What Works in Situational Awareness and Visibility: Reducing Time to Detect and Enhancing Business Outcomes with Splunk

SPONSORED BY

splunk>

WhatWorks is a user-to-user program in which security managers who have implemented effective Internet security technologies tell why they deployed it, how it works, how it improves security, what problems they faced and what lessons they learned. Got a story of your own? A product you'd like to know about? Let us know.

www.sans.org/whatworks

ABOUT ILLUMINA

Illumina is a global leader in genomics – an industry at the intersection of biology and technology. At the most fundamental level, we enable our customers to read and understand genetic variations. We strive to make our solutions increasingly simple, more accessible, and always reliable. As a result, discoveries that were unimaginable even a few years ago are now becoming routine – and are making their way into patient treatment. In 2014 we were chosen by MIT Technology review as the smartest company in the world.

ABOUT THE USER

Ryan Niemes has been in the IT industry since 1998, where he started as a UNIX administrator for SkyTel managing OSF/1, Tru64, and Solaris. He recevied his CCIE in 2001 during the old two-day lab, and started focusing on security. His first security specific role was at Fifth Third Bank, where he focused on network-based intrusions. He then moved to Germany to support the US Army 5th Signal Command's European theater. He later worked for Cisco as a Network Consulting Engineer for the US Marine Corps. In 2009, he achieved the CCDE & CISSP certifications. He started with Illumina in 2010 with a focus on networking & information security, eventually dedicating himself to the security team in 2015. He currently manages a team of information security professionals, whose focus is on automation, incident response, and security architecture.

ABOUT THE INTERVIEWER

John Pescatore joined SANS as director of emerging security trends in January 2013 after more than 13 years as lead security analyst for Gartner, 11 years with GTE, and service with both the National Security Agency, where he designed secure voice systems, and the U.S. Secret Service, where he developed secure communications and voice systems "and the occasional ballistic armor installation." John has testified before Congress about cyber security, was named one of the 15 most-influential people in security in 2008 and remains an NSA-certified cryptologic engineer.

SUMMARY

Detecting malicious activity more quickly and more accurately is key to reducing business impact of cybersecurity threats. This requires both visibility into alerts and logs from servers, PCs and network devices but also analytic tools to enable cybersecurity analysts to prioritize response and mitigation actions. A common success factor of those organizations who are not in the news for yet another breach is investment in the people, processes and technology to decrease time to detect and time to mitigate without negative impact to business operations. In this interview, learn why Illumnia chose the Splunk Platform.

Q Tell us a little bit about your background and your role at Illumina.

A I started about seven years ago with Illumina doing basically network administration. I was sort of their first networking/information security hire. We were a fairly small company at the time. And then five years later as our company grew in the molecular sequencing and array technology space, Jason Callahan started as our CISO, and I jumped over to doing specifically information security for Illumina. I'm out of the operational networking management and into just doing information security, and that really runs the gamut of incidence response; product evaluations, forensics, working with customers to secure their applications, all those different things. And we've increased the team size significantly since then. Like I said, it was just me at the time. Jason has since hired five additional folks to our team to augment and do various tasks that you would do in information security. I'm a CCIE. I worked for Cisco for a couple of years, and so my bread and butter is really UNIX administration and routing and switching.

Q You mentioned growth at Illumina over the past couple years. Give us an idea of the sort of scale, number of employees and number of locations.

A We were at around 1,800 or 2,000 employees when I started in 2010, and now we're about 6,500 employees. At the time, we had probably six or seven locations, and now I think we're up to about 15 locations around the world. Our HQ is in Southern California in San Diego, but we have a major manufacturing site in Singapore and a research and development site up in the Bay Area as well as one over in the UK.

Illumina's Security Problemand Candidate Solutions

Q So, let's drill down on the process you went through in looking at solutions. What were the problems or what were the operational drivers that persuaded you that you needed to go out and find a solution?

A Well, we didn't have anything at the time, really. We had very limited visibility into our network traffic. We had some of the typical security devices, things like we had some Cisco firewalls, and we had antivirus, but we didn't really have a consolidated view to look at all those platforms. We also decided to start implementing next generation firewalls. We started to do more with our security platform. We wanted to get operational visibility to just see how things were working to make sure that we could prove or disprove that the firewall was or was not blocking something and then use it, as we'll come up with some platform that would allow us to do some hunting.

We didn't start off with Splunk as the answer. We considered outsourcing all the major tasks of information security. We met with a couple of different partners to see what their offering was. We already had Splunk. I purchased it back when I started to do basic network troubleshooting and just log consolidation for all the network devices. And then we ultimately decided that we wanted to build the knowledge in-house for all of our detective controls as well as, obviously, our preventive controls. But, we wanted to be able to build that ourselves in Splunk and not outsource it. We really liked that Splunk could ingest all the different data sources and present us with a consolidated view. We started a security program years ago and after researching available options, (we) selected Splunk as our SIEM, specifically Splunk Enterprise Security software. We looked at managed security providers but wanted to keep the knowledge in-house.

Implementing the Splunk Platform

Q Let's go back in time a little bit. When you first got there, you bought and deployed Splunk for looking at some of the data from the operational side of things. Did you go through an evaluation there or some sort of competitive bake-off or anything or you had used Spunk previously?

A No, we didn't really go through any sort of bake-off--we haven't since, either. I have used other types of tools in the past, things like ArcSight and just grepping through logs.

Q So, you're at your starting point. You're using it for some log consolidation and looking at some of the network operational data. Walk us through what you had to do to sort of expand the use to do more than that.

A It's definitely been a journey. The biggest thing for us was learning how Splunk operates on the information as you ingest it. We made some missteps at first where we would just ingest everything and then try to run queries and make sense of it all. It didn't really work out for us well, and we use a model now where it's very specific to use cases.

As an example, we'll pick something that we want to know about our environment when it comes to specifically a security use case if we're seeing lots of IDS alerts all the time, but we don't really know if they've been successful or if the host that the IDS load triggers on was vulnerable to that particular attack. Our use case there was we want to find out if it is vulnerable and whether it was successful. What Splunk allows us to do is to break that problem down and say, 'okay, what are the log types we need? We need IDS logs. Okay, those are already in there.' Well, IDS logs don't have anything to do with the actual CVEs that are exposed, so we needed to figure out how to ingest CVE information into Splunk and

then correlate those using what Spunk calls a look-up table. We now know that an IDS alert that comes through and then all the CVEs that are associated with that particular IDS alert. Not all IDS alerts have a CVE, obviously, but if it does, then we know that information. And then we wanted to find out whether the host was exposed to that particular CVE. So, we're running a security center. We had Nessus installed everywhere. And so, we're getting reporting on all of our hosts. The next step was to make sure that all of the Nessus CVE information was ingested into Splunk.

At the time, it wasn't an easy way to do it, so we sort of hacked around and got it working. Now Splunk has created a much cleaner interface for ingesting that data. We just haven't converted everything over to it. But, once we got that data in there, then the next step was to build another look-up basically to provide the full correlation. We extracted all the information about the IP addresses and the CVEs, and then we just run a scheduled search every 15 minutes or every half an hour that does a lookup between the Palo Alto Networks or our Suricata or BRO logs and the CVE information as well as the IP address and CVE information from our Nessus scan data, and then we get an actionable alert there that says we saw an IDS alert, the host was either vulnerable or not. If it is vulnerable, we get an alert that an analyst needs to look at.

Q How long have you been up and running in broad operational use?

A We've been on Splunk Enterprise for six years. Enterprise Security, about two years.

Illumina's Security Use Cases

Q Let's discuss some of your use cases. Tell me about the key use cases that you use Splunk for.

A There are many security related uses that we use Splunk and Splunk Enterprise Security for. I'll go over a few use cases involving reducing malware infection rates, correlating vulnerability data with attacks, using Palo Alto WildFire to detect new malware in our environment and operationalizing threat intelligence for our use cases.

I will start with a vulnerability management use case: *Detecting exploit attempts against vulnerable hosts*. I correlate Nessus scan information and attack information from Palo Alto WildFire to create a search that triggers when an exploit attempt has happened against a host that is vulnerable to that exploit. Then I render it within Splunk Enterprise Security.

Our process goes like this:

1. What is the question we're trying to answer
2. Do we have the information in Splunk necessary to answer the question?
    a. If not, how do we get the information?
3. Once we have the information we need, what's the algorithm?
    a. Break the algorithm up into small chunks of discrete logic

In this particular case, we need the signatures of the exploits, which is supplied by Palo Alto Networks, and whether a host is vulnerable to the exploit, which is supplied by Tenable Nessus. In the case of Palo Alto Networks, we also needed to determine how to pull CVE information from the threat signatures. Luckily, the Splunk App for Palo Alto Networks contains that information, we just needed to configure the application to pull that information into a lookup table within Splunk. Next, we determined the algorithm necessary, which is to match up a threat signature, CVE, IP address with a Nessus CVE and IP address. Normally, lookup tables match on a single field, we needed to match on multiple fields. The other item we needed to add to the Nessus vulnerability information was a new field we simply called "is exploited", which is set to TRUE always. That way, a new field called is_exploited will be "TRUE" if all fields match when doing the lookup.

Q What's next?

A Another use case was operationalizing threat intel. Illumina uses lots of threat intel feeds, and the challenge we had was to correlate across all our devices.Splunk Enterprise Security has a threat intelligence framework which automates the task of ingesting threat feeds such as Facebook ThreatExchange data (and others) and it automatically performs correlation searches across ingested data. Additionally, we get hash listing from Symantec Endpoint Protection tool via Splunk DB Connect Application.

Q Any other examples?

A An interesting use case is detecting Patient-0 infection. For this we use correlation searches within Splunk in order to detect Palo Alto WildFire Patient 0 attacks and then use Splunk to help with the automated response and remediate.

This search is actually quite easy with Splunk. We search Palo Alto Networks logs for anything with the wildfire log_subtype, and look for anything that is malicious or grayware and make an assessment. We haven't fully automated this process yet but we'll likely either send an automated email to the user or kick off a malwarebytes scan remotely.

Q Do you have any operations related use cases?

A Yes, a regulatory use case is anomaly detection in manufacturing. FDA regulations require login activity monitoring, yet I want to limit access to infrastructure that I manage from other teams. My solution is to provide manufacturing with access to a Splunk Cloud instance, use built-in searches within Splunk Enterprise Security which is configured with hybrid search, so manufacturing has access to their logs only while from a SecOps perspective I have the full view.

The manufacturing devices all are configured with a specific index, lookup tables matching their hostnames and locations, and log collection for security related events. The manufacturing technicians have access to the cloud search head and are able to look at the dashboards I've configured for them and receive email alerts for failed login activity. The great thing about the cloud infrastructure is it's completely managed so I don't have to worry about uptime or upgrades and the like.

Illumina's Data Sources

Q For data sources, you mentioned a lot of the security control side of things, IDS and suricata and BRO, the vulnerability scanning from Nessus, sort of firewall lot type information from Palo Alto. Are there other information sources? Are you getting CISLab or other feeds right from IT servers or network devices and the like?

A Yes. Those are some of the use cases that we've had. Given the first level triage stuff is taken care of by systems like Symantec, we have been getting Symantec logs in there, and we saw things like malware attacks being blocked, pieces of malware being allowed by Symantec in some cases. But, the really interesting stuff to me is processes running on host or people running suspicious commands, things like PowerShell or TCPdump or IPconfig or mapping drives and that sort of stuff. My use case then was let me get process command line logging in here. We figured out how to get workstations logging that information, and then we installed this log forwarder on all the workstations at Illumina. If you had an Illumina workstation or a Mac, you were now getting all that process command line execution stuff. And what we do now, it's sort of ad hoc. We do have some dashboards built out that are looking for what I call "suspicious commands." We just check a dashboard on a daily basis and look for people running these commands, and then take action based on that information. We are logging process command line. We're logging very specific use cases.

Again, we wanted to monitor failed log-ins. I figured out the Windows EventID for failed log-ins, and we're white listing all those events from all of the domain controllers and all of our workstations. Probably about 30 percent of our servers are indexing that now. We have a large AWS presence, and we're working on getting the--all the AWS logs in here. Right now, we have CloudTrail logs, which is basically all our infrastructure logging. Some of the environments have a lot of LINUX hosts, and they've deployed the universal forwarder in those environments that we get all the authentication events. And then what other sources of data? Obviously, the network events, so all the syslog from the Cisco routers and switches isn't interesting to us typically, but we're still leaving it in there for the operational folks who want to use it.

Lessons Learned and Analytics

Q Are there any lessons learned, things you would do differently based on what you know now or lessons learned you would pass on to people?

A The biggest lesson for our security posture or our security use case is to take that specific use case model and make sure that for us and for the way that I've directed the team is that we're not going to ingest something into Splunk unless you're going to do something with it, right? A lot of times, some folks might want to just ingest everything and then just see what happens. For me, it's a very specific – I ask, "okay, what is it that you want to find out, how do we get the data into Splunk, how do we get it all mapped appropriately," and then "is it working" and then move on to the next thing because it can be very overwhelming if you just throw everything at it and you try to make sense of all the information. You've got to start small and just incrementally and take those operational wins whenever you can.

Q Are there any metrics you collect to say "here's improvements on our security posture" we've been able to demonstrate or show to management to justify continued funding, etc.?

A We are tracking notable events in Splunk--an enterprise security term, but basically, any time we run a search, any time we hit a use case like ransomware or suspicious log-ins, and we're able to take an automated response or an analyst response, we're tracking those in the same manner, and we're able to report on those differently. If I just bring up my dashboard, it tells us how many notables we've got. As an example, for the last 24 hours, we had 146 notable events on the network. Of those, we've been able to take an automated response action on 100 of those. That means that we've got 46 that we look at manually. In the automated response,

we're using some more open source tools and these are things that we've had to build in Splunk. In particular, we have written – using the Splunk add-on builder application – an interface to Google Rapid Response. We use Google Rapid Response on all of our workstations for live forensics. And so, if I see any sort of IOC and it triggers a notable in Splunk, I try to figure out how to get an automated response on that device. And if it's a workstation, what it will do is automatically tick off a malwarebytes remediation scan using GRR. It just runs it automatically for us. Specifically, with GRR you have the capability of running a python script. We have written a python script that automatically runs malwarebytes on the host.

Obviously, you can't automate everything, but we're trying to get to the point where we do automate everything. That's another big part of our use case, which is, "okay, you want to search for something in Splunk, what are you going to do with it once you've got that information?" Let's say you detect ransomware. What are you going to do with it? Are you going to just leave it until the morning, or are you going to have Splunk take some automated action for you? You can post something to HipChat, or you can set a pager duty alert, or you can create a service now ticket, or you can set an email, or you can run malwarebytes or whatever it is that you might have in your environment.

How Splunk Supports Illumina

Q You mentioned the products you're using and the cloud services. What about support? Are you using support directly from Splunk?

A Yes, we do use support directly from Splunk. We don't have a very big team, and I don't want to focus on managing Splunk. I want to focus on using Splunk. Because they have the cloud service, that's what we've decided to start leveraging more—that's where our big investment is now. This past year, we bought into the cloud environment, and that's what I'm directing all of my users to, that's what I'm trying to make sure that all of our searches run in the cloud so that we don't have to manage the environment. And the support there is very good. Basically, they have a portal. If we have some app that we want to install, we can open the ticket and Splunk support takes care of it for us in a couple of days - depends on the priority or urgency, obviously. But, they're usually pretty responsive. One thing to bear in mind, though, with the cloud environment is you don't have direct control over all the different aspects of Splunk. Like I mentioned before, there are these configuration files that you have to modify, and for the most part, they're exposed through the UI, but there are some that are not. You'll have to work with Splunk Cloud support if you want to make any changes to those. These are very advanced use cases, though.

Q There's the time you spend using Splunk as an analyst or a security person. Do you have any estimate of how much time it takes to take care of Splunk - full time equivalence for admin and other things you have to do to keep everything up and running?

A An average person – we have the 200 gig ingest per day – I would say, at a normal company, you'd want to have like a half an FTE working on it. But, if we're going all in with the cloud, it would be less effort than that. And because it's me, it's not that bad, not that big of a deal.