Internet Engineering Task Force
Internet-Draft
Intended status: Informational
Expires: September 22, 2016

R. Winter
NEC Laboratories Europe
M. Faath
University of Applied Sciences Augsburg
A. Ripke
NEC Laboratories Europe

March 21, 2016

Multipath TCP Support for Single-homed End-systems
draft-wr-mptcp-single-homed-07

Abstract

Multipath TCP relies on the existence of multiple paths between end-systems. These are typically provided by using different IP addresses obtained by different ISPs at the end-systems. While this scenario is certainly becoming increasingly a reality (e.g. mobile devices), currently most end-systems are single-homed (e.g. desktop PCs in an enterprise). It seems also likely that a lot of network sites will insist on having all traffic pass a single network element (e.g. for security reasons) before traffic is split across multiple paths. This memo therefore describes mechanisms to make multiple paths available to multipath TCP-capable end-systems that are not available directly at the end-systems but somewhere within the network.

Status of This Memo

Table of Contents

1. Introduction

The IETF has specified a multipath TCP (MPTCP) architecture and protocol where end-systems operate a modified standard TCP stack which allows packets of the same TCP connection to be sent via different paths to an MPTCP-capable destination ([RFC6824], [RFC6182]). Paths are defined by sets of source and destination IP addresses. Using multiple paths has a number of benefits such as an increased reliability of the transport connection and an effect known as resource pooling [resource_pooling]. Most end-systems today do not have multiple paths/interfaces available in order to make use of multipath TCP, however further within the network multiple paths are the norm rather than the exception. This memo therefore describes ways how these multiple paths in the network could potentially be made available to multipath TCP-capable hosts that are single-homed.

In order to illustrate the general mechanism we make use of a simple reference scenario shown in Figure 1.

```
                            +-------+
                            | DHCP  |
     +-------+      +----------+ Server|
     |       |      |          |       |
     | Host  +------+          +-------+
     |       |      |      +-------+        ISP 1
     +-------+      +------+       |----------
                           | Gatew.|
                           |       |----------
                           +-------+        ISP 2
```

Figure 1: Reference Scenario

The scenario in Figure 1 depicts e.g. a possible SOHO or enterprise setup where a gateway/router is connected to two ISPs and a DHCP server gives out leases to hosts connected to the local network. Note that both, the gateway and the DHCP server could be on the same device (similar to current home gateway implementations). Also, the two ISPs could really be two different access technologies (e.g. LTE and DSL) provided by a single ISP.

The host is running a multipath-capable IP stack, however it only has a single interface. The methods described in the following sections will let the host make use of the gateway's two interfaces without requiring modifications to the MPTCP implementation.

2. Approaches to Use Multiple Paths in the Network

All approaches in this document do not require changes to the wire format of MPTCP and both communicating hosts need to be MPTCP-capable. The benefit this approach has is that a) it has no implications on MPTCP standards, b) it will hopefully encourage the deployment of MPTCP as the number of scenarios where MPTCP brings benefits vastly increases and c) these approaches do not require complex middle-boxes to implement MPTCP-like functionality in the network as other approaches have suggested before.

2.1. Exposing Multiple Paths Through End-host Auto-configuration

Multipath TCP distinguishes paths by their source and destination IP addresses. Assuming a certain level of path diversity in the Internet, using different source and destination IP addresses for a given subflow of a multipath TCP connection will, with a certain probability, result in different paths taken by packets of different subflows. Even in case subflows share a common bottleneck, the

proposed multipath congestion control algorithm [RFC6356] will make sure that multipath TCP will play nicely with regular TCP flows.

In order to not require changes to the TCP implementation, we keep the above assumptions multipath TCP makes, i.e. working with different IP addresses to use different paths. Since the end-system is single-homed, all IP addresses are bound to the same physical interface. In our reference scenario in Figure 1, the host would e.g. receive more than one RFC1918 [RFC1918] private IP address from the DHCP server as depicted in Figure 2.

```
    Host                Gateway

                      +------------------+     ISP1
  +---------+         | src.             |
  |  virt. | 10.1.2.5 | 10.1.0.0/16  __.+-----------
  |        +---+      |        __.--'    |
  |  phys. |   |      |  __.--'      N   |
  |        +----------+.:_            A   |
  |        | 10.2.2.6 |   `-.._      T   |
  +---------+         | src.   `-.._     |     ISP2
                      | 10.2.0.0/16 `-..+-----------
                      |                  |
                      +------------------+
```

Figure 2: Gateway internals

The gateway that is shown in Figure 2 has received two IP addresses, one from each ISP that it is connected to (ISP1 and ISP2). The NAT that the gateway is implementing needs to "map" each private IP address of the host consistently to a one of the addresses received by the ISPs, i.e. each private IP to a different public IP. Packets sent by the host to the gateway are then routed based on the source address found in the packets as illustrated in the figure. In other words, depending on the source address of the host, the packets will either go through ISP 1 or ISP 2 and TCP will balance the traffic across those two links using its built-in congestion control mechanism.

The way the gateway has received its public IP addresses is not relevant. It could be via DHCP, IPCP or static configuration. In order to configure the hosts behind the gateway, we propose to make use of provisioning domains [RFC7556], more specifically one provisioning domain per external gateway interface (the two interfaces to ISP1 and ISP2 in Figure 2). The DHCPv6 specification for encoding provisioning domains can be found in [I-D.ietf-mif-mpvd-dhcp-support].

In order to signal to the host, that each provisioning domain will result in a different path towards the Internet, this memo introduces a new DHCP option called EXT_ROUTE, which will be included in each provisioning domain sent by the server. The option value will determine which external interface is used to sent the traffic when using the configuration information present in the respective provisioning domain.

Upon receipt of a DHCP offer including multiple provisioning domains, or multiple offers each including one or more provisioning domains, the client SHOULD create up to n virtual interfaces, where n is one less than the number of different EXT_ROUTE option values found in all received provisioning domains. Each virtual interface will contact the DHCP server and will request configuration information for the respective provisioning domains, excluding the configuration of the physical interface.

2.2. Heuristic Use of Multiple Paths

The auto-configuration mechanism above has the advantage that available paths and information on how to use them are directly sent to the end-host. In other words, there is an explicit signalling of the availability of multiple paths to the end-host. This has the advantage that the host can efficiently use these paths.

This method works well when multiple paths are available close to the end-host and means for auto-configuration are available. But that is not always the case. Another method to use different paths in the network without prior knowledge of their existence is to apply heuristics in order to exploit setups where Equal Cost Multi-path [RFC2991], a widely deployed technology [ECMP_DEPLOYMENT], or similar per-flow load-balancing algorithms are employed.

The ADD_ADDR option defined in [RFC6824] can be used to advertise the same address but a different port to open another subflow. Additionally, the MP_JOIN option can also be used to open another subflow with the same IP address and e.g. a different source port given that a different address ID is used. This means there are multiple scenarios possible (e.g. either sender-initated or receiver-initiated) where single-homed end-hosts can influence the 5-tuple (source and destination IP addresses and port numbers plus protocol number) which is often used as the basis for per-flow load balancing. Changing the 5-tuple will only with a certain probability result in using a different path unless the load-balancing algorithm that is used is known to the MPTCP implementation (an assumption we cannot generally make). This means that a number of subflows might end up on the same path. Fortunately, the MPTCP congestion control

algorithm will make sure that the collection of subflows on that path will not be more agressive than a single TPC flow.

3. Other scenarios and extensions

The reference scenario is only one conceivable setting. Other scenarios such as DSL broadband customers or mobile phones are conceivable as well. As an example, take the DSL scenario. The home gateway could be provided with multiple IP addresses using extensions to IPCP. The home gateway in turn can then implement the DHCP server and gateway functionality as described before. More scenarios will be described in future versions of this document.

4. Alternative approaches

One alternative is that a DHCP server always sends n offers, where n is the number of interfaces at the gateway to different ISPs. The client could then accept all or a subset of these offers. This approach seems interesting in environments where there are multiple DHCP servers, one for each ISP connection (think multiple home gateways). However, accepting multiple offers based on a single DHCP request is not standard's compliant behavior (at least for the DHCPv4 case). Also, to cater for a scenario that only contains a single DHCP server, server changes are needed in any case. Finally, correct routing is not always guaranteed in these scenarios.

An interesting alternative is the use of ECMP at the gateway for load distribution and let MPTCP use different port numbers for subflows. Assuming that ECMP is available at the gateway, this approach would work fine today. The only drawback of the approach is that it involves a little trial and error to find port numbers that actually hash to different paths used by ECMP [RFC2991].

5. Acknowledgements

Part of this work was supported by Trilogy (http://www.trilogy-project.org), a research project (ICT-216372) partially funded by the European Community under its Seventh Framework Program. The views expressed here are those of the author(s) only. The European Commission is not liable for any use that may be made of the information in this document.

6. IANA Considerations

One new DHCP options is required by this version of this document.

7. Security Considerations

TBD.

8. References

8.1. Normative References

[RFC1918] Rekhter, Y., Moskowitz, B., Karrenberg, D., de Groot, G., and E. Lear, "Address Allocation for Private Internets", BCP 5, RFC 1918, DOI 10.17487/RFC1918, February 1996, <http://www.rfc-editor.org/info/rfc1918>.

[RFC2991] Thaler, D. and C. Hopps, "Multipath Issues in Unicast and Multicast Next-Hop Selection", RFC 2991, DOI 10.17487/RFC2991, November 2000, <http://www.rfc-editor.org/info/rfc2991>.

[RFC7556] Anipko, D., Ed., "Multiple Provisioning Domain Architecture", RFC 7556, DOI 10.17487/RFC7556, June 2015, <http://www.rfc-editor.org/info/rfc7556>.

8.2. Informative References

[ECMP_DEPLOYMENT] Augustin, B., Friedman, T., and R. Teixeira, "Measuring Multipath Routing in the Internet", October 2011, <http://www.paris-traceroute.net/images/ton_2011.pdf>.

[I-D.ietf-mif-mpvd-dhcp-support] Krishnan, S., Korhonen, J., and S. Bhandari, "Support for multiple provisioning domains in DHCPv6", draft-ietf-mif-mpvd-dhcp-support-02 (work in progress), October 2015.

[RFC6182] Ford, A., Raiciu, C., Handley, M., Barre, S., and J. Iyengar, "Architectural Guidelines for Multipath TCP Development", RFC 6182, DOI 10.17487/RFC6182, March 2011, <http://www.rfc-editor.org/info/rfc6182>.

[RFC6356] Raiciu, C., Handley, M., and D. Wischik, "Coupled Congestion Control for Multipath Transport Protocols", RFC 6356, DOI 10.17487/RFC6356, October 2011, <http://www.rfc-editor.org/info/rfc6356>.

[RFC6824] Ford, A., Raiciu, C., Handley, M., and O. Bonaventure, "TCP Extensions for Multipath Operation with Multiple Addresses", RFC 6824, DOI 10.17487/RFC6824, January 2013, <http://www.rfc-editor.org/info/rfc6824>.

[resource_pooling]
Wischik, D., Handley, M., and M. Bagnulo Braun, "The Resource Pooling Principle", October 2008, <http://ccr.sigcomm.org/online/files/p47-handleyA4.pdf>.

Authors' Addresses

Rolf Winter
NEC Laboratories Europe
Kurfuersten-Anlage 36
Heidelberg 69115
Germany

Email: rolf.winter@neclab.eu

Michael Faath
University of Applied Sciences Augsburg
An der Hochschule 1
Augsburg 86161
Germany

Email: michael.faath@hs-augsburg.de

Andreas Ripke
NEC Laboratories Europe
Kurfuersten-Anlage 36
Heidelberg 69115
Germany

Email: andreas.ripke@neclab.eu