scone
Internet-Draft
Intended status: Standards Track
Expires: 30 October 2025

D. Wing
Cloud Software Group
T. Reddy
Nokia
S. Rajagopalan
Cloud Software Group
G. Mishra
Verizon Inc
M. Amend
Deutsche Telekom
L. Contreras
Telefonica
28 April 2025

Discovery of Network Rate-Limit Policies (NRLPs)
draft-brw-scone-rate-policy-discovery-03

Abstract

This document specifies mechanims for hosts to dynamically discover Network Rate-Limit Policies (NRLPs). This information is then passed to applications that might adjust their behaviors accordingly.

Networks already support mechanisms to advertize a set of network properties to hosts (e.g., link MTU (RFC 4861) and PREFIX64 (RFC 8781)). This document complements these tools and specifies a Neighbor Discovery option to be used in Router Advertisements (RAs) to communicate NRLPs to hosts. For address family parity, a new DHCP option is also defined. The document also discusses how Provisioning Domains (PvD) can be used to notify hosts with NRLPs.

Table of Contents

1. Introduction

To optimally deliver connectivity services via a network attachment, networks advertize a set of network properties [RFC9473] to connected hosts such as:

Link Maximum Transmission Unit (MTU): For example, the 3GPP [TS-23.501] specifies that "the link MTU size for IPv4 is sent to the UE by including it in the PCO (see TS 24.501). The link MTU size for IPv6 is sent to the UE by including it in the IPv6 Router Advertisement message (see RFC 4861)".

Section 2.10 of [RFC7066] indicates that a cellular host should honor the MTU option in the Router Advertisement (Section 4.6.4 of [RFC4861]) given that the 3GPP system architecture uses extensive tunneling in its packet core network below the 3GPP link, and this may lead to packet fragmentation issues.

MTU is cited as an example of path properties in Section 4 of [RFC9473].

Prefixes of Network Address and Protocol Translation from IPv6 clients to IPv4 servers (NAT64) [RFC8781]: This option is useful to enable local DNSSEC validation, support networks with no DNS64, support IPv4 address literals on an IPv6-only host, etc.

NAT is cited as an example of path properties (see "Service Function" bullet in Section 4 of [RFC9473]).

Traffic exchanged over a network may be subject to rate-limit policies for various operational reasons. [I-D.brw-scone-throughput-advice-blob] specifies a generic object (called, Throughput Advice) that can be used by mechanims for hosts to dynamically discover these network rate-limit policies. This information can then passed to applications that might adjust their behaviors accordingly.

Given that all IPv6 hosts and networks are required to support Neighbor Discovery [RFC4861], this document specifies a Neighbor Discovery option to be used in Router Advertisements (RAs) to communicate rate-limit policies to hosts (Section 4). The main motivations for the use of ND for such a discovery are listed in Section 3 of [RFC8781]:

* Fate sharing

* Atomic configuration

* Updatability: change the policy at any time

* Deployability

For address family parity, a DHCP option [RFC2132] is also defined for IPv4 in Section 5. Section 6 describes a discovery approach using Provisioning Domains (PvDs) [RFC8801].

These options are called: Network Rate-Limit Policy (NRLP).

Whether host-to-network, network-to-host, or both policies are returned in an NRLP is deployment specific. All these combinations are supported in this document. Also, the design supports returning one more NRLP instances for a given traffic direction.

Applications will have access to all available NRLPs and will, thus, adjust their behavior as a function of scope and traffic category indicated in a policy. Likewise, a host with multiple network attachments may use the discovered NRLPs to decide how to distribute its flows over these network attachments (prefer a network attachment to place an application session, migrate connection, etc.). That’s said, this document does not make any recommendation about how a receiving host uses the discovered policies.

Networks that advertize NLRPs are likely to maintain the policing in place within the network because of the trust model (hosts are not considered as trusted devices). Per-subscriber rate-limit policies are generally recommended to protect nodes against Denial of Service (DoS) attacks (e.g., Section 9.3 of [RFC8803]). Discussion about conditions under which such a trust model can be relaxed is out of scope of this document.

To enhance flexibility in applying rate-limiting policies and better accommodate diverse endpoint performance requirements, mechanisms such as solicited Router Advertisements (RAs) [RFC8273] and endpoint-specific DHCP responses can be used. These unicast responses enable granular signaling of rate-limit policies to individual endpoints, facilitating differentiated rate-limit configurations. However, this document does not prescribe how resources should be partitioned within local networks, as such considerations fall outside its scope.

This document does not assume nor preclude that other mechanisms, e.g., Low Latency, Low Loss, and Scalable Throughput (L4S) [RFC9330], are enabled in a bottleneck link. The reader may refer to I-D.brw-scone-manageability for a list of relevant mechanisms.

Refer to [NRLP-WIRE] for configuration examples to use NRLP.

2. Conventions and Definitions

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "NOT RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in BCP 14 [RFC2119] [RFC8174] when, and only when, they appear in all capitals, as shown here.

This document uses the terms defined in [I-D.brw-scone-throughput-advice-blob].

3. Common NLRP Parameters

The following common fields are present in all NRLP options:

3.1. Instance Flags (IF)

The format of this 8-bit flags is shown in Figure 1. This field is used to express some generic properties of the advice.

Figure 1: Flow flags Field

See Section 5 of [I-D.brw-scone-throughput-advice-blob] for the meaning of the R/D/S flags.

U are unassigned bits. These bits MUST be set to zero by senders and MUST be ignored by receivers.

3.2. Traffic Category (TC) & Throughput Parameters

The following parameters are used:

TC: See Section 5 of [I-D.brw-scone-throughput-advice-blob].

Committed Information Rate (CIR) (Mbps): See Section 5 of [I-D.brw-scone-throughput-advice-blob].

This is a mandatory parameter.

Committed Burst Size (CBS) (bytes): See Section 5 of [I-D.brw-scone-throughput-advice-blob].

This is a mandatory parameter.

4. IPv6 RA NRLP Option

4.1. Option Format

The format of the IPv6 RA NRLP option, with only mandatory fields included, is illustrated in Figure 2.

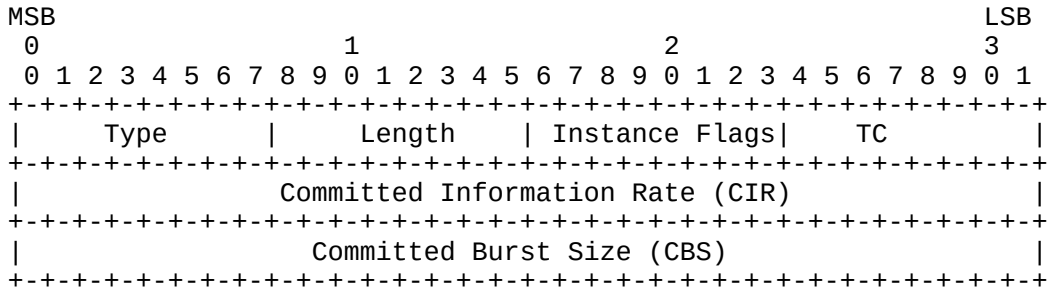

Figure 2: NRLP Option Format

The fields of the option shown in Figure 2 are as follows:

Type: 8-bit identifier of the NRLP option as assigned by IANA (TBD1) (see Section 8.1).

Length: 8-bit unsigned integer. The length of the option (including the Type and Length fields) is in units of 8 octets.

Refer to Section 3 for the meaning of the other parameters.

4.2. IPv6 Host Behavior

The procedure for rate-limit configuration is the same as it is with any other Neighbor Discovery option [RFC4861].

The host MUST be prepared to receive multiple NRLP options in RAs; each with distinct scope and/or application group.

If the host receives multiple NRLP options with overlapping scope/TC, the host MUST silently discard all these options.

If the receiving host has LAN capabilities (e.g., mobile CE or mobile handset with tethering), the following behavior applies:

* If an RA NRLP is advertised from the network, and absent local rate-limit policies, the device should send RAs to the downstream attached LAN devices with the same NRLP values received from the network.

* If local rate-limit policies are provided to the device, the device may change the scope or values received from the network to accommodate these policies. The device may decide to not relay received RAs to internal nodes if local policies were already advertized using RAs and those policies are consistent with the network policies.

Applications running over a host can learn the bitrates associated with a network attachment by invoking a dedicated API. The exact details of the API is OS-specific and, thus, out of scope of this document.

5. DHCP NRLP Option

Note that the base DHCP can only signal a rate policy change when the client first joins the network or renews its lease, whereas IPv6 ND can update the rate policy at the network's discretion. [RFC6704] specifies an approach for forcing reconfiguration of individual hosts without suffering from the limitations of the FORCERENEW design in [RFC3203].

5.1. Option Format

The format of the DHCP NRLP option is illustrated in Figure 3.

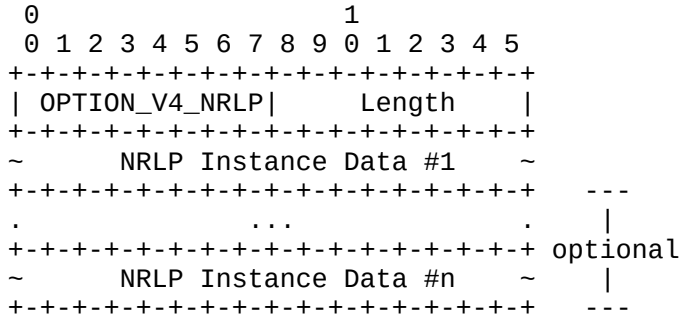

Figure 3: NRLP DHCP Option Format

The fields of the option shown in Figure 3 are as follows:

Code: OPTION_V4_NRLP (TBD2). (see Section 8.2).

Length: Indicates the length of the enclosed data in octets.

NRLP Instance Data: Includes a network rate-limit policy. The format of this field with only mandatory parameters is shown in Figure 4.

When several NRLPs are to be included, the "NRLP Instance Data" field is repeated.

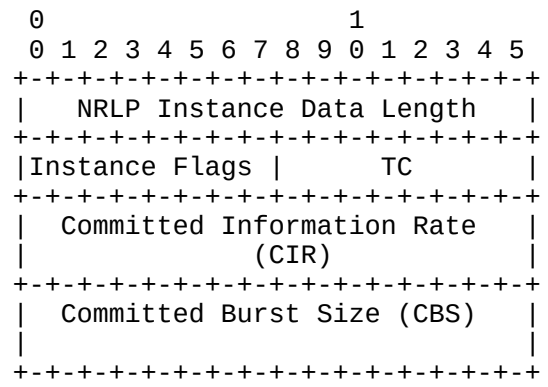

Figure 4: NRLP Instance Data Format with Mandatory Fields

The fields shown in Figure 4 are as follows:

NRLP Instance Data Length: Length of all following data in octets. This field is set to ’8’ when only the nominal bitrate is provided for an NLRP instance.

Refer to Section 3 for the meaning of the other parameters.

OPTION_V4_NRLP is a concatenation-requiring option. As such, the mechanism specified in [RFC3396] MUST be used if OPTION_V4_NRLP exceeds the maximum DHCP option size of 255 octets.

OPTION_V4_NRLP is permitted to be included in the RADIUS DHCPv4-Options Attribute [RFC9445].

5.2. DHCPv4 Client Behavior

To discover a network rate-limit policy, the DHCP client includes OPTION_V4_NRLP in a Parameter Request List option [RFC2132].

The DHCP client MUST be prepared to receive multiple "NRLP Instance Data" field entries in the OPTION_V4_NRLP option; each instance is to be treated as a separate network rate-limit policy.

6. Provisioning Domains

PvD may also be used as a mechanism to discover NRLP. Typically, the network will configured to set the H-flag so clients can request PvD Additional Information (Section 4.1 of [RFC8801]).

Figure 5 provides an example of the returned "application/pvd+json" to indicate a network-to-host NRLP for all subscriber traffic. The NRLP list may include multiple instances if distinct policies are to be returned for distinct traffic categories.

```
{
   "nrlp":[
      {
         "direction":0,
         "scope":0,
         "tc":0,
         "cir":50,
         "cbs":10000,
         "ebs":20000
      }
   ]
}
```

Figure 5: NRLP Example with PvD

7. Security Considerations

The techniques discussed in the document offer the following security benefit: An OS can identify the type of application (background, foreground, streaming, real-time, etc.) and enforce appropriate network policies, even if a misbehaving application tries to evade the rate-limit policies. If an application attempts to bypass rate-limiting by changing its 5-tuple or creating multiple flows, the OS can detect this and manage the application’s traffic accordingly.

7.1. ND

As discussed in [RFC8781], because RAs are required in all IPv6 configuration scenarios, RAs must already be secured, e.g., by deploying an RA-Guard [RFC6105]. Providing all configuration in RAs reduces the attack surface to be targeted by malicious attackers trying to provide hosts with invalid configuration, as compared to distributing the configuration through multiple different mechanisms that need to be secured independently.

RAs are already used in mobile networks to advertize the link MTU. The same security considerations for MTU discovery apply for the NRLP discover.

An attacker who has access to the RAs exchanged over an AC may:

Decrease the bitrate: This may lower the perceived QoS if the host aggressively lowers its transmission rate.

Increase the bitrate value: The AC will be overloaded, but still the rate-limit at the network will discard excess traffic.

Drop RAs: This is similar to the current operations, where no NRLP RA is shared.

Inject fake RAs: The implications are similar to the impacts of tweaking the values of a legitimate RA.

7.2. DHCP

An attacker who has access to the DHCP exchanged over an AC may do a lot of harm (e.g., prevent access to the network).

The following mechanisms may be considered to mitigate spoofed or modified DHCP responses:

DHCPv6-Shield [RFC7610]: The network access node (e.g., a border router, a CPE, an Access Point (AP)) discards DHCP response messages received from any local endpoint.

Source Address Validation Improvement (SAVI) solution for DHCP [RFC7513]: The network access node filters packets with forged source IP addresses.

The above mechanisms would ensure that the endpoint receives the correct NRLP information, but these mechanisms cannot provide any information about the DHCP server or the entity hosting the DHCP server.

8. IANA Considerations

8.1. Neighbor Discovery Option

This document requests IANA to assign the following new IPv6 Neighbor Discovery Option type in the "IPv6 Neighbor Discovery Option Formats" sub-registry under the "Internet Control Message Protocol version 6 (ICMPv6) Parameters" registry maintained at [IANA-ND].

Type	Description	Reference
TBD1	NRLP Option	This-Document

Table 1: Neighbor Discovery NRLP Option

Note to the RFC Editor: Please replace all "TBD1" occurrences with the assigned value.

8.2. DHCP Option

This document requests IANA to assign the following new DHCP Option Code in the "BOOTP Vendor Extensions and DHCP Options" registry maintained at [IANA-BOOTP].

Tag	Name	Data Length	Meaning	Reference
TBD2	OPTION_V4_NRLP	N	NRLP Option	This-Document

Table 2: DHCP NRLP Option

Note to the RFC Editor: Please replace all "TBD2" occurrences with the assigned value.

8.3. DHCP Options Permitted in the RADIUS DHCPv4-Options Attribute

This document requests IANA to add the following DHCP Option Code to the "DHCP Options Permitted in the RADIUS DHCPv4-Options Attribute" registry maintained at [IANA-BOOTP].

Tag	Name	Reference
TBD2	OPTION_V4_NRLP	This-Document

Table 3: New DHCP Option Permitted in the RADIUS DHCPv4-Options Attribute Registry

8.4. Provisioning Domains Split DNS Additional Information

IANA is requested to add the following entry to the "Additional Information PvD Keys" registry under the "Provisioning Domains (PvDs)" registry group [IANA-PVD]:

JSON key: "nrlp"

Description: "Network Rate-Limit Policies (NRLPs)"

Type: Array of Objects

Example:

```
{
   "nrlp":[
      {
         "scope":0,
         "direction":0,
         "tc":0,
         "cir":50,
         "cbs": 10000
      }
   ]
}
```

Reference: This-Document

8.5. New PvD Network Rate-Limit Policies (NRLPs) Registry

IANA is requested to create a new registry "PvD Rate-Limit Policies (NRLPs)" registry, within the "Provisioning Domains (PvDs)" registry group.

The initial contents of this registry are provided in Table 4.

JSON key	Description	Type	Example	Reference
scope	Specifies whether the policy is per host (when set to "1") or per subscriber (when set to "0)	Boolean	1	This-Document
direction	Indicates the traffic direction to which a policy applies	integer	1	This-Document
reliability	Specifies whether the policy is for both reliable and unreliable traffic (when	integer	1	This-Document

	set to "0"), for reliable (when set to "1"), or for unreliable traffic (when set to "2")				
tc	Specifies a traffic category to which this policy applies	Integer	0	This-Document	
cir	Committed Information Rate (CIR)	Integer	50	This-Document	
cbs	Committed Burst Size (CBS)	Integer	10000	This-Document	

Table 4: Initial PvD Network Rate-Limit Policies (NRLPs) Registry Content

Assignments must not be added directly to the "PvD Network Rate-Limit Policies (NRLPs)" registry. When a new attribute is added to the "SCONE Rate-Limit Policy Objects" Registry Group created by [I-D.brw-scone-throughput-advice-blob], a new JSON key is mirrored in the "PvD Network Rate-Limit Policies (NRLPs)" registry.

9. References

9.1. Normative References

[I-D.brw-scone-throughput-advice-blob]
Boucadair, M., Wing, D., Reddy.K, T., Rajagopalan, S., and L. M. Contreras, "Throughput Advice Object for SCONE", Work in Progress, Internet-Draft, draft-brw-scone-throughput-advice-blob-02, 13 December 2024, <https://datatracker.ietf.org/doc/html/draft-brw-scone-throughput-advice-blob-02>.

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, DOI 10.17487/RFC2119, March 1997, <https://www.rfc-editor.org/rfc/rfc2119>.

[RFC2132] Alexander, S. and R. Droms, "DHCP Options and BOOTP Vendor Extensions", RFC 2132, DOI 10.17487/RFC2132, March 1997, <https://www.rfc-editor.org/rfc/rfc2132>.

[RFC3396] Lemon, T. and S. Cheshire, "Encoding Long Options in the Dynamic Host Configuration Protocol (DHCPv4)", RFC 3396, DOI 10.17487/RFC3396, November 2002, <https://www.rfc-editor.org/rfc/rfc3396>.

[RFC4861] Narten, T., Nordmark, E., Simpson, W., and H. Soliman, "Neighbor Discovery for IP version 6 (IPv6)", RFC 4861, DOI 10.17487/RFC4861, September 2007, <https://www.rfc-editor.org/rfc/rfc4861>.

[RFC8174] Leiba, B., "Ambiguity of Uppercase vs Lowercase in RFC 2119 Key Words", BCP 14, RFC 8174, DOI 10.17487/RFC8174, May 2017, <https://www.rfc-editor.org/rfc/rfc8174>.

[RFC8801] Pfister, P., Vyncke, ., Pauly, T., Schinazi, D., and W. Shao, "Discovering Provisioning Domain Names and Data", RFC 8801, DOI 10.17487/RFC8801, July 2020, <https://www.rfc-editor.org/rfc/rfc8801>.

[RFC9445] Boucadair, M., Reddy.K, T., and A. DeKok, "RADIUS Extensions for DHCP-Configured Services", RFC 9445, DOI 10.17487/RFC9445, August 2023, <https://www.rfc-editor.org/rfc/rfc9445>.

9.2. Informative References

[IANA-BOOTP] IANA, "BOOTP Vendor Extensions and DHCP Options", <https://www.iana.org/assignments/bootp-dhcp-parameters/>.

[IANA-ND] IANA, "IPv6 Neighbor Discovery Option Formats", <https://www.iana.org/assignments/icmpv6-parameters/>.

[IANA-PVD] IANA, "Provisioning Domains (PvDs)", <https://www.iana.org/assignments/pvds/>.

[NRLP-WIRE] "Examples of Wire Format Options", <https://github.com/boucadair/draft-xxx-ac-rate-policy-discovery/blob/main/example-nrlp-wire-format.md>.

[RFC2865] Rigney, C., Willens, S., Rubens, A., and W. Simpson, "Remote Authentication Dial In User Service (RADIUS)", RFC 2865, DOI 10.17487/RFC2865, June 2000, <https://www.rfc-editor.org/rfc/rfc2865>.

[RFC3203] T'Joens, Y., Hublet, C., and P. De Schrijver, "DHCP reconfigure extension", RFC 3203, DOI 10.17487/RFC3203, December 2001, <https://www.rfc-editor.org/rfc/rfc3203>.

[RFC6105] Levy-Abegnoli, E., Van de Velde, G., Popoviciu, C., and J. Mohacsi, "IPv6 Router Advertisement Guard", RFC 6105, DOI 10.17487/RFC6105, February 2011, <https://www.rfc-editor.org/rfc/rfc6105>.

[RFC6704] Miles, D., Dec, W., Bristow, J., and R. Maglione, "Forcerenew Nonce Authentication", RFC 6704, DOI 10.17487/RFC6704, August 2012, <https://www.rfc-editor.org/rfc/rfc6704>.

[RFC7066] Korhonen, J., Ed., Arkko, J., Ed., Savolainen, T., and S. Krishnan, "IPv6 for Third Generation Partnership Project (3GPP) Cellular Hosts", RFC 7066, DOI 10.17487/RFC7066, November 2013, <https://www.rfc-editor.org/rfc/rfc7066>.

[RFC7513] Bi, J., Wu, J., Yao, G., and F. Baker, "Source Address Validation Improvement (SAVI) Solution for DHCP", RFC 7513, DOI 10.17487/RFC7513, May 2015, <https://www.rfc-editor.org/rfc/rfc7513>.

[RFC7610] Gont, F., Liu, W., and G. Van de Velde, "DHCPv6-Shield: Protecting against Rogue DHCPv6 Servers", BCP 199, RFC 7610, DOI 10.17487/RFC7610, August 2015, <https://www.rfc-editor.org/rfc/rfc7610>.

[RFC8273] Brzozowski, J. and G. Van de Velde, "Unique IPv6 Prefix per Host", RFC 8273, DOI 10.17487/RFC8273, December 2017, <https://www.rfc-editor.org/rfc/rfc8273>.

[RFC8781] Colitti, L. and J. Linkova, "Discovering PREF64 in Router Advertisements", RFC 8781, DOI 10.17487/RFC8781, April 2020, <https://www.rfc-editor.org/rfc/rfc8781>.

[RFC8803] Bonaventure, O., Ed., Boucadair, M., Ed., Gundavelli, S., Seo, S., and B. Hesmans, "0-RTT TCP Convert Protocol", RFC 8803, DOI 10.17487/RFC8803, July 2020, <https://www.rfc-editor.org/rfc/rfc8803>.

[RFC9330] Briscoe, B., Ed., De Schepper, K., Bagnulo, M., and G. White, "Low Latency, Low Loss, and Scalable Throughput (L4S) Internet Service: Architecture", RFC 9330, DOI 10.17487/RFC9330, January 2023, <https://www.rfc-editor.org/rfc/rfc9330>.

[RFC9473] Enghardt, R. and C. Krhenbhl, "A Vocabulary of Path Properties", RFC 9473, DOI 10.17487/RFC9473, September 2023, <https://www.rfc-editor.org/rfc/rfc9473>.

[TS-23.501] 3GPP, "TS 23.501: System architecture for the 5G System (5GS)", 2024, <https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=3144>.

Appendix A. Example of Authentication, Authorization, and Accounting (AAA)

Figure 6 provides an example of the exchanges that might occur between a DHCP server and an Authentication, Authorization, and Accounting (AAA) server to retrieve the per-subscriber NRLPs.

This example assumes that the Network Access Server (NAS) embeds both Remote Authentication Dial-In User Service (RADIUS) [RFC2865] client and DHCP server capabilities.

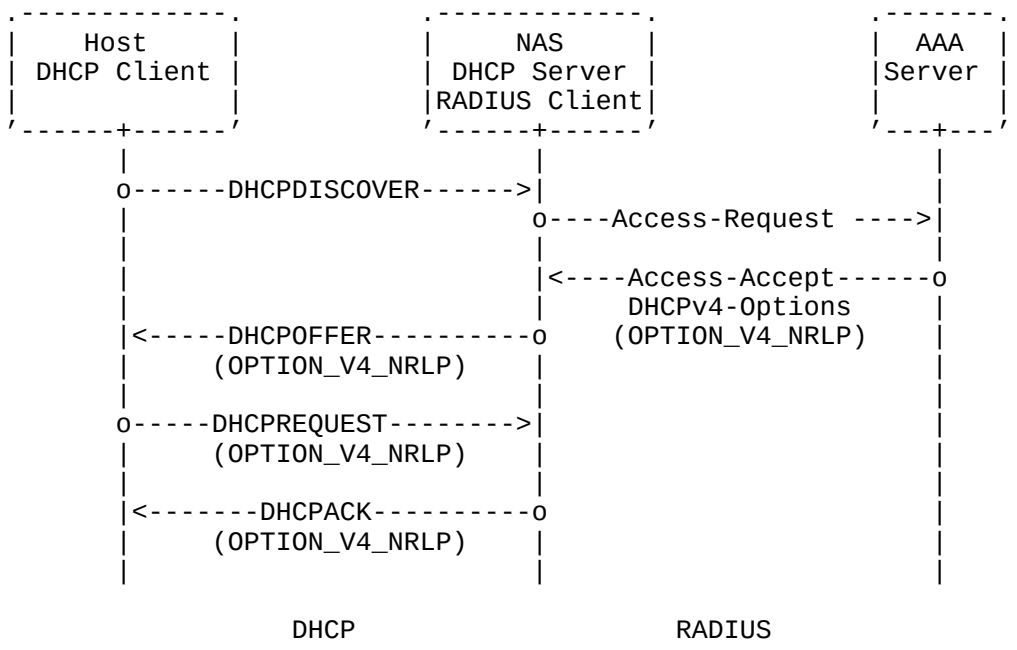

Figure 6: An Example of RADIUS NRLP Exchanges

Acknowledgments

Thanks to Tommy Pauly for the comment on PvD.

Authors' Addresses

Dan Wing
Cloud Software Group Holdings, Inc.
United States of America
Email: danwing@gmail.com

Tirumaleswar Reddy
Nokia
India
Email: kondtir@gmail.com

Sridharan Rajagopalan
Cloud Software Group Holdings, Inc.
United States of America
Email: sridharan.girish@gmail.com

Gyan Mishra
Verizon Inc
United States of America
Email: gyan.s.mishra@verizon.com

Markus Amend
Deutsche Telekom
Germany
Email: markus.amend@telekom.de

Luis M. Contreras
Telefonica
Spain
Email: luismiguel.contrerasmurillo@telefonica.com