Network Working Group
Internet Draft
Intended Status: Informational

Young Lee (Huawei)
Dave McDysan (Verizon)
Ning So (UTD)
Greg Bernstein (Grotto)
Tae Yeon Kim (ETRI)
Kohei Shiomoto (NTT)
Oscar Gonzalez de Dios (Telefonica)

April 20, 2011

Problem Statement for Network Stratum Query

draft-lee-network-stratum-query-problem-02.txt

Status of this Memo

Copyright Notice

Abstract

This document describes the general problem of network stratum query for application optimization. Network Stratum query is an ability to query the network from an application controller such as those used in Data Centers so that application controller decisions such as server assignment or virtual machine instantiation/migration could be performed with better knowledge of the underlying network conditions.

As application servers are distributed geographically across Data Centers, many application-related decisions such as which server to assign a new client or where to instantiate/migrate virtual machines will suffer from sub-optimality unless the underlying network conditions are factored in the decision process. The lack of network awareness may result in not meeting the end-user service objective for some key applications like video gaming/conferencing that require stringent latency and bandwidth requirement.

Table of Contents

1. Introduction..... 2
2. Network Contexts..... 4
3. Problem Statement 7
4. High-level requirements..... 9
4.1. Data Center-Network Stratum Communication (NS Query) Error! Bookmark not defined.
4.1.1. Application Profile..... 9
4.1.2. Network Load Data to be queried..... 10
4.1.3. Responses to NS Query from network to application. 10
5. Security Considerations..... 11
6. References..... 11
Author's Addresses..... 13
Intellectual Property Statement 13
Disclaimer of Validity..... 14

1. Introduction

Cross Stratum Optimization is a joint optimization effort in allocating resources to end-users that involves both the Application Stratum and Network Stratum.

The application stratum is the functional block which manages and controls application resources and provides application resources to a variety of clients/end-users. Application resources are non-network

resources critical to achieving the application service functionality. Examples include: application specific servers, storage, content, large data sets, and computing power. Data Centers are regarded as tangible realization of the application stratum architecture.

The network stratum is the functional block which manages and controls network resources and provides transport of data between clients/end-users to and among application resources. Network Resources are resources of any layer 3 or below (L1/L2/L3) such as bandwidth, links, paths, path processing (creation, deletion, and management), network databases, path computation, admission control, and resource reservation capability.

Application services offered by Data Centers by their very nature utilize application resources (e.g., servers, storage, memory, etc...) in Data Centers, and the underlying network resources provided by LANs, MANs, and carrier's transport networks.

As the application servers are distributed geographically across many Data Centers, decisions such as server assignment or new virtual machine instantiation/migration will suffer from sub-optimality unless the underlying network conditions are factored in the decision process. The lack of network awareness may result in not meeting the end-user service objective for some key applications like video gaming/conferencing that require stringent latency and bandwidth requirement.

This document describes the general problem of network stratum query (NS Query) in Data Center environments. Network Stratum query is an ability to query the network from application controller in Data Centers so that application server assignment or virtual machine instantiation/migration decision would be jointly performed based on both the application resource/load status and the network resource/load status.

The NS query is different from typical "horizontal" query capabilities in the network. The horizontal query in the network is carried by the head end (i.e., data source) that would "probe" the network to test the capabilities for data flows to/from particular point in the network. This is a horizontal scheme.

NS Query is a two-stage query that consists of two stages:

. A vertical query capability where an external point (i.e., the Application Control Gateway (ACG) in Data Center) will query the network (i.e., the Network Control Gateway (NCG)); and

. A horizontal query capability where the NCG to gather the collective information of a variety of horizontal schemes (IPPM, IGP, RIB, etc.) implemented in the network stratum.

NS Query does not re-invent the wheel on existing network capabilities but tries to reuse them where possible.

2. Network Contexts

Figure 1 shows a typical data center architecture where an end-user (the point of consuming resource) needs to be connected for its application (e.g., gaming) to a server located in one of the data centers geographically spread.

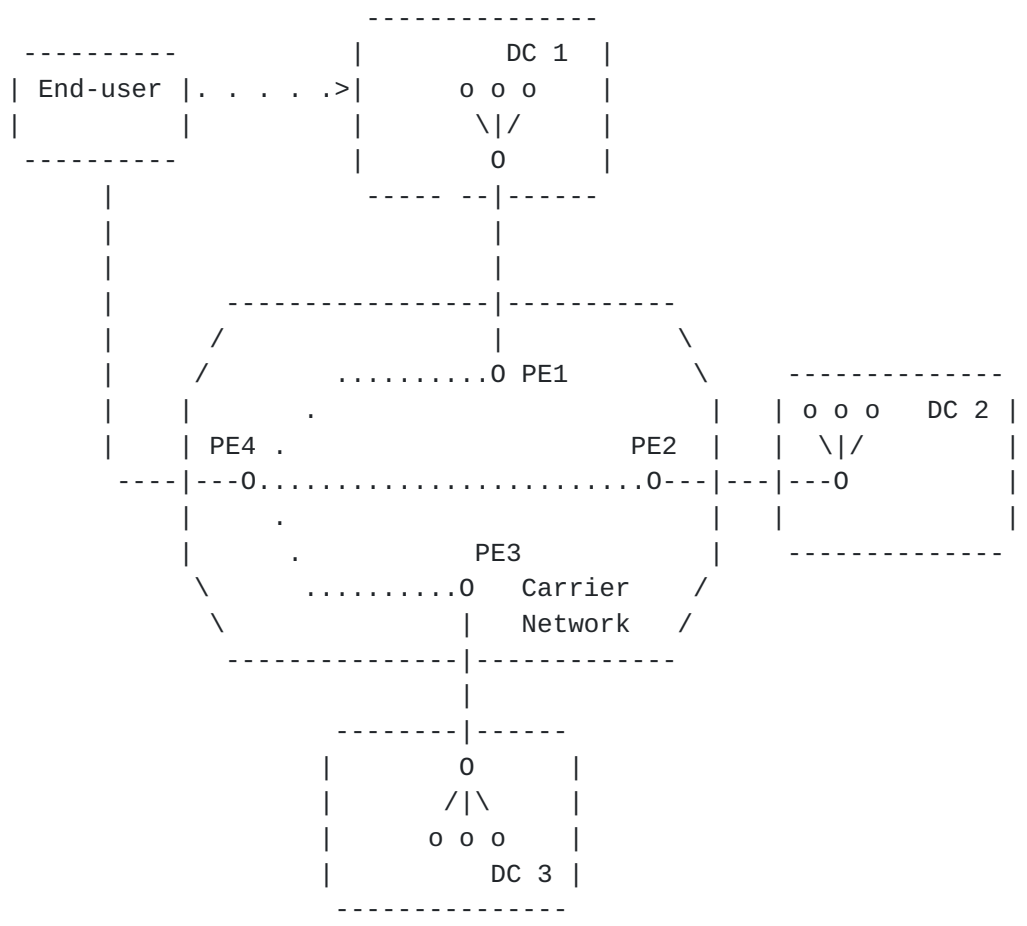

Figure 1. Data Center Architecture

Figure 1 shows that the user application can be served by any of the servers in DC1, DC2 or DC3. When the initial request arrives to the proxy server in DC1, the proxy server (aka, the load balancer) would ideally assign an "optimal" server based on both server resource/load status and the network resources/load status. This server assignment decision today, however, is limited due to the lack of network awareness in this decision making process in the application.

For example, the server close to the user in Data Center 1 may find a good server that can serve the application. Assume that this particular application requires x amount of minimum bandwidth guarantee and with less than y ms of latency limit. The route that serves Data Center 1 traffic to the end-user (PE1 - PE4) may not have enough capacity at a moment of service instantiation and therefore the service objective of the end-user may not be satisfied had such route been taken.

On the other hand, there may be good servers available in Data Centers 2 and 3 and their routes (PE2-PE4 and PE3-PE4) may have enough capacity to meet the service requirement.

This example illustrates the benefit of and the need for the joint optimization across the application and network strata. NS Query is the ability to query the network from an application to collect a certain level of network information. No such mechanisms exist in the today's Internet Protocol technologies.

Figure 2 shows the context of NS Query in a more detail within the overarching data center architecture shown in Figure 1.

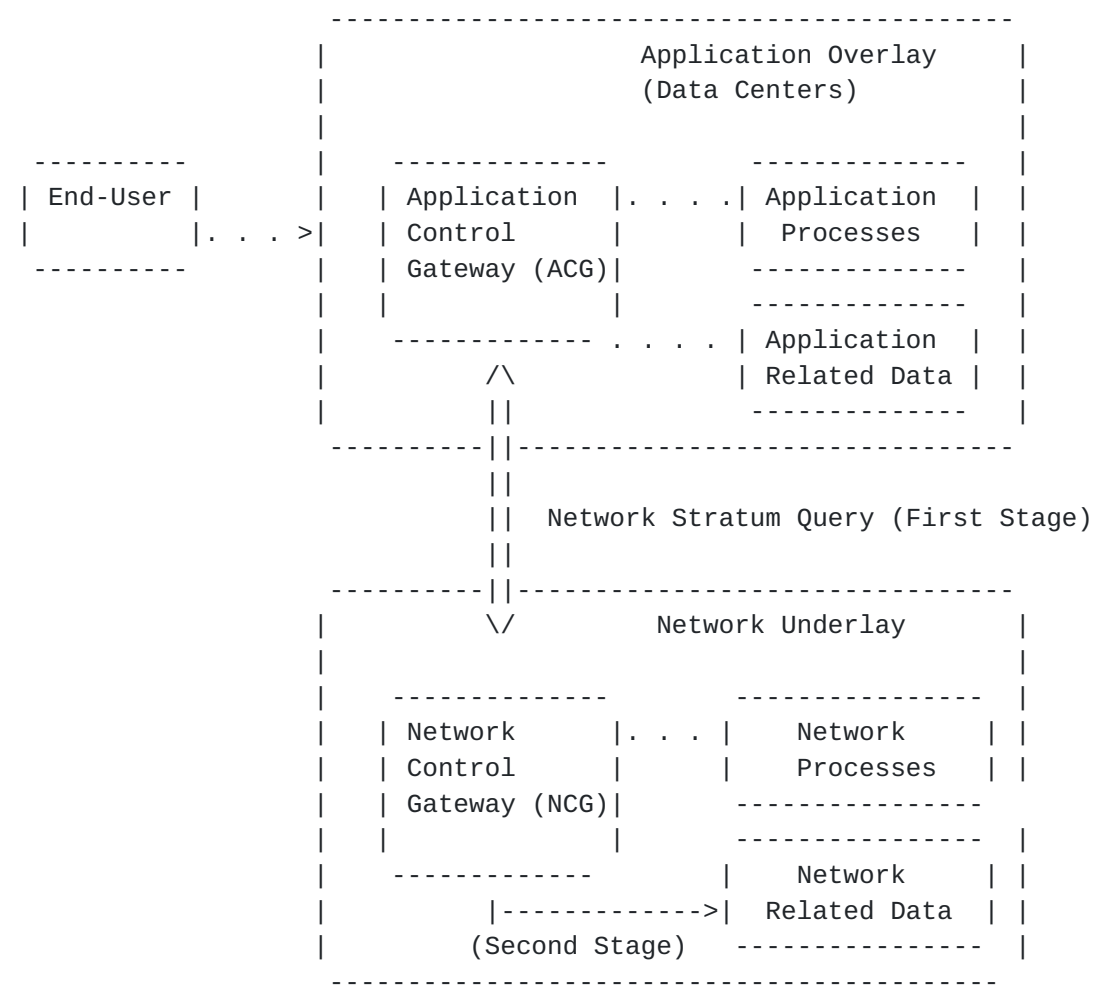

Figure 2. NS Query Architecture

Figure 2 shows key architectural components that enable NS Query capability. The Application Control Gateway (ACG) is the proxy gateway that interfaces with network and generate queries to network. The ACG can query various metric values that may contribute to meeting the overall service objective of an application. This is a vertical query (Stage 1).

In the network stratum, the Network Control Gateway (NCG) serves as the proxy gateway to the network. The NCG receives the query request from the ACG, probes the network to test the capabilities for data flow to/from particular point in the network, and gather the collective information of a variety of horizontal schemes (IPPM, IGP, MIB, TED, etc.) implemented in the network stratum. This is a horizontal query (Stage 2).

Further, the NCG provides the responses to the original query sent from the ACG. The data collected by the NCG needs to be abstracted. This abstraction is needed on two grounds.

First, the network does not usually reveal its details to the outside entity. Although the Data Center providers and the carriers are business partners in providing application services to the end-users and to the application providers (e.g., gaming providers), detail network data may not be leaked to the Data Centers, and vice versa.

Secondly, detail network data may not be understood by the application. Link or node level data in and of themselves may not help the application to process the detail data. For instance, latency or bandwidth on a link level is too detail for application to handle. Instead, latency or bandwidth on a route level (i.e., PE1 - PE4 in Figure 1) will help the application make its server selection/instantiation decision.

The abstraction function needs to be provided by the NCG. Note that NCG plays a head end role within the network probing/collecting network performance/management data (e.g., IPPM, MIB, etc.) or routing data [MRT] (e.g., LSDB, TED, BGP-RIB, etc.) and others. Once the basic data is collected, the NCG will need to abstract/summary before it sends to the application.

3. Problem Statement

3.1. Limitation of existing probing schemes

The current state-of-the art probing schemes from an external point are based on ping or trace route like mechanisms based on the assumption that the underlying transport network is L3 network and that the routing is simple IP forwarding.

In reality, the carrier's routing schemes are likely to include IP tunneling or MPLS tunneling on top of or in place of IP forwarding. In some cases, the actual network may be VPN, MPLS-TE or GMPLS-TE networks where trace route does not work.

This implies that network status estimation technique made from application stratum cannot be accurate. Thus, application resource allocation to end-users can suffer sub-optimality and fail to meet performance objective for the application.

3.2. Lack of vertical query schemes

Currently, the query in the network is carried by the head end (i.e., data source) that would "probe" the network to test the capabilities for data flows to/from particular point in the network. This is a horizontal scheme.

There is no standard "vertical" query scheme that allows an application control gateway in Data Center to query network stratum in a way suitable for a third party (i.e. an entity "outside" the network).

Due to the lack of standard vertical query scheme, there is a limitation on exchanging information between application and network that would increase efficiency of joint optimization across application to network. For instance, the ability to exchange the application profile information (defined in Section 4.1) or network capability information between application and network would increase efficiency of resource allocation across application to network.

3.3. Limitation of SNMP MIB network monitoring techniques

SNMP MIB monitoring techniques as defined in [RFC2261] and [RFC2265] do not provide mechanisms to guarantee synchronization of the data collection. This higher level of synchronization is necessary to service: a) application with stringent QoS and Bandwidth, or to b) better schedule massive quantities of small data flows.

In addition, SNMP MIB Network Monitoring lacks a whole network query capability. A whole network query is a query to gather information across many boxes simultaneously under the control of a single administration domain (AD) as defined in RFC 1136. A single AD means the single AS or multiple ASes under the control of a single AD.

3.4. Lack of abstraction mechanisms

Most of the information needed to provide NS Query is currently available from the network; however, it is not aggregated into a form suitable for use by the application stratum. For example from commonly monitored SNMP based link statistics and current routing tables one can easily compute average available bandwidth and many other statistical performance measures such as packet loss, latency, etc.

However, neither the raw SNMP nor routing table data should be delivered to the application stratum since (a) this reveals too much information concerning the carriers network, (b) presents too much information to transfer to each application. This warrants some works

on abstraction from network side to preserve the privacy of network stratum details from the application stratum.

4. High-level requirements

This section discusses high-level requirements to support NS Query in the Data Center environments.

The ACG plays the key role functioning as an application gateway to network and runs the NS Query. The ACG has access to the end-user profile for the application and the candidate servers' locations locally and remotely located. How the ACG access these information is beyond the scope of this work.

4.1. Application Profile

The application Stratum needs to provide the application profile to network.

Example service profile information that can be useful to network to understand is as follows:

- End user IP address;
- User access router IP address;
- Authentication Profile: Authentication Key;
- Bandwidth Profile: Minimum bandwidth required for the application;
- Connectivity Profile: P-P, P-MP, Anycast (Multi-destination);
- Directionality of the connectivity: unidirectional, bi-directional;
- Path Estimation Objective Function: Min latency, etc.

Additional profile information can be added depending on the network capability.

4.2. Network Load Data to be queried (First Satge)

For a given location mapping information (i.e., from the server location to end-user location), the query from an application can ask the following network load data:

- Type of networks and the technical capabilities of the networks;
- Bandwidth capabilities and availability;
- latency;
- jitter;
- packet loss;
- And other Network Performance Objective (NPO) as defined in section 5 of [ITU-T Y.1541].

Note that this can be asked in a different way. For example, the query can simply ask:

- Can you give me a route with x amount of b/w (from server to end-user) within y ms of latency?
- Can you give me a route with x amount of b/w (from server to end-user) with no packet loss?

4.3. A Whole Network Query capability (Second Stage)

Upon the request from application (specifically, the ACG in Figure 2), the network (specifically the NCG in Figure 2) should perform "a whole network query" of information.

A whole network query is a query to gather information across many boxes simultaneously under the control of a single administration domain (AD) as defined in RFC 1136. A single AD means the single AS or multiple ASes under the control of a single AD.

The scope of a whole network query can include the topology of the network, the bandwidth availability for the routes of interest, the capabilities and congestion of links and routes, and an indication of the contribution to delay and jitter that each link and route will contribute and so on.

4.4. Data Synchronization Mechanism

The ability to capture the data at the same instant should be provided.

4.5. Responses to NS Query from network to application

Given the network query from application, the network should provide the following mechanisms:

- For a given location mapping information from application (i.e., from the server location to end-user location) and the gathered information by the second stage query discussed in section 4.3., the network needs to present the requested information in a standard format and respond to the application.

The actual abstraction mechanism is beyond the scope of this document.

5. Security Considerations

TBD

6. IANA Considerations

This informational document does not make any requests for IANA action.

7. References

7.1. Informative References

[RFC2261] D. Harrington, et al., "An Architecture for Describing SNMP Management Frameworks," January, 1998.

[RFC2265] B. Wijnen, et al., "View-based Access Control Model (VACM) for the Simple Network Management Protocol (SNMP)," January, 1998.

[Y.2011] General principles and general reference model for Next Generation Networks, October, 2004.

[Y.2012] Functional Requirements and architecture of the NGN, April, 2010.

[MRT] L. Blunk, M. Karir, and C. Labovitz, "MRT routing information export format," draft-ietf-grow-mrt, work in progress.

Author's Addresses

Young Lee
Huawei Technologies
1700 Alma Drive, Suite 500
Plano, TX 75075
USA
Phone: (972) 509-5599
Email: ylee@huawei.com

Ning So
Univerity of Texas at Dallas
Email: ningso@yahoo.com

Dave McDysan
Verizon Business
Email: dave.mcdysan@verizon.com

Greg M. Bernstein
Grotto Networking
Fremont California, USA
Phone: (510) 573-2237
Email: gregb@grotto-networking.com

Tae Yeon Kim
ETRI
tykim@etri.or.kr

Kohei Shiomoto
NTT
Email : shiomoto.kohei@lab.ntt.co.jp

Oscar Gonzalez de Dios
Telefonica
Email : ogondio@tid.es

Intellectual Property Statement

represent that it has made any independent effort to identify any such rights.

Copies of Intellectual Property 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 any standard or specification contained in an IETF Document. Please address the information to the IETF at ietf-ipr@ietf.org.

Disclaimer of Validity

Acknowledgment

Funding for the RFC Editor function is currently provided by the Internet Society.