opsec
Internet-Draft
Intended status: Informational
Expires: 27 July 2026

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

23 January 2026

Encrypted Client Hello Deployment Considerations
draft-campling-ech-deployment-considerations-11

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.

Status of This Memo

Copyright Notice

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, including the Server Name Indication (SNI), which on-path middleboxes can see. This data could be used for various purposes, including the determination of policy decisions, in some cases for security or compliance reasons. As part of a wider initiative by the IETF 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 its peers. Developments like ECH mark a fundamental change in the architecture of the Internet, allowing opaque paths to be established from endpoints to 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 the potentially adverse security implications for end users. Given these considerations, the deployment of ECH 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 Virtual Private Networks (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. This includes 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, the introduction of ECH has 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 as there are use cases that benefit from its deployment;

* it does not 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 assessment of alternative options if 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, but with limited evidence, that "most of ’the unanticipated usage’ functions can, however, be realised 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.

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 three metadata sources: the certificate, the DNS name and the SNI. A middlebox may have used some or all of this metadata to determine the destination of the communication.

As part of the current initiative to provide complete end-to-end encryption, the visibility of this data has progressively diminished. Firstly the certificate was encrypted into TLS1.3, then encrypted DNS protocols such as DNS-over-TLS, DNS-over-HTTP and DNS-over-QUIC have been introduced to encrypt the DNS flow to its resolver, making it harder for middleboxes to use this information.

Even in situations where the DNS data can be accessed, it can be misleading (does it point to the real destination, or just the site hosting server name, or a proxy?). The SNI was invented to overcome some of the limitations of the DNS data by providing additional information. However, the SNI in itself may also be unreliable, which is why middleboxes start by not trusting it until they have validated the information that it provides (see Section 2.5 for more details).

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 explicitly or transparently.

* With the 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 to the end point that the request has been handled by the target destination directly.

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 rely on the SNI and not the DNS

Where multiple services are hosted on a shared server, the use of the DNS in isolation would be insufficient to establish a connection to the target service. As per the Introduction section of [RFC8744] "More and more services are collocated on multiplexed servers," so the SNI was introduced to allow "direct connections to the appropriate service implementation".

In other words, a proxy cannot rely only on the DNS to ensure the establishment of a connection, the SNI is simply required by design.

2.5. The unreliability of the SNI

SNI is not reliable by design, prompting the question why and how middleboxes are using it in the first place. It should be considered that, in security, in general, unreliability can be a useful source of information.

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

In other terms, the 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. Bad actors can easily exploit this to bypass security middle boxes. For example, malware can be coded to provide an SNI value that is mapped to a confidential category, such as personal finance or healthcare, to bypass inspection in a selective decryption regime.

2.5.1. How Unreliable is the SNI?

To illustrate the degree of unreliability of SNI data, two data sets were collected from SSL Session logs from a Symantec SSLV. The goal was to see how the prevalence of TLS sessions being established where the Server Name Indicator (SNI) was incorrect when 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 because these sessions are being improperly established.

None of the traffic in question was malicious. However, an improperly defined SNI could be used by an attacker to fool inspection devices to bypass security rules and measures.

2.5.1.1. The Datasets

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

The second dataset was from a telecommunications customer with Proxy Offloading. The log entries were from a 24-hour period and contained over five million log event entries. Since this customer was using a Symantec Edge Proxy aligned with SSLV, the session data was for explicit clients, and guest or Internet-of-things type traffic was a much lower percentage of total traffic. However, the existence of mismatched SNIs persisted.

2.5.1.2. Consumer Network Traffic

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

The top 20 services for mismatched SNIs in this sample included Google, Apple, Adware and IoT. Google DNS was the highest category with 5.3% of all mismatched sessions, followed by Samsung Smart things with 4.8%. Common services like Google, Apple, Adware and the remainder comprised 13%, 9.8%, 8%, and 10% respectively.

For matched traffic, Amazon Alexa was the biggest category with 25%, followed by Broadcom Cloud Proxy with 7.9%. Both Google and Apple services had a minority of properly established sessions.

2.5.1.3. Corporate Customer Traffic

Because the corporate dataset was proxy traffic, the session hygiene was much better, most new TLS sessions were properly established with matching SNIs/SANs. Note that the vast majority of this traffic was VPN-based, likely masking consumer-like traffic within the VPN tunnels.

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 highest category at 7.4% was Akamai, which could also be updated. Office applications and the remainder each individually accounted for 2.4% of the traffic.

2.5.1.4. Conclusions

* IoT and API based traffic is by far the largest offender for mismatched SNIs compared to browser-based initiated sessions.

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

* The consumer mismatched sessions were all using TLS v1.3, reaffirming 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.

* The existence of SNI mismatches has a side effect - policy rules have to be enumerated a second time 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 improve the performance of the security stack.

2.5.2. Middleboxes and TLS 1.2

When attempting to set up a connection, the user client generates the "ClientHello" with the SNI in plain text indicating 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 plain text mode. Middleboxes can then see and inspect both the SNI value in the "ClientHello" and the server certificate details in the "ServerHello" message. Middleboxes can then perform selective inspection based on the destination service details (the service requested by the user client) extracted from the server certificate.

Depending on the middlebox design, the correlation of different sources of information is used to confirm the real destination service. However, only information within the server certificate is sufficiently reliable to perform accurate web categorisation and then undertake selective inspection if required to initiate any necessary web or content filtering.

2.5.3. Middleboxes and TLS 1.3

TLS 1.3 offers significant improvements over TLS 1.2 in terms of Its security and privacy properties. 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. As the inspection capabilities of middleboxes are designed based on the server certificate, all vendors worked to adapt their capabilities to support TLS 1.3 (e.g ServerHello encryption).

2.5.3.1. How do providers undertake inspection?

The most common technique to support TLS 1.3 inspection is to get the SNI from the ClientHello (which is still shared in plain text format in TLS 1.3) before establishing a new full TLS session initiated from the proxy server to the destination server. Once the server certificate has been retrieved, the web category can be determined, after which selective inspection can be performed on the real TLS session initiated by the user client.

As the SNI is not reliable (see 2.5.1), proxies provisionally accept the SNI but do so without trusting it, then perform a range of checks to verify the data. This step-by-step approach enriches the evaluation and informs which policy to apply. This could end up with the proxy deciding to block the connection, or it may let the connection complete with a verified or corrected SNI.

2.5.4. Middleboxes and 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. In some cases, the DNS data may provide information, but only if it is not using an encrypted protocol; even then, this cannot be compared with the SNI.

So TLS 1.3 with ECH has an impact on security and compliance capabilities (including selective inspection), not because of the lack of visibility of the SNI (which is not reliable in isolation) due to encryption, but because other information about the destination server is not available and can’t be used to fetch and retrieve the server certificate (which is indeed reliable) to apply appropriate policies such as web categorisation.

The critical problem is not being able to identify the destination service in order to get the destination server certificate details.

3. Use Cases: The Impact of ECH on 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 the 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 on the Internet using 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 serious risks posed if the content 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 many schools have a very different budget for IT compared to enterprises and may 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 educational 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 section 2.3). 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 have to require more invasive software to be installed on third-party devices before they can be used, assuming that ensuring the establishment has the capability to comprehend and adequately manage these technologies. Alternatively, the establishment will have to prevent all third-party devices from operating. Neither option is without negative consequences.

4. Use cases: Child Online Protection

4.1. Context

In the context of Child Online Protection (COP), the primary aim for illegal content is removal of content at source. Blocking and filtering adds friction, but it is not the end result. Block lists reduce access while giving time to entities dedicated to Child Online Protection to work towards content removal.

In this context, the SNI is key blocking encrypted websites hosting illegal content, in particular when hosts are slow to remove content as is the case in countries where much of the illegal content is hosted.

But for legal content that is harmful, or not appropriate for young people or in the workplace, e.g. sexual content, gambling, self-harm, animal cruelty, the content cannot be removed if it is not illegal.

The only option is to detect and block. Implementing ECH is removing a key tool e.g. for schools to meet their statutory requirements to prevent their networks being used to access content that is harmful or inappropriate for children, same with employees in enterprise networks.

At the moment some countries either incentivise or make it mandatory that operators may offer network based service such as parental controls [DECRETO28] or have laws in the making [SREN]. These would typically be implemented as DNS and/or SNI based controls.

However these controls can be circumvented by children who know how to change their DNS parameters to point back to "adult" DNS services.

(TBD: add UK examples on suicide, IWF next report, the term CSAM, etc.)

4.2. Implications

As there is a vast global unawareness of ECH, few people in charge realize the problem posed by ECH and are caught by surprise to even consider mitigation approaches or a migration plan.

In Child Online Protection use cases, most of the time, there is little to no programmatic control, or control at all, over the endpoints or the networks, not to mention BYOD.

And even on the network, the IAB is taking a direction [NOEPSCAN] which doesn’t seem to give a chance to prepare a migration to an alternative solution on the endpoint.

(TBD: How encryption is hindering investigations when children want to report issues, etc.,)

4.3. Mitigations

When ECH is deployed, if it becomes impossible to maintain blocking or filtering at network level, mitigations may still be possible at the endpoint.

There are few attempts to provide solutions [VFDNSERRORSVIDEO], [VFDNSERRORSSLIDES] or [BRCMWEBEXT] which is based on the idea to inspect the destination before anything goes on the wire, within the web browser via a web extension.

This is not a panacea as a web extension, in residential user context and in particular in COP context needs to be voluntarily installed and so can be easily disabled. Moreover this is only web browser context.

In this particular case, the fact that web browsers do not exhibit standard APIs adds to the difficulty to the need to orchestrate the web extension with the operating system. An area where Regulators may consider to be prescriptive.

5. Use Cases: The Impact of ECH on private network contexts (enterprises or other organisations)

5.1. Context

5.1.1. The main requirements

Enterprises and other organisations need to protect themselves for a vast number of reasons, including to:

* Reduce their risks. In particular, as part of any Cyber Resilience strategy.

* Protect their reputation. The term reputation includes many aspects beyond an organisation’s traditional assets, such as its data.
  For example, a "successful" cyber attack may impact brand equity, market capitalisation and creditworthiness.

* Compliance with a growing diverse set of policies, regulations, certifications, labelling 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.

5.1.2. A degrading threat landscape

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

* COVID crisis generally accelerated the overall attack landscape. As the crisis forced many enterprises and organisations 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.

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

* EUR1B loss of revenue for French organisations from January to August 2022 [LOSSINREVENUE]

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

* Degradation by credit notation agencies [LOSSINCREDITSCORE]

Another implication arising from the COVID crisis is the acceleration of BYOD to support remote working. This has created two side effects for remote employees, contractors and third parties (the latter two may need to connect to one or more enterprise networks on a temporary basis):

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

5.2. 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 the capabilities to defend, both:

* Directly: a lack of visibility on a key metadata like the SNI will cause significant issues to enterprises and other organisations

* Indirectly: should ECH happen and should alternatives be provided, managing migrations to any alternative not requiring access to the SNI, in these conditions, is undesirable from a timing, resource, capacity and risk perspective.

5.3. Implications

5.3.1. Examples of regulatory implications

Regulators are accelerating their lawfare capabilities at an accelerated pace and new legislation is impacting 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. US regulators levied 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.

5.3.2. Impact of ECH deployment on Network Security Operations

The approach to endpoint control by enterprises 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 this may impact some use cases they need to implement, e.g. BYOD. The latter was accelerated, as per above, due to COVID requiring more flexibility in the extended workforce (employees, contractors, etc.).

Small and at least some medium businesses may not be able 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 opportunity for enterprises to disable ECH for their employees.

However, this doesn't provide an holistic solution. Indeed enterprises will need to consider several issues:

* Browsers which do not offer an option to disable ECH

* Browsers that will make ECH non-optional in the future

* Non-browser applications which are designed with software libraries that implement ECH without any option to disable it

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

* Adversaries leveraging ECH, e.g. to hide their command and control communications, such as for ransomware.

Whilst disabling ECH wherever possible provides one approach to mitigate ECH deployment issues, as per above, other mitigation approaches also 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)

5.3.2.1. Reminders on Network Security

Network Security is a set of security capabilities which is articulated as part of a defence strategy, e.g. Defence-in-Depth [NIST-DID], Zero Trust, Secure Access Service Edge or Security Service Edge (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 an indicator of compromise (see [I-D.draft-ietf-opsec-indicators-of-compromise]) has two main implications.

5.3.2.2. Implications from loss of Meta Data

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

* Categories of compromised sites cannot be applied any more, exposing employees and their organisations to potential cybersecurity risks; alternative approaches to block access to these sites need to be found

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

5.3.2.3. 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 specific security controls when needed, 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 counterproductive for privacy as it may require the proxy to decrypt the whole TLS connection.

5.3.3. Specific implications for SMBs

Small and Medium Businesses (SMBs) form a particularly vulnerable subset of enterprises and organisations and range from Small Office Home Office (SOHO, sometimes a one-person business) to medium sized business with strong variations depending on the country. For example, 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). Budget constraints may impose limits on the level support that is on offer to the organisation.

Therefore, even where technical solutions exist that allow them to meet their operational and any compliance obligations, affordability and expertise may present SMBs with significant difficulties.

6. Use Cases: The Impact of ECH on Public Network Service Providers

6.1. Context

Public network operators often have significant obligations placed upon them by national, regional or international legislators and regulators. These may cover aspects such as freedom of access to the Internet, the protection of fundamental rights and protection of the underlying infrastructure from malicious actors.

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, some countries rely on an existing "general" legal framework that is not specific to the Internet to conduct what is limited blocking or takedown of unlawful online material. It relies on self-regulation by network operators or limited political or legislative interventions in specific areas.

* The other approach is 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 defines the legal grounds when blocking or removal may be warranted, the administrative or judicial authority which has competence to take appropriate action and the procedures to be followed.

6.2. Content Takedown and Blocking

6.2.1. Definitions

It is useful 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 an 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.

The following comments apply to blocking only.

6.2.2. 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, a mixture of the above techniques are often needed.

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/or trusted flaggers 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 any Operator that does not comply with a 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 techniques have mainly been based on DNS or proxy blocking.

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

Given that the vast majority of Internet traffic is encrypted, the capability of blocking based on URLs 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 Domains. However, with the increased usage of external encrypted DNS services, DNS blocking is becoming less effective so operators need to use SNI analysis as well in order to fulfil legal obligations.

6.3. The implications of ECH for public network operators

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

If existing techniques for content blocking are rendered ineffective, operators may increasingly consider the use of IP blocking to meet any legal or regulatory obligations. The major downside of such an approach is that it is a relatively crude method, risking significant over-blocking, especially where multiple services are hosted on a shared server.

Operators that do not fulfil their legal obligations may be exposed to legal or regulatory remedies.

7. General issues

7.1. 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.

7.2. 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.)

7.3. Network management

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

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

(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?

)

7.5. Issues pushing protection to endpoints

The solution to lack of network visibility is often to move the security/safety tool to the client. This works acceptably well in some circumstances, for example on Windows desktops. However, many safety tools on Android and IOS (particularly 3rd party parental controls) rely on the self same network technology. Usually this is a proxy running locally, which obviates the issue s of unreliability and cost for a mobile device, but retains all the challenges of inspection on-network when it comes to issues like ECH.

7.6. 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?

)

8. Potential further development of this work

8.1. Potential development of this document

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

* There is a need for further clarification from the ECH draft, e.g. the links 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?

8.2. 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 from 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.

9. Conclusion

Access to SNI data is sometimes necessary for organisations operating private networks, such as those in the education and finance sectors, to protect their operations and 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 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 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.

10. 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.

11. IANA Considerations

This document has no IANA actions.

12. References

12.1. 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>.

12.2. 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>.

[BRCMWEBEXT]
"Symantec Browser Protection", n.d., <https://chromewebstore.google.com/detail/symantec-browser-protecti/hielpjjagjimpgppnopiibaefhfpbpfn>.

[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/>.

[DECRETO28]
"DECRETTO-LEGGE 30 aprile 2020, n.28", n.d., <https://www.gazzettaufficiale.it/eli/id/2020/04/30/20G00046/sg>.

[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-25, 14 June 2025, <https://datatracker.ietf.org/doc/html/draft-ietf-tls-esni-25>.

[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 Autorit des Marchs Financiers, "La cybercriminalit boursire dfinition, 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, "BAROMTRE 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>.

[SREN] "Projet de loi visant scuriser et rguler l’espace numrique", n.d., <https://www.assemblee-nationale.fr/dyn/16/dossiers/DLR5L16N47884>.

[VFDNSERRORSSLIDES]
"Slides of Vodafone presentation on Use of DNS Errors to improve Browsing User Experience With network based malware protection", n.d., <https://datatracker.ietf.org/meeting/116/materials/slides-116-dnsop-dns-errors-implementation-proposal-slides-116-dnsop-update-on-dns-errors-implementation-00>.

[VFDNSERRORSVIDEO]
"Video of Vodafone presentation on Use of DNS Errors to improve Browsing User Experience With network based malware protection", n.d., <https://www.youtube.com/watch?v=xh_uQo46yhE&t=2471s>.

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.

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 Cline for her work on cybersecurity financial impacts on enterprises.

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/