Network Working Group
Internet-Draft
Intended status: Informational
Expires: January 15, 2014

M. Bagnulo
UC3M
C. Paasch
UCLouvain
F. Gont
SI6 Networks / UTN-FRH
O. Bonaventure
UCLouvain
C. Raiciu
UPB
July 14, 2013

Analysis of MPTCP residual threats and possible fixes
draft-bagnulo-mptcp-attacks-00

Abstract

This documents performs an analysis of the residual threats for MPTCP and explores possible solutions to them.

Status of This Memo

Copyright Notice

Table of Contents

1. Introduction

This document provides a complement to the threat analysis for Multipath TCP (MPTCP) [RFC6824] documented in RFC 6181 [RFC6181]. RFC 6181 provided a threat analysis for the general solution space of extending TCP to operate with multiple IP addresses per connection. Its main goal was to leverage previous experience acquired during the design of other multi-address protocols, notably SHIM6 [RFC5533], SCTP [RFC4960] and MIPv6 [RFC3775] during the design of MPTCP. Thus, RFC 6181 was produced before the actual MPTCP specification (RFC6824) was completed, and documented a set of recommendations that were considered during the production of such specification.

This document complements RFC 6181 with a vulnerability analysis of the specific mechanisms specified in RFC 6824. The motivation for this analysis is to identify possible security issues with MPTCP as currently specified and propose security enhancements to address the identified security issues.

The goal of the security mechanisms defined in RFC 6824 were to make MPTCP no worse than currently available single-path TCP. We believe that this goal is still valid, so we will perform our analysis on the same grounds.

Types of attackers: for all attacks considered in this documents, we identify the type of attacker. We can classify the attackers based on their location as follows:

- Off-path attacker. This is an attacker that does not need to be located in any of the paths of the MPTCP session at any point in time during the lifetime of the MPTCP session. This means that the Off-path attacker cannot eavesdrop any of the packets of the MPTCP session.

- Partial time On-path attacker. This is an attacker that needs to be in at least one of the paths during part but not during the entire lifetime of the MPTCP session. The attacker can be in the forward and/or backward directions, for the initial subflow and/or other subflows. The specific needs of the attacker will be made explicit in the attack description.

- On-path attacker. This attacker needs to be on at least one of the paths during the whole duration of the MPTCP session. The attacker can be in the forward and/or backward directions, for the initial subflow and/or other subflows. The specific needs of the attacker will be made explicit in the attack description.

We can also classify the attackers based on their actions as follows:

- Eavesdropper. The attacker is able to capture some of the packets of the MPTCP session to perform the attack, but it is not capable of changing, discarding or delaying any packet of the MPTCP session. The attacker can be in the forward and/or backward directions, for the initial subflow and/or other subflows. The specific needs of the attacker will be made explicit in the attack description.

- Active attacker. The attacker is able to change, discard or delay some of the packets of the MPTCP session. The attacker can be in the forward and/or backward directions, for the initial subflow and/or other subflows. The specific needs of the attacker will be made explicit in the attack description.

In this document, we consider the following possible combinations of attackers:

- an On-path eavesdropper

- an On-path active attacker

- an Off-path active attacker

- a Partial-time On-path eavesdropper

- a Partial-time On-path active attacker

In the rest of the document we describe different attacks that are possible against the MPTCP protocol specified in RFC6824 and we propose possible security enhancements to address them.

2. ADD_ADDR attack

Summary of the attack:

Type of attack: MPTCP session hijack enabling Man-in-the-Middle.

Type of attacker: Off-path, active attacker.

Threat: Medium

Description:

In this attack, the attacker uses the ADD_ADDR option defined in RFC6824 to hijack an ongoing MPTCP session and enables himself to perform a Man-in-the-Middle attack on the MPTCP session.

Consider the following scenario. Host A with address IPA has one MPTCP session with Host B with address IPB. The MPTCP subflow between IPA and IPB is using port PA on host A and port PB on host B. The tokens for the MPTCP session are TA and TB for Host A and Host B respectively. Host C is the attacker. It owns address IPC. The attack is executed as follows:

1. Host C sends a forged packet with source address IPA, destination address IPB, source port PA and destination port PB. The packet has the ACK flag set. The TCP sequence number for the segment is i and the ACK sequence number is j. We will assume all these are valid, we discuss what the attacker needs to figure these ones later on. The packet contains the ADD_ADDR option. The ADD_ADDR option announces IPC as an alternative address for the connection. It also contains an eight bit address identifier which does not bring any strong security benefit.

2. Host B receives the ADD_ADDR message and it replies by sending a TCP SYN packet. (Note: the MPTCP specification states that the host receiving the ADD_ADDR option may initiate a new subflow. If the host is configured so that it does not initiate a new subflow the attack will not succeed. For example, on the Linux implementation, the server does not create subflows. Only the client does so.) The source address for the packet is IPB, the

destination address for the packet is IPC, the source port is PB’ and the destination port is PA’ (It is not required that PA=PA’ nor that PB=PB’). The sequence number for this packet is the new initial sequence number for this subflow. The ACK sequence number is not relevant as the ACK flag is not set. The packet carries an MP_JOIN option and it carries the token TA. It also carries a random nonce generated by Host B called RB.

3. Host C receives the SYN+MP_JOIN packet from Host B, and it alters it in the following way. It changes the source address to IPC and and the destination address to IPA. It sends the modified packet to Host A, impersonating Host B.

4. Host A receives the SYN+MP_JOIN message and it replies with a SYN /ACK+MP_JOIN message. The packet has source address IPA and destination address IPC, as well as all the other needed parameters. In particular, Host A computes a valid HMAC and places it in the MP_JOIN option.

5. Host C receives the SYN/ACK+MP_JOIN message and it changes the source address to IPC and the destination address to IPB. It sends the modified packet to IPB impersonating Host A.

6. Host B receives the SYN/ACK+MP_JOIN message. Host B verifies the HMAC of the MP_JOIN option and confirms its validity. It replies with an ACK+MP_JOIN packet. The packet has source address IPB and destination address IPC, as well as all the other needed parameters. The returned MP_JOIN option contains a valid HMAC computed by Host B.

7. Host C receives the ACK+MP_JOIN message from B and it alters it in the following way. It changes the source address to IPC and the destination address to IPA. It sends the modified packet to Host A impersonating Host B.

8. Host A receives the ACK+MP_JOIN message and creates the new subflow.

> At this point the attacker has managed to place itself as a MitM for one subflow for the existing MPTCP session. It should be noted that there still exists the subflow between address IPA and IPB that does not flow through the attacker, so the attacker has not completely intercepted all the packets in the communication (yet). If the attacker wishes to completely intercept the MPTCP session it can do the following additional step.

9. Host C sends two TCP RST messages. One TCP RST packet is sent to Host B, with source address IPA and destination address IPB and source and destination ports PA and PB, respectively. The other TCP RST message is sent to Host A, with source address IPB and destination address IPA and source and destination ports PB and PA, respectively. Both RST messages must contain a valid sequence number. Note that figuring the sequence numbers to be used here for subflow A is the same difficulty as being able to send the initial ADD_ADDR option with valid Sequence number and ACK value. If there are more subflows, then the attacker needs to find the Sequence Number and ACK for each subflow.

   At this point the attacker has managed to fully hijack the MPTCP session.

Information required by the attacker to perform the described attack:

In order to perform this attack the attacker needs to guess or know the following pieces of information: (The attacker need this information for one of the subflows belonging to the MPTCP session.)

- the four-tuple {Client-side IP Address, Client-side Port, Server-side Address, Servcer-side Port} that identifies the target TCP connection

- a valid sequence number for the subflow

- a valid ACK sequence number for the subflow

- a valid address identifier for IPC

TCP connections are uniquely identified by the four-tuple {Source Address, Source Port, Destination Address, Destination Port}. Thus, in order to attack a TCP connection, an attacker needs to know or be able to guess each of the values in that four-tuple. Assuming the two peers of the target TCP connection are known, the Source Address and the Destination Address can be assumed to be known.

We note that in order to be able to successfully perform this attack, the attacker needs to be able to send packets with a forged source address. This means that the attacker cannot be located in a network where techniques like ingress filtering [RFC2827] or source address validation [I-D.ietf-savi-framework] are deployed. However, ingress filtering is not as widely implemented as one would expect, and hence cannot be relied upon as a mitigation for this kind of attack.

Assuming the attacker knows the application protocol for which the TCP connection is being employed, the server-side port can also be assumed to be known. Finally, the client-side port will generally not be known, and will need to be guessed by the attacker. The chances of an attacker guessing the client-side port will depend on the ephemeral port range employed by the client, and whether the client implements port randomization [RFC6056].

Assuming TCP sequence number randomization is in place (see e.g. [RFC6528]), an attacker would have to blindly guess a valid TCP sequence number. That is,

RCV.NXT =< SEG.SEQ < RCV.NXT+RCV.WND or RCV.NXT =<
SEG.SEQ+SEG.LEN-1 < RCV.NXT+RCV.WND

As a result, the chances of an attacker to succeed will depend on the TCP receive window size at the target TCP peer.

We note that automatic TCP buffer tuning mechanisms have been become common for popular TCP implementations, and hence very large TCP window sizes of values up to 2 MB could end up being employed by such TCP implementations.

According to [RFC0793], the Acknowledgement Number is considered valid as long as it does not acknowledge the receipt of data that has not yet been sent. That is, the following expression must be true:

SEG.ACK <= SND.NXT

However, for implementations that support [RFC5961], the following (stricter) validation check is enforced:

SND.UNA - SND.MAX.WND <= SEG.ACK <= SND.NXT

Finally, in order for the address identifier to be valid, the only requirement is that it needs to be different than the ones already being used by Host A in that MPTCP session, so a random identifier is likely to work.

Given that a large number of factors affect the chances of an attacker of successfully performing the aforementioned off-path attacks, we provide two general expressions for the expected number of packets the attacker needs to send to succeed in the attack: one for MTCP implementations that support [RFC5961], and another for MPTCP implementations that do not.

Implementations that do not support RFC 5961

Packets = (2^32/(RCV_WND)) * 2 * EPH_PORT_SIZE/2 * 1/MSS

Where the new :

Packets:
Maximum number of packets required to successfully perform an off-path (blind) attack.

RCV_WND:
TCP receive window size (RCV.WND) at the target node.

EPH_PORT_SIZE:
Number of ports comprising the ephemeral port range at the "client" system.

MSS:
Maximum Segment Size, assuming the attacker will send full segments to maximize the chances to get a hit.

Notes:
The value "2^32" represents the size of the TCP sequence number space.
The value "2" accounts for 2 different ACK numbers (separated by 2^31) that should be employed to make sure the ACK number is valid.

The following table contains some sample results for the number of required packets, based on different values of RCV_WND and EPH_PORT_SIZE for a MSS of 1500 bytes.

Ports \ Win	16 KB	128 KB	256 KB	2048 KB
4000	699050	87381	43690	5461
10000	1747626	218453	109226	13653
50000	8738133	1092266	546133	68266

Table 1: Max. Number of Packets for Successful Attack

Implementations that do not support RFC 5961

Packets = (2^32/(RCV_WND)) * (2^32/(SND_MAX_WND)) * EPH_PORT_SIZE/2 * 1/MSS

Where:

Packets:
Maximum number of packets required to successfully perform an off-path (blind) attack.

RCV_WND:
TCP receive window size (RCV.WND) at the target MPTCP endpoint.

SND_MAX_WND:
Maximum TCP send window size ever employed by the target MPTCP end-point (SND.MAX.WND).

EPH_PORT_SIZE:
Number of ports comprising the ephemeral port range at the "client" system.

Notes:
The value "2^32" represents the size of the TCP sequence number space.
The parameter "SND_MAX_WND" is specified in [RFC5961].

The following table contains some sample results for the number of required packets, based on different values of RCV_WND, SND_MAX_WND, and EPH_PORT_SIZE. For these implementations, only a limited number of sample results are provided, just as an indication of how [RFC5961] increases the difficulty of performing these attacks.

Ports \ Win	16 KB	128 KB	256 KB	2048 KB
4000	91625968967	1431655765	357913941	559240

Table 2: Max. NUmber of Packets for Successful Attack

Note:
In the aforementioned table, all values are computed with RCV_WND equal to SND_MAX_WND.

2.1. Possible security enhancements to prevent this attack

1. To include the token of the connection in the ADD_ADDR option. This would make it harder for the attacker to launch the attack, since he needs to either eavesdrop the token (so this can no longer be a blind attack) or to guess it, but a random 32 bit number is not so easy to guess. However, this would imply that any eavesdropper that is able to see the token, would be able to launch this attack. This solution then increases the vulnerability window against eavesdroppers from the initial 3-way

handshake for the MPTCP session to any exchange of the ADD_ADDR messages.

2. To include the HMAC of the address contained in the ADD_ADDR option concatenated with the key of the receiver of the ADD-ADDR message. This makes it much more secure, since it requires the attacker to have both keys (either by eavesdropping it in the first exchange or by guessing it). Because this solution relies on the key used in the MPTCP session, the protection of this solution would increase if new key generation methods are defined for MPTCP (e.g. using SSL keys as has been proposed).

3. To include the destination address of the ADD_ADDR msg in the HMAC. This would certainly make the attack harder (the attacker would need to know the key). It wouldn't allow hosts behind NATs to be reached by an address in the ADD_ADDR option, even with static NAT bindings (like a web server at home). Probably it would make sense to combine it option 2) (i.e. to have the HMAC of the address in the ADD_ADDR option and the destination address of the packet.

4. To include the destination address of the SYN packet in the HMAC of the MP_JOIN message. This has the same problems than option 3) in the presence of NATs.

3. DoS attack on MP_JOIN

Summary of the attack:

Type of attack: MPTCP Denial-of-Service attack, preventing the hosts from creating new subflows.

Type of attacker: Off-path, active attacker

Threat: Low (? - as it is hard to guess the 32-bit token and still then the attacker only prevents the creation of new subflows)

Description:

As currently specified, the initial SYN+MP_JOIN message of the 3-way handshake for additional subflows creates state in the host receiving the message. This, because the SYN+MP_JOIN contains the 32-bit token that allows the receiver to identify the MPTCP-session and the 32-bit random nonce, used in the HMAC calculation. As this information is not resent in the third ACK of the 3-way handshake, a host must create state upon reception of a SYN+MP_JOIN.

Assume that there exists an MPTCP-session between host A and host B, with token Ta and Tb. An attacker, sending a SYN+MP_JOIN to host B, with the valid token Tb, will trigger the creation of state on host B. The number of these half-open connections a host can store per MPTCP-session is limited by a certain number, and it is implementation-dependent. The attacker can simply exhaust this limit by sending multiple SYN+MP_JOINs with different 5-tuples. The (possibly forged) source address of the attack packets will typically correspond to an address that is not in use, or else the SYN/ACK sent by Host B would elicit a RST from the impersonated node, thus removing the corresponding state at Host B. Further discussion of traditional SYN-flod attacks and common mitigations can be found in [RFC4987]

This effectively prevents the host A from sending any more SYN+MP_JOINs to host B, as the number of acceptable half-open connections per MPTCP-session on host B has been exhausted.

The attacker needs to know the token Tb in order to perform the described attack. This can be achieved if it is a partial on-time eavesdropper, observing the 3-way handshake of the establishment of an additional subflow between host A and host B. If the attacker is never on-path, it has to guess the 32-bit token.

Christoph: can you provide text about the birthday paradox and busy servers?

3.1. Possible security enhancements to prevent this attack

The third packet of the 3-way handshake could be extended to contain also the 32-bit token and the random nonce that has been sent in the SYN+MP_JOIN. Further, host B will have to generate its own random nonce in a reproducible fashion (e.g., a Hash of the 5-tuple + initial sequence-number + local secret). This will allow host B to reply to a SYN+MP_JOIN without having to create state. Upon the reception of the third ACK, host B can then verify the correctness of the HMAC and create the state.

4. SYN flooding amplification

Summary of the attack:

Type of attack: The attacker can use the SYN+MP_JOIN messages to amplify the SYN flooding attack.

Type of attacker: Off-path, active attacker

Threat: Medium

Description:

SYN flooding attacks [RFC4987] use SYN messages to exhaust the server’s resources and prevent new TCP connections. A common mitigation is the use of SYN cookies [RFC4987] that allow the stateless processing of the initial SYN message.

With MPTCP, the initial SYN can be processed in a stateless fashion using the aforementioned SYN cookies. However, as we described in the previous section, as currently specified, the SYN+MP_JOIN messages are not processed in a stateless manner. This opens a new attack vector. The attacker can now open a MPTCP session by sending a regular SYN and creating the associated state but then send as many SYN+MP_JOIN messages as supported by the server with different source address source port combinations, consuming server’s resources without having to create state in the attacker. This is an amplification attack, where the cost on the attacker side is only the cost of the state associated with the initial SYN while the cost on the server side is the state for the initial SYN plus all the state associated to all the following SYN+MP_JOIN.

4.1. Possible security enhancements to prevent this attack

1. The solution described for the previous DoS attack on MP_JOIN would also prevent this attack.

2. Limiting the number of half open subflows to a low number (like 3) would also limit the impact of this attack.

5. Eavesdropper in the initial handshake

Summary of the attack

Type of attack: An eavesdropper present in the initial handshake where the keys are exchanged can hijack the MPTCP session at any time in the future.

Type of attacker: a Partial-time On-path eavesdropper

Threat: Low

Description:

In this case, the attacker is present along the path when the initial 3-way handshake takes place, and therefore is able to learn the keys used in the MPTCP session. This allows the attacker to move away from the MPTCP session path and still be able to hijack the MPTCP session in the future. This vulnerability was readily identified at the moment of the design of the MPTCP security solution and the threat was considered acceptable.

5.1. Possible security enhancements to prevent this attack

There are many techniques that can be used to prevent this attack and each of them represents different tradeoffs. At this point, we limit ourselves to enumerate them and provide useful pointers.

1. Use of hash-chains. The use of hash chains for MPTCP has been explored in [hash-chains]

2. Use of SSL keys for MPTCP security as described in [I-D.paasch-mptcp-ssl]

3. Use of Cryptographically-Generated Addresses (CGAs) for MPTCP security. CGAs [RFC3972] have been used in the past to secure multi addressed protocols like SHIM6 [RFC5533].

4. Use of TCPCrypt [I-D.bittau-tcp-crypt]

5. Use DNSSEC. DNSSEC has been proposed to secure the Mobile IP protocol [dnssec]

6. Security considerations

This whole document is about security considerations for MPTCP.

7. IANA Considerations

There are no IANA considerations in this memo.

8. Acknowledgments

We would like to thank Mark Handley for his comments on the attacks and countermeasures discussed in this document.

9. References

9.1. Normative References

[RFC0793] Postel, J., "Transmission Control Protocol", STD 7, RFC 793, September 1981.

[RFC3972] Aura, T., "Cryptographically Generated Addresses (CGA)", RFC 3972, March 2005.

[RFC5961] Ramaiah, A., Stewart, R., and M. Dalal, "Improving TCP's Robustness to Blind In-Window Attacks", RFC 5961, August 2010.

[RFC6056] Larsen, M. and F. Gont, "Recommendations for Transport-Protocol Port Randomization", BCP 156, RFC 6056, January 2011.

[RFC6528] Gont, F. and S. Bellovin, "Defending against Sequence Number Attacks", RFC 6528, February 2012.

9.2. Informative References

[RFC6824] Ford, A., Raiciu, C., Handley, M., and O. Bonaventure, "TCP Extensions for Multipath Operation with Multiple Addresses", RFC 6824, January 2013.

[RFC6181] Bagnulo, M., "Threat Analysis for TCP Extensions for Multipath Operation with Multiple Addresses", RFC 6181, March 2011.

[RFC5533] Nordmark, E. and M. Bagnulo, "Shim6: Level 3 Multihoming Shim Protocol for IPv6", RFC 5533, June 2009.

[RFC4960] Stewart, R., "Stream Control Transmission Protocol", RFC 4960, September 2007.

[RFC3775] Johnson, D., Perkins, C., and J. Arkko, "Mobility Support in IPv6", RFC 3775, June 2004.

[RFC2827] Ferguson, P. and D. Senie, "Network Ingress Filtering: Defeating Denial of Service Attacks which employ IP Source Address Spoofing", BCP 38, RFC 2827, May 2000.

[I-D.ietf-savi-framework]
Wu, J., Bi, J., Bagnulo, M., Baker, F., and C. Vogt, "Source Address Validation Improvement Framework", draft-ietf-savi-framework-06 (work in progress), January 2012.

[RFC4987] Eddy, W., "TCP SYN Flooding Attacks and Common Mitigations", RFC 4987, August 2007.

[I-D.paasch-mptcp-ssl]
Paasch, C. and O. Bonaventure, "Securing the MultiPath TCP handshake with external keys", draft-paasch-mptcp-ssl-00 (work in progress), October 2012.

[I-D.bittau-tcp-crypt]

Bittau, A., Boneh, D., Hamburg, M., Handley, M., Mazieres, D., and Q. Slack, "Cryptographic protection of TCP Streams (tcpcrypt)", draft-bittau-tcp-crypt-03 (work in progress), September 2012.

[hash-chains] Diez, J., Bagnulo, M., Valera, F., and I. Vidal, "Security for multipath TCP: a constructive approach", International Journal of Internet Protocol Technology 6, 2011.

[dnssec] Kukec, A., Bagnulo, M., Ayaz, S., Bauer, C., and W. Eddy, "OAM-DNSSEC: Route Optimization for Aeronautical Mobility using DNSSEC", 4th ACM International Workshop on Mobility in the Evolving Internet Architecture MobiArch 2009, 2009.

Authors’ Addresses

Marcelo Bagnulo
Universidad Carlos III de Madrid
Av. Universidad 30
Leganes, Madrid 28911
SPAIN

Phone: 34 91 6249500
Email: marcelo@it.uc3m.es
URI: http://www.it.uc3m.es

Christoph Paasch
UCLouvain
Place Sainte Barbe, 2
Louvain-la-Neuve,   1348
Belgium

Email: christoph.paasch@uclouvain.be

Fernando Gont
SI6 Networks / UTN-FRH
Evaristo Carriego 2644
Haedo, Provincia de Buenos Aires 1706
Argentina

Phone: +54 11 4650 8472
Email: fgont@si6networks.com
URI: http://www.si6networks.com

Olivier Bonaventure
UCLouvain
Place Sainte Barbe, 2
Louvain-la-Neuve,   1348
Belgium

Email: olivier.bonaventure@uclouvain.be

Costin Raiciu
Universitatea Politehnica Bucuresti
Splaiul Independentei 313a
Bucuresti
Romania

Email: costin.raiciu@cs.pub.ro