

DNSOP  
Internet-Draft  
Intended status: Informational  
Expires: January 5, 2015

G. Deng  
N. Kong  
S. Shen  
CNNIC  
July 4, 2014

Approach on optimizing DNS authority server placement  
draft-deng-dns-authority-server-placement-00

Abstract

The geographical distribution of DNS authority servers highly affects the DNS query latency and financial costs. This document proposes an approach on optimizing the geographical placement of DNS authority servers so that the DNS query latency is highly reduced while the financial cost is within its budget.

Status of This Memo

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at <http://datatracker.ietf.org/drafts/current/>.

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."

This Internet-Draft will expire on January 5, 2015.

Copyright Notice

Copyright (c) 2014 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust's Legal Provisions Relating to IETF Documents (<http://trustee.ietf.org/license-info>) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of

the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

This document may contain material from IETF Documents or IETF Contributions published or made publicly available before November 10, 2008. The person(s) controlling the copyright in some of this material may not have granted the IETF Trust the right to allow modifications of such material outside the IETF Standards Process. Without obtaining an adequate license from the person(s) controlling the copyright in such materials, this document may not be modified outside the IETF Standards Process, and derivative works of it may not be created outside the IETF Standards Process, except to format it for publication as an RFC or to translate it into languages other than English.

## Table of Contents

1.	Introduction . . . . .	2
1.1.	Terminology . . . . .	4
2.	Use cases and scenarios . . . . .	4
2.1.	Geographically distributing all authority servers . . . . .	4
2.2.	Geographically distributing newly-added authority servers . . . . .	4
2.3.	Readjusting geographical distribution of all authority servers . . . . .	5
3.	Authority server placement approach . . . . .	5
3.1.	Problem statement . . . . .	5
3.2.	Optimization goals . . . . .	6
3.2.1.	Minimizing the average Domain Name Resolution Delay . . . . .	6
3.2.2.	Minimizing the maximal Domain Name Resolution Delay . . . . .	6
3.3.	Authority server placement algorithm . . . . .	6
3.3.1.	Best Efficiency Server Placement Algorithm . . . . .	6
3.3.2.	Best Fairness Server Placement Algorithm . . . . .	8
3.4.	Discussion . . . . .	8
4.	IANA consideration . . . . .	8
5.	Security considerations . . . . .	8
6.	References . . . . .	9
6.1.	Normative References . . . . .	9
6.2.	Informative References . . . . .	9
	Authors' Addresses . . . . .	9

## 1. Introduction

Domain Name System [RFC1035] is one of the most important components of Internet infrastructure and enables the association of the human memorable domain names and their corresponding information like routable IP addresses. It is revealed by Google that in average Internet users need hundreds of DNS lookups to be done in a typical browsing day [BROWSE]. Even for one page, multiple domain name

resolutions are needed since contents from many other domains are incorporated to one page. So the quality of service (QoS) of other Internet applications highly depends on the DNS system.

Also, with the development of telecommunication access technology, the network latency but not the bandwidth gradually becomes the main impediment for improving the QoS of web service, for the bandwidth price is becoming lower and lower. Moreover, for web service providers, their revenue and profit are greatly affected by network latency [Web.Latency]. For instance, Amazon estimates that the Internet latency inversely correlates with revenue and profit, and every 100 milliseconds increase in latency cuts profits by 1%. So shortening DNS query latency is an efficient way for improving the QoS of other Internet applications. Here the DNS query latency is defined as the time difference between the time when a stub resolver sends a DNS query and receives the corresponding response.

Nowadays, due to security threat like DDOS and the deployment of DNSSEC ([RFC4033] [RFC4034] [RFC4035]), the processing capacity of DNS authority servers needs to be increased and at the same time more bandwidth resource has to be added. And the launch of new gTLDs means that more DNS authority servers need to be deployed in the DNS hierarchy. However, to the best of our knowledge, there is still no rigorous authority server placement method yet and thus DNS operation engineers only rely on their personal experiences, which may lead to sub-optimal authority server distribution and thus long DNS query latency and high financial costs.

Fundamentally, there are two ways to shorten the DNS query latency; one is shortening the processing latency on DNS servers (both authority and recursive servers); the other is reducing the network latency between authority and recursive servers as well as that between stub and recursive servers. The processing latency on one authority server usually relates to the rate of incoming DNS queries (whose unit is query per second, QPS for short) and specifically larger the QPS is, longer the processing latency is. For network latency, since recursive servers are usually very near to stub resolvers, we just take the network latency between authority and recursive servers into consideration. For simplicity, we define Domain Name Resolution Delay (DNRD) as the difference between the time when one DNS recursive server sends one DNS query to and receives the corresponding response from a specific authority server. Then DNRD is almost the Round-Trip Time (RTT) between the DNS authority servers (of root domain, top level domain, second level domain et al.) and DNS recursive servers, plus the processing latency of that DNS query on the authority server. Finally, the main object of this document is to minimize the DNRD by reasonably select the location of authority servers without the financial costs exceeding

their budget. The financial costs of operating authority servers mainly consist of two parts: one is the fixed expense including equipment purchase cost, room rental cost, et al; the other is variable expense such as bandwidth rental cost, electricity fees, et al. Usually, different locations have different price level and then different geographical distribution of the same authority servers leads to different financial costs.

### 1.1. Terminology

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 [RFC2119].

## 2. Use cases and scenarios

Fundamentally, there are three main use cases on DNS authority server placement.

### 2.1. Geographically distributing all authority servers

DNS system works in a decentralized fashion and a new DNS zone can be created through the delegation from its parental domain. Several (or tens of even hundreds of ) authority servers having the authority for such new DNS zones have to be deployed for providing authoritative resolution service. For instance, due to the launch of new gTLDs, many authority servers have been being deployed. When the financial cost upper bound is given, the question is where to geographically place these authority servers so that the DNRD is minimized on the condition that the financial cost is within its upper bound. Here, at least two questions should be answered. One is which potential locations should be selected as actual locations; the other is how many authority servers should be placed on each actual location.

### 2.2. Geographically distributing newly-added authority servers

For some existing DNS zones, there are some running authority servers already. However, due to the reason like further reducing the DNRD or decreasing the financial cost, more authority servers will be added without changing the location of already deployed authority servers. Then the location of newly-added authority servers has to be carefully selected to reduce the DNRD as much as possible with a decreased budget.

### 2.3. Readjusting geographical distribution of all authority servers

Since now the authority servers are geographically placed by operation engineers by their personal experience, the authority server placement method may not be optimal. Then the DNRD may be too long and the financial cost may be too high, which calls for readjusting the geographical distribution of all currently deployed authority servers so that the DNRD is shortened and at the same time the financial cost is reduced.

## 3. Authority server placement approach

### 3.1. Problem statement

The problem of authority server placement is like this: given a set of potential locations (which number in thousands even tens of thousands) for placing authority servers, which locations (which usually number in less than ten or tens) should be selected to obtain a relatively low DNRD without the financial cost exceeding its upper bound. Towards this problem, at least two questions should be answered. One is which locations should be selected as the actual location for placing authority servers? The other is how many authority servers should be deployed on each selected location. Here, we assume the authority servers are homogeneous and thus have the same processing capacity. And multiple even tens of authority servers can be deployed at the same location for the convenience on DNS operation and zone file synchnization.

To answer these two questions, at least the following three kinds of data should be provided. The first is those relating to the DNS recursive servers, such as RTT between those potential locations and DNS recursive servers and the query rate of each DNS recursive servers. The second is those relating to the financial cost, like the bandwidth price, the electricity price, the room rent price, equipment maintenance price on each potential location and as well as the authority server purchase price and the total financial budget. The third is those relating to the authority server, such as the capacity of each authority server (whose unit can be handled queries per minute or hour) and the processing latency which relates to the QPS of the authority server. Since the number of recursive servers is very large, it is better to only choose a small part of recursive servers according to some policies (such as randomization or only choosing top N recursive servers) and just to measure the RTT between those selected recursive servers and potential locations.

With all above input data, the authority server placement approach should output one specific authority server placement scheme with lowest DNRD on condition that the actual financial cost is within its

budget. In fact, the authority server placement problem is one kind of optimization problem and thus some approximate optimization algorithm such as simulated-annealing algorithm [SA] can be used to solve this problem.

### 3.2. Optimization goals

Fundamentally, the goal of the authority server placement is: 1). Minimizing the average DNRD on the condition that the financial cost does not exceed its upper bound; 2). Minimizing the maximal DNRD with the financial cost within its upper bound.

#### 3.2.1. Minimizing the average Domain Name Resolution Delay

Average DNRD is calculated by averaging all DNRDs between authority servers and recursive servers which here only refers to those selected ones but not all the recursive servers just as mentioned before. The efficiency of a given DNS zone can be evaluated by its average DNRD. Specifically, lower the average DNRD is, better the efficiency is. Theoretically, when the financial cost is given, there is at least one DNS authority server placement scheme making the average DNRD be the minimal.

#### 3.2.2. Minimizing the maximal Domain Name Resolution Delay

Minimizing the average DNRD may shorten the DNRD of a large part of recursive servers but may prolong that of a small part of recursive servers with a high probability, so this strategy (namely Minimizing the average DNRD) does improve the efficiency but may not obtain good fairness. Towards this issue, the goal of authority server placement can be transferred from minimizing the average DNRD into minimizing maximal DNRD. Then the difference between the largest and smallest DNRD experienced by recursive servers is minimized, which means the recursive servers get almost the same DNRD and thus the fairness is achieved.

### 3.3. Authority server placement algorithm

#### 3.3.1. Best Efficiency Server Placement Algorithm

The input of this algorithm is as follows:

1. The data relating to DNS recursive servers:

- 1). The RTT between potential locations of authority servers and selected recursive servers. Here, the RTT data can be obtained through the network measurement technology like PING or TRACEROUTE [RFC4560].

2). The average query rate (whose unit can be queries per minute or hour) of each recursive servers, which can be obtained from the log of DNS recursive servers.

2. The data relating to financial cost:

- 1). The bandwidth price of each potential location.
- 2). The electricity price of each potential location.
- 3). The room rent price of each potential location.
- 4). The equipment maintenance price at each potential location.
- 5). The price of one authority server.
- 6). The total financial budget.

3. The data relating to authority servers:

- 1). The processing capacity of the authority server, like maximum QPS of one authority server.
- 2). The processing latency which relates to the QPS of the authority server.

The output of this algorithm is as follows:

- 1). The potential locations that is selected as the actual location to place authority servers;
- 2). The number of authority servers placed on each selected authority server location.
- 3). The recursive servers that each authority server serves.
- 4). The bandwidth should be purchased for each actual location of authority servers.

The methods used for solving this algorithm:

Enumeration can be used to generate the best output; however, enumeration will lead to very high computational complexity which make this intuitive idea impossible.

Some approximate optimization algorithm such as simulated-annealing algorithm [SA] can be used to obtain an output (that may not be the best one but we can accept) with a much lower computational

complexity compared with the enumeration. However, the performance of this kind of algorithms like simulated-annealing algorithm is not as good as we expected, thus the improvement of such algorithms is highly needed.

In fact, the problem of DNS authority server placement is NP-hard. For instance, the computational complexity of selecting 50 locations from 1000 potential locations is about  $C(1000, 50)$  which is as large as  $10^{64}$ !

### 3.3.2. Best Fairness Server Placement Algorithm

Best efficiency means the maximal DNRD is minimized. So the object of this algorithm is to make the maximal DNRD be as small as possible on condition that the financial cost does not exceed its upper bound. The input and output of this algorithm is the same as the above algorithm except the optimization object of this algorithm is to make the maximal DNRD be the lowest one. And of course some approximate optimization algorithms like simulated-annealing algorithm [SA] can be used to reduce the computational complexity.

### 3.4. Discussion

Above two server placement algorithms are applicable for those three mentioned scenarios. In the scenario two, only the location of one part of authority servers but not all needs to be fixed while in scenario one and three, the location of all authority servers needs to be fixed. So the algorithm used in scenario two can be seen as one specific case used in scenario one and three. Specifically, adding some restriction conditions to the algorithm used in scenario one and three can form the algorithm used in scenario two. In a word, the input and output of algorithms for these three scenarios are exactly the same, though the method of solving server placement algorithm in scenario two is a little different from that in scenario one and three.

## 4. IANA consideration

This document does not call for changes or additions to any IANA registry.

## 5. Security considerations

TBD.



## 6. References

### 6.1. Normative References

- [RFC1035] Mockapetris, P., "Domain names - implementation and specification", STD 13, RFC 1035, November 1987.
- [RFC4033] Arends, R., Austein, R., Larson, M., Massey, D., and S. Rose, "DNS Security Introduction and Requirements", RFC 4033, March 2005.
- [RFC4034] Arends, R., Austein, R., Larson, M., Massey, D., and S. Rose, "Resource Records for the DNS Security Extensions", RFC 4034, March 2005.
- [RFC4035] Arends, R., Austein, R., Larson, M., Massey, D., and S. Rose, "Protocol Modifications for the DNS Security Extensions", RFC 4035, March 2005.
- [RFC4560] Quittek, J. and K. White, "Definitions of Managed Objects for Remote Ping, Traceroute, and Lookup Operations", RFC 4560, June 2006.

### 6.2. Informative References

- [BROWSE] "A DNS number for faster browsing", December 2009, <<http://www.infoq.com/news/2009/12/Public-DNS-Google>>.
- [SA] "simulated-annealing algorithm", December 2009, <[http://en.wikipedia.org/wiki/Simulated\\_annealing](http://en.wikipedia.org/wiki/Simulated_annealing)>.
- [Web.Latency] Flach, Tobias., Dukkipati, Nandita., and Andreas. Terzis, "Reducing Web Latency: the Virtue of Gentle Aggression", September 2013, <Proceedings of ACM SIGCOMM 2013>.

## Authors' Addresses

Guangqing Deng  
CNNIC  
4 South 4th Street, Zhongguancun, Haidian District  
Beijing, Beijing 100190  
China

Phone: +86 10 5881 3430  
Email: dengguangqing@cnnic.cn

Ning Kong

CNNIC

4 South 4th Street, Zhongguancun, Haidian District

Beijing, Beijing 100190

China

Phone: +86 10 5881 3147

Email: nkong@cnnic.cn

Sean Shen

CNNIC

4 South 4th Street, Zhongguancun, Haidian District

Beijing, Beijing 100190

China

Phone: +86 10 5881 3038

Email: shenshuo@cnnic.cn