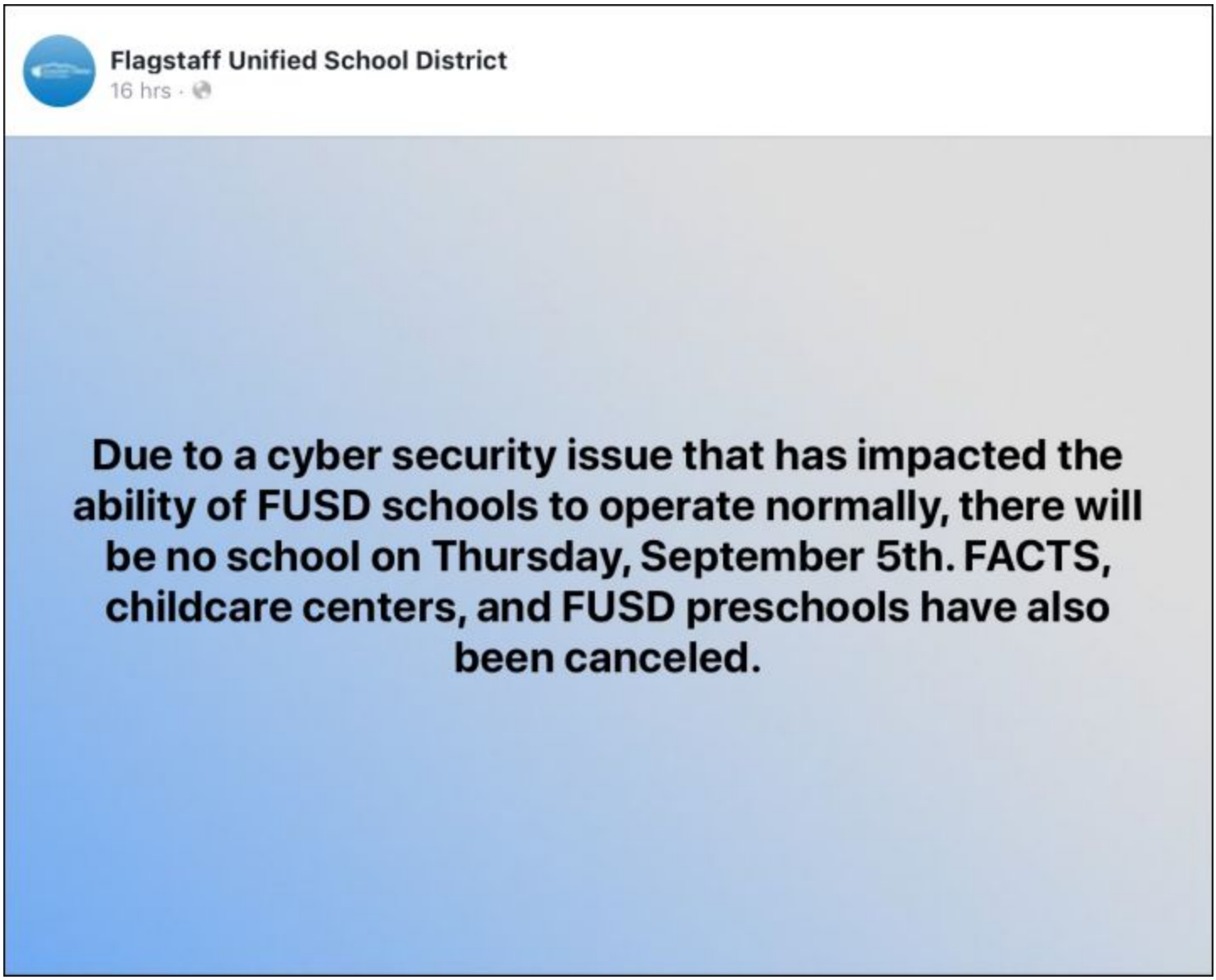
Security Now! #731 - 09-10-19

DeepFakes

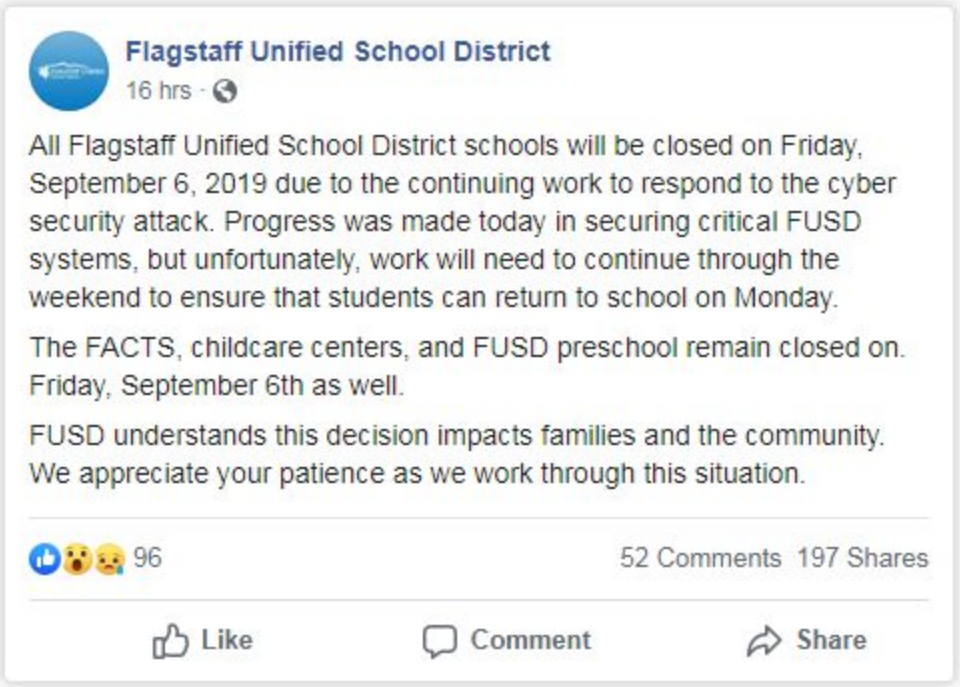
This week on Security Now!

This week we look at a forced two-day recess of all schools in Flagstaff, Arizona, the case of a Ransomware operator being too greedy, Apple's controversial response to Google's posting last week about the watering hole attacks, Zerodium's new payout schedule and what it might mean, the final full public disclosure of BlueKeep exploitation code, some potentially serious flaws found and fixed in PHP that may require our listener's attention, some SQRL news, miscellany, and closing-the-loop feedback from a listener. Then we take our first look on this podcast into the growing problem and threat of "DeepFake" media content.

All Flagstaff Arizona Schools Cancelled Thursday, August 5th

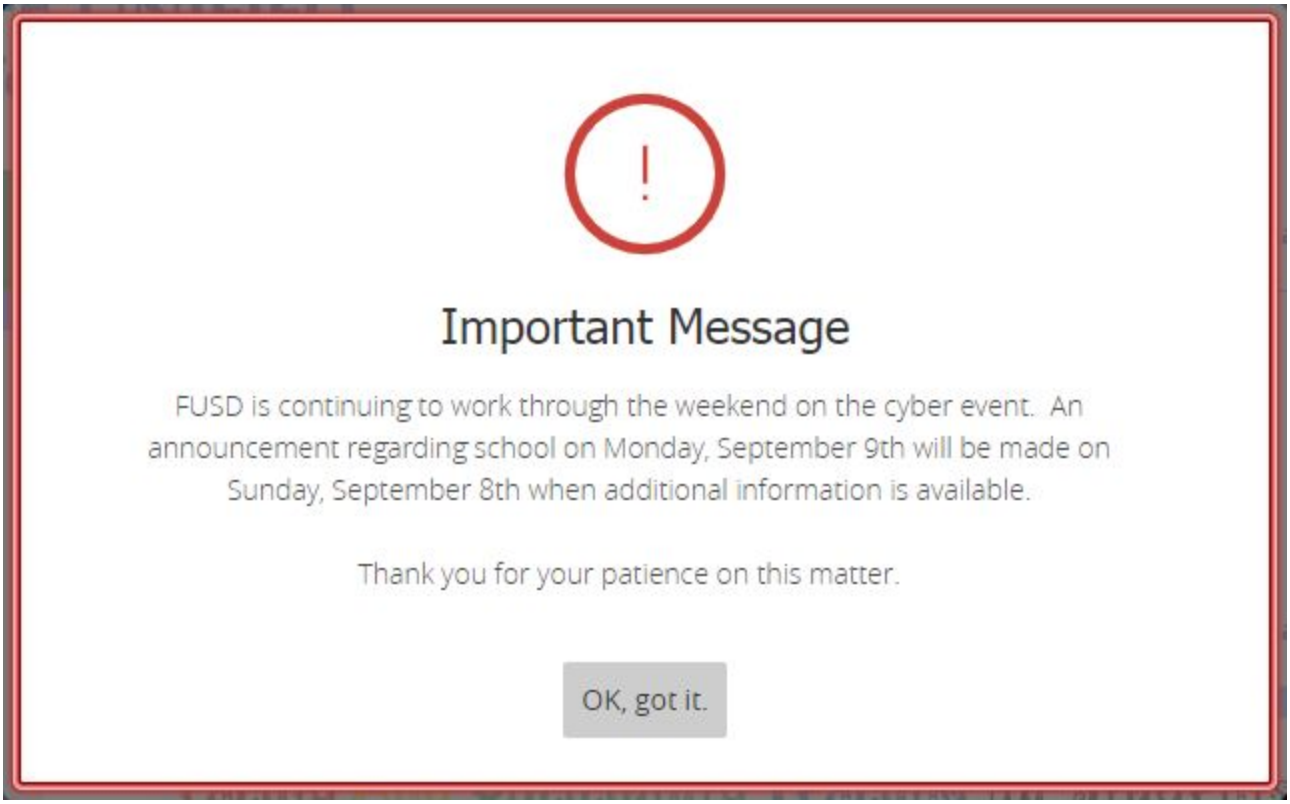


And not surprisingly, recess is extended through Friday:

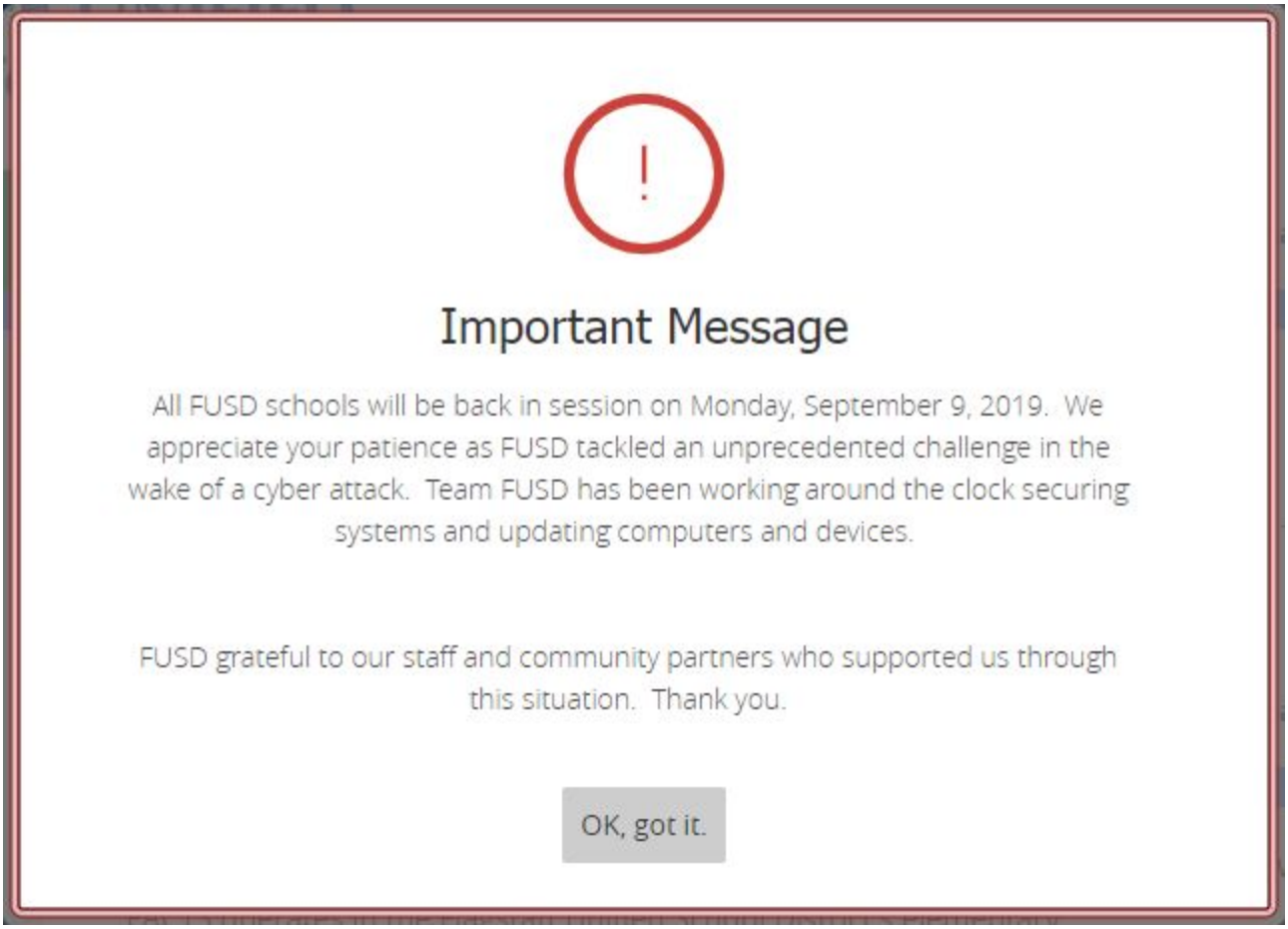


https://www.fusd1.org/facts https://www.facebook.com/FUSD1/

And Saturday…



And Sunday...



Security News

A lesson for greedy ransomware: Ask for too much... and you get nothing!
After two months of silence, last Wednesday Mayor Jon Mitchell of New Bedford, Massachusetts held their first press conference to tell the interesting story of their ransomware attack...

The city's IT network was hit with the Ryuk (ree-ook) ransomware which, by the way, Malwarebytes now places at the top of the list of file-encrypting malware targeting businesses. It'll be interesting to see whether So-Dino-Kee-Bee's affiliate marketing model is able to displace Ryuk.

But, in any event, very fortunately for the city of New Bedford, hackers breached the city's IT network and got Ryuk running in the wee hours of the morning following the annual 4th of July holiday. This may not sound "fortunate", but the Mayor said that the ransomware spread through the city's network and proceeded to encrypt the files of 158 workstations, which accounted for a mere 4% of the city's total fleet of 4,000 PCs. The attack would have been much worse, but most of the city systems were off at the time, which prevented the ransomware from spreading through the entirety of the city's network.

Since this year's 4th of July fell on a Thursday, the holiday was turned into a 4-day stretch, which was fortunate, since some IT staff discovered the ransomware the next day when most of the city's computers were still off. So they were able to move quickly to disconnect the infected computers from the city's network and contain the infection before it could cause even more harm.

At a press conference held last Wednesday, exactly two months following the attack, Mayor Jon Mitchell said: "While the attack was still underway, the city, through its consultants, reached out to the attacker, which had provided an email address. The attacker responded with a ransom demand specifically that it would provide a decryption key to unlock the encrypted files in return for a Bitcoin payment equal to $5.3 million."

At that moment the city didn't pay, primarily because it didn't have the funds. If it had paid it would have been the largest ransomware payment ever, dwarfing the previous record of $1 million which was paid by a South Korean web hosting firm.

Knowing they couldn't pay, Mayor Mitchell said the city decided to engage in a conversation with the hackers as a stalling tactic to give their IT staff more time to bolster the city's defenses and protect their network in the case the attackers might have been able to take additional action beyond running the ransomware.

At last Wednesday's press conference the Mayor said: "In light of these considerations, I decided to make a counter-offer using our available insurance coverage in the amount of $400,000, which I determined to be consistent with the other ransoms which had recently been paid by other municipalities."

However, the attacker declined to make a counter-offer and rejected the city's position outright. And with that, since the hackers wouldn't negotiate and since the city didn't have $5.3 million dollars anyway, they decided to restore from backups. The city's decision to restore from backups was easy thanks to the relatively low number of infected systems, and the fact that no critical systems had been impacted by the ransomware. And that, in turn, made managing the public pressure easier than in other municipalities where ransomware infections effectively crippled almost all city services.

Checking back in on Texas' 22 Sodinobiki victims…
We learn that the $2.5 million in ransom that was demanded by the attackers were declined and the 22 municipalities are in the process of bringing themselves back on line one by one.

Three weeks after the incident took place, the Texas Department of Information Resources (DIR) said that more than half of the impacted entities are now back to operations as usual. Some cities restored impacted systems from backups, while other rebuilt networks from scratch. This allowed municipalities to avoid paying ransom demands.

The incident responders who managed the ransomware infections at the 22 Texas municipalities have published advice this week that companies and government organizations can follow:

- Only allow authentication to remote access software from inside the provider's network
- Use two-factor authentication on remote administration tools and Virtual Private Network tunnels (VPNs) rather than remote desktop protocols (RDPs)
- Block inbound network traffic from Tor Exit Nodes
- Block outbound network traffic to Pastebin
- Use Endpoint Detection and Response (EDR) to detect Powershell (PS) running unusual processes.

Apple responds to Ian Beer's Project Zero posting
https://www.apple.com/newsroom/2019/09/a-message-about-ios-security/

> Last week, Google published a blog about vulnerabilities that Apple fixed for iOS users in February. We've heard from customers who were concerned by some of the claims, and we want to make sure all of our customers have the facts.
> First, the sophisticated attack was narrowly focused, not a broad-based exploit of iPhones "en masse" as described. The attack affected fewer than a dozen websites that focus on content related to the Uighur community. Regardless of the scale of the attack, we take the safety and security of all users extremely seriously.
>
> Google's post, issued six months after iOS patches were released, creates the false impression of "mass exploitation" to "monitor the private activities of entire populations in real time," stoking fear among all iPhone users that their devices had been compromised. This was never the case.
>
> Second, all evidence indicates that these website attacks were only operational for a brief period, roughly two months, not "two years" as Google implies. We fixed the vulnerabilities in question in February — working extremely quickly to resolve the issue just 10 days after we learned about it. When Google approached us, we were already in the process of fixing the exploited bugs.
>
> Security is a never-ending journey and our customers can be confident we are working for them. iOS security is unmatched because we take end-to-end responsibility for the security of our hardware and software. Our product security teams around the world are constantly iterating to introduce new protections and patch vulnerabilities as soon as they're found. We will never stop our tireless work to keep our users safe.

Last week the head of Threat Research for RiskIQ told *ZDNet* for their reporting on this that the attacks were, indeed, very targeted, and that Google was wrong in its initial assessment.

He later shared the same thoughts in a public tweet, and in it we see some filtering code:

The JS code above waits for 60 seconds after the page is initially loaded. It then retrieves a JSON structure containing the user's country of origin from the ip.nf site:

Here's documentation on the cool site: https://ip.nf/ Try this on yourself: https://ip.nf/me.json Returns a JSON structure containing the country: https://ip.nf/me.json?callback=jsons

With that, it invokes a function named "j_call_back" which checks the JSON structure for country == "China" and it ONLY injects a new <script> tag into the user's page DOM in the event that country == China.

Also… according to very good and comprehensive analysis and reporting by Volexity, the same targeted attacks have been waged against Android visitors: https://www.volexity.com/blog/2019/09/02/digital-crackdown-large-scale-surveillance-and-exploitation-of-uyghurs/
If Google knew that, it would have been more balanced to make note of the fact that it wasn't

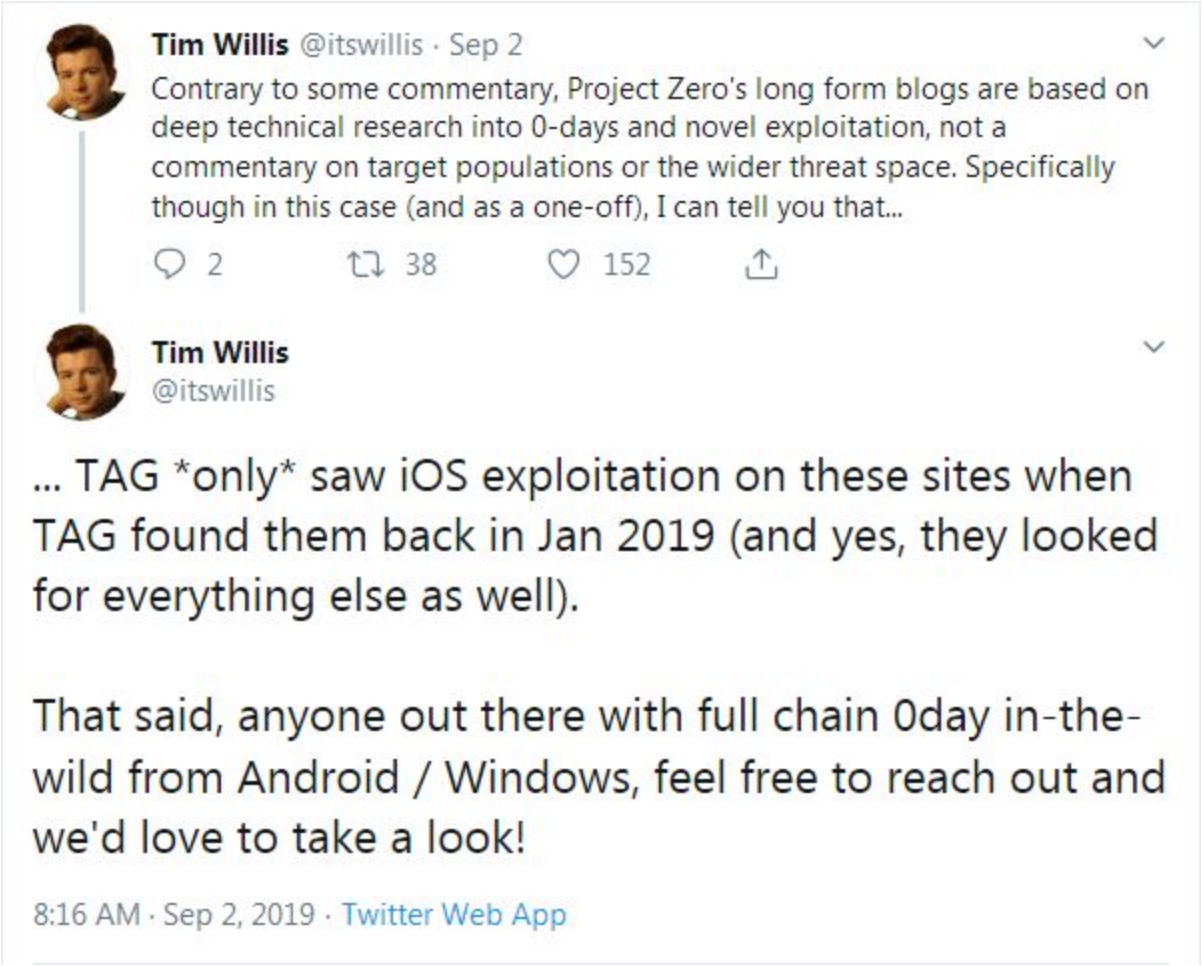
only iOS users. And since Google's blog posting which we covered last week, Google has been lambasted by the cyber-security community for only disclosing details about the coordinated campaign targeting iOS users, but not the one that targeted Android devices.

ZDNet noted that Google is standing by its research and its researchers. In a statement sent to ZDNet, Google said it stands by its original research, despite Apple's rebuttal:

> "Project Zero posts technical research that is designed to advance the understanding of security vulnerabilities, which leads to better defensive strategies. We stand by our in-depth research which was written to focus on the technical aspects of these vulnerabilities."

ZDNet wrote that Tim Willis, a Google Project Zero member said that this wasn't a sign of Google being duplicitous (and trying to sabotage a rival on the mobile OS market) but that Google researchers only saw the malicious code targeting iOS devices.

"[Google's Threat Analysis Group] only saw iOS exploitation on these sites when TAG found them back in Jan 2019," he said, "and yes, they looked for everything else as well."



Finally, ZDNet sums up the whole affair by writing: Willis' assertion, made on September 2, was confirmed three hours later, when Volexity published its report about the hacking campaign targeting Android users, where the company confirmed there was no overlap between the Android and iOS campaigns.

The conclusion here is that Apple is right in calling out Google, at least on the first point -- that the campaign wasn't an en-mass hacking spree aimed at all random iPhone users, but rather a very targeted operation.

Nevertheless, most of the cyber-security community also pointed out that Apple itself is pretentious because it failed to alert users when it learned of this hacking campaign back in February. Most tech companies clearly mark vulnerabilities that are under attack in security updates. But Apple has never done this, and it didn't mention back in February that some of the bugs it fixed were under active exploitation.

Yes, Google might have exaggerated its claims, but Apple isn't the victim here. The Uighur minority is, which, ZDNet writes... Apple failed to protect. And I'll note that this disparity in owning up to the full nature of vulnerabilities serves to artificially inflate Apple's security posture.

Dan Goodin concluded his excellent coverage of this for ArsTechnica by writing:

> "Another key criticism is that Apple's statement has the potential to alienate Project Zero, which according to a Google spokesman has to date privately reported more than 200 vulnerabilities to Apple. It's easy to imagine that it wasn't easy for Apple to read last week's deep-dive report publicly documenting what is easily the worst iOS security event in its 12-year history. But publicly challenging a key ally on such minor details with no new evidence does not create the best optics for Apple."
>
> "Apple had an opportunity to apologize to those who were hurt, thank the researchers who uncovered systemic flaws that caused the failure, and explain how it planned to do better in the future. It didn't do any of those things. Now, the company has distanced itself from the security community when it needs it most."

And speaking of which... https://zerodium.com/program.html#changelog

Changelog / Sep 3rd, 2019

Sep. 3, 2019 - Payouts for major mobile exploits have been modified. Changes are highlighted below:

Category	Changes
New Payouts (**Mobiles**)	**$2,500,000** - Android full chain (Zero-Click) with persistence (**New Entry**) **$500,000** - Apple iOS persistence exploits or techniques (**New Entry**)
Increased Payouts (**Mobiles**)	**$1,500,000** - WhatsApp RCE + LPE (Zero-Click) without persistence (previously: **$1,000,000**) **$1,500,000** - iMessage RCE + LPE (Zero-Click) without persistence (previously: **$1,000,000**)
Decreased Payouts (**Mobiles**)	**$1,000,000** - Apple iOS full chain (1-Click) with persistence (previously: **$1,500,000**) **$500,000** - iMessage RCE + LPE (1-Click) without persistence (previously: **$1,000,000**)
Desktops/Servers	No modifications.

Zerodium changes it payouts
TheHackerNews put it: "There's some good news for hackers and vulnerability hunters, though terrible news for Google, Android device manufacturers, and their billions of users worldwide."

It appears that the zero-day marketplace has recently shifted toward the Android operating system, with Zerodium suddenly bumping payouts to discoverers of the most severe class of Android 0-days up by a factor of 12.5! ... From its previous $200,000 to a whopping $2.5 million, thus moving it above the maximum $2 million payout for the equivalent “gold standard” exploit for iOS. In both cases these bounties are available to anyone who can find and provide a "full chain, zero-click, with persistence" zero-day.

But I wouldn’t consider this to be, as ThehackerNews put it “terrible news for Google, Android device manufacturers, and their billions of users worldwide.” I think that’s not quite the right way to look at it. Yes… If you were a member of an oppressed group or minority which a repressive regime might have a strong interest in monitoring, while also having the money to spend for the purchase the transient ability to do so, then, yeah; all other things being equal, the increased bounty on the Android side -- and it IS a significant whopping increase -- will increase the likelihood that hackers will be trying to pry into Android rather than iOS.

TheHackerNews notes that: "Just like other traditional markets, the zero-day market is also a game of supply, demand, and strategy, which suggests either the demand of Android zero-days has significantly increased or somehow Android OS is getting tougher to hack remotely, which is unlikely."

As we know, Zerodium is a controversial enterprise that purchases zero-day exploits from hackers, and then, so far as we know, almost certainly resells them to law enforcement agencies and nation-sponsored spies around the world. That's the only way that their economic model would work.

So... Zerodium's latest update notification indicates that it's looking for hackers who can develop full chain 0-click Android exploits. We should also note that there's at least some chance that Zerodium is also serving as a middleman for those desiring to acquire exploits. So, for example, a potential exploit purchaser might have said to Zerodium: "We'll pay 3 million dollars for 90-days of exclusive access to a zero-click persistent exploit for a fully-patched Android device." Whereupon Zerodium turns around to make the offer -- less their middleman commission -- to the hacker/cracker community. The public doesn't have much visibility into Zerodium's inner workings.

Zerodium's payout for the same type of zero-day exploit for iOS devices is $2 million, which places it at double what Apple has recently started offering hackers to responsibly report such severe exploits, described as "a zero-click kernel code execution vulnerability that enables complete, persistent control of a device's kernel." So, if I'm a hacker with such an exploit do I try to get $1 million from Apple or $2 million from Zerodium?

Besides Android exploits, Zerodium has also announced some app-targeted payouts... $500,000 for submitting new persistence exploits or techniques for iOS, and increased payouts of WhatsApp and iMessage exploits.

BlueKeep has Flown the Coop!
All details of the dreaded wormable BlueKeep exploit are now public.
A Metasploit module was released Friday... and since it was now public, our friend Marcus Hutchins (AMA Malware Tech Blog) posted a very clear expose' ...

https://www.malwaretech.com/2019/09/bluekeep-a-journey-from-dos-to-rce-cve-2019-0708.html

Marcus titled his posting: "BlueKeep: A Journey from DoS to RCE (CVE-2019-0708)"
"*Due to the serious risk of a BlueKeep based worm, I've held back this write-up to avoid advancing the timeline. Now that a proof-of-concept for RCE (remote code execution) has been release as part of Metasploit, i feel it's now safe for me to post this. This article will be a follow on from my previous analysis.*"

Marcus' previous analysis was posted back in May simply titled: "Analysis of CVE-2019-0708 (BlueKeep)" and in that posting he begins... "*I held back this write-up until a proof of concept (PoC) was publicly available, as not to cause any harm. Now that there are multiple denial-of-service PoC on github, I'm posting my analysis.*"

And, similarly, now that there is a publicly available remote code execution attack, Marcus has produced a wonderfully detailed description of the journey from a server crash (DoS) to an RCS - a remote code execution.

We have often noted that exploits typically begin their lives as a crash of the target system, thus a DoS in that services are being denied to other who wish to use the system... because it has crashed. But then that, if sufficient finesse is available, whatever it was that was invoking a crash can often be evolved into something that arranges to run an attacker's code.

Marcus picks up with posting #2 exactly where he left off with posting #1. I'm not going to get into any more detail here since it's truly down in the tech-weeds. But I have the links to both of his postings in the show notes for anyone who is interested in looking at the development of an RCS from a DoS...

https://www.malwaretech.com/2019/05/analysis-of-cve-2019-0708-bluekeep.html
https://www.malwaretech.com/2019/09/bluekeep-a-journey-from-dos-to-rce-cve-2019-0708.html

The most important takeaway for us is that the world now has access to everything needed for less skilled attackers to attack. The other worrisome aspect is that the attack allows the remote attacker to obtain the hashes of the credentials used by the other machines on the interior network, thus making this a very potent vulnerability. And we know that a huge number of machines remain unpatched.

Multiple Code Execution Flaws Found and Patched in PHP
As we know, PHP (Personal Home Page) continues to be the #1 most popular server-side web programming language in the world, laying claim to more than 78% of the Internet's servers.

So it's significant when the maintainers of PHP release updates to PHP which patch multiple high-severity vulnerabilities in PHP's core and bundled libraries... while nothing that the most severe of these could allow remote attackers to execute arbitrary code and compromise targeted servers.

So I wanted to give our listeners a heads up. It's not as if there's a publicly-exposed service that automatically exposes everyone. If that were the case this news would have been the title of the podcast and it would be a five alarm fire. But it is the case that some apps use these now-widely-known-to-be-vulnerable back-end PHP functions for various purposes, and in time exploits might be developed and leveraged against websites that use those add-on components.

Quoting from some of the tech press:

> Depending on the type, occurrence, and usage of the affected codebase in a PHP application, successful exploitation of some of the most severe vulnerabilities could allow an attacker to execute arbitrary code in the context of the affected application with associated privileges.
>
> Other the other hand, Failed attempts at exploitation will likely result in a denial of service (DoS) condition on the affected systems.
>
> The vulnerabilities could leave hundreds of thousands of web applications that rely on PHP open to code execution attacks, including websites powered by some popular content management systems like WordPress, Drupal and Typo3.
>
> Out of these, a 'use-after-free' code execution vulnerability, assigned as CVE-2019-13224, resides in Oniguruma, a popular regular expression library that comes bundled with PHP, as well as many other programming languages.
>
> A remote attacker can exploit this flaw by inserting a specially crafted regular expression in an affected web application, potentially leading to code execution or causing information disclosure.
>
> "The attacker provides a pair of a regex pattern and a string, with a multi-byte encoding that gets handled by onig_new_deluxe()," Red Hat says in its security advisory describing the vulnerability.
>
> Other patched flaws affect curl extension, Exif function, FastCGI Process Manager (FPM), Opcache feature, and more.
>
> Good news is that so far there is no report of any of these security vulnerabilities being exploited in the wild by attackers.
>
> The PHP security team has addressed the vulnerabilities in the latest versions. So users and hosting providers are strongly recommended to upgrade their servers to the latest PHP version

7.3.9, 7.2.22, or 7.1.32.

In most server system environments, like mine on GRC's server where I run PHP, PHP is a core component powering many of the site's features: My Wordpress blog, the SQRL XenForo forums which is a large PHP application, GRC's link shortener, and several other services are all PHP apps. Consequently, unlike an end-user application, there's no easy built-in pushbutton auto-update for PHP. The web server which invokes PHP needs to be stopped so that PHP is fully released and unloaded, then the PHP install needs to be updated, then the server needs to be restarted.

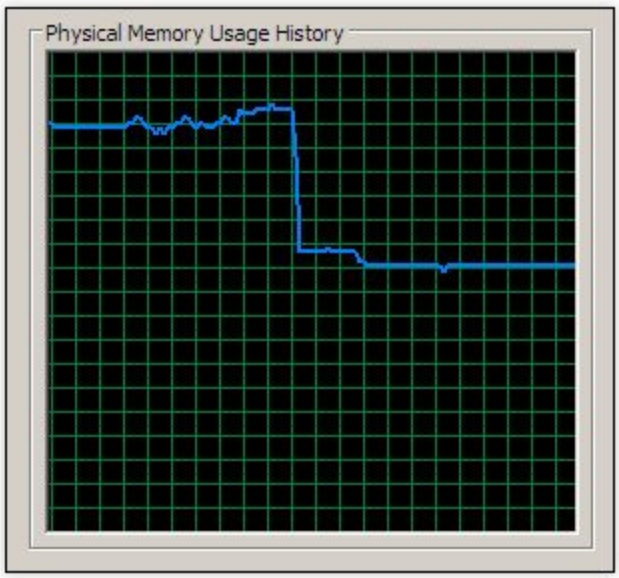
My point is... while it's not an emergency, if you are the responsible party you should plan to make some time to get that one done sometime when your various services can handle a brief update outage. You might sleep easier at night.

SQRL

The documentation project is completed.
Four documents, spanning 79 pages, which completely explain SQRL's feature, operation and implementation in sufficient detail for anyone to create SQRL clients and servers.

Miscellany

"UnloadTabs" in Firefox.
about:memory which issues garbage collection requests doesn't do it.
UnloadTabs takes up no space on the UI. It adds an option to the tabs context menu to unload this tab or all other tabs. That's perfect!



Watching TaskManager, you'll see an immediate huge drop, followed after a 10 to 20 seconds by additional memory releases.

SpinRite

Notre Poubelle (@notre_poubelle) / Saturday, 8:48pm
Hi Steve, I have a SpinRite question. I'm not currently a SpinRite customer. I have a hard drive that I suspect is having problems. I ran Windows chkdsk C: /f /r /x, a couple of times and it gets stuck for a very long time at the same percentage point. In this case, I actually don't care at all about the content of the hard drive. My question is, is it worth purchasing SpinRite given that I don't care about the contents of the hard drive? Or would it be wiser to just buy another hard drive. Assuming I did buy SpinRite and it fixed the hard drive, is this likely a temporary fix and I'll end up having to buy a hard drive anyway?

Closing The Loop

Steve S. @stevedsmi
I am an avid listener of Security Now, and wanted you to know that you are correct about the 22 towns that were infected with ransomware. A 3rd party was providing IT services, access to state DMV resources, utility payments, etc. The 3rd party had an OpenVPN connection to each municipality. Not sure where it started, but it spread everywhere. Keep up the good work! I enjoy the podcast.

OpenVPN: “On Demand” vs “Static network bridging”

DeepFakes

NYT: Fraudsters Used AI to Mimic CEO's Voice in Unusual Cybercrime Case
Scams using artificial intelligence are a new challenge for companies
https://www.wsj.com/articles/fraudsters-use-ai-to-mimic-ceos-voice-in-unusual-cybercrime-case-11567157402

Story begins: Criminals used artificial intelligence-based software to impersonate a chief executive's voice and demand a fraudulent transfer of €220,000 ($243,000) in March in what cybercrime experts described as an unusual case of artificial intelligence being used in hacking.

The CEO of a U.K.-based energy firm thought he was speaking on the phone with his boss, the chief executive of the firm's German parent company, who asked him to send the funds to a Hungarian supplier. The caller said the request was urgent, directing the executive...

Leo… we've never talked about this coming problem of "Deep Fake" audio and video, but it's been in the news a lot recently, and it really takes the whole concept of a fake eMail to another level.

In the very early days of eMail we might have taken a note apparently from our boss at face value and acted upon it. But those days are clearly long gone. And for some time, due to its power to convincingly fool and spoof, the term "Photoshopping" (of an image) is well established in contemporary vernacular. Now the crazy power of today's computational resources now moves us to an entirely new level of spoofing… we are beginning to be able to spoof complex time-varying signals -- both audio and video -- in real-time.

So today, upon receiving something important and unexpected we'd via eMail we might give our boss a call and say "hey… did you just send me an eMail asking me {whatever} ??" But notice how we confirm. We confirm by voice… because we still trust that. So the point here is… no one trusts a photo anymore, and we're on the brink of also losing our ability to trust real time electronic communication as well.

We're entering a future where it might be necessary for instructions to be: "If I ever tell you by eMail -- or even by phone -- to do something you don't expect, I want you to confirm it face-to-face. Just walk down the hall and stick you head in the door and ask "... did you just call me and ask me to {whatever}?" Think how much this changes our world.

Introducing the: Deepfake Detection Challenge
https://deepfakedetectionchallenge.ai/

The subheading reads: "Deepfake Detection Challenge invites people around the world to build innovative new technologies that can help detect deepfakes and tampered media. Identifying tampered content is technically challenging as deepfakes rapidly evolve, so we're working together to build better detection tools."

Facebook, Microsoft in partnership with academics from Cornell Tech, MIT, University of Oxford, UC Berkeley, University of Maryland, College Park and University at Albany-SUN have joined forces to sponsor a contest promoting research and development to combat deepfakes -- videos altered through artificial intelligence (AI) to mislead their viewers.

So we have the DFDC -- DeepFake Detection Challenge -- which aims to spur the industry to create technology that can detect and prevent deepfakes, according to a log post by Facebook's CTO Mike Schroepfer.

Mike's posting was titled: "*Creating a data set and a challenge for deepfakes*"

Data sets and benchmarks have been some of the most effective tools to speed progress in AI. Our current renaissance in deep learning has been fueled in part by the ImageNet benchmark. Recent advances in natural language processing have been hastened by the GLUE and SuperGLUE benchmarks.

"Deepfake" techniques, which present realistic AI-generated videos of real people doing and saying fictional things, have significant implications for determining the legitimacy of information presented online. Yet the industry doesn't have a great data set or benchmark for detecting them. We want to catalyze more research and development in this area and ensure that there are better open source tools to detect deepfakes. That's why Facebook, the Partnership on AI, Microsoft, and academics from Cornell Tech, MIT, University of Oxford, UC Berkeley, University of Maryland, College Park, and University at Albany-SUNY are coming together to build the Deepfake Detection Challenge (DFDC).

The goal of the challenge is to produce technology that everyone can use to better detect when AI has been used to alter a video in order to mislead the viewer. The Deepfake Detection Challenge will include a data set and leaderboard, as well as grants and awards, to spur the industry to create new ways of detecting and preventing media manipulated via AI from being used to mislead others. The governance of the challenge will be facilitated and overseen by the Partnership on AI's new Steering Committee on AI and Media Integrity, which is made up of a broad cross-sector coalition of organizations including Facebook, WITNESS, Microsoft, and others in civil society and the technology, media, and academic communities.

It's important to have data that is freely available for the community to use, with clearly consenting participants, and few restrictions on usage. That's why Facebook is commissioning a realistic data set that will use paid actors, with the required consent obtained, to contribute to the challenge. No Facebook user data will be used in this data set. We are also funding research collaborations and prizes for the challenge to help encourage more participation. In total, we are dedicating more than $10 million to fund this industry-wide effort.

To ensure the quality of the data set and challenge parameters, they will initially be tested through a targeted technical working session this October at the International Conference on Computer Vision (ICCV). The full data set release and the DFDC launch will happen at the Conference on Neural Information Processing Systems (NeurIPS) this December. Facebook will also enter the challenge but not accept any financial prize. Follow our website for regular updates.

This is a constantly evolving problem, much like spam or other adversarial challenges, and our hope is that by helping the industry and AI community come together we can make faster progress.

It's easy to understand FaceBook's interest in this, since they are a massive social media platform where we know that Russia was very actively posting false reports to stir things up before our last presidential election. And they and others will likely be doing so again.

Next month, on October 3rd in Boston, our panel will be talking about the challenges of establishing identity online… and that is exactly that.