

互联网架构关键资源可扩展研究*

刁玉平¹, 廖 铭², 刁永平³

- (1. 广东商学院信息学院, 广东 广州 510320;
2. 中国移动广州分公司, 广东 广州 510630;
3. 中国电信广州研究院, 广东 广州 510630)

摘 要: 互联网发展中需要解决的最为关键的两个问题就是互联网自治问题和互联网可扩展问题。结合互联网现实, 架构可扩展性研究提出了自主自治的可扩展域名体系和自治网络系统间的域名解释机制和层次结构; 提出了自治网络系统间的双向动态源/目的 NAT 可扩展的寻址方式; 提供了可扩展使用的关键网络资源如域名、地址等。可以最小的代价、无过渡期实现互联网的自主自治、可扩展。

关键词: 自治互联网; 可扩展 IP 网络; 域名; IP 地址

中图分类号: TP393.03 **文献标志码:** A **文章编号:** 0529-6579(2011)06-0053-05

Research on Key Resources Extension of Internet Architecture

DIAO Yuping¹, LIAO Ming², DIAO Yongping³

- (1. Information Science School of Guangdong University of Business Studies, Guangzhou 510320, China;
2. China Mobile Guangzhou Branch, Guangzhou 510630, China;
3. China Telecom Guangzhou Institute, Guangzhou 510630, China)

Abstract: The two key issues of developing internet are internet autonomous problem and internet extensible problems. With the reality of internet, extensible architecture research proposes layer structure between autonomous network systems independent extensible autonomous DNS architecture and DNS mechanism. It also put forward the extensible addressing manner with bidirection dynamic source/destination NAT mechanism between autonomous network systems, and provides enough key network resources such as domain name, IP address for extensible usage. Therefore, Autonomous Extensible Internet (AEIP) could be realized in least cost and without transition period.

Key words: autonomous internet; extensible IP network; DNS; IP address

从 20 世纪 70 年代至今, 互联网已经在世界范围内得到了巨大的发展。互联网创造了一个奇迹, 但其本身却存在着日益突出的问题, 给人们提出了极大的挑战, 其中最为关键的两个问题就是互联网自治的问题和互联网可扩展的问题。

互联网自治问题的实质是 IP 域名问题。互联网系统的核心是它的域名根服务器, 尽管现在网络很发达, 但实际支撑互联网运转的根服务器数量相

当有限。互联网始于美国, 因此这些根服务器一直由美国控制。通过控制根服务器, 可以达到控制全球各种域名的目的, 而且还可以对其他国家的网络使用情况进行监控。因此, 要想维护国家互联网安全, 必须打破互联网的垄断控制权, 拥有独立的根域名服务器, 这是全球众多国家的一致看法^[1-6]。

互联网可扩展问题的实质是 IP 地址问题。任何人在互联网发展初期都无法想象如今的网络会发

* 收稿日期: 2011-06-23

基金项目: 广东省科技计划资助项目 (2011B080701074)

作者简介: 刁玉平 (1973 年生), 女, 讲师; E-mail: gsddd@139.com

展到如此大的规模。互联网的迅速发展,使 IP 地址资源严重匮乏。业界提出的私有专网、动态分配、CIDR 和 NAT 技术仅可以减缓地址资源枯竭的速度,而采用 IPv6 还有诸多问题,进展极为缓慢,无法及时满足互联网快速发展的需求^[7-12]。

1 新型网络模型

围绕自治和可扩展这两个关键问题的解决,网络互联必须而且必将进入以现有互联网为基础,架构可扩展,具有自治架构,具有域名自主性,具有可复用地址空间的新时代。

可扩展的新型网络应该是多系统的网络互联架构,是以自治网络为单位系统的系统扩展,其中的每个单系统具有类互联网的自治架构。架构可扩展性需要考虑新型网络互联的体系架构、提出自治网络系统间的域名解释机制和层次结构,自治网络系统间的寻址方式和路由机制。

自治网络单系统具有独立的体系架构、域名解释机制和层次结构、寻址方式和路由机制,可以不依赖于外部系统而独立运作。自治网络具有自主管理的域名层次结构和域名规划;具有自主控制的根域名服务器和域名服务体系;具有自主可复用的网络地址空间;具有自主的网络寻址及路由拓扑规划。每个自治网络都具有类似现有互联网的整套体系和机制,关键网络资源如域名、地址等在不同自治网络中可以复用。自治可扩展互联网与现有互联网的不同之处是需要考虑自治网络之间的域名解释、寻址路由的互联互通。

2 可扩展自治域名体系

自治互联网 AIP 技术提供一种不需要改变互联网的底层支持技术和现有协议,就可以拥有完全独立自主的根域名服务器,实现互联网自治的技术。其中具有域名自主性的自治网络单系统的域名及域名解释服务是完全自主自治的。自治网络间的网际域名解释在各自域名服务体系控制下通过域名网关平等协商解释。同时, AIP 具有安全可扩展的架构,所需改造很小,可以通过新建、升级、改造或者单边行动的方式来实现互联网的自治^[13]。

自治互联网可扩展的域名体系主要规则是:

- 1) 每个自治 IP 网络本身具有完整的整套域名系统,支持本自治 IP 网络内部域名地址转换;本网网内域名解释按照传统方式。
- 2) 每个自治 IP 网络都具有区别于其它自治 IP 网络的名称编号,并以此作为本网域名被外部引用

时的缺省的网络域名后缀。任何一个网络节点的网内域名加上本自治 IP 网络的网络域名后缀就形成了该网络节点的网际域名。

3) 跨自治 IP 网络通信时,访问外网节点必须使用其网际域名,即在该外网节点网内域名后面添加其缺省的网际域名后缀。

据此,可以设计出如图 1 所示的自治互联网域名层次结构。可以看到,每个自治 IP 网络,如 A、B……,本身具有完整的整套域名系统并且互不干涉,不同自治 IP 网络内节点域名甚至可以重复如域名为 www. yahoo. com。跨自治 IP 网络通信时,需要采用网际域名,如 www. yahoo. com. A 就表示是自治 IP 网络 A 中的域名节点。为此,在每个自治域名层次结构树中都增加 ex (i) 的一级域名,以映射代表本自治 IP 网络可以通达的其它自治 IP 网络域名树。如当 ex (i) = B 时,表示可以通达的其它自治 IP 网络 B。

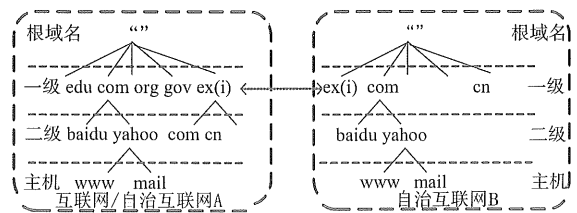


图 1 自治互联网域名层次结构图

Fig. 1 Autonomous internet DNS layer architecture

根据自治互联网的域名体系,可以构造出自治互联网的网络架构,如图 2 所示。每个自治互联网的根服务器,负责该自治网的所有的互联网域名解析工作。自治互联网中其它合法的 DNS 服务器默认都是指向这些根服务器。

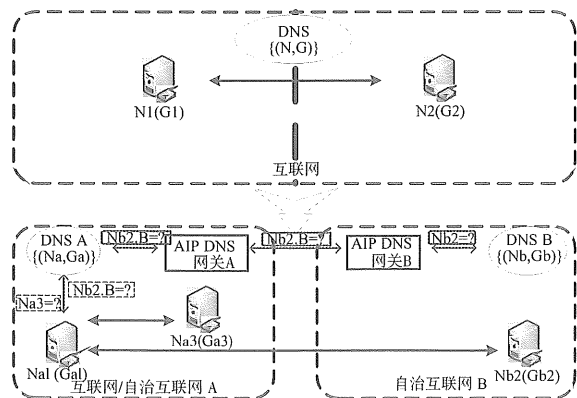


图 2 自治互联网架构

Fig. 2 Autonomous internet architecture

由于每个自治 IP 网络都相当于现在传统的互联网, 其内部域名解释和通信都不会有改变。唯一的改变是跨自治 IP 网络通信时, 域名需要增加目的自治 IP 网络的网际域名后缀。所以, 在每个自治 IP 网络中会增加一个称为“自治 IP 网络 DNS 网关”的设备 AIP DNS GW 以支持跨自治 IP 网络的域名解释, 即负责在每个自治域名层次结构树中增加的所有 $ex(i)$ 顶级域名的解释功能。一方面, 把本网的跨自治 IP 网络域名解释请求转发到目的自治 IP 网络, 接收到域名解释结果后返回给请求方; 另一方面, 接受外网对本自治 IP 网络的跨网域名解释请求, 在依据传统的方式得到本网 DNS 域名解释结果后反馈给外网请求方。

3 可扩展寻址机制

3.1 网络可扩展设计

自治互联网不仅可以解决互联网自治的问题, 而且分布式的架构使其架构上可扩展, 可以按需增加若干自治互联网实体, 即实现网络上可扩展。但是现行 IP 网络地址匮乏, 即将消耗分配完毕, 所以有必要进一步实现网络地址上的可扩展, 这样才能真正实现互联网的可扩展。基于互联网网络寻址及地址空间现状, 最合理的办法是单系统内部地址可复用、多系统互联地址统一规划保留。通过地址的分配调整, 可以使每个自治网络具有自主可复用的足够网络地址空间; 通过网络互联地址空间的规划使用, 可以按需扩展增加需要的自治网络数量, 使地址总数成倍数增加。所以, 可以在互联网寻址方式和路由机制保持不变的情况下, 通过 IPv4 地址空间的复用扩展, 保证提供足够使用的 IP 地址, 从而解决了现行互联网 IP 地址消耗殆尽的困境及由此带来的网络大规模升级、过渡及所有网络设备的更换更新等棘手繁杂的诸多问题^[14]。

这里考虑一种基于自治互联网架构, 采用本地网络地址为主、全局网络地址转换的技术来实现 IP 网络可扩展的方法-自治可扩展互联网 AEIP NAT。

针对本地网络地址具有动态的、一定集线比例的、双向的跨自治网络通信, 首先要求数据包的目的节点地址采用动态分配(特殊情况下可以少量的固定分配)的特定目的网络的全局网络地址(如通过网际域名解释请求获得), 这样源节点发出的跨网数据包可以首先路由到源 IP 节点所在网络的网络互通网关 AEIP NAT GW A 上; 该源网络的 AEIP NAT GW A 将为跨网发出的数据包进行源

地址 NAT 转换; 跨网数据包转发到目的 IP 节点所在网络的网络互通网关 AEIP NAT GW B 上时, 该目的网络的 AEIP NAT GW B 将为跨网进入的数据包进行目的地址 NAT 转换, 然后转发该数据包进入到目的网络, 并最终到达目的节点。全局网络地址可以由每个自治可扩展互联网 AEIP NAT 的 DNS GW 负责动态分配, 并保存和维护相应的全局网络地址-本地网络地址对表 (G, L) 和域名-全局网络地址对表 (N, G) 。任何一个全局网络地址-本地网络地址对如 (G_a, L_a) 将在其动态存活时期内发给 NAT GW 以进行 AEIP NAT 网际的地址转换。任何域名-全局网络地址对如 (N_b, G_b) 将在其动态存活时期内提供 DNS GW 进行网际的域名解释全局网络地址查询。

3.2 可复用地址空间

针对基于自治网络单系统的新型多系统可扩展网络架构, 可以设计一套可复用地址空间, 使互联网在系统扩展时总有一套现成的地址空间提供给新增加的自治网络单系统, 从而彻底解决互联网网络地址短缺问题, 扫除互联网扩展的障碍。

可复用地址空间包括三部分: 本地网络地址 L、全局网络地址 G 和私有网络地址 P。私有网络地址即原有的私有 IP 地址, 仍然保留, 作为自治网络内部可复用的私有网络地址子空间。

这三种类型的网络地址全部采用现有互联网的 32 位地址空间 C 适当划分而成, 而其中的全局网络地址子空间是不可复用的。现有互联网地址空间大小 $C = 2^{32}$, $P = 2^{24} + 2^{20} + 2^{16}$ 。故有

$$C = L + G + P \approx L + G \quad (1)$$

基于每个自治网络, 引入“本地网络地址”的概念, 它包括了互联网大部分的地址空间, 并且可以在不同的自治网络中复用。一般地, 网络节点都仅配置分配本地网络地址, 并在自治网络内部直接通信。这样, 每个自治网络都有约几十亿个 IP 地址; 而且可以根据需要拥有多个自治网络, 使可用 IP 网络地址成倍数增加, 提供“任何”可能需要的 IP 地址数量。而且就每个自治网络而言, 基本上与现有互联网现状一致, 无需升级改造现有节点, 内部通信独立于其它自治网络。假设需要 n 个这样的自治网络, 那么扩展后的地址空间总容量可以达到

$$C' = nL \quad (2)$$

不同的自治网络单系统之间采用“全局网络地址”互通。全球唯一的全局网络地址范围由全球协商统一规划管理(特殊情况下, 可以由互通的自治网

络间协商确定), 不同自治网络系统分配不同的全局网络地址范围 (可以采用部分现有或者保留的公有 IP 地址)。由于跨网通信量相对于整个自治网络具有较小的比例, 所以每个自治网络系统需要采用的全局网络地址可以按照需要的比例提供。

针对每个拥有本地网络地址容量为 L 的自治网络 (或者自治网络集合) i , 设其与外部系统互通的集线比为 Tr_i , 则所需的全局网络地址为 G_i , 则:

$$Tr_i = \frac{L}{G_i} \quad (3)$$

$$G = \sum_{i=1}^n G_i = \sum_{i=1}^n \frac{L}{Tr_i} \quad (4)$$

$$C \approx L + \sum_{i=1}^n \frac{L}{Tr_i} = (1 + \sum_{i=1}^n \frac{1}{Tr_i})L \quad (5)$$

根据经验, 一般可以假设自治网络的规划设置具有相似的集线比, 即 $Tr_i = Tr$, 则由上式可得:

$$C = (1 + \frac{n}{Tr})L \quad (6)$$

从而可得本地网络地址空间 L 及扩展地址总容量 C' 与现有 IP 地址空间 C 之间的关系:

$$L = \frac{Tr}{Tr + n} \cdot C \quad (7)$$

$$C' = nL = \frac{Tr \cdot n}{Tr + n} \cdot C \quad (8)$$

根据公式 (7)、(8) 对于给定的 $Tr, n, C'/C, L/C$ 是另一个变量 n, Tr 的增函数。我们分别尝试计算 $n = 1, 10, 10^2, 10^3, 10^4, 10^5, 10^6$ 及 $Tr = 1, 10, 10^2, 10^3, 10^4, 10^5, 10^6$ 时扩展地址总容量、本地网络地址空间与现有 IP 地址空间 C 的比值, 以衡量地址扩展的容量增幅以及自治网络单系统的本地网络地址空间大小。计算结果如表 1 和表 2 所示。

我们可以对这两个表格引入一些限定条件以排除不适合的扩展选项。首先, 由于世界上民族国家人口最高已经达到 10 亿的数量级别, 所以考虑单个自治网络的本地网络地址 L 达到 20 亿以上, 即 $L \geq 0.5C$; 其次, 为了获得足够大的扩展地址总容量, 要求至少 $C' \geq 5C$; 最后, 由于各个自治互联网单系统之间需要合理的集线比互通, 根据经验值, $Tr = 10^6$ 显然不合理。这样, 把两个表格中需要剔除值的对应单元格进行联动填充成灰黑色。同时对表 1 和表 2 的分析可见, 通过限定条件排除不适合的扩展选项后, 合理的网络扩展选项范围是收敛的。

进一步, 根据初步分析结果, 我们选取 $n = 10^3$ 及 $Tr = 10^3$ 为中间取值, 以 200 为步长向两边取值, 从而获得合理的自治可扩展互联网可扩展地

址空间取值样例如表 3 和表 4 所示。同时对应地给

表 1 扩展地址总容量倍数收敛

Table 1 Times convergence of total extended address space

Tr	n						
	1	10	10^2	10^3	10^4	10^5	10^6
1	0.5	0.9	1.0	1.0	1.0	1.0	1.0
10	0.9	5.0	9.1	9.9	10.0	10.0	10.0
10^2	1.0	9.1	50.0	90.9	99.0	99.9	100.0
10^3	1.0	9.9	90.9	500.0	909.1	990.1	999.0
10^4	1.0	10.0	99.0	909.1	5 000.0	9 090.9	9 901.0
10^5	1.0	10.0	99.9	990.1	9 090.9	50 000.0	90 909.1
10^6	1.0	10.0	100.0	999.0	9 901.0	90 909.1	500 000.0

表 2 本地地址容量比率收敛

Table 2 Quota convergence of local address space

Tr	n						
	1	10	10^2	10^3	10^4	10^5	10^6
1	0.500	0.091	0.010	0.001	0.000	0.0	0.0
10	0.909	0.500	0.091	0.010	0.001	0.0	0.0
10^2	0.990	0.909	0.500	0.091	0.010	0.0	0.0
10^3	0.999	0.990	0.909	0.500	0.091	0.0	0.0
10^4	1.000	0.999	0.990	0.909	0.500	0.1	0.0
10^5	1.000	1.000	0.999	0.990	0.909	0.5	0.1
10^6	1.000	1.000	1.000	0.999	0.990	0.9	0.5

表 3 扩展地址总容量倍数示例

Table 3 Times example of total extended address space

Tr	n						
	400	600	800	1 000	1 200	1 400	1 600
400	200.0	240.0	266.7	285.7	300.0	311.1	320.0
600	240.0	300.0	342.9	375.0	400.0	420.0	436.4
800	266.7	342.9	400.0	444.4	480.0	509.1	533.3
1 000	285.7	375.0	444.4	500.0	545.5	583.3	615.4
1 200	300.0	400.0	480.0	545.5	600.0	646.2	685.7
1 400	311.1	420.0	509.1	583.3	646.2	700.0	746.7
1 600	320.0	436.4	533.3	615.4	685.7	746.7	800.0

表 4 本地地址容量比率示例

Table 4 Quota example of local address space

Tr	n						
	400	600	800	1 000	1 200	1 400	1 600
400	0.500	0.400	0.333	0.286	0.250	0.2	0.2
600	0.600	0.500	0.429	0.375	0.333	0.3	0.3
800	0.667	0.571	0.500	0.444	0.400	0.4	0.3
1 000	0.714	0.625	0.556	0.500	0.455	0.4	0.4
1 200	0.750	0.667	0.600	0.545	0.500	0.5	0.4
1 400	0.778	0.700	0.636	0.583	0.538	0.5	0.5
1 600	0.800	0.727	0.667	0.615	0.571	0.5	0.5

出了曲面图图3和图4, 可以清楚了解可扩展地址空间选项, 并可以根据需要选取相应的 Tr 及 n 从而获得足够而合理的自治可扩展互联网的地址空间。

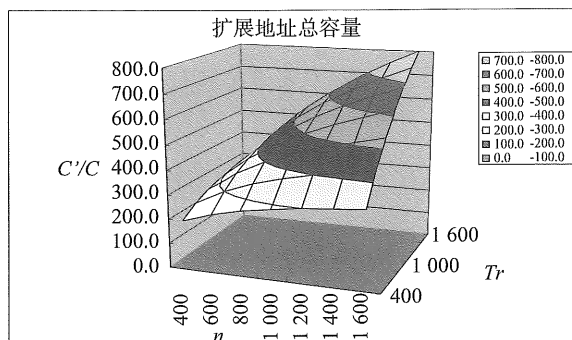


图3 扩展地址总容量倍数变化曲面

Fig. 3 Times curved surface of total extended address space

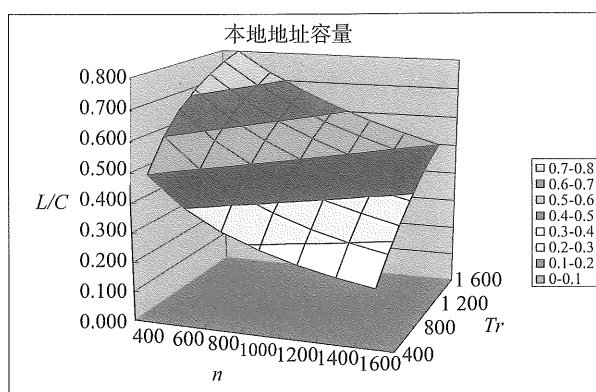


图4 本地地址容量比率变化曲面

Fig. 4 Quota curved surface of local address space

4 结论

对互联网的巨大需求促使人们必须研究解决阻碍互联网发展的“自治”和“可扩展”问题。本研究提出了可扩展的自治域名层次结构和解释机制、可扩展寻址方式; 提供任意可扩展的自治域名体系和足够的互联网地址空间(成千上万倍)。而

且可以即时平滑实现, 没有颠覆性过渡期的需要。期望基于自主知识产权、以最小代价来实现互联网的自治自治及可扩展, 以改变国际互联网的战略格局(可以单边行动实现), 维护国家安全。

参考文献:

- [1] RFC 1034. Domain names-concepts and facilities[S]. 1987.
- [2] RFC 1035. Domain Names: Implementation and Specification[S]. 1987.
- [3] CASTRO S, WESSELS D, FOMENKOV M, et al. A day at the root of the internet[C]. ACM SIGCOMM Computer Communications Review, 2008.
- [4] BROWNLEE N, CLAFFY K, NEMETH E. DNS Root/gTLD performance measurements[C]. Proc Passive and Active Measurement Workshop(PAM), 2001.
- [5] 尉迟学彪, 李晓东, 阎保平, 等. DNS 服务中的 Internet 访问行为测量研究[J]. 计算机工程与应用, 2009, 45(34): 85-88.
- [6] 王峰, 王恩海, 钱华林. DNS 与关键词服务请求分布研究[J]. 计算机工程, 2006, 32(5): 15-17.
- [7] RFC 791. Internet Protocol - DARPA Internet Program Protocol Specification[S]. 1981.
- [8] RFC 1918. Address Allocation for Private Internets[S]. 1996.
- [9] RFC 1518. An Architecture for IP Address Allocation with CIDR[S]. 1993.
- [10] RFC 2663. IP Network Address Translator (NAT) Terminology and Considerations[S]. 1999.
- [11] RFC 2460. Internet Protocol, Version 6 (IPv6) Specification[S]. 1998.
- [12] 林闯, 雷蕾. 下一代互联网体系结构研究[J]. 计算机学报, 2007, 30(5): 693-711.
- [13] 刁永平, 刁玉平, 廖铭. 自治互联网的实现: 中国, 201010277181.4[P]. 2010-12-15.
- [14] 刁永平, 刁玉平, 廖铭. 自治可扩展互联网的地址转换法实现: 中国, 201110031561.4[P]. 2011-07-06.