Going Dark:
catastrophic security and privacy losses due to loss of visibility by managed private networks

Paul Vixie, AWS Security (*)

NANOG, October 2022

(*) Views expressed here are mine alone

Abstract

- Effective modern site security is behavioral in nature. We cannot choose or exclude our endpoints nor validate their supply chains, and so to the extent that we manage digital risks posed by our endpoints we do it by watching the signals (packets and flows) they emit. Such observations are categorically untenable for investigative journalists and dissidents since the category is occupied by corrupt or authoritarian regimes or their national security apparatus -- as explained by E. Snowden in 2013 and as codified by the IETF in RFC 7258.
- Using the same protocols for mobile devices which accounted for most human-centric endpoint growth since 2010 as we do for fixed devices on networks controlled by families and businesses is disrupting our limited ability to secure the latter in order to defend against worst-case outcomes for the former. Several decades of unapologetic abuse by the powerful have led the IETF to reform the basic Internet protocol suite around TLS 1.3 with Encrypted Client Hello, DNS over HTTPS, and the replacement of TCP by the UDP-based QUIC protocol.
- In this new configuration, network operators will not be able to detect endpoint behavior changes corresponding to infection, takeover, poisoned software update, latent design dangers, predaceous grooming, insider corruption, or hundreds of other well-understood digital harms. Many such operators have not been warned about this "rules change" and deserve to have their expectations explicitly and immediately reset so that they can make new plans which will be practical in the next era. It is the goal of this presentation to enumerate those alarms.

Effective Modern Site Security

- Near oxymoronic – effective ≠ modern, for site security
- What there is, is behavioural, not pattern or privilege based
- Behavioural means we watch the signals (packets and flows) emitted
- Rule or law breaking is a local matter yet the Internet is global
- “Behavioural” is simultaneously too little and too much security

How Did Site Security Become Ineffective?

- Unbounded complexity
  - Most know little, few know much, nobody knows all
  - Hardware, software, protocols, configurations, in a virtual blender
  - Extreme churn in vendors, versions, patches, personnel, policies
- Confidence in safety is probabilistic
  - Backups, logging, verification, assurance – all expensive to do “well enough”
  - Many sites assume, and some sites know, that they are already breached
- Asymmetry of incentives
  - Attacks are a profit center whereas defense is a cost center

What's Happening Now (2013—2022)

- In 2013, E. Snowden famously traveled to Hong Kong and made some important disclosures, before traveling onward to Moscow
- The Internet Engineering Task Force (IETF) invited E. Snowden to give a plenary speech, which resulted in two Requests for Comment (RFCs)
  - RFC 7258 – Pervasive Monitoring Is An Attack
  - RFC 8890 – The Internet Is For End Users
- No carve outs were made for monitoring in the service of site security
- No distinction was made between end users, intruders, or "insiders"
- No awareness was shown of surveillance capitalism or malware

Simplicity Did Not Last But Serves Us Poorly

- Early Internet used then-radical informalism to create protocols
  - All any protocol had to do was mostly work and reach production quickly
  - This preceded the definition of “open source” but is fruit of the same tree
- Protocols for client ↔ server were the same as for server ↔ server
  - Problematic for DNS, SMTP, HTTP, and probably others
  - Nonrecognition of firewalls, NAT, and ALG makes this hard to fix
- Creeping featurism and complexity has amplified a mess at scale
  - The only good thing to say about it is that nothing else could have worked

Defense Now Requires Rule Breakage

- An endpoint operating system that does not trust its apps or its users has to allowlist, or denylist, or pervasively monitor
  - Same for a site security administrator
  - Same for an authoritarian government
- A passionate defender of human freedom in the face of wide spread abuse of end-user privacy by ISPs and governments must seek to disintermediate same
  - So: signal entropy must be maximized (i.e., “encrypt everything”)

On Information Entropy

If W is the number of microstates that can yield a given macrostate, and each microstate has the same a priori probability, then that probability is p = 1/W.

…

Some authors argue for dropping the word entropy for the H function of information theory and using Shannon's other term, "uncertainty", instead.

(Wikipedia)

RFC 8484 Can Speak For Itself

RFC 8484 DNS Queries over HTTPS (DoH) October 2018

…

1. Introduction

…

Two primary use cases were considered during this protocol's development. These use cases are *preventing on-path devices from interfering with DNS operations*, and also allowing web applications to access DNS information via existing browser APIs in a safe way consistent with Cross Origin Resource Sharing (CORS) [FETCH].

“What Is QUIC Manageability?”

3.1. Identifying QUIC Traffic

The QUIC wire image is not specifically designed to be distinguishable from other UDP traffic.

What does tcpdump see?

00:20:49.818784 IP6 2001:559:8000:ca::41.12684 > 2001:559:8000:cd::5.22:
Flags [S], seq 901949512, win 65535,
options [mss 1440,nop,wscale 6,sackOK,TS val 3799452908 ecr 0], length 0

00:20:49.819094 IP6 2001:559:8000:cd::5.22 > 2001:559:8000:ca::41.12684:
Flags [S.], seq 1775462878, ack 901949513, win 65535,
options [mss 1440,nop,wscale 6,sackOK,TS val 4233103442 ecr 3799452908], length 0

See Also Encrypted Client Hello (ECH)

- Transport Layer Security (TLS) is at the heart of HTTPS and similar
- TLS 1.2 is in near universal use, and works with next-gen firewalls
  - TLS 1.3 is now in final clearing stages, and won't
- The specific feature of interest is Encrypted Client Hello (ECH)
  - Was Encrypted Server Name Indicator (ESNI) but morphed
- A next-gen firewall will know the destinations' IP but not its name
  - Most destination IP's serve more than one ("many") names
  - It is the name that is of interest to a firewall operator

Ugly Choices for Site Security Administrators

- Network traffic is “going dark” due to political forces
  - There will be no good way to detect intrusion, exfiltration, malware, spam, DDoS, predacious grooming, or insider corruption
- Many sites cannot afford this fight, and will let the floodgates open
- Some sites cannot afford not to fight, and will take expensive action
  - For example, enforce the use of ALG (proxy) for all trans-gateway flows
  - Or perhaps, block all trans-gateway UDP, denylist known DoH, block TLS 1.3
  - Innovation (and chaos) can be expected here

A Possible Way Forward (“Works in Theory”)

- Let a network express its policy, reliably and securely, to endpoints
  - Allow cooperation, which might entail content transparency
  - Deny noncooperation, giving a clear choice to end users & apps
- This would mean endpoint <-> gateway protocols would be different from gateway <-> server protocols, as is already true of SMTP(S)
  - Will be seen by some as bad engineering
  - Will be seen by others as bad humanity
- Won’t work in practice (requires alignment of interests)

Impact on Malware Reversing / Analysis

- Malware, as an endpoint, now has rights
  - Can forbid monitoring, tracking, and interference (don’t fall back)
  - Defenders will have difficulty learning (or using) signalprints
- On the bright side, firewalls who block this stuff receive a boon
  - Managed private network looks to malware like a reverse engineer
  - Malware may not be willing to take that chance

When Endpoints Have Demands