域名镜像服务器部署分析

王 伟,李晓东,孙国念
WANG Wei,LI Xiao-dong,SUN Guo-nian

中国互联网络信息中心,北京 100080
CNNIC,Beijing 100080,China
E-mail:wangwei@cnnic.cn

WANG Wei,LI Xiao-dong,SUN Guo-nian.Analysis of DNS mirror server deployment.Computer Engineering and Applications,2008,44(7):161-163.

Abstract: DNS is one of the most important facilities on Internet,and DNS Anycasing is an important method to improve the security,stability and resolution performance of DNS.This paper proposes a method to optimize the DNS Anycasting deployment, based on the DNS operation mechanism and root DNS testing data in China.
Key words: DNS;Anycasting;address selection algorithm;correlation coefficient;regression analysis;clustering

摘 要:DNS是互联网最为重要的基础设施之一,DNS镜像技术是提升DNS系统安全性、稳定性和解析性能的重要方法。针对如何采用镜像技术优化DNS部署的问题,基于DNS的工作机制,以DNS根镜像服务器的实测数据为例,进行具体分析,给出一种实施方法。
关键词:DNS;Anycasting;地址选择算法;相关系数;回归分析;聚类
文章编号:1002-8331(2008)07-0161-03 **文献标识码:**A **中图分类号:**TP393

1 域名系统介绍

Internet域名服务系统(DNS)是一种分布式的等级制查询服务,用以在域名和互联网协议(IP)地址间进行翻译转换。Internet上的所有数据包和路由都基于IP地址,因此,DNS起到IP层与应用层间的桥梁作用[1]。

DNS的域分为不同等级。一般说来,顶级域包括如com、net、org的通用顶级域(gTLDs)和如cn、us等的国家顶级域(ccTLDs)。在顶级域之上,是DNS的根服务器,存储DNS体系中最高层次的zone file供各通用顶级域和国家顶级的记录信息。

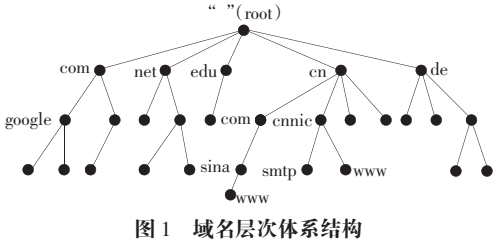

图1 域名层次体系结构

域名解析的工作原理和根服务器所起到的作用如下:

(1)客户机提出域名解析请求,并将该请求发送给本地的域名服务器(递归服务器)。

(2)当本地的域名服务器收到请求后,就先查询本地的缓存,如果有该纪录项,则本地的域名服务器就直接把查询的结果返回。

(3)如果本地的缓存中没有该纪录,则本地域名服务器就把请求发给根域名服务器,然后根域名服务器再返回给本地域名服务器一个所查询域(根的子域,如CN)的顶级域名服务器的地址。

(4)本地服务器再向查询返回的域名服务器(权威服务器)发送请求,收到该请求的服务器查询其数据库,并返回与此请求所对应的资源记录(下级域名服务器的地址或者域名所对应的IP地址等)。本地域名服务器将返回的结果保存到缓存。

(5)重复(4),直到找到正确的纪录。

(6)本地域名服务器把返回的结果保存到缓存,以备下一次使用,同时还将结果返回给客户机。

2 DNS镜像服务器

DNS协议本身在设计时,就考虑到了DNS的可用性问题,设定每个域的权威服务器可以有多个,递归服务器以一定的遍历顺序访问这些权威服务器,如果第一个失败,则顺序访问第二个,直到访问成功为止。通过这样的冗余协议设计,除非某域的所有权威服务器全部失效,对该域的DNS解析才会失败,这极大地提高了DNS的可用性。

但是,由于DNS协议本身数据单元的长度限制,上述多个权威服务器的数量不能多于13个,这极大地限制了这一冗余设计思路的实施,这种部署数量上的限制对根服务器、gTLD和ccTLD来说尤其突出。

作者简介:王伟,男,博士,副主任,主要研究方向为DNS运行;李晓东,男,博士,副研究员,技术总监,主要研究方向为DNS运行及下一代互联网地址资源技术;孙国念,男,高工,主要研究方向为DNS运行。
收稿日期:2007-06-29 **修回日期:**2007-10-09

DNS 服务器不论对全球互联网还是对单独一国来说,都是重要而关键的基础设施,DNS 服务器一直以来都是分布式拒绝服务攻击(DDOS)的目标。以根服务器为例,全球共有 13 个,已经达到了协议分布的最大值,但 2002 年 10 月发生的 DDOS 攻击直接导致根服务器的性能下降,证明要构建一个强壮、安全、高可用的 DNS 体系,13 个服务器的数量是远远不够的。

为此,根服务器管理组织采用了 Anycasting 镜像技术启动了全部 DNS 根镜像服务器全球部署计划(目前全球范围内共有根镜像服务器 111 个,详情可见 www.root-servers.net),其他 gTLD 和 ccTLD 的管理机构,也纷纷采用镜像技术优化自己管辖的 DNS 系统。

“Anycasting”是一项将单个服务器复制并分布部署到多点(即镜像)的技术,所有复制的服务器对外都是同一个 IP 地址,并包含同样的 DNS 记录数据。

Anycasting 技术的内容首先出现在 RFC1546 中[2],将 Anycasting 技术应用于 DNS 服务出现在 RFC3258 中[3]。RFC3258 描述了如何以同一 IP 配置不同权威服务器,路由选择将 DNS 的访问流量引导到网络拓扑最近的服务器上去。

使用镜像来实 DNS 部署能够解决一系列问题,具有如下意义:

(1)提高 DNS 整体解析性能:各区域 DNS 镜像将提高各自所在区域 DNS 的解析成功率和解析速度。

(2)提高 DNS 整体抗攻击能力:从本地发起的攻击只能到达本地镜像,全球其它镜像是不会受到影响的。这使得 DNS 的整体架构更富有弹性。

(3)提高紧急响应能力:部署镜像的一个非直接好处是在发生攻击时,更利于锁定攻击发起源,利于互联网管理者辨别和采取措施。越快定位攻击源,能越早制止攻击。

(4)提供域名信息安全的自给能力:镜像的建立,可保证 DNS 体系在一定区域内的完整部署。即使在因人为事故、自然灾害、突发战争等因素导致区域网络中断时,该区域内部镜像仍然能提供持续的域名解析。

(5)减少国际链路的开销:从国家规模减少路由器和链路资源的开销。由于 IP 路由协议按最近路径传递数据包,本地镜像根将减少占用昂贵国际链路的 DNS 流量。这一点对发展中国家或孤立地区尤其有益。

3 镜像技术对 DNS 访问性能提升的技术分析

本章以 DNS 根服务器为例,通过测试和分析,给出镜像技术提升 DNS 性能的原因,为进一步给出 DNS 镜像部署方法提供理论依据。

根据 DNS 的最初理论,对某域的多个权威服务器的访问应是随机的,即每个服务器得到的访问概率应是相同的。但事实上,根据 2004 年 CNNIC 在全国范围内组织的测试中,通过在 8 家 ISP 放置 DNS 测试机,测得实际上国内对各根服务器的访问性能和访问概率是不等的(见表 1)。

从到达百分比来看,对 F 根的访问最多,这是因为当时中国电信建立的 F 根镜像,而高版本 BIND 的地址选择算法[4],会根据对根的访问速度进行优化排序,访问速度最快的根得到的 DNS 请求最多。

表 1 测试结果中,F 根镜像访问速度最快,其得到的访问也最多,这是和理论相符的。进行统计分析,取到达各根时间倒数为序列 A,取访问百分比为序列 B,计算它们的相关系数[5],得:

表 1 某 ISP 访问根服务器的性能

Root	成功率/%	到达百分比/%	平均时间
A	87.56	3.05	302
B	69.75	0.01	200
C	71.55	1.80	330
D	93.49	4.90	252
E	89.07	1.06	247
F	95.56	47.76	26
G	89.69	1.02	262
H	90.02	1.11	403
I	82.01	1.25	263
J	79.22	5.63	136
K	83.74	0.51	523
L	92.25	5.83	241
M	62.79	17.76	93

ρ=0.981

近似线性的相关系数表明,访问时间越小,访问频率越高,这证明了“访问时间”与“访问百分比”之间的反比关系。依此方法,分别采用其他 7 家 ISP 的测试样本(见表 2),仍然可以得到上述相关特性的结论,证明此相关性的普遍性。

表 2 7 家 ISP 访问根服务器时间倒数与访问百分比的相关系数统计

	ISP1	ISP2	ISP3	ISP4	ISP5	ISP6	ISP7
ρ	0.874 2	0.993 2	0.994 6	0.845 1	0.984 1	0.928 8	0.990 4

因此,可以相信,若进一步提高对某特定 DNS 服务器 X 的访问速度,将等效地提高对 X 的访问频率,最终,随 X 所占访问比的增加,X 访问速度的提高将缩短 DNS 的整体访问时间。

以 ISP5 为例:由于 ISP5 在 2004 年下半年提高了与电信的带宽,因此其到达电信 F 根镜像的速度由 12 ms 缩短到 4 ms,相应地,对 F 根的访问比例从 60%提高到 90%以上,进而提高了 ISP5 访问整个根 DNS 的成功率和平均时间。

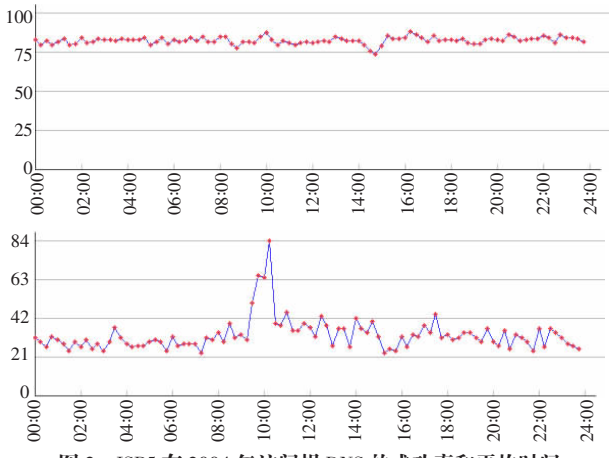

图 2 ISP5 在 2004 年访问根 DNS 的成功率和平均时间

综上,由于 DNS 协议中“优者先得”的选择机制,某域 DNS 镜像服务器的建立,可有力提升用户对该域的 DNS 访问速度和 DNS 访问成功率。

4 DNS 镜像服务器的部署分析

了解了 DNS“优者先得”选择机制对 DNS 镜像服务器的重要性后,仍然以 DNS 根镜像为例,分析如何更好地推进 DNS 镜像服务器的部署工作。

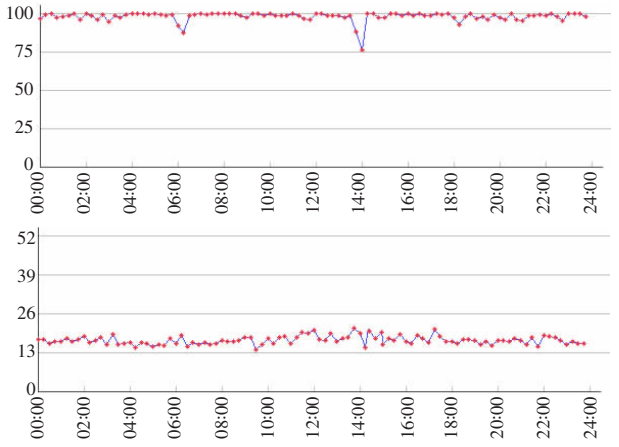

图 3　ISP5 在 2005 年访问根 DNS 的成功率和平均时间

一区域内对某域 DNS 镜像服务器的部署应遵循一个较优的方法，以尽量小的开销和投入，换得尽可能大的性能提升，即：此区域内网络各点对该域 DNS 的访问性能相差不大，均衡最优。

以根镜像为例，截止 2007 年 1 月，我国已在国内建立了三个根镜像（*F*、*I*、*J*），但目前的三个根镜像，并不能完全实现理想的根镜像服务器应达到的效果。

为说明这点，于 2007 年初再次进行了对 DNS 根服务器的访问效率测试，发现，不同根镜像对不同 ISP 的辐射效果是不一样的，不同 ISP、甚至不同地区用户得到的根镜像效率提升是不同的。统计结果如表 3。

可以看出，各 ISP 依然存在不同比例地对国外根服务器的访问，而这种比例对某些 ISP 来说还相当大，并没有完全起到提升国内网络性能和减少国际链路开销的目的。这是由下述可能原因造成的：

（1）各个网络对位于本网的根镜像查询时间较短，但鉴于互联互通基础较差，导致跨网查询性能不佳；

（2）根镜像的接入网络规模庞大，无法辐射到各个角落，平均性能不佳；

（3）根镜像的接入网络业务繁忙，造成实事上的网络拥塞，影响到根镜像性能（成功率和查询时间）的进一步提高。

一区域内某域 DNS 镜像服务器的部署数量和部署地点，是一个复杂的问题，既涉及到该区域网络安全的战略纵深规划，又涉及区域内部的 ISP 平衡和地区平衡，最终影响的是互联网用户终端 DNS 访问性能。本文中，不就战略规划和区域平衡进行分析，而只以上述根镜像部署问题为例，在根镜像访问性能测试的基础上，给出部署 DNS 镜像服务器的规划方法。

先利用回归分析[6]给出 DNS 测试样本中访问时间和访问百分比之间的精确函数关系：取到达各根时间为自变量序列 *X*，访问百分比为因变量序列 *Y*，采用最小残差法进行函数拟合。以 2004 年 8 家 ISP 的测试样本和 2007 年 7 地区测试样本进行上述拟合，从最小残差和曲线物理意义两方面证明，访问百分比与访问时间之间的函数关系为近似逆函数[7]，如图 4 所示。

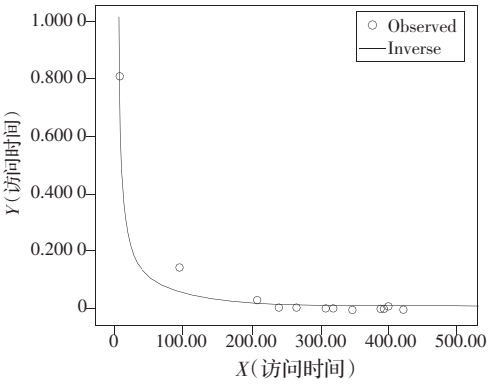

图 4　某 ISP 样本拟合函数曲线图

从减少国际链路开销的角度来看，理想状况下，对国内根镜像的访问比例应不低于 95%。那么，根据样本推出逆函数 $Y=b_0+b_1/x$，可计算出满足访问 Y=95%的 X 值。即，国内镜像访问速度必须小于此 X 数值。

理论上，链路质量越好的 ISP，其满足 95%访问率的访问时间应越小。由于国内各地区 ISP 的国际链路和国内互联传输质量不同，对 DNS 根服务器的访问性能和百分比不同，推算出的逆函数的常数项和回归系数也不同。对 2004 年 8 个样本和 2007 年 7 个样本分别计算，发现 2007 年的 X 值（平均 3.5 ms，最大 20 ms）明显小于 2004 年 X 值（平均 15 ms，最大 50 ms），这表明过去几年中，国内基础网络的建设和国际出口带宽的扩大，使得各 ISP 国际访问性能普遍提高，这很好地解释了 X 值的减小。

得到目前根 DNS 访问的理想性能 X=20 ms 这个指标后，本文提出根镜像部署的两种方法：

方法 1　实测法

进行全国范围的各 ISP 统一测试，以保证各测试点之间的 DNS 互访性能不大于 X 毫秒。

难度：需要在各 ISP 进行全国范围的测试，调动资源多，测试时间长。

优点：以结果为导向，循序渐进，根据测试结果逐步调整根镜像数量和地点。

方法 2　聚类法

如建成完善的国内 DNS 根镜像体系，从 BIND 优选机制来说，用户将访问满足 X 毫秒性能指标的根镜像，即，用户到达根镜像服务器的网络时间在 X 毫秒内（此处忽略 DNS 服务器的处理时间，为百微妙级）。基于此，提出算法如下：

（1）得到各 ISP 的网络拓扑。

表 3　不同 ISP 访问根服务器的百分比　%

ISP	*A*	*B*	*C*	*D*	*E*	*F*	*G*	*H*	*I*	*J*	*K*	*L*	*M*
ISP *a*	0.53	0.57	0.57	0.53	0.59	6.50	0.53	0.53	33.90	53.92	0.57	0.57	0.71
ISP *b*	1.99	2.03	2.47	1.90	2.19	38.87	2.04	1.73	38.07	1.88	2.14	2.57	2.11
ISP *c*	0.58	0.76	1.56	0.64	0.90	5.52	0.50	0.54	0.60	86.28	1.14	0.56	0.44
ISP *d*	6.81	2.20	2.70	1.76	2.84	23.66	2.98	1.65	7.94	28.54	4.71	7.06	7.14
ISP *e*	1.25	1.65	1.50	1.32	2.16	80.67	1.40	1.05	2.28	2.26	1.19	1.42	1.85
ISP *f*	1.95	2.58	2.38	2.29	2.25	25.54	2.40	2.09	49.11	1.21	2.31	2.67	3.24
ISP *g*	1.21	1.34	1.56	1.46	1.24	43.45	1.14	1.03	1.10	43.15	1.28	1.10	0.94

（下转 167 页）