

2 The SCION Architecture

DAVID BARRERA, LAURENT CHUAT, ADRIAN PERRIG,
RAPHAEL M. REISCHUK, PAWEŁ SZALACHOWSKI

This chapter provides an overview of SCION. The goals to be met by a secure Internet architecture were described in the previous chapter, but to recapitulate briefly, our main aim is to design a network architecture that offers highly available and efficient point-to-point packet delivery, even if some of the network operators and devices are actively malicious. The following chapters describe the SCION architecture in increasing detail.

SCION introduces the concept of an **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 depicted in Figure 2.1. An ISD is administered by multiple ASes, which form the **ISD core***. We refer to these as **core ASes***. An ISD usually also contains multiple regular ASes. The ISD is governed by a policy, called the **trust root configuration (TRC)***, which is negotiated by the ISD core. The TRC defines the roots of trust that are 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 indicates an acceptance of the ISD's TRC. Typically, 3–10 ISPs constitute an ISD core, and their associated customers participate in the ISD. 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, 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. 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. While ISDs may seem to lead to “Balkanization” and prevent an open Internet, they counter-intuitively provide openness and transparency, as we hope to elucidate

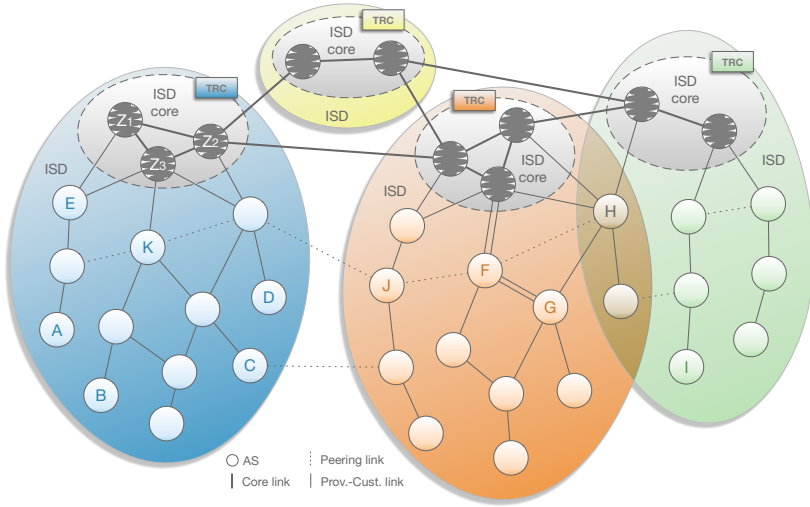


Figure 2.1: Autonomous systems (ASes) grouped into four ISDs. The core ASes are connected via core links. Non-core ASes are connected via customer-to-provider or peering links. AS *H* participates in two ISDs.

in this book (for more information on this point, please refer to the FAQ on Page 409).

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 2.2a). An ISD core AS announces a PCB and disseminates it as a policy-constrained multipath flood either *within* an ISD (to explore intra-ISD paths) or *amongst* core ASes (to explore inter-ISD paths). We refer to this process as *beaconing*. PCBs accumulate cryptographically protected AS-level path information as they traverse the network. This information (which we call hop fields (HF)) within received PCBs is chained together by sources to create a data transmission path segment that traverses a sequence of ASes. Packets thus contain AS-level path information, which avoid the need for border routers to maintain inter-domain forwarding tables. We refer to this concept as **packet-carried forwarding state (PCFS)***.

Figure 2.3 illustrates the chronological sequence of operations required to obtain a forwarding (i.e., end-to-end) path. During the *path exploration* or *beaconing* phase, ASes discover paths to core ASes. *Path registration* allows ASes to transform a few selected PCBs into path segments, and register these path segments with a path infrastructure (making them available for other ASes). The *name resolution* process translates a domain name into its associated **SCION address(es)***. The *path resolution* process allows end hosts to create an end-to-end forwarding path to a destination; it consists of (a) *path lookup*,

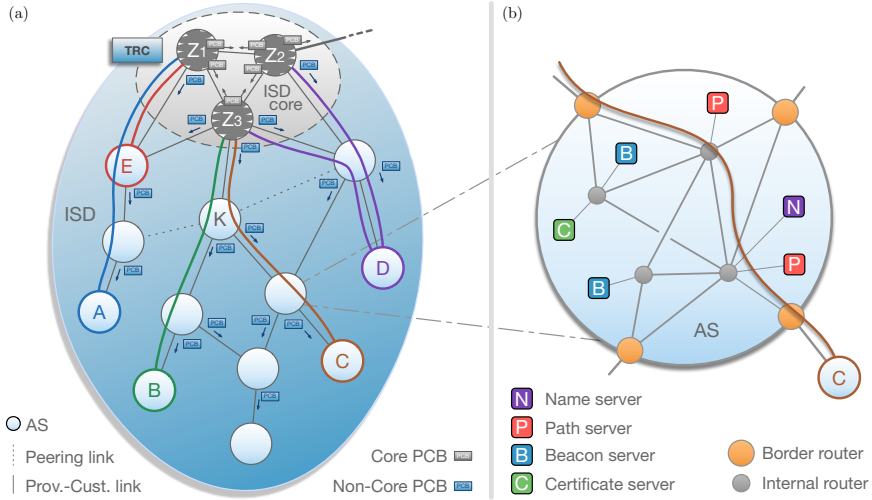


Figure 2.2: (a) ISD with path-segment construction beacons (PCBs) that are propagated from the ISD core to customer ASes, and path segments for ASes A, B, C, D, and E to the ISD core. (b) Magnified view of a SCION AS with its routers and servers. The path from AS C to the ISD core traverses two internal routers.

where the end host obtains path segments, and (b) *path combination*, where an actual forwarding path is created from the path segments. We discuss these phases in this chapter and describe them in more detail in the sections referred to in Figure 2.3.

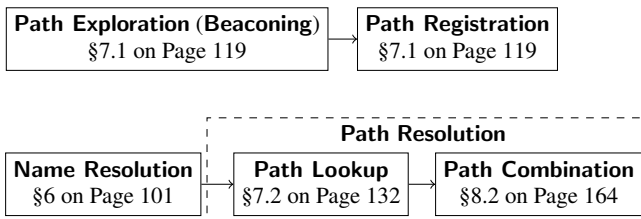


Figure 2.3: Process leading to the creation of a forwarding path.

Servers and Routers

Figure 2.2b shows the main AS components in SCION: **beacon servers*** discover path information, **path servers*** disseminate path information, **certificate servers*** assist with validating path information, and **name servers*** provide

name resolution from user-understandable names to SCION addresses. In addition, *border routers* provide the connectivity between ASes, while *internal routers* forward packets inside the AS.

Beacon servers are responsible for generating, receiving, and propagating PCBs (see Figure 2.2a) to construct path segments, a process we also refer to as beaconing. SCION supports two types of beaconing: intra-ISD beaconing (to construct path segments from a core AS to non-core ASes within an ISD) and inter-ISD beaconing (to construct path segments amongst core ASes within an ISD and across ISDs). Figure 2.4 shows how PCBs originate from a core AS beacon server and are propagated to non-core customer ASes. Non-core AS beacon servers receive these PCBs and re-send them to their customer ASes, which results in AS-level path segments. At every AS, information about the ingress and egress interfaces of the AS is added to the PCB. The ingress and egress interfaces identify the link to a neighboring AS. Periodically, a beacon server generates a set of PCBs, which it forwards to its customer ASes.

Inter-ISD beaconing in SCION is similar to BGP's route-advertising process, although in SCION the process is periodic and PCBs are flooded over policy-compliant paths to discover multiple paths between any pair of core ASes. SCION's beacon servers can be configured to implement all BGP route selection policies, as well as additional properties (e.g., control of upstream ASes) that BGP cannot express (see Section 10.9).

Name servers in SCION perform a similar task to DNS servers in today's Internet: translate a human-understandable name into a SCION address. SCION proposes the RAINS system for this purpose Chapter 6. Based on the (ISD, AS) tuple, end-to-end paths can be looked up and constructed. The end-host address and end-to-end path are then placed in the SCION packet header to enable delivery to a given destination.

Path servers store mappings from AS identifiers to sets of announced path segments, and are organized as a hierarchical caching system similar to today's DNS. Through beacon servers, ASes 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 keep cached copies of TRCs retrieved from the ISD core, keep cached copies of AS certificates, and manage keys and certificates for securing inter-AS communication. Certificate servers are queried by beacon servers when validating the authenticity of PCBs (i.e., when a beacon server lacks a certificate).

Border routers connect different ASes supporting SCION. The main task of border routers is to forward packets. In the case of a packet containing a service address, 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 (e.g., OSPF, SDN, MPLS), the internal routers do not need to be changed.

2.1 Control Plane

We will now discuss the control plane components and mechanisms in more detail. The control plane is responsible for discovering paths and making those paths available to end hosts.

2.1.1 Path Exploration and Registration

Inter-domain beaconing enables core ASes to learn paths to other core ASes. Through intra-domain beaconing, non-core ASes learn path segments leading to core ASes, which enable an AS to communicate with the ISD core. Figure 2.2a shows path segments from ASes *A*, *B*, *C*, and *D* 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 diversity; ASes often have several 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 2.4 demonstrates this feature: AS *F* receives two different PCBs via two different links from a core AS. Moreover, AS *F* uses two different links to send two different PCBs to AS *G*, each containing the respective egress interfaces. AS *G* extends the two PCBs and forwards both of them over a single link to its customer.

An AS typically receives several PCBs representing several path segments to various core ASes. Figure 2.2a shows two path segments for AS *D*, for example. There are three types of path segments:

- A path segment from a non-core AS to the core is an *up-segment*.
- A path segment from the core to a non-core AS is a *down-segment*.
- A path segment between core ASes is a *core-segment*.

However, path segments are typically bidirectional and thus support packet forwarding in both directions. In other words, 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. 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

importance of SCION's path transparency property; a source knows the exact set of ASes and ISDs traversed during the delivery of a packet.

2.1.2 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 first queries a name server to obtain the ISD-AS-address triplet of the destination. The host then 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.

Example. Consider a source host in ISD 1 sending a path lookup request to its local path server, which forwards the request to a core path server. If the requested path's destination AS is within ISD 1, the core path server responds by immediately sending up to k down-segments to the local path server. If the requested path's destination AS is in ISD 2, then the core path server first requests the corresponding down-segments from the core path server in destination ISD 2 before responding to the local path server. In both cases, the local path server returns up to k up- and down-segments to the requesting source (where k is a small integer set to 5 in the current implementation). If the up- and down-segments end at different core ASes, then core segments connecting the core ASes are returned as well.

2.1.3 PCB and Path-Segment Selection

Among the received PCBs, ASes must choose a set of PCBs to propagate further, and a set of path segments to register. These PCBs and path segments 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. To provide some concreteness to this description, we currently use $k = 5$, and send PCBs every 5 seconds to each neighbor over each provider-to-customer link. 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-AS PCBs only traverse core ASes. The same path selection metrics apply, where an AS attempts to forward the set of most desirable paths to its neighbors. A difference, however, is that an AS forwards k PCBs *per* source AS, with $k = 3$. The periodicity is also reduced; we forward PCBs once a minute or upon path changes. Similarly to BGP, this process is inherently not scalable (as the overhead grows linearly with the number of core ASes); however, as the number of ISDs and the corresponding number of core ASes is small, this approach is viable.

2.1.4 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:

- Beaconing occurs every few seconds, constantly establishing new working paths.
- The SCION control message protocol (SCMP) (SCION-equivalent of ICMP) is used for path-segment revocation. As described in detail in Section 7.3, failed links result in rapid erasure of affected path segments from path servers.
- 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 [8]), 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 frequent (although mostly brief) outages in the current Internet [131, 144].

2.1.5 Intra-AS Communication

Communication within an AS is handled by existing intra-domain communication technologies and protocols such as IP with Software-Defined Networking (SDN), or Multi-Protocol Label Switching (MPLS). Figure 2.2b on Page 19 shows one possible intra-domain path through the magnified AS.

2.2 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 an inter-domain 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.

An interesting aspect of this forwarding is enabled by the split of locator (the path towards the destination AS) and identifier (the destination address) [83]: the identifier can have any format that the destination AS can interpret, since only the destination needs to consider that local identifier. In other words, an AS 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 [7]) address. A nice consequence is that an IPv4 host can directly communicate with an IPv6 host over SCION.

In the next two sections, we describe how an end host combines path segments into an end-to-end forwarding path, and how border routers forward packets efficiently.

2.2.1 Path Combination

After name resolution and path lookup, the end host obtains path segments that need to be combined into an end-to-end path. A valid SCION **forwarding path*** can be created by combining up to three path segments, in the following ways (all combinations are illustrated with sample paths depicted in Figure 2.5):

- **Immediate combination of path segments** (e.g., $B \rightarrow D$): the last AS on the up-segment (core AS Z_3) is also the first AS on the down-segment. In this case, the simple combination of an up-segment and a down-segment creates a valid forwarding path.
- **AS shortcut** (e.g., $B \rightarrow C$): the up-segment and down-segment intersect at a non-core AS (e.g., K). In this case, a shorter forwarding path can be created by removing the extraneous part of the path.
- **Peering shortcut** (e.g., $A \rightarrow B$): a peering link (e.g., $L \rightarrow K$) 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.

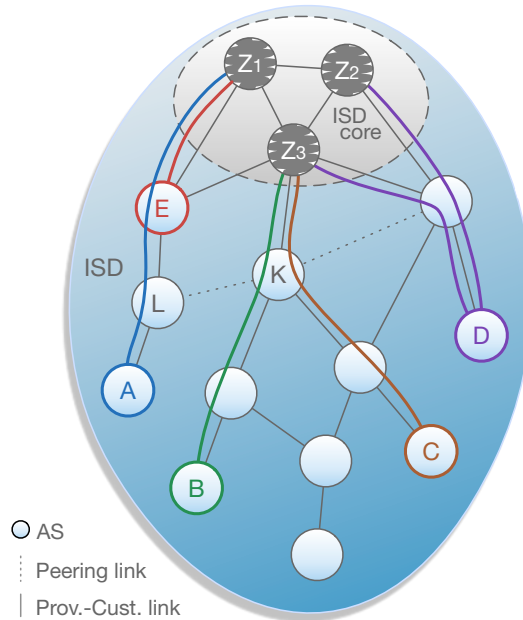


Figure 2.5: ISD with path segments for ASes A , B , C , D , and E .

- **Combination with a core-segment** (e.g., $A \rightarrow D$): the last AS on the up-segment is different from the first AS on the down-segment. This case requires an additional core-segment (e.g., $Z_1 \rightarrow Z_2$) to connect the up- and down-segment. If the communication remains within the same ISD ($A \rightarrow D$), a local ISD core-segment is needed; otherwise (e.g., $A \rightarrow I$ in Figure 2.1), an inter-ISD core-segment is required.
- **On-path** (e.g., $A \rightarrow E$): 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.

2.2.2 Forwarding

Routers can efficiently forward packets in the SCION architecture. In particular, the absence of inter-domain routing tables and the absence of complex longest IP prefix matching performed by current routers enables construction of faster and more energy-efficient routers, which we discuss in more detail in Chapter 14.

The SCION packet header contains a sequence of hop fields (HF), one for each AS that is traversed on the end-to-end path. During forwarding, each AS inspects its respective HF in the packet header. The HF contains interface numbers of the ingress and egress links, which are essentially descriptors of the links across which the packet is entering and exiting the AS. Figure 2.4 on Page 22 depicts how the HF information is assembled in the PCB as part of the beaconing process.

During packet forwarding, a SCION border router at the ingress point of the AS first verifies that the packet entered through the correct ingress interface corresponding to the information in the HF. If the packet has not yet reached the destination AS, the egress interface defines the egress SCION border router — in which case native intra-domain routing (e.g., OSPF, MPLS) is used to send the packet from the ingress SCION border router to the egress SCION border router.

2.3 Security Aspects

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

Similarly to BGPsec [158], each AS signs the PCBs 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 the hop field (HF). Excluding a few flag bits, the structure of the HF is 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.

2.3.1 Algorithm Agility

In terms of cryptographic mechanisms, SCION provides algorithm agility, so 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 PCB selection metric will favor creating paths where each AS on the path supports the new algorithm.

2.3.2 Authentication

Authentication in SCION is based on 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.

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). The TRC 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 serves as a way to bootstrap all authentications.

We now briefly discuss two properties offered by the TRC. *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 between them to sign each other's TRCs — thus, 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.

The authentication of control-plane messages has availability as the main requirement, since the control plane provides communication paths upon which

¹Our reason for not using self-certifying identifiers [7, 180] for long-term identities is their inherent inability to be revoked and the complexities involved with key updates. For short-term identities, however, we do appreciate their features.

other mechanisms rely. Once end-to-end communication is established, additional entities can be contacted to achieve a more secure authentication of end entities (e.g., web servers). The Attack-Resilient PKI (ARPKI) [23] is a highly secure PKI system based on **log servers*** that keep a public log of all certificates to monitor CAs' operations. In turn, CAs and validators verify the content of log servers. By requiring multiple signatures on certificates, and by adding signatures on all operations, we create a situation where multiple malicious trusted entities within the same ISD are needed to perform a man-in-the-middle attack on a single domain. To further increase security, we combine ARPKI with PoliCert, which enables domains to specify their detailed security policy [235]. By storing the domain policies in an ARPKI log, policy consistency and integrity are ensured. In concert, ARPKI and PoliCert achieve a high level of security, as all PKI attacks we have witnessed in the past decade would have been avoided in this framework.

The ISDs and the ARPKI system used in SCION address the problem of CA compromise, as a CA's authority is scoped to the ISDs in which the CA is active, and as multiple trusted entities need to be compromised to perform a successful man-in-the-middle attack. Moreover, the SCION trust roots update mechanism enables revocation within tens of seconds, enabling quick recovery from compromise.

More details on SCION's authentication infrastructure are provided in Chapter 4.

2.3.3 SCION Control Message Protocol (SCMP)

The control plane includes the SCION Control Message Protocol (SCMP), which is similar to the current Internet control message protocol (ICMP), 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, see Section 12.5). In DRKey, each AS uses a local secret key known to its 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.

2.3.4 DDoS Defenses

SCION offers several complementary defenses against link-flooding DDoS attacks, which frequently disrupt daily-life communication (e.g., by exploiting vulnerabilities of IoT devices [138] and launching attacks against IT-security blogger Brian Krebs in September 2016 [140], or against the DNS infrastructure causing outages for Twitter, Spotify, and Reddit in October 2016 [137]).

The SCION architecture comes by default with three mechanisms that provide a strong defense against DDoS attacks:

- **Non-registered (or hidden) path segments:** An AS can prevent an adversary from sending traffic to it by not publicly announcing its down-segment on the path servers. A destination thus cannot be reached, unless it explicitly permits a sender to send traffic. This approach, referred to as off-by-default [19], is explained in more detail in Section 7.2.5.
- **Short-lived paths:** Each SCION path segment has an expiration time, which is set in a PCB to provide several hours of validity. A careful administrator of an AS can let a path segment age and only announce it briefly before the expiration time. For instance when a path segment p that expires within 5 minutes is publicly announced at a path server, then p can only be used to attack the destination AS for at most 5 minutes. The approach here is to publicly announce only short-lived path segments, and to provide longer-lived path segments only to trusted and verified senders.
- **Multipath communication:** Because SCION uses multipath communication by default, an adversary has to congest *all* paths instead of only the single path that is currently used. This approach will prevent attacks that are unable to congest all network paths simultaneously: for example consider a multi-homed domain with two providers, with a 1 Gbps link to each provider. In the current Internet, usually only one of the links will be the active link that carries all incoming and outgoing traffic. If the attacker has a capacity of 1.5 Gbps for example, it can congest that link. Once the victim attempts to change to the other link, the attack traffic will simply follow and congest the alternative link. With multipath communication, however, whichever link the adversary clogs, the other link will still be available and thus communication is always ensured. In summary, multipath communication forces the adversary to *simultaneously* clog *all* paths that are available to the victim, which requires a larger attack capacity and access to *all* paths.

Furthermore, SCION offers two extensions to improve availability and defend against DDoS attacks:

- The SIBRA extension (Chapter 11) enables fine-grained inter-domain bandwidth allocations to guarantee availability even during large-scale DDoS attacks. SIBRA enables fine-grained temporal access, in which

so-called ephemeral paths expire within tens of seconds, putting a rapid stop to a misbehaving sender.

- The OPT extension (Chapter 12) provides source authentication to prevent attacks with spoofed source addresses. Spoofed victim source addresses are used in reflection-based amplification attacks to disguise the attacker's identity and to redirect the response traffic to the actual victim [214].

2.4 Use Cases

SCION improves many aspects of the current Internet. This section highlights some of the applications and use cases that demonstrate unique properties and benefits offered by the new architecture.

2.4.1 High-Availability Communication

Highly available communication is important in many contexts, in particular for critical infrastructures such as financial networks and industrial control systems used for power distribution. Internet outages have been known to wreak havoc on day-to-day operations, for example preventing ATM withdrawals or payment terminal operations [238]. SCION's control-plane isolation through ISDs, its stable data plane, and its multipath operation all contribute to higher availability.

Business continuity refers to the uninterrupted operation of an organization. Business continuity is currently highly dependent on communication. We can witness the increasing inter-connectedness required for business operations when network outages cause a disruption of a surprising number of operations. For instance, when Telecom Malaysia wrongly announced 179,000 IP prefixes to Level3, it caused global outages for 2 hours, even affecting ATM operations in Sweden [238].

Here are a few examples of sectors where availability is crucial:

- Financial services require highly available communication networks, for instance for the distribution of stock market data, real-time market trading, or transaction processing. While critical communication is often sent over leased lines, it is not economical to pervasively use leased lines between all communicating parties. In this setting, SCION can offer higher availability than the legacy Internet at a lower price than a leased line.

High availability for communication is also important in blockchain applications such as bitcoin mining, where a disconnected mining pool does not learn of newly mined coins and is wasting processing on finding irrelevant coins. Similarly, a disconnected mining pool cannot post

its found coins, which will likely be ignored once connectivity is re-established. Both of these cases occur with high probability if the mining pool's computation capacity is less than half the total mining capacity, which is the case for individual mining pools [12].

- Critical command-and-control infrastructures — such as air-traffic, power-grid, or power-plant control systems, or public safety emergency communication — require very high communication availability. Communication disruptions can lead to outages with significant cost for industry and danger for society.
- Governments require high communication availability especially during crisis situations. Examples of critical communication include cables to foreign embassies, law enforcement communication, or access to databases for verifying documents at a country's border.

With SCION's resilience against network-layer DDoS attacks, prevention of prefix hijacking, and data-plane isolation, communication over the regular SCION network can achieve a level of availability that approximates the availability of leased lines. In addition, the SIBRA extension, as described in Chapter 11, offers an extended level of availability through a concept we call DILL, which stands for dynamic inter-domain leased line. DILLs provide a lower bound on the guaranteed bandwidth at inter-domain scale, regardless of the bandwidth requirements of other ASes.

2.4.2 Path Transparency, Path Control, and Compliance with Traffic Flow Regulations

Packets do not always directly reach their destination via the shortest path. Instead, in current practice, many Internet paths take detours. While some extreme cases of detours are due to prefix hijacking [63, 160, 162], most detours are taken for economic reasons or are simply due to the preferred connectivity of ISPs. As a consequence, traffic that would be expected to stay within a geographic area is often routed through nearby countries. For example, paths connecting sources and destinations within Switzerland are sometimes routed through Frankfurt or London, or traffic that would be expected to stay within continental Europe is routed through London.

Path control and transparency are important properties when a sender wants to influence and learn about the ASes that sensitive data will traverse (for legal, secrecy, or safety reasons). For instance, banking or medical data, which is typically bound by strict data privacy regulations, can be constrained in SCION to traverse only selected authorized ASes: a source knows the AS-level path that a packet will follow based on the hop fields in the packet header. Such packet-carried forwarding state in the packet header provides not only transparency, but also path control by letting the source node select the paths amongst a set of paths provided by path servers. Path transparency and control

enable an organization to achieve compliance with laws or regulations that require traffic to be constrained within a jurisdiction. These properties can be further strengthened by SCION's OPT extension (Chapter 12). In a nutshell, OPT provides the receiver with a cryptographically verifiable guarantee that a sequence of ASes were all traversed in the correct order.

2.4.3 Inter-domain Traffic Engineering

In the legacy Internet, only rudimentary forms of inter-domain path control and traffic engineering are possible. For outgoing traffic, one can at best control the next ISP, but only if an AS is multi-homed. A little more path control is available to direct incoming traffic, as an AS can decide to which upstream ISP to send a BGP update. However, to achieve high availability for outages, an IP prefix should be announced to each upstream ISP. AS path pre-pending is a technique that enables a very limited form of path control for incoming traffic; but this technique will not be available in a secure version of BGP, for instance in BGPsec [157, 158].

In intra-domain networks, software-defined networking (SDN) has revolutionized path control; for example, Google has achieved higher network utilization with their B4 system [124]. Analogous to B4's intra-domain path control, SCION makes inter-domain path control available through path registration. An AS can select the down-segments that are announced to the path servers. Hidden paths can be used, which are only communicated to senders who are selected to use them (as discussed in Section 2.3.4). Much path control is available to the sender, who can select which end-to-end path the packet will follow. We anticipate that this level of path control creates a strong reason for adopting SCION.

2.4.4 High-Speed Web Browsing

Current congestion control hinders high-speed communication because the sender and receiver require time to determine their sending rate and to continuously perform congestion control. Consequently, the sending rate is usually below the maximum possible rate. In SCION, through the SIBRA extension (Chapter 11), the sender performs a resource reservation with its initial packet, and the receiver will likely obtain a reservation with a high sending rate, which it can immediately start to use on the reverse path. With such a reservation, a given bandwidth is provided, so no congestion control is needed; consequently, the web server can immediately start sending data at a high rate to the browser.

2.4.5 Mobility Support

With the proliferation of mobile devices, supporting reliable communication can be challenging since these devices frequently connect to and disconnect from (sometimes several) networks. SCION supports high availability and mobility through multipath communication. Moreover, SCION provides a header extension to inform the other party of new down-segments, such that a mobile device that obtains a new address as it connects to a new network can inform the other party about its new down-segment. Failing paths are discarded and new paths are dynamically discovered transparently to users and applications. One challenge, however, is that both sender and receiver might simultaneously move to a new network, and all the previously established paths might fail at the same time. In this unlikely scenario, a name resolution server and a path server need to be contacted to fetch fresh down-segments for the other party [220].

2.5 Incentives for Stakeholders

While SCION offers a wide assortment of security, availability, and performance benefits over current-generation networks, its lack of direct compatibility with BGP may lead to adoption resistance. This resistance can stem from the notion that the cost of changing to the new architecture will be higher than the benefits obtained, or that it is risky to take on a new architecture that may not find widespread adoption. In this section, we discuss deployment incentives to dissipate such reservations.

2.5.1 End Users

End users in SCION benefit primarily from *higher throughput* afforded by the use of native multipath communication, and from *lower latency* due to path control and packet-carried forwarding state. SCION paths are selected based on performance metrics, which translate to better quality of service (such as audio, video, and file transfers) and generally shorter transfer delays. Although the increased size of SCION packets sacrifices goodput, we anticipate that the continuous path optimization of SCION's multipath system will compensate for the higher overhead.

End users also benefit from *higher availability* (i.e., fewer Internet outages) again due to the multipath communication that is used by default. Even if the user's local ISP does not employ SCION, it is possible to provide the benefits of multipath communication via access tunnels as described in Section 10.1.2. Moreover, the SCION-IP gateway (Section 10.3) provides an incremental deployment approach, which enables users to use SCION without requiring changes to software on their devices.

Path control gives users *higher assurance* when performing security-critical tasks such as online banking or shopping. Using SCION, users gain transparency over the communication path to the destination server, allowing them to include or exclude specific paths traversing ASes that are not trusted.

The SCION end-to-end public-key infrastructure offers strong assurance that a contacted web site is the correct entity — fending off man-in-the-middle attacks that could eavesdrop on or alter information sent on a TLS connection. As a consequence, users can perform secure transactions over the Internet with higher confidence.

Finally, SCION extensions (such as Hornet [49] and SIBRA) provide users with a range of additional benefits, such as *high-speed anonymous communication* and *guaranteed bandwidth*.

2.5.2 ISPs

ISPs can create new revenue streams by offering services based on SCION. ISPs that enable SCION can create services for customers who demand higher availability than BGP can provide, but who cannot afford dedicated leased lines. In addition to lower operating cost, SCION gives early adopters increased resilience to network attacks, higher availability, and better path control. ISPs may even offer SCION services to customers of other ISPs through access tunnels. SIBRA, for example, enables inter-domain traffic guarantees, which ISPs cannot offer today unless they operate a global network.

Since SCION PCB propagation policies are more expressive than is possible in BGP, ISPs benefit from finer control of traffic traversing their domain (see Section 10.9), which can help with traffic engineering.

SCION's path transparency properties can provide evidence to regulators and customers that ISPs are not violating network neutrality [194].

Finally, SCION's ISDs and secure operation help to minimize the impact of an ISP's configuration errors, which can simplify ISPs' operations.

2.5.3 Businesses

Businesses or corporations using SCION benefit from path management for incoming and outgoing traffic, path transparency and control, attack resilience, and highly available communication. One particular advantage is that through SCION, a business can ensure that traffic does not leave an ISD. This is important for complying with data privacy laws, which vary from country to country. For example, a recent European Union (EU) ruling declared that companies with an EU presence had to comply with EU data privacy laws, and could no longer make use of "safe harbor" when storing data on servers in approved countries [62]. It is unclear whether forwarding and caching data also falls under this ruling, but SCION allows businesses to specify their traffic policies.

While control over outgoing traffic has so far proven to be an attractive incentive for businesses, control over inbound traffic should also provide an attractive feature. Corporations offering network services to a restricted set of clients (e.g., banks) may want to allow incoming traffic only from those authorized clients or through authorized ISPs. SCION paths are flexible enough to allow this by distributing certain paths to specific authorized entities, rather than announcing them globally.

2.5.4 Governments

Governments using SCION can benefit from the same advantages as businesses, but additionally benefit from avoiding the use of a global trust root. As shown in Section 13.8, a global trust root provides a kill switch that can cause entire networks to be taken offline, which could be particularly damaging in the case of government networks. Like businesses, governments will also value the path control facility that will ensure their traffic traverses ISPs they trust.

The open-source nature of the SCION codebase allows governments to build their own hardware to reduce their reliance on untrusted foreign manufacturers. The codebase can also be inspected and maintained by trusted developers.

2.6 Deployment

Deployability plays a key role in the success of any network architecture. To this end, we have designed SCION to be deployable (by both ISPs and end users) without requiring substantial changes to the existing infrastructure.

2.6.1 Incremental Deployment

As a minimum, an ISP needs to deploy only a single border router capable of encapsulating and decapsulating SCION traffic as it leaves, enters, or traverses its network. SCION ASes must also deploy certificate, beacon, name, and path servers. These servers can run on commodity hardware and can optionally be replicated for increased availability. The current version of the SCION codebase uses IP for internal AS communication, which allows the use of existing intra-domain networking infrastructure and configuration.

We envision that ISDs will grow organically within an area with homogeneous trust. Tier-1 ISPs within those ISDs would become core ASes. SCION facilitates the evolution of ISD and AS structure through efficient updates to the TRC.

Deployment of SCION to end-user sites (e.g., homes or businesses) is designed to require little effort as well, initially needing no changes to hosts or internal network communication equipment. For initial deployment, we achieve

customer-friendly conditions through a gateway device that can be installed in a network to enable both SCION and standard Internet communication. The SCION-IP gateway replaces a home access router and transparently enables any type of communication (legacy IPv4/IPv6 or SCION), as described in Section 10.3.

2.6.2 Deployment Caveats and SCION Disadvantages

The deployment and structure of ISDs is hard to predict, as is which ASes within an ISD will or should become core ASes. We envision that among a group of ASes that deploy a top-level ISD, the AS or ASes that can form peering agreements with core ASes in other ISDs should become core ASes in their own ISD. However, SCION itself does not require or impose strict rules regarding the allocation of ISDs; ISDs can overlap, which means an AS can belong to several ISDs. Sub-ISDs are possible as well, offering the flexibility to start an ISD without needing to peer with core ASes of other ISDs and enabling finer-grained control over routing isolation and authentication. In this context, the important properties SCION offers are path control and transparency: as long as communicating hosts can select and inspect the paths of their packets, the question of ISD partitioning is of secondary nature.

A challenge that could arise is that each AS will attempt to be its own ISD or will want to be part of the ISD core. While too many top-level ISDs will pose a problem for SCION scalability, we observe that economically sound decisions will lead to larger ISDs due to economies of scale — because the startup costs of a core AS are higher than those of a non-core AS, the operation of a large ISD will amortize the cost over more non-core ASes. Moreover, ASes preferentially associate with larger ISDs, which can offer better connectivity to other ISDs as well as to other ASes within the ISD. On the other hand, ISD growth is limited to the extent that entities can agree on the ISD's TRC (i.e., roots of trust). Finally, ASes desiring to be part of the ISD core are assessed in the same way in which current ASes assess peering: an AS is permitted into the core if the current core ASes deem it to be large enough to fulfill core AS duties (which include, for example, participating in beacon and path server replication).

SCION ASes need to manage cryptographic keys, which requires additional effort to securely administer. As a security architecture, every AS has to have a public-private key pair, and obtain a signature on the public key. Although managing cryptographic keys can be a challenge for some ASes, it is a necessity for any secure network architecture. In our development, we are building systems to simplify the overhead of managing cryptographic keys, for instance through our CASTLE system [169], which offers a local low-rate CA environment built from off-the-shelf components. To further mitigate the risks associated with the

management of cryptographic keys, SCION reduces the effect of key loss and compromise by offering approaches for resilience and quick recovery.

As expected in architectures with PCFS, packet headers are necessarily larger. Larger headers place a limit on goodput, since payload space is traded for header space. The current SCION codebase implements the HF as an 8-byte field. Since every AS on an end-to-end path has to be represented through a corresponding HF, the overhead increases linearly with the number of ASes on the path. However, given that the average AS path in today's Internet is four hops long (and decreasing) [66, 141], the overhead introduced by SCION should not exceed around 50 bytes per packet on average. The performance penalty of transmitting more packets appears reasonable since per-packet forwarding performance can be faster than for forwarding-table-based architectures. While the default header size has not turned out to be a performance disadvantage in our testing environment, many of the proposed SCION extensions further increase the header size.

Due to path dissemination and registration dynamics, SCION beacon and path servers can incur a high overhead under specific circumstances. For example, if a given link's state were to fluctuate frequently between available and unavailable (due to error, hardware fault, or an adversary), the beacon server would need to constantly update the set of paths that include that link, and serve new paths excluding that link. We expect that this case will be rare, but also easily detectable. Additionally, higher quality (uptime, availability) links will have a higher probability of selection, minimizing the impact of rapid path fluctuations.

We believe that the basic building blocks of SCION are relatively straightforward to understand and provide many beneficial properties for applications. However, as more extensions and alternative PKIs are added to the architecture, the operational complexity of the architecture increases correspondingly. We believe that this additional complexity is worth the security, efficiency, and availability guarantees provided by the extensions. It is ultimately up to the networking and research community to decide which of these extensions will be deemed worthwhile for pervasive deployment.

2.6.3 SCION Network Deployment

We have deployed a global SCION network, which we are actively using to vet SCION's functionality and security. The current network has about 50 border routers and servers deployed in ASes around the world, with new nodes joining the network on a weekly basis. The deployment status as of December 2016 is described in more detail in Section 10.1.4. Details and requirements for sponsoring a SCION node can be found on our website. The SCION testbed, enabling any researcher to use the SCION network, is described in more detail in Section 10.7.

2.7 Extensions

SCION's extensible architecture enables new systems that can take advantage of the novel properties and mechanisms provided. As compared to the current Internet, most of the benefits can be afforded through the use of PCFS, path transparency, and control. We briefly describe three systems that have been built as extensions to SCION.

Path validation. SCION, through its use of PCFS, paves the way for the Origin and Path Trace (OPT) mechanism (Chapter 12). OPT enables the sender, receiver, and routers to cryptographically verify the path that the packet traversed. By leveraging the DRKey mechanism (Section 12.5), routers can efficiently derive their key, verify the path, and update the path validation fields.

Anonymity and privacy. 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. Proposals such as LAP [113] and HORNET [49] leverage SCION's infrastructure to offer high-bandwidth and low-latency anonymous communication.

DDoS defense. The hierarchical organization of ASes into a manageable number of ISDs enables neighbor-based contracts between pairs of core ASes, which in conjunction with path segments inside the ISDs allows for establishing efficient bandwidth guarantees between any two end hosts (more details are presented in Chapter 11 and Section 13.7.1). Such bandwidth guarantees are provided by the SIBRA extension to 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. The SIBRA extension offers powerful mechanisms for DDoS defense, as it guarantees a lower bound on the bandwidth between any pair of ASes [22], which cannot be lowered even by a large-scale botnet using new types of DDoS attacks such as Crossfire [129] and Coremelt [231].


2.8 Main Contributions


The SCION architecture introduces many new concepts and contributions. Although prior work has proposed related concepts and methods, many of which we build upon, we believe that SCION has advanced the state of the art

by creating a coherent architecture that can be deployed and used in practical environments.





Throughout the book, we highlight some chapters or sections with a diamond symbol in the title to indicate research, engineering, and deployment contributions that we believe are particularly important and interesting. In the remainder of this section, we briefly describe these contributions.

2.8.1 Isolation Domains

The concept of network partitioning and hierarchical domains has been considered since the early days of the Internet [34, 46, 47, 56, 57, 127, 134, 232, 242, 259]. In addition to the scalability sought by previous approaches, SCION's concept of  *isolation domains (ISDs)* (Chapter 3) provides strong security guarantees including meaningful trust roots and the absence of global kill switches. Isolation domains provide control-plane isolation, trust root scoping, and data-plane transparency. Most SCION protocols and extensions rely and build on these properties.

As a design principle, SCION does not require any globally trusted party, and ISDs can operate independently and autonomously. However, there must be a way for them to join the network and be discovered. To this end, in Chapter 5, we present the  *ISD coordination mechanism*, which operates in a distributed fashion without any globally trusted entities. With our mechanism, individual trust decisions made solely by ISDs enable global trust verification, similarly to the PGP web of trust [267], although operating in the constrained environment of large-scale ISPs. The mechanism is based on the rule that trust validations follow routing paths (i.e., commercial relationships). To balance the design tradeoffs, our system allows inconsistencies but makes them visible. It enables determination of network topology and connectivity, from any point of the network, without any central global entity.

2.8.2 Authentication

Another main contribution is SCION's  *authentication infrastructure*, which leverages the properties offered by isolation domains (Chapter 4).  *TRCs* contain the roots of trust of the SCION authentication infrastructure (Section 4.2.1), providing scoped trust, fast and flexible trust root updates, and transparent trust relationships.  *The control-plane PKI* (Section 4.2.3) is a high-availability PKI and is designed to secure SCION's control plane. It ties TRC and certificate distribution to the dissemination of PCBs, thus *removing any circular dependencies* between routing and control-plane PKI operations, which results in efficiency and high availability. On the other hand,  *the end-entity PKI* focuses on achieving high security (see details in Section 4.4). It leverages two recent proposals (i.e., ARPKI [23, 24] and PoliCert [235]). First, it provides

resilience against a selectable number of compromised trusted parties. Second, it allows domain owners to express flexible policies on their TLS certificates and connections.

The control-plane PKI provides network-level authentication, enabling in-network and end host source authentication, which in turn facilitates construction of a variety of secure network protocols. ♦ *The OPT protocol* (Chapter 12) is a source authentication and path validation scheme. It enables end hosts to enforce path compliance according to their path selection, and moreover, it achieves high-speed and stateless operation on routers. OPT relies on ♦ *the DRKey scheme* (Section 12.5), an efficient key derivation mechanism. DRKey allows network entities (e.g., border routers) to derive symmetric keys (shared with destinations) with a negligible computation overhead and without keeping per-destination state. Due to these properties, we use DRKey for ♦ *the authentication of SCMP messages* (SCMP being SCION’s equivalent of ICMP — see Section 4.2.5 and Section 7.6). To the best of our knowledge, it is the first Internet-scale control message protocol with authenticated messages.

As a consequence of scoped trust and isolated control plane, SCION ensures an ♦ *absence of global kill switches* (Section 13.8). No entity can cause an outage of an ISD by performing an operation outside the ISD (such as the revocation of an important key).

2.8.3 Novel Mechanisms and Protocols

Due to its architecture, SCION can intrinsically support multiple novel mechanisms and protocols. For instance, ♦ *RAINS* provides a next-generation name resolution system (Chapter 6). The control plane allows the definition of ♦ *flexible path policies*, enabling implementation of BGP route policies and definition of policies that cannot be expressed in BGP (Section 10.9). Furthermore, SCION’s ♦ *data plane* (Chapter 8) provides *highly efficient and secure packet forwarding*. The forwarding path is encoded within each packet and is cryptographically protected. To make a forwarding decision, the border router checks whether the relevant information is fresh and was authorized by its AS. To this end, efficient symmetric cryptography is used. Moreover, the cryptographic mechanisms required are widely supported by modern hardware; thus, a high-speed SCION border router can be built on commodity hardware.

Another example is the ♦ *AS-level anycast infrastructure* (Section 7.5), which provides a service-oriented infrastructure enabling a packet to be delivered to the nearest server of a given service. This infrastructure is an especially powerful mechanism when used for building services that can take advantage of hierarchical caching.

Although path infrastructures have also been explored in other Internet architectures, SCION introduces a novel ♦ *secure path revocation system* (Section 7.3). Our path revocation system works on the link level. Its main novelty

is a traffic-driven fault detection and failed-link revocation mechanism. The revocations are disseminated as responses to data packets that encounter a failed link. In this design, the system quickly disseminates revocations only to entities that have used failed paths, thus avoiding the overhead of informing entities that do not use those paths. To the best of our knowledge, it is the first secure and practical inter-domain link revocation scheme. The scheme also provides *authenticated failed-link localization*.

2.8.4 Resource Allocation

Another main contribution is 💎 *SIBRA* (Chapter 11), a SCION extension that implements *global bandwidth resource allocation*. SIBRA's main objective is to provide DDoS attack defense, and it is realized through end-to-end bandwidth allocation. The system provides *botnet-size independence*, a property that no prior DDoS defense system could achieve. A main feature of SIBRA is its per-flow stateless **fastpath*** packet forwarding.

2.8.5 Deployment and Evolvability

Finally, SCION makes the following deployment contributions. 💎 *The SCION-IP gateway* (Section 10.3) provides an easy and flexible way of *interconnecting SCION with the current Internet*. It supports a variety of connection and deployment variants. The gateway can be used by ISPs, organizations, or individual users to bootstrap and benefit from the deployment of SCION even for their legacy clients and legacy IP communication.

Taking into consideration the lessons learned from Internet deployment, SCION is designed to support and deploy new mechanisms. Flexible *extension mechanisms* are built into both the data and control planes (Section 15.1.4 and Section 15.3.4), which enables the architecture to evolve. Furthermore, in the spirit of evolvability and maintenance, SCION supports 💎 *algorithm agility* (Section 17.1), which is crucial in the context of cryptographic algorithms (as over time they become weaker or become vulnerable to a newly discovered attack).

SCION: A Secure Internet Architecture

Perrig, A.; Szalachowski, P.; Reischuk, R.M.; Chuat, L.

2017, XX, 432 p. 120 illus., 78 illus. in color., Hardcover

ISBN: 978-3-319-67079-9