Transparent Security Technology and Potential Data Privacy Implications

Purpose

To provide a non-technical audience (e.g., privacy officers, data protection officers, and other privacy professionals) with an overview of Transparent Security technology and the personal data used by the technology.

Transparent Security Technology Summary

CableLabs' Transparent Security technology enables the near real-time identification and mitigation of Distributed Denial of Service (DDoS) attack traffic at its source. DDoS attacks cost the broadband industry billions of dollars each year in malicious traffic delivery costs, traffic scrubbing, and service downtime. As an increasing number of devices connect to the network, there are more vulnerable points of attack for malicious actors. CableLabs created Transparent Security is an open source solution for identifying and mitigating DDoS attacks and the devices (e.g., IoT) that are the source of those attacks. Transparent Security allows operators to block DDoS traffic on their networks much closer to its source and the point of attack. In comparison to traditional DDoS mitigation approaches, Transparent Security may protect the network from both internal and externally generated DDoS threats, through monitoring ingress and egress traffic at every point in the network, and reacting quickly to new threats at an unprecedented level of granularity.

Transparent Security is enabled through a P4-based programmable data plane and in-band network telemetry (INT) data. P4 fundamentally changes how network systems are designed.[1] With P4, network design requirements are defined; a P4 program is written to implement those requirements, define how packets are processed by the network; and then the P4 program is compiled and deployed to control the forwarding elements in the network. P4 brings all the benefits of software engineering (composing programs, debugging, code coverage, provable behavior, model checking, etc.) to the design of network systems, all the way down to the physical wire.[2] INT is a method for adding telemetry data to every packet at multiple points on the network. This approach can help determine things such as the path of a packet or the source device emitting packets. With INT, the packet only needs to be inspected at the edge of the network, which reduces the overhead and generates less traffic compared with sampling at multiple points in the network.[3]

Transparent Security leverages the P4's ability to inspect every packet and add additional INT data to the packet header. Using this detailed data about each packet, including source device, exact route through the network, and travel duration, Transparent Security identifies attack traffic and then blocks that traffic in less than a second. The detailed machine learning and other data analytics may be done within operator's network or through a third-party cloud application. Transparent Security is enabled by currently available programmable chips that can process packets at line speed and be deployed at any point in the network, from the core network to residential and business customer premises.

Current solutions are deployed at interconnection points, which can only mitigate external DDoS attacks, can take several minutes to recognize the attack, are only capable of identifying large areas of the network that are infected, and cannot monitor egress traffic. These solutions require the traffic to be routed out-of-band and through a 'scrubber'

---

[1] P416 Language Specification, The P4 Language Consortium, v.1.2.1 (June 11, 2020), https://p4.org/p4-spec/docs/P4-16-v1.2.1.html#sec-benefits-of-p4.

[2] Cablelabs/transparent-security, GitHub, available at https://github.com/cablelabs/transparent-security; and Randy Levensalor, *Vaccinate Your Network to Prevent the Spread of DDoS Attacks*, CableLabs (Oct 2, 2019), https://www.cablelabs.com/vaccinate-your-network-to-prevent-the-spread-of-ddos-attacks. *See also*, Jen Rexford and Nick McKeown, *Let's Get Started*, P4 Language Consortium (May 29, 2015), https://p4.org/p4/lets-get-started.html.

[3] Randy Levensalor, *Vaccinate Your Network to Prevent the Spread of DDoS Attacks*, CableLabs (Oct 2, 2019), https://www.cablelabs.com/vaccinate-your-network-to-prevent-the-spread-of-ddos-attacks; and In-band Network Telemetry (INT) Dataplace Specification, version 2.1, GitHub (Nov. 11, 2020), available at https://github.com/p4lang/p4-applications/blob/master/docs/INT_v2_1.pdf.

system in order to detect malicious traffic, which increases response time and introduces additional latency into the network. The ability to quickly identify compromised devices and DDoS packets reduces malicious traffic delivery expenses for operators, service interruptions for subscribers, and costly out-of-band mitigation efforts deeper in the core network.

Transparent Security provides substantial improvements over these current solutions through the collection and processing of detailed network data that in many instances may qualify as personal data.

Possible Types of Data Collected by Transparent Security

Specifically, Transparent Security captures a copy of the full header of each packet that originates or terminates on the operator's network. To facilitate this high-speed capture, Transparent Security copies the first fixed number of bits of each packet and then may capture a portion of the data payload in cases where the header is shorter than average. The following table provides a list of the information that is contained in the packet header and a description of the data field.

Table 1. Example Transparent Security Header Fields and Type of Data Collected

Header Fields	**Type of Data Collected**
Source IP Address	Unique identifier linking the packet to a specific subscriber or subscriber household
Destination IP Address	Unique identifier potentially linking the packet to a specific application, website, or service (e.g., Netflix, Bank of America, United Healthcare)[4]
MAC Address	May identify the subscriber's device type manufacturer
Node ID	May create a persistent unique identifier (must be unique within the service provider domain) for a specific local network or network element that could be connected with a specific customer. Such as a gateway device on the customer's premises or a networking device inside the operator's network.
Traffic characteristics (e.g., buffer occupancy, drop reason, hop latency, etc.)	May identify unique traffic patterns associated with specific programs, services, or devices that may be present on the home network
Any OSI networking model Layer 2 through Layer 7 headers, including the application header, presentation header, session header, transport header, network header, data link header, and parts of the physical data payload	May identify unique traffic patterns associated with specific programs, services, or devices may be present on the home network

Table 1 shows header fields used by Transparent Security technology along with the type of data derived from these fields. This degree of data granularity is exactly what allows network operators to identify and mitigate DDoS traffic patterns, although the data required to conduct this analysis potentially includes data that may qualify as personal data. As shown in Table 1, the DDOS analysis requires the header of each packet, which typically includes source and destination IP addresses and device MAC addresses, among other data fields.

The Transparent Security prototype implementation provides that a packet fragment including all original packet headers, INT data, and the truncated payload where the total number of bytes of the telemetry report shall not exceed the configured number of bytes (reference implementation currently set at 200). The size of the packet fragment can be configured per deployment and may vary based the packet type (i.e., IPv4, IPv6).[5] While the exact size of the packet fragment will depend on the implementation, it will likely include elements of the packet such as the L3 header (e.g., IP), L4 header (e.g., UDP, TCP, SCTP), and L4 header (e.g., http, https) and possibly even more information such as fragments of the payload which may or may not be encrypted. In many cases this includes the packet destination address along with the identity of the device itself. Additionally, some portion of the potentially unencrypted packet payload can be swept into the fragment depending on the configuration and actual lengths of headers.

[4] See, e.g., Simran Patil, *What Can You Learn from an IP Address?* APNIC (Aug. 23, 2019), https://blog.apnic.net/2019/08/23/what-can-you-learn-from-an-ip-address/.
[5] *Id*.

The protype Transparent Security architecture employs a series of telemetry reports to deliver the INT data and standard header data to an analytics engine for identifying potential attacks.[6] The telemetry reports use two kinds of reports in order to identify and mitigate attacks; tracked flows[7] used to sample the forwarded traffic and drop reports[8] used to determine if an attack is being mitigated. The hop-by-hop INT header will follow the header as described in section 4.7 of the specification.[9] The hop-by-hop metadata record will be updated at each hop on the network which supports the INT header (at P4-enable elements). Each metadata record corresponds to a bit field in the instruction set and is 4 octets long. The only bit required to be set for Transparent Security is “bit0”, however other bits can be set as indicated by the INT specification. Bit0 is set at the customer's gateway, the gateway enters its ID as the switch ID.[10] In the prototype implementation, a packet is generated from a device on a customer premises and the gateway on the customer premises inserts the MAC address from the source filed in the Ethernet frame. The gateway then adds its own switch ID (which is configurable) and the packet is sent through further network switches, where they each add their switch IDs to establish the packet’s pathway through the network. The device INT metadata thus includes, at least, an originating device ID MAC address, a gateway switch ID, and one or more subsequent switch IDs. Depending on real-world design, implementation, and business decisions, the analytics engine data processing may require sending the telemetry reports and header data to a third-party vendor or service provider.

[6] This will leverage the following specification from the P4 Organization; *See* Telemetry Report Format Specification, The P4.org Applications Working Group, v2.0 (Oct. 8,2020), available at https://github.com/p4lang/p4-applications/blob/master/docs/telemetry_report_v2_0.pdf.

[7] The sampled reports will be categorized as "Tracked Flows" in section 2.2 of the specification. These will only be generated by the sink switches that will typically, be core switches or edge routers. The tracked flows reports will be used to identify when an attack begins and its source and will generate reports in the INT-MD (eMbed Data) mode as defined in section 2.4.2 of the specification.

[8] Drop reports will be generated by all devices which are mitigating the attack. This can be used to track the effectiveness of the mitigation and when the attack has stopped.

[9] INT Hop-by-Hop Metadata Header Format in the current INT specification

[10] Bit0: Switch ID: Unique identifier for the switch (4 octets).