Internet Engineering Task Force
Internet-Draft
Intended status: Informational
Expires: September 14, 2018

R. Winter
University of Applied Sciences Augsburg
M. Faath
Conntac GmbH
F. Weisshaar
University of Applied Sciences Augsburg
March 13, 2018

Privacy considerations for protocols relying on IP broadcast and multicast
draft-ietf-intarea-broadcast-consider-09

Abstract

A number of application-layer protocols make use of IP broadcasts or multicast messages for functions such as local service discovery or name resolution. Some of these functions can only be implemented efficiently using such mechanisms. When using broadcasts or multicast messages, a passive observer in the same broadcast/multicast domain can trivially record these messages and analyze their content. Therefore, designers of protocols that make use of broadcast/multicast messages need to take special care when designing their protocols.

Table of Contents

1. Introduction

Broadcast and multicast messages have a large (and to the sender unknown) receiver group by design. Because of that, these two mechanisms are vital for a number of basic network functions such as auto-configuration or link-layer address lookup. Also application developers use broadcast/multicast messages to implement things such as local service or peer discovery. It appears that an increasing number of applications make use of it as suggested by experimental results obtained on campus networks including the IETF meeting network [TRAC2016]. This trend is not entirely surprising. As [RFC0919] puts it, "The use of broadcasts [...] is a good base for many applications". Broadcast and multicast functionality in a subnetwork are therefore important as a lack thereof renders the protocols relying on these mechanisms inoperable [RFC3819].

Using broadcast/multicast can become problematic if the information that is being distributed can be regarded as sensitive or when the information that is distributed by multiple of these protocols can be correlated in a way that sensitive data can be derived. This is clearly true for any protocol, but broadcast/multicast is special in at least two respects:

(a) The aforementioned large receiver group, consisting of receivers unknown to the sender. This makes eavesdropping without special privileges or a special location in the network trivial for anybody in the same broadcast/multicast domain.

(b) Encryption is difficult when broadcast/multicast messages are used, for instance because a non-trivial key management protocol might be required. When encryption is not used, the content of these messages is easily accessible, making it easy to spoof and replay them.

Given the above, privacy protection for protocols based on broadcast or multicast communication is significantly more difficult compared to unicast communication and at the same time invading the privacy is much easier.

Privacy considerations of IETF-specified protocols have received some attention in the recent past (e.g. [RFC7721] or [RFC7819]). There is also general guidance available for document authors on when and how to include a privacy considerations section in their documents and on how to evaluate the privacy implications of Internet protocols [RFC6973]. RFC6973 also describes potential threats to privacy in great detail and lists terminology that is also used in this document. In contrast to RFC6973, this document contains a number of privacy considerations especially for protocols that rely on broadcast/multicast, intended to reduce the likelihood that a broadcast/multicast protocol can be misused to collect sensitive data about devices, users and groups of users in a broadcast/multicast domain.

The above mentioned considerations particularly apply to protocols designed outside the IETF - for two reasons. For one, non-standard protocols will likely not receive operational attention and support in making them more secure, e.g. what DHCP snooping does for DHCP. But because these protocols are typically not documented, network equipment does not provide similar features for them. The other reason is that these protocols have been designed in isolation, where a set of considerations to follow is useful in the absence of a larger community providing feedback and expertise to improve the protocol. In particular, carelessly designed protocols that use broadcast/multicast can break privacy efforts at different layers of

the protocol stack such as MAC address or IP address randomization [RFC4941].

1.1. Types and usage of broadcast and multicast

In IPv4, two major types of broadcast addresses exist, the limited broadcast which is defined as all-ones (255.255.255.255, defined in section 5.3.5.1 of [RFC1812]) and the directed broadcast with the given network prefix of an IP address and the host part of all-ones (defined in section 5.3.5.2. of [RFC1812]). Broadcast packets are received by all nodes in a subnetwork. Limited broadcasts never transit a router. The same is true for directed broadcasts by default, but routers may provide an option to do this [RFC2644]. IPv6 on the other hand does not provide broadcast addresses but solely relies on multicast [RFC4291].

In contrast to broadcast addresses, multicast addresses represent an identifier for a set of interfaces that can be a set different from all nodes in the subnetwork. All interfaces that are identified by a given multicast address receive packets destined towards that address and are called a multicast group. In both IPv4 and IPv6, multiple pre-defined multicast addresses exist. The ones most relevant for this document are the ones with subnet scope. For IPv4, an IP prefix is reserved for this purpose called the Local Network Control Block (224.0.0.0/24, defined in section 4 of [RFC5771]). For IPv6, the relevant multicast addresses are the two All Nodes Addresses, which every IPv6-capable host is required to recognize as identifying itself (see section 2.7.1 of [RFC4291]).

Typical usage of these addresses include local service discovery (e.g. Multicast DNS (mDNS) [RFC6762] and Link-Local Multicast Name Resolution (LLMNR) [RFC4795] make use of multicast), autoconfiguration (e.g. DHCPv4 [RFC2131] uses broadcasts and DHCPv6 [RFC3315] uses multicast addresses) and other vital network services such as address resolution or duplicate address detection. But besides these core network functions, also applications make use of broadcast and multicast functionality, often implementing proprietary protocols. In sum, these protocols distribute a diverse set of potentially privacy sensitive information to a large receiver group and to be part of this receiver group, the only requirement is to be on same subnetwork.

1.2. Requirements Language

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 [RFC2119].

2. Privacy considerations

There are a few obvious and a few not necessarily obvious things designers of protocols utilizing broadcast/multicast should consider in respect to the privacy implications of their protocol. Most of these items are based on protocol behavior observed as part of experiments on operational networks [TRAC2016].

2.1. Message frequency

Frequent broadcast/multicast traffic caused by an application can give away user behavior and online connection times. This allows a passive observer to potentially deduce a user's current activity (e.g. a game) and it allows to create an online profile (i.e. times the user is on the network). The higher the frequency of these messages and the duration of time these messages are sent, the more accurate this profile will be. Given that broadcasts/multicasts are only visible in the same broadcast/multicast domain, these messages also give the rough location of the user away (e.g. a campus or building).

This behavior has e.g. been observed by a synchronization mechanism of a popular application, where multiple messages have been sent per minute via broadcast. Given this behavior, it is possible to record a device's time on the network with a sub-minute accuracy given only the traffic of this single application installed on the device. But also services used for local name resolution in modern operating systems utilize broadcast/multicast protocols (e.g. mDNS, LLMNR or NetBIOS) to announce for example resources regularly which also allow tracking the online time of a device.

If a protocol relies on frequent or periodic broadcast/multicast messages, the frequency SHOULD be chosen conservatively, in particular if the messages contain persistent identifiers (see next subsection). Also, intelligent message suppression mechanisms such as the ones employed in mDNS [RFC6762] SHOULD be implemented. The lower the frequency of broadcast messages, the harder passive traffic analysis and surveillance becomes.

2.2. Persistent identifiers

A few protocols that make use of broadcast/multicast messages observed in the wild make use of persistent identifiers. This includes the use of host names or more abstract persistent identifiers such as a universally unique identifiers (UUID) or similar. These IDs, which e.g. identify the installation of a certain application might not change across updates of the software and can therefore be extremely long lived. This allows a passive

observer to track a user precisely if broadcast/multicast messages are frequent. This is even true in case the IP and/or MAC address changes. Such identifiers also allow two different interfaces (e.g. WiFi and Ethernet) to be correlated to the same device. If the application makes use of persistent identifiers for multiple installations of the same application for the same user, this even allows to infer that different devices belong to the same user.

The aforementioned broadcast messages from a synchronization mechanism of a popular application also included a persistent identifier in every broadcast. This identifier never changed after the application was installed and it allowed to track a device even when it changed its network interface or when it connected to a different network.

Persistent IDs are considered bad practice in general for broadcast and multicast communication, as persistent application layer IDs will make efforts on lower layers to randomize identifiers (e.g. [I-D.huitema-6man-random-addresses]) useless. When protocols that make use of broadcast/multicast need to make use of IDs, these IDs SHOULD be rotated frequently to make user tracking more difficult.

2.3. Anticipate user behavior

A large number of users name their device after themselves, either using their first name, last name or both. Often a host name includes the type, model or maker of a device, its function or it includes language specific information. Based on data gathered during experiments performed at IETF meetings and at a large campus network, this appears currently to be prevalent user behavior [TRAC2016]. For protocols using the host name as part of the messages, this clearly will reveal personally identifiable information to everyone on the local network. This information can also be used to mount more sophisticated attacks, when e.g. the owner of a device is identified (as an interesting target) or properties of the device are known (e.g. known vulnerabilities). Host names are also a type of persistent identifier and therefore the considerations in Section 2.2 apply.

Some of the most commonly used operating systems include the name the user chooses for the user account during the installation process as part of the host name of the device. The name of the operating system can also be included, revealing therefore two pieces of information, which can be regarded as private information if the host name is used in broadcast/multicast messages.

Where possible, the use of host names and other user-provided information in protocols making use of broadcast/multicast SHOULD be

avoided. An application might want to display the information it will broadcast on the LAN at install/config time, so the user is at least aware of the application's behavior. More host name considerations can be found in [RFC8117]. More information on user participation can be found in [RFC6973].

2.4. Consider potential correlation

A large number of services and applications make use of the broadcast/multicast mechanism. That means there are various sources of information that are easily accessible by a passive observer. In isolation, the information these protocols reveal might seem harmless, but given multiple such protocols, it might be possible to correlate this information. E.g. a protocol that uses frequent messages including a UUID to identify the particular installation does not give the identity of the user away. But a single message including the user's host name might just do that and it can be correlated using e.g. the MAC address of the device's interface.

In the experiments described in [TRAC2016], it was possible to correlate frequently sent broadcast messages that included a unique identifier with other broadcast/multicast messages containing usernames (e.g. mDNS, LLMNR or NetBIOS), but also relationships to other users. This allowed to reveal the real identity of the users of many devices but it also gave some information about their social environment away.

A designer of a protocol that makes use of broadcast/multicast needs to be aware of the fact that even if - in isolation - the information a protocol leaks seems harmless, there might be ways to correlate that information with information from other protocols to reveal sensitive information about a user.

2.5. Configurability

A lot of applications and services relying on broadcast/multicast protocols do not include the means to declare "safe" environments (e.g. based on the SSID of a WiFi network and the MAC addresses of the access points). E.g. a device connected to a public WiFi will likely broadcast the same information as when connected to the home network. It would be beneficial if certain behavior could be restricted to "safe" environments.

A popular operating system e.g. allows the user to specify the trust level of the network the device connects to, which for example restricts specific system services (using broadcast/multicast messages for their normal operation) to be used in trusted networks

only. Such functionality could implemented as part of an application.

An application developer making use of broadcasts/multicasts as part of the application SHOULD make the broadcast feature, if possible, configurable, so that potentially sensitive information does not leak on public networks, where the threat to privacy is much larger.

3. Operational considerations

Besides changing end-user behavior, choosing sensible defaults as an operating system vendor (e.g. for suggesting host names) and the considerations for protocol designers mentioned in this document, there is something that the network administrators/operators can do to limit the above mentioned problems.

A feature commonly found on access points e.g. is to manage/filter broadcast and multicast traffic. This will potentially break certain applications or some of their functionality but will also protect the users from potentially leaking sensitive information. Wireless access points often provide finer-grained control beyond a simple on/off switch for well-known protocols or provide mechanisms to manage broadcast/multicast traffic intelligently using e.g. proxies (see [I-D.ietf-mboned-ieee802-mcast-problems]). These mechanisms however only work on standardized protocols.

4. Summary

Increasingly, applications rely on protocols that send and receive broadcast and multicast messages. For some, broadcasts/multicasts are the basis of their application logic, others use broadcasts/multicasts to improve certain aspects of the application but are fully functional in case broadcasts/multicasts fail. Irrespective of the role of broadcast and multicast messages for the application, the designers of protocols that make use of them should be very careful in their protocol design because of the special nature of broadcast and multicast.

It is not always possible to implement certain functionality via unicast, but in case a protocol designer chooses to rely on broadcast/multicast, the following should be carefully considered:

- IETF-specified protocols, such as mDNS [RFC6762], SHOULD be used if possible as operational support might exist to protect against the leakage of private information. Also, for some protocols privacy extensions are being specified, which can be used if implemented. E.g. for DNS-SD privacy extensions are documented in [I-D.ietf-dnssd-privacy]

- Using user-specified information inside broadcast/multicast messages SHOULD be avoided, as users will often use personal information or other information aiding attackers, in particular if the user is unaware about how that information is being used

- The use of persistent IDs in messages SHOULD be avoided, as this allows user tracking, correlation and potentially has a devastating effect on other privacy protection mechanisms

- If one really must design a new protocol relying on broadcast/ multicast and cannot use an IETF-specified protocol, then:

  * the protocol SHOULD be very conservative in how frequently it sends messages as an effort in data minimization

  * it SHOULD make use of mechanisms implemented in IETF-specified protocols that can be helpful in privacy protection such as message suppression in mDNS

  * it SHOULD be designed in a way that information sent in broadcast/multicast messages cannot be correlated with information from other protocols using broadcast/multicast

  * it SHOULD be possible to let the user configure "safe" environments if possible (e.g. based on the SSID) to minimize the risk of information leakage (e.g. a home network as opposed to a public Wifi)

5. Other considerations

Besides privacy implications, frequent broadcasting also represents a performance problem. In particular in certain wireless technologies such as 802.11, broadcast and multicast are transmitted at a much lower rate (the lowest common denominator rate) compared to unicast and therefore have a much bigger impact on the overall available airtime [I-D.ietf-mboned-ieee802-mcast-problems]. Further, it will limit the ability for devices to go to sleep if frequent broadcasts are being sent. A similar problem in respect to Router Advertisements is addressed in [I-D.ietf-v6ops-reducing-ra-energy-consumption]. In that respect broadcasts/multicast can be used for another class of attacks that is not related to privacy. The potential impact on network performance should nevertheless be considered when designing a protocol that makes use of broadcast/multicast.

6. Acknowledgments

We would like to thank Eliot Lear, Joe Touch and Stephane Bortzmeyer for their valuable input to this document.

This work was partly supported by the European Commission under grant agreement FP7-318627 mPlane. Support does not imply endorsement.

7. IANA Considerations

This memo includes no request to IANA.

8. Security Considerations

This document deals with privacy-related considerations of broadcast- and multicast-based protocols. It contains advice for designers of such protocols to minimize the leakage of privacy-sensitive information. The intent of the advice is to make sure that identities will remain anonymous and user tracking will be made difficult.

It should be noted that certain applications could make use of existing mechanisms to protect multicast traffic such as the ones defined in [RFC5374]. Examples of such applications can be found in Appendix A. of [RFC5374]. Given the required infrastructure and assumptions about these applications and the security infrastructure, many applications will not be able to make use of such mechanisms.

9. References

9.1. Normative References

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.

9.2. Informative References

[I-D.huitema-6man-random-addresses]
Huitema, C., "Implications of Randomized Link Layers Addresses for IPv6 Address Assignment", draft-huitema-6man-random-addresses-03 (work in progress), March 2016.

[I-D.ietf-dnssd-privacy]
Huitema, C. and D. Kaiser, "Privacy Extensions for DNS-SD", draft-ietf-dnssd-privacy-00 (work in progress), October 2016.

[I-D.ietf-mboned-ieee802-mcast-problems]
Perkins, C., McBride, M., Stanley, D., Kumari, W., and J. Zuniga, "Multicast Considerations over IEEE 802 Wireless Media", draft-ietf-mboned-ieee802-mcast-problems-01 (work in progress), February 2018.

[I-D.ietf-v6ops-reducing-ra-energy-consumption]
Yourtchenko, A. and L. Colitti, "Reducing energy consumption of Router Advertisements", draft-ietf-v6ops-reducing-ra-energy-consumption-03 (work in progress), November 2015.

[RFC0919] Mogul, J., "Broadcasting Internet Datagrams", STD 5, RFC 919, DOI 10.17487/RFC0919, October 1984, <http://www.rfc-editor.org/info/rfc919>.

[RFC1812] Baker, F., Ed., "Requirements for IP Version 4 Routers", RFC 1812, DOI 10.17487/RFC1812, June 1995, <http://www.rfc-editor.org/info/rfc1812>.

[RFC2131] Droms, R., "Dynamic Host Configuration Protocol", RFC 2131, DOI 10.17487/RFC2131, March 1997, <http://www.rfc-editor.org/info/rfc2131>.

[RFC2644] Senie, D., "Changing the Default for Directed Broadcasts in Routers", BCP 34, RFC 2644, DOI 10.17487/RFC2644, August 1999, <http://www.rfc-editor.org/info/rfc2644>.

[RFC3315] Droms, R., Ed., Bound, J., Volz, B., Lemon, T., Perkins, C., and M. Carney, "Dynamic Host Configuration Protocol for IPv6 (DHCPv6)", RFC 3315, DOI 10.17487/RFC3315, July 2003, <http://www.rfc-editor.org/info/rfc3315>.

[RFC3819] Karn, P., Ed., Bormann, C., Fairhurst, G., Grossman, D., Ludwig, R., Mahdavi, J., Montenegro, G., Touch, J., and L. Wood, "Advice for Internet Subnetwork Designers", BCP 89, RFC 3819, DOI 10.17487/RFC3819, July 2004, <http://www.rfc-editor.org/info/rfc3819>.

[RFC4291] Hinden, R. and S. Deering, "IP Version 6 Addressing Architecture", RFC 4291, DOI 10.17487/RFC4291, February 2006, <http://www.rfc-editor.org/info/rfc4291>.

[RFC4795] Aboba, B., Thaler, D., and L. Esibov, "Link-local Multicast Name Resolution (LLMNR)", RFC 4795, DOI 10.17487/RFC4795, January 2007, <http://www.rfc-editor.org/info/rfc4795>.

[RFC4941] Narten, T., Draves, R., and S. Krishnan, "Privacy Extensions for Stateless Address Autoconfiguration in IPv6", RFC 4941, DOI 10.17487/RFC4941, September 2007, <http://www.rfc-editor.org/info/rfc4941>.

[RFC5374] Weis, B., Gross, G., and D. Ignjatic, "Multicast Extensions to the Security Architecture for the Internet Protocol", RFC 5374, DOI 10.17487/RFC5374, November 2008, <http://www.rfc-editor.org/info/rfc5374>.

[RFC5771] Cotton, M., Vegoda, L., and D. Meyer, "IANA Guidelines for IPv4 Multicast Address Assignments", BCP 51, RFC 5771, DOI 10.17487/RFC5771, March 2010, <http://www.rfc-editor.org/info/rfc5771>.

[RFC6762] Cheshire, S. and M. Krochmal, "Multicast DNS", RFC 6762, DOI 10.17487/RFC6762, February 2013, <http://www.rfc-editor.org/info/rfc6762>.

[RFC6973] Cooper, A., Tschofenig, H., Aboba, B., Peterson, J., Morris, J., Hansen, M., and R. Smith, "Privacy Considerations for Internet Protocols", RFC 6973, DOI 10.17487/RFC6973, July 2013, <http://www.rfc-editor.org/info/rfc6973>.

[RFC7721] Cooper, A., Gont, F., and D. Thaler, "Security and Privacy Considerations for IPv6 Address Generation Mechanisms", RFC 7721, DOI 10.17487/RFC7721, March 2016, <http://www.rfc-editor.org/info/rfc7721>.

[RFC7819] Jiang, S., Krishnan, S., and T. Mrugalski, "Privacy Considerations for DHCP", RFC 7819, DOI 10.17487/RFC7819, April 2016, <http://www.rfc-editor.org/info/rfc7819>.

[RFC8117] Huitema, C., Thaler, D., and R. Winter, "Current Hostname Practice Considered Harmful", RFC 8117, DOI 10.17487/RFC8117, March 2017, <https://www.rfc-editor.org/info/rfc8117>.

[TRAC2016] Faath, M., Weisshaar, F., and R. Winter, "How Broadcast Data Reveals Your Identity and Social Graph", 7th International Workshop on TRaffic Analysis and Characterization IEEE TRAC 2016, September 2016.

Authors' Addresses

Rolf Winter
University of Applied Sciences Augsburg
Augsburg
DE

Email: rolf.winter@hs-augsburg.de

Michael Faath
Conntac GmbH
Augsburg
DE

Email: faath@conntac.net

Fabian Weisshaar
University of Applied Sciences Augsburg
Augsburg
DE

Email: fabian.weisshaar@hs-augsburg.de