ETSI GR NGP 014 V1.1.1 (2019-10)

Next Generation Protocols (NGP);
Preferred Path Routing (PPR) for Next Generation Protocols

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-0014

Keywords

mobility, QoS, routing support

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

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 focuses on using Preferred Path Routing (PPR) for next generation architectures towards various access, transport and DC networks and beyond. The basic concept of PPR consists of a novel path routing paradigm for various data planes that allows to provide dynamic traffic engineering optionally with resource reservation along the path. PPR can be deployed in a way that supports current architectures, while also enabling more optimal future architectures. The work aims to examine and propose recommendations to improve and simplify the network infrastructure to support optimized source routing aligned with SDN/NFV infrastructure natively by adopting PPR. In addition, the present document may require the development of new protocols and/ or modification of existing protocols.

Introduction

The present document provides recommendations toward new protocols and/or modification of existing ones in the context of PPR.

1 Scope

The present document provides brief overview of existing routing mechanisms, traffic engineering and proposes next generation source routing which can support multiple and extensible data planes with hard SLA guarantees with dynamic resource reservations along the path. It explores new TE framework for high-precision transport networks by signalling simple linear paths as well as graph structures to provide compact and yet provide better scalability of overall paths in the network.

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] ISO/IEC 10589:2002 (Second edition): "Intermediate system to intermediate system intra-domain-routing routine information exchange protocol for use in conjunction with the protocol for providing the connectionless-mode Network Service (ISO 8473)".

[i.2] IETF RFC 2328: "OSPF Version 2", Moy, J., April 1998.

NOTE: Available at https://www.rfc-editor.org/info/rfc2328.

[i.3] IETF RFC 5340: "OSPF for IPv6", Coltun, R., Ferguson, D., Moy, J., and A. Lindem, July 2008.

NOTE: Available at https://www.rfc-editor.org/info/rfc5340.

[i.4] IEEE Global Communications Conference (GLOBECOM): "The Segment Routing Architecture", C. Filsfils, N. K. Nainar, C. Pignataro, J. C. Cardona, P. Francois, 2015, San Diego, CA, 2015.

[i.5] draft-ietf-spring-segment-routing-mpls-13: "Segment Routing with MPLS data plane", A. Bashandy, C. Filsfils, S. Previdi, B. Decraene, S. Litkowski, R. Shakir, (work in progress), IETF, April 2018.

[i.6] draft-ietf-6man-segment-routing-header-12: "IPv6 Segment Routing Header (SRH)", S. Previdi, C. Filsfils, J. Leddy, S. Matsushima, D. Voyer, (work in progress), IETF, April 2018.

[i.7] IETF RFC 5440: "Path Computation Element (PCE) Communication Protocol (PCEP)", J.P. Vasseur, J.L. Le Roux, March 2009.

NOTE: Available at https://www.rfc-editor.org/info/rfc7752.

[i.8] draft-ietf-isis-segment-routing-msd-10"Signaling MSD (Maximum SID Depth) using IS-IS", J. Tantsura, U. Chunduri, S. Aldrin, L. Ginsberg, (work in progress), IETF, April 2018.

[i.9] IETF RFC 5714: "IP Fast Reroute Framework", Shand, M. and S. Bryant, January 2010.

[i.10] IETF RFC 7490: "Remote Loop-Free Alternate (LFA) Fast Reroute (FRR)", Bryant, S., Filsfils, C., Previdi, S., Shand, M., and N. So, April 2015.

NOTE: Available at http://www.rfc-editor.org/info/rfc7490.

[i.11] IEEE Global Communications Conference (GLOBECOM): "Preferred Path Routing - A Next-Generation Routing Framework Beyond Segment Routing", U. Chunduri, A. Clemm, R. Li, 2018, Abu Dhabi, UAE, December 2018.

[i.12] IEEE Hi-Precision Networks (HIPNET): "Preferred Path Routing (PPR) Graphs Beyond signalling of paths to Networks", T. Eckert, Y. Qu, U. Chunduri, 2018, Rome, Italy, 2018.

[i.13] IETF RFC 2025: "The Simple Public-Key GSS-API Mechanism (SPKM)".

[i.14] IETF RFC 3209: "RSVP-TE: Extensions to RSVP for LSP Tunnels".

[i.15] IETF RFC 7810: "IS-IS Traffic Engineering (TE) Metric Extensions".

[i.16] IETF RFC 7471: "OSPF Traffic Engineering (TE) Metric Extensions".

[i.17] IETF RFC 8402: "Segment Routing Architecture".

[i.18] draft-ietf-rtgwg-segment-routing-ti-lfa-01: "Topology Independent Fast Reroute using Segment Routing".

[i.19] IETF RFC 8300: "Network Service Header (NSH)".

NOTE: Available at https://tools.ietf.org/html/rfc8300.

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:

3GPP	3rd Generation Partnership Project
5G	Fifth Generation Mobile Networks
BGP	Border Gateway Protocol
CPU	Central Processing Unit
CSR	Cell Site Router
DB	DataBase
DC	Data Centre
DCI	Data Centre Interconnect
EH	IPv6 Extension Headers
FEC	Forwarding Equivalence Class
FIB	Forwarding Information Base
FRR	Fast ReRoute
GPRS	General Packet Radio Service
GTP	GPRS Tunneling Protocol
HW	Hardware
IGP	Interior Gateway Protocol
IP	Internet Protocol
IS-IS	Intermediate System to Intermediate System
LDP	Label Distribution Protocol

LFA	Loop Free Alternative
LSA	Link State Advertisement (OSPF)
LSP	Link State PDU (IS-IS)
LTE	Long Term Evolution
MPLS	Multi-Protocol Label Switching
MSD	Maximum SID Depths
MSDC	Massive Scale Data Centre
MTU	Maximum Transmission Unit
NGER	Next Generation Explicit Routing
NH	NextHop
NR	New Radio
OAM	Operations, Administration and Maintenance
OSPF	Open Shortest Path First
PCE	Path Computation Element
PDE	Path Description Element
PE	Provider Edge
PPG	Preferred Path Graph
PPR	Preferred Path Routing
PPR-ID	Preferred Path Route Identifier
PPR-TE	Preferred Path Route- Traffic Engineering
QFI	QoS Flow Identifier
RQI	Reflective QoS Indicator
RSVP	Resource Reservation Protocol
SDN	Software-Defined Networks
SID	Segment IDentifier
SLA	Service Level Agreement
SP	Service Provider
SPF	Shortest Path First
SR	Segment Routing
SRH	Segment Routing Header
TCP	Transmission Control Protocol
TE	Traffic Engineering
TLV	Type Length Value
UDP	User Datagram Protocol
UE	User Equipment
UPF	Use Plane Functionality (5G)
WI	Work Item

4 Background

4.1 Overview

Routing is a fundamental concept in packet networks. This clause provides background about routing technologies that are dominant today and points out certain shortcomings. This sets the stage for the introduction of PPR in the subsequent clause.

4.2 Shortest-path routing

Much of today's routing is based on the concept of Shortest Path Routing, based on the idea to always attempt to route packets along a path that is the "shortest", or of the least cost.

Research of SPF algorithm (invented by Dijkstra) started in 1970's and variations/modifications of this core algorithm is widely deployed with link state protocols or IGPs (OSPFv2 [i.2], IS-IS [i.1], OSPFv3 [i.3]). In IGPs, a directed graph is computed with flooded link state information (LSA/LSP DB) with links having configured weights/metrics. SPF Algorithm calculates a tree of shortest path from self to all other nodes in the network with candidate list of nodes kept sorted by weight. Shortest (best) value in the candidate and downloaded to the routing table with the computed immediate Next-Hop (NH). IP routing table only needs NH to each advertised prefix from all the nodes while LSA/LSP tree has all the paths. In the example below a shortest path from Rs is shown to a prefix advertised from Rd is shown with NH set to R11. Similar to Rd all Rs would compute NHs for all prefixes advertised from all the nodes in the network.

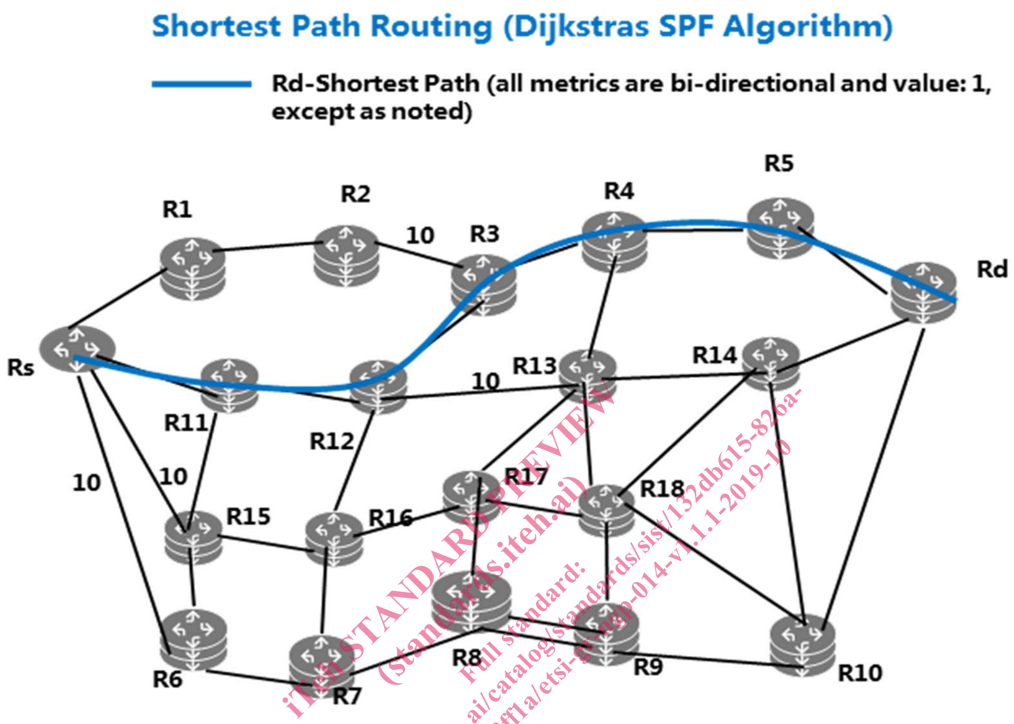

Figure 1: Concept of a shortest path

One drawback of shortest-path routing concerns that it is not always the shortest path that may be preferred, as there may be different cost metrics and other considerations (such as load balancing, ease of failover service levels, or robustness of path). As a result, other technology has begun to appear that allows to route on paths other than shortest paths. One such technology is Segment Routing (see clause 4.3).

4.3 Segment Routing

Segment Routing [i.4] is a novel source routing approach, which enables packet steering with a specified path in the packet itself. This is defined for MPLS (with a set of stacked labels) and IPv6 (path described as list of IPv6 addresses in SRHeader), and data planes called SR-MPLS [i.5] and SRv6 [i.6] respectively.

SR simplifies the Multi-Protocol Label Switching (MPLS) control plane by distributing Segment Identifiers (SIDs) for routing prefixes, which are constitute MPLS global labels into Interior Gateway Protocols. This allows source routing to be achieved by representing the network path with stacked SIDs on the data packet without any changes to the MPLS data plane. In addition to MPLS, as specified above, SR also introduces an IPv6 Extension Header (EH) for use with the IPv6 data plane, resulting in SRv6. In SRv6, a segment is encoded as an IPv6 address, with a new type of IPv6 Routing Header (EH) called SRH. A set of segments is encoded as an ordered list of IPv6 addresses in SRH to represent the path of the data packet.

Segments and source routes can be computed by a controller with knowledge of the network topology and subsequently provision the network with end-to-end SR paths. A controller could include e.g. a Path Computation Element (PCE) [i.7] or another type of SDN controller. Using a controller allows to perform different optimization and customizations to paths that take into account different constraints. This also obviates the need for traditional MPLS control plane protocols like LDP and RSVP, reducing the number of protocols that need to be deployed in a network.

To illustrate, Figure 2 depicts an example of an SR Path. To represent this path, a stack of 8 labels from node R15 to Rd is needed. In fact, more than 8 labels may be required to accommodate entropy labels, which become necessary for better load balancing of traffic across MPLS networks.

However there are some issues/drawbacks with SR:

1) The additional path overhead with complete path using SIDS on the data packet in various SR deployments may cause the following issues:

   - HW Capabilities: Not all nodes in the path can support the ability to push or read label stack needed Maximum SID Depth (MSD) [i.8] to satisfy user/operator requirements. Alternate paths which meet these user/operator requirements may not be available.

   - Line Rate: Potential performance issues in deployments, which use SRH data plane with the increased size of the SRH with 16 byte SIDs.

   - MTU: Larger SID stacks on the data packet can cause potential MTU/fragmentation issues.

   - Header Tax: Some deployments, such as 5G, require minimal packet overhead in order to conserve network resources. Carrying 40 or 50 octets of data in a packet with hundreds of octet of header would be an unacceptable use of available bandwidth.

2) Another limitation of SR concerns the fact that while it allows a data packet to steer through a custom path, by itself it cannot guarantee that proper QoS along the path needed. The ability to manage resource reservations or to provide traffic engineering attributes are not in SR's scope.

3) A more subtle issue concerns the ability to conduct performance measurements and collect traffic accounting statistics using SR-MPLS. Because labels on data packets refer only to individual path segments, attributing statistics of any particular packet to a path or flow is inhibited and difficult to perform efficiently.

4) SR cannot be applied to native IPv4/IPv6 data planes. While SR can be supported with MPLS without any changes in the data plane, use with IPv6 requires an SRH extension header, whose support requires hardware upgrades across the network. While SR is considered as a potential alternative for backhaul transport networks (like 5G), non-support for native IP data planes imposes a significant hurdle on SR adoption, as many cellular networks around the world still use native IPv4 and IPv6 data planes. As path steering capability is an essential component for network slicing in 5G backhaul transport, lack of this capability forces operators to upgrade the hardware for SRH support.

5) Last but not least SR also defines complex FRR approach with Topology Independent LFA (TI-LFA) [i.18]. Here, the post convergent backup path does not reflect the original SR path QoS characteristics. This is because alternative path is computed in a distributed fashion by the network nodes using LFA/RLFA [i.9] and [i.10] algorithms which can only give a loop free shortest path to the destination.

5 Preferred path routing concept and architecture

5.1 Overview

PPR is a new source routing paradigm where a prefix is signalled in a routing domain (control plane) along with a data plane identifier as well as path description on how the packets has to be forwarded when actual data traffic with the data plane identifier is seen. This builds on existing IGPs and fits well with SDN paradigm as the needed path can be crafted dynamically based on various inputs from a central entity.

5.2 PPR Core Concept

Traditionally routing in network is based on shortest path computations (through Interior Gateway Protocols or IGPs [i.1], [i.2] and [i.3]) for all prefixes in the network. As explained, Segment Routing allows to compute custom paths (other than shortest paths) that are subsequently represented by a sequence of segment identifiers in a packet, leading to another set of problems.

Preferred Path Routing (PPR) enables route computation based on a specific path described along with the prefix as opposed to shortest path towards the prefix. The key change that is required concerns how the next hop is computed for the prefix. Instead of using the next hop of the shortest path towards the destination, the next hop towards the next node in the path description is used. PPR is a novel architecture to signal explicit path and per-hop processing requirement and optionally including QoS or resources to be reserved along the path.

PPR is concerned with the creation of a routing path as specified in the PPR-Path which is advertised in IGPs along with a data plane identifier (PPR-ID). With this, any packet destined to the PPR-ID would use the PPR-Path instead of the IGP computed shortest path to the destination indicated by the PPR-ID. In other words, packets destined to the PPR-ID may use the PPR-Path instead of the IGP computed shortest path. This works as follows: IGP nodes process the PPR-Path. If an IGP node finds itself in the PPR-Path, it sets the next-hop towards the PPR-ID according to the PPR-Path.

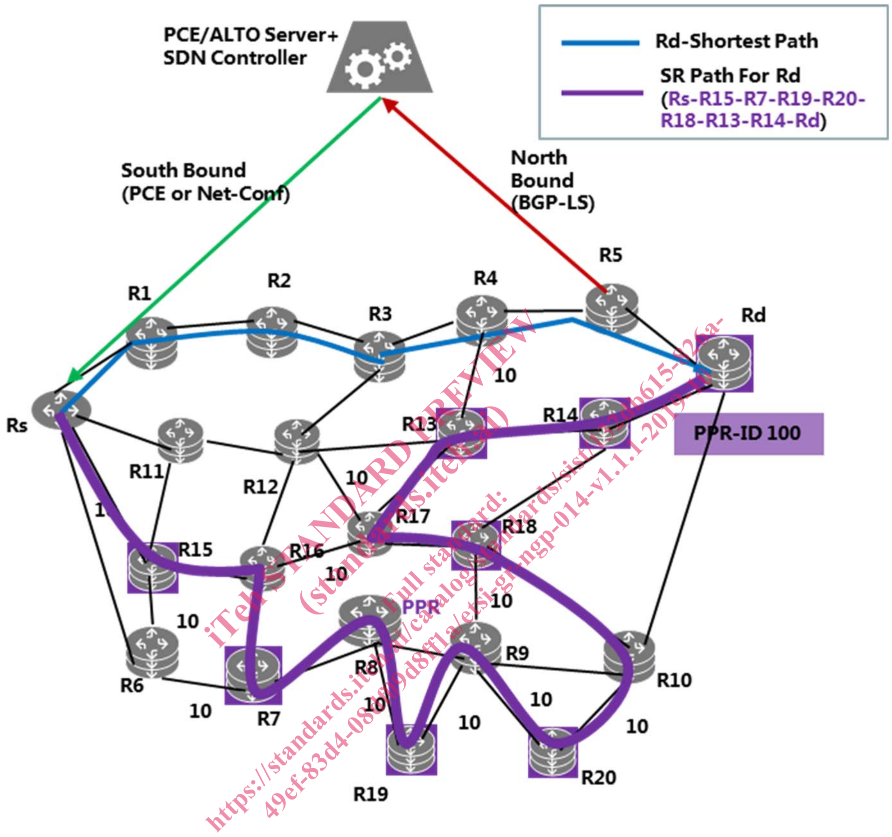

Figure 2: Shortest Path, Source Routed Path

Consider the network depicted earlier in Figure 2 to describe the operation of PPR. Node Rs is an ingress or head-end node, while node Rd serves as egress or as another head-end node. Assume the bi-directional IGP link metric for all the links connecting any two node to be of value 1 except from some links with value 10, as shown explicitly. Rs may be configured to receive TE or explicit source routed path information from a central entity (PCE or Controller). The received path comprises PPR information. A PPR is identified using a PPR-ID, which can also relate to sources that are attached to node Rs: traffic from those sources may have to use a specific PPR-ID. (It is also possible to have a PPR provisioned locally for non-TE needs, e.g. for purposes of FRR or to chain services.) The PPR path information is encoded as an ordered list of PPR-Path Description Elements (PDEs) from source to a destination node in the network. The PPR-PDE information represents both topological and non-topological segments and specifies the actual path towards a Forwarding Equivalence Class (FEC) or Prefix by Rd.

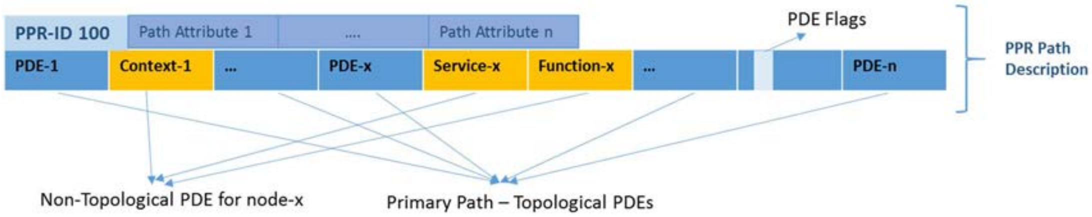

Figure 3: PPR Path Structure

Considering the same example, the SR path Rs-R15-R7-R19-R20-R18-R13-R14-Rd can be attached with a PPR-ID, say 100. Once the path and PPR-ID are signalled in an underlying IGP as a PPR, only nodes that find themselves in the path description have to act on this path. For example, after completing its shortest path computation as usual, R15 finds that its node information is encoded as PPR-PDE in the path. As a result, it adds an entry to its Forwarding Information Base (FIB) with PPR-ID 100 as the incoming label (assuming the data plane type in PPR TLV is MPLS) and sets the NH as the shortest path NH towards the next PPR-PDE (R7), which in this case is the link towards R16. If instead R15 had added a shortest path route entry in the FIB for Rd, it would have added by setting NH as link towards R11 (shortest path metric to reach Rd). This process continues on every node as represented in the PPR path description.

Inter-Area Scenarios:

- The above can be extended inter-area scenarios. The below diagram (Figure 4) represents one such scenario. In this 2 IS-IS levels are used each having separate north bound and south bound communication end points with PCE/SDN controller. It is expected PPR path for each level is computed and given to the ingress nodes Rs and Rd for L1 and L2 respectively. Node Rd is acting as an L1/L2 router and some of the prefixes in L2 are leaked to L1 (including Rb). Now in L1 area, the path advertised by Rs (shown in purple) would be id for prefix Rb, hosted by L1/L2 router Rd. At Rd when it receives the PPR for Rb in L2, the same is advertised in L2 area (in orange), in this example it happened to be a strict path Rd-Ra-Ra15-Ra16-Ra17-Ra13-Ra14-Rb. It is important to note the L1/L2 router functionality of separate path advertisement in respective levels (no leaking of path information).

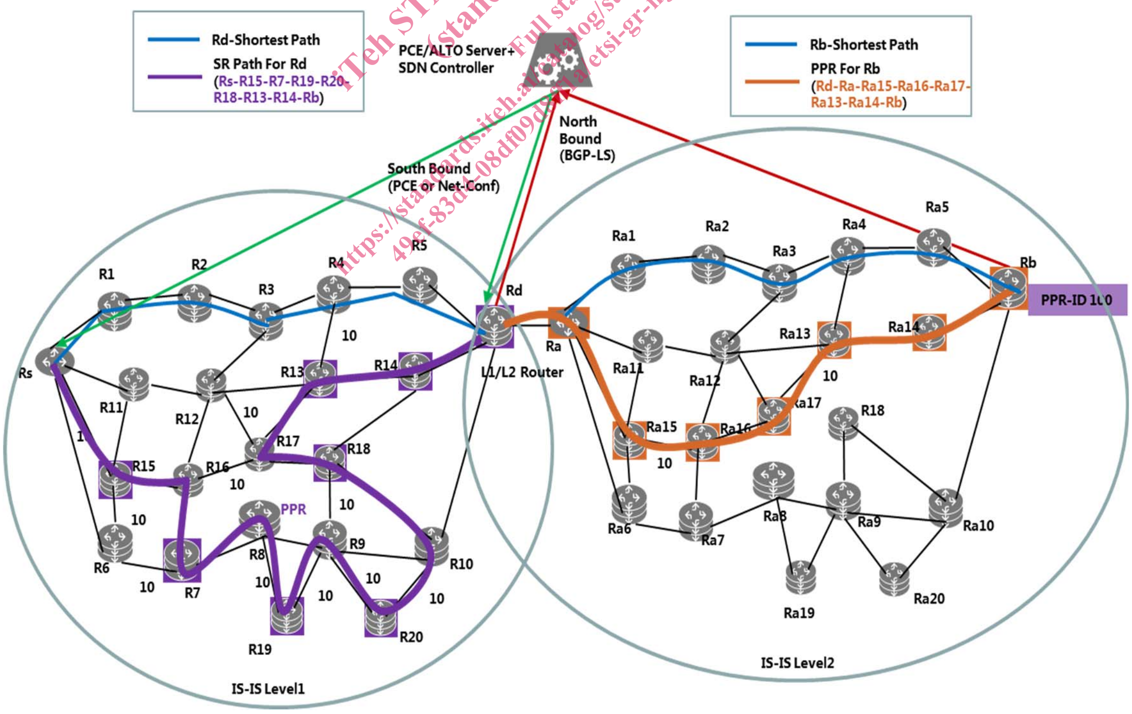

Figure 4: PPR with Inter-Area Scenario (IS-IS Example)