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

GROUP REPORT

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

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

From the summary of the above analysis, and the evolution history of routing protocols in Internet, it is not hard to have conclusions as below:

- The more complicated the solution, the lower the possibility that the industry will deploy it in short term.
- Scalability is a key factor for a new algorithm to be considered.
- An Intra-domain multi-path solution is easier than an Inter-domain solution.
- For Inter-domain multi-path solution, a non-BGP based algorithm is harder to be adopted.
- Multi-path and security are separate issues.

5 Visions for Future Internet to Support Multi-Path

5.1 IP Evolution and Future Internet

The Internet has evolved for over 40 years. There are many technologies developed, from software to hardware/silicon and from control plane to data plane. However, the network layer or IP (IPv4 and IPv6) and its associated protocol/technologies are relatively steady and still dominant. IP is the default and mandatory feature for network equipment, desktop/laptop, smart phone, etc. From the perception of predictable future, IP should be at least one of the most important technology in Future Internet, if not the dominant. However, from the perception of future progression, limitations inherent in the original IP design spur the research into alternatives to improve performance, scalability and security.

The following list is the key IP associated technology/protocol widely deployed so far:

- L3 Intra-domain: RIP, ISIS, OSPF, BGP, MPLS, L2/L3 VPN, IPSec.
- L3 Inter-domain: BGP.
- L4: TCP, UDP, QUIC, BBR.

The development of IP technologies in industry including both ISP and enterprise has indicated:

- BGP is the key protocol for inter-domain routing.
- ISIS/OSPF are the key protocol for intra-domain routing.
- MPLS is widely used within ISP's AS for the multiple services such as VPN, tunnelling, PWE, leased line, etc. But MPLS is not widely deployed cross different ISP domains even the protocol supports it.

Some new hot topics have been emerged recently due to the fast growth of the industries: 5G, Data Center, Cloud, MEC, etc. These topics include mobility, virtualization, NFV, IGP flood reduction, BGP for intra-domain routing, etc.

Disruptive research for multi-path architecture includes RINA and SCION.

5.2 Why Not Deployed

There are so many proposed multi-path solutions in academic research in recent decades, but none of them are deployed commercially. This may be due to many reasons, but below should be major reasons:

- Market driven:
  - The market requirement for multi-path is not strong enough to stimulate the service providers to seriously consider the investment in network for new services derived from multi-path.

- Deployment complexity between ISPs:
  - To achieve the benefit of multi-path as discussed in clause 4.1, multiple service providers have to cooperate together to provide Disjoint Path to obtain the end-to-end effect to customers. This cooperation is not only between different ISPs within one country, but also cross different countries and regions. There are probably legal issues involved.
- The deployment history of running protocols between different ISPs is that the ISPs prefer the protocol is:

  a) as simple as possible;

  b) easy for an ISP to enforce routing policy;

  c) exposes an ISP's internal network information as less as possible.

  The preference is coming from the consideration of security, business mode, and cost of operation and management. For example, complicated protocols may result in the difficulty of billing, high OPEX, and weakness of security.

  All above factors lead to the reality that the routing protocol between different ISPs is typically BGP. There is some inter-domain protocol other than BGP, such as inter-domain MPLS for VPN, but they are only deployed across different AS within one ISP, and rarely between different AS of different ISP.
- Technology maturity:
  - So far, there is no widely accepted technology in both academia and industry for multi-path. No big-scale experiments were done to prove the technical completeness, scalability, performance, and robustness of any technology.
- Alternative solution:
  - The multi-path benefits are always partially achievable through other alternative solutions. One of the simplest solutions is the network expansion. By deploying more network devices and higher speed of links to obtain more network capacity, the issues that have driven the need of multi-path solution such as for MPTCP can always be mitigated.

Even above aspects have been the major obstacles for the deployment of multi-path solution, the pressure from market has been mounted from 5G and IOT. By expanding network capacity only, the physical limit will be reached sooner or later. Multi-path solution needs to be ready, and industry will gradually start to deploy.

5.3 Prospect of Multi-path Technology

5.3.1 Evolution

The present document predicts the multi-path support will gradually become as an important feature in Internet with the growth of new market in 5G, MEC, IOT, etc. The evolution will have the milestones as:

1) The 1st milestone is that multi-path will be available for a special end-to-end scenario. This scenario will need very light network provisioning, but without the need of any new technologies. The scenario includes the characters:

   a) The client communication end has multiple access network connected to internet through LTE, Wi-Fi, DSL, etc. The access network may belong to different ISP.

   b) The ISP's network connecting to two end points are connected directly with multiple links or connected to an IXP with multiple links. This is shown in figure 1.

2) The 2nd milestone is for a special end-to-end scenario; this scenario is the same the scenario in 1st milestone but with the difference as followings. This scenario will need very light network provisioning, and need new protocol at access network side:

   a) The client communication end has only one access network connected to internet.

b) The ISP's network of two end points is connected directly with multiple links or connected to an IXP with multiple links. This is shown in figure 2.

3) The 3rd milestone is for a more general end-to-end scenario. This scenario may need new protocol at access network side, and extensive changes to current protocols between ISP's network between access networks:

a) The client communication end has one or multiple access network connected to internet.

b) The ISP's network of two end points is connected to a transit networks with multiple links. The transit networks are composed of one ISP's networks or multiple ISP's networks. This is shown in figure 3 and figure 4 respectively.

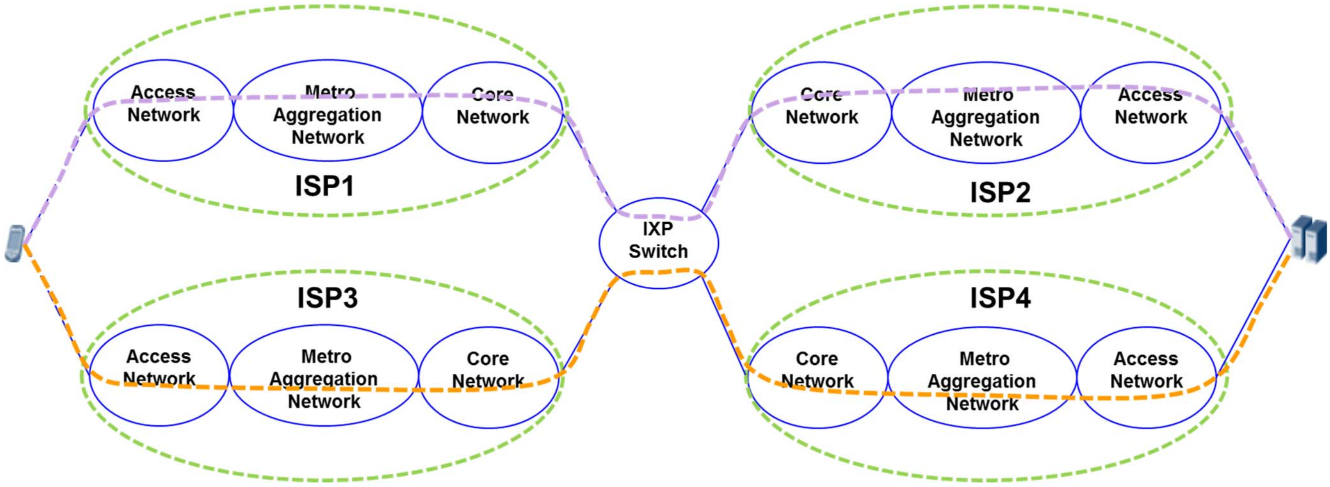

Figure 1: Two access ISP for each end-user and two access ISP are connected by IXP

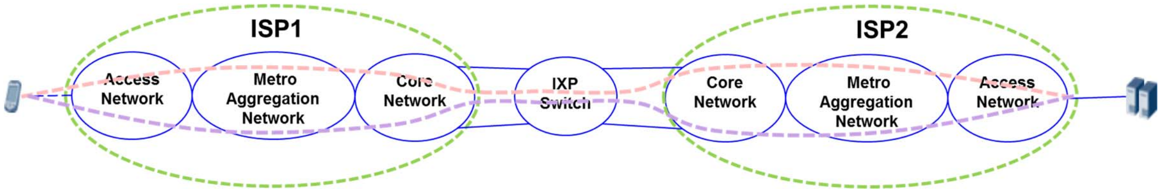

Figure 2: One access ISP (ISP1 = ISP2) for each end-user and two access ISP are connected by IXP

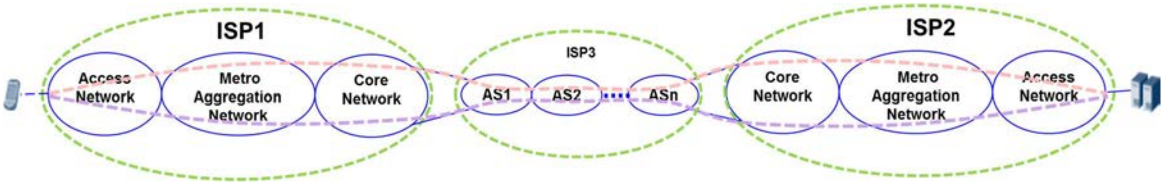

Figure 3: Two access ISP (ISP1 != ISP2) for each end-user and two access ISP are connected by one transit ISP

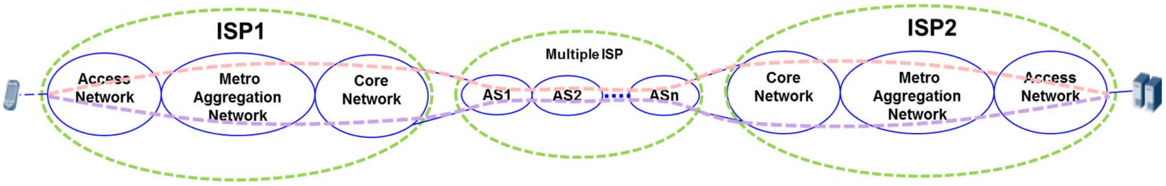

Figure 4: Two access ISP (ISP1 != ISP2) for each end-user and two access ISP are connected by multiple transit ISP

5.3.2 Framework

Figures 5 and 6 demonstrate the framework for the multipath technologies supported in a Future Internet. The framework is composed of the network topology, new component MSP, the interfaces and the key technologies:

1) The network topology can be divided as the segment of access networks and the segment of transit networks:

   a) The access networks connecting to end-user devices:

      - These networks normally belong to an individual ISP and are shown as the networks of ISP1 and ISP2 in figure 5. They may include access network, metro aggregation network, and core network. Multiple AS domains may exist. i.e. each access network, aggregation network and core network have its own ASN. All these ASN are owned by one ISP.

      - One user device may connect to multiple ISP's network described as above that can be shown in figure 1. Figure 5 only shows one ISP's networks connecting to end-use device.

   b) The transit networks between two ISP's access networks described above. There are two scenarios for these networks:

      - The transit networks belong to one ISP. One or multiple AS domains may exist, see the ISP3 in figure 5.

      - The transit networks belong to multiple ISPs. Multiple AS domain exist, see figure 6.

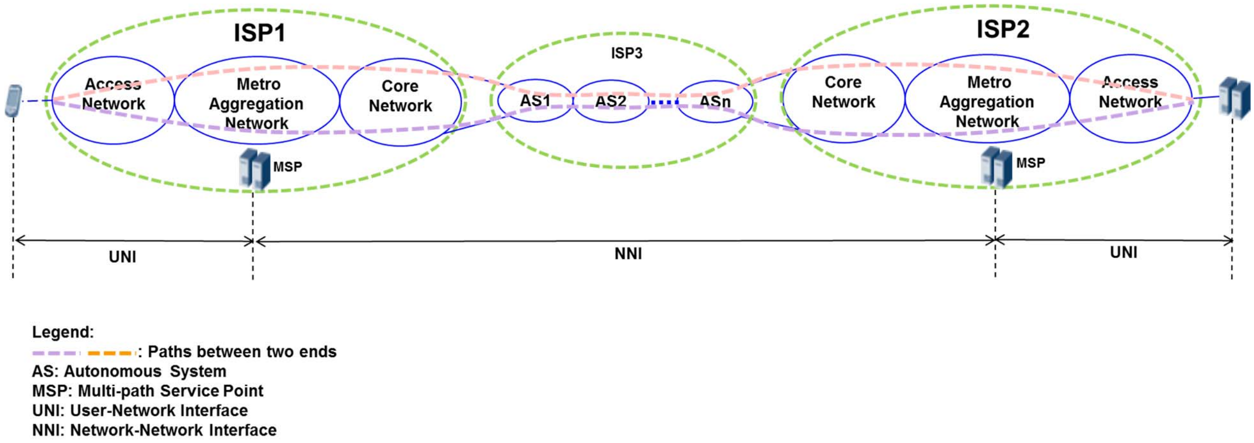

Figure 5: UNI and NNI for a specific Internet Architecture

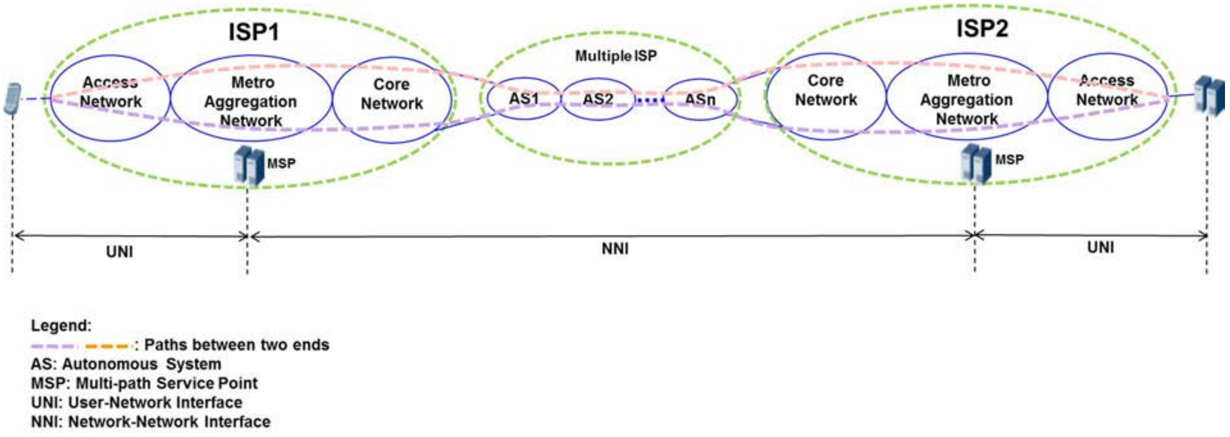

Figure 6: UNI and NNI for a general Internet Architecture

MSP is Multi-path Service Point. It provides a new functionality for multiple path service. MSP can be an independent server or attached to another network or computer device.

The key functions provided by MSP is the interface between user and network:

- User-network interface between network and user for multi-path info exchange, more details are described in clause 6.4.1.
- Network-network interface between networks for multi-path info exchange, more details are described in clause 6.4.2.

There are four key technologies in the framework listed as below, and described in detail in clause 5.3.3:

1) Protocols for Intra-domain to obtain the multi-path info.
2) Protocols for Inter-domain to obtain the multi-path info.
3) Protocols (UNI) between network and user for multi-path info exchange.
4) Protocols (NNI) between network and network for multi-path info exchange.

5.3.3 Prospected Solutions

The present document proposes solutions for multi-path in future Internet. The multi-path here means the Disjoint Multi-path. The whole solution assumes that the current Internet architecture will not be changed in the foreseeable future (5 ~ 10 years), this includes following technologies are not fundamentally changed or removed:

1) Internet is still composed of many connected domains; each domain is represented by an Autonomous System (AS) Number. One ISP may have one or more networks and each network has one AS assigned. The AS described here is the public AS number assigned by IANA^{TM}. In fact, there could have private AS in one public AS domain.
2) IPv4 and IPv6 packet format do not have fundamental changes.
3) Routing protocols for intra-domain, and for inter-domain are still dominated by IGP (ISIS and OSPF) and BGP. Other technology, such as PCE, SDN, are still only used in an intra-domain or within one ISP.
4) IGP and BGP fundaments are not changed in message exchanging mechanism, state machine, etc.
5) The data plane for intra-domain can be either by native IP forwarding or by other technologies such as Ethernet or MPLS switch.
6) The data plane for Inter-domain still use the native IPv4 or IPv6.

5.3.4 Prospected Key Technologies

The four key technologies in the framework described in clause 5.3.2 are the backbone of the prospected solutions. Due to the limit of the document, the present document only gives the direction of research for those key technologies instead of detailed proposes.

1) Protocols for Intra-domain to provision multi-path and obtain the multi-path info:
   - There are two categories for these protocols or technologies:
     a) One is based on the extension of distributed protocols like IGP or iBGP. More search is needed for the support of multi-path by traditional protocols. Path Splicing is a good candidate for this regard.
     b) Another is based on the central computation entity or controller like PCE or SDN. For the solution of this category, the major work exists on the performance and scalability optimizations. There is already deployment for such technologies. For example, use PCE integrated with BGP-LS to compute multi-path for multi-domains; use SDN and topology discovery for intra-domain path calculation.
   - In general, some current technologies are able to provide limited multi-path provisioning, just better solution and enhancement or performance optimization are needed.

2) Protocols for Inter-domain to provision multi-path and obtain the multi-path info:

- eBGP is still the only candidate to support multi-path for Inter-domain, and it is the most critical and challenge part for multi-path technology to be adopted in future Internet.
- Currently, multi-path eBGP (IETF RFC 6774 [i.12], IETF RFC 7911 [i.13], draft-ietf-idr-add-paths-guidelines [i.14], etc.) can provide some level of multi-path support, but they are far from the global deployment. They cannot guarantee the multi-paths are disjoint cross AS domains, and many other limits exist.
- Since there is no sign in industry to adopt any new technologies described in clause 4.5, the vision for the new technology for inter-domain multi-path support is that:
    - New technologies may still be based on the current BGP architecture but with big extensions to provide multi-path information between inter-domain.
    - Intensive research is needed to solve the issues in multi-path info exchange, global population and sync, performance, scalability, security, etc.
- By using the current eBGP multi-path protocol, the end-to-end multi-path feature is feasible for some special transit network topologies. This include scenarios described in clause 5.3.1:

    a) The transit networks belong to one ISP. One or multiple AS domains may exist, see figure 5.

- For this type of network topology, careful design of eBGP with multi-path support and use of protocols described below would make the multi-path available to end-user.

3) Protocols between network and user for multi-path info exchange:

- The protocol is for the multi-path info and requirement exchange between user and ISP. The detailed protocol is addressed in clause 6.4.1.

4) Protocols between network and network for multi-path info exchange:

- The protocol is for the multi-path info and requirement exchange between ISPs. The detailed protocol is addressed in clause 6.4.2.

6 Framework to Support end-to-end Multi-Path in Current Internet Architecture

6.1 Overview

In clause 5.3, the framework to support multi-path in future Internet is discussed, it is also concluded that the multi-path is feasible for some special network topologies by using enhanced current technologies. Figures 1 to 3 show the typical topologies in current Internet that can provide the end-to-end multi-path support by provisioning the current protocol and introducing new protocol of UNI and NNI.

In clause 6.3, two scenarios are analysed for that special network topologies. One is that Internet Exchange Point (IXP) [i.15] is used to connect two ISP's access network; Another is that there is only one ISP's network (transit ISP network) between two access networks. More details about network deployment and new protocols will also be illustrated.

6.2 IXP

Figures 1 and 2 show the topology that Internet Exchange Point (IXP) is used.

Internet exchange point (IX or IXP) is the physical infrastructure through which IXP members, Internet service providers (ISPs) and content delivery networks (CDNs), can exchange IP routes and Internet traffic between their networks.

Basic IXP facility will provide either L1/2 or L3 inter-connection for different IXP members. Members of an IXP can establish direct BGP peering relationship by setup one or multiple eBGP (external BGP) sessions. The global IP routes or the public achievable IP prefix of different members are exchanged through eBGP. After the eBGP routing convergency, the Internet traffic can travel through different IXP members.

Currently, the major IXP in different continents already attracted the presence of most of the major ISP and content providers. In theory, if all ISP and content providers have a point of presence (POP) in one IXP, then all end-users can access all content provider network or communicate with each other through the IXP without going through a third part of the transit ISP.

When there is only IXP between two end users (figure 2), the multi-path support will be simplified to be the issue that how to provide multi-path to two end-user devices connected to one ISP. This is because IXP is essentially an equivalence to the direct connection of two access networks belonging to one ISP. Details will be discussed in clause 6.3.1.

6.3 Multi-path Support for Special Network Topologies

6.3.1 ISP Provides Internet Access for End-user

This is for ISP that has access network connected to the end-user, such as networks of ISP1 and ISP2 shown in figures 5 and 6.

For the purpose of providing multi-path for an end-user connected to an access ISP network, the ISP should do:

1) Be aware of if it has multiple links to outside of ISP's network, for example, multiple links connected to an IXP or another transit ISP's network.

2) Provision the separate IP forwarding path or tunnel between the end-user and PE having the links to outside of ISP network. This is demonstrated as the dot lines in figure 7. The separate IP forwarding path should be disjoint with the default IP path from/to outside of ISP network.

3) There are different ways to provision the non-default path and multi-path for native IP forwarding or IP over MPLS tunnel, such as static routing configuration, segment routing, RSVP-TE, SDN programming the path, etc.

4) After the provision is done, the information is sent to the MSP. MSP for an ISP provides functions to exchange multi-path info with end-user by UNI or with another ISP's MSP by NNI. MSP could be co-located with the device that provides IPv6 neighbour discovery or DHCPv6 service.

5) After the provision is done, the multi-path eBGP session with neighbour ISP's BGP peer can be configured on the PE router properly to receive and advertise the multi-path info with its BGP peers. The neighbour ISP could be either the access ISP connecting the content provider or end-user, or the transit ISP between two access ISP.

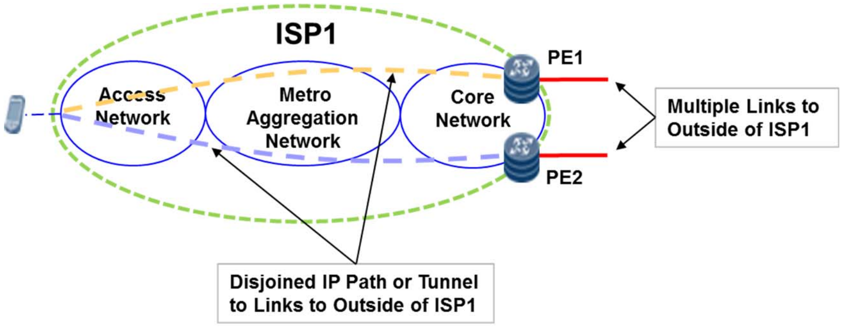
ISP1
Access Network
Metro Aggregation Network
Core Network
PE1
PE2
Multiple Links to Outside of ISP1
Disjoined IP Path or Tunnel to Links to Outside of ISP1

Figure 7: Network Provisioning for Access Network ISP

6.3.2 ISP Provides Transit Service for Other Access ISP

This is for ISP that has transit networks connected to the access network of end-user, for example, the networks of ISP3 between ISP1 and ISP2 shown in figure 5.

For the purpose of providing multi-path for a transit ISP network, the ISP should do:

1) Be aware of if it has multiple links to outside of ISP's network, for example, multiple links connected to outside of transit ISP's network shown in figure 8.

2) Provision the separate IP forwarding path or tunnel between the PEs having links to outside of ISP network. This is demonstrated as the dot lines in figure 8. The separate IP forwarding path should be disjoint with the default IP path from/to outside of ISP network.

3) There are different ways to provision the non-default path and multi-path for native IP forwarding or IP over MPLS tunnel, such as static routing configuration, segment routing, RSVP-TE, SDN programming the path, etc.

4) After the provision is done, the multi-path eBGP session with neighbour ISP's BGP peer can be configured on the PE router properly to receive and advertise the multi-path info with its BPG peers.

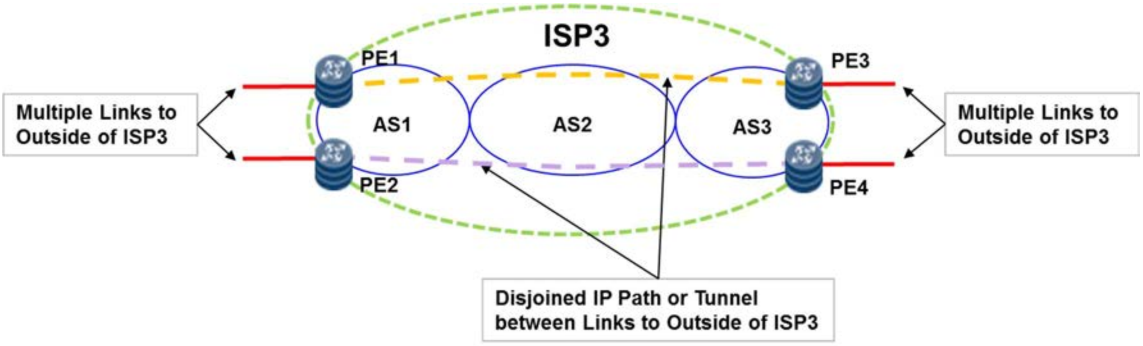
ISP3
PE1
PE2
PE3
PE4
AS1
AS2
AS3
Multiple Links to Outside of ISP3
Multiple Links to Outside of ISP3
Disjoined IP Path or Tunnel between Links to Outside of ISP3

Figure 8: Network Provisioning for Transit Network ISP

6.4 New Interfaces

6.4.1 User-Network Interface

User-network interface (UNI) is between end-user device and the MSP. It is responsible for the message exchange about the multi-path related information. The information includes:

1) User Service Expectation. It is the message from end-user device to the MSP to describe user's expectation in network service, such as Maximum/Minimum bandwidth required, maximum latency, etc.

2) Path Quality Info. It is the message from MSP to end-user device to notify each path's network quality parameters, such as path index, maximum/minimum bandwidth for a direction, maximum latency, etc.

3) Path Aware Info. It is the message from MSP to end-user device to notify each path's network proxy parameters, such as path index, proxy address.

4) Path Segment Info. It is the message from MSP to end-user device to notify each path's network segment parameters, such as path index and associated list of segments. This message is used when the segment routing is used for multi-path data plane.

5) Path Label Info. It is the message from MSP to end-user device to notify each path's MPLS label parameters, such as path index and associated list of MPLS labels. This message is used when the MPLS is used for multi-path data plane.

6) Data Plane Info. It is the message from MSP to end-user device to notify end-use to send data by the specified data plane, native IPv6, SRv6 or MPLS.

UNI interface could be a new defined protocol, or through the extension of existing protocols such as DHCPv6 or IPv6 Neighbour Discovery.

6.4.2 Network-network interface

Network-network interface (NNI) is between two access network's MSP. It is responsible for the message exchange about the multi-path related information. The information includes:

1) Network Address. It is the prefix a MSP supports for the multi-path feature.

2) Path Quality Info. It is each path's network quality parameters, such as path index, maximum/minimum bandwidth for a direction, maximum latency, etc.

NNI interface could be a new defined protocol, or through the extension of existing protocols such as BGP.

Annex A: Change History

Date	Version	Information about changes
1/2019	1.1.1	Initial Version

History

Document history		
V1.1.1	November 2019	Publication