High(er) Fidelity Software Supply Chain Attack Detection

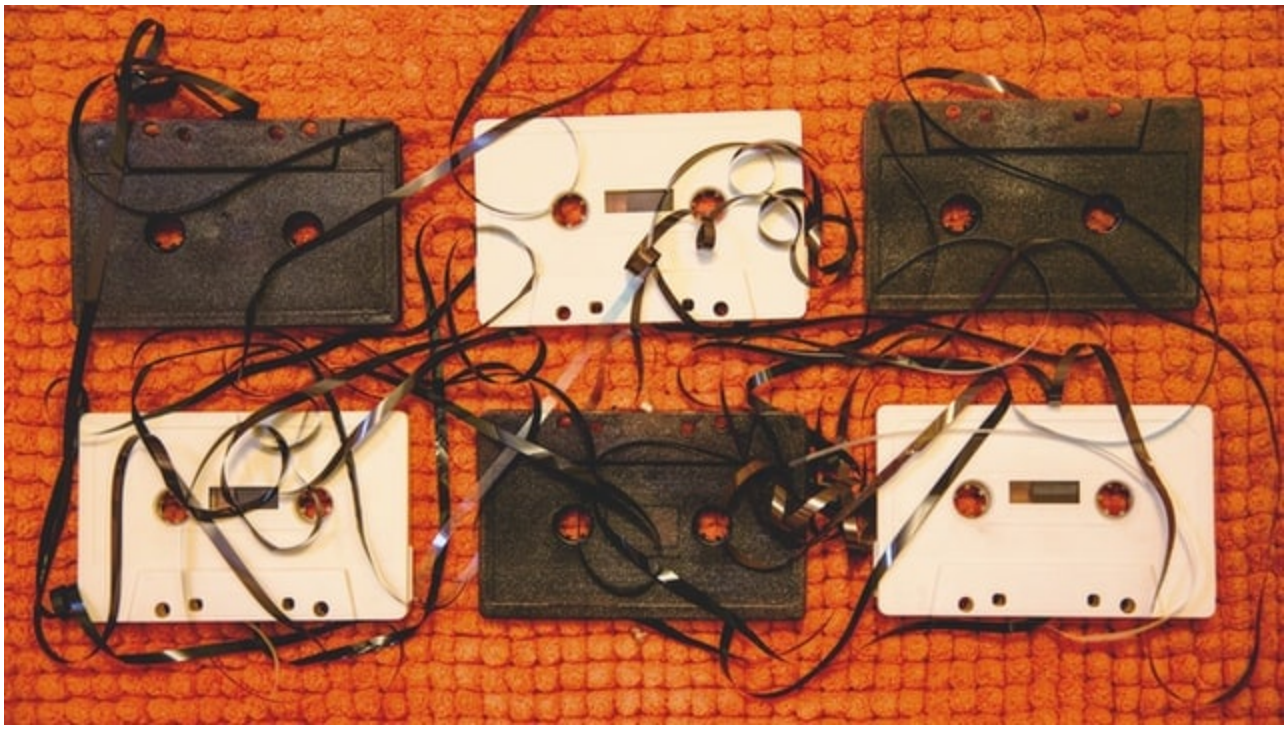
splunk.com/en_us/blog/security/high-er-fidelity-software-supply-chain-attack-detection.html

October 26, 2021

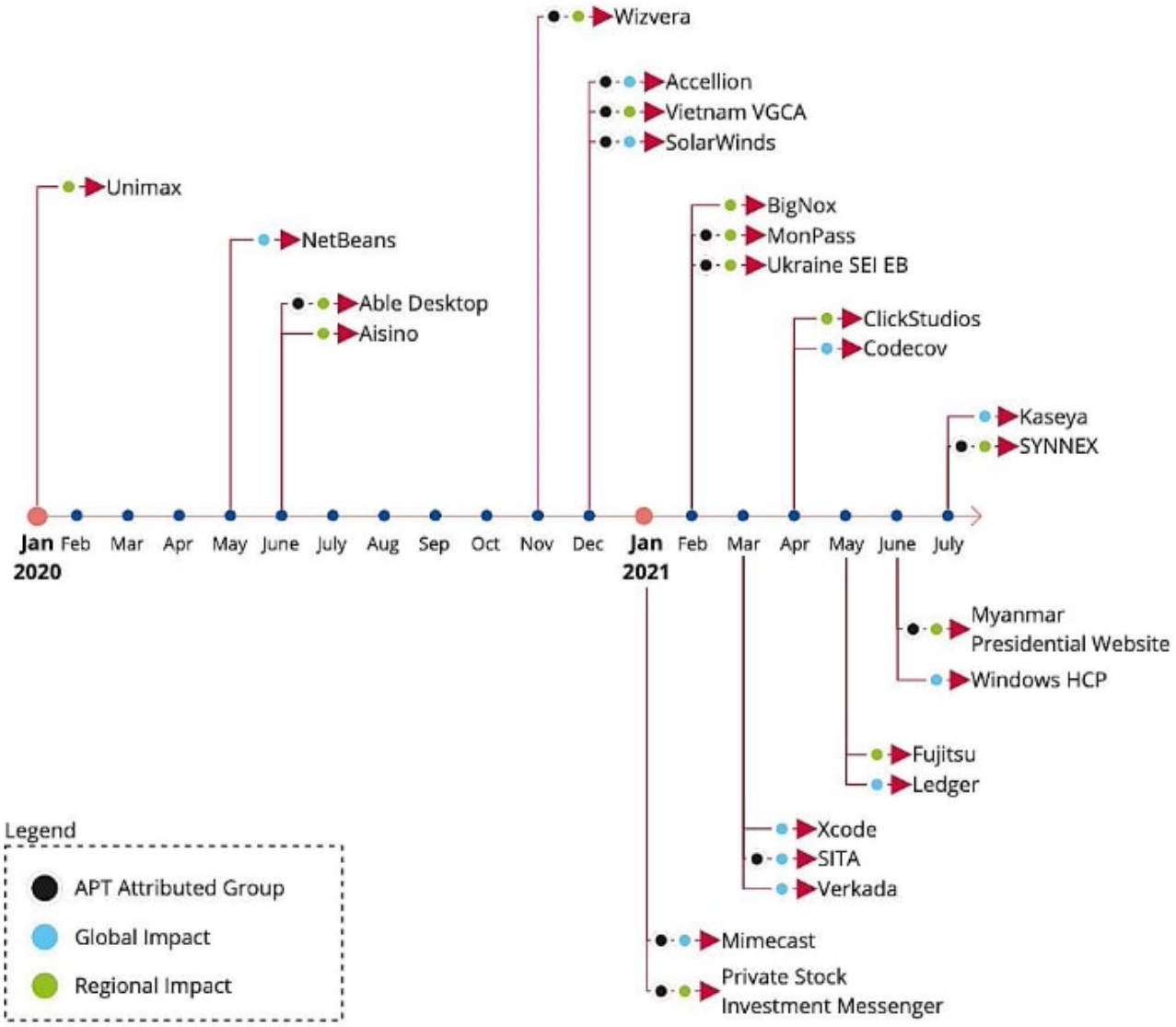
By Marcus LaFerrera October 26, 2021

Over the last year, many of us have been introduced to the term “Software Supply Chain”. For better or worse, it is now part of our defense vernacular and won’t be going away any time soon. If anything, it has consumed us in many ways and has been the cause of many nights of lost sleep. Well, that could just be us on the SURGe team here at Splunk.

Last year, after the Solarwinds Orion supply chain attack, we decided to take a closer look



at software supply chain attacks. We knew it would be a challenging subject to tackle, especially considering the wide swath of software, services, infrastructure, people, and all of the other bits and bobs that make up our computing life. It turns out that software supply chain attacks are far more common than we'd like to think. The European Union Agency for Cybersecurity (EINSA) recently published a report detailing about 25 known software supply chain attacks between January 2020 and July of 2021 alone.



Source: https://www.enisa.europa.eu/publications/threat-landscape-for-supply-chain-attacks

Over the last few months, we've embarked on some research to explore methods to potentially detect supply chain attacks and shed some light on this dark corner of cyber defense. We purposefully limited our scope to methodologies that had the potential to be useful in detecting abnormal activity on critical servers that are terrifying to look at sideways, nevermind patching. Why would we limit our scope so much? Well, it turns out detecting supply chain attacks isn't easy. It's actually pretty hard. We also wanted to ensure that whatever the end result was, it would be useful for a majority of our readers, not just the wizards, as Ryan Kovar, leader of the SURGe team, would say.

After much deliberation, we came to the conclusion that focusing on JA3 and JA3s hashes would be a fun and perhaps rewarding path forward. What is this magical alphanumeric string I speak of? I'm glad you asked. We won't go into much detail here, but suffice it to say it is a fairly simple, yet ingenious, method to fingerprint TLS negotiations between a client and a server. I know, you want to know more. Well, we recommend that you check out

some amazing research by Lee Brotherson and even more from the fine folks over at Salesforce that open-sourced their JA3 code. After reading that, you should come away with a great understanding of TLS fingerprinting and how JA3 works.

Can You At Least Show Me Something?

Of course! Let's go over the useful information rather than just talk about our inspirational journey. What would a Splunk blog post be without some screenshots of Splunk? One of the queries we developed uses Splunk's anomalydetection command. In this query we're going to simply try to find some anomalous TLS sessions in our Zeek logs coming from our critical server netblock, 192.168.70.0/24.

```
sourcetype="bro:ssl:json" ja3="*" ja3s="*" src_ip IN (192.168.70.0/24)

| anomalydetection method=histogram action=annotate pthresh=0.0001 src_ip, ja3, ja3s

| stats sparkline max(log_event_prob) AS "Max Prob", min(log_event_prob) AS "Min
Prob", values(probable_cause) AS "Probable Causes", values(dest_ip) AS "Dest IPs",
values(server_name) AS "Server Names", values(ja3) AS "JA3", values(src_ip) as
"Source IPs" count by ja3s

| table "Server Names", "Probable Causes", "Max Prob", "Min Prob", "Dest IPs", ja3s,
"JA3", "Source IPs", count

| sort "Min Prob" asc
```

Once we run this query, we'll be able to see that there are some anomalies in the data we should probably take a look at. Namely, the TLS sessions that are reaching out to manic.imperial-stout.org and update.lunarstiiiness.com.

Events Patterns Statistics (25) Visualization

20 Per Page ▾ Format Preview ▾ < Prev 1 2 Next >

	Server Names	Probable Causes	Max Prob	Min Prob	Dest IPs	ja3s	JA3	Source IPs	count
1	sls.update.microsoft.com	ja3s	-15.2435	-15.2911	20.54.89.106	17e97216fa7f4ec8c43090c6eed97c25	bd0bf25947d4a37404f0424edf4db9ad	192.168.70.19 192.168.70.227	2
2	storecatalogrevocation.storequality.microsoft.com	ja3s	-15.2435	-15.2911	104.93.156.139 23.14.171.52	35af4c8cd9495354f7d701ce8ad7fd2d	bd0bf25947d4a37404f0424edf4db9ad	192.168.70.19 192.168.70.227	2
3	settings-win.data.microsoft.com	ja3s	-14.5745	-14.6221	52.137.106.217 52.167.17.97	3ffaa1393a2bf5ecfc7b6b2323452f2d	bd0bf25947d4a37404f0424edf4db9ad	192.168.70.19 192.168.70.227	4
4	ad.froth.ly login.live.com login.microsoftonline.com	ja3s	-14.3562	-14.4038	192.168.70.227 40.126.29.5 40.126.29.6 40.126.29.7 40.126.29.8	7d8fd34fdb13a7fff30d5a52846b6c4c	bd0bf25947d4a37404f0424edf4db9ad	192.168.70.19 192.168.70.227	5
5	manic.imperial-stout.org	ja3	-14.3577	-14.3577	161.35.19.170	0eec924176fb005dfa419c80ab72d27c	54328bd36c14bd82ddaa0c04b25ed9ad	192.168.70.19	10
6	clients2.google.com storage.googleapis.com update.googleapis.com	ja3s	-14.1772	-14.2248	142.250.217.112 142.250.69.195 142.250.69.206 142.250.69.208 142.251.33.67	eca9b8f0f3eae50309eaf901cb822d9b	bd0bf25947d4a37404f0424edf4db9ad	192.168.70.19 192.168.70.227	6
7	softlines-trova.s3-us-west-2.amazonaws.com	ja3s	-13.7452	-13.7452	52.218.237.129	704239182a9091e4453fdbfe0fd17586	5b305623f54f097036bebf649e702c4d	192.168.70.19	1
8	update.lunarstiiiness.com	ja3s	-12.7091	-12.7091	143.244.189.78	15af977ce25de452b96affa2addb1036	5b305623f54f097036bebf649e702c4d	192.168.70.19	3
9	images-na.ssl-images-amazon.com m.media-amazon.com	ja3s	-12.7091	-12.7091	104.71.132.13	15c4d139d9f284ce5a6e4380e77c1f5c	5b305623f54f097036bebf649e702c4d	192.168.70.19	3
10	web.vortex.data.microsoft.com	ja3s	-12.6615	-12.6615	64.4.54.254 65.55.44.109	9cac3f41e89d651cd76e799381601768	5b305623f54f097036bebf649e702c4d	192.168.70.227	3
11	www.amazon.com	ja3s	-12.4295	-12.4295	104.71.134.207	cb101004a95f86e96902c2919db762c7	5b305623f54f097036bebf649e702c4d	192.168.70.19	4
12	update.lunarstiiiness.com	ja3s	-12.3883	-12.3883	143.244.189.78	b653c251b0ee54c3088fe7bb997cf59d	3b5074b1b5d032e5620f69f9f700ff0e	192.168.70.19	2

To make these results even more useful, we can also include an allow list in our query. This will help to ensure some of the more benign hosts that we would expect to see (such as legitimate software update sessions) aren't included in our results.

Events Patterns **Statistics (6)** Visualization

20 Per Page ▾ Format Preview ▾

	Server Names	Probable Causes	Max Prob	Min Prob	Dest IPs	ja3s	JA3	Source IPs	count
1	ad.froth.ly	ja3	-18.5422	-18.5422	192.168.70.227	7d8fd34fdb13a7fff30d5a52846b6c4c	bd0bf25947d4a37404f0424edf4db9ad	192.168.70.19	1
2	update.lunarstiiiness.com	ja3	-16.4701	-16.4701	143.244.189.78	15af977ce25de452b96affa2addb1036	5b305623f54f097036bebf649e702c4d	192.168.70.19	3
3	update.lunarstiiiness.com	ja3s	-15.8991	-15.8991	143.244.189.78	b653c251b0ee54c3088fe7bb997cf59d	3b5074b1b5d032e5620f69f9f700ff0e	192.168.70.19	2
4	manic.imperial-stout.org	ja3s	-14.8327	-14.8327	161.35.19.170	ec74a5c51106f0419184d0dd08fb05bc	3b5074b1b5d032e5620f69f9f700ff0e	192.168.70.19	6
5	manic.imperial-stout.org	ja3	-14.1078	-14.1078	161.35.19.170	0eec924176fb005dfa419c80ab72d27c	54328bd36c14bd82ddaa0c04b25ed9ad	192.168.70.19	10
6					18.236.219.157	e35df3e00ca4ef31d42b34bebaa2f86e	c34a54599a1fbaf1786aa6d633545a60	192.168.70.19 192.168.70.227	8318

Now we're getting somewhere. All of our suspicious queries have bubbled their way right to the top of our list. Right where suspicious data belongs.

Great Teaser, Is There More We Can See?

We spent a lot of time not only conducting our research, but we also wrote a beautiful white paper that provides far more detail than this blog post. Not enough? No problem. Ryan Kovar and I are also presenting on this very topic at Splunk's annual conference, .conf21. You can check out our talk Hunting the Known Unknown: Supply Chain Attacks (SEC1745C) to see us in all of our on-screen glory. Still want more? Did I mention that we've written a white paper? You can check out all of the queries we've developed, along with a few bits of code we developed specifically for this research. One of them being a nifty way to generate anomalous TLS sessions that mimic Zeek SSL logs.

Do you want even more? Well then, you've come to the right place. You can also play along today, as in right now, with Boss of the SOC (BOTS). A software supply chain scenario was designed and built out by John Stoner, which was integral to our research efforts.

If software supply chain in the DevSecOps realm is more your cup of tea, we'd highly suggest you take a look at Dave Herrald and Chris Riley's talk from this years' .conf21, Enabling DevSecOps and Securing the Software Factory with Splunk (SEC1108C). They put together a lot of terrific content to better understand how Splunk can be used in your DevSecOps pipeline to better defend against and detect software supply chain attacks.

One Step Ahead

Software supply chain attacks are not going away. As our network defenses improve, adversaries must move up the chain to stay a step ahead of our defenses. This cat and mouse game is the nature of network defense. It's why we lose sleep, working to think of new and novel ways to help everyone detect adversaries and stay safe. Our research presented here sadly isn't a silver bullet. It won't solve the software supply chain problem

today, or at all. What it will do however, is provide everyone with another tool and means of staying ahead of adversaries. With any luck, it will also help start a conversation and inspire others to explore interesting avenues of detection, ensuring we can all sleep better at night.

Posted by

Marcus LaFerrera

US | JP | Pentagon | DARPA | Splunk