Network Working Group
Internet-Draft
Intended status: Informational
Expires: 26 August 2025

H. Song
Futurewei Technologies
D. Eastlake
Independent

22 February 2025

The Architecture of Network-Aware Domain Name System (DNS)
draft-song-network-aware-dns-06

Abstract

This document describes a framework which extends the Domain Name System (DNS) to provide network awareness to applications. The framework enables DNS system responses that are dependent on communication service requirements such as QoS or path without changes in the format of DNS protocol messages or application program interfaces (APIs). The different enhancement methods and use cases are discussed.

Table of Contents

1. Introduction

Different application flows have different requirements on networking, such as bandwidth, delay, jitter, reliability, security, and so on. Many requirements are critical for the quality of service. Users are more willing to pay for premium services (e.g., Virtual Reality (VR)) than before. Meanwhile, today's networks have advanced beyond the best-effort model and are capable of providing per-flow services to meet various application requirements (e.g., QoS) by means of programmability, resource management (e.g., network slicing), traffic engineering, and path regulation (e.g., segment routing and service function chaining).

However, a clear gap exists. Applications usually only care about the abstract requirements ("WHAT") instead of the actual measures for networks to meet such requirements ("HOW"). So far there is no direct means for networks to tell applications their capabilities because the applications do not know what to do about them. On the other hand, due to the limitation of the commonly available network socket API, it is also difficult for applications to convey their service requirements to networks. Currently, if any service that different from "best effort" is desired, one either assumes the requirements can be expressed to network controllers through some out-of-band manner or resorts to encoding the requirements into the packets (e.g., options in IPv6 extension headers network tokens [I-D.yiakoumis-network-tokens]). We need a simpler and more extensible way to set up the service contract.

We define a framework to support network awareness through DNS. Requirements for network services can be incorporated into DNS queries from a host (e.g., as specified in [I-D.eastlake-dnsop-expressing-qos-requirements]) and the returned information enables access to services meeting those requirements. For example, by including new semantics representing a service commitment embedded in the returned IP addresses (i.e., semantic addressing [I-D.farrel-irtf-introduction-to-semantic-routing]). Alternatively, a richer format for expressing how to obtain the desired service would be to query for a service binding (SVCB) RR [RFC8460].

The Domain Name System (DNS) is a distributed database that stores data under hierarchical domain names and supports redundant servers, data caching, and security features. The data is formatted into resource records (RRs) whose content type and structure are indicated by the RR Type field. A typical use of DNS is that, by running the DNS protocol, a host gets the IP addresses stored at a domain name from DNS servers through a DNS resolver. Many other types of data besides IP addresses can be stored in and returned by the DNS.

In a nutshell, the application’s service requirements are embedded into the DNS queries from a host. The DNS replies either provide semantic IP addresses or data that help construct the packet header or headers signaling the special packet handling in networks. The application flow packets may use the existing socket API to send the packet. Network devices, after capturing such packets, would decode the semantics and apply any special packet handling accordingly.

This document describes the architecture, requirements, and use cases of the Network-Aware DNS. The details on DNS query encoding and semantic addressing/data in DNS replies will be described in other documents.

1.1. Terminology and Acronyms

The following terminology and acronyms are used in this document.

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.

API - Application Program Interface

DNS - Domain Name System

RR - Resource Record [RFC8499]. The unit of data stored in the DNS.

Semantic Addressing - Encoding extra semantics beyond the destination ID in an address

2. Architecture

The architecture of the Network Aware DNS where the required DNS reply information can be statically stored in the DNS is shown in Figure 1.

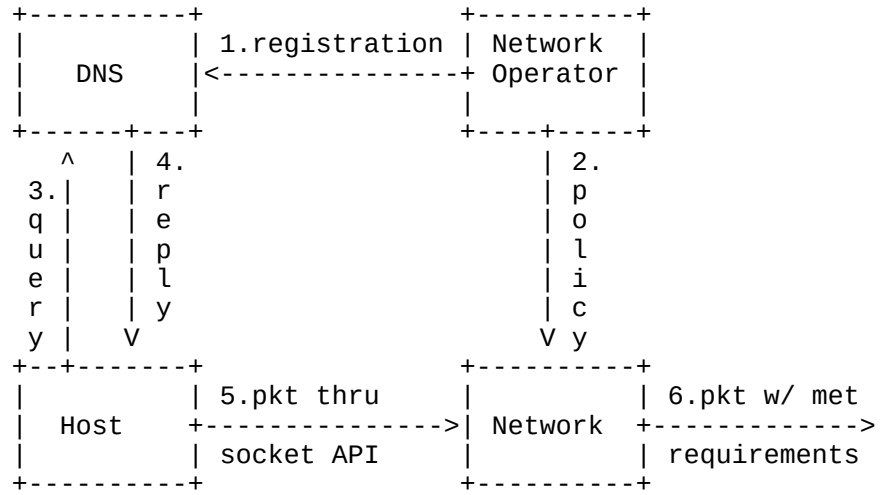

Figure 1: Static Architecture

The architecture of the Network Aware DNS where the required DNS reply information is dynamically computed is shown in Figure 2.

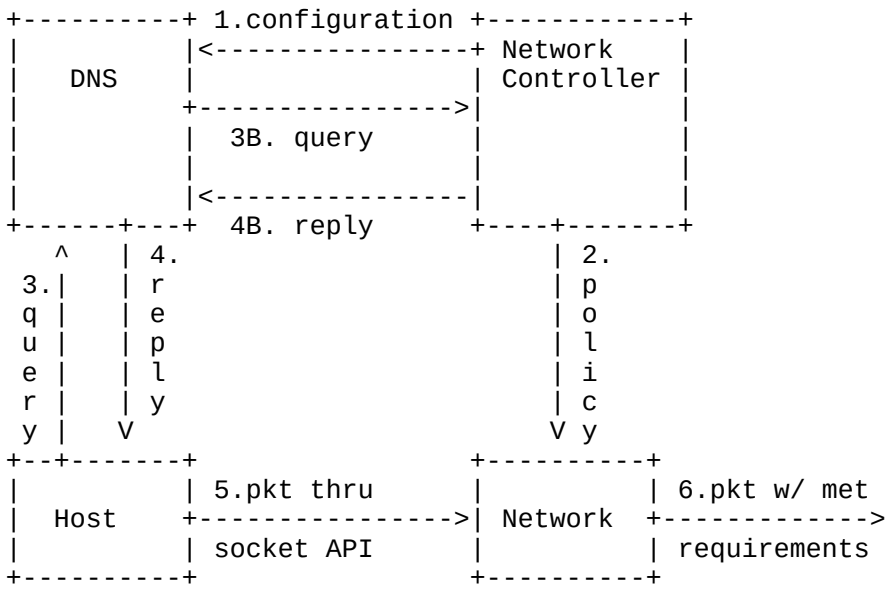

Figure 2: Dynamic Architecture

Figure 2: Dynamic Architecture

The procedure steps are explained as follows:

1. The network operator registers the semantic addresses/data associated with a name in authoritative DNS servers in form of RRs. In addition to the location, the semantics represent the commitments for network to meet certain service requirements. The semantic addresses or data can be dynamically computed or statically configured by the network operator.

2. Meanwhile, the packet processing policy corresponding to each semantic address/data is configured to the network devices such as routers. How the network meets the service requirements is opaque to host applications.

3. A host application, when conducting a DNS query to a name, would also express its service requirement. A host application can also be ignorant of this service requesting scheme; in this case, the normal DNS query is used and the best-effort results are returned. In the static case, this is answered directly by a normal DNS server. In the dynamic case, the query is redirected, through a CNAME or zone referral, to a network controller that can emulate a DNS server and which dynamically computes the response.

4. If the query with service requirements can be satisfied by some RRs in DNS, the result will be returned to the host, either directly in the static case or from the network controller in the dynamic case; otherwise, a normal DNS response, or either an error or the best effort result, will be returned.

5. Once the host application receives the reply, assuming the reply is not an error, it simply uses the address (or assembles the header fields as directed by the semantic data) to forward the packet through a standard socket API. The semantic address or data may be cached at the host for the lifetime of the flow. Alternatively, the DNS response TTL may indicate the period of time for which the semantic address will provide the service assurances, and the application may again query the DNS at or shortly before the end of the time to refresh the semantic address/data or obtain a new address or data that will be effective for a future interval; however, it is not common for TTL information to be returned to an application doing a DNS query.

6. The network devices would process the packets based on the configured policies if the packets carry semantic addresses and/or header fields. Using a semantic address/data other than for the best effort service might be subject to extra cost based on some service agreement.

We enforce some requirements on the architecture to make it practical for incremental deployment.

* No new protocol is introduced to enable the architecture.

* As an infrastructural system and protocol, DNS is stable. No change to DNS architecture and protocol is made. However, within the DNS framework, we explore the freedom to introduce new semantics and new RR types to encode semantic data.

* Similarly, it is hard to change the ubiquitous network socket APIs, so we just rely on the existing ones.

* The system would be better used in limited domains where the network operator owns not only the networks but also the proper name servers. In some cases, it is also possible to extend the scope into multiple domains if the packet processing to meet the service requirements can be coordinated cross domains.

* The semantic address or data should be per application or per flow based. So each application or flow may need its own DNS name resolution even for the same service. Most applications can still use the conventional best effort service without noticing any change.

In the more dynamic architecture, DNS queries with service requirements can be dynamically sent to a controller maintained by the network operator when received by a resolver, allowing network operator to generate on-demand semantic addresses or data for the name server, which will eventually return the information back to the host application.

3. Obtaining Needed Information from DNS

A host application can have three methods to obtain information from the DNS to enable the application to meet its service requirements. These methods are as follows:

Method 1: It sends a requirement-encoded name to ask for an IP address type RR (e.g., AAAA) and expects the semantics to meet the requirements to be embedded in the returned addresses. The encoding method is described in [I-D.eastlake-dnsop-expressing-qos-requirements].

Method 2: It sends a normal name to ask for a different type of RR and the semantic data in the returned RRs represents the means to meet the service requirements (e.g., [I-D.eastlake-dnsop-svcb-rr-tunnel]).

Method 3: Combining 1 and 2, it sends a requirement-encoded name to ask for a different type of RR, which might be in addition to or lead to (such as the SRV type RR) an IP address type RR, and the semantic data in the RR represents the means to meet the service requirements.

This architecture can support multiple use cases using one of the above methods. Below are some examples.

E2E SRv6: This use case may use method 2. We can support true end-to-end SRv6 service where a Segment List (SL) is acquired from DNS using the RR Type specified in [I-D.eastlake-dnsop-rrtype-srv6] and an SRH (Segment Routing Header) is directly inserted in the IPv6 packet header. While the SRH determines the packet’s forwarding path, different packet handling and QoS treatment can also be applied to the packet along the path.

Semantic Addressing: This use case may use method 1. Due to the

abundance of IPv6 addresses, each name can be assigned multiple addresses with each representing some special network services. While the network devices are configured or programmed to be able to interpret and process the semantics embedded in addresses, different services can be applied to flows for the same destination. The details are described in a companion draft.

Service Header Fields: This use case may use method 3. Some service-defining header fields (e.g., DSCP in IPv4 header and traffic class and flow label in IPv6 header) can be used to indicate QoS or other service requirements. Such semantic data can also be provided by DNS replies in form of RRs. The details are described in a companion draft.

Other Semantic Data: This use case may apply the method 2 or 3. Some services may have other means to be encapsulated into a packet (e.g., IPv6 Extension Header). The required information can also be returned by DNS reply as semantic data.

4. Security Considerations

TBD

5. IANA Considerations

This document requires no IANA actions.

6. Acknowledgments

The comments and suggestions of the following are gratefully acknowledged:

* TBD

7. References

7.1. Normative References

[I-D.eastlake-dnsop-expressing-qos-requirements]
Eastlake, D. E. and H. Song, "Expressing Quality of Service Requirements (QoS) in Domain Name System (DNS) Queries", Work in Progress, Internet-Draft, draft-eastlake-dnsop-expressing-qos-requirements-05, 18 September 2024, <https://datatracker.ietf.org/doc/html/draft-eastlake-dnsop-expressing-qos-requirements-05>.

[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/info/rfc2119>.

[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/info/rfc8174>.

7.2. Informative References

[I-D.eastlake-dnsop-rrtype-srv6]
Eastlake, D. E. and H. Song, "The IPv6 Segment Routing (SRv6) Domain Name System (DNS) Resource Record", Work in Progress, Internet-Draft, draft-eastlake-dnsop-rrtype-srv6-07, 19 February 2025, <https://datatracker.ietf.org/doc/html/draft-eastlake-dnsop-rrtype-srv6-07>.

[I-D.eastlake-dnsop-svcb-rr-tunnel]
Eastlake, D. E. and H. Song, "A Domain Name System (DNS) Service Parameter and Resource Record for Tunneling Information", Work in Progress, Internet-Draft, draft-eastlake-dnsop-svcb-rr-tunnel-06, 20 October 2024, <https://datatracker.ietf.org/doc/html/draft-eastlake-dnsop-svcb-rr-tunnel-06>.

[I-D.farrel-irtf-introduction-to-semantic-routing]
Farrel, A. and D. King, "An Introduction to Semantic Routing", Work in Progress, Internet-Draft, draft-farrel-irtf-introduction-to-semantic-routing-04, 25 April 2022, <https://datatracker.ietf.org/doc/html/draft-farrel-irtf-introduction-to-semantic-routing-04>.

[I-D.yiakoumis-network-tokens]
Yiakoumis, Y., McKeown, N., and F. Sorensen, "Network Tokens", Work in Progress, Internet-Draft, draft-yiakoumis-network-tokens-02, 22 December 2020, <https://datatracker.ietf.org/doc/html/draft-yiakoumis-network-tokens-02>.

[RFC8460] Margolis, D., Brotman, A., Ramakrishnan, B., Jones, J., and M. Risher, "SMTP TLS Reporting", RFC 8460, DOI 10.17487/RFC8460, September 2018, <https://www.rfc-editor.org/info/rfc8460>.

[RFC8499] Hoffman, P., Sullivan, A., and K. Fujiwara, "DNS Terminology", RFC 8499, DOI 10.17487/RFC8499, January 2019, <https://www.rfc-editor.org/info/rfc8499>.

Authors' Addresses

Haoyu Song
Futurewei Technologies
2220 Central Expressway
Santa Clara, CA 95050
United States of America
Email: haoyu.song@futurewei.com

Donald Eastlake
Independent
2386 Panoramic Circle
Apopka, FL 32703
United States of America
Phone: +1-508-333-2270
Email: d3e3e3@gmail.com