opsec
Internet-Draft
Intended status: Informational
Expires: 27 July 2024

A. J. Campling
419 Consulting Limited
P. Vixie
Red Barn
D. Wright
UK Safer Internet Centre
A. Taddei
S. Edwards
Broadcom

24 January 2024

Encrypted Client Hello Deployment Considerations
draft-campling-ech-deployment-considerations-08

Abstract

(Editorial note: to be updated as the text in the main body of the document is finalised) This document is intended to inform the community about the impact of the deployment of the proposed Encrypted Client Hello (ECH) standard that encrypts Server Name Indication (SNI) and other data. Data encapsulated by ECH (ie data included in the encrypted ClientHelloInner) is of legitimate interest to on-path security actors including those providing inline malware detection, parental controls, content filtering to prevent access to malware and other risky traffic, mandatory security controls etc.

The document includes observations on current use cases for SNI data in a variety of contexts. It highlights how the use of that data is important to the operators of both public and private networks and shows how the loss of access to SNI data will cause difficulties in the provision of a range of services to end-users, including the potential weakening of cybersecurity defences. Some mitigations are identified that may be useful for inclusion by those considering the adoption of support for ECH in their software.

Table of Contents

1. Introduction

1.1. Background

In order to establish its handshake, the TLS protocol needs to start with a first handshake message called the Client Hello. As this handshake message is in clear text, it exposes metadata, e.g. the Server Name Indication (SNI) which allow middleboxes on path to make policy decisions, in particular but not only for security reasons. As part of a wider initiative to achieve end-to-end encryption, a proposed extension to TLS 1.3 called Encrypted Client Hello (ECH) [I-D.draft-ietf-tls-esni] is attempting to encrypt all the remaining metadata in the clear.

There are use cases where encryption of the SNI data may be a useful precaution to reduce the risk of pervasive monitoring that offers some benefits (e.g Enterprises offering services for their own customers will appreciate that their customers' privacy be better protected). However ECH presents challenges for other use cases (e.g. Enterprises needing network security controls for compliance reasons).

The Internet was envisaged as a network of networks, each able to determine what data to transmit and receive from their peers. Developments like ECH mark a fundamental change in the architecture of the Internet, allowing opaque paths to be established from endpoints to commercial services, some potentially without the

knowledge or permission of the device owners. This change should not be undertaken lightly given both the architectural impact on the Internet and potentially adverse security implications for end users. Given these implications, it certainly should not be undertaken without either the knowledge of or consultation with end users, as outlined in [RFC8890].

Whilst it is reasonable to counter that VPNs also establish opaque paths, a primary difference is that the use of a VPN is a deliberate act by the user, rather than a choice made by client software, potentially without either the knowledge and/or consent of the end-user or device owner.

[RFC7258] discusses the critical need to protect users' privacy when developing IETF specifications and also recognises that making networks unmanageable to mitigate pervasive monitoring is not an acceptable outcome.

[RFC8404] discusses current security and network operations as well as management practices that may be impacted by the shift to increased use of encryption to help guide protocol development in support of manageable and secure networks. As [RFC8404] notes, "the implications for enterprises that own the data on their networks or that have explicit agreements that permit the monitoring of user traffic are very different from those for service providers who may be accessing content in a way that violates privacy considerations".

1.2. Scope, objectives and limits of this document

This document considers the implications of ECH for private, edge and public networks using the examples of education establishments, enterprises and public operators. It addresses the limitations of [RFC8744] by providing more information about the issues posed by the introduction of ECH due to the loss of visibility of SNI data on private networks building on the report from a roundtable discussion [ECH_Roundtable].

The objective of this document is to detail some operational impacts of ECH. It will focus specifically on the impact of encrypting the SNI data by ECH, but it should be noted that other elements in the client hello may also be relevant for some on-path security methods.

The data encapsulated by ECH is of legitimate interest to on-path security actors including those providing inline malware detection, firewalls, parental controls, content filtering to prevent access to malware and other risky traffic, mandatory security controls (e.g. Data Loss Prevention) etc. Beyond network security, there are various operational impacts of different types e.g. network management, content filtering, etc.

Whilst this document identifies operational issues:

* it does not consider solutions nor question the development of the ECH proposal itself

* it doesn't attempt to be exhaustive,

* it will start by focusing on one category of middleboxes: proxies.

2. General considerations about the encryption of the Client Hello

2.1. About encrypting the Server Name Indication (SNI)

[RFC8744] describes the general problem of encrypting the Server Name Identification (SNI) TLS extension. The document includes a brief description of what it characterises as "unanticipated" usage of SNI information (section 2.1) as well as a brief (two paragraph) assessment of alternative options in the event that the SNI data is encrypted (section 2.3).

The text in [RFC8744] suggests that most of the unanticipated SNI usage "could also be implemented by monitoring DNS traffic or controlling DNS usage", although it does then acknowledge the difficulties posed by encrypted DNS protocols. It asserts, with limited evidence, that "most of 'the unanticipated usage' functions can, however, be realized by other means", although without considering or quantifying the affordability, operational complexity, technical capability of affected parties or privacy implications that might be involved. It is unclear from the document whether any stakeholders that may be impacted by the encryption of SNI data have been consulted; it certainly does not appear to be the case that any such consultation has taken place.

The characterisation of "unanticipated usage" of SNI data could be taken to imply that such usage was not approved and therefore inappropriate in some manner. The reality is that the development of the Internet has many examples of permissionless innovation and so this "unanticipated usage" of SNI data should not be dismissed as lacking in either importance or validity.

2.2. Why are middleboxes using the SNI?

For middleboxes to be able to perform their job they need to identify the destination of the requested communication. Before TLS1.3 a middlebox could rely on, at least, 3 metadata sources: The certificate, the DNS name and the SNI. A middlebox may have used some or all of these metadata to determine the destination in the best possible way. Yet, as part of the current initiative to complete end-to-end encryption, the certificate was encrypted into TLS1.3, then DoH/DoT/DoQ are encrypting the DNS flow to its resolver making it harder for middleboxes to use this information. Even if the DNS data can be accessed, it can be misleading in some situations (does it point to the real destination, or just the site hosting

server name, or a proxy?) and the SNI was invented to overcome some of the limitations of the DNS data by providing additional information. However, the SNI in itself may be unreliable which is why middleboxes start by non-trusting it until they have validated the information that it provides.

2.3. The case of Proxies

A proxy server is a server application that acts as an intermediary between a client requesting a resource and the server providing that resource. Instead of connecting directly, the client directs the request to the proxy server which evaluates the request before performing the required network activity. Proxies are used for various purposes including load balancing, privacy and security.

Proxies can be used explicity or transparently.

In explicit proxy model, proxies are accessed by configuring a user's application or network settings, with traffic diverted to the proxy rather than the target destination.

With "transparent" proxying, the proxy intercepts packets directed to the destination, making it seem as though the request is handled by the target destination itself.

A key advantage of transparent proxies is that they work without requiring the configuration of user devices or software. They are commonly used by organisations to provide content filtering for devices that they don't own that are connected to their networks. For example, some education environments use transparent proxies to implement support for “bring your own device” (BYOD) without needing to load software on third- party devices.

Transparent proxies use SNI data to understand whether a user is accessing inappropriate content without the need to inspect data beyond the SNI field. Because of this, encryption of the SNI field, as is the case with ECH, will disrupt the use of transparent proxies, requiring far more intrusive data inspection to be undertaken instead.

2.4. Why relying on the SNI and not the DNS

As per the Introduction section of [RFC8744] "More and more services are colocated on multiplexed servers," the SNI was introduced to allow "direct connections to the appropriate service implementation".

By design, a proxy cannot rely only on the DNS to ensure establishing the connection, the SNI is simply required by design.

2.5. About the unreliability of the SNI

SNI is by design not reliable therefore prompting the question for

why and how middelboxes are using it in the first place. Before anything, in Security, in general, unreliability is a useful source of information in itself.

Referring to [RFC6066], TLS extenisons, including SNI, are designed to be backward compatible. This means that if the server doesn't recognize the SNI value, the TLS handshake should continue anyway.

In other terms, SNI value can 1) be empty or 2) have an alternative name which is different from the real name of the destination server without impacting the establishment of the TLS session. This can also be easily exploited by bad actors indeed to bypass security middle boxes. E.g A malware could just be coded to provide an SNI value that is mapped to finance or healthcare categories to bypass inspection (it is not that easy, but there is a way to do it).

2.5.1. With TLS 1.2

User client generates “ClientHello” with SNI in plain text to the destination server. If accepted, the server responds to the user client request with the “ServerHello” message containing the intended server certificate alongside other encryption-related information in the plain text mode as well. Middleboxes can then see/inspect 1/ the SNI in the “ClientHello” and 2/ the server certificate details in the “ServerHello” message. Middleboxes can perform selective inspection based on the destination service details (service requested by the user client) extracted from the server certificate. I would like to note that depending on the middlebox design, it is more about a correlation of different source of information to confirm the right destination service. But still only information within the server certificate is reliable to perform accurate web categorization and then do selective inspection of any kind of web/content filtering.

2.5.2. With TLS 1.3

TLS 1.3 improves significantly over the TLS 1.2 in terms of security and privacy aspects. More specifically, in terms of privacy, it overcomes the plain text server certificate exchange issue by masking the server’s host identity through the encrypted server certificate. So as middleboxes inspection capabilities are designed based on the server certificate, all vendors worked on how to adapt their capabilities to support TLS 1.3 (e.g ServerHello encryption but not only).

So how do providers support TLS 1.3 inspection then?

The most common technique is first to get the SNI from the ClientHello (which is still shared in plain text format.) and then replays / establish a new full TLS session initiated from the proxy server itself to the destination server to retrieve the server certificate details before determining the web category; Once

identified, selective inspection can be performed on the real TLS session initiated by the user client.

As the SNI is not reliable, proxies accept the SNI asis but do without trusting it, then they perform checks at various level to verify this SNI and they step by step enrich the evaluation, therefore bringing more possibilites to interpret which policy to apply. This could end up with the proxy deciding to block the connection, or the proxy to let the connection happened with a verified or corrected SNI.

See Appendix for some initial data gathering on the reliability of the SNI.

2.5.3. With TLS 1.3 with ECH extension

The entire legacy ClientHello message (Inner ClientHello) is encrypted, encapsulated and sent as part of the new ClientHello wrapper message (Outer ClientHello); So middleboxes cannot identity the destination service anymore and cannot replay the TLS session to the destination server because it has no clue about it. (May be DNS knows but not all the time, see other issue here created by myself).

So TLS ECH has an impact on enterprise security and compliance (including selective inspection) not because of SNI (which is not reliable) encryption but because there is no other information about the destination server is available and that can be used to fetch and retrieve the server certificate (Which is indeed reliable) to perform web categorization.

We don’t care about SNI and its reliability, we need just to know the destination service to get the destination server certificate details.

3. The Education Sector

3.1. Context

Focusing specifically on the education sector, the primary issue caused by ECH is that it is likely to circumvent the safeguards applied to protect children through content filtering, whether in the school or home environments, adding to adverse impacts already introduced through the use of encrypted DNS protocols such as DNS over HTTPS [RFC8484].

Content filtering that leverages SNI information is used by education establishments to protect children from exposure to malicious, adult, extremist and other content that is deemed either age-inappropriate or unsuitable for other reasons. Any bypassing of content filtering by client software on devices will be problematic and may compromise duties placed on education establishments. For example: schools in England and Wales have obligations to provide "appropriate filtering

systems" [KCSE]; schools in the US use Internet filters and implement other measures to protect children from harmful online content as a condition for the receipt of certain federal funding, especially E-rate funds [CIPA].

3.2. Why Content Filtering Matters to Schools

The impact that ineffective content filtering can have on an educational institutions should not be underestimated. For example, a coroner in the UK in 2021 ruled that a school's failure to prevent a pupil from accessing harmful material online on its equipment contributed to her taking her own life [Coroner]. In this particular instance, the filtering software installed at the school was either faulty or incorrectly configured but the case highlights the harmful risks posed if the filtering is bypassed by client software using ECH.

3.3. Mitigations

Whilst it may be possible for schools to overcome some of the issues ECH raises by adopting similar controls to those used by enterprises, it should be noted that most schools have a very different budget for IT compared to enterprises and usually have very limited technical support capabilities. Therefore, even where technical solutions exist that may allow them to continue to meet their compliance obligations, affordability and operational expertise will present them with significant difficulties.

Absent funding and technical expertise, schools will need to consider the best way forward that allows them to remain compliant. If client software does not allow ECH to be disabled, any such software that implements support for ECH may need to be removed from school devices and replaced, assuming that suitable alternatives are available. This will have a negative impact on budgets and may be operationally challenging if institutions have made a significant investment in the deployment and use of particular applications and technologies.

There are instances where policies in education establishments allow for the use of equipment not owned by the institution, including personal devices and the devices of contractors and site visitors. These devices are unlikely to be configured to use the institution's proxy but can nevertheless connect to the school network using a transparent proxy (see below). Transparent proxies used for filtering will typically use SNI data to understand whether a user is accessing inappropriate data, so encrypting the SNI field will disrupt the use of these transparent proxies.

3.4. Implications

In the event that transparent proxies are no longer effective, institutions will either have to require more invasive software to be

installed on third party devices before they can be used along with ensuring they have the capability to comprehend and adequately manage these technologies or will have to prevent those devices from operating. Neither option is desirable.

4. Impact of ECH in private network contexts (Enterprises or other organizations)

4.1. Context

4.1.1. The main requirements

Enterprises and Organizations need to protect themselves for a vast number of reasons, mainly:

* Reduce their Risks. And in particular as part of any Cyber Resilience strategy.

* Protect their Reputation. The term Reputation includes many aspects way beyond the traditional enterprises and organization assets (data, etc.).

* Comply to a growing diverse set of Policies, Regulations, Certifications, Labeling and Guidelines. These requirements are growing in both scope and complexity as they are added to by various national and regional bodies around the world.

4.1.2. A degrading threat landscape

In addition, the general threat landscape which was already very large (see [I-D.draft-mcfadden-smart-threat-changes]), has significantly increased in three ways:

* COVID crisis generally accelerated the overall attack landscape. Indeed as the crisis forced many enterprises and organizations to accelerate their digital transformation, it increased the opportunity for cyber criminals and nation states to launch more attacks, leverage innovations to their advantage, better select their targets, increase their efficiency and increase their rewards, in particular with Ransomware based attacks.

* The Supply Chain is under stress as per the [SOLARWIND] attack

* Nation State attacks are continuing to evolve, for example as noted to those linked to the current Ukraine crisis.

Attacks are now damaging enterprises and other organizations with ransomware being the number 1 issue by a considerable margin. The attacks are increasing in severity, to the extent that this is now being measured at macroscopic level in some countries:

* EUR1B loss of revenue for French organizations from January to

August 2022 [LOSSINREVENUE]

* Loss in capitalisation between 1-5% [LOSSINCAP]

* Degradation by credit notation agencies [LOSSINCREDITSCORE]

Another implication from the COVID crisis is the acceleration of BYOD with the current reliance on remote working. This has created two side effects for remote employees, contractors and third parties that need to connect to one or more enterprise networks on a temporary basis:

* need to use a VPN access to the corporate network, which brings all the benefits (e.g. protected access to corporate network) and risks that VPNs may open (e.g. lateral movement when the end point is compromised),

* need to access a cloud proxy which requires an agent to be installed on the device to steer the traffic to the right place.

4.2. Additional considerations

In such circumstances, requiring software or custom configurations to be installed on those devices may be problematic (see [I-D.draft-taddei-smart-cless-introduction]).

This is why network security solutions are required and this is why the use of ECH to prevent access to the SNI data makes it impossible for blue teams to defend (see the next sections for details).

Finally there is a global shortage of cybersecurity personnel. Any expansion of technical requirements, for example to mitigate the operational challenges through the introduction of ECH, will exacerbate the problem.

All the above conditions are weighing on capabilities to defend, both:

* Directly: a lack of visibility on a key meta data like the SNI will cause significant issues to enterprises and organizations

* Indirectly: should ECH happen and should alternative be provided, managing migrations to any alternative not requiring access to the SNI, in these conditions, is undesirable from a timing, resources, capacities and risks perspectives.

4.3. Implications

4.3.1. Examples of regulatory implications

Regulators are accelerating their lawfare capabilities at accelerated pace and new legislations is impacting on the actions of enterprises

with increased precision. The EU GDPR had ripple effects such as requiring Financial Institutions to use selective decrypt in order to implement Data Loss Prevention. The recent indication that US regulators are in the process of levying fines of $200m each on a number of institutions because they were unable to track all communications by their employees using WhatsApp or Signal , [Bloomberg], creates new auditability constraints. It is with growing concern that an ECH enabled ecosystem may clash with future regulatory requirements.

4.3.2. Impact of ECH deployment on Network Security Operations

Enterprises approach to endpoint control varies significantly depending on size, use cases and a vast number of other factors.

For example, large enterprises generally exert control over their endpoints, yet to the limits of some use cases they may need to implement, e.g. BYOD. The latter was accelerated, as per above, due to COVID forcing more flexibility in the extended work force (employees, contractors, etc.).

On the contrary, small and some medium businesses may not be in the position to control their endpoints to the same extent (see specific implications for SMBs section below).

As some Browser makers made the use of ECH optional, this gives a first approach for enterprises to disable ECH for their employees.

However this doesn't provide an holistic solution. Indeed enterprises will need to consider a number of issues:

* Browsers which do not offer an option to disable ECH

* Browsers that will make ECH non optional in the future

* Non-browsers applications which are designed to use libraries enforcing ECH, without any option to disable it

* All the range of BYOD use cases where enterprises do not control the endpoint

* Adversaries leveraging ECH e.g. to hide their command and control communications, e.g. in Ransomware cases.

Whilst, disabling ECH wherever possible provides one approach to mitigate ECH deployment issues, as per above, other mitigations approaches need to be offered to enterprises.

(Editor's note: we need to describe how to strip the RRs to force a global disabling of ECH, yet mindful it might not be sufficient if an adversary finds a way to not use the enterprise DNS resolver)

4.3.2.1. Reminders on Network Security

Network Security is a set of security capabilities which is articulated as part of a defense strategy, e.g. Defense In Depth [NIST-DID], Zero Trust, SASE/SSE, etc. and can trigger and enable other security capabilities such as sandboxing, Data Loss Prevention, Cloud Access Service Broker (CASB), etc. One constituency is a Web Proxy, combining both a TLS proxy and an application level (HTTP) proxy.

In the same way that [I-D.draft-ietf-opsec-ns-impact] showed the impact of TLS1.3 on operational security, a loss of visibility of the SNI as indicator of compromise (see [I-D.draft-ietf-opsec-indicators-of-compromise]) has two main implications

4.3.2.2. Implications from loss of Meta Data

The loss of visibility of the SNI, at TLS level, will prevent transparent proxies from applying corporate policies to manage risk and compliancy. Typical examples:

* categories of compromised sites cannot be applied anymore, exposing employees and their organisations to potential cybersecurity risks; alternative approaches to block access to theses sites need to be found

* corporate lists of excluded sites for compliance or policy reasons need alternatives ways to be blocked.

4.3.2.3. Implications from loss of Selective Decrypt

TLS proxies also have the ability to selectively intercept, avoiding any visibility into or modification of the original application protocol payload - but such selective intercept relies heavily on knowledge of the origin content server hostname, which can be extracted in plaintext from the TLS ClientHello SNI (server name) field.

This capability allows the application proxy, in particular an HTTPS proxy to engage efficiently specific security controls, e.g. Data Loss Prevention, Sandboxing, etc.

The loss of SNI visibility will make it more difficult for corporate user flows to be intercepted, with it becoming impossible for BYOD use cases.

This will create inefficiencies, will require more resources and will increase security risks. It will also be counter productive for privacy as it may require the proxy to decrypt the whole TLS connection.

4.3.3. Specific implications for SMBs

Small and Medium Business (SMBs) form a particularly vulnerable subset of enterprises and organizations and span from Small Office Home Office (SOHO, sometimes a one person business) to Medium Business with strong variations depending on the country (a 50 employee company is considered the upper range of SMB business in developing countries while it is up to 25,000 in some developed countries).

Similarly to the above education use case and irrespective of definitions, many SMBs have very limited in-house capabilities to defend themselves, with security often outsourced to Managed Security Service Providers (typically network operators, mid range and small service providers).

5. Public Network Service Providers

5.1. Context

In Public Networks the national, regional and international legislator has to balance between freedom of access to the information on the one hand, and safety of the Internet and the protection of other fundamental rights on the other hand.

There are 2 main approaches:

* First, there are countries which do not have any specific legislation on the issue of blocking, filtering and takedown of illegal Internet content: there is no legislative or other regulatory system put in place by the state with a view to defining the conditions and the procedures to be respected by those who engage in the blocking, filtering or takedown of online material. In the absence of a specific or targeted legal framework, several countries rely on an existing "general" legal framework that is not specific to the Internet to conduct what is, generally speaking, limited blocking or takedown of unlawful online material. Here the approach has been differentiated in relying on self regulation from the private sector or limited political or legislative intervention to specific areas.

* The other approach has been to set up a legal framework specifically aimed at the regulation of the Internet and other digital media, including the blocking, filtering and removal of Internet content. Such legislation typically provides for the legal grounds on which blocking or removal may be warranted, the administrative or judicial authority which has competence to take appropriate action and the procedures to be followed.

5.2. Mitigations

5.2.1. Current approaches and procedures

In relation to specific areas where the public interest has to be protected more strongly, such as child abuse crimes, terrorism, criminality and national security, many states have a framework for the urgent removal of Internet content regarding the above materials without the need of a court order. In such circumstances, administrative authorities, police authorities or public prosecutors are given specific powers to order Internet access providers to block access without advance judicial authority. It is common to see such orders requiring action on the part of the Internet access provider within 24 hours, and without any notice being given to the content provider or host themselves.

Particularly in relation to material concerning child abuse and other serious crimes, many countries adopt a “list” system, whereby a central list of blocked URLs or domain names are maintained and updated by the relevant administrative authority. This is notified to the relevant Internet access providers, who are required to ensure that blocking is enforced. Additionally in some states the authorities can request the removal of content that infringes intellectual property, privacy or defamation rights. In this case the removal need to be requested by a court order.

Generally speaking, the grounds relied on broadly correspond to the interests protected under Article 10(2) of the European Convention of Human Rights (ECHR), namely: the protection of national security, territorial integrity or public safety, the prevention of disorder or crime, the protection of health or morals, the protection of the reputation or rights of others, and the prevention of the disclosure of information received in confidence. From the methodology we have to distinguish between blocking or takedown of content.

* The blocking, filtering or prevention of access to Internet content are generally technical measures intended to restrict access to information or resources typically hosted in another jurisdiction. Such action is normally taken by the Internet access provider through hardware or software products that block specific targeted content from being received or displayed on the devices of customers of the Internet access provider.

* Takedown or removal of Internet content, on the other hand, will instead broadly refer to demands or measures aimed at the website operator (or "host") to remove or delete the offending website content or sub content.

In these considerations we will refer to blocking only.

5.2.2. The blocking use case

This can be achieved through a number of techniques, including the blocking of the Domain Name System (DNS), the analysis of the SNI

field or the Uniform Resource Locator (URL). Given the increasing adoption of encryption techniques often a mixture of the above techniques is needed.

In particular for the most serious crimes such as child abuse or national security many countries adopt a “list” methodology, where a central list of blocked Domains or URLs is maintained by the authorities and updated on a regular basis (daily or even hourly) and shared with Public Network Operators that have to enforce the blocking.

In many jurisdictions there are legal consequences for the Operator not complying with the blocking order.

Technically the blocking can be implemented using techniques that have been adapted over time as new technologies have been introduced.

Historically depending on the content of the list the technique have been based on DNS or proxy blocking.

DNS is effective on Domains (the whole domain is blocked), while proxy is effective either on Domain (for encrypted traffic) or URL (for unencrypted traffic).

Given that nowadays the vast majority of traffic is encrypted, the capability of blocking based on URL is limited to a small portion of traffic and proxy blocking is as effective as that based on the DNS.

Theoretically DNS blocking would be the preferred option for operators given the more limited investments necessary to implement blocking of the Domains, but given the increased usage of external encrypted DNS services DNS blocking is becoming less effective and operators need to use SNI analysis as well in order to fulfil legal obligations.

5.3. Implications

The adoption of ECH will cause additional problems and limit the possibility of implementing operators fulfilling their legal blocking obligations, exposing the population to illegal content related to crimes such as Child Sex Abuse Material (CSAM), malware and other malicious content, and possibly even content deemed to be detrimental to National Security.

In addition, operators that do not fulfil their legal obligations may be exposed to legal or regulatory remedies.

6. General issues

6.1. Threat Detection

[RFC8404] identifies a number of issues arising from increased

encryption of data, some of which apply to ECH. For example, it notes that an early trigger for DDoS mitigation involves distinguishing attacker traffic from legitimate user traffic; this become more difficult if traffic sources are obscured.

The various indicators of compromise (IoCs) are documented in [I-D.draft-ietf-opsec-indicators-of-compromise], which also describes how they are used effectively in cyber defence. For example, section 4.1.1 of the document describes the importance of IoCs as part of a defence- in-depth strategy; in this context, SNI is just one of the range of indicators that can be used to build up a resilient defence (see section 3.1 in the same document on IoC types and the 'pyramid of pain').

In the same Internet-Draft, section 6.1 expands on the importance of the defence in depth strategy. In particular, it explains the role that domains and IP addresses can play, especially where end-point defences are compromised or ineffective, or where endpoint security isn't possible, such as in BYOD, IoT and legacy environments. SNI data plays a role here, in particular where DNS data is unavailable because it has been encrypted; if SNI data is lost too, alongside DNS, defences are weakened and the attack surface increased.

6.2. Endpoint security limits

(Editorial note: Elaborate on endpoint security complications as [I-D.draft-taddei-smart-cless-introduction] as well as [MAGECART] [MITB] [MITB-MITRE] [MALVERTISING] showed that in some cases, the only way to detect an attack is through the use of network-based security. The loss of visibility of the SNI data will make it much harder to detect attacks. The endpoints components (operating system, applications, browsers, etc.) cannot be judge and jury.)

6.3. Network management

(Editorial note: this is a placeholder for future issues)

6.4. Future operational deployment issues due to the introduction of the Client Facing servers themselves

(Editorial note: this is a placeholder for future issues;

* Consolidation considerations - the use of ECH may accelerate the move of content away from standalone servers and on to CDNs, reducing infrastructure resilience.

* What happens if Client Facing servers are controlled by malicious actors?

* The Client Facing servers are acting as a new category of middleboxes. In this shift left movement, until the attack surface is minimal and complexities are removed, you have to rely

on third parties for inspection. In these conditions, on which basis can they be more trusted than any other middleboxes? Is this creating a concentration problem?

)

6.5. Migration issues

(Editorial note: this is a placeholder for future issues;

* If ECH is enforced what are the solutions to all the above problems and what are the migration paths?

)

7. Potential further development of this work

7.1. Potential development of this document.

This section lists potential development of this work in particular for the General Issues section.

* There are need for further clarifications from the ECH draft, e.g. The link between the Client Facing and the backend servers are not clear enough and need further description. It can't be just 'left to the implementation'. The action is still underway and feedback to the TLS working group will be provided.

* Will there be any impact to the DNS by adding so many new RRs?

7.2. Potential development outside of the scope of this document

This document infers a number of ideas that could be relevant for other groups and in other deliverables. In particular regarding what type of solutions could be considered

* There is a need to address the apparent disconnect between user privacy and security, it should be possible to provide both rather than one compromising the other.

* What prevents a Client Facing server providing security solutions to protect the data path?

* Given some of the many challenges there is the opportunity to review the current ECH proposal from the perspective of a respectful inspection protocol.

8. Conclusion

Access to SNI data is sometimes necessary in order for institutions, including those in the education and finance sectors, to discharge their compliance obligations. The introduction of ECH in client

software poses operational challenges that could be overcome on devices owned by those institutions if policy settings are supported within the software that allows the ECH functionality to be disabled.

(Editorial note: these two below paragraph need revision towards the end of the development of this draft)

Third-party devices pose an additional challenge, primarily because the use of ECH will render transparent proxies inoperable. The most likely solution is that institutions will require the installation of full proxies and certificates on those devices before they are allowed to be connected to the host networks. They may alternatively determine that such an approach is impractical and instead withdraw the ability for network access by third-party devices.

An additional option that warrants further consideration is the development of a standard that allows a network to declare its policy regarding ECH and other such developments. Clients would then have the option to continue in setting up a connection if they are happy to accept those policies, or to disconnect and try alternative network options if not. Such a standard is outside of the scope of this document but may provide a mechanism that allows the interests and preferences of client software, end-users and network operators to be balanced.

9. Security Considerations

In addition to introducing new operational and financial issues, the introduction of SNI encryption poses new challenges for threat detection which this document outlines.

This I-D should help improve security in deployments of ECH.

10. IANA Considerations

This document has no IANA actions.

11. References

11.1. Normative References

[RFC6066] Eastlake 3rd, D., "Transport Layer Security (TLS) Extensions: Extension Definitions", RFC 6066, DOI 10.17487/RFC6066, January 2011, <https://www.rfc-editor.org/rfc/rfc6066>.

[RFC8484] Hoffman, P. and P. McManus, "DNS Queries over HTTPS (DoH)", RFC 8484, DOI 10.17487/RFC8484, October 2018, <https://www.rfc-editor.org/rfc/rfc8484>.

11.2. Informative References

[Bloomberg]
Spezzati, S., Robinson, M., and L. Beyoud, "Wall Street's Record Fines Over WhatsApp Use Were Years in the Making", 16 August 2022, <https://www.bloomberg.com/news/articles/2022-08-16/wall-street-sticker-shock-whatsapp-fines-were-years-in-making>.

[CIPA] FCC, "Children's Internet Protection Act (CIPA)", 30 December 2019, <https://www.fcc.gov/consumers/guides/childrens-internet-protection-act/>.

[Coroner] Henderson, "Prevention of future deaths report", 26 November 2021, <https://www.judiciary.uk/publications/frances-thomas-prevention-of-future-deaths-report/>.

[ECH_Roundtable]
419 Consulting, "Encrypted Client Hello - Notes from an ECH Roundtable", 18 August 2021, <https://419.consulting/encrypted-client-hello/>.

[I-D.draft-ietf-opsec-indicators-of-compromise]
Paine, K., Whitehouse, O., Sellwood, J., and A. S, "Indicators of Compromise (IoCs) and Their Role in Attack Defence", Work in Progress, Internet-Draft, draft-ietf-opsec-indicators-of-compromise-04, 3 February 2023, <https://datatracker.ietf.org/doc/html/draft-ietf-opsec-indicators-of-compromise-04>.

[I-D.draft-ietf-opsec-ns-impact]
Cam-Winget, N., Wang, E., Danyliw, R., and R. DuToit, "Impact of TLS 1.3 to Operational Network Security Practices", Work in Progress, Internet-Draft, draft-ietf-opsec-ns-impact-04, 26 January 2021, <https://datatracker.ietf.org/doc/html/draft-ietf-opsec-ns-impact-04>.

[I-D.draft-ietf-tls-esni]
Rescorla, E., Oku, K., Sullivan, N., and C. A. Wood, "TLS Encrypted Client Hello", Work in Progress, Internet-Draft, draft-ietf-tls-esni-17, 9 October 2023, <https://datatracker.ietf.org/doc/html/draft-ietf-tls-esni-17>.

[I-D.draft-mcfadden-smart-threat-changes]
McFadden, M., "BCP72 - A Problem Statement", Work in Progress, Internet-Draft, draft-mcfadden-smart-threat-changes-04, 22 January 2022, <https://datatracker.ietf.org/doc/html/draft-mcfadden-smart-threat-changes-04>.

[I-D.draft-taddei-smart-cless-introduction]

Taddei, A., Wueest, C., Roundy, K. A., and D. Lazanski, "Capabilities and Limitations of an Endpoint-only Security Solution", Work in Progress, Internet-Draft, draft-taddei-smart-cless-introduction-03, 13 July 2020, <https://datatracker.ietf.org/doc/html/draft-taddei-smart-cless-introduction-03>.

[KCSE] DfE, "Keeping children safe in education 2021", 1 November 2021, <https://419.consulting/encrypted-client-hello/>.

[LOSSINCAP] Neyret, A. and Autorité des Marchés Financiers, "La cybercriminalité boursière – définition, cas et perspectives", 10 October 2019, <https://www.amf-france.org/sites/default/files/2020-02/etude-sur-la-cybercriminalite-boursiere-_-definition-cas-et-perspectives.pdf>.

[LOSSINCREDITSCORE] Deloitte, "Beneath the surface of a cyberattack – A deeper look at business impacts", 2016, <https://www2.deloitte.com/content/dam/Deloitte/global/Documents/Risk/gx-risk-gra-beneath-the-surface.pdf>.

[LOSSINREVENUE] ANOZR WAY, "BAROMÈTRE ANOZR WAY DU RANSOMWARE", 4 September 2022, <https://anozrway.com/wp-content/uploads/dlm_uploads/2022/09/ANOZR-WAY_Barometre-Ransomware_edition-septembre-2022.pdf>.

[MAGECART] Wikipedia, "Magecart", 3 April 2022, <https://en.wikipedia.org/wiki/Web_skimming#Magecart>.

[MALVERTISING] Wikipedia, "Malvertising", 2 June 2022, <https://en.wikipedia.org/wiki/Malvertising>.

[MITB] OWASP, "Man-in-the-browser attack", n.d., <https://owasp.org/www-community/attacks/Man-in-the-browser_attack>.

[MITB-MITRE] MITRE, "Browser Session Hijacking - T1185", 25 February 2022, <https://attack.mitre.org/techniques/T1185/>.

[NIST-DID] NIST, "Glossary - defense-in-depth", n.d., <https://csrc.nist.gov/glossary/term/defense_in_depth#:~:text=Definition(s)%3A,and%20missions%20of%20the%20organization.>.

[RFC7258] Farrell, S. and H. Tschofenig, "Pervasive Monitoring Is an Attack", BCP 188, RFC 7258, DOI 10.17487/RFC7258, May

2014, <https://www.rfc-editor.org/rfc/rfc7258>.

[RFC8404] Moriarty, K., Ed. and A. Morton, Ed., "Effects of Pervasive Encryption on Operators", RFC 8404, DOI 10.17487/RFC8404, July 2018, <https://www.rfc-editor.org/rfc/rfc8404>.

[RFC8744] Huitema, C., "Issues and Requirements for Server Name Identification (SNI) Encryption in TLS", RFC 8744, DOI 10.17487/RFC8744, July 2020, <https://www.rfc-editor.org/rfc/rfc8744>.

[RFC8890] Nottingham, M., "The Internet is for End Users", RFC 8890, DOI 10.17487/RFC8890, August 2020, <https://www.rfc-editor.org/rfc/rfc8890>.

[SOLARWIND] Symantec, a Division of Broadcom Software Group, "SolarWinds (Sunburst) Attack What You Need to Know", December 2020, <https://symantec.broadcom.com/en/solarwinds-sunburst-attacks>.

Appendix A. Acknowledgment

In memory of Simon Edwards who passed away in the night of 8th-9th of January 2023.

In addition to the authors, this document is the product of an informal group of experts including the people listed in the Contributors list in Appendix.

Appendix B. Initial data illustrating SNI unreliability

B.1. Data collections

In this appendix a couple data sets were collected from SSL Session logs from a Symantec SSLV. The goal was to see how prevalent are TLS sessions being established where the Server Name Indicator (SNI) was incorrect as compared to the Subject Alternative Name (SAN) contained within the Server Certificate. Applications and browsers that are establishing these mismatched connections have TLS Hygiene issues. In other words these sessions are being improperly established. None of the traffic was for nefarious means. However, an improperly defined SNI can be used to fool inspection devices to bypass security rules and measures.

The first set was based on consumer traffic which includes Internet of Things, Social Media, and Corporate access traffic. The dataset of session log entries were over 63K event entries over a 24 hour period.

The second dataset was from a telecommunications customer with Proxy

Offloading. The log entries were for a period of 24 hours and contained over 5M log events entries. Since this customer is using a Symantec Edge Proxy aligned with SSLV. The session data is for explicit clients. Guest or Internet of things type traffic is a much lower percentage of total traffic. However the existence of Mismatched SNIs persisted.

B.2. Consumer Network Traffic

For consumer based network traffic Mismatched SNIs were very prevalent. Actually out of new sessions the majority of the sessions were with mismatched SNIs as compared to properly matched SNIs. These were the result of many many short lived TLS sessions that persistently 'phone home'. 22% of all traffic was mismatched as compared to only 4% was properly matched. The rest of the log activity was non session related.

The services that were in the top 20 offenders for mismatched SNIs includes Google, Apple, Adware, IoT. Google DNS dominated the counts at 5.3% of all mismatched sessions. A close second at 4.8% was Samsung Smart things. But, if I add up common services like Google, Apple, Adware, and the remainder: 13%, 9.8%, 8%, and 10% respectively.

For matched traffic Amazon Alexa stood tall at 25%, with a runner up for Broadcom Cloud Proxy at 7.9%. Both Google and Apple did have a minority of properly established sessions. In other words some of their stuff is clean.

B.3. Corporate Customer Traffic

Since the traffic was proxy traffic the session hygiene was much better. The majority of new TLS sessions were properly established with matching SNIs/SANs. The vast majority of this traffic was VPN based. Likely masking out consumer like traffic invisible within the VPN tunnels. For instance there was 0.8% Google.com traffic reported. But, there was a small percentage ~1.3% Google API traffic that was mismatched. I would have expected more search engine activity.

Looking across the distribution of domain names for mismatched sessions. 29.6% of the traffic was related to Corporate applications. These applications could be updated and corrected. The next runner up coming in at 7.4% was for Akamai which also could be updated. Office applications and the remainder each individually accounted for 2.4% of the traffic.

B.4. Take away observations

* IoT and API based traffic is by far the largest offender as compared to Browser based initiated sessions.

* Long-lived TLS session counts were dwarfed by the chatter of the API calls using short lived yet pervasively reporting. For instance there were new sessions at a rate of every 20 seconds per IoT device.

* The consumer mismatched sessions were all using TLS v1.3. Which reaffirmed the need to decrypt TLS v1.3 traffic. These sessions, if established without TLS interception, may have gone unreported by NGFW, which makes policy decisions on SNI vs SAN. Conversely, the corporate traffic was a close split between TLS v1.3 and v1.2.

* The presence of VPN tunnels masked a clearer picture of the corporate traffic usage.

* SNI Mismatches are more prevalent in the wild than first thought.

* SNI Mismatches has a side cause which policy rules have to be enumerated twice for category matching. And the second category matching is more intensive since it has to enumerate the entire SAN list which can be very large.

* Fixing the session hygiene for corporate owned applications could potentially improve performance of the security stack.

Contributors

Eric Chien
Broadcom
Email: Eric.Chien@broadcom.com
URI: https://www.linkedin.com/in/eric-chien-66b4b258/

Eric contributed to the analysis of the Man in the Browser attacks.

Gianpaolo Scalone
Vodafone
Email: gianpaolo-angelo.scalone@vodafone.com
URI: https://www.linkedin.com/in/gianpaoloscalone/

Contributed the research on the conflicts of ECH with local legislations to block.

Daniel Engberg
Skandinaviska Enskilda Banken AB (SEB)
Email: daniel.engberg@seb.se
URI: https://www.linkedin.com/in/daniel-engberg-1561aaa/

Validate the issues for his organization.

Celine Leroy

Eight Advisory
Email: celine.leroy@8advisory.com
URI: https://www.linkedin.com/in/celine-leroy-1a534252/

Thank you to Céline for her work on cybersecurity financial impacts on enterprises.

Daniel Engberg
Skandinaviska Enskilda Banken AB (SEB)
Email: daniel.engberg@seb.se
URI: https://www.linkedin.com/in/daniel-engberg-1561aaa/

Validate the issues for his organization.

Gianpiero Tavano
Broadcom
Email: Gianpiero.Tavano@broadcom.com
URI: https://www.linkedin.com/in/gianpiero-tavano-5b975383/

Review the text, provided feedback and reminded us on the budgetary issues

Roelof duToit
Broadcom
Email: roelof.dutoit@broadcom.com
URI: https://www.linkedin.com/in/roelof-du-toit-a66831/

Roelof contributed many things including research, former I-D, text, the newly setup github, etc.

Diego Lopez
Telefonica
Email: diego.r.lopez@telefonica.com
URI: https://www.linkedin.com/in/dr2lopez/

Diego contributed in several aspects including MCPs.

Gary Tomic
Broadcom
Email: gary.tomic@broadcom.com
URI: https://www.linkedin.com/in/garytomic/

Gary contributed many things including research, keep us on scope, critique for when issues where not impacted by ECH as we initially thought.

Bob Blair
Broadcom
Email: bob.blair@broadcom.com
URI: https://www.linkedin.com/in/bob-blair-8b7273/

Bob contributed to several reviews, many calls, and the whole appendix A.

Pascal Paisant
BNP Paribas
Email: pascal.paisant@bnpparibas.com
URI: https://www.linkedin.com/in/pascal-paisant-727a531/

Pascal contributed to several parts, in particular in the general SNI section, on enterprises section and on migration issues.

Zied Turki
ZT Consulting
Email: zied.turki@ztconsulting.fr
URI: https://www.linkedin.com/in/zied-turki/

Zied contributed to several parts, in particular the rationale on SNI unreliability.

Authors' Addresses

Andrew Campling
419 Consulting Limited
Email: Andrew.Campling@419.Consulting
URI: https://www.419.Consulting/

Paul Vixie
Red Barn
Email: paul@redbarn.org
URI: http://www.redbarn.org/

David Wright
UK Safer Internet Centre
Email: david.wright@swgfl.org.uk
URI: https://saferinternet.org.uk/

Arnaud Taddei
Broadcom
1320 Ridder Park Dr
San Jose, CA 95131
United States of America

Phone: 41795061129
Email: Arnaud.Taddei@broadcom.com
URI: https://www.linkedin.com/in/arnaudtaddei/

Simon Edwards
Broadcom
1320 Ridder Park Dr
San Jose, CA 95131
United States of America
Email: Simon.Edwards@broadcom.com
URI: https://www.linkedin.com/in/simononsecurity/