ETSI GR NGP 015 V1.1.1 (2019-11)

Next Generation Protocol (NGP);
Recommendation for Network Layer Multi-Path Support

Disclaimer

The present document has been produced and approved by the Next Generation Protocols (NGP) ETSI Industry Specification Group (ISG) and represents the views of those members who participated in this ISG.
It does not necessarily represent the views of the entire ETSI membership.

Reference

DGR/NGP-0015

Keywords

internet, IP, path

ETSI

650 Route des Lucioles
F-06921 Sophia Antipolis Cedex - FRANCE

Tel.: +33 4 92 94 42 00 Fax: +33 4 93 65 47 16

Siret N° 348 623 562 00017 - NAF 742 C
Association à but non lucratif enregistrée à la
Sous-Préfecture de Grasse (06) N° 7803/88

Important notice

The present document can be downloaded from:
http://www.etsi.org/standards-search

The present document may be made available in electronic versions and/or in print. The content of any electronic and/or print versions of the present document shall not be modified without the prior written authorization of ETSI. In case of any existing or perceived difference in contents between such versions and/or in print, the prevailing version of an ETSI deliverable is the one made publicly available in PDF format at www.etsi.org/deliver.

Users of the present document should be aware that the document may be subject to revision or change of status. Information on the current status of this and other ETSI documents is available at https://portal.etsi.org/TB/ETSIDeliverableStatus.aspx

If you find errors in the present document, please send your comment to one of the following services:
https://portal.etsi.org/People/CommiteeSupportStaff.aspx

Copyright Notification

Contents

Intellectual Property Rights

Essential patents

IPRs essential or potentially essential to normative deliverables may have been declared to ETSI. The information pertaining to these essential IPRs, if any, is publicly available for **ETSI members and non-members**, and can be found in ETSI SR 000 314: *"Intellectual Property Rights (IPRs); Essential, or potentially Essential, IPRs notified to ETSI in respect of ETSI standards"*, which is available from the ETSI Secretariat. Latest updates are available on the ETSI Web server (https://ipr.etsi.org/).

Pursuant to the ETSI IPR Policy, no investigation, including IPR searches, has been carried out by ETSI. No guarantee can be given as to the existence of other IPRs not referenced in ETSI SR 000 314 (or the updates on the ETSI Web server) which are, or may be, or may become, essential to the present document.

Trademarks

The present document may include trademarks and/or tradenames which are asserted and/or registered by their owners. ETSI claims no ownership of these except for any which are indicated as being the property of ETSI, and conveys no right to use or reproduce any trademark and/or tradename. Mention of those trademarks in the present document does not constitute an endorsement by ETSI of products, services or organizations associated with those trademarks.

Foreword

This Group Report (GR) has been produced by ETSI Industry Specification Group (ISG) Next Generation Protocols (NGP).

Modal verbs terminology

In the present document "**should**", "**should not**", "**may**", "**need not**", "**will**", "**will not**", "**can**" and "**cannot**" are to be interpreted as described in clause 3.2 of the ETSI Drafting Rules (Verbal forms for the expression of provisions).

"**must**" and "**must not**" are **NOT** allowed in ETSI deliverables except when used in direct citation.

Executive summary

The present document analyses the existing technologies proposed for multi-path for Internet, provides the visions for future Internet to support multi-path. It also proposes a framework to support the end-to-end multi-path in the current Internet without fundamental changes, the framework covers the most scenarios of the current network topologies for Internet.

Introduction

ETSI ISG NGP is tasked with reviewing networking technologies, architectures and protocols for the next generation of communication systems.

1 Scope

The present document reports the analysis of different multi-path ideas for network layer or IP layer. It includes the problem statement, the benefits of multi-path for networking, the existing research and technologies.

It also gives the visions for future Internet that will support network layer multi-path, also proposes the framework to support multi-path in current Internet without dramatically changing the architecture of Internet.

2 References

2.1 Normative references

Normative references are not applicable in the present document.

2.2 Informative references

References are either specific (identified by date of publication and/or edition number or version number) or non-specific. For specific references, only the cited version applies. For non-specific references, the latest version of the referenced document (including any amendments) applies.

NOTE: While any hyperlinks included in this clause were valid at the time of publication, ETSI cannot guarantee their long-term validity.

The following referenced documents are not necessary for the application of the present document, but they assist the user with regard to a particular subject area.

[i.1] ETSI GS NGP 001 (V1.1.1): "Next Generation Protocols (NGP); Scenario Definitions".

[i.2] Xuewu Xu, et.al.: "3D Holographic Display and its Data Transmission Requirement", in 2011 International Conference on Information Photonics and Optical Communications.

[i.3] K. Argyraki and D. R. Cheriton: "Loose source routing as a mechanism for traffic policies", in Proc. Future Directions in Network Architecture, 2004.

[i.4] D. Andersen, H. Balakrishnan, F. Kaashoek and R. Morris: "Resilient overlay networks", in Proc. SOSP, 2001.

[i.5] Wen Xu and Jennifer Rexford: "MIRO: Multi-path Interdomain Routing", in SIGCOMM'06, September 11-15, 2006, Pisa, Italy.

[i.6] Igor Ganichev, Bin Dai, P. Brighten Godfrey, Scott Shenker: "YAMR: Yet Another Multipath Routing Protocol", in ACM SIGCOMM Computer Communication Review, Volume 40, Number 5, October 2010.

[i.7] Murtaza Motiwala, Megan Elmore, Nick Feamster and Santosh Vempala: "Path Splicing", in SIGCOMM'08, August 17-22, 2008, Seattle, Washington, USA.

[i.8] Xiaowei Yang, David Clark and Arthur W. Berger: "NIRA: A New Inter-Domain Routing Architecture", in IEEE/ACM TRANSACTIONS ON NETWORKING, VOL. 15, NO. 4, AUGUST 2007.

[i.9] P. Brighten Godfreyy, Igor Ganichevz, Scott Shenkerzx and Ion Stoica: "Pathlet Routing", in SIGCOMM'09, August 17-21, 2009, Barcelona, Spain.

[i.10] David Barrera, Laurent Chuat, Adrian Perrig, Raphael M. Reischuk, Pawel Szalachowski: "The SCION Internet Architecture".

NOTE: See https://netsec.ethz.ch/publications/papers/SCION-CACM.pdf.

[i.11] IETF Path Aware Networking Research Group (panrg).

NOTE: See https://datatracker.ietf.org/rg/panrg/about/.

[i.12] IETF RFC 6774: "Distribution of Diverse BGP Paths".

NOTE: See https://tools.ietf.org/html/rfc6774.

[i.13] IETF RFC 7911: "Advertisement of Multiple Paths in BGP".

NOTE: See https://tools.ietf.org/html/rfc7911.

[i.14] draft-ietf-idr-add-paths-guidelines: "Best Practices for Advertisement of Multiple Paths in IBGP".

NOTE: See https://tools.ietf.org/html/draft-ietf-idr-add-paths-guidelines-08.

[i.15] International Telecommunication Union (ITU): "Internet Exchange Points (IXPs)".

NOTE: See https://www.itu.int/en/wtpf-13/Documents/backgrounder-wtpf-13-ixps-en.pdf.

3 Definition of terms, symbols and abbreviations

3.1 Terms

Void.

3.2 Symbols

Void.

3.3 Abbreviations

For the purposes of the present document, the following abbreviations apply:

AS	Autonomous System
ASN	Autonomous System Number
BBR	Bottleneck Bandwidth and Round-trip propagation time
BGP	Border Gateway Protocol
BGP-LS	BGP - Link State
DSL	Digital Subscriber Line
eBGP	external Border Gateway Protocol
ECMP	Equal-Cost Multi-Path
iBGP	internal Border Gateway Protocol
IETF	Internet Engineering Task Force
IGP	Interior Gateway Protocol
IOT	Internet Of Things
IP	Internet Protocol
IRTF	Internet Research Task Force
ISIS	Intermediate System to Intermediate System
ISP	Internet Service Provider
IT	Information Technology
IX	Internet eXchange
IXP	Internet eXchange Point
MEC	Mobile Edge Cloud
MIRO	Multi-path Interdomain Routing
MITM	Man-In-The-Middle attack
MPLS	MultiProtocol Label Switching
MPTCP	Multi-Path TCP
MSP	Multipath Service Point

NFV	Network Function Virtualisation
NIRA	New Inter-domain Routing Architecture
NNI	Network-Network Interface
OPEX	Operating Expense
OSPF	Open Shortest Path First
PANRG	Path Aware Networking Research Group
PCE	Path Computation Element
PE	Provider Edge
PKI	Public Key Infrastructure
POP	Point Of Presence
PWE	Pseudo Wire Emulation
QUIC	Quick UDP Internet Connections
RFC	Request for Comments
RINA	Recursive InterNetwork Architecture
RIP	Routing Information Protocol
SDN	Software Defined Networking
TCP	Transmission Control Protocol
UDP	User Datagram Protocol
UNI	User-Network Interface
VPN	Virtual Private Network

4 Introduction

4.1 Problem Statement

Today's Internet is based on the Internet Protocol (IPv4/IPv6). The network layer technologies used in the Internet includes many IP or related protocols for control plane, and IP forwarding in data plane. One of the major characters of Internet is it only provides one path at network layer from end-to-end for any IPv4/IPv6 destination, this path sometimes is also called as Default Path, Best Path, Shortest Path, etc. In other words, there is only one path in current Internet for any unicast IP packet to travel from end to end.

For Next Generation Protocol for future network as defined in [i.1], it is still assumed that only one path is used even for multi-homing and mobility situation.

Within one Autonomous System (AS) domain, the default path is calculated and populated by an Interior Gateway protocol (IGP) such as ISIS and OSPF. Crossing different AS domains, the default path is governed by the Border Gateway Protocol (BGP). The problems associated with the current one path strategy and the benefits of multi-path are stated below from different aspects of networking:

- Best path criteria:
  - The default path may be only the best path by one criterion, but not by other criteria. For example, IGP calculates the path based on the link cost that has different definition such as link speed or distance. But it cannot reflect some dynamic traffic related factors such as total bandwidth used, current available bandwidth, the latency for a hop, congestion status for a link, etc. It is desired that there is multiple path available, each path may have different objectives. For example, the Default Path is for the best-effort traffic, and another path is for the latency sensitive service.

- Connection backup:
  - One path cannot provide backup for end-to-end Internet connection, but multi-path can. The current Internet end-to-end backup strategy for any network outages such as link failure, node failure, re-routing due to routing policy changes, etc., relies on the routing protocol recovery mechanism. This mechanism normally involves the path re-calculation and routing information population and sync. The time consumed is normally high especially for BGP since BGP needs to populate any update globally. It is known that the corresponding BGP recovery time could stretch into hundreds of seconds or more for isolated Internet outages and lead to high packet drop rates. Some local protection and backup technologies, such as MPLS Fast Reroute and IP Fast Reroute, can only be used in restricted scenarios and cannot provide end-to-end protection for Internet. If there are multiple path, the end user can switch the traffic to another path when one path is failed. Since an application can detect path failure quickly by lost packet, this can dramatically reduce the packet drop due to network outages in Internet.

- Resource utilization:
  - One path may lead to lower network resource utilization, but multi-path may lead to higher utilization. In most of network, the traffic is not evenly distributed in all nodes and links. Theoretically, it is almost impossible to design or provision this kind of perfect network. As a mitigation, more distribution of traffic definitely will improve the utilization. Some protocol mandates more than one path to transmit traffic and can greatly improve the total throughput for applications. For example, MPTCP will only be beneficial to an end user if there is more than one path to distribute TCP traffic. With MPTCP running over multiple disjoint path, the obtained TCP throughput for end user application will be higher than one path, and naturally the user can get the redundancy or failure protection feature in case any path is failed.

- Network throughput:
  - One path may not support ultra-high throughput, but multi-path may support. Using multi-path can not only improve the efficiency of network resource utilization, but also provide support for applications that requires ultra-high throughput such as holographical display. The network bandwidth requirement for holographical display can reach the level of Tbps [i.2], and this has exceeded the speed of most of individual network link. Without multi-path to aggregate to achieve higher throughput, the application is not viable.

- Security issues:
  - Multi-path can benefit the security of application and network. When there is multi-path between two security end points, either two end-hosts or two network devices, the current security mechanism (PKI or IPSec) can be enhanced. Due to the presence of multi-path, two security end points can distribute the security related messages to multi-path, thus the possibility that whole messages are eavesdropped will be reduced dramatically. Without the complete security message, it is highly impossible to do the MITM (Man-in-the-middle attack) and other security damages. The security messages can be the exchanged key information or authentication information.

4.2 Current State

Multi-path has been a hot research topic for quite long time. Clause 4.4 will describe more details for the multi-path definitions and its impacts. Clause 4.5 gives the review of some typical proposals that can lead to multi-path support.

It should be noted that there are technologies to support some features like the multi-path discussed in the present document, but they are different, such as:

- Equal-cost multi-path (ECMP):
  - This is a technology to support multiple equal cost path. The equal cost path is only locally significant. For example, one router can choose different next hop or interface that leads to different path with the same cost. The selection is based on some policy controlled by network operator or pre-defined algorithm locally on the router, and the router is not aware of full properties of multi-path except the next hop. ECMP has been widely used within data center network, IGP domain, and between two BGP domains.

- Static configured multi-path:
  - For some scenarios, multi-path can be provided by careful pre-planning, designing and configuration for an IGP domain. Normally this needs a central controller like Network Management tool, PCE or SDN controller. The multi-path is only for local use within the network. This scheme is hard to be used in crossing different administrative domains due to complexity in management, security, business model, etc.

4.3 Proposed Targets

The present document proposes the Internet supports multi-path with the following targets, and it is obvious that above technologies in clause 4.2 cannot satisfy all requirements:

- There are no disruptive technologies introduced for multi-path support. It is based on the current Internet architecture by using new protocols or extension of existing protocols. All technologies are backward compatible.
- The multi-path includes both equal-cost and non-equal-cost paths.
- The multi-path is end-to-end in Internet.
- End-user can select one or multi-path to use, and ISP can direct the traffic to expected path(s).
- The property of each path is visible to end user, the property may (but not always) include:
  - Network topology of a path, such as node list and links associated with a node.
  - The path quality information, such as reliability, minimum and maximum bandwidth/latency/jitter.
  - The monetary information, such as cost of unit throughput, cost of different categories of latency or jitter.

4.4 Multiple Path Definitions

The present document introduces following multi-path definitions for the purpose of distinguishing different path:

- Complete Disjoint Paths:
  - When two paths do not share any network device, they are called Complete Disjoint Paths. Complete Disjoint Paths are the best to obtain more bandwidth by MPTCP and obtain the backup path protection when there is any node or link fails in any path.
- Partial Disjoint Paths:
  - When two paths share one or more network device, but do not share any L2 link, they are called Partial Disjoint Paths. Partial Disjoint Path are the best to obtain more bandwidth by MPTCP but may not be the best to obtain the backup protection. The failure of the shared node may make the backup path protection invalid.
- Joint Paths:
  - When two paths share one or more network L2 links, they are called Joint Paths. Joint Paths are not the candidate multi-path to obtain more bandwidth by MPTCP, and to obtain the backup path protection. When the shared link get congestion, the total bandwidth of MPTCP will be shrank to one TCP session can get; the backup path protection will only be effective when the failure does not happen on the shared node or links.

4.5 Review of Existing Technologies

4.5.1 Existing Proposals

There are many technologies for multi-path support, below are listed the important ones:

- Source routing [i.3].
- Overlay network [i.4].
- MIRO [i.5].
- YAMR [i.6].
- Path Splicing [i.7].
- NIRA [i.8].
- Pathlet [i.9].
- SCION [i.10].
- Activities in IRTF PANRG [i.11]:
    1) Currently the Internet architecture assumes a separation between the end hosts and the network between the endpoints. In the network, control plane protocols make routing decisions without considerations of endpoints. Endpoints have very little information about the network topology, and how the traffic is carried over the network.
    2) In 2017, the Internet Research Task Force (IRTF) created the Path Aware Networking Research Group (PANRG). PANRG is intended to extend the path knowledge from network control plane to the edge. So, endpoints can discover paths, and associate certain properties to path, further make a selection among all available paths.

4.5.2 Summary of the Existing Proposals

Table 1 is the analysis for existing proposals in terms of "Basics to obtain the multi-path info", "Multi-path Complexity" and the "Scalability".

Table 1

Technology	Basics to obtain the multi-path info	Multi-path Complexity	Scalability
Source Routing	Rely on a separate controller (PCE/SDN/etc.) to provide the multi-path info and provisioning	Depends on the algorithm running on controller	Limited by controller
Overlay Network	Reply on a separate controller (PCE/SDN/etc.) to provide the multi-path info and provisioning	Depends on the algorithm running on controller	Limited by controller
MIRO	Inter-domain: BGP based Intra-domain: Not addressed	Medium	Good, Similar to BGP
YAMR	Inter-domain: BGP based Intra-domain: Not addressed	Medium	Good, Similar to BGP
Path Splicing	Inter-domain: BGP based Intra-domain: Multi IGP instance	Medium	Good, Similar to BGP
NIRA	Inter-domain: Not BGP, New scheme based on hierarchical architecture of ISPs Intra-domain: Not addressed	High	Worse than BGP
Pathlet	Inter-domain: Not BGP, New algorithm like IGP Intra-domain: Not addressed	Medium	Worse than BGP
SCION	Complete new architecture and scheme for both intra-domain and inter-domain routing	High	Worse than BGP