EPISODE 789

[INTRODUCTION]

[00:00:00] JM: Malware is malicious software that makes money for the creator of that software. Malware can appear on to a user's computer if that user visits a malicious websites or installs malicious software by accident. There are many types of malware.

Spyware sits on your machine and logs your data in order to sell it. Ransomware can lock your computer and demand that you pay money to unlock it. Adware serves you random popup ads that you don't want to see.

Cryptojacking is a newer form of malware. Cryptojacking software uses your computer to mine Bitcoin and other cryptocurrencies. Cryptojacking can occur when you visit a website that is running JavaScript that executing along with the rest of the webpage. When you visit a website with a cryptojacker, your computer will become slower because your CPU is being taken over to mine cryptocurrency.

Cryptojacking can occur anywhere that code runs and there's a lot of code running on cloud providers. Cloud providers themselves are very secure, but a cloud provider cannot force its customers to be secure. Users who host an insecure application on a cloud provider make it infected with a cryptojacker. If I host a large complex website on a cloud provider and I'm serving millions of users, I'm already paying a lot in cloud costs. But when my application gets infected with a cryptojacker, my costs could shoot up, and if I don't know why my costs are increasing, I might leave the cloud provider.

Esteban Vargas joins the show today. He's the cofounder of SafeTalpa, a company that provides defense against cryptojackers for cloud providers. Esteban joins the show to explain how cryptojackers work and why cloud providers have trouble defending against them.

An update on FindCollabs, the new company that I'm starting, we increased the prizes for the FindCollabs Hackathon to a total of $5,000. If you're thinking about starting a business or you

have an app, if you have wireframes or designs for, or you're an artist, or you're a creative person of some kind who's looking for collaborators, check out FindCollabs.

The FindCollabs Hackathon has a first place price of $4,000 and a second place price of $1,000. There are also runner up prizes and the winners will appear on SE Daily to talk about their projects. The judges for the FindCollabs Hachathon are mostly investors. They are venture investors and seed investors. So if you have a cool business idea, this is a great place to get some feedback from investors, and potentially win some money to fund your interesting business or your artistic project.

The Hackathon will run until April 14th at midnight, so you have plenty of time to find collaborators and build something awesome. You can go to findCollabs.com/hackathon to find out more about the Hackathon, and if you check out FindCollabs, please send any feedback you have about FindCollabs to me, or you can post it in various places across the FindCollabs website. FindCollabs is a platform for online collaboration and it was built for people like you, the people who listen to Software Engineering Daily. So your feedback is by definition useful for me.

Also, one last announcement is that we will be having a Software Engineering Daily in-person meet up on April 3rd in San Francisco. Haseeb Qureshi and I will sit down for a conversation about subjects related to cryptocurrency, investing and philosophy. Haseeb has been a frequently requested guest to come back on the show. I'm looking forward to talking to him once again. In the details for that will be at softwareengineeringdaily.com/meetup, though they are not posted quite yet as of today. So you can check back on that link and see if they're posted yet, softwareengineeringdaily.com/meetup. Haseeb is also one of the judges for the FindCollabs Hackathon. So if you want some feedback on any cryptocurrency related ideas, FindCollabs is a great place to post it.

So with that, let's get on to today's episode.

[SPONSOR MESSAGE]

[00:04:23] JM: Failure is unpredictable. You don't know when your system will break, but you know it will happen. Gremlin prepares for these outages. Gremlin provides resilience as a

service using chaos engineering techniques pioneered at Netflix and Amazon. Prepare your team for disaster by proactively testing failure scenarios.

Max out CPU, black hole or slow down network traffic to a dependency. Terminate processes and hosts. Each of these shows how your system reacts allowing you to harden things before a production incident. Checkout Gremlin and get a free demo by going to gremlin.com/sedaily. That's gremlin.com/sedaily to get your free demo of how Gremlin can help you prepare with resilience as a service.

[INTERVIEW]

[00:05:22] JM: Esteban Vargas, you are the cofounder of SafeTalpa. Welcome to Software Engineering Daily.

[00:05:27] EV: Hi, Jeffrey. Thanks for having me. As I told you, I'm a fan of this podcast and it's an honor for me to be here.

[00:05:33] JM: Absolutely. Well, it's an honor to have you on. I want to start by talking about the subject of malware. How do you define that term, malware?

[00:05:43] EV: Okay. It's basically any kind of cyber threat that damages the operating system or some other client software such as a browser or even like some desktop application. So the most common term that we hear actually is virus, but virus is a subset of malware. Other types of malware could be worms, which are the one that propagates through a Wi-Fi network. Ransomware, which ransomware when a hacker encrypts some file that is valuable to you, and in order to decrypt it and to get back the file, you have to pay a ransom, like deposit some money on a Bitcoin wallet or something like that to be able to recover the file.

A new kind of malware is cryptojacking, the unconsented use of computer resources to mine cryptocurrency. So malware is all sets of the sum of all these virus, ransomware, etc.

[00:06:36] JM: It seems like modern malware is taking a different form than it was taking when I grew up. So when I was growing up, many people had malware on their consumer operating

systems and they would get it from clicking on links that they shouldn't click on or installing some kind of software that they got over email, some EXC software. These days it seems like there's less malware on consumer operating systems. Is that the case?

[00:07:04] EV: Actually, no. It's just changing. So of course, like the big companies manufacturing the systems that we use, like Google and the operating system makers such as Apple and Microsoft, they're doing a better job adding cyber security as a future towards your products, right?

For example, Trojans, they still exist, but they're not as common as they were 10, 15 years ago, right? Taking it from that perspective, you're right. But threats are evolving and cryptojacking is a clear case of that. With the popularity of cryptocurrency, hackers found a way to find a gap hole in how our current security systems are architectured and they're making a lot of money of it. Actually, it's more profitable for them, because for example let's say ransomware. If I get affected by ransomware, I have to do all the job of contacting the hacker, "Hey, what's your Bitcoin wallet? How can I deposit the ransom?"

Whereas cryptojacking, it just makes the money when people navigate on the web, right? So threats are evolving to bypass current cyber security host and to be every time more profitable. So they keep existing, it's just that they got to adapt to how the manufacturers adapt their new systems. But I do agree on one thing. The industry is more conscious about cyber security now and the big companies have like a real big interest now in building cyber security natively on their systems.

[00:08:41] JM: How has cryptocurrency affected the ecosystem of malware?

[00:08:47] EV: Well, cryptojacking is for end users and for cloud providers which I want to talk a lot about that, how it affects cloud providers. As for end users, it's currently the attack with the greatest trend. Basically, yeah, even if the price of Bitcoin and Monero is at a low, there's still a motivation for the cyber criminals to get easy money. So yeah, there's a big correlation between this hyper on cryptocurrency and malware attacks.

[00:09:21] JM: You've mentioned this term a couple of times, cryptojacking. What is cryptojacking?

[00:09:26] EV: Okay. So cryptojacking is the unconsented use of your computer or computer resources at a broader level to my cryptocurrency.

[00:09:34] JM: And who does cryptojacking affect?

[00:09:36] EV: Okay, yeah. So let me tell you a story. We started thinking that the one that suffer the most were like companies of any size, because cryptojackers tend to act as a worm, which means that for example I'm an employee at some agency and I do some creative work. So I got to download music for the creative work I'm doing to have some background sound for the video or something, and I go to an illegal download site and I download the song but that site is a contingent cryptojacker.

So the cryptojacker then will propagate throughout their whole Wi-Fi network I'm part of, which inside the whole office will be infected by these without even logging into the illegal download site that an end user logged into, right? The greatest cost here is that they hire IT support. They might also spend some money on equipment change [inaudible 00:10:36] — oh and higher electricity bills. When this happens to us, more business located in the U.S. because we also have to adjust, for example, the cost of electricity to different countries and different areas. When this happens to us, more business in the U.S., the financial impact is between 400 and $3,000 mixing IT support, equipment change and electricity bills. But then we realized that actually cloud providers are the ones that suffer in the order of magnitude of the millions of the dollars.

So let me explain, let's say I'm a cybercriminal and I deploy that cryptojacking site that contains a cryptojacking script with any cloud providers. There are hundreds of them now. Let's say the cloud provider has a really big and important customer. Let's say it's a bank, and I'm sharing infrastructure with the bank. So because my script will consume a lot of the cloud provider's computer resources, the bank will notice that there's a problem with their cloud infrastructure. What happens then is that the bank blacklists that cloud provider, and that is a huge revenue stream being blocked for the cloud provider.

Not only that, there are damages that could happen as well like for infrastructure damage, high electricity bills, but the main component of the financial burden generated for cloud providers is that they will get blacklisted. Yeah, that's basically an insight that we have in conversations with various cloud providers.

Because of that discovery, we actually decided to shift away from building a Chrome extension for end users and we're now building an API that cloud providers can integrate into their infrastructure, because they are the actor in this whole ecosystem that suffers the most when a cryptojacking script is deployed.

[00:12:32] JM: Let's explore some of the dimensions of cryptojacking and then we'll get into the meat of this problem as it applies to cloud providers. As you said, these are the primary victims. But the use case you described, like let's say an office worker. I work in an office of 100 people. Maybe it's like a marketing company and during my lunch break I'm like, "I really want to hear that Taylor Swift song and I want to download it, because for some reason I'm downloading music." I mean, some people do still download music. I remember this happening in high school, a relative of mine went to download music and they paid the price. Their Windows machine could not take that downloaded music file. Of course, exactly what I'm referring to is the fact that they got a virus in the process of "downloading" the music.

That still happens today. You search download Taylor Swift and it takes you to this page where it's like, "Okay. Download Shake It Off," and then it's like it's very hard to figure out where you're clicking. You're clicking on this big popup that's kind of like a download thing, or maybe it's like a torrent site and you're confused like, "Why am I on this site? I just want to download the actual MP3, not this torrent site."

In any case, there's this vast network of different companies that maliciously try to get people to download various types of malware by telling them, "Oh! You're downloading an MP3 file," and what you're describing is something that an office worker could easily fall victim to, is that they go to click this download on a song thing and then they accidentally install cryptojacking software and they install it on their work computer and then it manages to infect the entire Wi-Fi network. I guess everybody that's on the Wi-Fi network.

Can you explain to me how does that happen? So if I'm an IT worker in an office environment, I download this cryptojacking software. It's going to start using my CPU resources to mine cryptocurrency, which is annoying enough. But how does it manage to propagate it across the entire network?

[00:14:39] EV: Okay. So a really important there. There are two kinds of cryptojacking, in-browser cryptojacking and not in-browser cryptojacking. The first one is the one that is trending right now. What actually gets to your computer hijack to steal resources used for crypto mining is not some binary. It's not the actual song, the thing that is the cryptojacker. [inaudible 00:15:02]. It's client side JavaScript what executes that cryptojacking. Putting binaries into some file was the way to cryptojacking in the past, but it didn't emerge in popularity as much as in-browser cryptojacking.

For example, Pirate Bay and other sites have actually changed – I don't know if I should call this business, but let's just call it that way. They changed their business model from adware to cryptojacking, because it's more profitable for them basically. It just happens that the JavaScript that they injected into their client side code is what executes the cryptojacking operation. So it's not the binaries. It's not the file that is downloaded.

[SPONSOR MESSAGE]

[00:15:56] JM: The blockchain is a new computer science primitive. It allows us to build applications that we could not have built before, and we're in the early days of blockchain applications. It's a great time to get started. Blockstack is an open computing protocol for building applications where users truly own their data. They own their identity and even their content and connections. With a Blockstack ID, users can have a more transparent identity system rather than the modern internet identity systems that are closely tied to advertising.

At blockstack.org/sedaily, you can learn how to build decentralized applications easily. Blockstack is open source, it's free, and it's an application stack that won't serve you ads or demonetize online media personalities, or be subject to the whims of an individual CEO.

Developers who build on Blockstack can even get paid to build better applications using Blockstack via the app mining program. To find out about Blockstack including these programs, you can go to blockstack.org/sedaily. You can learn how to build decentralized applications that are private and secure and easy to build, thanks to Blockstack.

Cryptocurrency are a huge unexplored space. If you're a developer, there's no better time to get started. Just so you know, it's not easy to build decentralized applications today much like it was not easy to build internet applications in 1994, but we know that things get easier overtime and Blockstack is one of the easier ways to develop on the decentralized internet today.

So if you're getting started, it's a great place to go. Go to blockstack.org/sedaily and learn more about how to build decentralized applications.

[INTERVIEW CONTINUED]

[00:18:03] JM: So we're talking just about in-browser cryptojacking here. So how does it makes it way? So if I'm the user, I go to this sketchy site and my browser starts getting hijacked to mine cryptocurrency. How does it make it to other people's computers in the same network?

[00:18:21] EV: Well, first thing. Most of the cases are from sketchy sites, like illegal download sites. Yes, sketchy sites. But it's not always the case. High profile cases include Tesla and Starbucks. For example, what happened on Tesla is that some cyber criminals bypassed the Tesla's cloud security and injected the cryptojackers into their Kubernetes clusters knowledgeable that Tesla has a really recruit website. So whenever someone would log on to Tesla's site, they would mine cryptocurrency.

Yeah, basically like once bypassed that infrastructure, it's the best thing for them to get a profit. So it's not always on sketch sites. What happened on Starbucks, for example, is that someone injected the cryptojacker on – You know when to go a Starbucks and you can ask for your receipt number or something to be able to get into the Wi-Fi. Well, a hacker injected the cryptojacking script into that site. Whenever someone used a Starbucks Wi-Fi, they would mine cryptocurrency for the hacker. So, yeah, what I want to tell you, this study, it's not always sketchy sites, the cases of in-browser cryptojacking.

[00:19:34] JM: So the Tesla one, that's fascinating. I had heard about this Kubernetes cryptojacking situation at Tesla, but I didn't understand exactly the attack vector. So what you're describing is that Tesla had an unsecured Kubernetes cluster. The Hacker got into the Kubernetes cluster and let's say they found the server, like some Node.js server that's serving the frontend website, like when you go to tesla.com, and they altered the code for that, let's say, node backend service and they made the code send a cryptojacking script to the frontend user who goes to tesla.com. So that now when I go to Tesla.com, I'm mining Bitcoin for this random person.

[00:20:21] EV: Yeah, that's exactly the flow of all these. Yeah.

[00:20:25] JM: Okay. The Starbucks example is also interesting, but I think we should move on. I want to understand how this cryptojacking script actually works. So I go to tesla.com, I've got this malicious script that's executing in my browser. What's actually going on? How is that execution leading to putting money into – Eventually putting money into the wallet of some hacker?

[00:20:51] EV: Well, this is the deepest – I haven't produced software jacking script myself, but what I can tell you at the deepest level of understanding I have. So you import a JavaScript library called Coinhive for example. It's fairly available on these sites that, by the way, contain cryptojacking scripts. All they do is import Coinhive [inaudible 00:21:12] equals a new miner, then you put in your API key and Coinhive does all on the background.

[00:21:20] JM: Coinhive is the open source cryptojacker, right?

[00:21:23] EV: Yeah, and it's the most popular one right now. Yeah, that's right. But now, what's behind the Coinhive code, like I don't have an answer to that to be honest. Yeah, that's basically what's happening. But I can tell you how we are detecting it and maybe that can also give an insight to you.

We inspired ourselves on a paper by some cyber security PhD that explains how analyzing certain variables were relevant to in-browser cryptojacking, which is actually what we're

providing cloud providers. Yeah, as I told you, the most basic cryptojacking mechanism is importing Coinhive and [inaudible 00:22:05] equals new miner, blah-blah-blah, and then you deploy that and you have a cryptojacker. It's something that you do not even need to have technical knowledge about.

Well, a prevention mechanism could be just analyze every JavaScript file or HTML that has a script tag inside it and detect if it has the word coin have in it somewhere, or a Coinhive import might be better, right? That would be a really basic detection mechanism, but we understand that cryptojackers are going to evolve and as all cyber threats are going to evolve and ones are going to emerge. So we cannot depend on such basic defense.

So what we're doing is basing ourselves on the paper, which basically analyzes three things. First, network packets. So the cyber security researchers found that on a network package you get certain variables, like home address, foreign address. You could still do like a really hard coded solution just blacklisting certain IP addresses, but there are other things such as the size of the network packet, the network protocol of that packet and combining all these variables that you get when you analyze a network packet. You can actually build unsupervised machine learning algorithm that clusters the packet into [inaudible 00:23:30] into three clusters, which basically are benign packages, malicious packets that are not cryptojacking, but we're not going to cover them on this paper, and cryptojacking packets. Yeah, analyzing network packets, it's one thing that tells you if a cryptojacking script is taking place or not.

Second is CPU power. So as I told you, a cryptojacking script steals CPU power, and if you analyze the peaks, you basically do time series analysis to your CPUs usage graph. You can also get a lot of insight if a cryptojacking attack is taking place or not. Finally, it's analyzing the source code of the JavaScript code.

As I told you, a really simple analysis of what a code would be, check if Coinhive is being ported or not, but there are more sophisticated ways to detect that for every single case. For example, they analyze the software complexity measures. So one of them is source lines of code, but the more complex ones. [inaudible 00:24:34] complexity, which is basically like which function calls another function. Basically understand how that code is threated and how it's so interconnected.

Like get an insight of how complex the code is. Those code complexity measures also give an insight to see if a cryptojacking attack is taking place or not.

[00:24:57] JM: So if I understand correctly, there's this open source library, Coinhive, and anybody could use it to install a cryptojacker. But if you just install it in its normal code form, then it's going to be very easy to detect. So people probably use some kind of obfuscation or minification to – Like they transpile the JavaScript to some kind of minified obfuscated cryptojacker code. Is that right?

[00:25:27] EV: We think that cyber criminals are going to do that once solutions like ours get more popular. But right now you can actually just search for a cryptojacking tutorial on YouTube, and a tutorial uploaded like in the last few months, not in January 2018, like uploaded in the recent months. It will just tell you to do it without minification, without obfuscation. Yeah, right now it's actually pretty explicit.

[00:25:54] JM: Well, sure. I mean, you might be catching people that are doing it pretty explicitly, but the fact that you're not necessarily seeing people make YouTube videos about how to obfuscate your JavaScript doesn't mean that people aren't doing it.

[00:26:07] EV: Oh, yeah! More organized cyber criminals will obfuscate their code of course. But what I want to tell you is that people, or less sophisticated cyber criminals shall I say, are deploying these cryptojacking scripts really explicitly and it's happening right now this way explicitly.

[00:26:24] JM: Okay. Let's get into this. So you've said that the cloud providers are big victims of this. How do the cloud providers suffer from cryptojacking?

[00:26:34] EV: It's basically like I told you before. Their cloud infrastructure will be blacklisted by their big customers, like banks and all those. Banks, devops team or – Yes, someone inside the bank, or their automated cloud, health, check systems detect that that cloud provider is having some issues with the user of their computing resources.

Once that thing is detected by the bank, the bank will blacklist the cloud provider. So the cloud provider which makes money by a number of calls and – Yeah, basically by variable usage, will have that revenue stream blocked because of such blacklisting. That's like the main thing. Stop proceeding a huge amount of revenue, and there are like more indirect things such as damage to infrastructure, the fact that they're consuming a lot more. So if they're a level one provider, that will mean a high electricity bill, and if they are a provider above level one, that will basically mean a higher check from the cloud provider that is providing cloud to the cloud provider, because there are cloud providers who hire other cloud providers and are basically innovating in use edge or something of that. So cloud providers also provide to other cloud providers. Yeah, if it's level one, it means a high electricity bill. If it's above level one, it means a higher cloud bill.

[00:28:06] JM: So level one is like AWS, and then level two is like a ZEIT or a Spotinst.

[00:28:12] EV: Exactly. That's absolutely right. Yeah.

[00:28:15] JM: Okay. Well, let's talk through level one first, because I want to understand how does AWS end up with cryptojacking scripts on their servers. I mean, if I'm operating like a Tesla, they were probably AWS servers that the Tesla Kubernetes cluster was running on. But anyway, I should just ask that. If I'm an engineering at AWS, I discover a cryptojacking script running on one of my servers. How does cryptojacking software manage to get on to an AWS server and how does that differ from the path of me going to a malicious site and the cryptojacking script just executing in my browser?

[00:28:56] EV: There are two paths. First, a cybercriminal bypassing site with good intensions such as Tesla and then injecting the script, but the other path is me just building a site and deploying it via AWS. There are no barriers to that, and that's basically the path. The path for the end user, yeah, there's basically logged in to one of those sites that have such scripts injected on to them.

[00:29:22] JM: So how do you know that this is a prevalent problem at the cloud providers?

[00:29:27] EV: Yeah. So one of those cloud providers explained this to us with a real good analogy. Remember when your mom in 2000s or even in the 90s browsed over on the web and she got all these weird toolbars on her browser –

[00:29:44] JM: Okay. Just to be clear, my dad was actually the one who downloaded that song earlier. So I don't want to make this a gendered-malware discussion. But your point is well taken.

[00:29:56] EV: Yeah, [inaudible 00:29:56].

[00:29:58] JM: Please continue.

[00:29:59] EV: Yeah. So tool bar with all those weird icons, that person – Yeah, the browser will be slower, the computer will be slower, but that person can still navigate on the web despite having those weird toolbars, Internet Explorer, or Google Chrome, or whatever.

But the cloud provider mainly because of the financial reasons has to do an enormous effort to remove the equivalent of those tool bars. En end user can survive with the weird icons on their toolbar. The cloud provider cannot. Yeah, the cloud provider has to do whatever it takes to remove the equivalent in the analogy of such weird toolbars.

[00:30:44] JM: So help me understand why this is an economic problem for the cloud provider, because let's say I'm running jeffswebsite.com and my infrastructure is running on AWS, I've got backend serving infrastructure on AWS, and there's a vulnerability in my site and somebody installs a cryptojacker on jeffswebsite.com, on the infrastructure. Isn't that cost just going to go to me, not AWS?

[00:31:11] EV: No. No, because cloud computing basically means renting a server. Like in really simplified terms, that's what it means, and that's the whole point of cloud computing. That you don't have to buy a physical server just to host jeff.com. That is really costly. Instead, you rent AWS, a portion of a server that they have, that they physically own, and you share that infrastructure with other AWS customers. But the big customer will have more advanced

monitoring mechanisms to detect if the infrastructure that's providing cloud to them is having issues. That's really the thing comes.

Of course, jeff.com, you as Jeff the owner of jeff.com, yes, you will pay some bill, but you're getting a profit. So it's not actually a problem for you. It's a problem for AWS because of the detection mechanisms that really big customers have.

[00:32:13] JM: I'm still having trouble understanding this. So if I'm operating jeffswebsite.com, and what I'm saying is a malicious attacker manages to install cryptojacking software on the infrastructure that I am renting from AWS, how does that negatively impact AWS? I would imagine that I would be the one who would be billed.

[00:32:34] EV: Yeah, yeah. You'll get billed higher, but you're making money as well. It ends up being a profit for you.

[00:32:41] JM: How am I making money?

[00:32:42] EV: Yeah, yeah, because – Not in dollars, but in Bitcoin or Monero, yeah.

[00:32:45] JM: No. I'm saying there's somebody, the malicious, who has installed cryptojacking software on jeffswebsite.com. I thought that was the attack vector.

[00:32:53] EV: Oh, okay. So you're taking it from the perspective of what Tesla suffered basically.

[00:32:58] JM: Yes.

[00:33:01] EV: Okay, yeah. Yeah, in that case, the owner of jeff.com would also get a higher bill, but it's just a higher bill. The damage isn't as big as carrying one of your important customers. That's why the numbers are bigger for the cloud provider than for jeff.com.

[00:33:23] JM: So are you saying that AWS's own core service infrastructure or a cloud provider like AWS could potentially be hacked and have cryptojacking software installed across the entire cloud provider infrastructure?

[00:33:40] EV: Well, I don't know exactly how AWS's architecture on the inside, but it's happening. People are deploying – I don't know if – Deploy single [inaudible 00:33:52] deploy affects AWS's whole infrastructure. It probably doesn't, because they have reasons and all that and maybe other mechanisms to mitigate, like issues such as these and other issues, but to affect at least a part of their infrastructure.

[00:34:09] JM: What I'm trying to understand is how a cryptojacking script gets on to cloud provider infrastructure, because cloud providers are typically – I mean, at least the big ones are very, very secure.

[00:34:23] EV: Yeah, but as a person, if you open up an AWS account and deploy whatever I want. The problem is that, yeah, they do have a lot of monitoring mechanisms, but not one particularly focused on preventing that a cryptojacking script is deployed. So that's the thing. Anyone is free to deploy a cryptojacking script at any time.

[00:34:45] JM: So you're saying if I launch an AWS EC2 instance and then I start mining cryptocurrency on it, that's going to cost AWS money? I don't understand how that works. I mean, aren't I just paying the server costs to run that server and so AWS is charging me for running that software and they still make a profit?

[00:35:07] EV: Yeah. If jeff.com was AWS's only customer in the world, it wouldn't cost money to AWS, or yeah, [inaudible 00:35:16], but not in that order of magnitude. If you were the only customer AWS had, but AWS has customers of all sizes including the really, really big ones with a big box. One of those customers with the big box cuts their revenue stream they give to AWS that AWS has problems, because they stopped perceiving that huge amount of revenue.

[00:35:43] JM: I'm still having trouble understanding it, because if I – So let's say I'm a user, I spin up an EC2 instance and I just start mining Bitcoin with that instance. Well, first of all, is there anything problematic about that?

[00:35:58] EV: Technologically, no. Legally, yes, or at least in some places, but technologically there's no barrier to do that.

[00:36:05] JM: Okay. So what's the difference between me spinning up a cryptocurrency mining node, like I just said, versus committing cryptojacking. How does cryptojacking manifest there? If I set up a node and I just start mining cryptocurrency, that's one thing. That's not going to harm AWS.

[00:36:25] EV: Okay. Yeah.

[00:36:27] JM: I'm just trying to understand the difference. What exactly is causing this noisy neighbor problem that you're describing?

[00:36:32] EV: Yeah, there's people using – Mining cryptocurrency, they're absolute consent. So your question is what the difference between that and cryptojacking, which is unconsented.

[00:36:44] JM: Yeah.

[00:36:45] EV: To be honest, I owe that answer to you. I never thought about that question before for cloud providers, and now we have to investigate that. But yeah, I don't have an answer to that at the moment to be honest.

[00:36:58] JM: Let's say there's a customer who operates who a bank and you're saying that the bank gets infected with cryptojacking software on AWS and that they don't realize that. So that's why they take their business off of AWS.

[00:37:12] EV: I mean, not forever, but temporarily, yeah.

[00:37:15] JM: Okay. So really the problem here is that companies get infected with cryptojacking software and they don't realize it?

[00:37:23] EV: That's right.

[00:37:24] JM: Okay. So I'm a bank, I'm running my servers on a giant cloud provider and somehow my infrastructure gets infected with cryptojacking software. I'm trying to figure out what's going on and I decided to take my software off of AWS or whatever giant cloud provider I'm hosted on because I just don't know what's going on. I don't know why it's getting so expensive to host my infrastructure. In fact it's because there's cryptojacking software installed on my software.

[00:37:55] EV: Yeah. I mean, not only your front end necessarily, but somewhere in your infrastructure, yeah.

[00:38:01] JM: So this is actually happening where people are getting infected with cryptojacking and they end up with really, really high cloud bills. So they have to migrate their infrastructure off of the cloud because it's just so expensive.

[00:38:14] EV: Yes, or not off the cloud, but to some other cloud.

[00:38:18] JM: Wow!

[00:38:20] EV: It's a better way of putting it.

[00:38:22] JM: When they do that, do they manage to get rid of the cryptojacking software?

[00:38:28] EV: That's really is a case by case thing. There are companies that react faster than others. Yeah, we cannot generalize for all.

[00:38:40] JM: How severe is this? How many companies are getting infected with cryptojacking software and running up gigantic cloud bills?

[00:38:47] EV: Well, I don't know that number, but I can give you the following number; 33,000 sites reported in 2018 known to contain cryptojacking scripts worldwide reported. There are tons that are not reported because they use better obfuscation mechanism or whatever, or because they were just simply not discovered, but 33,000 reported. Combined, they sum up over 1 billion

users, those 33,000 sites. The number of sites, that's where 33,000 was reported to be growing at an 18% monthly growth rate.

[00:39:23] JM: Of course, here we're talking about frontend sites, right? We're not necessarily talking about the backend serving layer.

[00:39:29] EV: Yeah, we're talking about frontend. Yeah.

[00:39:31] JM: Okay, and we don't have any idea how many of those are hosted on cloud providers, I guess. So I think this is a good place to introduce what you're working on, which is pretty interesting. So your company, SafeTalpa, is an API for understanding if a site is infected with a cryptojacker. Explain what SafeTalpa does.

[00:39:51] EV: Well, that's exactly an API that provides security for cloud providers initially focused on this modern threat called cryptojacking. Yeah, we're basically an API that you can integrate your infrastructure with just a few lines of code and we do all the detection behind the scenes. They're really complex. The detection operation, we do it behind the scenes.

[SPONSOR MESSAGE]

[00:40:23] JM: I've been going to O'Reilly Software Conferences for the last four years ever since I started Software Engineering Daily. O'Reilly conferences are always a great way to learn new technologies, network with people, you get to eat some food, and the 2019 O'Reilly Software Architecture Conferences are for anyone interested in designing an engineering software more intelligently.

Software architecture is coming to Silicon Valley for the first time June 10th through 13th in San Jose. You can get a 20% discount on your ticket to Software Architecture Conference by going to softwarearchitecturecon.com/sedaily and entering discount code SE20.

Software architecture is a great place to learn about microservices, domain-driven design, software frameworks, management, many other topics, and there are lots of great networking

opportunities to get better at your current job, or you could find a new job altogether. Software architecture is also coming to Berlin, November 4th through 7th.

I've met great people at every O'Reilly conference I've gone to. So whether you're thinking about going to San Jose or going to Berlin for a software architecture conference, it's a great place to meet people who love software. Go to softwarearchitecturecon.com/sedaily and find out more about the software architecture conference. It's highly educational and your company will probably pay for it, but if they don't, you can get 20% off by going to softwarearchitecturecon.com/sedaily and use promo code SE20.

Thanks to O'Reilly for supporting us since the beginning with passes to your conferences early on when we had barely any listeners, but O'Reilly still gave us free passes and it was really kind. So thank you to O'Reilly, and if you are interested in going to software architecture conference, go to softwarearchitecturecon.com/sedaily and use promo code SE20.

[INTERVIEW CONTINUED]

[00:42:43] JM: So describe the usage for a typical customer. So somebody is purchasing SafeTalpa. They're using your API. What are they doing?

[00:42:49] EV: Basically what they do when they purchase [inaudible 00:42:52] or Stripe, which is integrate an API to their codebase.

[00:42:57] JM: What is the path of that API request?

[00:43:01] EV: For example, our first [inaudible 00:43:03], for example for this API, was analyzing URLs. So you have a function that receives one parameter that is a URL, and that URL returns a yes or a no like benign of malign, and that's just with a single line of code. What we're doing behind the scenes for that verification is that we're querying the least of reported URLs to see that URL is part of that list or not. Yeah, that's like our first [inaudible 00:43:35] that's really basic, but really useful.

What the next feature that we're working on is analyzing JavaScript files. So analyzing the code complexity of the JavaScript file that it passes a parameter we do the [inaudible 00:43:48] have that complexity and analysis and of the code complexity measures on the background. All you have to do is write that single line of code, like analyze code and pass the JavaScript file as a parameter and then we'll do the analysis and tell you yes or no, it's good or bad.

[00:44:05] JM: And you're selling this product to cloud providers or to companies that run on cloud providers?

[00:44:11] EV: Okay. So we haven't released publicly yet. So it's more like a preselling and it's for cloud providers. Yeah, we've been asked that a lot and it's basically because, as I told you, the cloud providers are the ones that suffer the most in this whole ecosystem and story of cryptojacking deploy. So we think that we're better serving – Our customers would be served if that customer is the cloud provider, not any other [inaudible 00:44:41], or at least in 2019 of course. The future can change. But in 2019, that's the answer.

[00:44:47] JM: Okay. So I think I understand. So if I'm operating a cloud provider, I would want to be able to make API calls to the production sites that are running on my cloud provider, because I want to be able to detect if cryptojackers are running on these companies domains, because if a cryptojacker is running there, I want to be able to send them an email and say, "Hey! You've got a problem with your infrastructure. You've got a cryptojacker running on it. You should probably figure out what's going on."

[00:45:16] EV: Yeah, that's right. The whole point of making an API is that. Maybe you want to send an email saying, "Sir, please correct this, or you might just want to take it down immediately without asking." That's something that you're able to customize, and that's the whole point of building an API.

[00:45:33] JM: This is a brilliant business.

[00:45:34] EV: Yeah. It's really hard, but yet. We're starting this really long journey.

[00:45:40] JM: Just to clarify, what you said about the first layer cloud providers versus layer cloud providers, now I'm thinking about like – So we just had a show today with Netlify. Netlify is an example of a second layer cloud provider. They're built on AWS and GCP. They host a bunch of different websites. I think a lot of the people who deploy websites to Netlify are – It's got a real big appeal to newer users, newer developers. Newer developers, the thing is they might be consuming really random GitHub repositories. If they're consuming these random GitHub repositories, one of the vulnerabilities that I read about that you had written is that let's say you've got an open source GitHub repository. It's for some like really minor tool, like some kind of stringify kind of tool and it's just your open source repository. Let's say you host this open source repository. Maybe there's been 10 contributors to it and somebody you don't recognize makes a poll request to it, and like they fix a minor bug and you're like, "Wow! You fixed a minor bug. You also committed like this strange piece of code that I don't really understand. I don't really know why that is in the code, but man! Because you fixed this bug, I'm just going to accept your poll request and forget about it.

Then several months later you can imagine somebody who is a new user, who's a new coder, they're looking for a solution to some kind of stringify problem and they find your open source repository or your NPM package and they're like, "Oh! Yeah, I'll install this NPM package. It solves my problem," and this person then deploys to Netlify. Then all of a sudden it turns out that that line of code that you didn't understand was a cryptojacker, and now this user has installed a cryptojacker on their Netlify website and that's either going to cause sad times for this new developer, or it's going to sad times for their users or it's going to cause side times for Netlify, or it's going to cost sad times for the infrastructure that Netlify if built upon. If there was a SafeTalpa API request to this website, then hopefully it would at least identify that, "Hey, there is a vulnerability that is mining cryptocurrency on your website.

[00:47:55] EV: Yeah, that's exactly our vision. You just reminded me of a story I read once on Hacker News about some open source project that was supposedly like to change the colors of your internal or something like that, but in the background was an evil AI terminator style that would take over the world and steal everybody's passwords or something like that. It's not a science fiction at it sounds. This might actually happen.

As you said, someone might commit a change to some open source GitHub repository with some weird line of code that is actually a makeup for some malicious script cryptojacking or of some other kind. So yeah, that's our vision, to prevent that happening.

[00:48:40] JM: Do you know, doesn't this already happened with like ad tech, where people you end up getting a flashlight app or you end up getting some NPM package that like is viewing ads in the background and it's – So it's generating ad revenue.

[00:48:58] EV: I haven't heard about that. I haven't heard about it, but it sounds really interesting.

[00:49:02] JM: Okay. All right. Yeah, we've done some shows about advertising fraud. You may find those interesting. So we're nearing the end of our time. This is a really interesting technology. I want to talk a little bit about Latin America, because I've read your Twitter a little bit. How does the startup ecosystem in Latin America compared to that of the U.S.?

[00:49:20] EV: A lot of things here, but we got a very important yesterday though. So for example, the whole deal for Latin America in 2018 was somewhere between $2 billion and $3 billion, okay? Yesterday SoftBank announced a new fund focused on just Latin America of $5 billion. Yeah, we're emerging, but I think – We're betting our life on it, that we're going to be at least somewhere near China in the years someday. Yeah, things are starting to improve a lot.

We already have unicorns here. We have two unicorns in both Columbia, for example, LifeMiles and Rappi. Rappi, you might have heard of them, Y Combinator baked, Sequoia, Andreessen Horowitz, etc., and there are delivery for like – But not only for groceries, but for every new groceries, restaurants, etc., etc. These guys of Rappi have been inspiring a whole new generation of entrepreneurs and of software developers as well, because software – Colombia and Latin America and software developers in the past were lured more towards going to Europe or North America to work for a big, giant, like Google or Microsoft of something like that instead of staying here in Latin America and starting a startup or working for a startup here.

So yeah, that's starting to change and we still have to compete with the U.S. salaries which are way higher. Of course, it's suggested that we have really low living cost here and life is nice

here and people might want to be near their college friends, near their family. Yeah, Latin America have to compete with that, not versus their U.S. counterparts, not with the money, but like showing them the fact that they can keep growing happy here in their home city, home country and with other things as well.

Equity, that's another thing as well. By the way, of course you can have like a Delaware C Corp here. Stripe Atlas has lower [inaudible 00:51:22]. So having a Delaware C Corp [inaudible 00:51:25] your equity to your employees and investors is really easy. Yeah, you also compete with equity, but I think that the most important thing that is used to convince developers to work for you instead of going to North America or Europe is telling them that you're going to work on an interesting technical problem to solve.

Yeah, a lot of school [inaudible 00:51:44], which is totally fine. You need something a product that people absolutely need and it's giving millions of jobs to an underserved population. But we are now starting to see more hard tech stuff. For example, another startup from Bogota, which actually is emerging, now [inaudible 00:52:03] bought now. They're building robots for automated full delivery. So this guy started – Like here in Colombia University and then in Chile, Mexico, etc., basically big and on demand delivery up focused on university campuses.

Then they went to UC Berkley, opened UC Berkley there, but they realized that a human delivery person costs a lot in the U.S. So they said, "You know what? We're going to automate this. We're going to build robots with [inaudible 00:52:32] computer vision and are a really cool technology. [Inaudible] really has made a really good job [inaudible 00:52:39] talent.

Yeah, we're not starting to see more companies solving really – Or harder technical problems. I think those are the main two. Oh and another thing, we do have a lot of Latin American VCs, but unfortunately most of like startups are something common now in Latin America, but not a cyber security startup. Yeah, cyber security is still like a really abstract thing for most business to fund here. So we're actually going this summer to the U.S. to raise our seed round.

[00:53:10] JM: Okay. I want to talk more about the investment side of things. After the show hopefully I can introduce you to some people. I think you got a really interesting business.

[00:53:18] EV: Thanks, man. Thanks.

[00:53:19] JM: Yeah, absolutely. Last question I want to ask you both because you're on the younger side of things and you're in Latin America, where cryptocurrency is actually seeing some of its earliest use cases and actually like I think positive. Positive like really needed use cases at least as far as I read in the news.

[00:53:39] EV: And bad ones as well, like drug lords them. Good cases and bad cases as well, unfortunately. Yeah, that's the truth.

[00:53:49] JM: Unfortunately, yeah. But why don't you lay out – What's your vision for how cryptocurrency are going to affect the world. You hear such divergent opinions from people these days. What do you think happens in the next 5 to 10 years?

[00:54:04] EV: First of all, I think that we're all guessing. I don't think there's still much data to back up like a real strong opinion. But I do have an opinion, and my opinion is the following. I was like 5-years-old when the .com bubble happened. So I don't have a memory from those days, but I've read from that episode of history, the .com bubble. What I've read is that like everyone was pouring hundred million dollar rounds to startups who didn't have a product and companies with IPO, 6 months software incorporating and just having a .com in their name would increase their Nasdasq value like twice or something like that. That was [inaudible 00:54:49] of course and a lot of people lost tons of cash in that and their jobs and it was horrible. But it left us a really cool thing called the web and the internet. Yeah, that changed the world for a lot of good.

I think that these cryptocurrency bubble, because I think it is a bubble, will leave us with something good. I think that that something good is going to be watching technology. I think that whole decentralization thing and the transparency that blockchains gives to [inaudible 00:55:24] and things like that, I think that that's going to be the good side of this bubble.

[00:55:30] JM: Esteban, it's been really fun talking to you. I appreciate you coming on Software Engineering Daily.

[00:55:33] EV: Thanks, Jeff. Thank you so much for your time.

[END OF INTERVIEW]

[00:55:40] JM: This podcast is brought to you by wix.com. Build your website quickly with Wix. Wix code unites design features with advanced code capabilities, so you can build data-driven websites and professional web apps very quickly. You can store and manage unlimited data, you can create hundreds of dynamic pages, you can add repeating layouts, make custom forms, call external APIs and take full control of your sites functionality using Wix Code APIs and your own JavaScript. You don't need HTML or CSS.

With Wix codes, built-in database and IDE, you've got one click deployment that instantly updates all the content on your site and everything is SEO friendly. What about security and hosting and maintenance? Wix has you covered, so you can spend more time focusing on yourself and your clients.

If you're not a developer, it's not a problem. There's plenty that you can do without writing a lot of code, although of course if you are a developer, then you can do much more. You can explore all the resources on the Wix Code's site to learn more about web development wherever you are in your developer career. You can discover video tutorials, articles, code snippets, API references and a lively forum where you can get advanced tips from Wix Code experts.

Check it out for yourself at wicks.com/sed. That's wix.com/sed. You can get 10% off your premium plan while developing a website quickly for the web. To get that 10% off the premium plan and support Software Engineering Daily, go to wix.com/sed and see what you can do with Wix Code today.

[END]