

NSIS Working Group
Internet-Draft
Intended status: Informational
Expires: January 10, 2008

L. Cordeiro
M. Curado
E. Monteiro
V. Bernardo
D. Palma
University of Coimbra
F. Racaru
M. Diaz
C. Chassot
LAAS
July 9, 2007

GIST Extension for Hybrid On-path Off-path Signaling (HyPath)
draft-cordeiro-nsis-hypath-04

Status of this Memo

By submitting this Internet-Draft, each author represents that any applicable patent or other IPR claims of which he or she is aware have been or will be disclosed, and any of which he or she becomes aware will be disclosed, in accordance with Section 6 of BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF), its areas, and its working groups. Note that other groups may also distribute working documents as Internet-Drafts.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

The list of current Internet-Drafts can be accessed at
<http://www.ietf.org/ietf/lid-abstracts.txt>.

The list of Internet-Draft Shadow Directories can be accessed at
<http://www.ietf.org/shadow.html>.

This Internet-Draft will expire on January 10, 2008.

Copyright Notice

Copyright (C) The IETF Trust (2007).

Abstract

In a multi-domain Internet that offers QoS guaranties for

applications, there is the need of signaling among the domain entities that are responsible for the management of QoS. Because different domains have different network protocols and topologies, the HyPath approach uses the NSIS protocol and interactions with the local routing protocols to have an off-path signaling in hybrid environments.

Table of Contents

1.	Introduction	3
1.1.	Terminology and Abbreviations	3
2.	Off-path signaling, state of the art	4
2.1.	Off-path signaling proposals	4
2.1.1.	SIBBS (Simple Inter-domain Bandwidth Broker Protocol)	5
2.1.2.	COPS-SLS	6
3.	HyPath	8
3.1.	HyPath signaling	10
3.2.	Non-NSIS domains	12
3.3.	Usage of external functions	12
3.3.1.	BGP domain scenario	13
3.4.	Heterogeneous solution	16
3.5.	NSIS architecture with HyPath	16
3.5.1.	HyPath on the Resource Manager	17
3.5.2.	HyPath in the Border Router	18
3.5.3.	The HyPath additional information	19
3.6.	HyPath multiple domain example	26
4.	Acknowledgments	28
5.	Security Considerations	28
6.	Conclusion	29
7.	Open issues	29
8.	Normative References	29
	Authors' Addresses	30
	Intellectual Property and Copyright Statements	33

1. Introduction

During past years, we assisted to a common rise of new technologies in the telecommunication and computer science fields. This evolution led to the emergence of new types of applications involving among others, multimedia like VoIP, VoD, tele-engineering, telemedicine, etc. These applications have new constraints and requirements concerning Quality of Service (QoS) parameters such as delay and jitter. Therefore, new services are required besides those given by the actual Internet.

Nowadays, all packets in the Internet receive the same treatment. However, some data flows need special handling in order to satisfy the application requirements, and thus it is necessary to address QoS issues. The internet is an interconnection of networks, comprising different domains, called Autonomous Systems (AS), managed independently, especially in what concerns QoS strategies. In order to support QoS for communications over several domains, intra and inter-domain QoS signaling appears to be inevitable.

Our work aims at a context of a multi-domain Internet that offers QoS guarantees for applications. Inside a domain, the QoS is managed through central entities, that are in charge of installing and handling QoS based on internal rules. This concept was introduced in the DiffServ domains, and is associated with Bandwidth Brokers [1]. Currently, new requirements arise: signaling MUST take place, not only among devices strictly on the data path, but also among these central entities, that we call hereafter Resource Managers (RM).

Several signaling protocols have been proposed, and recently, the IETF NSIS working group [2] has proposed a new signaling architecture. The goal of the NSIS framework is to manipulate the network state related to data flows assuming that the messages will be processed on the nodes which also handle the data flows themselves ("path-coupled signaling"). This document discusses an NSIS multi-domain, multi-service, RM based Internet that allows off-path signaling. The main issue addressed in this document is the interoperability between NSIS and non-NSIS domains.

1.1. Terminology and Abbreviations

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in [3].

The following additional terms are used:

- o e2e: end-to-end
- o QoS: Quality of Service
- o Non-NSIS domain: an administrative domain where only the RM is NSIS aware.
- o NSIS domain: an administrative domain where the RM and, at least, all border routers are NSIS aware.
- o RM: Resource Manager, central entity of a domain in charge of the QoS management.

2. Off-path signaling, state of the art

In the off-path approach, entities participating in the signaling process are not bound to the path followed by data flows. The most common example is when particular entities inside a domain, which have special responsibilities (QoS, policy control, servers, etc.), MUST be signaled. These devices are NOT strictly on the data path; nevertheless the signaling protocol MUST arrive to interact with these devices. Off-path signaling has advantages, as presented in [4] :

- o independence between the signaling plane and the forwarding plane;
- o introduction of flexibility allowing entities such as proxies to be signaled even if they are not on the data path;
- o functioning with new routing protocols or traffic engineering mechanisms (QoS routing, q-BGP, etc.);
- o better adapted for mobility.

On the other hand, off-path signaling MUST answer new challenges such as discovering the next hop and synchronization with IGP (Interior Gateway Protocol) and EGP (Exterior Gateway Protocol) routing protocols.

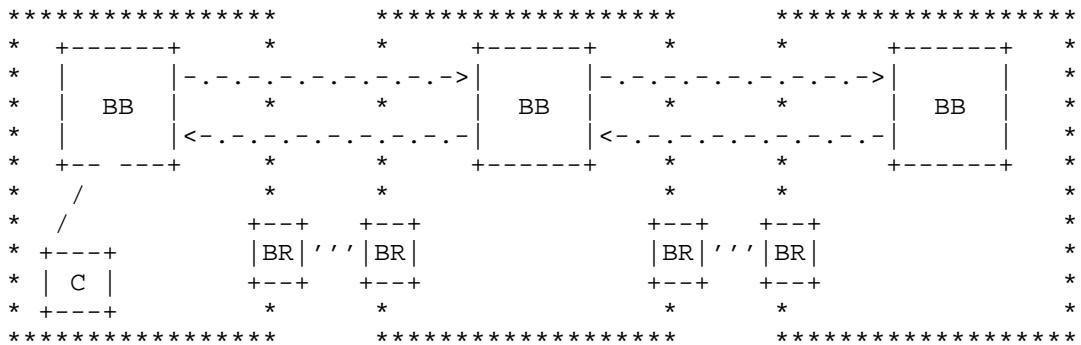
2.1. Off-path signaling proposals

Several protocols have been proposed for the off-path signaling in a bandwidth broker-based multi-domain DiffServ model. The next subsections will present an overview of some of them.

2.1.1. SIBBS (Simple Inter-domain Bandwidth Broker Protocol)

The SIBBS protocol has been defined by the QBone Signaling Workgroup and it aims to be used on DiffServ bandwidth broker-based domains. In the QBone testbed, each network is a differentiated service (DiffServ) domain, supporting one or more globally well known forwarding services, built from fundamental DiffServ blocks. SIBBS, as described in Figure 1, is a very simple protocol to be used between bandwidth brokers. It contains two principal Policy Decision Units (PDUs):

- o RAR (Resource Allocation Request)
- o RAA (Resource Allocation Answer)



<-----> = signaling message between BB
 ----- = message between client and BB
 C = client
 BB = Bandwidth Broker BR = Border Router

Figure 1: SIBBS protocol

The RAR message includes a globally well-known service ID, information related to the QoS request (class of service and bandwidth), a destination IP address, a source IP address, an authentication field, and the other parameters of the service. The sender can be the client host, a BB or a proxy. The RAA message contains the answer to an RAR PDU. The communication between BB is supposed to be reliable, i.e. using TCP.

When receiving an RAR message, a BB:

- o Authenticates if the request is indeed from a peer bandwidth broker;
- o Determines the egress router (interface) from its (inter-domain) routing tables;
- o Checks if the requested resources fall within the SLS;
- o Ensures if there are sufficient resources within the domain to support the flow from the ingress border router;
- o Determines whether the flow may be accepted according to the policies of the domain.

If the required resources are available, the request is propagated recursively through the inter-domain path to the last BB. This last BB returns an RAA message to its immediate upstream BB and the process is continued until the originating BB. This process is concluded with an admission of the QoS request. Resources are confirmed by means of refresh messages, sent periodically.

In order to perform the configuration the BB MUST have access to the border routers. SIBBS does not specify a particular protocol, but may use protocols such as COPS, DIAMETER, SNMP.

2.1.2. COPS-SLS

COPS-SLS [5][6] is an extension of the COPS (Common Open Policy Service) protocol [7] for SLS management in a multi-domain environment. COPS is a client/server protocol designed for the management of policy based networks. The basic model of COPS is presented in Figure 2

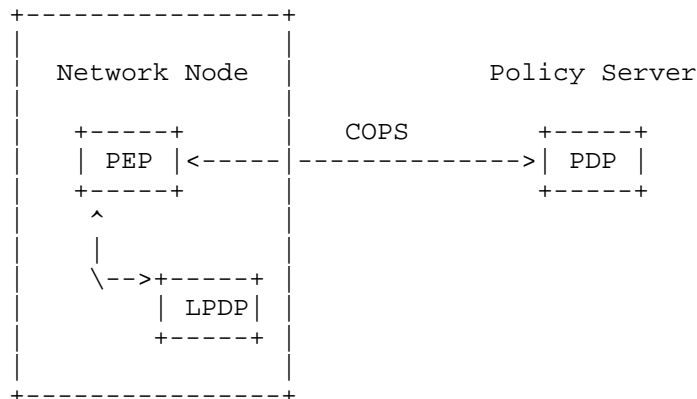


Figure 2: COPS protocol

The Policy Decision Point (PDP) is the central entity in charge of making the decisions (for itself or for other elements of the network). The Policy Enforcement Point (PEP) is where the policies are applied, such as a router. The optional Local Policy Decision Point (LPDP) can be used by the device to make local policy decisions in the absence of a PDP.

COPS is a request/response protocol that allows a PEP (router) to interrogate its PDP about the action to perform once an event has occurred (for instance, if a signaling message arrived).

COPS-PR is an extension of COPS with the goal to force the application of a policy in the PEP without any prior request.

COPS-SLS behaves like SIBBS: a request is propagated from one BB to the other in each domain of the data path. Each BB has a double role:

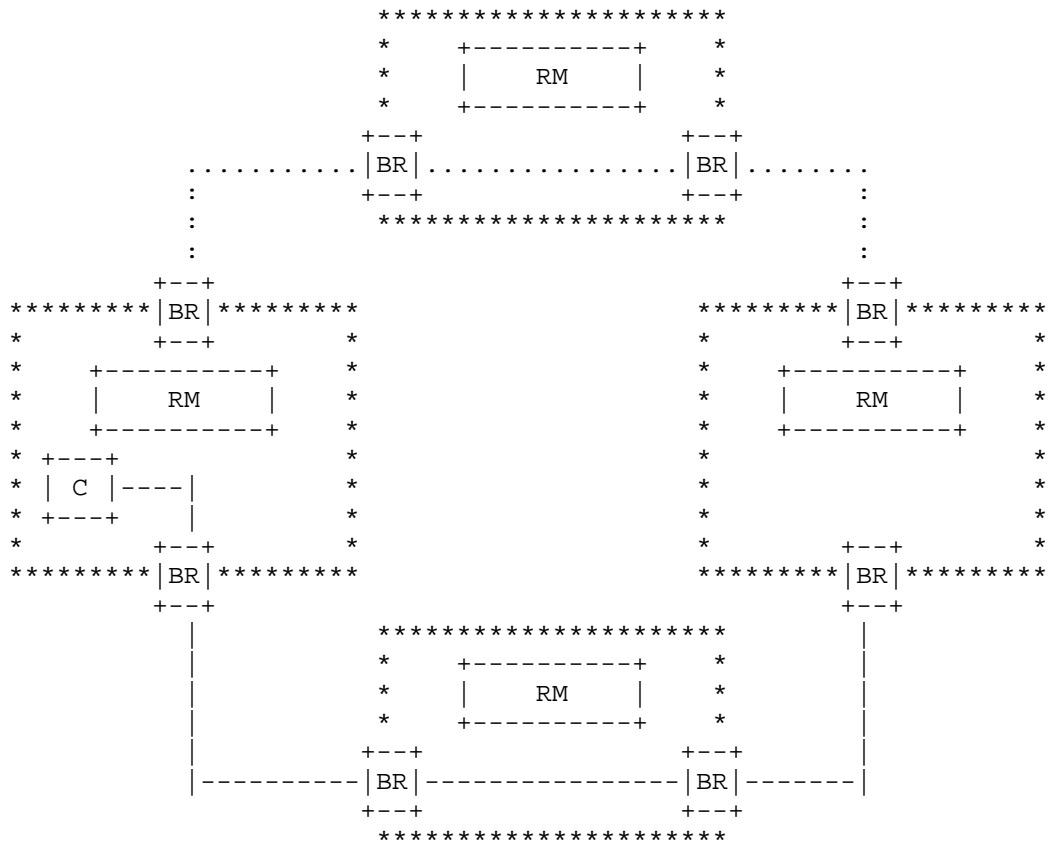
- o PDP for the upstream domain (BB which sends the request) and
- o PEP for the next BB domain.

Compared to SIBBS, COPS-SLS adds some features to the protocol, as renegotiation of classes of service in case of failure of admission control. The communication between the BB and border routers is assured by the COPS-PR protocol. COPS-SLS does not provide any specification on the discovery of the next BB or on the identification of border routers.

3. HyPath

The requirements for an hybrid on-path off-path approach, for end-to-end (e2e) signaling across NSIS and non-NSIS domains, are not fully solved by the NSIS protocol as it is being defined currently in the IETF NSIS working group. There is the need to have network signaling between specific entities in domains (not only the routers in the data path like the normal on-path solution). This is the case of QoS network signaling when resource managers are responsible for the domain QoS. In these situations the entities to be signaled are the RM entities and not only the network elements (routers). Figure 3 shows an example of the normal NSIS signaling from a source RM to a destination RM.

With the usual behavior, the NSIS protocol [2] does not signal the RM servers in the data path, and it does not force the signaling to follow the same path as the user data (because the source and destination are different and the domains can have different routing policies based on local source IP addresses). As presented in Figure 3, the signaling message may not follow the same inter-domain path from the sender domain to the receiver domain. If this situation occurs, the resource reservation will not be properly done on the data path.



----- signaling started by the client
..... signaling started by the RM
C = client, RM = Resource Manager, BR = Border Router

Figure 3: Normal NSIS signaling

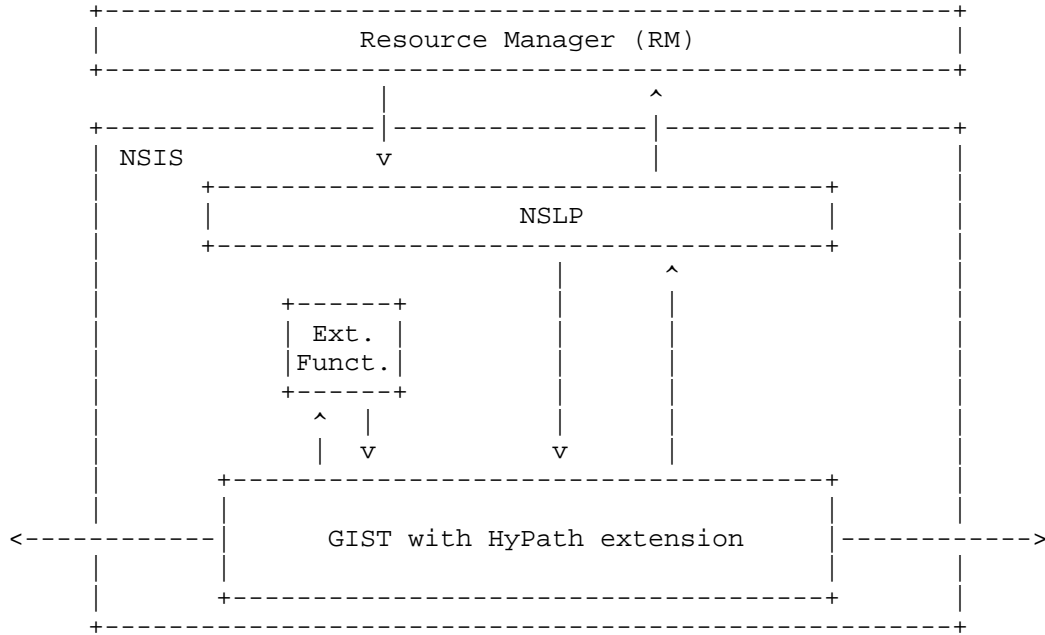
The major requirements to achieve e2e network signaling are the following:

- o signaling messages MUST follow the same path as the user data;
- o all the RMs in the data path MUST be signaled.

The NSIS protocol as it is being defined in the IETF, cannot solve these two major requirements simultaneously. Therefore, GIST needs to be adapted in order to fulfill the above requirements, either by creating a new Message Routing (MRM) or a new TLV Object. The

proposal for the extension of the NTLF layer is named GIST Extension for Hybrid On-path Off-Path Signaling (HyPath).

With this approach the communication between layers, NSLP [2] and NTLF [8], works without changing their specifications. Figure 4 describes an example of the NSIS framework architecture using the HyPath.



HyPath extension on the NSIS architecture
 Ext. Funct. = External Functions (eg. BGP processing)

Figure 4: NSIS architecture with HyPath

In order to fulfill these e2e requirements the GIST processing MUST include the processing of this new proposal.

3.1. HyPath signaling

HyPath signaling provides a solution for the hybrid on-path off-path signaling. To better understand the new signaling, Figure 5 illustrates the NSIS protocol with HyPath.

reached and the signaling stops in the RM server.

With this architecture all the requirements to achieve e2e network signaling are met and no changes are needed in the definitions of the NSLP. The changes needed on the GIST layer are described later in section Section 3.5.

3.2. Non-NSIS domains

The drawback of the approach described in Section 3.1 is that all border routers of all domains MUST be NSIS aware and HyPath compliant. Even though in theory this is a reasonable assumption, in practice we cannot guarantee this scenario. For that reason, we define an heterogeneous solution (described in more detail in Section 3.4) that works when border routers are not NSIS aware (non-NSIS domains) and the only information available is provided by the routing protocol of the domain.

Since it is not possible to count on the NSIS interception in the border router, the solution has to rely on the routing protocol.

In non-NSIS domains, when the RM intends to send a signaling message, HyPath uses an external function (described in Section 3.3) to discover the local egress border router on the data path and the next RM IP address. With this information, an NSIS message with the NSLP payload and the HyPath additional information (described in Section 3.5) is sent directly to the RM of the next domain.

Using again the external function to discover the local egress border router on the data path and the next RM IP address, the NSIS signaling message is sent to the RM of the next domain. The procedure described is repeated until the last domain is reached.

In this approach, the signaling messages pass through all the RMs in the data path, even if they are not on the data path.

The disadvantage of this approach is the extensive usage of the external functions. Since these functions are used in all non-NSIS domains, this approach would have an impact on the processing time and on the amount of resources used. The information about the local egress border router is obtained from the normal AS association procedure.

3.3. Usage of external functions

The motivation for off-path signaling was described in the beginning of this document. If the signaling is decoupled from the data path (but still path-related signaling), two general problems need to be

solved:

- o The RM MUST discover the ingress and egress points through which the data path will pass in its domain; this information is needed in order to continue the NSIS signaling and to perform an admission control between the ingress and the egress border routers and on the inter-domain link;
- o In non-NSIS domains, the Resource Manager of the next domain MUST be identified in order to propagate the request.

For these reasons HyPath needs to use external functions to gather the required information.

The interface between HyPath and the external functions MUST be as described next:

- o GetEgressBRAddress(Source IP Address, Destination IP Address)
 - * This interface is responsible for requesting the IP address of the data egress border router of the first domain. The interface provides the data source and destination IP addresses and returns the egress border router IP address;
- o GetNextRMAddress(Source IP Address, Destination IP Address)
 - * This interface is responsible for requesting the next RM IP address in non-NSIS scenarios. The interface provides the data source and destination IP addresses and returns the RM IP address;
- o IsNextDomainNSIS(Source IP Address, Destination IP Address)
 - * This interface is responsible for requesting the type of the next domain. The interface provides the data source and destination IP addresses and returns the next domain type (NSIS or non-NSIS);

The external functions are outside the purpose of this specification, but an example of a BGP domain scenario is described next.

3.3.1. BGP domain scenario

In a BGP domain the HyPath external function must interact with the BGP protocol for the local and inter domain routing.

In this scenario the RM has access to the routing tables of the border routers of its domain, and is able to question the BGP tables. This

interrogation is implemented as a request/response protocol via telnet or ssh. The main information in the BGP routing table after rejecting an unacceptable route is:

- o Accessible destination network list (IP prefixes);
- o For each prefix:
 - * next router address (next-hop) in the adjacent domain; this information is carried in the messages inside the AS (i- BGP session);
 - * list of Autonomous Systems successively traversed (AS path), from adjacent domains to the AS destination domain for the destination network;
- o For each border router: address of neighbor routers with whom it has established BGP session (neighbor) which are either border routers or Router Reflectors [9].

The approach to discover ingress and egress border routers is the following:

- o Discover the ingress border router;
- o Discover the egress border router.

The discovery process of the ingress border router depends on the type of the actual and upstream domains, as described next.

If the actual and the upstream domains are NSIS domains, the ingress router is easy to retrieve. It is the border router that intercepts the NSIS message and redirects it to the RM.

If the actual domain is an NSIS domain and the upstream domain is a non-NSIS Domain, the ingress router is retrieved from the message received by the RM. In this case, the upstream RM sends the message directly to the local RM as explained in Section 3.5.1. The upstream RM interrogates the BGP table of its ingress border router and retrieves the address of the next domain.

Relying on the intra-domain routing information (topology, traffic engineering) and BGP interactions, the RM of the AS2 retrieves the ingress border router in the next domain. In the upstream AS2 domain the BGP R21 tables contain:

Network	Next-Hop	Path
*>ip form AS3	R31	i AS3

The RM of the AS2 interrogates this table and retrieves the next-hop address of R31. This operation is done only with adjacent non-NSIS domains. Then, it inserts this address in the signaling message to be sent to the RM of AS3. A particular case is when the IP address of the next-hop is not distributed through the internal routing protocol (for instance, if a private IP is used), and the new request MUST be addressed to the egress border router.

When this RM receives the message, it already has the IP address of the ingress border router. If this address is not the loopback address (a private IP address, for instance), it can obtain the loopback address *b* from local BGP and topology configuration.

Dealing with a non-NSIS domain, is a similar procedure to the one presented when dealing with an NSIS domain.

All border routers communicate in the i-BGP session in order to discover the egress border router inside an AS. The egress border router is discovered using the BGP routing table of the ingress border router. If we deal with a full mesh iBGP (all border routers are connected on iBGP), then the egress border router is a neighbor. As an alternative, if the domain uses Route Reflectors, either the attribute `ORIGINATOR_ID`, or the domain topology can be used to find the egress border router for the data path.

For non-transit traffic (i.e. traffic originating inside the domain) the Resource Manager can use a database (similar to TED for the PCE Element [10][11]) where the cartography of the domain (network topology) is stored.

The RM may obtain the IP address of a peer RM in an adjacent domain through the Service Level Agreement (SLA) between two adjacent domains which contains the loopback address of the RM present in the domain. As the administrator is aware of the SLAs, it can configure the RM with all peer RM addresses.

Another solution follows the SIBBS proposal that suggests to retrieve the Bandwidth Broker address via a DNS mechanism (the BB for each domain is to be named `bb. <domain_name >` and put it in a CNAME record in the DNS). Instead of using the domain name, we propose to associate the AS number to an RM IP address.

When an RM needs to obtain the next RM IP address, it checks the BGP table to find the AS path to the destination. In the AS path, it finds the next AS number and, based on one of the mechanisms presented before, it does the correspondence AS number <-> RM IP address.

In this section we described some mechanisms that allow for the solution of major issues of the off-path signaling path, namely to discover the next hop to signal and how to interact with external routing protocols such as BGP.

3.4. Heterogeneous solution

The presented NSIS and non-NSIS solutions are able to fulfill the objectives, but as discussed in the previous sections, they have some disadvantages. On one hand, NSIS domains need to use HyPath in the border routers. On the other hand, in non-NSIS domains an intensive usage of external functions that extensively access the routing protocol is need.

A new solution is to integrate the two approaches presented. Particularly, in this new scheme, in NSIS domains it is used the NSIS solution and in non-NSIS domains it is used the non-NSIS solution.

The hybrid approach raises a problem when there is interaction between domains that have different solutions implemented (between NSIS and non-NSIS domains). If the NSIS approach is used when the border router sends the message to the destination (through the data path) the signaling message is never intercepted by the next domain. Therefore, the next domain RM is never signaled.

To solve this problem, the NSIS domain MUST check the type of the next domain before sending any signaling message. This information is obtained from the normal AS association procedure. If the next domain is a non-NSIS domain, the message MUST be sent as described in Section 3.1, otherwise it is sent as described in Section 3.2.

This approach implies that NSIS domains connected with non-NSIS domains need to determine the type of the next domain. This procedure might increase somehow the complexity of the solution. If an NSIS domain is only connected to other NSIS domains the solution is much more simple and light weight.

3.5. NSIS architecture with HyPath

As described in the previous sections, with the usage of HyPath for the hybrid on-path off-path approach for e2e signaling across NSIS and non-NSIS domains requires that the RMs and the routers, especially the border routers, support HyPath.

The main HyPath functionalities are the following:

- o In the Resource Manager
 - * Discovery of the egress border router of the first domain
 - * Discovery of the ingress border router after a non-NSIS domain
 - * RM signaling
 - * Message reception and decoding
 - * Sending messages
- o In the egress border router
 - * Start RM signaling in the first domain
- o In the ingress border router
 - * Message interception and sending them to the local RM
 - * Reception of the local RM response message and continuation of RM signaling

These functionalities are described in more detail in the next subsections.

3.5.1. HyPath on the Resource Manager

The HyPath in the RMs is responsible for changing the destination address of the signaling message in order to assure that the right RM is signaled.

In the first domain (the domain where the network signaling starts), HyPath discovers the egress border router on the data path using an external function. If the next domain (discovered using the external function) is an NSIS domain, the message is sent to the egress border router. Otherwise, the ingress border router and the IP address of the RM of the next domain in the data path MUST be discovered using again an external function. Afterwards, the message is sent directly to the IP address of the next domain RM.

If a domain is not the first domain, it means that the NSIS message has already been received, and that the ingress border router information is included in the HyPath. If the next domain and the current domain are NSIS domains, the message is sent to the ingress border router (IP address in the MRM) to be forwarded through the

same path as the data. If the next domain is a non-NSIS domain, then again, an external function MUST be used to discover the ingress border router and the IP address of the RM of the next domain in the data path. Afterwards, the message is sent directly to the IP address of the next domain RM.

The messages to be sent upstream use the normal GIST state to allow the usage of already established states and associations.

In the RM, HyPath requires an RM state table to store the previous RM IP address when the previous or the local domains are non-NSIS domains. In all other cases HyPath does not require any state.

3.5.2. HyPath in the Border Router

HyPath in the border router has two different functionalities depending if it is an egress or ingress border router. In the first domain, the border router acts as the egress router where the signaling merges with the data path. From this point forward, if the message is always sent to the end user, the signaling path will follow the same path as the data path. In the other domains, the border router acts as an ingress border router where NSIS messages are intercepted.

In the border router, if the messages are received from the local RM (HyPath messages), they are forwarded to the end user. If the messages are intercepted, they are forwarded to the local RM.

If the received message is from the local RM and the current network is neither the source nor the destination network, the message direction MUST be set with the original direction field of the HyPath additional information. The Border Router Address field of the HyPath MUST be set with the address of the current machine and the border router flag MUST be set.

If the received message is to be sent to the final destination and if the border router flag is set, the Border Router Address field in the HyPath additional information MUST be checked. If the Border Router Address belongs to the local domain, the message is forwarded, only updating the Border Router Address field. If the Border Router Address is not from the local domain the message MUST be sent to the local RM. In this case the source address MUST be set to the BR address, the destination address MUST be set to the local RM address and the Border Router Address field MUST be updated with the local BR address.

3.5.3. The HyPath additional information

Taking into account the requirements of HyPath, additional information MUST be included in the GIST messages. There are two possible approaches to fulfill these requirements:

- o MRM solution;
- o TLV solution.

These two approaches are described in the following sub-sections.

3.5.3.1. MRM solution

The current MRMs [8] do not solve the problems describe in the previous sections.

There are two possible approaches to adapt GIST with the previously described functionalities through the MRMs:

- o A new MRM;
- o A Path-Coupled MRM extension.

The creation of a new MRM is one of the approaches to have an hybrid on-path off-path approach for e2e signaling with NSIS. The creation of a new GIST MRM adapts the GIST signaling with the described HyPath functionalities in an independent way of the others MRMs.

With this approach, the HyPath MRM takes the following format:

```

0                                     1                                     2                                     3
0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
                                     +-----+-----+-----+-----+
                                     |IP-Ver |   Type   |D|O|B|Res|
+-----+-----+-----+-----+
//                                     Source Address                                     //
+-----+-----+-----+-----+
//                                     Destination Address                               //
+-----+-----+-----+-----+
//                                     Original Source Address                         //
+-----+-----+-----+-----+
//                                     Original Destination Address                   //
+-----+-----+-----+-----+
:                                     Border Router Address                           :
+-----+-----+-----+-----+

```

Type: type of HyPath message.

- 1 - Standard message: a message that follows the data path and it is intercepted in the border routers
- 2 - Border router message: a message that it is sent from the border router to the RM
- 3 - Resource Manager message: a message that is sent either to the border router of the actual domain or to the RM of the next domain, depending on the type of the next domain (NSIS or non-NSIS)

Flag D: direction flag

Flag O: original direction flag

Flag B: means that the Border Router Address field is present

Source Address: data source address

Destination Address: data destination address

Original Source Address: original data source address

Original Destination Address: original data destination

Border Router Address: border router address to where the Resource Manager MUST send the next message

This new MRM needs to be processed by GIST in a different way from the other MRMs. The Source and Destination Addresses are used by GIST to send the messages, but unlike most of the others MRMs these

two addresses are changed to perform the off-path signaling. The Original Source and Original Destination addresses are the data flow addresses that are kept unchanged. Taking into account the HyPath architecture, the Border Router Address is used to keep the border router address for the off-path signaling and to be used in the Source Address and Destination Address fields to perform the off-path signaling.

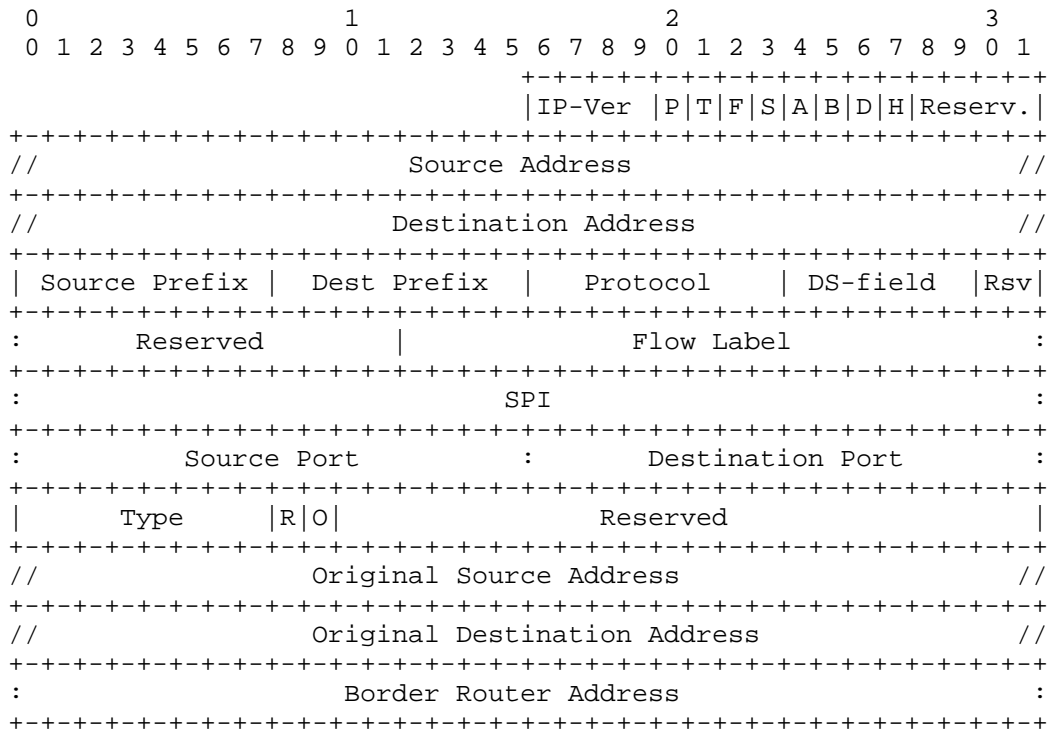
The HyPath MRM is contains also three flags that can influence the GIST message processing. The D flag, like in other MRMs, indicates to GIST the direction of the message. The O flag, is used to store the D flag value whenever off-path is used in the signaling. If the B flag is set, HyPath message processing knows that the Border Router Address is present, otherwise this field is not checked.

The other approach for HyPath is to extend the Path-Coupled MRM [8] with the new characteristics required to make an hybrid on-path off-path e2e signaling.

By adding additional fields to the current Path-Coupled MRM it is possible to maintain the usual path-coupled functionalities and, when required, have the additional HyPath functionalities. This can be achieved by adding a new flag, H Flag, in the reserved area of the MRM (after the already specified D Flag). NSIS entities that are configured to process HyPath messages (usually RM and border routers) can search for this H Flag in the Path-Coupled MRM, while other entities can simply ignore it and process the message normally (ignoring all extra HyPath fields in the MRM).

When the H Flag is present, new fields MUST be included on the MRM. These fields consist of a message type, followed by a flag stating if the border router address field exists, a flag to depict the message real direction and a reserved field for future use. Next, there are two fields containing the source and destination addresses of the message. Finally, there is an optional field containing the address of the border router of the domain.

The Path-Coupled MRM with the HyPath extension to support hybrid on-path off-path signaling takes the following format.



Type: type of HyPath message.

- 1 - Standard message: a message that follows the data path and it is intercepted in the border routers
- 2 - Border router message: a message that it is sent from the border router to the RM
- 3 - Resource Manager message: a message that is sent either to the border router of the actual domain or to the RM of the next domain, depending on the type of the next domain (NSIS or non-NSIS)

Flag H: means that the HyPath information is present
 Flag R: means that the Border Router Address field is present
 Flag O: original direction flag

Original Source Address: original data source address
 Original Destination Address: original data destination
 Border Router Address: border router address to where the resource

manager MUST send the next message

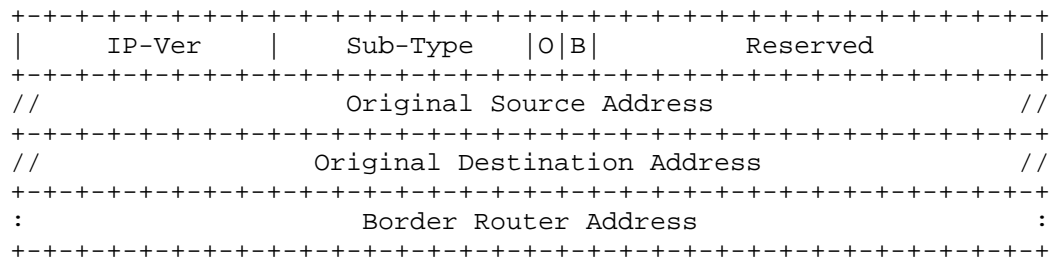
3.5.3.2. TLV solution

The HyPath Payload Object is a new TLV that allows to store the previously referred data, instead of creating a new MRM or changing the existing one.

This approach will ensure that all data is understood by GIST and that there is no way that important off-path information is discarded by HyPath non-compliant peers. A description of this object is presented below.

Type: TBD by IANA

Length: Variable



The A/B extensibility flags follow the definition in the GIST specification and MUST be used as Forward ('10').

This allows the new TLV to be forwarded by HyPath unaware peers. The GIST processing remains the same so there's no concern about non-compliant HyPath peers that could reject unidentified data.

Sub-Type: type of HyPath message.

- 1 - Standard message: a message that follows the data path and it is intercepted in the border routers
- 2 - Border router message: a message that it is sent from the border router to the RM
- 3 - Resource Manager message: a message that is sent either to the border router of the actual domain or to the RM of the next domain, depending on the type of the next domain (NSIS or non-NSIS)

Flag O: original direction flag defined like the D flag on the GIST MRM

Flag B: means that the Border Router Address field is present

Original Source Address: original data source address

Original Destination Address: original data destination

Border Router Address: border router address to where the resource manager MUST send the next message

This new TLV Object needs to be processed by GIST peers that support HyPath. The Source and Destination Addresses of the MRM are changed to perform the off-path signaling. The Original Source and Original Destination addresses are the data flow addresses that are kept unchanged and stored on this object. Taking into account the HyPath architecture, the Border Router Address is used to keep the border router address for the off-path signaling and to be used in the Source Address and Destination Address fields to perform the off-path signaling.

The HyPath TLV is also composed by two flags that can influence the GIST message processing. The O flag is used to store the MRM D flag value whenever off-path is used in the signaling. If the B flag is set, HyPath message processing knows that the Border Router Address is present, otherwise this field is not checked.

Any other HyPath specific processing remains as it would be, with the MRM approaches.

3.5.3.3. Conclusion

The two MRM solutions proposed previously (the new MRM and the Path-Coupled MRM extension), are able to provide an hybrid on-path off-path approach for e2e signaling. Both of these approaches allow the communication between the two NSIS layers, NSLP [2] and NTLP [8], without changing their specifications. The GIST changes to allow these solutions are few and SHOULD NOT influence the other functionalities.

With the new HyPath MRM approach there is no dependency of other MRMs and there is no need to change their processing in GIST. On the other hand, this solution is only feasible if all NSIS Entities support HyPath MRM.

With the Path-Coupled MRM Extension approach, the NSIS entities that

Comparing with the MRM solutions, the TLV Object has only a single word overhead on the transmitted data, since it has its own header.

On the other hand it gives flexibility to make future updates on the necessary fields, and it SHOULD represent less changes on the existing GIST implementations.

3.6. HyPath multiple domain example

To better understand how HyPath works in a multiple domain situation, a three domains example is described next. This example is composed by three networks (A, B and C), three RMs (RM A, RM B and RM C), four border routers (BR A1, BR B1, BR B2 and BR C1) and two users (S, source user, and D, destination user).

Network A is composed by the source user, S, RM A and BR A1. Network B is composed by RM B and two BRs, BR B1 and BR B2. Network C is composed by the destination user, D, RM C and BR C1. Network A connects to network B through BR A1 and BR B1 and network B connects to network C through BR B2 and BR C1. The connection between the network A and network C MUST be through the network B. The data path between the source user and the destination user is the following sequence: S -> BR A1 -> BR B1 -> BR B2 -> BR C1 -> D.

Figure 11 illustrates the three domains network described with the signaling messages required to make a signaling between the source and the destination users. In this case the source user, S, requests the local RM, using one signaling protocol (independent from HyPath) that initiates HyPath signaling to the destination user.

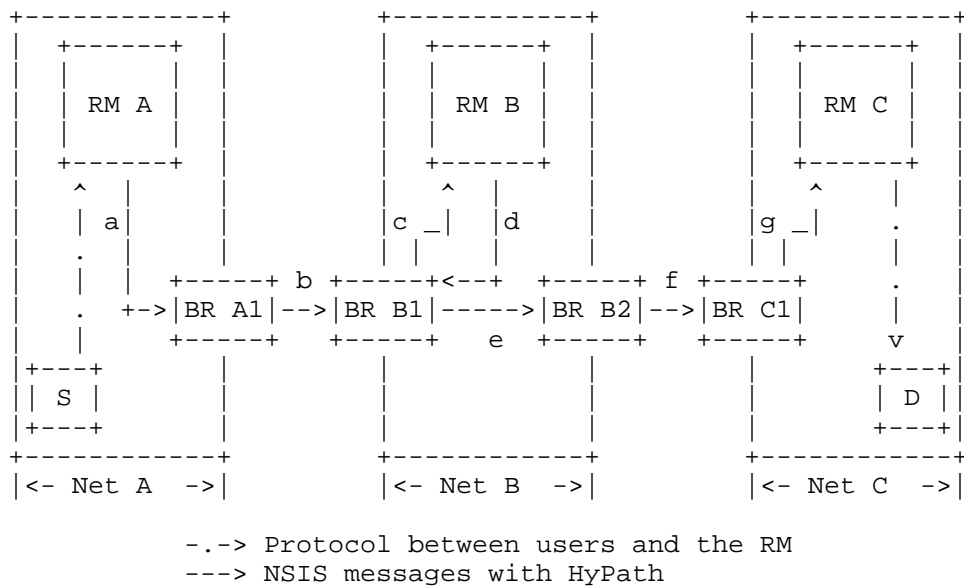


Figure 11: HyPath architecture in the border routers

HyPath signaling starts in the RM A, where the NSLP requests HyPath to send a message from S to D. Before starting the signaling, the local data path egress border router needs to be discovered (using an external function). With this information, the RM A sends the message to the respective egress border router, BR A1.

When the BR A1 receives the message, knows it MUST be injected in the data path. BR A1 changes the received MRM with the original information available there (source IP address, destination IP address and direction) and sends the message towards the destination. This is the first message that follows the data path, like as if it had been generated by the source user.

In the network B (the next domain) the message is intercepted by the ingress border router (BR B1). This border router checks if the message needs to be sent to the local RM or not. Since the message was received from another domain (by checking the BR Address field) the message is then forwarded to the local RM, the RM B. The MRM information is changed in order to send the message to the local RM. Moreover, the source is changed to the BR IP address and the destination is changed to the local RM IP address. The direction of the message is also set to downstream.

In the RM B, when HyPath receives a message, sends it to the

respective NSLP. If NSLP sends a response message, it is sent back to the respective BR, the BR B1. Before sending the message, the destination IP address MUST be changed with BR IP address information. One message that has no border router field set (in case it is the first message) is treated as a new signal. With the BR B1 IP address, the message is sent upstream (using the already established states) to BR B1.

When a message is received in BR B1 from the RM B (because of the message type) the message is injected in the network like in BR A1. The message is then sent to the destination.

In network B the message is intercepted by the egress border router, the BR B2. In this BR, the message does not need to be sent to the local RM because the previous node is a local node (BR B1). This means that no processing is needed and the message is forwarded to the destination. Only the BR IP address MUST be updated.

In network C the message is again intercepted by the ingress border router, BR C1. The ingress border routing procedure is the same as described above, and the message is sent to the local RM (RM C).

In the RM C the node is treated as the last node because the destination user is in the RM network. Here, the HyPath signaling stops and other type of signaling (protocol independent from the HyPath like the Session Initiation Protocol [13]) can interact with the destination user.

When the NSLP requests a response message to an HyPath message in the RMs, the procedure is similar to the one described previously. The message SHOULD be sent upstream and, if new associations are required, the GIST layer SHOULD support their creation upstream.

4. Acknowledgments

The authors would like to thank Luis Conceicao for his contribution on the implementation and testing of GIST and HyPath.

5. Security Considerations

This section describes the security considerations related to the HyPath and this will be discussed and clarified later.

6. Conclusion

This draft presented an architecture in the context of an NSIS multi domain Internet that aims at off-path signaling when an hybrid solution is required (for instance if NSIS is not implemented in all domains).

Currently, an increasing number of applications claim special treatment for their packets in order to satisfy new requirements in terms of delay, loss, jitter, etc. Inside an AS, the QoS management is often delegated to a central entity which has a global view of network topology. This entity is also aware of QoS availability inside and on the inter-domain links of the domain. In order to signal these entities, which are not on the data-path, this draft proposes a solution called Hybrid Path, to involve the central entities on the signaling in the NSIS context.

Two solutions for hybrid on-path off-path signaling are presented in this draft. One of these solutions proposes the creation of a new MRM for GIST or an extension to the existing Path-Coupled MRM. The second solution presents a new TLV Object that allows the usage of HyPath with no changes on the already defined objects. Both SHOULD have a minimal impact on the NSIS architecture. A choice between these solutions MUST be made.

This work, an Hybrid on-path off-path approach for e2e signaling across NSIS and non-NSIS domains, aims specifically at the EuQoS project (<http://www.euqos.eu>), but also for all network signaling that needs to signal specific entities in all domains in the data path.

7. Open issues

This section describes the open issues related to the HyPath and this will be discussed and clarified later.

8. Normative References

- [1] Blake, S., Black, D., Carlson, M., Davies, E., Wang, Z., and W. Weiss, "An Architecture for Differentiated Services", RFC 2475, December 1998.
- [2] Hancock, R., Karagiannis, G., Loughney, J., and S. Van den Bosch, "Next Steps in Signaling (NSIS): Framework", RFC 4080, June 2005.

- [3] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.
- [4] Hancock, R., "A Problem Statement for Partly-Decoupled Signalling in NSIS", draft-hancock-nsis-pds-problem-04 (work in progress), October 2006.
- [5] Trang, N., Boukhatem, N., Doudane, Y., and G. Pujolle, "COPS-SLS - a service level negotiation protocol for the internet", Editor IEEE Communication Magazine, vol. 40, no5, May 2002.
- [6] Trang, N., Boukhatem, N., Doudane, Y., and G. Pujolle, "COPS-SLS usage for dynamic policy-based QoS management over heterogeneous IP networks", Editor IEEE Communication Magazine, vol. 17, no3, May 2003.
- [7] Durham, D., Boyle, J., Cohen, R., Herzog, S., Rajan, R., and A. Sastry, "The COPS (Common Open Policy Service) Protocol", RFC 2748, January 2000.
- [8] Schulzrinne, H. and R. Hancock, "GIST: General Internet Signalling Transport", draft-ietf-nsis-ntlp-13 (work in progress), April 2007.
- [9] Bates, T., Chandra, R., and E. Chen, "BGP Route Reflection - An Alternative to Full Mesh IBGP", RFC 2796, April 2000.
- [10] "Path Computation Element (PCE) Charter", URL <http://www.ietf.org/html.charters/pce-charter.html>, 2005.
- [11] Farrel, A., "A Path Computation Element (PCE) Based Architecture", draft-ietf-pce-architecture-05 (work in progress), April 2006.
- [12] Manner, J., "NSLP for Quality-of-Service Signaling", draft-ietf-nsis-qos-nslp-14 (work in progress), June 2007.
- [13] Rosenberg, J., Schulzrinne, H., Camarillo, G., Johnston, A., Peterson, J., Sparks, R., Handley, M., and E. Schooler, "SIP: Session Initiation Protocol", RFC 3261, June 2002.

Authors' Addresses

Luis Cordeiro
University of Coimbra
Polo II - Pinhal de Marrocos
Coimbra 3030-290
Portugal

Email: cordeiro@dei.uc.pt

Marilia Curado
University of Coimbra
Polo II - Pinhal de Marrocos
Coimbra 3030-290
Portugal

Email: marilia@dei.uc.pt

Edmundo Monteiro
University of Coimbra
Polo II - Pinhal de Marrocos
Coimbra 3030-290
Portugal

Email: edmundo@dei.uc.pt

Vitor Bernardo
University of Coimbra
Polo II - Pinhal de Marrocos
Coimbra 3030-290
Portugal

Email: vmbern@student.dei.uc.pt

David Palma
University of Coimbra
Polo II - Pinhal de Marrocos
Coimbra 3030-290
Portugal

Email: palma@student.dei.uc.pt

Florin Racaru
LAAS
Avenue du Colonel Roche
Toulouse 31077
France

Email: fracaru@laas.fr

Michel Diaz
LAAS
Avenue du Colonel Roche
Toulouse, 31077
France

Phone:
Fax:
Email: diaz@laas.fr
URI:

Christophe Chassot
LAAS
Avenue du Colonel Roche
Toulouse, 31077
France

Phone:
Fax:
Email: chassot@laas.fr
URI:

Full Copyright Statement

Copyright (C) The IETF Trust (2007).

This document is subject to the rights, licenses and restrictions contained in BCP 78, and except as set forth therein, the authors retain all their rights.

This document and the information contained herein are provided on an "AS IS" basis and THE CONTRIBUTOR, THE ORGANIZATION HE/SHE REPRESENTS OR IS SPONSORED BY (IF ANY), THE INTERNET SOCIETY, THE IETF TRUST AND THE INTERNET ENGINEERING TASK FORCE DISCLAIM ALL WARRANTIES, EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO ANY WARRANTY THAT THE USE OF THE INFORMATION HEREIN WILL NOT INFRINGE ANY RIGHTS OR ANY IMPLIED WARRANTIES OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE.

Intellectual Property

The IETF takes no position regarding the validity or scope of any Intellectual Property Rights or other rights that might be claimed to pertain to the implementation or use of the technology described in this document or the extent to which any license under such rights might or might not be available; nor does it represent that it has made any independent effort to identify any such rights. Information on the procedures with respect to rights in RFC documents can be found in BCP 78 and BCP 79.

Copies of IPR disclosures made to the IETF Secretariat and any assurances of licenses to be made available, or the result of an attempt made to obtain a general license or permission for the use of such proprietary rights by implementers or users of this specification can be obtained from the IETF on-line IPR repository at <http://www.ietf.org/ipr>.

The IETF invites any interested party to bring to its attention any copyrights, patents or patent applications, or other proprietary rights that may cover technology that may be required to implement this standard. Please address the information to the IETF at ietf-ipr@ietf.org.

Acknowledgment

Funding for the RFC Editor function is provided by the IETF Administrative Support Activity (IASA).

