Bootstrapping Privacy Services in Today's Internet

Taeho Lee
ETH Zürich
kthlee@inf.ethz.ch

Christos Pappas
ETH Zürich
pappasch@inf.ethz.ch

Adrian Perrig
ETH Zürich
aperrig@inf.ethz.ch

ABSTRACT

Internet users today have few solutions to cover a large space of diverse privacy requirements. We introduce the concept of privacy domains, which provide flexibility in expressing users' privacy requirements. Then, we propose three privacy services that construct meaningful privacy domains and can be offered by ISPs. Furthermore, we illustrate that these services introduce little overhead for communication sessions and that they come with a low deployment barrier for ISPs.

CCS CONCEPTS

• **Networks** → *Security protocols*; *Network privacy and anonymity*;

KEYWORDS

Privacy, Privacy Domain, ISP-level Privacy Services

1 INTRODUCTION

In 1983, Germany established the principle of informational self-determination for data processing in the context of a German constitutional ruling [1]. The principle empowers an individual to determine i) what personal information is disclosed and to whom, which reflects Westin's description of privacy [2], and ii) how the disclosed information is used. Although this ruling was made well before the Internet era, the principle is of special importance for online privacy, and there have been efforts to apply it in the context of today's Internet [3].

Unfortunately, in recent years, there have been legal developments that undermine privacy and diminish confidence in today's constitutions about the way they handle disclosed information. For example, in the U.S., ISPs can now sell their customers' history without obtaining customers' consent [4]. Also, we know that governments massively collect Internet traffic for thorough and detailed analysis [5]. Therefore, the Internet community is turning to technical solutions to limit the disclosure of personal information.

However, today's deployed solutions provide limited options. On one side, encryption protocols hide the application-layer payload (e.g., TLS), yet they still reveal the addresses of the communicating hosts. On the other side, anonymity solutions such as Tor provide stronger privacy guarantees, but at a high performance cost. There is a large tradeoff space between privacy and performance that remains unfulfilled. For example, a user may want to only prevent a web server from linking together the user's requests, but using Tor will probably lead to an unacceptable quality of experience.

Our goal is to explore practical and readily deployable solutions for the diverse privacy requirements of Internet users. Therefore, we take a user-centric approach to privacy, and our first step is to consider users' diverse privacy requirements. Following the definition of privacy [2]—what personal information is disclosed and to whom—we introduce the term of *privacy domains*. A privacy domain is defined by the entities (e.g., ISPs) and the privacy-sensitive information (e.g., source address) that is revealed to these entities for a user's communication session. Privacy domains help us in bridging the gap between users' high-level privacy requirements and the more actionable technical requirements.

Our second step is to identify simple and common networking practices that can be used to realize privacy domains. We identify three such practices—encryption, address translation, and tunneling—that can be composed to construct a range of privacy domains.

Then, based on these common networking practices, we propose three privacy services that can be offered by ISPs: i) An address-hiding service that enables customers to use a different IP address for every traffic flow. ii) An ISP-level tunneling service that channels traffic between source and destination ISPs over an encrypted point-to-point tunnel, where the two tunnel end points are operated by the source and destination ISPs. iii) An ISP-level VPN service that remote hosts of other ISPs can use.

We argue that ISPs are in an ideal market position to offer such privacy services. They already have high-capacity infrastructure, which they can use to offer services at a large scale. Furthermore, they have the required know-how and experience in deploying and operating large systems. Our initial evaluation results indicate that the services can be offered at a low overhead, even on today's commodity hardware. In addition, our proposed services have a low deployment barrier, offering incentives for first movers.

Contributions. The main focus of the paper is to explore privacy from a new perspective, rather than proposing new privacy-enhancing technologies. We leverage existing technologies to enable ISP-based privacy services, which can fill the large tradeoff space between privacy and performance. We make contributions in three directions:

- We introduce a concept—privacy domains—to express users' privacy requirements at a high-level, irrespective of the underlying implementation.
- We describe common networking practices and techniques that can serve as building blocks to implement privacy domains.
- We present simple privacy services (with a preliminary feasibility analysis) that can be offered by ISPs.

2 PRIVACY DOMAINS

We observe that users have diverse privacy requirements, yet few options or tools to achieve them. Our starting point is to introduce the concept of privacy domains, which enables us to argue about disclosed information at a higher layer of abstraction without considering a specific privacy architecture nor a specific adversary model; existing work focuses on architecture-specific analysis and considers specific threat models [6, 7].

For example, a common privacy requirement is to prevent entities that observe traffic from creating a history of user activity (e.g., a list of hosts that the user communicated with). This high-level requirement is translated to a more technical requirement that no entity can observe the source and destination addresses of a communication session at the same time. To facilitate this translation, we define *privacy domains*.

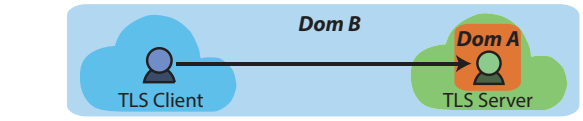
A *privacy domain* is a virtual domain that consists of entities with which a user shares the same subset of her privacy-sensitive information. Thus, a privacy domain is defined by a set of a user's privacy-sensitive information and the entities that have access to this information.

- **Entities.** We consider entities in a privacy domain from a sender's view-point, i.e., entities that assist in transferring packets from the sender to the destination. Such entities include the ISP of the sender, the transit ISPs on the path to the destination, the destination ISP, and the destination host. More specifically, with the terms "sender" and "destination" we refer to the producer and consumer of traffic; this traffic can be forwarded through one or more intermediate waypoints, e.g., tunnel end points; however, we do not consider the waypoints as senders or destinations in our model.
- **Privacy-sensitive Information.** We consider information that is revealed about the sender through sent packets. Specifically, we consider the source/destination host addresses and ISPs, the transport-layer headers, and the plaintext payloads.

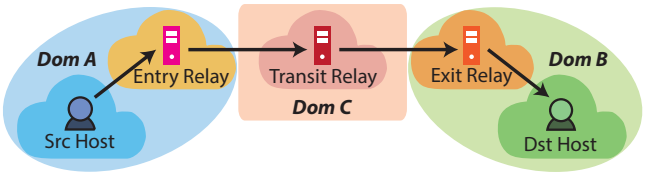
From a security perspective, privacy domains may remind one of threat models. Threat models describe which entities try to compromise what privacy-sensitive information and their capabilities to achieve their goal. For example, passive adversaries can only observe traffic, whereas active adversaries can also modify, delay, and even drop traffic. Furthermore, threat models often define what side-channel information can be used to obtain private information. Our goal, however, is different: privacy domains make it clear for Internet users what information about themselves would be shared with whom, making it easier for non-experts to make more informed privacy-related decisions. However, further study is needed to investigate how to help everyday users to make informed privacy-related decisions.

As an illustration, we describe the privacy domains that are created when using the TLS protocol and Tor. When a client connects to a server over the *TLS protocol*, two privacy domains are created (Figure 1a). Domain *A* includes the server to which the client shares all its privacy-sensitive information. Domain *B* consists of all other entities that observe traffic, i.e., all ISPs on the path including the client's ISP. The client shares all privacy-sensitive information with this domain, except for the (encrypted) payload.

When a client connects to a server using the *Tor network* [8], three privacy domains are created (Figure 1b) assuming that a typical three-hop circuit is used (i.e., a circuit with an entry, a transit, and an exit relay). Domain *A* consists of the Tor entry relay, the entry relay's ISP, the client's ISP, and all transit ISPs between these two ISPs. The client shares only her address with domain *A*. Domain *B* consists of the Tor exit relay and its ISP, the destination host and its ISP, and all transit ISPs between the two ISPs. The client shares with domain *B* the address of the destination host,



(a) TLS Protocol.



(b) A 3-hop Tor Circuit.

Figure 1: Privacy domains created by a) the TLS protocol and b) a 3-hop Tor Circuit. Clouds represent ISPs; for simplicity, we do not show transit ISPs; they would exist on inter-cloud arrows.

the transport-layer header and the payload. Note that the combination of source and destination addresses is not present in any of the two domains. This property enables a user to hide her history of user activity from all entities—at least in theory. Lastly, Domain *C* consists of the transit relay and its ISP, the transit ISPs between the ISPs of the entry and transit relays, and transit ISPs between the ISPs of the transit and exit relays. The client does not share any privacy-sensitive information with this domain.

By defining privacy domains, our goal is to clarify and articulate information disclosure, not to evaluate the privacy guarantees of a certain mechanism. For example, when a user leverages Tor, the combination of source and destination addresses is not present in any of the domains, however, there are sophisticated side-channel attacks that a strong adversary can launch to infer communicating source-destination pairs [9–13]. Existing approaches can be used to evaluate the privacy guarantees offered by certain mechanisms [14].

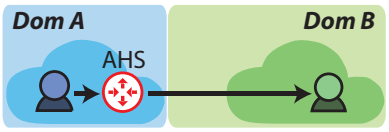
3 OVERVIEW

Our next step is to propose technical approaches that can realize privacy domains, i.e., translating from the high-level privacy requirements to more actionable networking practices and technologies. We start with basic building blocks that can help in constructing privacy domains. Then, we present three privacy services that are based on these blocks and can be offered by ISPs.

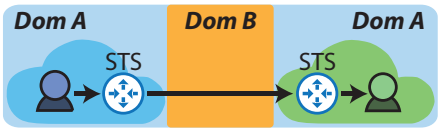
3.1 Building Blocks

Encryption masks privacy-sensitive information from unwanted parties. In its most common use case—the TLS protocol—encryption masks the payload so that only the destination can see it in plaintext. Alternately, encryption can also mask the transport-layer headers, thus hiding the application being used, as happens in IPsec transport mode.

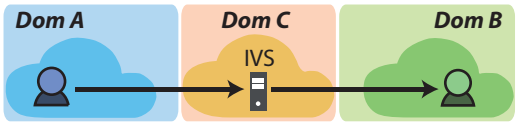
As a generic concept, encryption creates two privacy domains: one domain that is defined by the entities that have the decryption keys and therefore have access to all privacy-sensitive information; and the other domain that consists of all remaining entities that have access only to the unencrypted information in the packets.



(a) ISP-level Address Hiding Service (Section 3.2.1).



(b) ISP-level Tunneling Service (Section 3.2.2).



(c) ISP-level VPN Service (Section 3.2.3).

Figure 2: Privacy domains created by the ISP-based privacy services. For simplicity, we do not show transit ISPs; they would exist on inter-cloud arrows.

Address translation is used so that multiple users can share the scarce IPv4 address space. A side-effect of address translation is that an observer located after the translation point cannot identify the original source host of a packet.

Address translation creates two privacy domains: one domain that consists of the entities that know the original source address of the sender (and all other privacy-sensitive information), and one domain that consists of the entities that see all information except for the original source address.

Tunneling bridges two networks by creating a virtual point-to-point link; it is typically used to provision a service that the underlying network cannot support (e.g., supporting IPv6 over the IPv4 network). Tunneling by itself does not define new privacy domains, but can be used for this purpose. More specifically, it enables the sender to specify a waypoint that the traffic should follow towards the destination. Then, with the help of the waypoint and in combination with the previously mentioned building blocks new privacy domains can be defined (see Sections 3.2.2 and 3.2.3).

3.2 ISP-based Privacy Services

We describe privacy services that ISPs can offer; the services are based on the building blocks described, and thus compatible with today's networking practices. For each service, we describe the privacy domains that are constructed (Figure 2) and use cases that are based on common privacy requirements.

3.2.1 ISP-level Address Hiding Service (AHS). We start with an address-hiding service (AHS) that is based on *source-address translation* and can be offered by ISPs to its hosts. When subscribers' packets exit the ISP's boundary, the source addresses are replaced with other addresses in the ISP's address pool. Specifically, the source address is replaced with a different address for every outgoing flow. When subscribers' packets enter the ISP's boundary, the reverse translation takes place so that the packets are forwarded to the intended recipients. Hosts remain agnostic of AHS in that they purchase the service but they do not need to upgrade the OS nor run specialized applications.

We are not the first to propose address shuffling by ISPs. In 2009, Raghavan et al. [15] proposed the use of a tweakable block cipher to enable ISPs to shuffle their IP addresses; we build on the same motivation, but construct mechanisms that provide a higher degree of flexibility to ISPs (Section 4.1).

Privacy Domains. The AHS creates two privacy domains for a subscribing host (Figure 2a). Domain *A* consists of the host's ISP, and the host shares all its information with this domain. Domain *B* consists of all other entities; the host shares all information except for the source address, which is instead shared at the granularity of an AS.[1]

Use Cases. This service is useful for users who want to hide the history of their online activity, i.e., which hosts they contact; the online activity is still disclosed to the user's ISP.

3.2.2 ISP-level Secure Tunneling Service (STS). We propose ISPs to offer a secure tunneling service (STS) by setting up encrypted tunnels with other ISPs; source and destination ASes are the tunnel end points. The traffic is encrypted by the source AS when it enters the tunnel and then decrypted by the destination AS when it exits the tunnel. Similar to AHS, hosts remain agnostic of the tunneling service and do not need to upgrade.

Privacy Domains. This service is composed of two basic building blocks: tunneling and encryption. Tunneling specifies two waypoints on the path to the destination, and the waypoints en/decrypt the traffic, hiding it from other entities on the path.

STS creates two privacy domains for a subscribing host (Figure 2b). Domain *A* consists of the source ISP, the destination ISP, and the destination host; the subscriber shares all its information with this domain. Domain *B* consists of all transit ISPs, and the subscriber shares its address and the address of the destination host at the AS granularity.

Use Cases. ISP-level tunnels provide an additional security measure for traffic that is already encrypted. Services that exchange sensitive information typically perform encryption with TLS at the application layer, though this is not always sufficient. Protocol and implementation vulnerabilities of popular TLS libraries have enabled decryptions at a large scale (e.g., the Heartbleed attack [16] and compression attacks [17–19]). Moreover, the lack of forward secrecy in some TLS deployments can lead to compromised plaintexts, if a long-term key is compromised—about 30% of 200K popular TLS-enabled websites still do not fully support forward secrecy [20].

In addition, ISP-level tunnels can provide a layer of security for unencrypted webpages. Today's trend moves towards pervasive encryption and TLS is gaining traction; however, we are far from universally encrypted traffic—less than 20% of the top 10k websites and less than 0.1% of all websites have TLS enabled by default [21].

ISP-level tunnels can also harden today's privacy protocols. For example, Tor is known to be vulnerable against traffic correlation attacks when an adversary can observe traffic at the entry and exit points of the Tor network [22]. Even worse, an adversary can

[1]We use the term "ISP" when referring to services, and the term "AS" mostly for protocol-level details.

launch BGP prefix hijacking attacks to position itself on the path of inbound and outbound traffic [13]. ISP-level tunnels between source AS and the AS of the Tor entry node, or between the AS of the Tor exit node and the destination AS can enhance resilience against the correlation attacks.

STS alleviates the consequences of today's insecure inter-domain routing. More precisely, adversaries often launch BGP prefix hijacking attacks to attract traffic [23–26]. They can then analyze traffic patterns (who talks with whom), modify and inspect unencrypted traffic, or store encrypted traffic with the prospect of breaking it in the future. Using ISP-level tunnels does not prevent BGP hijacks, but limits the capabilities of attackers through strong network-layer encryption for all traffic in the tunnel.

3.2.3 ISP-level VPN Services (IVS). We propose ISPs to offer VPN services to hosts of other ISPs, similar to the VPN services that already exist on the market. The motivation for this proposal is that despite the increased interest in VPN services, the market is littered with low quality services. Reasonable performance is offered only by premium services that cost around $10 per month—a considerable fraction of an Internet connection's monthly cost [27]. Furthermore, most such services suffer from critical vulnerabilities (e.g., IPv6 traffic leakage and DNS hijacking) that disclose the identity and traffic payloads of users [28].

We argue that ISPs are in a better market position to offer VPN services at a fraction of what VPN services cost today. ISPs have high-capacity infrastructure and experience in deploying and operating services at a large scale. Furthermore, ISPs already manage large blocks of IP addresses, which are necessary to mask customer identities through address translation; it is commonly the case that today's VPN services point customers to connect to servers at a different location due to their IP address scarcity at some locations. Thus, we believe ISPs can leverage their experience and economies of scale to gain an extra source of revenue.

Privacy Domains. This service combines all three building blocks: tunneling, encryption, and address translation. Tunneling is used to specify the VPN provider as a waypoint for the traffic towards the destination host, and encryption hides the traffic from entities on the path to the VPN provider. Then, address translation replaces the subscriber's source address with another address in the provider's address pool.

Three privacy domains are created for a subscribing host (Figure 2c). Domain *A* contains the source ISP and all transit ISPs leading to the VPN provider, and the subscriber shares only its source address with this domain. Domain *B* contains all transit ISPs between the VPN provider and the destination ISP, the destination ISP, and the destination host; the subscriber shares all information but its source address. Domain *C* consists of the ISP offering VPN service, and the subscriber shares all its information with this domain.

Use Cases. Users may use VPN services for different purposes: i) to bypass geolocation restrictions, ii) to circumvent governmental censorship, and iii) to hide their activity from their ISPs by encrypting traffic and hiding the destination.

4 FEASIBILITY STUDY

Our goal is to provide deployable privacy services for today's Internet, and therefore we build on well-established practices. Although this ensures interoperability with today's protocols, there are open problems when deploying at a large scale. We address such challenges and provide a preliminary evaluation to assess the feasibility of our proposal.

4.1 ISP-level Address Hiding

We propose ISPs to shuffle IP addresses of hosts subscribed to the AHS. When the ISP receives an outgoing packet, a *translation gateway* creates a mapping from the packet's source IP and port number to a new IP address and port number. This new IP address belongs to one of the ISP's address blocks; the ISP can use the bits in the IP's host portion together with the bits of the source port in order to multiplex different flows behind a few IP addresses. When the ISP receives an incoming packet, the gateway performs the reverse translation so as not to break bidirectional communication. This process may remind one of a carrier-grade NAT [29, 30], yet there are multiple challenges that we address in the following.

Coordination of Translation Gateways. An ISP will operate multiple translation gateways, and they must all perform identical translations. First, translations must be performed on the original data path to minimize latency overhead; rerouting traffic to a centralized location would cause a latency inflation. Therefore, an ISP operates multiple gateways, but under the constraint of identical translations. This is to prevent two different inputs from generating a same output. Also, the translation of an outgoing packet and the reverse translation for the incoming packet may be performed by different gateways due to asymmetric routing.

One naive approach to achieve identical translations is to exchange mapping tables between gateways. However, this solution is not viable since such mappings must be distributed for every new flow and to all gateways. We leverage cryptography to perform the translation without keeping per-flow state [15]: translation gateways encrypt the source address and port tuple and generate a new tuple, under the constraint that the network prefix belongs to the ISP. All gateways share the same key so that all (reverse) translations are identical; the state stored at every gateway is just the encryption key. This approach satisfies all the constraints mentioned.

Privacy vs. Traffic Engineering. The privacy benefits of AHS result from ISP's large IP blocks, which can provide a satisfying level of anonymity. However, ISPs deaggregate these address blocks in their BGP announcements to perform fine-granular traffic engineering. Although this practice still allows address shuffling, it reduces the anonymity set to the size of an advertised block.

We cannot eliminate this conflict, but we can provide flexibility to ISPs in picking the desired tradeoff. Therefore, we design a translation mechanism that enables shuffling for prefix blocks of arbitrary size. This task raises a challenge, since most cryptographic primitives operate on input of fixed length, e.g., 128-bit block for AES and 32-/64-/128-bit blocks for RC5. Furthermore, we need a secure cryptographic primitive that prevents an adversary to infer the original addresses by observing the translated addresses.

Our translation scheme is based on FF3 encryption, which is an instance of format-preserving encryption [31]. Equation 1 shows the performed translation, with k being a secret key known only to the ISP. This scheme generates a new source address and port tuple ($saddr'$, $sport'$) for every new flow. When the corresponding

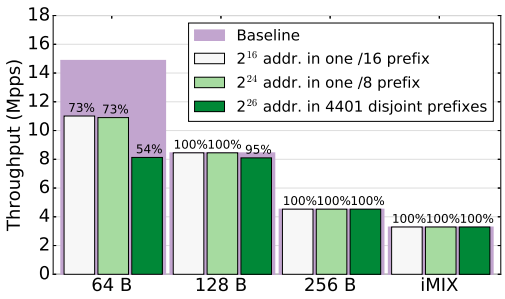


Figure 3: Forwarding performance of a translation gateway.

incoming packet arrives, the translation gateway performs the decryption, and then XORs the result with the source address and port of the incoming packet.

$$(saddr', sport') = FF3_k((saddr, sport) \oplus (daddr, dport)) \quad (1)$$

We utilize FF3 because it provides the required flexibility: it encrypts a plaintext of some format and length into a ciphertext of identical format and length, allowing us to shuffle variable-length address blocks. Furthermore, FF3 provides the security guarantees of conventional block ciphers [32] and is approved by NIST. Lastly, FF3 is efficient; it is based on AES as the underlying block cipher, which is implemented in hardware even on commodity CPUs.

Processing Overhead. We quantify the processing overhead of our flexible address-translation scheme. We have implemented the AHS service on the Data Plane Development Kit (DPDK) [33], running on a commodity server equipped with a 10 Gbps NIC and an Intel XEON E5 CPU. We evaluate three cases of shuffling a varying number of addresses: i) 2^{16} addresses in one /16 prefix, ii) 2^{24} addresses in one /8 prefix, and iii) 2^{26} addresses of the 4401 disjoint prefixes of an ISP (AS 4130). Furthermore, we evaluate forwarding performance for multiple packet sizes and for a representative mix of Internet traffic (iMIX [34] with 340 bytes avg. size); the baseline for our measurements is the performance of typical IP forwarding without additional processing.

Figure 3 shows the forwarding performance. For 64-byte packets, performance degrades by about 25% for shuffling an /8 and a /16 prefix. However, the decline is higher for the 2^{26} addresses because the blocks are disjoint and must be linearized. For larger packet sizes almost all cases perform optimally. The evaluation shows the efficiency of a single translation gateway—it can handle at line-rate a fully-saturated 10 Gbps link with typical Internet traffic patterns.

4.2 ISP-level Secure Tunnels

We propose that ISPs set up pairwise encrypted tunnels using existing protocols, such as IPsec in tunnel mode [35] and the Resource Public-Key Infrastructure (RPKI) [36]. IPsec in tunnel mode creates a virtual private network between two remote networks: each network deploys an IPsec gateway and the exchanged traffic is tunneled as ciphertext between the gateways. The RPKI consists of publicly accessible repositories of resource certificates, which prove that an AS is the owner of a resource (e.g., an IP prefix).

Although we leverage existing technologies, there are multiple challenges with respect to scalability, performance, and security. First, since there are thousands of ASes, it is hard to manually configure pairwise tunnels at a large scale. Then, how can we automate tunnel establishment so that key negotiations are performed dynamically?

Second, given the volume of traffic that ISPs forward, it is a considerable overhead to encrypt/decrypt even a fraction of an ISP's traffic. Furthermore, using a single IPsec gateway to serve all traffic of all established tunnels with other ISPs is not possible. Therefore, we describe an intra-domain architecture that is capable of supporting ISP-level tunnels.

4.2.1 Dynamic Tunnel Configuration. ISPs interested in deploying pairwise tunnels could initially coordinate and exchange required information manually. However, this approach does not scale as more and more ISPs start offering the service. To support automated tunnel configuration, two steps are needed: i) discover deploying peers, and ii) establish a security association with the peer.

Peer Discovery. The first step is to enable ASes to advertise support for the ISP-level tunnels to other ASes. To this end, we leverage resource certificates and RPKI to disseminate the additional information needed for ISP-level tunnels.

A resource certificate verifies that an AS owns a certain prefix and is therefore authorized to make a BGP announcement for that prefix [37]. Typically, a resource certificate contains the AS number, the public key of the AS, and a list of prefixes that the AS owns; the certificate is signed by the private key of a regional or a local Internet registry. We augment such certificates with additional information for the tunnel establishment; such information includes cryptographic material for key exchanges, security parameters, and addresses of the IPsec gateways. Presence of this information indicates that the ISP supports the tunneling service.

The process of peer discovery is then performed as part of route-origin authorization [38]: when ASes receive BGP announcements from their peers, they consult RPKI and verify that the advertised prefix is valid through the corresponding resource certificate. At this stage ASes can also learn if the advertised AS supports tunnels.

Key Exchange. The second step for automated tunnel configuration is to establish a security association (SA). A SA is a structure that contains all necessary parameters for the corresponding tunnel: the tunnel end-point addresses, the encryption and authentication algorithms, and the symmetric key used for encryption/decryption.

In order to set up a SA, IKEv2 [39] can be used. The protocol can perform mutual authentication between two IPsec gateways based on X.509 certificates; it also negotiates the security parameters for a SA. To facilitate authentication and key exchange, we leverage again RPKI and the resource certificates: the published X.509 resource certificate contains a public value that is used to perform an authenticated Diffie-Hellman key exchange and derive the symmetric key.

4.2.2 Intra-domain Architecture. In addition to mechanisms for setting up tunnels, an ISP will need mechanisms inside its network to support encryption/decryption of tunneled traffic. We benchmark the operation of a single IPsec gateway, and we describe how to support multiple gateways with a low management overhead.

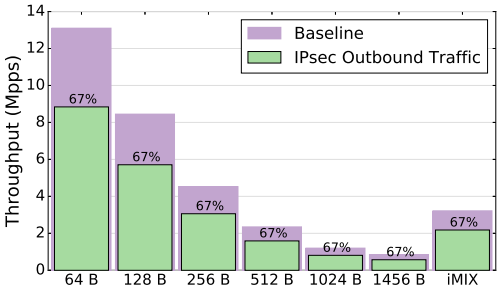


Figure 4: Forwarding performance of an IPsec gateway.

Gateway Evaluation. A deploying ISP may have to tunnel a considerable portion of traffic if many customers opt for the service. To put the processing overhead into context, we benchmark the performance of a gateway assuming it has a tunnel with every AS.

We use DPDK and we run an IPsec gateway on the same commodity server. Then, we analyze a BGP routing information base[2] to generate realistic SAs. More precisely, we generate 56,223 SAs—one for every AS—and we construct over 200k routing policies so that each packet is forwarded over the correct tunnel based on the destination address. For encryption and authentication we use the AES block cipher in GCM mode with 128-bit keys [40].

Figure 4 shows the forwarding performance of the gateway for the outbound direction and for multiple packet sizes, including the iMIX (inbound direction has the identical performance); the performance baseline is forwarding without any additional processing. The results show that performance is at approximately 67% of the baseline for all cases. This constant performance decline is observed because as we increase the packet length the packet rate decreases for a fully utilized link. At the same time, the length of the plaintext to be encrypted increases so that these two factors cancel each other out. The experiment shows the worst case in that we have established the maximum number of tunnels. However, the result also indicates that ISPs may have to expand their infrastructure if many users opt for the service.

Supporting Multiple Gateways. An ISP will need multiple gateways to support the traffic demands. However, this raises a challenge if the gateways of the two ISPs have to maintain a mesh of tunnels between them.

We introduce the concept of a *logical* IPsec gateway which is decoupled from the underlying *physical* IPsec gateways. That is, an ISP advertises a single tunnel end point and only one IPsec tunnel is established per remote ISP. Then, the state of the tunnel is shared between all physical IPsec gateways so that they can perform identical operations.

We make the following design choices to realize a logical IPsec gateway. First, we define a *designated* IPsec gateway[3] that represents all physical IPsec gateways in an ISP and is identified by an IP address within the ISP's prefix blocks; the same address that is published in the resource certificate. The designated gateway establishes the ISP-level tunnels with the designated gateways of other ISPs, and then it disseminates all necessary state to the physical IPsec gateways of the ISP. Second, we leverage IP anycast so that a single IP address is used for all physical IPsec gateways. This enables ISPs to support a tunnel end point from multiple locations and at the same time perform load-balancing among these locations by adjusting their intra-domain routing protocols.

[2]http://data.ris.ripe.net/rrc06/2017.04/

[3]Similar to the designated routers in OSPF [41] and the IS-IS [42] protocols.

4.3 ISP-level VPN

We propose that ISPs leverage their infrastructure to offer VPN services to customers of other ISPs.

As mentioned in Section 3.2.3, a VPN service consists of two main functionalities: 1) an encrypted tunnel that transfers packets between a VPN customer and a VPN gateway, and 2) address translation that replaces the customer's IP address with an IP owned by the VPN provider.

We have already addressed the challenges of the two functionalities, albeit in a different context. In Section 4.2, we have described how to set up ISP-level tunnels using IPsec; the same approach can be used for the tunnel between a VPN customer and the gateway. There are various implementations of VPN tunnels (e.g., VPN over IPsec and VPN over SSL), but they are conceptually the same. Furthermore, authentication, key exchange and tunnel configuration is typically performed through an application that is provided by the VPN provider and is installed on the customer's device.

In Section 4.1, we have shown how to perform ISP-level address translation. A VPN provider can use the same approach to multiplex multiple customers behind the source-port number and the host portion of its IP address block.

5 COMPOSITION OF PRIVACY SERVICES

In Section 3, we described how three building blocks can offer meaningful privacy services. Here we show three examples (Figure 5) of how proposed services can be composed to create additional privacy services.

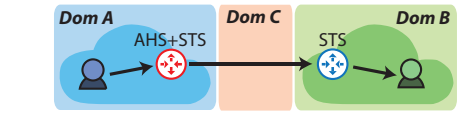
5.1 AHS+STS Composition

Consider a host that subscribes to the address-hiding service and the secure tunneling service of its ISP. This combination provides additional privacy benefits that are not offered by these services in isolation by enabling: i) to hide also the host's address from the destination ISP and host (not offered by secure tunneling), and ii) to hide the transport-layer header and payload from all transit ISPs (not offered by address hiding).

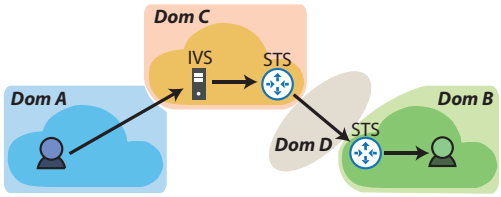
Privacy Domains. Three privacy domains are formed (Figure 5a). Domain *A* consists of the source ISP, and the host shares all its information. Domain *B* consists of the destination host and ISP, and the host shares all its information except for the source address, which is instead revealed at the AS granularity. Domain *C* consists of all transit ISPs, and the host shares the source and destination addresses at the AS granularity.

5.2 IVS+STS Composition

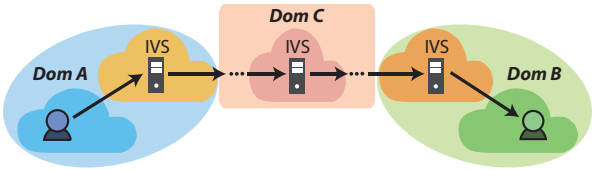
Another meaningful service composition is the following: a host buys the VPN service from a remote ISP and also the secure tunneling service from that ISP. Again this combination offers privacy benefits that cannot be achieved by a single service: i) the transport-layer header and payload are hidden from the host's ISP



(a) AHS+STS composition (Section 5.1).



(b) IVS+STS composition (Section 5.2).



(c) Seq of IVS composition (Section 5.3).

Figure 5: Three examples of service compositions. For simplicity, we do not show transit ISPs; they would exist on inter-cloud arrows.

(not offered by secure tunneling), and ii) the host hides its payload from all transit ISPs between the VPN provider and the destination ISP and host (not offered by the VPN service).

Privacy Domains. It creates the following four privacy domains (Figure 5b). Domain *A* consists of the source ISP, and the transit ISPs to the remote VPN ISP, and the host shares only his address with this domain. Domain *B* consists of the destination host and ISP; with this domain, the host shares the payload, the transport header, and the destination address. Domain *C* consists of the remote VPN ISP, and the host shares all information. Finally, Domain *D* consists of the transit ISPs between the remote VPN ISP and the destination ISP; with this domain, the host shares only the destination address at the AS granularity.

5.3 A Sequence of IVSes

A Tor-esque privacy service can be created by using a sequence of IVSes, where each IVS acts as a Tor relay. Specifically, a host creates an onion by encrypting his message in layers and forwards through a series of IVSes, each of which peels off an encryption layer and forwards to the next IVS.

Using a sequence of IVSes allows users to find a right balance between anonymity and performance, as studied by the past anonymous routing proposals [43–46]. For instance, a user could use the IVSes offered by the ISPs on the shortest path between the source and the destination ISPs to construct a Tor-esque privacy service that has low latency.

Privacy Domains. This privacy service creates the same privacy domains (Figure 5c) that are described for Tor in Section 2.

Challenge. We envision that an IVS is a paid service that requires explicit subscription; hence, a client needs to authenticate to all IVSes that it uses. This authentication leads to a privacy implication: all IVS-offering ISPs that are part of the Tor-esque privacy service can identify the sender based on client's authentication credentials. In particular, the last hop IVS-offering ISP (equivalent to the exit relay in Tor) can identify the source and the destination at the same time.

One possible approach to solve this problem is based on anonymous credentials. In this approach, ISPs could form an alliance and allow a host who subscribes to any member ISP to use the privacy services offered by all member ISPs. The alliance provides hosts with anonymous credentials that prove that the host is authorized to use the privacy services without revealing any further information about the host. Then, hosts use their anonymous credentials to authenticate to the IVSes.

6 DISCUSSION

6.1 Deployment Incentives

We believe that our proposed services are in line with market incentives of today's ISPs. A major deployment hurdle for many proposals is the lack of incremental deployability so that the proposals are valuable only if all, or almost all, ISPs adopt (e.g., BGPsec [47]); there are no benefits for first-movers, leading to a chicken and egg problem. This is not the case for two of our three proposed services: the address-hiding service and the VPN service can be offered to end users without requiring global adoption of new protocols. Users could buy these services directly from their ISPs, and the ISPs would offer these services independently and without coordination with other ISPs.

The secure-tunneling service has a higher deployment barrier in that it requires coordination among the interested ISPs. However, setting up a tunnel requires coordination only between two ISPs—not universal coordination; and, there are two ways for ISPs to coordinate with each other. 1) An ISP can negotiate bilateral contracts with other ISPs with which its customers frequently communicate, similar to roaming agreements between telecommunication providers. 2) We can envision ISP alliances, similar to Global Telco Security Alliance [48, 49] and Ngena [50, 51] where a group of ISPs with common interests come together and cooperate towards their common goal; within the alliance, a member ISP can establish tunnels with all other member ISPs without any further negotiation. In fact, a similar business model is seen at Internet Exchange Points (IXPs) that offer multi-lateral peering agreements: ISPs in the IXP peer with all other ISPs without separate negotiations;[4] this peering strategy is becoming increasingly popular due to its simplicity by avoiding excessive negotiations among ISPs.

In addition, we believe there are strong incentives for ISPs to adopt the secure-tunneling service: on one hand, large content providers are concerned about large-scale surveillance that degrades their customers' privacy. On the other hand, residential ISPs can offer value-added services to their customers by setting up such tunnels with large content providers.

[4] http://www.openpeering.nl/publicpeering.shtml.

6.2 Implications based on Government Policies

Different regulations and policies of different countries can have impact on how the ISPs treat users' privacy-sensitive data. For example, in the U.S., ISPs can now sell their customer's history without their consent [4]. Also, some governments pressure ISPs to release user data, e.g., to tackle illegal activities on the Internet.

Privacy domains cannot express such policy details, since they are technical specifications on disclosure of privacy-sensitive information. We believe that supplementary information, e.g., potential privacy risks due to policies of countries from which privacy services are used, should be provided to users when they select the privacy services for their communication.

6.3 Trusting ISP Operations

Given the diminishing confidence in ISPs regarding how they treat their customer's privacy, users may desire assurance that ISPs do provide the privacy services that they sell. The simplicity of obtaining assurances depends on the privacy service. For the ISP-level VPN service (IVS) and the ISP-level Address Hiding Service (AHS), an interested user can contact lookup services that provide their IP address information (e.g., whatismyip.com) to check if her address is properly altered.

Unfortunately, however, obtaining assurance for the ISP-level secure tunneling service (STS) is more difficult, since there is no direct way for users to check. This is because, unlike the other two services, users' packets would be delivered unmodified to the destinations. Instead, users need to rely on a third-party to obtain assurance. For example, governments that monitor ISP practices can provide the assurance to the users. Alternately, there can be monitoring infrastructures, operated by non-profit organizations, from which ISP operations can be monitored, e.g., traffic have been encrypted and tunneled.

7 RELATED WORK

The most closely related work to ours is the proposal from Raghavan et al. that suggests a tweakable block cipher for address shuffling by ISPs [15]. We have extended their initial idea by proposing ISPs to offer multiple privacy services with additional privacy benefits. Moreover, from a technical perspective, our address-hiding service provides a higher degree of flexibility to ISPs: our shuffling method works for IP address blocks of arbitrary length, compared to the tweakable block cipher that requires exactly a /16 IP address block. In a similar fashion, when using carrier-grade NATs (CGNs), an ISP shares a few public IP addresses with a large customer base [29, 30]. CGNs, however, require per-flow state and therefore suffer from scalability issues.

8 CONCLUSION & FUTURE WORK

Our final goal is to embrace user-defined privacy through a privacy-as-a-service architecture (Figure 6). In this architecture, the user specifies privacy and performance requirements at a high level; then, the architecture translates them to technical requirements using privacy domains, and chooses among a set of ISP-offered privacy services. These privacy services should be readily deployable in today's Internet, and the corresponding privacy service providers should be able to offer these services in a cost-efficient manner.

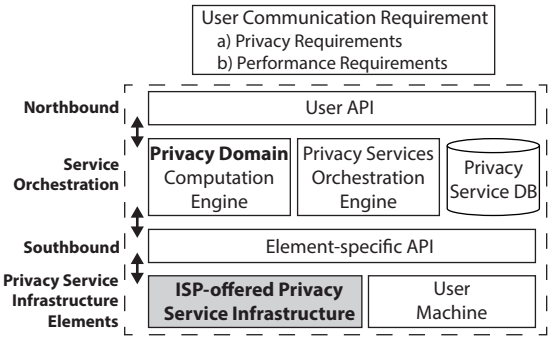


Figure 6: Privacy-as-a-Service Architecture.

In this paper, we make a first step towards our goal (highlighted element in Figure 6). We have introduced a new concept—privacy domains—in order to capture the diverse privacy requirements. We propose three privacy services that can be offered by ISPs: an address-hiding service, a secure tunneling service, and a VPN service. All services can be easily deployed and our initial evaluation indicates that the overhead is within reach for the existing infrastructure.

9 SOURCE CODE AND DATA RELEASE

We make all source code and data used for the experiments at https://github.com/kthlee86/privacy-domain. For more details about the source code and how to run the experiments, we direct the readers to the README.md file in the repository.

10 ACKNOWLEDGMENT

We would like to thank our shepherd, Barath Raghavan, and the anonymous reviewers for their insightful feedback. Also, we would like to thank Woojik Chun for the initial discussions and motivation that led to this paper. The research leading to these results has received funding from the European Research Council under the European Union's Seventh Framework Programme (FP7/2007-2013)/ERC grant agreement 617605; from the IITP grant funded by the Korean government (MSIT) (No.2016-0-00157, Development of self-certifying ID based trustworthy networking technology). We also gratefully acknowledge support by ETH Zürich and the Zürich Information Security Center (ZISC), and by Intel for their equipment donation that enabled the high-capacity experiments.

REFERENCES

[1] A. Rouvroy and Y. Poullet, "The Right to Informational Self-Determination and the Value of Self-Development," in *Reinventing Data Protection?* Springer Netherlands, 2009, ch. 2, pp. 45–76.
[2] A. F. Westin, *Privacy and Freedom.* Scribner, 1967.
[3] S. Fischer-Hübner, C. Hoofnagle, K. Rannenberg, M. Waidner, I. Krontiris, and M. Marhöfer, "Online Privacy: Towards Informational Self-Determination on the Internet (Dagstuhl Perspectives Workshop 11061)," Dagstuhl, Tech. Rep. 2, 2011.
[4] "Senate Votes to Let ISPs Sell Your Web Browsing History to Advertisers," http://bit.ly/2nNetnc, 2017.
[5] "NSA Collecting Phone Records of Millions of Verizon Customers Daily," http://bit.ly/2brf9H0, 2013.
[6] C. Diaz, S. Seys, J. Claessens, and B. Preneel, "Towards Measuring Anonymity," in *Proceedings of the Workshop on Designing Privacy Enhancing Technologies (PETS)*, 2003.
[7] A. Serjantov and G. Danezis, "Towards an Information Theoretic Metric for Anonymity," in *Proceedings of the Workshop on Designing Privacy Enhancing Technologies (PETS)*, 2003.

[8] R. Dingledine, N. Mathewson, and P. Syverson, "Tor: The Second-generation Onion Router," in *Proceedings of the USENIX Security Symposium*, 2004.

[9] Y. Zhu, X. Fu, B. Graham, R. Bettati, and W. Zhao, "On Flow Correlation Attacks and Countermeasures in Mix Networks," in *Proceedings of the Workshop on Designing Privacy Enhancing Technologies (PETS)*, 2004.

[10] X. Wang, S. Chen, and S. Jajodia, "Tracking Anonymous Peer-to-Peer VoIP Calls on the Internet," in *Proceedings of the ACM Conference on Computer & Communications Security (CCS)*, 2005.

[11] V. Shmatikov and M.-H. Wang, "Timing Analysis in Low-latency Mix Networks: Attacks and Defenses," in *Proceedings of the European Symposium on Research in Computer Security (ESORICS)*, 2006.

[12] P. Syverson, G. Tsudik, M. Reed, and C. Landwehr, "Towards an Analysis of Onion Routing Security," in *Proceedings of the Workshop on Designing Privacy Enhancing Technologies (PETS)*, 2001.

[13] Y. Sun, A. Edmundson, L. Vanbever, O. Li, J. Rexford, M. Chiang, and P. Mittal, "RAPTOR: Routing Attacks on Privacy in Tor," in *Proceedings of the USENIX Security Symposium*, 2015.

[14] S. W. L. Meiser, "Quantitative Anonymity Guarantees for Tor," Ph.D. dissertation, Saarland University, 2016.

[15] B. Raghavan, T. Kohno, A. C. Snoeren, and W. David, "Enlisting ISPs to improve online privacy: IP Address Mixing by Default," in *Proceedings of the Privacy Enhancing Technologies Symposium (PETS)*, 2009.

[16] "The Heartbleed Bug," heartbleed.com, 2014.

[17] "CRIME Attack Uses Compression Ratio of TLS Requests as Side Channel to Hijack Secure Sessions," http://bit.ly/2rn7QbT, 2012.

[18] "A Perfect CRIME? Only TIME Will Tell," http://ubm.io/2rn5qtU, 2013.

[19] "The BREACH Attack," http://www.breachattack.com/, 2013.

[20] "SSL Pulse," https://www.trustworthyinternet.org/ssl-pulse/, 2017.

[21] "SSL by Default Usage Statistics," http://bit.ly/2rnj1S7, 2017.

[22] A. Johnson, C. Wacek, R. Jansen, M. Sherr, and P. Syverson, "Users Get Routed: Traffic Correlation on Tor by Realistic Adversaries," in *Proceedings of the ACM Conference on Computer & Communications Security (CCS)*, 2013.

[23] "Russian-controlled Telecom Hijacks Financial Services' Internet Traffic," http://bit.ly/2pp44fp, 2017.

[24] "Iran's Porn Censorship Broke Browsers as Far Away as Hong Kong," http://bit.ly/2s4E90s, 2017.

[25] "BackConnect's Suspicious BGP Hijacks," http://bit.ly/2rnhBa8, 2016.

[26] "Large Hijack Affects Reachability of High Traffic Destinations," http://bit.ly/2qMW6Rm, 2016.

[27] H. Yi, "This is how Internet speed and price in the U.S. compares to the rest of the world," http://to.pbs.org/1z5AGNQ, April 2015.

[28] V. C. Perta, M. V. Barbera, G. Tyson, H. Haddadi, and A. Mei, "A Glance through the VPN Looking Glass: IPv6 Leakage and DNS Hijacking in Commercial VPN Clients," in *Proceedings on Privacy Enhancing Technologies (PoPETs)*, 2015.

[29] S. Jiang, D. Guo, and B. Carpenter, "An Incremental Carrier-Grade NAT (CGN) for IPv6 Transition," RFC 6264 (Informational), IETF, Jun. 2011. [Online]. Available: http://www.ietf.org/rfc/rfc6264.txt

[30] I. Yamagata, Y. Shirasaki, A. Nakagawa, J. Yamaguchi, and H. Ashida, "NAT444," https://tools.ietf.org/id/draft-shirasaki-nat444-06.txt, 2012.

[31] M. Bellare, T. Ristenpart, P. Rogaway, and T. Stegers, "Format-Preserving Encryption," in *Proceedings of the Workshop on Selected Areas in Cryptography*, 2009.

[32] M. Dworkin, "Recommendation for Block Cipher Modes of Operation: Methods for Format-Preserving Encryption," *NIST Special Publication*, vol. 800, 2016.

[33] "Data Plane Development Kit," http://dpdk.org, Sep 2015, retrieved on 1/2016.

[34] A. Morton, "IMIX Genome: Specification of Variable Packet Sizes for Additional Testing," RFC 6985 (Informational), IETF, Jul. 2013. [Online]. Available: http://www.ietf.org/rfc/rfc6985.txt

[35] S. Kent and K. Seo, "Security Architecture for the Internet Protocol," RFC 4301, IETF, 2005.

[36] R. Bush and R. Austein, "The Resource Public Key Infrastructure (RPKI) to Router Protocol," RFC 6810, IETF, 2013.

[37] G. Huston, G. Michaelson, and R. Loomans, "A Profile for X.509 PKIX Resource Certificates," RFC 6487, IETF, 2012.

[38] M. Lepinski, S. Kent, and D. Kong, "A Profile for Route Origin Authorizations (ROAs)," RFC 6482, IETF, 2012.

[39] C. Kaufman, P. Hoffman, Y. Nir, P. Eronen, and T. Kivinen, "Internet Key Exchange Protocol Version 2 (IKEv2)," RFC 7296, IETF, 2014.

[40] D. A. McGrew and J. Viega, "The Galois/Counter Mode of Operation (GCM)," http://goo.gl/9sl9kK, 2004.

[41] J. Moy, "OSPF Version 2," RFC 2328 (INTERNET STANDARD), IETF, Apr. 1998, updated by RFCs 5709, 6549, 6845, 6860, 7474. [Online]. Available: http://www.ietf.org/rfc/rfc2328.txt

[42] ISO, "Intermediate System to Intermediate System intra-domain routing information exchange protocol for use in conjunction with th protocol for providing the connectionless-mode network service (ISO 8473)," *International Standard*, vol. 10589, 2002.

[43] C. Chen, D. E. Asoni, D. Barrera, G. Danezis, and A. Perrig, "HORNET: High-speed Onion Routing at the Network Layer," in *Proceedings of the ACM Conference on Computer & Communications Security (CCS)*, 2015.

[44] M. Akhoondi, C. Yu, and H. V. Madhyastha, "LASTor: a low-latency AS-aware tor client," *Transactions on Networking (TON), IEEE/ACM*, 2014.

[45] H.-C. Hsiao, T. H.-J. Kim, A. Perrig, A. Yamada, S. C. Nelson, M. Gruteser, and W. Meng, "LAP: Lightweight Anonymity and Privacy," in *Proceedings of the IEEE Symposium on Security and Privacy (S&P)*, 2012.

[46] V. Liu, S. Han, A. Krishnamurthy, and T. Anderson, "Tor Instead of IP," in *Proceedings of the ACM Workshop on Hot Topics in Networks (HotNets)*, 2011.

[47] R. Lychev, S. Goldberg, and M. Schapira, "BGP Security in Partial Deployment: Is the Juice Worth the Squeeze?" in *Proceedings of the ACM Conference on SIGCOMM*, 2013.

[48] "Etisalat, Singtel, Softbank, and Telefonica create Global Cyber Security Alliance," https://goo.gl/tEhM1H, 2018, retrieved on 6/2018.

[49] E. Yu, "Four telcos set up global cybersecurity group," goo.gl/GaqiHd, 2018, retrieved on 6/2018.

[50] "ngena," www.ngena.net, 2018, retrieved on 6/2018.

[51] K. Compton, "ngena: Creating Next-Gen Global Business Networks," https://blogs.cisco.com/digital/ngena, 2018, retrieved on 6/2018.