Network Working Group T. Hardie, Ed.
Internet-Draft October 28, 2016
Intended status: Informational
Expires: May 1, 2017

Path signals
draft-hardie-path-signals-00

Abstract

TCP's state mechanics uses a series of well-known messages that are exchanged in the clear. Because these are visible to network elements on the path between the two nodes setting up the transport connection, they are often used as signals by those network elements. In transports that do not exchange these messages in the clear, on-path network elements lack those signals. This document discusses the nature of the signals as they are seen by on-path elements and reflects on best practices for transports which encrypt their state mechanics.

Status of This Memo

Copyright Notice

Table of Contents

1. Terminology

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in RFC 2119 [RFC2119].

2. Introduction

TCP [RFC0793] uses handshake messages to establish, maintain, and close connections. While these are primarily intended to create state between two communicating nodes, these handshake messages are visible to network elements along the path between them. It has been common over time for certain network elements to treat the exchanged messages as signals which related to their own functions.

A firewall may, for example, create a rule that allows traffic from a specific host and port to enter its network when the connection was

initiated by a host already within the network. It may subsequently remove that rule when the communication has ceased. In the context of TCP handshake, it sets up the pinhole rule on seeing the initial TCP SYN acknowledged and then removes it upon seeing a RST or FIN & ACK exchange. Note that in this case it does nothing to re-write any portion of the TCP packet; it simply enables a return path that would otherwise have been blocked.

When a transport encrypts the headers it uses for state mechanics, the signal path elements inferred from examination is no longer available. Their behavior in its absence will depend on which signal is not available, on the default behavior configured by the path element administrator, and by the security posture of the network as a whole.

3. Signals Type Inferred

The following list of signals which may be inferred from transport state messages includes those which may be exchanged during sessions establishment and those which derive from the ongoing flow. Some of these signals are derived from the direct examination of packet trains, such as using a sequence number gap pattern to infer network reliability; others are derived from association, such as inferring network latency by timing a flow's packet inter-arrival times. This list is not exhaustive, and it is not the full set of effects due to encrypting data and metadata in flight.

3.1. Session establishment

One of the most basic inferences made by examination of transport state is that a packet will be part of an ongoing flow; that is, an established session will continue until messages are received that terminate it. Path elements may then make subsidiary inferences related to the session.

3.1.1. Session identity

Path elements that track session establishment will typically create a session identify for the flow, commonly using a tuple of the visible information in the packet headers. This is then used to associate other information with the

3.1.2. Routability and Consent

A second common inference is that the session establishment provides is that the communicating pair of hosts can each reach each other and are interested in continuing communication. The firewall example given above is a consequence of the inference of consent; because the

internal host initiates the connection, it is presumed to consent to return traffic. That, in turn justifies the pinhole.

3.1.3. Resource Requirements

An additional common inference is that network resources will be required for the session. These may be requirements within the network element itself, such as table entry space for a firewall or NAT; they may also be communicated by the network element to other systems. For networks which use resource reservations, this might result in reservation of radio air time, energy, or network capacity.

3.2. Network Measurement

Some network elements will also use transport messages to engage in measurement of the paths which are used by flows on their network. The list of measurements below is illustrative, not exhaustive.

3.2.1. Path Latency

There are several ways in which a network element may measure path latency using transport messages, but two common ones are examining exposed timestamps and associating sequence numbers with a local timer. These measurements are necessarily limited to measuring only the portion of the path between the system which assigned the timestamp or sequence number and the network element.

3.2.2. Path reliability and consistency

A network element may also measure the reliability of a particular path by examining sessions which expose sequence numbers; retransmissions and gaps are then associated with the path segments on which they might have occurred.

4. Options

The set of options below are alternatives which optimize very different things. Though it comes to a preliminary conclusion, this draft intends to foster a discussion of those tradeoffs and any discussion of them must be understood as preliminary.

4.1. Do not restore these signals

It is possible, of course, to do nothing. The transport messages were not necessarily intended for consumption by on-path network elements and encrypting so they are not visible may be taken by some as a benefit. Each network element would then treat packets without these visible elements according to its own defaults. While our

experience of that is not extensive, one consequence has been that state tables for flows of this type are generally not kept as long as those for which sessions are identifiable. The result is that heartbeat traffic must be maintained to keep any bindings (e.g. NAT or firewall) from early expiry.

4.2. Replace these with network layer signals

It would be possible to replace these implicit signals with explicit signals at the network layer. Though IPv4 has relatively few facilities for this, IPv6 hop-by-hop headers [RFC7045] might suit this purpose. Further examination of the deployability of these headers may be required.

4.3. Replace these with per-transport signals

It is possible to replace these implicit signals with signals that are tailored to specific transports, just as the initial signals are derived primarily from TCP. There is a risk here that the first transport which develops these will be reused for many purposes outside its stated purpose, simply because it traverses NATs and firewalls better than other traffic. If done with an explicit intent to re-use the elements of the solution in other transports, the risk of ossification might be slightly lower.

4.4. Create a set of signals common to multiple transports

Several proposals use UDP[RFC0768] as a demux layer, onto which new transport semantics are layered. For those transports, it may be possible to build a common signalling mechanism and set of signals, such as that proposed in "Transport-Independent Path Layer State Management" [I-D.trammell-plus-statefulness].

This may be taken as a variant of the re-use of common elements mentioned in the section above, but it has a greater chance of avoiding the ossification of the solution into the first moving protocol.

5. Recommendation

Fundamentally, this paper recommends that implicit signals should be replaced with explicit signals, but that a signal should exposed to the path only when the signal's originator intends that it be used by the network elements on the path. For many flows, that may result in signal being present, but it allows them to be present when needed.

Discussion of the appropriate mechanism(s) for these signals is continuing but, at minimum, any method should meet the principles set out in the security considerations below.

6. IANA Considerations

This document contains no requests for IANA.

7. Security Considerations

Addition of visible signals to the path allows network elements along the path to act. If the network element is controlled by an attacker, those actions can include dropping, delaying, or mishandling the constituent packets of a flow.

Note that actions that do not benefit the flow or the network may be perceived as an attack even if they are conducted by a responsible network element. Designing a system that minimizes the ability to act on signals at all by removing as many signals as possible may reduce this possibility. This approach also comes with risks, principally that the actions will continue to take place on an arbitrary set of flows.

Addition of visible signals to the path also increases the information available to an observer and may, when the information can be linked to a node or user, reduce the privacy of the user.

This document recommends three basic principles:

- o Cryptographic contexts should be available on any flow, derived from ubiquitous end-system cryptographic capabilities.

- o Anything exposed to the path should be done with the intent that it be used by the network elements on the path.

- o Intermediate path elements should not add visible signals which identify the user, origin node, or origin network [I-D.hardie-privsec-metadata-insertion].

8. Acknowledgements

In addition to the editor listed above, this document incorporates contributions from Brian Trammel, Mirja Kuehlwind, and Joe Hildebrand. These ideas were also discussed at the PLUS BoF, sponsored by Spencer Dawkins. The ideas around the use of IPv6 hop-by-hop headers as a network layer signal benefited from discussions with Tom Herbert. The description of UDP as a demuxing protocol comes from Stuart Cheshire.

All errors are those of the editor.

9. References

9.1. Normative References

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, DOI 10.17487/RFC2119, March 1997, <http://www.rfc-editor.org/info/rfc2119>.

9.2. Informative References

[I-D.hardie-privsec-metadata-insertion]
Hardie, T., "Design considerations for Metadata Insertion", draft-hardie-privsec-metadata-insertion-03 (work in progress), October 2016.

[I-D.trammell-plus-statefulness]
Kuehlewind, M., Trammell, B., and J. Hildebrand, "Transport-Independent Path Layer State Management", draft-trammell-plus-statefulness-00 (work in progress), October 2016.

[RFC0768] Postel, J., "User Datagram Protocol", STD 6, RFC 768, DOI 10.17487/RFC0768, August 1980, <http://www.rfc-editor.org/info/rfc768>.

[RFC0793] Postel, J., "Transmission Control Protocol", STD 7, RFC 793, DOI 10.17487/RFC0793, September 1981, <http://www.rfc-editor.org/info/rfc793>.

[RFC7045] Carpenter, B. and S. Jiang, "Transmission and Processing of IPv6 Extension Headers", RFC 7045, DOI 10.17487/RFC7045, December 2013, <http://www.rfc-editor.org/info/rfc7045>.

Author's Address

Ted Hardie (editor)

Email: ted.ietf@gmail.com