Adhering to the end-to-end principle even more than the current Internet yields highly available point-to-point communication.

BY DAVID BARRERA, LAURENT CHUAT, ADRIAN PERRIG, RAPHAEL M. REISCHUK, AND PAWEL SZALACHOWSKI
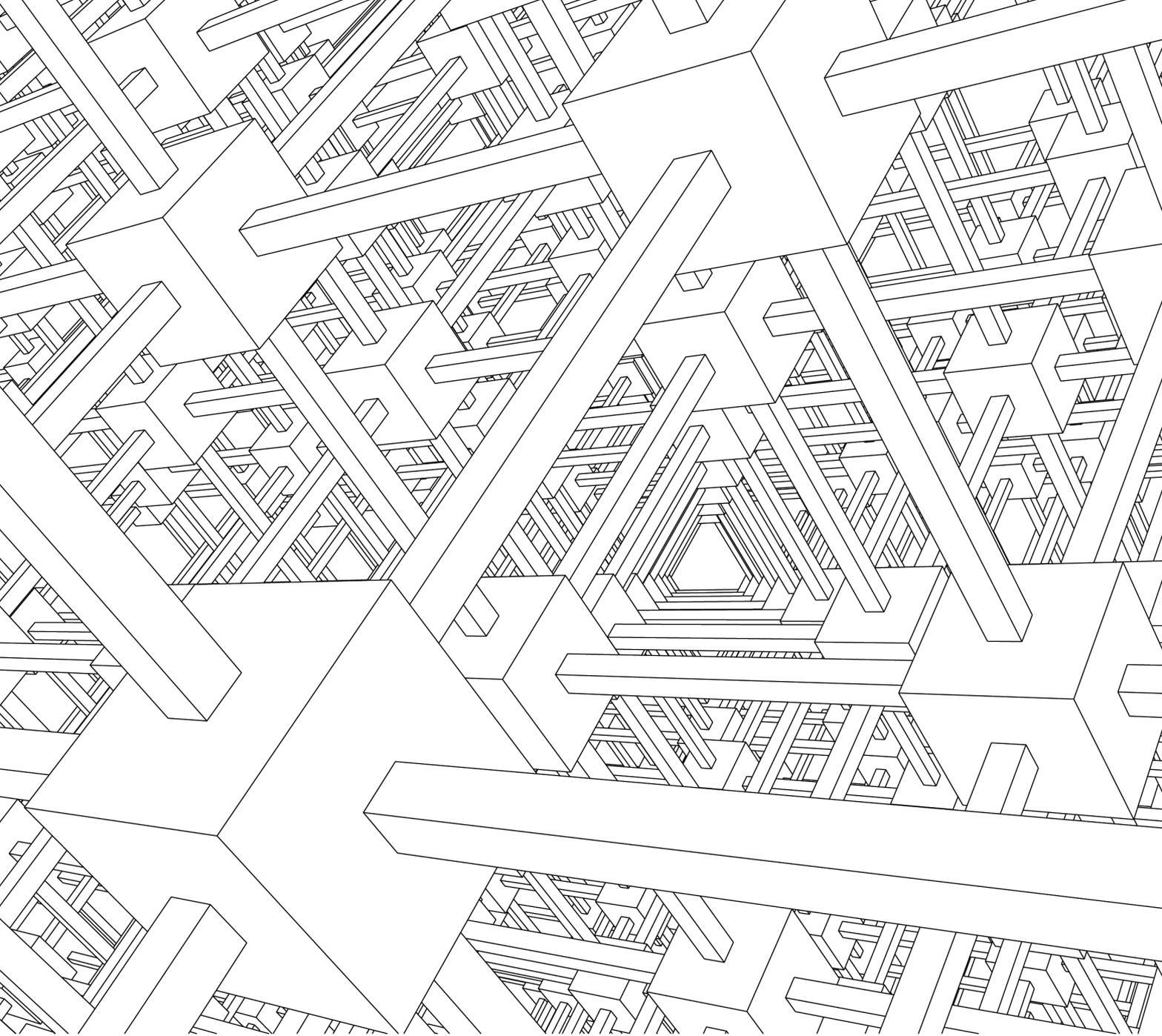
The SCION Internet Architecture

THE INTERNET HAS been successful beyond even the most optimistic expectations. It permeates almost every aspect of our society and economy worldwide. This success has created universal dependence on communication, as many of the processes underpinning modern society would grind to a halt if it were unavailable. However, the state of the safety and availability of the Internet is far from commensurate with its importance.

Although we cannot conclusively determine what the impact of even a one-minute outage of Internet connectivity would be, anecdotal evidence suggests that even a brief outage would have a profound negative effect on governmental, economic, and societal operations.[11] Making matters worse, the Internet is not designed primarily for high availability in the face of malicious actions by adversaries. Recent patches to improve Internet security and availability are indeed constrained by the design of the current
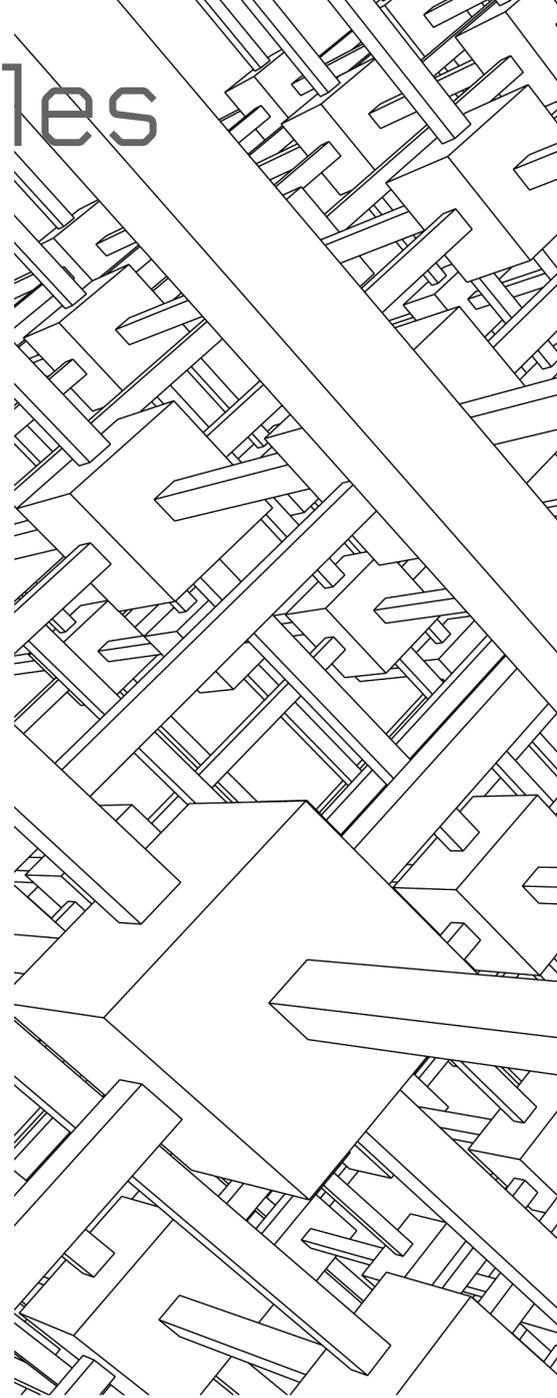
» key insights

- **Patching the current Internet is an undesirable long-term solution; a clean-slate redesign of inter-domain routing would provide many benefits and is surprisingly simple to deploy using legacy protocols for intra-domain communication.**

- **SCION's isolation domains offer control-plane isolation and scoped trust; rather than restrict communication, they provide transparency for path selection, packet forwarding, and authentication.**

- **SCION's packet-carried forwarding state eliminates the need for inter-domain routing table lookups, improves forwarding performance, and supports multipath communication; packet-carried forwarding state gives path control to senders, providing scalability, security, and availability benefits.**

Internet architecture. A new Internet architecture must offer availability, security by design, and incentives for deployment, as well as address economic, political, and legal issues at the design stage.

Such features require a completely new cohesive architecture that provides one fundamental building block—highly available point-to-point communication—on which other proposed Internet architectures that provide content-centric,[15,21] extensibility-centric,[14] or mobility-centric[23] properties can build.

This article describes SCION, or Scalability, Control, and Isolation On Next-generation networks, an inter-domain network architecture designed to address these issues, covering SCION's goals, design, and functionality, as well as the results of six years of research we have conducted since our initial publication.[28]

Objectives

We begin with the high-level goals an inter-domain point-to-point communication architecture must be able to accomplish.

Availability in the presence of adversaries. Our aim 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 the endpoints, and 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 must thus provide mechanisms to counteract malicious operations. An "off-path adversary" could launch a hijack attack to attract traffic to flow through network elements under its control. Such traffic attraction can take several 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 (such as by inducing congestion). An adversary controlling a large botnet could also perform distributed denial-of-service (DDoS) attacks, congesting selected network links. And an adversary could interfere with the discovery of legitimate paths (such as 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 are ensured of being able to traverse certain Internet service providers (ISPs) and avoid others.

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

The beneficial aspects of path control include:

Separation of network control plane and data plane. Ensuring that forwarding cannot be retroactively influenced by control plane operations (such as routing changes);

Enabling multipath communication. Improving availability by allowing senders to select multiple paths to their destinations; and

Defending against network attacks. Including DDoS and traffic interception by rogue networks, since destinations can observe a packet's traversed path in the packet header.

Particular care must be taken for the proper handling of the fragile aspects of communication, including:

Respecting ISPs' forwarding policies. By offering policy-compliant paths from which senders can choose;

Preventing malicious path creation. Including paths that contain loops;

Ensuring scalability of path control. By allowing sources to select paths from among a relatively small set, as opposed to full-edged source routing; and

Enabling ISP traffic engineering. Despite end hosts' path control, giving ISPs the ability to balance their load across the links to their neighbor autonomous systems (ASes).

**Transparency and control over trust roots.** Roots of trust are used to verify entities in the current Internet, as in verification of a server's public key in a Transport Layer Security (TLS) certificate or of a Domain Name System (DNS) response in DNSSEC (DNS Security Extensions).[5] Transparency of trust roots provides end hosts and users knowledge of the complete set of trust roots relied upon for entity-certificate validation. Enumerating trust roots is difficult due to intermediate certification authorities that are trusted implicitly. Control over trust-root selection enables trust agility, allowing users to readily select or exclude the roots of trust they wish to rely upon.

Efficiency and scalability. Despite the lack of availability and transparency, the current Internet also suffers from 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[24] or even days.[8] Moreover, routing tables have reached the limit of their scalability due to multihoming and prefix de-aggregation or announcement of more-specific IP address spaces. Increasing memory size for routing tables is problematic, as the underlying hardware is expensive and power-hungry, accounting for approximately one-third of a router's total power consumption.

Security and high availability usually come at a cost, resulting in less efficiency and potentially diminished scalability. High performance and scalability are, however, 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, most notably with respect to BGP and to the growing size of routing tables.

One approach for achieving 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 greater packet processing speeds and simpler router architectures compared to today's IP routers. Avoiding state on routers also prevents state-exhaustion

Figure 1. ASes grouped into four ISDs. Core ASes are connected through core links. Non-core ASes are connected through customer-to-provider or peering links. Some ASes are contained in multiple ISDs.
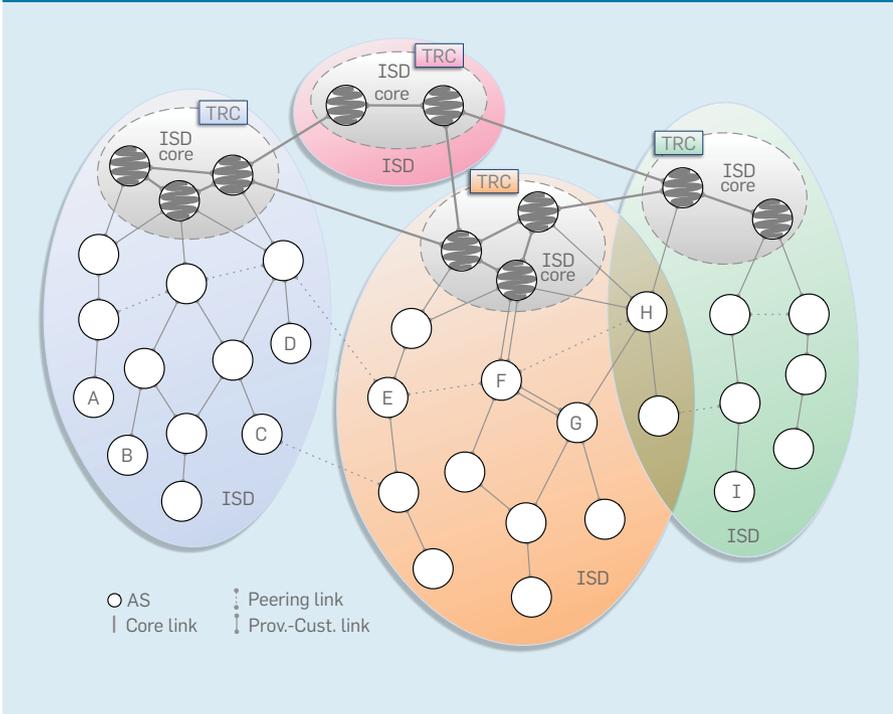
attacks[26] and state inconsistencies across routers. Our goal of efficiency and scalability is in line with the end-to-end principle, which states that a function should be implemented at the network layer in which it can operate most effectively.[25] Since the end host has the most information about its own internal state, network functions related to that state (such as error detection and correction, acknowledgment of receipt, and retransmission) 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; that is, SCION adheres to the end-to-end principle even more than the current Internet.

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

Support for global but heterogeneous trust. Given the diverse nature of the constituents in the current Internet, with its multiple legal jurisdictions and interests, an important challenge is how to scale authentication of entities (such as AS ownership for routing, name servers for DNS, and domains for TLS) to the global environment. The roots of trust of currently prevalent public key infrastructure (PKI) models (monopoly and oligopoly) do not scale to a global environment because mutually distrustful 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 inherently distrustful entities.

Deployability. A new Internet architecture should offer a multitude of features that incentivize its deployment. We thus aim for SCION to provide high availability even under control-plane and data-plane attacks (thanks to built-in DDoS defenses), path transparency and control, trust-root transparency and control, robustness to configura-

tion errors, fast recovery from failure, high forwarding efficiency, and multipath forwarding. Economic and business incentives are also critical, making it possible for ISPs to define new business models and sell new services.

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

Foundation for other architectures. To achieve a simple, scalable, secure, efficient architecture, we now focus on the most basic communication mode: point-to-point communication. Other architectures that provide support for higher-level properties (such as for content distribution,[15,21] extensibility,[14] and mobility[23]) all require a working point-to-point communication infrastructure.

SCION Architecture

SCION introduces the concept of isolation domain (ISD), a building block for achieving high availability, transparency, scalability, and support for heterogeneous trust, constituting a logical grouping of ASes, as outlined in Figure 1.

An ISD is administered by multiple ASes that form the ISD core; we refer to them as "core ASes." The ISD is governed by a policy we call "trust root configuration" (TRC), which is negotiated by the ISD core. The TRC 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 acceptance of the ISD's TRC. We envision ISDs spanning areas with uniform legal environments that provide enforceable contracts. If two ISPs have a contract dispute they are unable to resolve by themselves, such a legal environment would provide an external authority to resolve the dispute. All ASes within an ISD also agree on the TRC, or 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. 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 support heterogeneous trust environments.

SCION includes two levels of routing—intra-ISD and inter-ISD—that use "path-segment construction beacons" (PCBs) to explore routing paths, as outlined in Figure 2a.

A core AS announces a PCB and disseminates it as a policy-constrained multi-path flood either within an ISD (to discover intra-ISD paths) or among core ASes (to discover inter-ISD paths), a process we call "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, a concept we call "packet-carried forwarding state" (PCFS).

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

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

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

Figure 2. SCION components at different scales: (a) SCION ISD with PCBs propagated from the ISD core down to customer ASes, and path segments for ASes *A*, *B*, *C*, *D*, and *E* to the ISD core; and (b) magnified view of an AS with its routers and servers. The path from AS *C* to the ISD core traverses two internal routers.
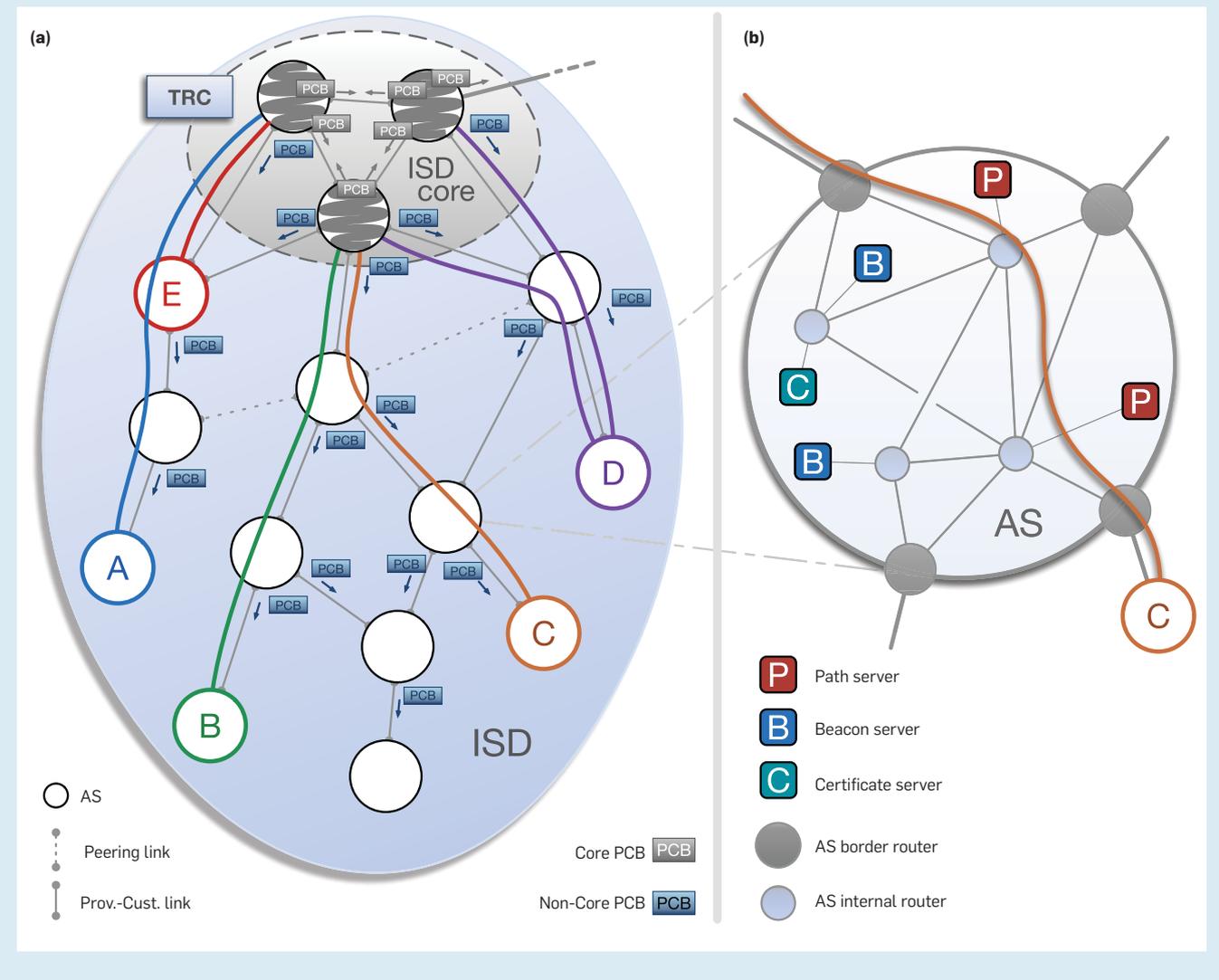
inside ASes. We did not include name servers in Figure 2b, as their operation is similar to today's DNS.

Beacon servers are responsible for disseminating PCBs, as in 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 these PCBs and re-send them to their customer ASes, resulting in AS-level path segments. Figure 3 outlines PCBs 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 generates a set of PCBs 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 it is periodic and

PCBs are flooded through multiple paths 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 (such as control of upstream ASes) BGP is unable to 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, uploading 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. Beacon servers require certificate servers when validating the authenticity of PCBs.

Border routers forward packets between ASes supporting SCION. In the case of a control packet, the border router forwards it to the appropriate server, and, in the case of a data packet, forwards it either to a host inside the AS or toward the next border router.

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

Path exploration and registration. Through inter-domain beaconing, core ASes discover paths to other core ASes. Through intra-domain beaconing,

ASes discover path segments leading to core ASes that enable an AS to communicate with the ISD core; Figure 2a outlines path segments from ASes *A*, *B*, *C*, *D*, and *E* to the core. The beaconing process is asynchronous; that is, 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-grain path diversity; ASes often have several diverse connection points, and 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 outlines this feature; AS *F* receives two different PCBs via two different links from the core. Moreover, AS *F* uses two different links to send two different PCBs to AS *G*, each with the respective egress interfaces. AS *G* extends the two PCBs, forwarding both over a single link to its customer.

An important requirement of the architecture is that SCION also supports peering links between ASes. Consistent with AS policies in the current Internet, PCBs do not traverse peering links, though peering links are announced, along with a regular path in a PCB. Figure 3 outlines 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 without going through the core. SCION also supports peering links that cross ISD boundaries, highlighting the importance of SCION's path-transparency property; a source host knows the exact set of ASes and ISDs traversed during the delivery of each packet.

An AS typically receives several PCBs representing path segments to various core ASes. Figure 2a outlines two path segments for AS *D*. We call a path segment that leads toward an ISD core an "up-segment" and a path segment that leads from the ISD core to an AS a "down-segment," though 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 sequence of ASes, 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. 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 prefers to be reached and register them 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.

Path lookup. To reach a remote destination, a host first queries a SCION name server to obtain the `<ISD, AS, end-host address>` triplet of the destination. The ISD and AS identifiers are needed to perform a path lookup, and the end-host address is used by the destination AS to deliver the packet to the destination host. To obtain up-segments to reach its ISD core, a host performs a path lookup at its local path server. To obtain down-segments to

reach the destination, the host queries the local path server with the destination's `<ISD, AS>` tuple. If the local path server has no cached down-segments, it will automatically query the destination AS's core path server.

PCB and path-segment selection. The PCBs to propagate and path segments to register are selected by each AS 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 like anonymous forwarding) or link property (such as 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.

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.
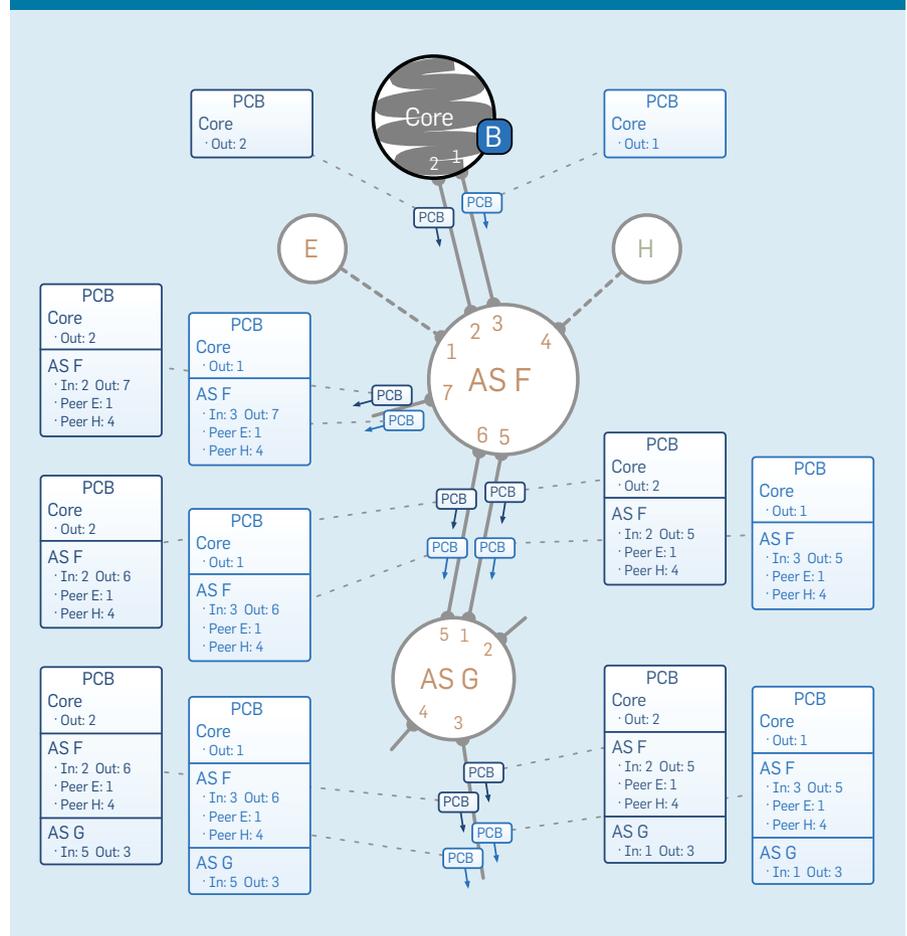
"Policy compliance" refers to the requirement that the path adheres to the AS's routing policy. Based on past PCBs that were sent, a beacon server scores the current set of candidate path segments and sends the *k* best segments as the next PCB. SCION intra-ISD beaconing can scale to networks of arbitrary 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 is similar to intra-ISD beaconing, except inter-ISD PCBs traverse only ISD core ASes. The same path-selection metrics apply in which an AS attempts to forward the set of most-desirable paths to its neighbors. Like BGP, the process is inherently not scalable, but, as the number of ISDs and the corresponding number of core ASes is small, the approach is viable for SCION.

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

Intra-AS communication. Communication within ASes is handled through existing intra-domain communication protocols (such as IP, Open Shortest Path First, Multiprotocol Label Switching, and Software-Defined Networking).

We explicitly seek high efficiency such that packet-forwarding latency and throughput are at least as fast as current IP forwarding.

Figure 2b outlines one possible intra-domain path through the magnified AS.

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 (augmented with ingress and egress interface identifiers for each AS) without having to inspect 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 be able to forward it to the appropriate local host(s).

An interesting aspect of forwarding is enabled by the split of "locator" (the path toward 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. A domain can thus select an arbitrary addressing format for its hosts (such as a 4B IPv4, 6B medium access control, 16B IPv6, 20B accountable IP, and AIP[3]). A nice consequence is that an IPv4 host can communicate with an IPv6 host directly through SCION.

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

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

▸ Immediate combination of path segments (such as $B \rightarrow D$ 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 (such as $A \rightarrow B$ in Figure 2a). A peering link exists between the two segments, so a shortcut via the peering link is possible. As in the case of the AS shortcut, the extraneous path segment is cut off; the peering link could also traverse to a different ISD;

▸ AS shortcut (such as $B \rightarrow C$ 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 the same way or the intersection of both segments is omitted from the path;

▸ Combination with a core-segment (such as $A \rightarrow D$ 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 ($A \rightarrow D$), an intra-ISD core-segment is needed; otherwise, an inter-ISD core segment is needed; and

▸ On-path (such as $A \rightarrow E$ 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 the host chooses a forwarding path, it is encoded in the SCION packet header, making inter-domain routing tables unnecessary for border routers; both the egress and the ingress interface of each AS on the path are encoded as PCFS in the packet header. The destination can respond to the source by inverting the end-to-end path from the packet header or perform its own path lookup and combination.

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

As in BGPsec,[19] each AS signs the PCB it forwards, enabling PCB validation by all entities. To ensure path correctness, the forwarding information within each PCFS also needs to be cryptographically protected, but signature verification would hamper efficient forwarding. Each AS thus uses a secret symmetric key that is shared among beacon servers and border routers and 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 are (by default) all encoded within an 8B field we refer to as a "hop field" (HF). The structure of the HF is largely at the discretion of each AS and requires no coordination with any other AS, as long as the AS itself can determine 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, for example, a router is connected via the same outgoing interface to three different neighboring ASes, three different egress-interface identifiers would be assigned by network administrators. The HF's expiration time can be set to the granularity of seconds or hours, depending on path type. For this article, we consider only the common case where paths are long-lived and HFs have an expiration time of approximately 12 hours.

Algorithm agility. In terms of cryptographic mechanisms, SCION includes built-in algorithm agility, meaning cryptographic methods are easily updated and exchanged. The MAC validation of HFs is per-AS, so an AS can independently (without interaction with any other entity) update its keys or cryptographic mechanisms. SCION supports multiple signatures by an AS, meaning an AS can readily deploy a new signature algorithm and start adding those signatures as well. A component of the selection metric favors creating paths where each AS on the path supports the new algorithm.

Authentication. Authentication in SCION is based on digital certificates that bind identifiers to public keys and carry digital signatures that are verified by roots of trust. One notable challenge is how to achieve trust agility to enable flexible selection of trust roots, resilience to private key compromise, and efficient key revocation.[20]

A central question we have had to address is how to structure the Internet's trust roots. The current Internet follows two trust models: monopoly and oligopoly. In the monopoly model, a single root of trust is used for authentication. The DNSSEC PK[15] or the Resource Public-Key Infrastructure (RPKI)[2] used in BGPsec are examples of the monopoly model, as 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, the misuse of which enables forging a certificate for an arbitrary entity, and its revocation can result in a kill-switch for all its entities. The oligopoly model fares no better; instead of a single root of trust, the oligopoly model relies on several roots of trust, all equally and completely trusted. Instead of a 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 approximately 1,500 trusted signing certificates with approximately 300 roots of trust.[1,12] Attacks reported since 2011 against authorities (such as Comodo, DigiNotar, and GlobalSign) demonstrate how compromise of a single trusted certificate authority enables issuing server certificates for any domain, including those with which 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 TRC, which has a version number, a list of public keys that serve as roots of trust for various purposes, and policies governing the number of signatures required for performing different types of actions. The TRC serves as a way to bootstrap all authentications within SCION.

The TRC provides important properties. Trust agility enables users to select trust roots used to initiate certificate validation. Users can thus select an ISD they believe maintains a non-compromised set of trust roots. A challenge with trust agility is how to maintain global verifiability of all entities, regardless of the user's selection. SCION of-

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 it indeed provides.

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

Anonymity and privacy. PCFS also provides advantages for privacy. For example, with PCFS and path transparency, the source is able to select paths that appear more trustworthy (such as 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. The High-speed Onion Routing at the Network Layer (HORNET)[10] leverages SCION's path-selection infrastructure to deliver high-bandwidth, low-latency anonymous communication.

Highly available communication. Critical infrastructure (such as financial networks and industrial control systems used for power distribution) requires a high degree of availability. Internet outages have been known to disrupt day-to-day operations by, for example, preventing ATM withdrawals or payment terminal operations.[27] Numerous such outages are due to the malicious or erroneous announcement of IP address spaces, or "prefix hijacking." Perhaps the most well-known example is the 2008 hijack of YouTube by Pakistan Telecom for the purpose of censorship, resulting in a global outage of YouTube.[9] In fact, hijacks affecting 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 other ISDs; 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 the Scalable Internet Bandwidth Reservation Architecture (SIBRA),[6] preventing DDoS attacks at the architectural level; independent of the number of distributed bots, end hosts gain protection against Internet-wide link-flooding attacks, a major threat in the current Internet. SIBRA provides ISDs with dynamic bandwidth guarantees to permanently enable communication. Critical infrastructures can additionally keep some network paths to a destination secret, preventing an adversary from even sending traffic to that destination because the cryptographic HFs 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 it can use immediately on the reverse path. With such a reservation, no congestion control is needed; consequently, Web servers can start sending content immediately at a high rate to the client.

Mobility support. With the ongoing proliferation of mobile devices, supporting reliable communication can be a challenge for any architecture, as these devices frequently connect and disconnect from (sometimes multiple) 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 discovered dynamically.

Protection from forged TLS certificates. The government of Iran in 2011 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 own citizens. Iran is suspected of having mounted the attack on the DigiNotar certificate authority (CA) that signed these certificates. ISDs and the Attack Resilient Public-Key Infrastructure (ARPKI)[7] system used in SCION prevent such attacks, as a CA's authority is scoped to the ISDs in which the CA is active. Moreover, in the ARPKI, multiple trusted entities must 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 the compromise.

fers this property by requiring all ISDs with a link between them to sign each other's TRCs; 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 the current Internet, trust roots are revoked manually or through operating system or browser updates, often requiring a week or more before a large fraction of the Internet population has seen the 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 it needs a newer TRC can contact the AS from which it has received the PCB. Following distribution of PCBs, an entire ISD updates the TRC within tens of seconds.

SCION Control Message Protocol. The SCMP is similar to ICMP in the current Internet but is authenticated and adapted to SCION. One challenge we have had to address in the design of SCMP is how to enable efficient authentication of SCMP messages, as the naïve 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. The SCION architecture thus makes use of an efficient symmetric key derivation mechanism called the "Dynamically Re-creatable Key" (DRKey)[17] in which 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." Hardware implementations of modern block ciphers enable faster computation than a memory lookup from DRAM, and such dynamic key derivation can thus result in a speedup even 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 this key needs to be fetched only infrequently. As a consequence, SCION provides fully secured control messages with minimal overhead.

Deployment

As of April 2017, we had deployed a global SCION testbed we use to vet SCION's functionality and security, including deployment nodes in five continents with four ISDs and 15 ASes, including ISPs—KDDI, Swisscom, and SWITCH—and financial and academic institutions. SCION's open-source code and information for how to deploy a SCION node is available at http://www.scion-architecture.net/

Obtaining SCION's full benefits requires a direct connection among multiple ASes. When a direct link is not pos-

sible, remote ASes can be connected via IP tunnels, but their communication depends on the BGP routing protocol. As the testbed expands, we expect more participants will connect directly to benefit from SCION's full feature set.

To use SCION, ISPs at a minimum must deploy a border router capable of "encapsulating" and "decapsulating" SCION traffic as it leaves or enters their networks. SCION ASes must also deploy certificate, beacon, name, and path servers that can run on commodity hardware. Deploying SCION in homes or businesses is designed to require little effort, initially with no changes to existing software or networking stacks or replacement of end-user network devices. This ready connection 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, notably the startup Anapaya Systems, which offers SCION routers.

Conclusion

SCION is an 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 following six years of effort, SCION is still in its infancy in terms of deployment. While requiring relatively small changes by ISPs and domains, broadening adoption is SCION's foremost goal. We expect the benefits for various stakeholders will provide strong incentives for 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 let us 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 this problem space enables us ask not how a future Internet can achieve what the current Internet has

already achieved, but rather what additional features can and should a future Internet offer. We anticipate the insight into the possible applications of a secure, dynamic, highly available network will help engage the network community to leverage SCION for its applications and contribute to the project.

Our 2017 book *SCION: A Secure Internet Architecture* describes the architecture in more detail, including authentication, name resolution, deployment, operation, extensions, and specifications.[22] **C**

References
1. Abadi, M., Birrell, A., Mironov, I., Wobber, T., and Xie, Y. Global authentication in an untrustworthy world. In *Proceedings of the 14th Workshop on Hot Topics in Operating Systems* (Santa Ana Pueblo, NM, May 13–15). Usenix Association, Berkeley, CA, 2013.
2. American Registry for Internet Numbers. Resource Public Key Infrastructure (RPKI); https://www.arin.net/resources/rpki/
3. Andersen, D.G., Balakrishnan, H., Feamster, N., Koponen, T., Moon, D., and Shenker, S. Accountable Internet Protocol (AIP). In *Proceedings of ACM SIGCOMM* (Seattle, WA, Aug. 17–22). ACM Press, New York, 2008.
4. Andersen, D.G., Balakrishnan, H., Kaashoek, M.F., and Morris, R. Resilient overlay networks. In *Proceedings of the ACM Symposium on Operating Systems Principles* (Chateau Lake Louise, Banff, Canada, Oct. 21–24). ACM Press, New York, 2001.
5. Arends, R., Austein, R., Larson, M., Massey, D., and Rose, S. *DNS Security Introduction and Requirements.* RFC 4033 (Proposed Standard), 2005; https://www.ietf.org/rfc/rfc4033.txt
6. Basescu, C., Reischuk, R.M., Szalachowski, P., Perrig, A., Zhang, Y., Hsiao, H.-C., Kubota, A., and Urakawa, J. SIBRA: Scalable Internet Bandwidth Reservation Architecture. In *Proceedings of Network and Distributed System Security Symposium* (San Diego, CA, Feb. 21–24). Internet Society, Reston, VA, 2016.
7. Basin, D., Cremers, C., Kim, T. H.-J., Perrig, A., Sasse, R., and Szalachowski, P. ARPKI: Attack Resilient Public-Key Infrastructure. In *Proceedings of the ACM Conference on Computer and Communications Security* (Scottsdale, AZ, Nov. 3–7). ACM Press, New York, 2014.
8. BBC News. Asia communications hit by quake. Dec. 27, 2006; http://news.bbc.co.uk/2/hi/asia-pacific/6211451.stm
9. Brown, M. Pakistan Hijacks YouTube; http://research.dyn.com/2008/02/pakistan-hijacks-youtube-1/
10. Chen, C., Asoni, D., Barrera, D., Danezis, G., and Perrig, A. HORNET: High-speed onion routing at the network layer. In *Proceedings of the ACM Conference on Computer and Communications Security* (Denver, CO, Oct. 12–16). ACM Press, New York, 2015.
11. Dübendorfer, T., Wagner, A., and Plattner, B. An economic damage model for large-scale Internet attacks. In *Proceedings of the 13th IEEE International Workshops on Enabling Technologies: Infrastructure for Collaborative Enterprises* (University of Modena and Reggio Emilia, Italy, June 14–16). IEEE Press, 2004.
12. Electronic Frontier Foundation. SSL Observatory, 2010; https://www.eff.org/observatory
13. Farinacci, D., Fuller, V., Meyer, D., and Lewis, D. *The Locator/ID Separation Protocol (LISP).* RFC 6830, 2013; https://tools.ietf.org/html/rfc6830
14. Han, D., Anand, A., Dogar, F., Li, B., Lim, H., Machado, M., Mukundan, A., Wu, W., Akella, A., Andersen, D.G., Byers, J.W., Seshan, S., and Steenkiste, P. XIA: Efficient support for evolvable internetworking. In *Proceedings of the Ninth USENIX Symposium on Networked Systems Design and Implementation* (San Jose, CA, Apr. 25–27). USENIX Association, Berkeley, CA, 2012.
15. Jacobson, V., Smetters, D.K., Thornton, J.D., Plass, M.F., Briggs, N.H., and Braynard, R.L. Networking named content. In *Proceedings of the Fifth International Conference on Emerging Networking Experiments and Technologies* (Rome, Italy, Dec. 1–4). ACM Press, New York, 2009.
16. Katz-Bassett, E., Scott, C., Chones, D., Cunha, I.,

17. Valancius, V., Feamster, N., Madhyastha, H., Anderson, T., and Krishnamurthy, A. LIFEGUARD: Practical repair of persistent route failures. In *Proceedings of ACM SIGCOMM* (Helsinki, Finland, Aug. 13–17). ACM Press, New York, 2012.
17. Kim, T. H., Basescu, C., Jia, L., Lee, S.B., Hu, Y., and Perrig, A. Lightweight source authentication and path validation. In *Proceedings of ACM SIGCOMM* (Chicago, IL, Aug. 17–22). ACM Press, New York, 2014.
18. Kushman, N., Kandula, S., and Katabi, D. Can you hear me now? It must be BGP. *ACM SIGCOMM Computer Communication Review 37*, 2 (Apr. 2007), 75–84.
19. Lepinski, M. and Turner, S. *An Overview of BGPsec.* IETF draft, May 8, 2012; http://tools.ietf.org/html/draft-ietf-sidr-bgpsec-overview-02
20. Matsumoto, S., Reischuk, R.M., Szalachowski, P., Kim, T.H.-J., and Perrig, A. Authentication challenges in a global environment. *ACM Transactions on Privacy and Security 20*, 1 (Feb. 2017), 1–34.
21. Palo Alto Research Center. The CCNx Project (Content-Centric Networking); http://blogs.parc.com/ccnx/
22. Perrig, A., Szalachowski, P., Reischuk, R.M., and Chuat, L. *SCION: A Secure Internet Architecture.* Springer, Berlin, Germany, 2017.
23. Raychaudhuri, D., Nagaraja, K., and Venkataramani, A. MobilityFirst: A robust and trustworthy mobility-centric architecture for the future Internet. *ACM SIGMOBILE Mobile Computing and Communications Review 16*, 3 (July 2012), 2–13.
24. Sahoo, A., Kant, K., and Mohapatra, P. BGP convergence delay under large-scale failures: Characterization and solutions. *Computer Communications 32*, 7 (May 2009), 1207–1218.
25. Saltzer, J.H., Reed, D.P., and Clark, D.D. End-to-end arguments in system design. *ACM Transactions on Computer Systems 2*, 4 (Nov. 1984), 277–288.
26. Schuchard, M., Vasserman, E.Y., Mohaisen, A., Kune, D.F., Hopper, N., and Kim, Y. Losing control of the Internet: Using the data plane to attack the control plane. In *Proceedings of the Network and Distributed System Security Symposium* (San Diego, CA, Feb. 6–9). Internet Society, Reston, VA, 2011.
27. Toonk, A. Massive route leak causes Internet slowdown. BGPmon, June 12, 2015; http://www.bgpmon.net/massive-route-leak-cause-internet-slowdown/
28. Zhang, X., Hsiao, H.-C., Hasker, G., Chan, H., Perrig, A., and Andersen, D.G. SCION: Scalability, control, and isolation on next-generation networks. In *Proceedings of IEEE Symposium on Security and Privacy* (Oakland, CA, May 22–25). IEEE Press, 2011.

David Barrera (david.barrera@inf.ethz.ch) is a postdoc in the Network Security Group at ETH Zürich in Switzerland.

Laurent Chuat (laurent.chuat@inf.ethz.ch) is a Ph.D. student in the Network Security Group at ETH Zürich in Switzerland.

Adrian Perrig (aperrig@inf.ethz.ch) is a professor in the Department of Computer Science and leads the Network Security Group at ETH Zürich in Switzerland and an ACM Fellow.

Raphael M. Reischuk (reischuk@inf.ethz.ch) is a senior IT-security researcher at ETH Zürich in Switzerland focusing on network and Web security.

Pawel Szalachowski (psz@inf.ethz.ch) is a senior researcher in the Network Security Group at ETH Zürich in Switzerland.