familiar with Scuba. If you could, maybe just give a brief explanation of what Scuba is, and how DAS is quite different, and beyond what Scuba offers.

Dana Tamir: That's a great question. A lot of our customers confuse Scuba to be our vulnerability assessment solution, where Scuba is really just a small Java utility that you can install on desktops to assess a vulnerable database server. It assesses one database server at a time, and it's really not an enterprise solution. You cannot get any consolidated reports. You cannot run vulnerability scans across the environment. It doesn't actually have any discovery or classification capabilities in it. It's a pretty basic solution. SecureSphere Discovery and Assessment Server is an enterprise-based solution that helps you understand the environment, the databases in it, the data that they're hosting, and the vulnerabilities to

that data. It actually supports a much more in-depth data risk management approach than Scuba.

The combined analysis of discovery classification vulnerability assessment, when you combine all that together, you have the foundation for any data security project, where Scuba just provides you with very limited information about vulnerabilities on a certain platform.

Brian Contos: Of course, Scuba is freely available for download from our site for anyone that's interested. Dana, you mentioned a little bit earlier DAS. DAS, of course, can be used as a standalone solution, but also with integration with other SecureSphere solutions. What are some of the value-adds of using it by itself, and what are some of the enhanced value-adds of using it with some of the additional SecureSphere solutions?

Dana Tamir: Deploying DAS on its own supports any discovery and assessment projects you have. It provides you with the needed visibility and awareness to do a few things. First of all, properly scope out any data security and compliance projects. By knowing where the data is, where PCI regulated data is, or where SOX regulated data is, you can then make sure that you scope those projects, that you are aware of those projects for those databases as part of the project. It helps you analyze risks to sensitive data, and helps you plan, prioritize any system remediation efforts, because now you know which systems host critical data, and which are vulnerable to attacks. Then you can put them at the top of the list for remediation.

When it's part of an enterprise data risk management initiative, SecureSphere provides more automation for implementation of a sound security and compliance program, and the specific controls that need to be put in place. For example, some of the things you can do with SecureSphere DAS, is automatically map, scan the environment, map out databases, and based on the data they contain, apply audit policies to those databases, and apply security policies to those databases. So automatically you're then including all those databases in your compliance and security efforts.

Another thing you can do is, because you know which objects contain sensitive data, you can now set very granular audit policies to monitor access specifically to that data, versus the entire database. That provides you with much better visibility, and much better ability to consume the results, and understand patterns in accessing that data.

Thirdly, you can manage database vulnerabilities by actively mitigating through virtual patching, which I mentioned before. So you can apply database firewalls to control access to sensitive data, and mitigate any vulnerabilities that may expose that data to a data breach.

Brian Contos: Most people understand how a network vulnerability scanner works. You have a system or group of systems and you perform a ping, ICMP, TCP, whatever, to see if the machine's up, checking the ports to see what's running, and then once it sort of fingerprints and knows what the system is, running a level of checks based on patch levels, audit config if you have credentials, and known vulnerabilities. Database is a little bit different. Could you explain a little bit of the technical side of exactly how DAS discovers systems, and then, in turn, how it walks through the information and classifies it, etc.

Dana Tamir: Absolutely. The scans are run from a central, pre-configured appliance. We provide DAS either as a virtual appliance, or on Imperva proprietary hardware. The appliance actually uses NMAP, which a lot of you are familiar with, to map out the databases and the servers out there. Now, the additional value we bring is, first of all, we keep

historical information, so we now know if a discovered server is new on the network, or has been seen before. Then, we use system credentials to run the assessments on the server, run classification on the server.

The classification process itself is the process where we sample the objects in the database, and we look for any specific patterns, any specific names in the object, in the columns that define the object, to see if they might point to the existence of sensitive data. And based on that, we define -- if the object is sensitive or not, we can then define the object policies appropriately.

Brian Contos: So, Dana, you mentioned the credential side of things, and actually being able to go through and step to the actual data on the database. How does that work, and how do you deal with false positives and false negatives when you're looking at the data? Obviously every company's different. One company might have credit cards. One company might have stuff called "Top Secret," or "Classified," or something else. How does that all work?

Dana Tamir: DAS uses a few different methods to classify the data on the database. What it does is, it connects database credentials, which are provided by the user, and it looks for any names based on a dictionary of names, and the object name, as well as the columns, to see if there's any suspicious word that may indicate that this object contains sensitive data. And then it samples the data itself using regular expressions and algorithms to verify that the data itself is indeed sensitive. So, that means that not every table would contain the words "credit card" would be necessarily sensitive. We actually sample the data. We're looking for credit card information, and we're validating that information as well.

So, for credit cards, specifically, if we find a 16-digit number, we don't automatically flag it as sensitive, but we actually use the Luhn algorithm, which is used by credit card companies to verify that these 16 digits actually do validate through the algorithm and confirm that is a credit number. Only then we will alert that this object is indeed sensitive.

We use a similar mechanism with social security numbers. We have the ability to detect personal identifiable data. We can detect bank account numbers. A lot of different data types including custom data types, which the customer can provide to the solution to scan.

Brian Contos: Yes, I think that's probably one of the most interesting pieces. If you consider some recent events. Just in October there was a gentleman at Ford Motor Company that stole some information and wanted to bring it to a competitor in China. I believe he was a product manager or something along those lines, and he accessed sensitive information.

It wasn't credit cards. It wasn't PII. It wasn't healthcare. But it had to do with R & D and manufacturing models for new vehicles and parts for vehicles. So certainly that would be very different than what a healthcare organization has.

I think one of the strongest pieces of DAS is the graphical user interface. What do you usually find is the learning curve for people to get comfortable with DAS and actually be able to use reports and get value out of it?

Dana Tamir: It's very quick. The UI is very intuitive. There is a specific tab for discovery, where you define the policies as well as the review the results. The discovery tab also contains the classification policies and results, so you pretty much define the network area which you want to run discovery on. You define the servers where you want to run classification. You define which data types you're looking for, and you get the results. There

are interactive views available for users to analyze the results. You can apply different filters to focus on specific servers or specific data types, and then in the data risk tab there is a very nice graphical UI dashboard type that allows you to drill down into different areas of the network or based on data types, which is pretty unique.

So you can say, "I'm interested in PCI data. Show me the risk to PCI data." And you look at all the servers that may contain PCI data, and you drill down into those servers, and you look at specific vulnerabilities that put that data at risk, and then you analyze these vulnerabilities.

You can decide to take mitigation actions. You actually have workflow integrated into the product so you can track mitigation efforts from the product itself. You can track the status of each vulnerability from the product itself. Very easy to use. Very nice.

Brian Contos: Last question. Who uses this? Is this DBA, a sys admin, auditors, security professionals? And do you need to be a DBA or database expert at some level to make use of this tool?

Dana Tamir: Excellent question. Though DBAs can absolutely use the tool, and they should be aware of vulnerabilities that put their databases at risk, you do not need to be a DBA to use it. It does not require any programming skills, any database administration skills, and the UI is very interactive. We actually have a lot of auditors who are interested in this, because it provides very easy-to-use reports. We have IT groups that are interested in this, because it provides visibility that allows them to manage and control those servers. So anyone from security to compliance, to risk, auditors, DBAs, IT managers -- it's a great tool for all of these users.

Brian Contos: Fantastic. So there's a lot of great information out there on DAS. Where can people go to get more details?

Dana Tamir: We have a website, <http://www.imperva.com>. We have a section under products that talks about discovery assessment server. The capabilities are also defined under solutions, where we talk about the specific use cases, the discovery classification use case versus vulnerability assessment versus data risk management. A lot of information is out there. We also have data sheets and white papers available in our resources area including webinars that we've run on this topic.

Brian Contos: Great. Well, thanks so much, Dana.

Dana Tamir: Thanks, Brian.

Brian Contos: If you would like to learn more about this subject and Imperva, visit <http://www.imperva.com>, check out our blog at <http://blog.imperva.com>, follow us on Twitter at <http://www.twitter.com/imperva>, 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