Security Now! #880 - 07-19-22
RetBleed

This week on Security Now!
This week we start with a quick update on last week's Rolling Pwn problem. Then we look at the state of IPv4 space depletion and the rising price of an IPv4 address. We have an interesting report on the Internet's failed promise, Facebook's response to URL-tracker trimming, Apple's record-breaking Lockdown Mode bounty, ClearView Ai's new headwinds, a new feature being offered by ransomware gangs, the return of Roskomnadzor, last Tuesday's patches and some feedback from our listeners. Then we look at the details of the latest way of exfiltrating secrets from operating system kernels thanks to insecurities in Intel and AMD micro-architecture implementations. Yes, some additional bleeding.

Hmm....

Security News

The Rolling Pwn, take II

We have a follow up to last week's Honda-centric story where the Honda engineers made the mistake of allowing their system which resynchronizes their autos to their remotes, by allowing them to move back to a previous state. Resyncing would have been fine if the resync was only to forward to a later state. But there's no safe way to allow an earlier state to be restored.

Last week when we covered this, the spokesperson for Honda told The Record that hackers would need "sophisticated tools and technical know-how to mimic Remote Keyless commands and gain access to certain vehicles of ours." It didn't hit me until just now that that statement makes no sense. If hackers did not have "sophisticated tools and technical know-how to mimic Remote Keyless commands" in the first place, then NO rolling codes of any sort would be needed at all. It's specifically because hackers DO HAVE "sophisticated tools and technical know-how to mimic Remote Keyless commands" that it's necessary to design a system — which Honda failed to do — which would defeat hackers who had "sophisticated tools and technical know-how to mimic Remote Keyless commands."

But in any event, that's not why we're back here this week. In additional dialog spurred by last week's revelations, Honda also said: "Honda regularly improves security features as new models are introduced that would thwart this and similar approaches." And the spokesperson added that all "completely redesigned" 2022 and 2023 model year vehicles have an "improved system" that addresses the issue. Saying: "Currently this includes 2022 Civic, 2022 MDX and 2023 HR-V. Our newer system transmits codes that immediately expire, which would prevent this type of attack from being successful," the Honda spokesperson explained.

What's confusing is that the original hacking team used their system to crack ten Honda's with four of them being 2022 year models and one of those four being the Honda Civic, which this spokesperson claims fixes this problem. Also note that **all** rolling codes immediately expire. That's the entire point of having them rolling. They are inherently meant to be single-use codes.

The good news is, hacking cars is fun and doing so is an easy way to generate headlines, which is the only payoff most researchers seek or receive. Since the hardware required to do this is now available inexpensively off the shelf, we can be pretty sure that automakers' past laziness with regard to their auto's true security will no longer go unnoticed and will be making future headlines whenever and wherever it is found to be lacking.

The great IPv4 Address Space Depletion.

We've had a lot of fun through the years watching the saga of diminishing IPv4 address space. According to SIDN, the Netherlands' official domain registrar, IPv4 address price has doubled in the past year. Back in 2015 IPv4 space was selling for $5 per IP. By this time last year that $5 had grown to between $25 and $30 last year. And today, one year later, we're at $50 to $60 each. Contrast this to IPv6 where there is essentially no practical limit to address availability. IPv6 addresses are not only free, but they are so freely available that ISPs are handing out large chunks of IPv6 space to each of their residential subscribers. It's probably difficult to find a better example of an entrenched unwillingness to change than for IPv4 space to be selling at such a premium, and for its cost to be rising exponentially.

Confronting Reality in Cyberspace: Foreign Policy for a Fragmented Internet

The Council on Foreign Relations just published a monster 116-page report titled: “Confronting Reality in Cyberspace: Foreign Policy for a Fragmented Internet”

A pull quote from the article headlines the Executive Summary. It reads:
“*The utopian vision of an open, reliable, and secure global network has not been achieved and is unlikely ever to be realized. Today, the internet is less free, more fragmented, and less secure.*”

Though this report is obviously US Centric, I think that everyone will find this interesting. Here’s the report’s Executive Summary:

> *The global internet—a vast matrix of telecommunications, fiber optics, and satellite networks—is in large part a creation of the United States. The technologies that underpin the internet grew out of federal research projects, and U.S. companies innovated, commercialized, and globalized the technology. The internet’s basic structure—a reliance on the private sector and the technical community, relatively light regulatory oversight, and the protection of speech and the promotion of the free flow of information—reflected American values.*
>
> *Moreover, U.S. strategic, economic, political, and foreign policy interests were served by the global, open internet. Washington long believed that its vision of the internet would ultimately prevail and that other countries would be forced to adjust to or miss out on the benefits of a global and open internet.*
>
> *The United States now confronts a starkly different reality. The utopian vision of an open, reliable, and secure global network has not been achieved and is unlikely ever to be realized. Today, the internet is less free, more fragmented, and less secure.*
>
> *Countries around the world now exert a greater degree of control over the internet, localizing data, blocking and moderating content, and launching political influence campaigns. Nation-states conduct massive cyber campaigns, and the number of disruptive attacks is growing. Adversaries are making it more difficult for the United States to operate in cyberspace. Parts of the internet are dark marketplaces for vandalism, crime, theft, and extortion.*
>
> *Malicious actors have exploited social media platforms, spread disinformation and misinformation, incited disparate forms of political participation that can sway elections, engendered fierce violence, and promoted toxic forms of civic division.*
>
> *At the same time, the modern internet remains a backbone for critical civilian infrastructure around the world. It is the main artery of global digital trade. It has broken barriers for sharing information, supports grassroots organization and marginalized communities, and can still act as a means of dissent under repressive government regimes.*
>
> *As the Internet of Things (IoT) expands in coming years, the next iteration of the network will connect tens of billions of devices, digitally binding every aspect of day-to-day life, from heart monitors and refrigerators to traffic lights and agricultural methane emissions.*
>
> *The United States, however, cannot capture the gains of future innovation by continuing to pursue failed policies based on an unrealistic and dated vision of the internet.*

The United States needs a new strategy that responds to what is now a fragmented and dangerous internet. The Task Force believes it is time for a new foreign policy for cyberspace. The major findings of the Task Force are as follows:

- *The era of the global internet is over.*
- *U.S. policies promoting an open, global internet have failed, and Washington will be unable to stop or reverse the trend toward fragmentation.*
- *Data is a source of geopolitical power and competition and is seen as central to economic and national security.*
- *The United States has taken itself out of the game on digital trade, and the continued failure to adopt comprehensive privacy and data protection rules at home undercuts Washington's ability to lead abroad.*
- *Increased digitization increases vulnerability, given that nearly every aspect of business and statecraft is exposed to disruption, theft, or manipulation.*
- *Most cyberattacks that violate sovereignty remain below the threshold for the use of force or armed attack. These breaches are generally used for espionage, political advantage, and international statecraft, with the most damaging attacks undermining trust and confidence in social, political, and economic institutions.*
- *Cybercrime is a national security risk, and ransomware attacks on hospitals, schools, businesses, and local governments should be seen as such.*
- *The United States can no longer treat cyber and information operations as two separate domains.*
- *Artificial intelligence (AI) and other new technologies will increase strategic instability.*
- *The United States has failed to impose sufficient costs on attackers.*
- *Norms are more useful in binding friends together than in constraining adversaries.*
- *Indictments and sanctions have been ineffective in stopping state-backed hackers.*

The Task Force proposes three pillars to a foreign policy that should guide Washington's adaptation to today's more complex, variegated, and dangerous cyber realm.

***First**, Washington should confront reality and consolidate a coalition of allies and friends around a vision of the internet that preserves—to the greatest degree possible—a trusted, protected international communication platform.*

***Second**, the United States should balance more targeted diplomatic and economic pressure on adversaries, as well as more disruptive cyber operations, with clear statements about self-imposed restraint on specific types of targets agreed to among U.S. allies.*

***Third**, the United States needs to put its own proverbial house in order. That requirement calls for Washington to link more cohesively its policy for digital competition with the broader enterprise of national security strategy.*

The Executive Summary finished by listing 16 major recommendations. Several stood out to me as being worthy of note:

- Agree to and adopt a shared policy on digital privacy that is interoperable with Europe's General Data Protection Regulation (GDPR).
- Declare norms against destructive attacks on election and financial systems.
- Negotiate with adversaries to establish limits on cyber operations directed at nuclear command, control, and communications (NC3) systems.
- Hold states accountable for malicious activity emanating from their territories.
- Clean up U.S. cyberspace by offering incentives for internet service providers (ISPs) and cloud providers to reduce malicious activity within their infrastructure.

This podcast is not the place to go more deeply into this report. But as I was scanning it and reading some of the many more interesting details, I kept thinking that our listeners would really find some of the report's detail interesting. So I've included the link to the 116-page report PDF in the show notes, and the PDF is this week's shortcut of the week: https://grc.sc/880

https://www.cfr.org/report/confronting-reality-in-cyberspace/download/pdf/2022-07/CFR_TFR80_Cyberspace_Full_SinglePages_06212022_Final.pdf

The Summary ends:

> *A free, global, and open internet was a worthy aspiration that helped guide U.S. policymakers for the internet's first thirty years. The internet as it exists today, however, demands a reconsideration of U.S. cyber and foreign policies to confront these new realities. The Task Force believes that U.S. goals moving forward will be more limited and thus more attainable, but the United States needs to act quickly to design strategies and tactics that can ameliorate an urgent threat.*

Wow.

Facebook has started encrypting its link URLs.

Yep. We were just talking last week about how Firefox v102 had added a feature to strip some of the tracking information from URLs it was querying before handing them over to a web server. It would do this, when enabled, for a small set of URLs that it recognized and felt conformable altering on the fly. And we noted that Firefox was apparently being conservative about what they were stripping since the Brave browser was reported to be significantly more aggressive.

While discussing this last week I commented that although I loved the idea of removing tracking identifiers from URLs, the whole thing felt flaky and unclean to me, since modifying a link's URL is inherently trouble prone and because it would be so easy for Facebook, for example, to change the token name of the value in the URL link. Then all browsers would need to update their URLexception handlers and we'd be back in a cat and mouse game.

Well, all of that hand wringing has been rendered moot because Facebook's links have suddenly transformed into opaque blobs. And this should not be a surprise to anyone. It should have been obvious to everyone that Facebook would not be happy having anyone mucking around with its URL links. The composition of any URL is entirely up to the creator of the URL. Back in 1994, RFC 1738, whose lead author was Tim Berners-Lee at CERN, made clear that a URL is inherently an opaque token that only needs to have any meaning to the server that receives it. Once upon a time, URLs tended to directly reflect the hierarchy of the receiving server's file system. And that file system was often organized by a human in some reasonable structure. But as pages became more and more dynamic, being assembled on-the-fly by server-side PHP, ASP or JSP code after querying a backend database, the primary reason URLs have remained at all understandable to humans is that they have been a source of signals for Internet indexing search engines. We'd like Google to learn something about a page's link from its textual content, so that's often been preserved. But we've increasingly seen URLs being cluttered up with things like GUIDs which only have meaning to server-side processes. Amazon's URLs have a short code near the front, surrounded by long hyphenated descriptive strings which describe the product. All of that superfluous text is only there for search engines to pick up on. Amazon has no need for it and completely ignores it. Since those massive multiline Amazon URLs are annoying to share, one of my favorite tricks is to strip everything out of an Amazon URL other than an anchoring /dp/ followed by the 10-character product ID. That results in a short URL that always works: https://www.amazon.com/dp/B01DYKSJL4

In any event, all of Facebook's content is obviously all being assembled on the fly, driven by code and a massive backend database. So the construction of their URLs is arbitrary and in no way reflects anything other than whatever their code wants. So Facebook apparently decided, for whatever reason — and it should come as no surprise to anyone — that it was tired of having 3rd-party browsers messing around with its links as an increasing number of browsers and anti-tracking privacy-enhancing add-ons had started doing. So now, no one who doesn't know how to unscramble or decrypt a Facebook link can see what's going on. They have truly become opaque blobs.

Since older pre-encrypted links are going to still be around, probably forever, I'm sure that all incoming links are checked to see whether they are old-style in-the-clear format or the new opaque-blob format. If they're old school, they're accepted as is. If they are obfuscated they're first decrypted then handled.

It's clear, as it always should have been, that any anti-tracking privacy enforcement we're going to obtain will need to be created by policy and not by technology.

Crack iOS 16's "Lockdown Mode", earn $2 million

As we discussed previously, Apple officially launched their very interesting new "Lockdown Mode" feature for iOS 16 during this year's WWDC. This idea makes so much sense because it's the insane "it'll do anything" breadth of features — many often unneeded, unwanted and unused — which, nevertheless, hugely increase any device's attack surface. So, simply turning off all of that unwanted and unneeded excess for individuals for whom security trumps the ability to receive cat videos from strangers, seems like an obvious win. When Leo and I were talking about this last week, it seemed like something we would both be included to at least take out for a spin, since I can easily live without receiving unsolicited cat videos. It will probably go a long way toward limiting the victimization by commercial malware such as “Pegasus.” And because Apple agrees with this, they’ve decided to put their money where their mouth is by offering the industry’s largest bounty ever — **$2 million** — to anyone able to reproducibly crack into an iOS 16 device when it’s in Lockdown Mode. I think that’s cool... and I’ll bet that a bounty of that size will likely give those who have enjoyed finding jailbreaks just for the fun of it some new incentive.

ClearView AI faces some new headwind

Get this: ClearView AI has now essentially been fined by Greece’s privacy authority, the Hellenic Data Protection Authority (HDPA) for violating parts of Europe’s infamous GDPR. The fine which has been levied is a hefty 20 million Euros and what's galling, even to me, is that it's not due to any use or abuse of ClearView AI's admittedly controversial facial recognition database technology. It's just because ClearView AI exists, Greece doesn't like the idea, and the GDPR gives them the right to fine.

A 22-page decision demands that Clearview AI stop processing biometric data on individuals in Greece and said the company must delete all the data it has already collected. The decision stems from a complaint filed by a number of privacy organizations which questions Clearview AI’s practice of scrapping selfies and photos from public social media accounts to assemble its facial recognition database which is rapidly growing toward 10 billion facial images. As we know since we've been tracking this interesting edge case since they emerged several years ago, ClearView AI sells its facial recognition tools to law enforcement agencies around the world and says it wants to reach 100 billion images in the coming years. It's also the case that ClearView has been at work in Ukraine, helping to identify both deceased Ukrainian citizens and Russian soldiers.

The problem that ClearView AI faces surrounds consent. More and more privacy regulations are requiring consent but ClearView's autonomous image scraping technology is inherently consent-free.

What I thought was interesting is that while Greece's Hellenic Data Protection Authority has levied this hefty fine, ClearView AI has never had any contact with either Greece's citizens or its law enforcement agencies. They simply share the same planet.

Clearview AI said it does not have a place of business in Greece or the EU and it does not have any customers in Greece or the EU. The company also claimed its product has never been used in Greece, and “does not undertake any activities that would otherwise mean it is subject to the GDPR.” One of the several privacy groups which filed the initial complaint explained that the fine and the ruling made clear that the GDPR is applicable because Clearview AI uses its software to monitor the behavior of people in Greece, even though the company is based in the U.S. and does not offer its services in Greece or the EU. The privacy organization said: “Collecting images for a biometric search engine is illegal.” (period.)

One thing that made me just shake my head is that ClearView has made it clear that they're happy to stay clear of regions that don't want their services. Yet the Greek authority also ordered Clearview to appoint a representative in the EU, to enable EU citizens to exercise their rights more easily and so regulators have a contact person in the EU.

I don't mean to sound overly sympathetic toward ClearView AI. But this does sort of seem to be a gray area. All of the images it's collecting are public. Anyone can view them. Just like the web pages that Google crawls across and indexes which allows us to later locate the information we seek. So it's clear that the difference is that pictures of people's faces are considered biometric data by these regulations and are not being regarded any differently than fingerprints or DNA.

If someone followed us around, dusting everything we touched to lift our fingerprints, that would likely annoy us. The fact that ClearView AI's image collection is unseen doesn't render it any less noxious in the eyes of privacy regulators.

One country another another is lowering the boom on ClearView AI. We previously talked about the UK's 7.5 million Euro fine in May, similar rulings have recently been made by France and Italy, and Austria is said to be preparing a similar ruling.

Ransomware gangs are getting into the searchable database game, too...

And speaking of searchable databases, several ransomware and extortion groups have started creating searchable databases of information stolen during attacks. As we know, it's not news that ransomware groups have been extorting organizations with the threat of leaking the data they've stolen. But that stolen data has typically been left sitting on the leak sites where it's effectively buried on the dark web.

However, over the last month, two ransomware groups AlphV, Karakurt and LockBit have debuted features on their leak sites which allow visitors to search through troves of data by company name or other signifiers.

A senior staff research engineer at Tenable has confirmed that all three groups have incorporated some kind of searchable database functionality into their leak sites. And if we've seen anything, it's that if an idea is useful it will be quickly picked up and mimicked by other groups.

Emsisoft's threat analyst Brett Callow said that the tactic was designed to further increase the pressure on organizations by weaponizing their customers and business partners. Callow said: “The gangs likely believe that making the data available in this way will result in more companies paying due to a perceived increase in the potential for reputational harm. And they may be right.” He added that in the past, companies have been able to dodge accountability for the leaks by claiming there is “no evidence user data has been misused” – which is a line seen in hundreds of breach notification letters over the last few years. Callow notes that such “Soothing statements like that aren’t really possible when people know their personal information was exfiltrated, compiled into an individual downloadable pack and made available online.”

Maybe Google will start indexing it!

Roskomnadzor strikes again!

Moscow has imposed a $358 million dollar fine on Google over Google's failure to filter out information from its search results that Russia's Internet watchdog “Roskomnadzor” has demanded be removed. I should note that the amount of the fine is much more fun when expressed in Russia's much less valuable Rubles. That would be 21 Billion (very small) rubles.

Anyway, Roskomnadzor announced that Google, and its subsidiary YouTube, have failed to remove the following materials multiple requests:

- Information about the course of the "special military operation" in Ukraine, which discredits the Armed Forces of the Russian Federation.
- Content promoting extremism and terrorism.
- Content promoting harmful acts for the life and health of minors.
- Information that promotes participation in unauthorized mass actions.

Ah, a single free and open Internet isn't always the best thing for everyone. I suppose this is what the Council on Foreign Relations meant when they said that the dream had not come true and the sooner we in the West — and the U.S. specifically — wake up and smell the packets, the better.

I guess Roskomnadzor realizes that Google is too useful to block outright, or they would have. They've tried over and over to enforce sanctions based on various parts of Russia's Code of Administrative Offenses. Last month, Roskomnadzor fined Google only $1.2 million dollars (a measly 68 million rubles). But as the fines remain unpaid, the multiple violations qualify for it to be based upon a piece of the action. In this case, up to 10% of Google's annual Russian revenue.

Russian users of Google Search and YouTube will also now encounter a warning about Google's violation of the law and they won't be allowed to place advertisements or use them as information sources.

So Russia is attempting to squeeze Google in the wallet. And, for what it's worth, it's working. Google's paid services are disappearing and being withdrawn. After Russia's invasion of Ukraine and the so-called “anti-fake news laws” which were enacted in the country (which amounted to don't say anything we don't like), Russia's Google subsidiary, Google LLC, filed for bankruptcy, claiming an inability to continue business after a series of massive fines and, ultimately, asset confiscation.

Loyal Russians presumably think: “Well, that's just those corrupt Westerners getting what they deserve.”

Last Tuesday's Patches

Speaking of getting what we deserve... last Tuesday Windows users received patches to (hopefully) fix a total of **84** individual flaws across Microsoft's sprawling software base, one being a true 0-day privilege of elevation bug which was being actively exploited in the wild.

The patches' demographic break down was:

- 52 Elevation of Privilege Vulnerabilities
- 12 Remote Code Execution Vulnerabilities
- 11 Information Disclosure Vulnerabilities
- 5 Denial of Service Vulnerabilities
- 4 Security Feature Bypass Vulnerabilities

There were no reports of big meltdowns following last Tuesday's updates, so nothing big and obvious was messed up this month. A handful of bugs are no more. Well... except for any new ones that may have been introduced. Maybe we'll get to those eventually.

SpinRite

A DM'd Tweet from yesterday contained a fun SpinRite testimonial:

Paul Jolley / @j0113y

> *Last week one of our power stations reported they needed to restore a GEM80 PLC. They had two separate backups on 3.5" floppy diskettes but neither would read. Configuration control (knowing what code is running on programmable devices performing process control in an OT environment) is very important in our industry so they were in a pickle. They tried a number of ways to read the floppies using various freeware and were unsuccessful. So I offered to try SpinRite as a last resort. I took delivery of the floppies this morning and set version v5.0 to work on Level 2, it managed to recover about 90% of the file required from the first floppy, then from the second floppy (which had a totally corrupt file system) I was able to "cat" the entire device in Linux to a file and subsequently extract the same file contents. Combining the recovered data from both floppies provided full coverage and thanks to you I was the hero.*

What was interesting was Paul's reference to SpinRite v5.0. As I've previously mentioned, for some confounding apparently mystical reason, v5.0 is superior to v6.0 for the recovery of diskettes. I've stared at its code for days and I have no idea why. But in testing both, v5.0 consistently produces superior results.

The other thing that was interesting was that being DM's, I had our previous DM thread. Back on February 13th of 2021, so about a year and a half ago, Paul DM'd to ask:

> *Happy to contact GRC support but thought I could quickly ask you first...I listened to a recent podcast where you said SpinRite 6.0 owners could download 5.0 if they want by simply changing the download url. I was interested because at our site we still use floppy diskettes 💾 so I looked up my purchase email and followed the link to the download page where it asks for my transaction code then generates links that don't have a version in the url. I must be missing something or didn't understand what you meant on the podcast!*

Paul later Tweeted:

> *Was just about to contact Greg this morning, when I found the answer was on the FAQ page at the bottom grc.com/sr/faq.htm#sr5*

So, I don't know whether he used SpinRite back then, or, it sounds more likely that he just wanted to be prepared for what happened yesterday.

Closing The Loop

Michael Swanson / @nexstarmike

Hi Steve - I just listened to SN 879 and regarding the use of a VPN when traveling (or even at a coffee shop), I prefer to use a "travel router" like the TP-Link N300. I connect the travel router to whatever Internet service is available and whatever devices I bring with me (laptop, tablet, phone, Roku, etc.) connect to the WiFi network of the travel router. All my devices are then behind a full NAT firewall. Added security - the travel router is using Google DNS to prevent DNS hijack and it is also possible to set the router to be a VPN client to many VPN services (and thus tunnel through to any VPN exit point including my home network if desired). And, the WiFi network on my travel router has the same SSID as my home network so all my devices connect automatically thinking they are at home.

IcyvRan TocVuc / @IcyvRan

Hi Steve, one solution if one does not trust a wifi hotspot, is setting up a Raspberry Pi at home with Tailscale, and configuring it as an "exit node" https://tailscale.com/kb/1103/exit-nodes/

Tailscale:

*"Exit nodes" capture **all** your network traffic, which is typically not what you want. The exit node feature lets you route all non-Tailscale internet traffic through a specific device on your network. The device routing your traffic is called an "exit node."*

By default, Tailscale acts as an overlay network: it only routes traffic between devices running Tailscale, but doesn't touch your public internet traffic, such as when you visit Google or Twitter. This is ideal for most people, who need secure communication between sensitive devices (company servers, home computers), but don't need extra layers of encryption or latency for their public internet connection.

*However, there may be times when you **do** want Tailscale to route your public internet traffic: in a cafe with untrusted Wi-Fi, or when traveling overseas and needing access to an online service (such as banking) only available in your home country. By setting a device on your network as an exit node, you can use it to route all your public internet traffic as needed, like a consumer VPN.*

Taskel @Taskel7

FYI on the Quantum-Resistant algorithms CRYSTALS-Kyber and CRYSTALS-Dilithium. Kyber crystals are what is used in lightsabers in Star Wars so there was something for Star Trek and Star Wars fans there.

Lethal Dosage / @LethalDosage1

> *I logged into twitter for the first time in 4 year to poke fun at you! You are losing geek points! The first Star Wars movie was not episode IV a new hope, it was "Star Wars". "Episode IV A New Hope" was added later. Watch the original intro, only 2 minutes long:*
> https://www.youtube.com/watch?v=iXDnFYu91vY
> *"Star Wars (1977) original opening crawl" (1,431,123 views)*

That all look authentic. But it's old and grainy as hell and in this day and age it could easily have been edited. So I did digging around the Internet and got the whole story:

In the beginning, there was just "Star Wars." But then, fans of the most popular science fiction movie of all time were thrown a hyperspace curveball: The film known as just "Star Wars" wasn't the beginning of a story, it was the middle. And four years after the original film hit theaters, it was re-released, this time being called "Star Wars Episode IV: A New Hope." Here's what happened:

In March of 1978, science fiction author Leigh Brackett dies, and George Lucas takes over the writing of movie #2, "The Empire Strikes Back", a task which he shares with Lawrence Kasdan.

Next, Lucas decides that there's a bigger backstory to all of Star Wars, which means that Empire isn't "part two," but instead, "" So, in 1980, "The Empire Strikes Back" identifies itself as the confusing "Episode V" — which totally blew everyone's mind at the time and resulted in no end of confusion. Then in 1981, Star Wars is re-released as Episode IV to make everything line up.

What's confusing about all this is that I definitely saw the original Star Wars in 1977. Having been born in 1955 I was 22... and I recall that afternoon, 45 years ago, sitting in the theater. And I distinctly recall anxiety and consternation being created by Star Wars' episode numbering. But I guess the anxiety must have been created when "The Empire Strikes Back" identified itself as Episode V, rather than the original Star Wars identifying itself from the start as Episode IV.

davepope / @davepope

> *FYI, my 2013 Ford key fob has bidirectional comms. It has a light that shows me red or green if the remote start was successful or not. No idea if it does the handshake you mention in the episode though*

RetBleed

Well, Jason, you seem to be getting the “Bleeding” podcasts. Three weeks ago, with you, we tackled the “HertzBleed” attack. And today we’re back with “RetBleed.”

“RET” (short for “return”) is the universal name of the CPU instruction that’s placed at the end of a subroutine to cause the subroutine to terminate its execution and for control to be returned to the instruction following the one that invoked the subroutine. In essence, the instruction tells the CPU to “return” to the point where the subroutine was called. In stack-based processors, subroutines are often provided with parameters which they use for their work. These values or pointers will be placed onto the stack for the subroutine’s use before the subroutine is called. And subroutines may place some of their own local temporary data onto the stack as well. (How many times has this podcast used the term “stack buffer overflow?”) And when the processor’s return instruction is executed, all of this stack-based data will typically be discarded, simply by being.

RetBleed is the brainchild of two researchers from ETH Zurich, who have been behind a number of previous clever attacks. Their paper on RetBleed will be delivered a few weeks from now in a Technical Session of the USENIX Security '22 conference. They Responsibly disclosed their discovery to Intel and AMD back in February of this year, and it emerged from embargo last Tuesday the 12th of July.

I’m going to start by just reading their paper’s Abstract so that we can get an overall feel for what this is, then we’ll break it down:

> *Modern operating systems rely on software defenses against hardware attacks. These defenses are, however, as good as the assumptions they make on the underlying hardware. In this paper, we invalidate some of the key assumptions behind retpoline, a widely deployed mitigation against Spectre Branch Target Injection (BTI) that converts vulnerable indirect branches to protected returns. We present RETBLEED, a new Spectre-BTI attack that leaks arbitrary kernel memory on fully patched Intel and AMD systems. Two insights make RETBLEED possible: first, we show that return instructions behave like indirect branches under certain microarchitecture-dependent conditions, which we reverse engineer. Our dynamic analysis framework discovers many exploitable return instructions inside the Linux kernel, reachable through unprivileged system calls. Second, we show how an unprivileged attacker can arbitrarily control the predicted target of such return instructions by branching into kernel memory. RETBLEED leaks privileged memory at the rate of 219 bytes/s* [with 98% accuracy] *on Intel Coffee Lake and 3900 bytes/s* [with >99% accuracy] *on AMD Zen 2.*

There are a few things to observe here. One is that this is another instance of the lesson that attacks never get worse, they only ever get better. When we started off with the Spectre and Meltdown speculative execution attacks, they were purely theoretical. But they didn’t remain that way for long. Before long researchers were discovering how to use these attacks to probe the contents of memory that they had absolutely no access to.

Essentially, they deliberately created a road that would not be taken, but which the CPU would speculatively prepare to take anyway. In doing so it would preload the contents of some memory into its cache. Then they would probe the cache to see what the CPU had cached in preparation for that untaken road. In this manner, they would get the CPU to access memory they could not legally access. Access violations were never triggered because speculation never triggers access violations. This all amounted to some very clever manipulations of the insanely complex micro-architectures that have been incrementally added, generation after generation, to modern processors, all in the name of squeezing out every last cycle of performance.

The problem that's the subject of this paper, and of much sudden scurrying around (for example, Linus has just delayed the next Linux kernel release by one week due to this) has the name "branch target injection" it's also known as Spectre variant 2. There are essentially two available mitigations for this sort of speculation side-channel leakage: "Retpoline" (which is a contraction of return and trampoline) and IBRS which stands for "Indirect Branch Restricted Speculation."

Just over three years ago, the SUSE Linux blog posted an article titled "Removal of IBRS mitigation for Spectre Variant2" and what was written then is interesting in light of today's events:

> As the Meltdown and Spectre attacks were published at the beginning of January 2018, several mitigations were planned and implemented for Spectre Variant 2. Spectre Variant 2 describes an issue where the CPUs branch prediction can be poisoned, so the CPU speculatively executes code it usually would never try to. For instance userspace (attacker controlled) code could make the kernel code speculatively execute Spectre code gadgets that disclose secret kernel information, via Flush+Reload disclosure methods.
>
> Two major mitigations were proposed:
>
> A CPU feature called "Indirect Branch Restricted Speculation" (IBRS) that would not use branch predictions from lower privilege levels to higher ones. Or, software workarounds called "retpolines" and "RSB stuffing". These can fully replace the IBRS mitigation. On Intel Skylake there is the theoretical possibility that these software mitigations are not sufficient, but so far research has not shown any holes. [That was true three years ago, but as we know today, it is no longer true.]
>
> SUSE backported the IBRS patches to our kernels for the initial release of mitigations and enabled them, as the "retpoline" mitigations were not yet ready. SUSE pushed the "retpoline" mitigation some months later after support in the compiler and kernel became available, but left in the IBRS mitigation. As of today [again, this was three years ago], the "retpoline" and "RSB stuffing" software workarounds provide the same level of mitigations that IBRS provides. While IBRS support continued in the SUSE kernel, it was not accepted by the Linux upstream kernel community, and it was also shown to cause performance degradation.
>
> As "retpoline" and "RSB stuffing" completely mitigate the Spectre Variant 2 issue for the Linux Kernel, SUSE decided, with guidance from Intel, to remove the IBRS patches from our kernel releases. While on Intel Skylake there exists a theoretical possibility that the software mitigations are not complete, so far no research has shown exploitable scenarios. Should research show any exploitable scenarios there, SUSE will re-enable the IBRS mitigation on these chipsets.

So now that research **has** shown exploitable scenarios, I'm sure that's what they've been doing. This means that the clever "no hardware required" Retpoline hack that Google originally invented to protect their Chromium browser core from these attacks worked for about three years until enough time, focus and reverse engineering was applied by some dedicated researchers to hack past the imperfect mitigation and turn a theoretical vulnerability into a very real threat.

Meanwhile, the day before yesterday, on Sunday, Linus posted into the Linux Kernel 5.19-rc7 thread. He wrote:

> It's a Sunday afternoon, I wonder what that might mean..
>
> Another week, another rc. We obviously had that whole "Retbleed" thing, and it does show up in both the diffstat and the shortlog, and rc7 is definitely bigger than usual.
>
> And also as usual, when we've had one of those embargoed hw issues pending, the patches didn't get the open development, and then as a result missed all the usual sanity checking by all the automation build and test infrastructure we have. So no surprise - there's been various small fixup patches afterwards too for some corner cases.
>
> That said, last week there were two other development trees that independently also asked for an extension, so 5.19 will be one of those releases that have an additional rc8 next weekend before the final release. We had some last-minute btrfs reverts, and there's also a pending issue with an Intel GPU firmware.
>
> When it rains it pours.
>
> Not that things really look all that bad. I think we've got the retbleed fallout all handled (knock wood), and the btrfs reverts are in place. And the Intel GPU firmware issue seems to have a patch pending too (or we'll just revert). So it's not like we have any huge issues, but an extra week is most definitely called for.
>
> Linus

I loved Intel's description of this problem, CVE-2022-29901. It starts out *"Non-transparent sharing"* ya gotta love that. Somewhere in their technical press-release department, someone called out: "Hey anyone!, I need a term for "Leakage" that doesn't sound like a bad thing. And someone replied: How about "non-transparent sharing." The writer said "Perfect!", returned to his keyboard and wrote: *"Non-transparent sharing of branch predictor targets between contexts in some Intel processors may allow an authorized user to potentially enable information disclosure via local access."*

The good news is, not all processors will be affected. The ETH Zurich researchers said they tested the Retbleed attack in practice on AMD Zen 1, Zen 1+, Zen 2, and Intel Core generation 6–8. This essentially means Intel CPUs from 3 to 6 years old, and AMD processors from 1 to 11 years will likely be affected.

Fortunately, the industry is getting better about addressing these sorts of problems and patches for Retbleed were incorporated into this month's Patch Tuesday in both OS & cloud infrastructure updates from all the major providers.

This leaves us with the performance hit that comes with disabling performance-enhancing but inherently exploitable features. We've talked about this from the first glimmer of the first of these many micro-architectural side channel vulnerabilities. Since all of these fancy features were invented to speed up the execution of real world code, taking them out or shutting them down means some performance loss.

The ETH researchers noted that installing these patches will have an impact on the CPU's performance metrics on affected processors between 14% and 39%. And another issue they found in AMD processors that they named Phantom JMPs (CVE-2022-23825) might even come with a 209% performance overhead.

The ETH researchers concluded their paper by writing:

> *We showed how return instructions can be hijacked to achieve arbitrary speculative code execution under certain microarchitecture-dependent conditions. We learned these conditions by reverse engineering the previously-unknown details of indirect branch prediction on Intel and AMD microarchitectures and its interaction with the Return Stack Buffer. We found many vulnerable returns under these conditions, using a new dynamic analysis framework which we built on top of standard Linux kernel testing and debugging facilities. Furthermore, we showed that an unprivileged process can control the destination of these kernel returns by poisoning the Branch Target Buffer using invalid architectural page faults. Based on these insights, our end-to-end exploit, RETBLEED, can leak arbitrary kernel data as an unprivileged process running on a system with the latest Linux kernel with all deployed mitigations enabled. Our efforts led to deployed mitigations against RETBLEED in the Linux kernel.*