#165 - Modernizing Our SOC Ingest (with JP Bourget)

[00:00:00]

[00:00:12] **G Mark Hardy:** Hello, and welcome to another episode of CISO Tradecraft, the podcast that provides you with the information, knowledge, and wisdom to be a more effective cybersecurity leader. My name is G Mark Hardy, I'm your host for today, and I've got a special guest who we've actually got a little bit of in common.

Go Bills! we're both from Buffalo, and I'll let JP tell you a little bit more about how we're even closer than that. but our episode today is on security data pipeline, or how do we modernize our SOC ingest. and so JP Bourget, welcome to the CISO Tradecraft podcast.

[00:00:45] **JP Bourget:** Thank you. Nice to be here.

[00:00:48] **G Mark Hardy:** as I mentioned, we had a little bit in common, obviously we're both from Buffalo. I went to Buffalo last week for mom's birthday and, so by the time this episode drops, hopefully Buffalo's already beat Kansas City, but we, we've had a pretty good [00:01:00] run, but Beyond that, tell us a little bit about yourself, where you went to school and cool stuff that you've done because you've got a really fascinating background.

[00:01:07] **JP Bourget:** Yeah. So it's funny. I usually start out with RIT, college, but in this case, we actually grew up in like literally the same neighborhood and, in Snyder, New York, which is a part of Amherst, outside of Buffalo. And we went to the same grammar school, although I think you were 20 years ahead of me.

And maybe 18, 17 years ahead of me. and I went to high school with your sister, at Amherst Central. And, yeah, which is just crazy to me. And, like I knew who you were way before you knew who I was. Cause you were doing, was a hacker jeopardy back in the day at DefCon, like my first

[00:01:42] **G Mark Hardy:** Hacker Jeopardy! Yep.

[00:01:47] **JP Bourget:** with Vital, Vanna, and all that stuff, and I'll never forget, I'll never forget those days, when I was 21, I moved out of Buffalo, I

went to Rochester, did some restaurant management crap, and, that led me to not wanting to work in restaurants for 30 [00:02:00] years, and, so I went back to school.

when I was 24, I went to a community college in Rochester called Monroe Community College, led me to RIT, which, led me to, a Bachelor's in IT and then a Master's in Cybersecurity. And up until this past weekend, I had never met anybody who had a cyber security degree prior to me in 2008 in a master's and I met somebody who had one from JMEU, James Madison University in 2002, which is like bonkers to me,

but

[00:02:31] **G Mark Hardy:** I used to teach there as adjunct faculty, and it was interesting when they stood that program up. I think it was 97 or 98. I remember, the gentleman who had done so, he said, G Mark, would you like to come help us, teach this program? I said, sure. He said, where's your PhD from? I said, I don't have a PhD.

He said, okay, where's your master's from? At the time, I didn't have one. I got a couple now. And he said, you could design this entire program, but I can't let you teach because of the way that the [00:03:00] rules are in the university setting. And so it wasn't until years later that I came back. It's adjunct faculty, but yeah, it was a, one of the very first programs that had cybersecurity in the title.

And what I used to tell the students is that you're going to be the only person sitting at the table. If you're even at the, the executive table is a C level with cybersecurity in your degree. So as a result, you need to know this stuff. So sorry for the inject. But yeah, they've been around for quite a while and did it, did quite

[00:03:25] **JP Bourget:** no, it was just, and I, it's, I just, I didn't know. So, I knew that Norwick University in Connecticut had a degree and, but anyway, so, fast forward a couple of years, I graduated in 08, I was an adjunct in RIT for a while. That in 2012, in 2011, Bruce, and Heidi and I had, basically had a really crazy miscommunication.

And for those who don't know, Bruce Potter and Heidi Power from SchmooCon, is the least of what they do. But, I had sent

Bruce and,

[00:03:58] **G Mark Hardy:** should we tighten up the overlap further? My [00:04:00] son used to babysit for

him.

[00:04:01] **JP Bourget:** oh, yeah, yep, that's all, yeah, and, I emailed Bruce one day, and I'm like, hey, we should do a bike ride at DefCon, and I get a reply, 20 minutes later, Dude, DEFCON's in a month. There's no way we can ride a bike to DEFCON. I'm like, no, I was just saying, let's do a bike ride there.

And so that turned into, let's, bike across the country, or let's bike to DEFCON, then we'll might as well finish the ride. And, in 20, in 2012, I went to my boss and I said, hey, I need next summer off. Another Christ the King grab. and, he's no, you can't have that summer off. So I'm like, all right, and I went to DEF CON, I met Mike Murray from MadSecurity who, rest his soul.

and he got me, into the consulting world. So I did a couple of try before you buy gigs and, I had been teaching and all that. one thing led to another and January 23rd, end of 2012, quit my job, [00:05:00] flew 108, 000 miles around the world, did the bike ride. and then in 2013, while I was, I was flying around the world, I guess un fixing send deployments, I don't know how else to put it.

And, and, I, observed that people, or teams, security teams, they have the people, the process. the skills, the tools, the technology to really like respond to alerts and handle incidents. we didn't know it at the time, but we started building what I would call an I, at the time it was an IRP platform or instant response platform. And, as I started to get clued into the market, there's a couple. A couple folks doing that, like there's the old school, JIRA's been around forever, but there's RequestTracker, RTIR, which is like their IR module. There was, Resilient. They were Code 3 systems at the time that became Resilient. And so, we started building this, case management for the SOC product. and so I, ended up [00:06:00] moving to DC, went through an accelerator called Mach 37.

Which was awesome.

[00:06:07] **G Mark Hardy:** what? When did you do that? 'cause I did mine in spring of 2014.

[00:06:10] **JP Bourget:** I did mine in the fall of 2014.

[00:06:13] **G Mark Hardy:** That's right. You're F 14. You got the cool name for

[00:06:15] **JP Bourget:** Yeah. Forgot that we both did that together. Yeah. that's

[00:06:19] **G Mark Hardy:** This is one of these. It's an Yeah.

[00:06:21] **JP Bourget:** Yeah. and that, that's where I met Gal, Who, did the prior data, pipeline, show with you.

And,

[00:06:30] **G Mark Hardy:** Yeah. So Little Inject right here. CISO Tradecraft Episode #118. Data engineering with Gal Shpantzer from February 27th, 2023. It's up on our website.

[00:06:38] **JP Bourget:** Yeah, and I would say Gal predates me in, data pipeline work. but we definitely, talk all the time. He has a lot of great thoughts on the problem space. but it wasn't really a problem space ten years ago. anyways, fast forward to 2020, we ended up, [00:07:00] I guess we'll stop in 2015 for one second.

Started my company with Syncurity, and, Syncurity. And we, ended up, Building that IRP product, we raised money, did Mach 3 7, and that, in 2017, I hired a growth CEO,

And

in 2015 And

16 was like when the SOAR market started to become a thing. So the SOAR stands for Security Orchestration Automation Response, and, really like Gartner would say that the four pillars of that are orchestration, automation, case management, and ticketing. And, I've spoken on other podcasts about like how interesting, like, how the SOAR, market has had a last mile problem for a while. And there's, in my opinion, MSSPs are some of the best beneficiaries of SOAR, but they really just need orchestration as a service. They don't actually need the full SOAR suite, because they already have ServiceNow or JIRA.

They're doing ticketing, they're doing threat intel.

[00:07:58] **G Mark Hardy:** So,

lemme throw a little quick time out on the field [00:08:00] right now. So let's throw a little bit background. For those who are not using SOAR, help 'em understand a little bit about what you know. Why would you have SOAR? What does it do for you? And under what circumstances would, if you didn't have it, you would say, guys, we need to go ahead and look at this technology.

[00:08:14] **JP Bourget:** Yeah, so, here's what I'll say is I believe That well, so SOAR stands for Security Orchestration Automation Response. And so, which started out as, hey, like the ticking systems that existed in 2010 in the 2010s wasn't competent for security use cases. and before we get into actual SOAR, I'm going to share what I think is the most.

Successful orchestration platform out there, which is HubSpot. So HubSpot took literally like a bunch of people who hate process. They don't want to enter stuff. They don't want to use the system and they made it so easy for sales guys and marketing guys. to essentially, automate their world and get the data that you needed and just make, really good things happen.

And, [00:09:00] they've solved the problem that many other industry, industries have really struggled to, to wrap their heads around it and do elegantly. so I look at them as, in 2014, they were way better than what Sora is today, in, in my opinion. And so, think of the same type of workflow automation in the security operations center.

our marquee use case, man, it was 10 years ago, was phishing triage. So, we would have a, you'd have this user submitted phishing email box. We would consume that email, we'd pick it apart, we'd, look for indicators, we'd look for malware, we'd sandbox it, and we, we'd provide a disposition, or the goal is to either convict or acquit an alert, and so we would provide a disposition on that, and then that turned into, what if it was convicted, and it is a true puzzle, then we'd send that into an incident workflow, and so that, it started out with that, and then all of a sudden you want to [00:10:00] start to automate all that stuff.

And when

[00:10:04] **G Mark Hardy:** It's automating your playbooks. You've got security

playbook, and if you had, you tear a cheat and do this, Now we're saying, Hey, let all this stuff take place. It's going to be faster. It's going to be automated, which means it's going to be done the same way every single time, which means you're not going to miss stuff.

You're not going to forget something on a checklist

[00:10:20] **JP Bourget:** And I'm guarded in saying what's the perfect use case, because what I've experienced when a team that's used to doing things manually tries to move to automation, there's two in between steps that nobody realizes. The first one is that you actually have to have well documented processes. And a lot of folks have pretty decent documented processes.

But once you have well documented processes, you then need to go re document them and figure out what are these micro decisions you're making along the way that the computer can't, the system can't infer. So once you figure that out and you have these atomic actions of everything that needs to happen, then you can actually start to turn that into automation.

And, [00:11:00] you want to,

in?

[00:11:02] **G Mark Hardy:** There's an important point you'd made there that SOAR is not just go ahead and, stick disk one into drive A, hit enter. And you've automated your entire response. It requires a lot of customization, a lot of understanding to be able to get this raw tool set there. it's a box of tools, but you've got to go ahead and put them in the right order for it to do what you want it to do.

And as you tune it and refine it and something changed or modified, you've got to tweak it. But it really puts your, SOC on steroids in terms of when something hits the fan, so to speak. you can go ahead and go a lot quicker and perhaps with even fewer people that happen to be there because it's 2 in the morning.

[00:11:39] **JP Bourget:** Yeah, and so, I describe that as like the last mile problem. So like a lot of sword vendors have, they have all these playbooks integrations and things. And, those are great. But as soon as your JIRA has a custom field in it might break it. So there's all these things that could possibly go wrong.

And as a [00:12:00] vendor, as a, as running a startup, that made customer acquisition costs expensive, or else there's lots of pro serve that have to go on in order to get those use cases deployed. And so what I'll say is that if you can get over those, if you can get past the last mile problem, if you have the money to invest in actually automating the right use cases, because a lot of folks want to automate use cases that might not have the ROI on them.

they're sexy use cases, but they only run five times a month. There might be highly visible, but, what's the point of spending 100, 000 to automate something we do 30 times a year versus something we do 3, 000 times a year? it's, more mundane. So, yeah, so anyways, in 2020, sold Syncurity, I haven't done, I've done a lot of SOAR advisory work, but I haven't, touched a lot of SOAR, over the past few years.

some of my team has done a lot of work with it, but I, started a company called BlueCycle and, we try to do four things. so we do a lot of SecOps maturity [00:13:00] and advisory, I'm sorry, SecOps maturity work. We go in and advise MSSPs and enterprises, and then we're trying to, really look at what they're doing today.

What, let's say the top 10 percent of the, The players in the, blue team world are doing, and it's slightly different for MSSPs and enterprises where what one's, one's becoming more economically motivated, motivated in the enterprise, but they've traditionally been risk motivated, where MSSP is motivated by economics and, number of alerts per day and margins and, different things, but, at the end of the day, they still care about a lot of the same things.

And yeah. we, go in and look at what folks are doing and we, have a thought model we put in front of them and we try to help them, help meet them where they are and identify some low hanging fruit or maybe, what can we do in year one? What, is your 1, 3, 5 year plan look like on, getting to, that next level of not just [00:14:00] efficiency, but things like, are you doing detection engineering in source control?

Is there a process? and one of the things I love to talk about is, the next thing we do is secure data pipeline modernization, and I'm going to characterize two parts of it. So first, the first off is go back to 15 years ago when, Splunk used to have these ads saying, hey, you should send everything to Splunk.

I'm going to, it's like the game of how much data can you get into Splunk or whatever your SIEM is and now we've gone to today where, injustice, utility, it's, it's a consumption model. And, so what happened after 15 years of doing that? a lot of enterprises have essentially made their system analysis in Splunk, also their system of record, and that's the most expensive way you could make a system of record considering the other technologies that are available today.

[00:14:54] **G Mark Hardy:** Now explain briefly system of record.

[00:14:56] **JP Bourget:** So if I think of, ways to store data, [00:15:00] any system that the data is hot, it's queryable instant, almost instantly, the cost to store and operate and query that data is higher. but then I have all these other things like Snowflake, S3, Blob Storage, basically Blob Storage, or Block Storage up in the cloud.

That's pennies on the dollar compared to the cost of throwing it in my system of analysis. When I think of in the SOC, or in, in, in DefenderLand, we generally need a set of stuff to, query and trigger alerts. and then we need stuff for, let's say, long term retention, like maybe it's compliance.

a lot of times it's compliance combined with the ability to go investigate things from, greater than a year ago. so one of the things we talk a lot about is, hey, if we are, defining security alert use cases, those inform one or more detection use cases, that informs our log ingest use cases, or like what we [00:16:00] need to actually ingest.

And so when we can start to put rigor around that, there's more to it, but we're able to actually be smart about what we're sending to like our most expensive ingest versus what we're sending to our, system of record, our long term storage that we can go get later when we need it.

so that's, a preview to some of the security data pipeline modernization work we do where, we use a product called Cribl most of the time, but we can do it in other ways. the way, the reason we like using Cribl is because if I, know what you like.

pipeline modernization in Splunk, and I'm doing transforms and reduction in Splunk or in, in the SIEM tool.

I need a much more expensive person after, to really maintain that, so I still need a senior Splunk guy to maintain that. Where if we can do that in a tool like Cribl, we can have a junior engineer, curating and maintaining that data pipeline. Because it's like a low code set of functions.

Where, if I'm doing it in a SIEM, I [00:17:00] still need that, person with 10 years experience who understands. Not just like the data, but like Linux and services and the network, all that kind of stuff.

[00:17:09] **G Mark Hardy:** it's not just Splunk saying hey I want to make more money off of my people or we're going to make it so complicated you have to have an engineer but it's the nature of the fact that this is an operating SIEM and

it's got all these moving parts. As compared to something that's in cold storage or close to it in a data lake where there's a lot less going on and therefore, as you said, you can use a junior person.

Is that kind of a fair analysis?

[00:17:30] **JP Bourget:** kindness. So, more of think of it this way. what I need to do ingest traditionally, I would have to configure my products to send to one or more places. Most products can only send to one place. if you look at the old, so yeah, Cisco Firewalls may only send to lots of places for a while, but if you go to a lot of SaaS apps, you can send to one destination.

today, I think you can send a lot more, but, it hasn't always been the case. the other problem is if I need to switch a destination, I need to go through [00:18:00] change control. I need to actually change configuration of a product. if you put, Cribl in the middle, essentially, once we're sending to Kribble, we can do whatever we want with that data, including multiplex it, without touching the IT infrastructure, without touching the security infrastructure.

you'll have one piece of, one event that comes in. It goes down like a table that will filter. Hey, if it's Cisco SA, sends this pipeline, which may change it from, Ceph to JSON or Syslog to JSON, and then send it to Splunk, or then send it to Sentinel. then it can go down again, and then, oh, we, this is, to send the, to the data lake, or to S3.

We're just going to send the raw event out. And and then, we want to do a SIEM bake off, and, try out another product. we can just add another route and then send it to that destination without having to modify all the other destinations. And what's really cool is not only can I filter and multiplex it, I can, [00:19:00] like I can do things like say, Hey, I need these Windows event IDs to go to the SIEM.

I want all my Windows Events IDs to go to S3. but I can also replay that back from S3, like in 18 months, like it just happened. And I can also reduce the size of the log. So if you look at a Windows event, every Windows event has like those two paragraphs at the bottom of the event log. we don't really need to ingest those into our SIEM, right?

in some cases, the design of, the legacy design of things has, created unneeded expense. But, my favorite example is, this comes up over and over again. It's, a Palo Alto log. Like a pan traffic event that has 35 fields, the stock price is like 13 of them, maybe 14 of them.

And so, we can, reduce the, size of the event by 50, 60 percent at ingest. and so look, I'm never going to say that could help you reduce your SIEM bill, but we can definitely help you plateau it and get more value for your current spike.

[00:19:58] **G Mark Hardy:** So there's actually an [00:20:00] efficiency that's gained from a couple of things. Number one is, of course, is reducing the amount of stuff you're going to store. And then secondly, where you're going to store it in terms of its availability. Then actually in a third way, the technical requirements for the person who's gonna be doing it, having to be this much seniority and this much experience versus somebody here could do an adequate job and then this part of their career pipeline where they get really good at that and then they can qualify for better stuff.

It sounds, sound like a win.

[00:20:26] **JP Bourget:** yeah, for sure. And, we're seeing this at some scale you wouldn't believe. we, a lot of folks will start out like, with 500 megs or 500 gigs or terabyte of daily ingest. But like we've seen, people doing Either the petabytes of ingest, which is just crazy.

so, one of the things that was really interesting to me that's come up through this whole data, security data pipeline thing is, with Cribl or not, or any other, like, you have vectors, Datadog, you actually you've heard of [00:21:00] Logstash, there's all these other tools that can get data out of systems too, but, but they don't work with everything.

And so, we've started to build a lot of, custom code, not custom code, but like really, We have a lot of, integration, like patterns to consume a lot of the major, like SaaS products, like things like Salesforce that are wonky to get data out. So I might need to go and say, Hey, what reports do you have for me?

And then it'll give me a list of links to CSV files that I have to go transport, right? things that like, like when you get to really high volume polling, like your octalogs. you might be getting so many events per second that you actually need to put some resiliency in there to make sure you're not missing those 10 events that was somebody, breaching 2FA or something.

so, we do a lot, my team and I do a lot of work around, essentially building the AWS or the Azure infrastructure [00:22:00] to, at MSSPs and enterprises to do ingested scale. And then a lot of times we'll just then send it to Cribl. But we're able to, add a lot of, precision and, guarantees that we're not dropping logs and we're doing it in a lossless fashion, when it comes to, pulling APIs.

and then the other thing we've been working on, which, I think is interesting is we've started to build a, like we're doing like integration factories for, vendors. as we've done so much of this integration work, we've shipped probably a thousand instances of integration in the past, two years. and so it turns out that a lot of startups, they don't, like they have product folks, so then they go find a product where it can fit, and then all of a sudden, then they hire sales folks. And they start to need, essentially like they have inbound requests for integrations. these product guys, they know the product, but they don't know the security data ecosystem.

They don't know all the [00:23:00] APIs. And I almost look at us as we're like a dev shop, but we're also security SMEs. Where, we know the data, we, we've made the mistakes, we've fixed them, we've made other mistakes, we've, gotten past those, and we know a lot of the, a lot of the main tools, like probably the top 200 tools we've worked with, many times, so we can help accelerate their go to market, really, actually, I say we can accelerate their time integration. so they can land customers and, we've both done startups and, one of the key things in early startups, key friction, how do I say this? A common friction point in startups, trajectories is their ability to integrate with third party products to add value to what they're trying to do.

[00:23:48] **G Mark Hardy:** So it's really not just a matter of grabbing an API and saying, Hey, we'll just plug it in there. And it's just, real simple red wire to red wire, blue wire to blue wire. because if you're building something as a small business owner, you [00:24:00] actually have to build your own device and you have to decide, Hey, are these inputs and outputs sufficient to do what you want to do?

If you want to do something beyond that and your API that's providing your inputs, doesn't give you what you need. How do you solve that problem? You said, Hey, I, for whatever reason, maybe I wanted something that's just not being captured in the logs or a prior decision was made. to say, Hey, we're going to create skinny logs, which is great.

But the problem with the skinny log is a little bit later. You say, Hey, I need something that's a little bit more robust and it's just not there. Sometimes it's if it's a lossy compression. Now, as you said, the Microsoft paragraph at the bottom, if it's the same paragraph for the same message up here, I only really need to take the code.

I don't need the whole paragraph explaining what that code is. So that's, loss less, right? I can compress that and I don't lose a single bite of knowledge

because I can rehydrate it, so to speak, and put it back in there. But it's an interesting challenge when you say, I've got financial or bandwidth or storage constraints.

I'm going to have to limit what I'm going to [00:25:00] store. And then, would I have to reinvestigate something or there's a long, slow attack that all of a sudden I realize, Wow, this bad actor might have been here for a while. any tricks that you found out to be able to come back and do that? Because I know there's some interesting stuff on Microsoft with respect to their telemetry.

And we're talking about that in our pre show, about, maybe Microsoft had held back a little bit because we found some of these federal government intrusions that took place and the answer seemed to be, Oh, if you were subscribing to our E5 and our P2 licenses, you would have had all that telemetry, but because you had a cheaper package, we, just didn't give it to you.

and, so there's, two questions there. One for the developer of a product, how do you get the information you don't have? And the second one, which I want to delve into a little bit, Is that what happens if your vendor basically says, Oh, we'll give you that, but it's going to cost you and that exceeds your budget.

Is it bad on the vendor for trying to upsell you for something that would be necessary? It's Oh, [00:26:00] you need an engine with that car after you've paid for it? Sorry about that. Is this something that it's a kind of a standard practice? what are your thoughts,

[00:26:07] **JP Bourget:** Yeah. so I heard, four things. I think the key thing I heard is would we think of a vendor or other, or just like an enterprise trying to defend what's going on.

we're always looking for the, to, find the most value out of the situation for them. if I'm a, if I'm a SOC, and, a lot of this is religion too, so I will a lot of times recommend, hey, you should keep these logs, and send them to your SIEM, because, there's only 500 a day, but they, provide very interesting context.

Where, in a lot of cases, I believe firewall logs can go right to S3, and you can really easily write code to go retrieve them with 32 times a day you need them. Now, if you look at a Splunk bill, or a SIEM bill, a third, it's [00:27:00] very common for a third of the ingest to just be like, network traffic, or firewall logs.

So what we'll do is we'll say, Hey, we're gonna take power logs, but we're gonna, we're only gonna take the ones that tell us that there's risk. Or, like, a, wildfire alert, or a, a, malicious URL or something. We're gonna send everything to the system, or, to the, data lake, and we'll, create when we need it.

Now, that's not okay for a lot of people. Not everybody's there yet. which is okay. but we, always try to think of what's the, best way to use the money we have, use the resources we have to get the best outcome for that particular team. so the next thing is, you pointed out like, hey, You were talking about hey, we might have some logs that, that because we're reducing them, we're losing some data or we're sending a crappy, ASCII key log format or something. like what's really interesting is that as APIs have matured over the past five years, and people have really interesting requirements, [00:28:00] in some cases, like it's a graph based API.

where you can pivot around and pull back what you need immediately. But, there's many times where you have to do two or three calls to the API to get what you need out of it. Now, what's happened in especially the VR space, is that's matured into a query language where I can say, give me everything that includes this, but filters that out, but does this too.

but, in a lot of cases, we still need to go like query a product, like for a device ID that has, this known. that has like a no invulnerability on it or no alert on it. But then we need to go get context on that device ID, and then we need to get context on the alert. And that might, sometimes that's one big fat JSON payload.

In other cases, we have to go do a bunch of queries to do that. So, knowing that, and, matching that with whoever's the requirements are, whether it's a vendor or SOC or whatever, that's where the rubber meets the road. And, I, anybody can figure this out. A lot of times it's about speed to outcome and we can help, [00:29:00] we generally are able to do that in a day or two, where a lot of times it'll take teams months.

Just need to get access to the products. So, that, that makes it interesting too.

[00:29:09] **G Mark Hardy:** talking really about security pipeline or data pipeline approach, which is basically what we're talking about. How do you set up this plumbing? How do we take all the possible event or log generating things, select from them with the knowledge of the enterprise, what it is going to contain information that we care about, Parse those out into some areas.

Some things that we say, like firewall logs, just throw them right out of the S3 bucket because if you need them every now and then, you can look at this up or look up that. You might have something that could be a lot more intensive in terms of having to take a careful look at it. And things like that.

And some stuff you might just want to throw into Glacier and say, yeah, if you ever have to thaw it out, we'll look at

[00:29:51] **JP Bourget:** well, look, there's a lot of nuances there's G Mark. So if I have Splunk Cloud, I get, [00:30:00] like the standard retention periods one year. a lot of folks will go and send to our, Deep's, Deep Glacier, or whatever it's called, Glacier Storage, right away, because they know that they're, why pay for it for a year if they can get it in Splunk, right?

Now, another nuance would be, hey, we're a small shop, and we don't have the resources to go pull stuff from S3 all the time. then you might send your firewall logs to your SIEM, because you just need them there, and because you're not a huge shop, it's not that cost prohibitive. But if you're a 30, 000 endpoint, organization all over the world, like your, network traffic is to be a lot bigger than a 300 seat company. So, there's a lot of nuance to it and and it all depends on the security team and, what, where, what they're comfortable with and the CISO and what they're comfortable with, and really the history of what's happened to that company. So, if they've had a breach and they didn't have their firewall logs and that would have told them who patient zero was, then I'm pretty sure they're probably going to invest in, [00:31:00] in having those firewall logs pretty handy.

If they've never had that particular experience before, or if they have a developer who can write the tool in Splunk or Sentinel to go pull those logs on demand, whenever there's a particular type of alert, then they're probably going to do, put it to the more cost effective store.

[00:31:17] **G Mark Hardy:** That makes good sense. So I think the lesson from here that I'm trying to distill out, it's not one size fits all.

[00:31:23] **JP Bourget:** It's, it's not

[00:31:23] **G Mark Hardy:** It's really independent, both in your size, your scope, what you're

doing,

[00:31:27] **JP Bourget:** your appetite for investment.

Yeah, and regulatory requirements. Yeah, and your appetite for investment across certain parts of your security organization.

[00:31:34] **G Mark Hardy:** And in your experience, have you found that organizations that choose to do a more robust logging and ability to go ahead and provide more rapid retrieval? Do they do that because it's good? for a security planning purpose or they do it because something went wrong and then typically you only see people responding when they realize oops we could have done this better we better do it better because someone's looking over our shoulder.

[00:31:59] **JP Bourget:** you know, all the [00:32:00] answers. It's all on the board. a lot of, a buddy of mine who's a CISO is I have never had more money to spend than after we got breached. a lot of times organizations, the CISO gets a blank check for the next year because of the, the shock and awe of the ransomware event or whatever it might be.

but, that short term blank check doesn't necessarily solve the systematic problems in the organization. And if So the way I look at the, I guess the way I look at this is that it really depends on what the, culture and the makeup of the board is. And are they deriving this? so is this driven by a culture or is it driven by compliance and regulations?

[00:32:47] **G Mark Hardy:** And so for a CISO who is looking at this saying listening to this episode I realize that I probably need to do something a little bit different than what I'm doing now but I need to build a business case for this because [00:33:00] in the whole idea about collecting logs and storing them is quite honestly, if nothing ever horrible happened, that's a diode.

All this logging information goes into something and it builds up and it builds up and it's like your mom's attic or the junk room in the basement or something like that. If you'd never have to go dig anything out of it, so what? Who cares? It's only when you need it. And then in my opinion, to create that business case, obviously a lot easier, as you said.

If something happens in the insurance industry, they say nothing helps sell insurance like the house across the street being on fire.

But the other thought is, that how would a CISO or a business leader in the cybersecurity realm be able to create that effective business case in the absence

of a nasty event so that they are better positioned to helpfully mitigate the impact of a future event such that they can respond faster and more accurately.

, so, I have two thoughts when you say that the [00:34:00] first one is that

[00:34:01] **JP Bourget:** You can't count what you don't measure. however that saying goes, right? So if I'm not collecting logs, I actually have no evidence that we're actually safe. Not only do I have no evidence to, what patient zero was, if we had an incident, I don't have any proof that we actually didn't have an incident. So, I think that. The other thing I think of is, So, Cribl defines, the term observability as the capability to ask a question of your logs that you don't know you need to answer. And so the idea being is that very frequently, and this isn't just in security logs, but this is in like lots of logs, like application logs, and traces and metrics. I don't know what I need to ask tomorrow, considering the cost of long term storage, data lake type, is so cheap, it's, not that hard to make the business case for making that spend, [00:35:00] so that I can go ask that question tomorrow.

[00:35:04] **G Mark Hardy:** Now, here's what I'm thinking. And again, you can correct me on this. If you take a look at the developments that we've seen, most notably in the last 12 months or so, with that, magic two letter word, AI, we're now wondering, how about from a solution set, instead of saying, hey, I can go pull all these things out of the basement and go rummage through the boxes of that, oh, found the, find the thing I need, using that as a analogy for a human going through or taking a look at it with a search tool.

Could we have AI constantly scouring this stuff, so not only are you able to say, Hey, go fetch, but as all this stuff gets poured into there, to be able to start to build the gestalt of what's going on, to say, hey, time out, human. this, and then this, You've sent out all those logs in here.

You never said they were special, but they're indicative of a breach. there's a lateral movement going on, and it's low and slow, and it's sneaky and quiet, and you missed it, but I caught it. Are you seeing that taking place in the industry or not? Right now, or if not, do we [00:36:00] want to go on Shark Tank?

[00:36:02] **JP Bourget:** no, it's too late to go on Shark Tank. I see, the industry's going in that direction. and, one of the things I'm doing a lot of research for, my clients is, how should we rethink this data lake architecture so that we can load all this data into vector databases? So that we can load it into, essentially, so we can store it in a way that these LLMs, or, these learning models can, go and take advantage of this data.

So I think the last thing to chat about, is the idea of a tax or like a premium or enterprise spend to get security functionality. And, and you pointed out the situation where the Chinese, I think it was, had, found a way to get some certain keys in the Azure or M365 ecosystem. And,

[00:36:54] **G Mark Hardy:** This federal, U. S. federal government was the client. and

[00:36:58] **JP Bourget:** yeah, [00:37:00] And what ended up happening was, if you were like an E5, which is like the top, Microsoft license spend you can make, you were able to get. Those, you had those logs, but otherwise, if you were like, it even just an E3 plus security or something, you didn't get those events. And people are super happy about that as in Not.

and what ended up happening is Microsoft made those events, available. I think the E3 level. so if you were making that, that investment

[00:37:33] **G Mark Hardy:** right. It went from E5 to E3. It was Microsoft Purview Audit that was really, you had to subscribe to pay for that. Plus, had a limit of 90 days. They've extended that to 180, but a lot of that was a do from a push from CISA. Okay, so CISA has done some great stuff and

I've actually got an invitation out.

I've been talking to Jen Easterly's office to see if in her crazy busy schedule I might be able to get her on our podcast to talk about a lot of the work that they're doing. But, they're doing a lot of amazing stuff with the Cybersecurity and [00:38:00] Infrastructure Security Agency, for those who didn't know what CISA stands for.

But anyway, yeah, back to this and about the tax on that. So you want to be careful that Information may be available, but a vendor might say, Yeah. but, how much you got. and so just something to be aware of.

[00:38:17] **JP Bourget:** yeah. Now that look like it's funny, I, have a small business. I don't have a hundred users in my environment and, but I'm a security guy. I want all the, I want all the bells and whistles, but like a lot of times I pay 99 a seat to get SSO or to get like the security logging API or whatever it might be.

And like I'm a pretty big believer that I don't think it needs to be at the entry free tier, it needs to be like the. The business or the pro tier, like that core security

functionality that helps protect the environment and look like it's good for not just the vendor or not just the user, but good for the vendor too, because if people are implementing SSO and [00:39:00] two factor and those types of things, like they're not going to be, it lessens the chance they're going to be part of a breach as a

[00:39:06] **G Mark Hardy:** under promise, over

[00:39:07] **JP Bourget:** vendor and get that

[00:39:07] **G Mark Hardy:** deliver, and your customers will love you.

[00:39:10] **JP Bourget:** Yeah. And so, there's this, site called SSO Tax, which, which is like just a funny, it's a shaming list of the vendors that charge crazy amounts of money to enable SSO. so yeah, like I get it. I don't expect everything to be free. People have to eat, but I don't expect it to be 99 a month to do it.

you can get for free on Gmail.

[00:39:35] **G Mark Hardy:** And that's a very good point, because I think what we're finding out is that multi factor authentication, single sign on, go a long, way toward keeping the bad guys out. Now, what it doesn't do, what it does is going to help against things like password spraying and things such as that. But it doesn't necessarily help with hacking the user.

So I worked a recent case where a user had, got a little fake email. It looked like it was from [00:40:00] DocuSign or it was from Microsoft saying, Hey, your password is going to expire in five days. Clicked on it, took them to some Microsoft lookalike site. Hey, enter your ID and your password. You need to log in.

Okay, they got that. And the next thing was, hey, there goes the MFA to their phone, as you would expect. And they keyed in the MFA. what's happening is the attacker is essentially a man in the middle. And you're like, thank you for coughing up your ID, your password, and your MFA creds. I'm in. And then, of course, the CISO, we, on the security side, I get an alert a few minutes later from Microsoft saying Danger, Will Robinson, there's somebody logging in for this really weird country that you can't pronounce, and so in the go and doing triage and things like that.

So what we're finding then is our technical Skills are technical defenses. Those levels are going up and the attackers aren't going to give up, get a haircut, go

commute into work and pay their taxes. They're going to just find a new attack vector and they're looking at the attack surfaces being, the wetware, the [00:41:00] carbon based part of the network instead of the silicon.

And a lot of what we're talking about here, that's out of scope, but it's just a reminder to folks that even if you had the perfect technology. They say nothing is foolproof because, fools are so ingenious and that happens. any last thoughts you have before we wrap up for today?

[00:41:18] **JP Bourget:** to wrap up what you're just saying just now is you're literally like selling me zero trust. Or the need for zero trust.

[00:41:24] **G Mark Hardy:** And Zero Trust, if you listen to the episode I just recorded this past week on the 7 Deadly Lies, the problem with Zero Trust is what? Is that if I'm in there and I convince somebody to let me do a quick assist session, because I'm masquerading as IT, I am now in on an approved device, with an approved user, an approved login, with all the rights and privileges, and Zero Trust just collapsed.

Because Zero Trust is really assigned trust. I've assigned trust to a device, to a user, to a particular session. But if somebody can get in there and actually piggyback in on the back of an existing [00:42:00] legitimately logged in user on a legitimate session, zero trust out the window.

[00:42:05] **JP Bourget:** Can I do that from another country?

[00:42:07] **G Mark Hardy:** You can do it for anywhere you want. You ever try using QuickAssist? Next time one of us is in Europe, we'll do it. I've done QuickAssist. I was over in Eastern Europe last month and had to do that with somebody and it worked just fine.

[00:42:19] **JP Bourget:** Yeah. so look, like I think my takeaway of that is like security is a balance between risk management and usability and every business, if I could leave CISOs listening to this with a takeaway is that understanding that. So if you're working at the NSA, I can almost guarantee you the usability of those networks is much more difficult than, the usability of Macy's.

Or at different places. And so it's always a balance between what the organization requires and what the users demand. But as a CISO, you have to tow that line and you have to sell security in the, in, in exchange for usability a [00:43:00] lot of times. so

[00:43:03] **G Mark Hardy:** And it is what it is.

[00:43:06] **JP Bourget:** what it is, but

[00:43:08] **G Mark Hardy:** JP, this has been great. great catching up with you this past weekend at ShmooCon. Glad we had a chance to do this. Again, if you, we were talking about all the back and forth. but, for our listeners at CISO Tradecraft, we create weekly podcasts to help our listeners increase their knowledge of cybersecurity.

Now, if you love learning from us, try hiring us for consulting services. CISO Tradecraft does have a variety of virtual CISO services that we can offer, such as Helping you create or update your cyber security strategy. Leading tabletops with executives to simulate how your organization would respond to a cyber attack.

Performing process improvement activities to fix processes that need improvement such as third party procurement, vulnerability management, or third party risk assessments. And additionally, we also perform custom coaching to help you along with your career to become a stronger CISO. If any of these items are interesting to you and you'd like some more information or some help on that, reach out to us.

Go to CISOTradecraft. com or connect to us [00:44:00] directly on LinkedIn. So thank you very much for listening in today. Our guest today has been JP Bourget with BlueCycle. net, and we've been talking about security data pipeline, modernizing our SOC ingest. And so this hopefully would be useful to you if you like CISO Tradecraft.

Follow us on LinkedIn. We got more than just podcasts. If you're watching us on YouTube, great. If not, go ahead and subscribe. I think you'll find that we can add a little bit of that, to your day as well. And lastly, this is your host G Mark Hardy. And, until next time, go bills. and stay safe out there.