Network Working Group
Internet-Draft
Intended status: Standards Track
Expires: July 26, 2020

L. Iannone
Telecom Paris
D. von Hugo
Deutsche Telekom
B. Sarikaya
Denpel Informatique
E. Nordmark
Zededa

January 23, 2020

Privacy issues in Identifier/Locator Separation Systems
draft-iannone-pidloc-privacy-00

Abstract

There exist several protocols and proposals that leverage on the Identifier/Locator split paradigm, having some form of control plane by which participating nodes can share their current Identifier-to-Location information with their peers. This document explores some of the privacy considerations for such a type of system.

Status of This Memo

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at https://datatracker.ietf.org/drafts/current/.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

This Internet-Draft will expire on July 26, 2020.

Table of Contents

1. Introduction

When the IP address is separated, one way or another, into an identifier and a locator, there is typically the need to be able to look up an identifier to find possible locators which can be used to reach the identified endpoint. If such a system (think a distributed database) was publicly available, then this would introduce additional privacy considerations which do not exist in the absence of the ID/locator split. Think for instance if identifiers are assigned to devices such as mobile phones which have a strong binding with an individual. Having the location of such identifier publicly available implies make the individual whereabouts public.

Without an ID/locator split, a device is already providing its IP address (in the form of a source address) to any network device along the path, and also to the remote endpoint. That endpoint in particular might use IP geolocation databases to get a pretty good idea of where its peer is located, for instance to offer information and/or advertising relevant to that location.

However, in such scenario, when a device (e.g., a laptop or smartphone connected over WiFi) moves (e.g., from home to a coffee shop) the IP address changes. This makes it harder for network devices along the paths to realize that it is the same mobile device. And if the mobile device is not retaining cookies or logged into websites, those remote peers would also have some difficulty determining whether it is the same mobile device. Furthermore, a mobile device which is using typical cellular network technologies ends up with an IP address, at least as seen by remote peers outside of the cellular network, which is associated with the cellular operator but does not necessarily indicate a particular location of the mobile device.

Note that even if the IP address isn't always useful to track a mobile device today, there are several mechanisms higher in the stack which can do this. For instance cookies or SSL sessions, applications which share GPS location, or operators who offer additional location information (for instance based on which cellular base station a mobile device is using) to business partners.

Promising proposals are Identifier Locator (Id-Loc) separation systems like, Identifier-Locator Network Protocol (ILNP) [RFC6740], Locator/ID Separation Protocol (LISP) [I-D.ietf-lisp-rfc6830bis] [I-D.ietf-lisp-rfc6833bis], Virtual eXtensible LAN [RFC7348], and others.

Architectures and protocols for these approaches are already documented in detail and are under continuous evolution in different WGs. This document on the other hand attempts to identify potential issues with respect to real-world deployment scenarios, which may demand for implementations of the above-mentioned Id-Loc systems. In particular, this document overviews issues related to threats due to privacy violation of devices and their users, as well as location detection and movement tracking, where specific countermeasures may be needed.

2. Keywords and Terminology

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "NOT RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in BCP 14 [RFC2119] [RFC8174] when, and only when, they appear in all capitals, as shown here.

Identifier: An identifier is information allowing to unambiguously identify an entity or an entity group within a given scope. An identifier is the equivalent of an End-point IDentifier (EID) in The

Locator/ID Separation Protocol (LISP). It may or may not be visible in communications.

Locator: A locator is a routable network address. It may be associated with an identifier and used for communication on the network layer according to identifier locator split principle. A locator is the equivalent of a Routing Locator (RLOC) in LISP or an IP address in other cases.

3. Identifier Locator Split Protocols and Use-Cases

3.1. Identifier Locator Separation Protocols

Identifier represents a communication end-point of an entity and may not be routable. Locator also represents a communication end-point, however, it is a routable network address. Because entities identified by an Identifier can move the association between Identifiers and Locators may be ephemeral. A database called a mapping system needs to be used for Identifier to Locator mapping. Identifiers are mapped to locators for reachability purposes. A mapping system has to handle mobility by updating the identifier to locator mappings in the database.

To start the communication, a device needs to know the identifier of the destination, hence it relies on a identifier lookup process to obtain the associated locator(s). Note that both identifier and locator may be carried in clear in packet headers, depending on the specific technology used and the level of security/privacy enforced.

Usage of identifiers readily available for public access raises privacy issues. For public entities, it may be desirable to have their fully qualified domain names or host names available for public lookups by the clients, however, this is not the case in general for all identifiers, e.g. for individuals roaming in a mobile network.

3.1.1. ILNP

Identifier-Locator Network Protocol (ILNP) [RFC6740] is a host-based approach enabling mobility using mechanisms that are only deployed in end-systems and do not require any router changes.

3.1.2. VxLAN

Virtual Extensible LAN (VXLAN) [RFC7348] is a network virtualization technology that attempts to address the scalability problems associated with large cloud computing deployments. It uses a VLAN-like encapsulation technique to encapsulate layer 2 Ethernet frames within layer 4 UDP datagrams, using 4789 as the default IANA-assigned

destination UDP port number. VXLAN endpoints, which terminate VXLAN tunnels and may be either virtual or physical switch ports, are known as VXLAN tunnel endpoints (VTEPs) and can be considered the locators of the devices in the extended VLAN.

3.1.3. LISP

Locator/Id Separation Protocol (LISP) [I-D.ietf-lisp-rfc6830bis] [I-D.ietf-lisp-rfc6833bis] is based on a map-and-encap approach, which provides a level of indirection for routing and addressing performed at specific ingress/egress routers at the LISP domain boundaries. Such border routers performing LISP encapsulation at the packet's source stub network are indicated as Ingress Tunnel Routers (ITRs), while border routers at the packet's destination stub network are called Egress Tunnel Routers (ETRs), all of them are indicated by the general term xTRs. In order to obtain mappings used for encapsulation operation, xTRs query the mapping system in order to obtain all mappings related to a certain EID only when necessary (usually, but not exclusively, at the beginning of a new flow transmission). The LISP control plane protocol [I-D.ietf-lisp-rfc6833bis] allows to support several different mapping systems (e.g., LISP+ALT [RFC6836] and LISP-DDT [RFC8111]). More than that, it can actually also be applied to various other data plane protocols.

3.2. Use Cases

The collection of use-cases shall serve as an overview of possible Loc-ID split application and help in identifying different issues in privacy and security in generic Identifier Locator Split approaches.

3.2.1. Industrial IoT

Sensors and other connected things in the industry are usually not personal items (e.g. wearables) potentially revealing an individuals sensitive information. Yet, industrial connected objects are business assets which should be detected/accessed only by authorised intra-company entities. Since the huge amount of these things (massive IoT) as well as the typical energy and bandwidth constraints of battery-powered devices may pose a challenge to traditional routing and security measures.

In Industrial IoT, there are very strong reasons to not share the ID/Locator binding with third parties, i.e. retain the privacy. This can be achieved in a number of ways such as: using an ID/locator system but using some fixed anchor point as a locator; injecting routing prefixes for the ID prefixes into the normal routing system and use proxy indirection; providing limited ID/Locator exposure.

These are just examples, more approaches should be explored in order to find which one is the most suitable in the context of industrial IoT.

3.2.2. 5G Use Case

Upcoming new truly universal communication via so-called 5G systems will demand for much more than (just) higher bandwidth and lower latency. Integration of heterogeneous multiple access technologies (both wireless and wireline) controlled by a common converged core network and the evolution to service-based flexile functionalities instead of hard-coded network functions calls for new protocols both on control and user (data) plane. While Id-Loc approach would serve well here, the challenge to provide a unique level of security and privacy even for a lightweight routing and forwarding mechanism - allowing for ease of deployment and migration from existing operational network architecture - remains to be solved.

3.2.3. Cloud Use Case

The cloud, i.e. a set of distributed data centers for processing and storage connected via high-speed transmission paths, is seen as logical location for content and also for virtualized network function instances and shall provide measures for easy re-location and migration of these instances deployed as e.g. containers or virtual machines. Id-Loc split routing protocols are proposed for usage here as in VXLAN [RFC7348] and LISP [I-D.ietf-lisp-rfc6830bis] [I-D.ietf-lisp-rfc6833bis] while the topology of the cloud components and logical correlations shall be invisible from outside.

In a cloud, an upstream IP address does not necessarily belong to the actual service location, but a gateway or load balancer. So, the locator or also ID reveal the location with the accuracy of a data center, not the function taking a service request. This issue also manifests itself in today's 4G cellular networks (LTE/EPS) as the Packet Data Network Gateways (PGWs) [3GPP] interfacing the Internet are often realised already virtually in a data center, binding UEs' IP addresses which are from the network of the data center.

4. Assumptions

We assume that there are benefits associated with sharing ID to locator mappings with some peers sometimes. Those benefits can be

- Lower latency and higher bandwidth: If two peer devices have some locators which are topologically closer, then sharing all the locators means that the devices can find a shorter path (fewer hops and/or shorter round-trip time), or find a path, which offers higher throughput, then if the devices only shared some form of default

locator.

- Higher availability and robustness: If two peer devices share all their locators, then if there is some network outage the devices can autonomously discover a working path using the different locator pairs.

However, those benefits do not imply that it is a good idea to always share all of the locators with everybody. That would make tracking by third parties trivial.

A device can obfuscate itself by, instead of using a single long-lived identifier, using multiple short-lived identifiers. In that case the value to the ID/locator binding for any particular identifier would be lower. However, this assumes that the device can ensure no relation between the different identifiers it is using, either concurrently or over time. Also, some of the benefits above implicitly assume that there can be some long-lived sessions or associations between pairs of identifiers. For instance, if a device would need to go fetch the current identifier of its peer from some removed system, then it might not experience improved robustness since that fetch might depend on the failed external connectivity. Thus we believe that we can explore the core of the ID/locator privacy issue by looking at long-lived identifiers.

5. Threats against Privacy

There seem to be at least two different privacy threats relating to ID/locator mapping systems.

5.1. Location Privacy

If a third party can at any time determine the IP location of some identifier, then the device can at one point be IP geolocated at home, and later a coffee shop.

5.2. Movement Privacy

If a third party can determine that an identifier has changed locator(s) at time T, then even without knowing the particular locators before and after, it can correlate this movement event with other information (e.g., security cameras) to create a binding between the identifier and a person.

6. Not everybody all the time

In order to see the benefits about but minimizing the privacy implication one can explore limiting to which peers and when the ID/locator binding are exposed.

A few initial examples help illustrate this.

6.1. Optimized Routing

If some operator of a network where there is a large amount of mobility wants to ensure efficient routing, then an ID/locator split approach might make sense. Such a system can potentially be limited to the set of devices (routers etc.) which are under the operators control. If this is the case, then the ID/locator mapping system can provide access control so that only those trusted devices can access the mappings.

Note that from a privacy perspective this isn't any different than the same operator using a link-state routing protocol to share host routes for all the mobile devices. In that case all participants in the link-state protocol can determine the location (attached to which router) and notice any mobility events. Of course, there are significant non-privacy differences between those two approaches.

Exposing the ID/locator mapping to attached devices (e.g., any mobile devices which wouldn't be trusted to participate in the link-state routing counterpart approach), will change the privacy implications.

6.2. Family and Friends

There are cases where it is quite reasonable to share location information with other family members or friends. For instance, young children might run applications which enable their parents to track them on their way to/from school. And I might share my location with friends so we can more easily find each other while out in town.

Today such location sharing happens at an application layer using GPS coordinates. But while such sharing is in effect, it wouldn't be unreasonable to also consider sharing IP locators to make it more efficient or more robust to e.g., route a video feed from one device to another.

6.3. Business Assets

In the area of Industrial IoT there are cases where an asset owner might want to ensure that their assets can communicate efficiently and robustly. In many cases those assets might be decoupled from any persons, but there can still be strong reasons to not share the ID/locator binding with third parties, such as enabling competitors to determine the number of deployed devices in a particular IP prefix.

7. Boundary between ID/locator part and rest of Internet

If the access to the ID/locator mapping is restricted as suggested above, then most of the potential peer devices would not have access to the ID/locator mappings. This means that there has to be a demarcation point between the part of the network which can access the ID/locator mappings for a particular identifier and the one which can not. There might be several choices how to handle this such as still using an ID/locator system but pointing to a locator for some fixed anchor point, or injecting routing prefixes for the ID prefixes into the normal routing system, or not providing any stable locators across this boundary; only allow ephemeral IP addresses per session or otherwise limited exposure.

8. Security Considerations

This document discusses privacy considerations, but does not explore any security considerations.

9. IANA Considerations

There are no IANA actions needed for this document.

10. References

[I-D.ietf-lisp-rfc6830bis]
Farinacci, D., Fuller, V., Meyer, D., Lewis, D., and A. Cabellos-Aparicio, "The Locator/ID Separation Protocol (LISP)", draft-ietf-lisp-rfc6830bis-30 (work in progress), January 2020.

[I-D.ietf-lisp-rfc6833bis]
Farinacci, D., Maino, F., Fuller, V., and A. Cabellos-Aparicio, "Locator/ID Separation Protocol (LISP) Control-Plane", draft-ietf-lisp-rfc6833bis-27 (work in progress), January 2020.

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

[RFC6740] Atkinson, RJ. and SN. Bhatti, "Identifier-Locator Network Protocol (ILNP) Architectural Description", RFC 6740, DOI 10.17487/RFC6740, November 2012, <https://www.rfc-editor.org/info/rfc6740>.

[RFC6836] Fuller, V., Farinacci, D., Meyer, D., and D. Lewis, "Locator/ID Separation Protocol Alternative Logical Topology (LISP+ALT)", RFC 6836, DOI 10.17487/RFC6836, January 2013, <https://www.rfc-editor.org/info/rfc6836>.

[RFC7348] Mahalingam, M., Dutt, D., Duda, K., Agarwal, P., Kreeger, L., Sridhar, T., Bursell, M., and C. Wright, "Virtual eXtensible Local Area Network (VXLAN): A Framework for Overlaying Virtualized Layer 2 Networks over Layer 3 Networks", RFC 7348, DOI 10.17487/RFC7348, August 2014, <https://www.rfc-editor.org/info/rfc7348>.

[RFC8111] Fuller, V., Lewis, D., Ermagan, V., Jain, A., and A. Smirnov, "Locator/ID Separation Protocol Delegated Database Tree (LISP-DDT)", RFC 8111, DOI 10.17487/RFC8111, May 2017, <https://www.rfc-editor.org/info/rfc8111>.

[RFC8174] Leiba, B., "Ambiguity of Uppercase vs Lowercase in RFC 2119 Key Words", BCP 14, RFC 8174, DOI 10.17487/RFC8174, May 2017, <https://www.rfc-editor.org/info/rfc8174>.

[3GPP] 3GPP TS 23.401: "General Packet Radio Service (GPRS) enhancements for Evolved Universal Terrestrial Radio Access Network (E-UTRAN) access", <https://www.3gpp.org/ftp/Specs/latest/>

Authors' Addresses

Luigi Iannone
Telecom Paris

Email: ggx@gigix.net

Dirk von Hugo
Deutsche Telekom

D-64295 Darmstadt
Germany

Email: Dirk.von-Hugo@telekom.de

Behcet Sarikaya
Denpel Informatique

Email: sarikaya@ieee.org

Erik Nordmark
Zededa
Santa Clara, CA
USA

Email: nordmark@sonic.net