The SCION Internet Architecture

An Internet Architecture for the 21st Century

David Barrera, Laurent Chuat, Adrian Perrig, Raphael M. Reischuk, Pawel Szalachowski

Network Security Group, ETH Zurich, Switzerland

March 6, 2017

1. INTRODUCTION

The Internet has been successful beyond even the most optimistic expectations. It permeates and intertwines with almost all aspects of our society and economy. The success of the Internet has created a dependency on communication as many of the processes underpinning the foundations of modern society would grind to a halt should communication become unavailable. However, much to our dismay, the current state of safety and availability of the Internet is far from commensurate with its importance.

Although we cannot conclusively determine what the impact of a 1-minute, 1-hour, 1-day, or 1-week outage of Internet connectivity on our society would be, anecdotal evidence indicates that even short outages have a profound negative impact on governmental, economic, and societal operations [11]. To make matters worse, the Internet has not been primarily designed for high availability in the face of malicious actions by adversaries. Recent patches to improve Internet security and availability have been constrained by the design of the current Internet architecture. A new Internet architecture should offer availability, security by design, provide incentives for deployment, and consider economic, political, and legal issues at the design stage.

We believe addressing these issues requires a new cohesive architecture that provides one fundamental building block: highly available point-to-point communication, which other proposed future Internet architectures that provide content-centric [9, 15], extensibility-centric [14], or mobility-centric [23] properties could build upon.

This article describes **SCION** (**S**calability, **C**ontrol, and **I**solation **O**n **N**ext-generation networks), an inter-domain network architecture designed to address these issues. We discuss SCION's goals, design, and functionality, as well as the results of 5 years of research conducted since the initial publication [29].

Key Insights

- **Clean-slate vs. ad-hoc:** Patching today's Internet has proven to be an undesirable long-term solution. A clean-slate redesign of inter-domain routing provides many benefits and is surprisingly simple to deploy by using legacy protocols for intra-domain communication.
- **Isolation yet transparency:** SCION's Isolation Domains (ISDs) offer control-plane isolation and scoped trust for end-entity authentication. Instead of restricting communication, ISDs provide transparency for path selection, packet forwarding, and end-entity authentication.
- **End-to-end focus:** Network-layer functions are performed by end hosts through path selection; providing scalability, security, and availability benefits.
- **Packet-carried forwarding state:** PCFS removes the need for inter-domain routing table lookups, provides path control to senders, improves forwarding performance, and supports multipath communication.

2. OBJECTIVES

In this section, we present high-level goals that an inter-domain point-to-point communication architecture should accomplish.

Availability in the presence of adversaries. Our overarching goal is to offer a point-to-point communication infrastructure that remains highly available even in the presence of distributed adversaries: as long as an attacker-free path between endpoints exists, that path can be discovered and used with guaranteed bandwidth between these endpoints, which is an exceedingly challenging property to achieve.

An *on-path adversary* may drop, delay, or alter packets instead of forwarding them, or inject packets into the network. The architecture hence needs to provide mechanisms to counteract malicious operations. An *off-path adversary* could launch hijack attacks to attract traffic to flow through network elements under its control. Such traffic attraction can take various forms; for instance, an adversary could announce a desirable path to a destination by using forged paths or attractive network metrics. Conversely, the adversary could render paths not traversing its network less desirable (e.g., by inducing congestion). An adversary controlling a large botnet could also perform distributed denial-of-service (DDoS) attacks, congesting selected network links. Finally, an adversary could interfere with the discovery of legitimate paths (e.g., by announcing bogus paths).

Transparency and control. When the network offers *path transparency*, end hosts know (and can verify) the forwarding path taken by network packets. Applications that transmit sensitive data can benefit from this property, as packets can be ensured to traverse certain Internet service providers (ISPs) and avoid others.

In addition to path transparency, we aim to achieve end-host *path control*, a stronger property that allows receivers to select the incoming paths through which they are reachable, and that allows senders to select the end-to-end path. This seemingly benign requirement has various repercussions — beneficial but also fragile if implemented incorrectly.

The beneficial aspects of path control are: (a) *Separation of network control plane and data plane*, which ensures that forwarding cannot be retroactively influenced by control-plane operations (e.g., routing changes). (b) *Enabling of*

multipath communication, which improves availability by allowing senders to select multiple paths to their destinations. (c) *Defending against network attacks* such as DDoS and traffic interception by rogue networks, since destinations can observe the packet's traversed path in the packet header.
The fragile aspects that need to be handled with particular care are: (a) *Respecting ISPs' forwarding policies* by offering policy-compliant paths from which senders can choose. (b) *Preventing malicious path creation* such as paths that contain loops. (c) *Ensuring scalability of path control* by allowing sources to select paths among a relatively small set (as opposed to full-fledged source routing). (d) *Enabling ISP traffic engineering* despite end hosts' path control, to provide the ISP with the ability to balance their load across the links to their neighbor ASes.

Transparency and control over trust roots. Roots of trust are used for the verification of entities in today's Internet. For example, verification of the server's public key in a TLS certificate, or verification of a Domain Name System (DNS) response in DNSSEC [5]. Transparency of trust roots provides end hosts and users knowledge of the complete set of trust roots relied upon for entity certificate validation. Such enumeration of trust roots is difficult today due to intermediate certification authorities (CAs) that are implicitly trusted. Control over trust root selection enables *trust agility* [20], allowing *users* to easily select or exclude the roots of trust they want to rely upon.

Efficiency and scalability. Despite the lack of availability and transparency, today's Internet also suffers from a number of efficiency and scalability deficiencies: for instance, the Border Gateway Protocol (BGP) has scaling issues in cases of network fluctuations, where routing protocol convergence can take minutes [25] or even days [1]. Moreover, routing tables have reached the limits of their scalability due to multihoming and prefix de-aggregation (i.e., announcement of more specific IP address spaces). Increasing the memory size for routing tables is challenging as the underlying hardware is expensive and power-hungry, consuming around a third of the total power consumption of a router.

Security and high availability usually come at a cost, resulting in lower efficiency and potentially diminished scalability. High performance and scalability, however, are required for economic viability. We thus explicitly seek *high efficiency* such that packet forwarding latency and throughput are at least as fast as current IP forwarding. Moreover, we seek *improved scalability* compared to the current Internet, in particular with respect to BGP and to the growing size of routing tables.

One approach to achieve efficiency and scalability is to avoid router state wherever possible. We thus aim to place state into packet headers and protect that state cryptographically. Since modern block ciphers such as AES can be computed faster than performing DRAM memory lookups, packet-carried state can enable higher packet processing speeds and simpler router architectures compared to today's IP routers. Avoiding state on routers additionally prevents state-exhaustion attacks [27] and state inconsistencies across routers.

Our goal of efficiency and scalability is in line with the *end-to-end principle*, which states that a function should be implemented at the network layer in which it can most effectively operate [26]. Since the end host has the most information about its internal state, functions such as bit error recovery, duplicate suppression, or delivery acknowledgments are handled by the end host. Moreover, SCION end hosts are involved in path selection, as they have the knowledge of preferred or undesirable network paths. In other words, SCION adheres to the end-to-end principle even more than the current Internet.

Extensibility. To future-proof SCION, we design the core architecture and code base to be *extensible*, such that additional functionality can be easily built and deployed. SCION end hosts and routers should (without overhead or expensive protocol negotiations) discover the minimum common feature set supported by all intermediate nodes.

Support for global but heterogeneous trust. Given the diverse nature of constituents in the Internet of the 21st century with its diverse legal jurisdictions and interests, an important challenge is how to scale authentication of entities (e.g., AS ownership for routing, name servers for DNS, or domains for TLS) to a global environment. The roots of trust of currently prevalent PKI models (monopoly and oligopoly) do not scale to a global environment because mutually distrusting entities cannot agree on a single trust root (monopoly model), and because the security of a plethora of roots of trust is only as strong as its weakest link (oligopoly model). We thus seek a trust architecture that supports *meaningful trust roots in a global environment* with mutually distrusting entities.

Deployability. Incentives for deployment are important to overcome the resistance for upgrading today's core Internet infrastructure. A multitude of features is necessary to offer the initial impulse: high availability even under control-plane and data-plane attacks (e.g., built-in DDoS defenses), path transparency and control, trust root transparency and control, high efficiency, robustness to configuration errors, fast recovery from failures, high forwarding efficiency, multipath forwarding, etc. Economic and business incentives are also of critical importance; ISPs should be able to define new business models and sell new services.

Migration to the new architecture should require minimal added complexity (and cost) to the existing infrastructure. Deployment should be possible by utilizing the internal switching infrastructure of an ISP, and only require installation or upgrade of a few border routers. Moreover, configuration of the new architecture should be similar to the existing architecture, such as in the configuration of BGP policies, minimizing the amount of additional personnel training.

Foundation for other architectures. To achieve a simple, scalable, secure and efficient architecture, we focus on the most basic communication mode: point-to-point communication. Other architectures that provide support for higher-level properties, such as support for content distribution [9, 15], extensibility [14], or mobility [23], all require a working point-to-point communication infrastructure.

3. THE SCION ARCHITECTURE

SCION introduces the concept of *isolation domain* (ISD), which is a fundamental building block for achieving the properties of high availability, transparency, scalability, and support for heterogeneous trust. An ISD constitutes a logical grouping of *autonomous systems* (ASes), as illustrated in Figure 1.

An ISD is administered by multiple ASes, which form the *ISD core*. We refer to these ASes as core ASes. The ISD is

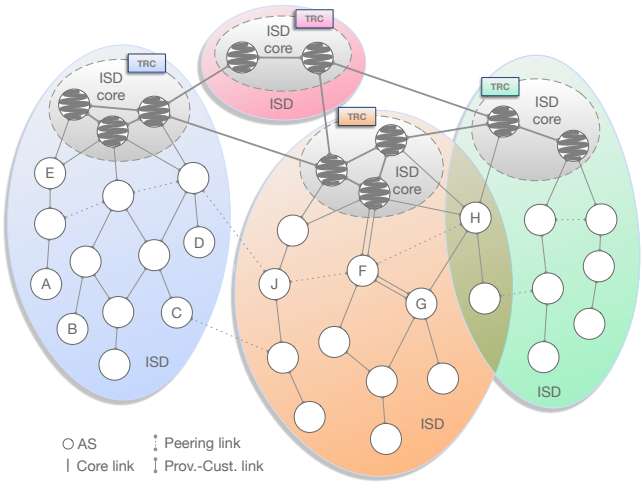

Figure 1: Autonomous systems (ASes) grouped into four ISDs. Core ASes are connected via core paths. Non-core ASes are connected via customer-to-provider or peering links. Some ASes are contained in multiple ISDs.

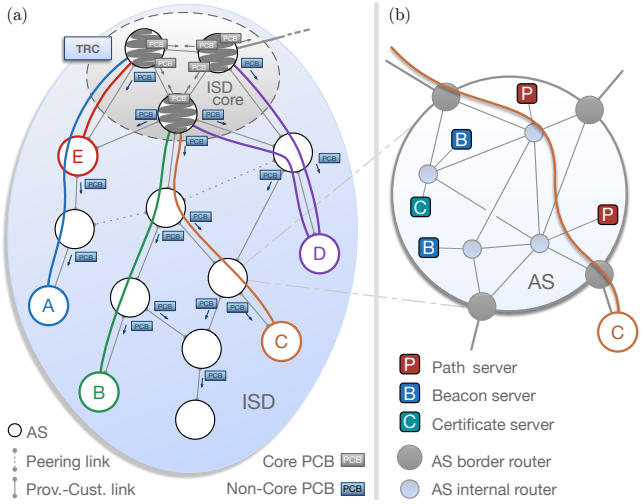

Figure 2: (a) SCION ISD with path construction beacons (PCBs) that are propagated from the ISD core down to customer ASes, and path segments for ASes A, B, C, D, and E to the ISD core. (b) Magnified view of an AS with its routers and servers. The path from AS C to the ISD core traverses two internal routers.

governed by a policy, called *trust root configuration* (TRC), which is negotiated by the ISD core. The TRC file defines the roots of trust used to validate bindings between names and public keys or addresses.

An AS joins an ISD by purchasing connectivity from another AS in the ISD. Joining an ISD constitutes an acceptance of the ISD's TRC file. We envision that ISDs will span areas with uniform legal environments that provide enforceable contracts. If two ISPs have a contract dispute they cannot resolve by themselves, such a legal environment can provide an external authority to resolve the dispute. All ASes within an ISD also agree on the TRC file, i.e., the entities that operate the trust roots and set the ISD policies. One possible model is thus for ISDs to be formed along national boundaries or federations of nations, as entities within a legal jurisdiction can enforce contracts and agree on a TRC file. ISDs can also overlap, so an AS may be part of several ISDs. Although an ISD ensures isolation from other networks, the central purpose of an ISD is to provide transparency and to support heterogeneous trust environments.

SCION uses two levels of routing, intra-ISD and inter-ISD. Both levels utilize *path-segment construction beacons* (PCBs) to explore routing paths (see Figure 2a).

A core AS announces a PCB and disseminates it as a policy-constrained multipath flood either *within* an ISD (to discover intra-ISD paths) or *among* core ASes (to discover inter-ISD paths). We refer to that process as *beaconing*. PCBs accumulate cryptographically protected AS-level path information as they traverse the network. These protected contents within received PCBs are chained together by sources to create a *path segment* that enables packets to traverse a sequence of ASes. Packets thus contain AS-level path information avoiding the need for border routers to maintain inter-domain routing tables. We refer to this concept as *packet-carried forwarding state* (PCFS).

Through beaconing, ASes learn paths between themselves and core ASes. Path registration allows ASes to turn a few selected beacons into path segments, and make them available to other ASes. Then, path resolution allows end hosts to create a *forwarding path* to the destination. This process consists of (a) *path lookup*, where an end host obtains path segments to the destination, and (b) *path combination*, where a forwarding path is created from the path segments.

3.1 Control Plane

The control plane is responsible for discovering paths and making those paths available to end hosts.

3.1.1 Servers and Routers

Figure 2b shows the main AS components that perform control-plane operations in SCION: *beacon servers* discover path information, *path servers* disseminate path information, and *certificate servers* assist with validating path information. In addition, *border routers* provide the connectivity between ASes, while *internal routers* forward packets inside the AS.

Beacon servers are responsible for disseminating PCBs (see Figure 2a). Beacon servers in a core AS generate intra-ISD PCBs that are sent to non-core ASes of the ISD. Non-core AS beacon servers receive such PCBs and re-send them to their customer ASes, which results in AS-level path segments. Figure 3 shows PCBs that are propagated from the ISD core down to customer ASes. At every AS, information about the AS's interfaces is added to the PCB. The beacon servers run a fault-tolerant protocol to ensure state consistency across all local servers. Periodically, a master beacon server generates a set of PCBs that it forwards to its customer ASes. In the case of inter-ISD communication, the beaconing process is similar to BGP's route advertising process, although the process is periodic and PCBs are flooded multipath over policy-compliant paths to discover multiple paths between any pair of core ASes. SCION's beacon servers can be configured to implement current BGP policies, as well as additional properties (e.g., control of upstream ASes) that BGP cannot express.

Path servers store mappings from AS identifiers to sets of such announced path segments, and are organized as a

hierarchical caching system similar to today's DNS. ASes, through the master beacon servers, select the set of path segments through which they want to be reached, and upload them to a path server in the ISD core.

Certificate servers store cached copies of TRCs retrieved from the ISD core, store cached copies of other ASes' certificates, and manage keys and certificates for securing intra-AS communication. Certificate servers are queried by beacon servers when validating the authenticity of PCBs.

Border routers connect different ASes supporting SCION. The main task of border routers is to forward packets. In the case of a control packet, the border router forwards it to the appropriate server, and in the case of a data packet the border router forwards it either to a host inside the AS or towards the next border router.

Since SCION can operate using any communication fabric inside an AS, the internal routers do not need to be changed.

3.1.2 Path Exploration and Registration

Through inter-domain beaconing, core ASes learn paths to other core ASes. Through intra-domain beaconing, ASes learn path segments leading to core ASes, which enable an AS to communicate with the ISD core. Figure 2a shows path segments from ASes A, B, C, D, and E to the core. The beaconing process is asynchronous, i.e., the PCB generation is local, based on a per-AS timer and PCBs are not propagated immediately upon arrival.

Paths are represented at AS-level granularity, which by itself is insufficient for fine-grained path diversity; ASes often have several diverse connection points, and thus a disjoint path is possible despite the AS sequence being identical. For this reason, SCION encodes AS ingress and egress interfaces as part of the path, exposing a finer level of path diversity. Figure 3 demonstrates this feature: AS F receives two different beacons via two different links from the core. Moreover AS F uses two different links to send two different beacons to AS G, each containing the respective egress interfaces. AS G extends the two beacons and forwards both of them over a single link to its customer.

An important requirement is that SCION also supports peering links between ASes. Consistent with AS policies in the current Internet, PCBs do not traverse peering links. However, peering links are announced along with a regular path in a PCB. Figure 3 shows how AS F includes its two peering links in the PCB. If the same peering link is announced in two path segments, then the peering link can be used to shortcut the end-to-end path (i.e., without going through the core). SCION also supports peering links that cross ISD boundaries, which highlights the importance of SCION's path transparency property; a source knows the exact set of ASes and ISDs traversed during the delivery of a packet.

An AS typically receives several PCBs representing several path segments to various core ASes. Figure 2a shows two path segments for AS D. We call a path segment that leads towards an ISD core an *up-segment*, and a path segment that leads from the ISD core to an AS a *down-segment* — although path segments are typically bi-directional and thus support packet forwarding in both directions. More precisely, up-segments and down-segments are invertible: by flipping the order, an up-segment is converted to a down-segment and vice versa. Path servers learn up-segments by extracting them from PCBs they obtain from the local beacon servers.

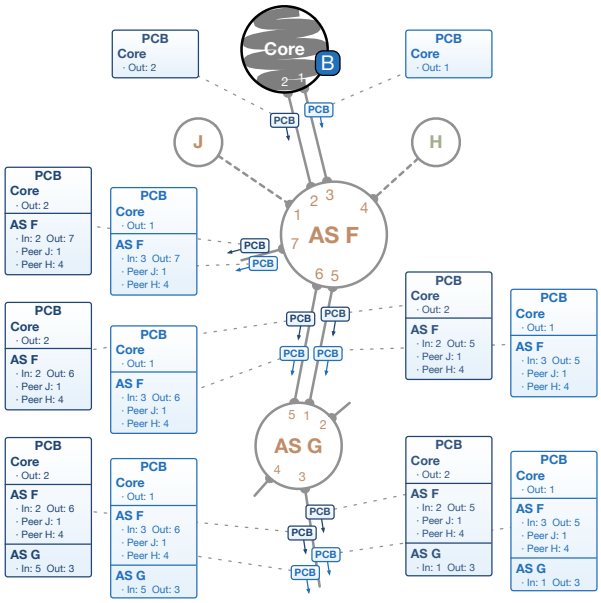

Figure 3: Intra-ISD PCB propagation from the ISD core down to customer ASes. For the sake of illustration, the interfaces of each AS are numbered with consecutive integer values. In practice, each AS can choose any encoding for its interfaces; only the AS itself needs to understand its encoding.

Path servers in core ASes also store *core-segments* to reach other core ASes.

The beacon servers in an AS select the down-segments through which the AS desires to be reached, and register these path segments at the core path servers. When links fail, segments expire, or better segments become available, the beacon servers keep updating the down-segments registered for their AS.

3.1.3 Path Lookup

To reach its ISD core, a host performs a path lookup at its local path server, fetching up-segments. To reach a remote destination, a host queries its path server for the down-segment of the destination AS. If the local path server has no cached entry for the down-segment, it will query the destination AS's core path server.

3.1.4 Beacon and Path-Segment Selection

The beacons to propagate and path segments to register are selected based on a path quality metric with the goal of identifying consistent, diverse, efficient, and policy-compliant paths. *Consistency* refers to the requirement that there exists at least one property along which the path is uniform, such as an AS capability (e.g., anonymous forwarding) or link property (e.g., low latency). *Diversity* refers to the set of paths that are announced over time being as path-disjoint as possible to provide high quality multipath options. *Efficiency* refers to the length, bandwidth, latency, utilization, and availability of a path, where more efficient paths are naturally preferred. *Policy compliance* refers to the requirement that the path adheres to the AS's routing policy. Based on past PCBs that were sent, a beacon server scores the current set of candidate path segments and sends the k best segments as the next PCB. SCION intra-ISD beaconing can scale to

networks of arbitrary size, because each inter-AS link carries the same number of PCBs regardless of the number of PCBs received by the AS.

Inter-ISD beaconing operates similarly to intra-ISD beaconing, except that inter-ISD PCBs only traverse ISD core ASes. The same path selection metrics apply, where an AS attempts to forward the set of most desirable paths to its neighbors. Similar to BGP, this process is inherently not scalable, however, as the number of ISDs and the corresponding number of core ASes is small, this approach is viable.

3.1.5 Link Failures

Unlike in the current Internet, link failures are not automatically resolved by the network, but require active handling by end hosts. Since SCION forwarding paths are static, they break when one of the links fails. Link failures are handled by a three-pronged approach that typically masks link failures without any outage to the application and rapidly re-establishes fresh working paths: (a) Beaconing occurs every few seconds, constantly establishing new working paths. (b) The *SCION control message protocol* (SCMP), a SCION-equivalent of ICMP, is used for link revocation. (c) SCION end hosts use multipath communication by default, thus masking link failures to an application with another working path. As multipath communication can increase availability (even in environments with very limited path choice [4]), SCION beacon servers actively attempt to create disjoint paths, SCION path servers make an effort to select and announce disjoint paths, and end hosts compose path segments to achieve maximum resilience to path failure. Consequently, we expect that most link failures in SCION will be unnoticed by the application, unlike the numerous short outages in the current Internet [16, 18].

3.1.6 Intra-AS Communication

Communication within ASes is handled by existing intra-domain communication protocols such as IP, Open Shortest Path First (OSPF), Multiprotocol Label Switching (MPLS), and Software-Defined Networking (SDN). Figure 2b shows one possible intra-domain path through the magnified AS.

3.2 Data Plane

While the control plane is responsible for providing end-to-end paths, the data plane ensures packet forwarding using the provided paths. A SCION packet minimally contains a path; source and destination addresses are optional in case the packet's context is unambiguous without addresses. Consequently, SCION border routers forward packets to the next AS based on the AS-level path in the packet header (which is augmented with ingress and egress interface identifiers for each AS), without inspecting the destination address and also without consulting a routing table. Only the border router at the destination AS needs to inspect the destination address or packet purpose to forward it to the appropriate local host(s).

An interesting aspect of this forwarding is enabled by the split of locator (the path towards the destination AS) and identifier (the destination address) [13]: since only the destination AS needs to consider the local identifier, the identifier can have any format the destination can interpret. Therefore, a domain can select an arbitrary addressing format for its hosts, e.g., a 4-byte IPv4, 6-byte medium access control, 16-byte IPv6, or 20-byte accountable IP (AIP [3]) address. A nice consequence is that an IPv4 host can directly communicate with an IPv6 host over SCION.

Routers can efficiently forward packets in the SCION architecture. In particular, the absence of inter-domain routing tables and the absence of complex longest prefix matching performed by current routers enables construction of faster and more energy-efficient routers. During forwarding, a border router would first verify that the packet entered through the correct ingress interface. If the packet has not yet reached the destination AS, the egress interface defines the next hop.

3.2.1 Path Combination

End-to-end communication is enabled by a combination of up to three path segments that form a SCION *forwarding path*. After path lookup, depending on the returned segments, a forwarding path can be created as follows:

- **Immediate combination of path segments** (e.g., $B \rightarrow D$ in Figure 2a): the last AS on the up-segment (ending at a core AS) is the same AS as the first AS on the down-segment (starting at a core AS). In this case, the simple combination of an up-segment and a down-segment creates a valid forwarding path.
- **Peering shortcut** (e.g., $A \rightarrow B$ in Figure 2a): a peering link exists between the two segments, so a shortcut via the peering link is possible. As in the *AS shortcut* case, the extraneous path segment is cut off. The peering link could be traversing to a different ISD.
- **AS shortcut** (e.g., $B \rightarrow C$ in Figure 2a): the up-segment and down-segment intersect at a non-core AS. In this case, a shorter forwarding path can be created by removing the extraneous part of the path. The special case where the source's up-segment contains the destination AS is treated in the same way, i.e., the intersection of both segments is omitted from the path.
- **Combination with a core-segment** (e.g., $A \rightarrow D$ in Figure 2a): the last AS on the up-segment is different from the first AS on the down-segment. This case requires an additional core-segment to connect the up- and down-segment. If the communication remains within the same ISD ($A \rightarrow D$), a intra-ISD core-segment is needed; otherwise, an inter-ISD core-segment is required.
- **On-path:** (e.g., $A \rightarrow E$ in Figure 2a): the destination AS is directly on the path to the ISD core, so a single up-segment is sufficient to create a forwarding path.

Once a forwarding path is chosen, it is encoded in the SCION packet header, which makes inter-domain routing tables unnecessary for border routers: both the egress and the ingress interface of each AS on the path are encoded as packet-carried forwarding state (PCFS) in the packet header. The destination can respond to the source by inverting the end-to-end path from the packet header, or it can perform its own path lookup and combination.

3.3 Security Aspects

For protection against malicious entities and to provide secure control and data planes, SCION is equipped with an arsenal of security mechanisms.

Similar to BGPsec [19], each AS signs the PCB it forwards. This signature enables PCB validation by all entities. To ensure path correctness, the forwarding information within each packet-carried forwarding state (PCFS) also needs to be cryptographically protected, but signature verification would hamper efficient forwarding. Thus, each AS uses a

secret symmetric key that is shared among beacon servers and border routers and is used to efficiently compute a message authentication code (MAC) over the forwarding information. The per-AS information includes the ingress and egress interfaces, an expiration time, and the MAC computed over these fields, which is (by default) all encoded within an 8-byte field that we refer to as *hop field* (HF). The structure of the hop field is largely at the discretion of each AS and requires no coordination with any other AS — as long as the AS itself can extract how to forward the packet on to the next AS.

The specified ingress and egress interfaces uniquely identify the links to the previous and following ASes. If a router is connected via the same outgoing interface to three different neighboring ASes, three different egress interface identifiers would be assigned. The HF's expiration time can be set on the granularity of seconds or hours, depending on the path type. For the discussion of this overview, we only consider the common case where paths are long-lived and HFs have an expiration time on the order of 12 hours.

3.3.1 Algorithm Agility

In terms of cryptographic mechanisms, we built in algorithm agility, meaning that cryptographic methods can be easily updated and exchanged. The MAC validation of hop fields is per-AS, so an AS can independently (without interaction with any other entity) update its keys or cryptographic mechanisms. We support multiple signatures by an AS, thus, an AS can readily deploy a new signature algorithm and start adding those signatures as well. A component of the path-segment and beacon selection metric will favor creating paths where each AS on the path supports the new algorithm.

3.3.2 Authentication

Authentication in SCION is based on digital certificates, which bind identifiers to public keys and carry digital signatures that are verified by roots of trust, i.e., public keys that are axiomatically trusted.[1] One challenge is how to achieve trust agility to enable flexible selection of trust roots, resilience to private key compromise, and efficient key revocation [21].

A central question is how to structure the trust roots. Today's Internet follows two trust models: monopoly and oligopoly. In the *monopoly* model, a single root of trust is used for authentication. The DNSSEC PKI [5] or the Resource Public-Key Infrastructure (RPKI) [24] used in BGPSEC are examples of the monopoly model as they both essentially rely on a single public key that serves as a root of trust to verify all subsequent entities. The monopoly model suffers from two main drawbacks: all parties must agree on a single root of trust, and the single root of trust represents a single point of failure: its misuse enables forging a certificate for an arbitrary entity, and its revocation can result in a kill-switch for all its entities. The *oligopoly* model does not fare much better — instead of a single root of trust, the oligopoly model relies on several roots of trust, all of which are equally and completely trusted. Instead of one single point of failure in the monopoly model, the oligopoly model thus exposes several points of failure. The prime example is the TLS PKI, featuring on the order of 1 500 trusted signing certificates with about 300 roots of trust [2, 12]. Recently reported attacks have demonstrated that the compromise of a single trusted certificate authority enables issuing server certificates for any domain, including those with whom there is no business relationship.

SCION allows each ISD to define its own set of trust roots, along with the policy governing their use. Such scoping of trust roots within an ISD greatly improves security, as compromise of a private key associated with a trust root cannot be used to forge a certificate outside the ISD. An ISD's trust roots and policy are encoded in the *trust root configuration* (TRC) file. The TRC file has a version number, a list of public keys that serve as roots of trust for various purposes, and policies governing how many signatures are required for performing different types of actions. The TRC file serves as a way to bootstrap all authentications.

The TRC file provides important properties. *Trust agility* enables the selection of trust roots used to initiate the validation of certificates. A user can thus select an ISD that she believes maintains a non-compromised set of trust roots. A challenge with trust agility is to maintain global verifiability of all entities, regardless of the user's selection. SCION offers this property by requiring all ISDs with a link among them to sign each other's TRC files — as long as a network path exists, a validation path exists along that network path. *Efficient revocation of trust roots* is the second important property. In today's Internet, trust roots are revoked manually, or through OS or browser updates, often requiring a week or longer until a large fraction of the Internet population has observed such revocations. There is also a long tail of devices and installations that apply revocations very late or never. In SCION, PCBs carry the version number of the current TRC, and the updated TRC is required to validate that PCB. An AS that realizes that it needs a newer TRC can contact the AS from whom it has received the PCB. Following the distribution of PCBs, an entire ISD updates the TRC within tens of seconds.

3.3.3 SCION Control Message Protocol (SCMP)

The control plane includes the SCION Control Message Protocol (SCMP), which is similar to ICMP in the current Internet but authenticated and adapted to SCION. One challenge in the design of SCMP was how to enable efficient authentication of SCMP messages, as the naive approach of adding a digital signature to SCMP messages could create a processing bottleneck at routers when many SCMP messages would be created in response to a link failure. We thus make use of an efficient symmetric key derivation mechanism called *Dynamically Re-creatable Key* (DRKey) [17]. In DRKey, each AS uses a local secret key known to SCION border routers to derive on-the-fly a per-AS secret key using an efficient pseudorandom function (PRF). Hardware implementations of modern block ciphers enable faster computation than a memory lookup from DRAM, and therefore such dynamic key derivation can even result in a speedup over fetching the key from memory. For verification of SCMP messages, the destination AS can fetch the derived key through an additional request message from the originating AS, which is protected by a relatively slow asymmetric operation. However, local caching ensures that this key only needs to be fetched infrequently. As a consequence, SCION provides fully secured control messages with minimal overhead.

[1] Our reason for not using self-certifying identifiers [3] for long-term identities is their inherent inability for revocation and the complexities involved with key updates. For short-term identities, however, we do appreciate their features.

4. DEPLOYMENT

As of 2017, we have deployed a global SCION testbed which we are actively using to vet SCION's functionality and security. The testbed includes deployment nodes in 5 continents with 4 ISDs and 15 ASes including ISPs (KDDI, Swisscom, SWITCH) as well as financial and academic institutions. Our open-source code and information for deploying a SCION node can be found at the following address:

`www.scion-architecture.net`

Obtaining SCION's full benefits requires a direct connection between ASes. When a direct link is not possible, remote ASes can be connected via IP tunnels, but their communication will depend on the BGP routing protocol. As the testbed expands, we expect that more participants will connect directly to benefit from SCION's full feature set.

To use SCION, ISPs at a minimum need to deploy a border router capable of encapsulating and decapsulating SCION traffic as it leaves or enters their network. SCION ASes must also deploy certificate, beacon, and path servers. These servers can run on commodity hardware. Deploying SCION to homes or businesses is designed to require little effort, initially requiring no changes to existing software, networking stacks, nor the replacement of end user network devices. This is achieved through a gateway device that transparently switches communication over to SCION if the remote endpoint is also SCION-enabled. Several companies are currently exploring commercialization of these technologies, in particular the startup Fresh Start Networks, which offers SCION routers.

5. CONCLUSION

We have revisited SCION, a future Internet architecture that provides security, availability, transparency, control, scalability, and more (see the sidebar "The Future Looks Bright with SCION"). SCION offers numerous advantages over the current Internet and supports other future Internet proposals as an underlying building block for highly reliable point-to-point communication.

Despite its research maturity after 5 years of work, SCION is still in its infancy in terms of deployment. While requiring relatively small changes by ISPs and domains, broadening adoption is currently SCION's foremost goal. We expect that the benefits for various stakeholders will provide strong incentives to drive adoption, leading to islands of SCION deployment. In the long term, connections and mergers among islands will enable ever-increasing numbers of native SCION end-to-end connections.

Working on SCION has offered us the opportunity to consider Internet architectures from a clean-slate perspective. The absence of limiting constraints (imposed by the current Internet environment) has been particularly rewarding, as the deep exploration of a problem space enabled us ask not *how can a future Internet achieve what the current Internet has?*, but rather *what additional features can/should a future Internet offer?* We anticipate that the insight into the possible applications of a secure, dynamic, and highly-available network will help engage the network community to leverage SCION for their applications, and contribute to the project.

Our book "SCION: A Secure Internet Architecture" describes in more detail numerous aspects of SCION, including authentication, name resolution, deployment, operation, extensions, and specifications. [22].

The Future Looks Bright with SCION

The SCION inter-domain network architecture enables new systems that can take advantage of the isolation, scalability, and transparency properties provided.

Path validation. SCION, through its use of packet-carried forwarding state, paves the way for the Origin and Path Trace (OPT) mechanism [17]. OPT enables senders, receivers, and routers to cryptographically verify the exact path the packets have traversed, with negligible overhead. OPT allows the transmission of banking or medical data, which is typically bound to strict data privacy regulations, to be constrained to traverse only selected authorized ASes.

Anonymity and privacy. Packet-carried forwarding state (PCFS) also provides advantages for privacy: with PCFS and path transparency, the source is able to select paths that appear more trustworthy (e.g., those that do not traverse certain ASes). In addition, the packet header can be further obfuscated such that ASes on the path cannot learn identifying details about the source or the destination, unless they are immediately connected to one of them. HORNET [10] leverages SCION's path selection infrastructure to offer high bandwidth and low latency anonymous communication.

Highly available communication. Critical infrastructures such as financial networks and industrial control systems used for power distribution require a high degree of availability. Internet outages have been known to wreak havoc on day-to-day operations, for example preventing ATM withdrawals or payment terminal operations [28]. Numerous such outages are due to the malicious or erroneous announcement of IP address spaces, which is also known as prefix hijacking. Perhaps the most famous case is the hijack of YouTube by Pakistan for internal censoring, resulting in a global outage of YouTube [8]. In fact, hijacks that impact only a small portion of the Internet happen on a daily basis. SCION's control-plane isolation through ISDs, its stable data plane, and its multipath operation all contribute to dramatically higher availability. With ISDs, misconfigurations and attacks in one ISD do not affect others; digitally signed route announcements prevent unauthorized injection of routes; and digitally signed path distribution allows verification of paths by the sender.

DDoS prevention. Bandwidth guarantees are enabled by SIBRA [6], which prevent DDoS attacks at the architectural level: independent of the number of distributed bots, end hosts obtain protection against Internet-wide link-flooding attacks, one of the major threats in today's Internet. SCION provides ISDs with dynamic bandwidth guarantees to permanently enable communication. Critical infrastructures can additionally keep some network paths to a destination secret, thus preventing an adversary from even sending traffic to that destination because the cryptographic OFs are necessary to use a path, but are unknown to an adversary.

High-speed web browsing. Through the SIBRA extension, the sender performs a resource reservation with its initial packet, and the receiver will likely obtain a reservation with a high sending rate that it can immediately use on the reverse path. On such a reservation, no congestion control is needed; consequently, web servers can immediately start sending content at a high rate to the client.

Mobility support. With the proliferation of mobile devices, supporting reliable communication can be challenging since these devices frequently connect and disconnect from (sometimes several) networks. SCION supports high availability through multipath communication and provides a header extension to inform the other party of new down segments as it connects to a new network. Failing paths are discarded and new paths are dynamically discovered.

Protection from forged TLS certificates. The government of Iran infamously used compromised roots of trust to

create rogue TLS certificates for Google and Yahoo services to perform man-in-the-middle attacks on its citizens. Iran is suspected to have mounted the attack on the DigiNotar CA, who signed these certificates. ISDs and the ARPKI [7] system used in SCION prevent such attacks, as a CA's authority is scoped to the ISDs where the CA is active in. Moreover, in ARPKI, multiple trusted entities need to be compromised to perform a successful man-in-the-middle attack, and revocation of trust roots is possible within a minute, enabling quick recovery from compromise.

6. REFERENCES

[1] Asia communications hit by quake. `http://news.bbc.co.uk/2/hi/asia-pacific/6211451.stm`, december 2006.

[2] M. Abadi, A. Birrell, I. Mironov, T. Wobber, and Y. Xie. Global authentication in an untrustworthy world. In *Proceedings of Workshop on Hot Topics in Operating Systems (HotOS)*, 2013.

[3] D. G. Andersen, H. Balakrishnan, N. Feamster, T. Koponen, D. Moon, and S. Shenker. Accountable internet protocol (AIP). In *Proceedings of ACM SIGCOMM*, 2008.

[4] D. G. Andersen, H. Balakrishnan, M. F. Kaashoek, and R. Morris. Resilient overlay networks. In *Proceedings of ACM Symposium on Operating Systems Principles (SOSP)*, 2001.

[5] R. Arends, R. Austein, M. Larson, D. Massey, and S. Rose. DNS Security Introduction and Requirements. RFC 4033 (Proposed Standard), 2005.

[6] C. Basescu, R. M. Reischuk, P. Szalachowski, A. Perrig, Y. Zhang, H.-C. Hsiao, A. Kubota, and J. Urakawa. SIBRA: Scalable internet bandwidth reservation architecture. In *Proceedings of Network and Distributed System Security Symposium (NDSS)*, 2016.

[7] D. Basin, C. Cremers, T. H.-J. Kim, A. Perrig, R. Sasse, and P. Szalachowski. ARPKI: Attack resilient public-key infrastructure. In *Proceedings of the ACM Conference on Computer and Communications Security (CCS)*, 2014.

[8] M. Brown. Pakistan Hijacks YouTube. `http://research.dyn.com/2008/02/pakistan-hijacks-youtube-1/`, 2008.

[9] CCNx - Content Centric Networking. `http://www.ccnx.org`, 2015.

[10] C. Chen, D. 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 and Communications Security (CCS)*, 2015.

[11] T. Dübendorfer, A. Wagner, and B. Plattner. An economic damage model for large-scale internet attacks. In *13th IEEE International Workshops on Enabling Technologies: Infrastructure for Collaborative Enterprises*, 2004.

[12] Electronic Frontier Foundation. SSL Observatory. `https://www.eff.org/observatory`, 2010.

[13] D. Farinacci, V. Fuller, D. Meyer, and D. Lewis. The locator/id separation protocol (LISP). RFC 6830, 2013.

[14] D. Han, A. Anand, F. Dogar, B. Li, H. Lim, M. Machado, A. Mukundan, W. Wu, A. Akella, D. G. Andersen, J. W. Byers, S. Seshan, and P. Steenkiste. XIA: Efficient support for evolvable internetworking. In *Proceedings of the 9th USENIX Symposium on Networked Systems Design and Implementation (NSDI)*, 2012.

[15] V. Jacobson, D. K. Smetters, J. D. Thornton, M. F. Plass, N. H. Briggs, and R. L. Braynard. Networking named content. In *Proceedings of the 5th international conference on emerging networking experiments and technologies (CoNEXT)*, 2009.

[16] E. Katz-Bassett, C. Scott, D. Choffnes, I. Cunha, V. Valancius, N. Feamster, H. Madhyastha, T. Anderson, and A. Krishnamurthy. LIFEGUARD: Practical repair of persistent route failures. In *Proceedings of ACM SIGCOMM*, 2012.

[17] T. H. Kim, C. Basescu, L. Jia, S. B. Lee, Y. Hu, and A. Perrig. Lightweight source authentication and path validation. In *Proceedings of ACM SIGCOMM*, 2014.

[18] N. Kushman, S. Kandula, and D. Katabi. Can you hear me now?! it must be BGP. *ACM SIGCOMM Computer Communication Review*, 2007.

[19] M. Lepinski and S. Turner. An overview of BGPSEC. IETF Draft, `http://tools.ietf.org/html/draft-ietf-sidr-bgpsec-overview-02`, 2012.

[20] M. Marlinspike. SSL and the future of authenticity. `http://blog.thoughtcrime.org/ssl-and-the-future-of-authenticity`, 2011.

[21] S. Matsumoto, R. M. Reischuk, P. Szalachowski, T. H.-J. Kim, and A. Perrig. Authentication Challenges in a Global Environment. *ACM Transactions on Privacy and Security (TOPS)*, 2017.

[22] A. Perrig, P. Szalachowski, R. M. Reischuk, and L. Chuat. *SCION: A Secure Internet Architecture*. Springer, 2017.

[23] D. Raychaudhuri, K. Nagaraja, and A. Venkataramani. MobilityFirst: A robust and trustworthy mobility-centric architecture for the future internet. *ACM SIGMOBILE Mobile Computing and Communications Review*, 2012.

[24] Resource Public Key Infrastructure (RPKI). `https://www.arin.net/resources/rpki/`, 2015.

[25] A. Sahoo, K. Kant, and P. Mohapatra. BGP convergence delay under large-scale failures: Characterization and solutions. *Computer Communications*, 32(7), 2009.

[26] J. H. Saltzer, D. P. Reed, and D. D. Clark. End-to-end arguments in system design. *ACM Transactions on Computer Systems*, 2(4), 1984.

[27] M. Schuchard, E. Y. Vasserman, A. Mohaisen, D. F. Kune, N. Hopper, and Y. Kim. Losing control of the Internet: Using the data plane to attack the control plane. In *Proceedings of Network and Distributed System Security Symposium (NDSS)*, 2011.

[28] A. Toonk. Massive route leak causes internet slowdown. `http://www.bgpmon.net/massive-route-leak-cause-internet-slowdown/`, 2015.

[29] X. Zhang, H.-C. Hsiao, G. Hasker, H. Chan, A. Perrig, and D. G. Andersen. SCION: Scalability, control, and isolation on next-generation networks. In *Proceedings of IEEE Symposium on Security and Privacy*, 2011.