

基于覆盖网络的应用层选播服务的设计与研究

张园园,余镇危

(中国矿业大学(北京校区),北京 100083)

摘 要: 针对网络层的选播通信存在着可扩展性和无状态服务等根本问题,在现有的应用层选播研究基础上,提出了一种基于覆盖网络的应用层选播的结构,并把这种结构从代理节点的部署和链路的选取上进行求解,并通过实验对此拓扑结构进行了验证.

关键词: 覆盖网络;选播服务;ALAON;代理部署;链路选取

中图分类号: TP393

文献标识码: A

文章编号: 1000-7180(2011)12-0072-03

Research and Design of Application Level Anycast Based on Overlay Network

ZHANG Yuan-yuan, YU Zhen-wei

(China University of Mining & Technology, Beijing 100083, China)

Abstract: In allusion to scalability and stateless connection of network anycast layer, this article is proposed new overlay architecture based on research on application anycast level, which is divided into two aspects: overlay agent node placement and overlay link selection. The results of simulation give verification about the network topology.

Key words: overlay network;anycast communication;ALAON; agent placement;link selection

1 引言

选播是在 RFC1546 中作为一种新的网络通信模型被提出的,它与单播、组播并列称为未来网络的三种通信方式之一.选播是这样定义的:用一个选播地址来标识属于不同节点的多个接口,发送给这个地址的数据包将会被传送到具有该地址标识的、根据路由协议距离度量“最近”的一个接口上.选播(Anycast)可以有效地平衡网络负载,减少路由和链接数量及简化主机配置等网络应用,来提高网络的服务质量(Quality of Service, QoS).它可以在网站镜像、DNS、负载均衡、移动多播,并行计算等领域发挥重要的作用.

起初选播的研究集中在网络层,即通过网络的拓扑结构由路由器来完成选路,具有高效性.但是它本身也存在两个局限性限制了网络层选播的发展:其一,在网络层实现选播过程中可能产生“无状态服

务”问题,由于选播数据包在路由到“最优”服务节点过程中不依赖于以前的数据包这个自然属性,当请求的数据包属于面向连接的应用或者被分段,那么选播服务的效率反而可能低于单播.其二,选播路由表不遵从单播的路由汇聚规则,因为共享某个 Anycast 地址的选播组成员节点可能分散在 Internet 的不同地方.若用单播路由协议对每个 Anycast 地址作为独立的路由表项进行处理,则可能导致路由表的膨胀.这便是 Anycast 存在的可扩展性问题.

因此近年来人们把研究目光转向了应用层.与 IP 选播相比,它具有易于部署和实现,无需修改路由信息,能根据网络动态变化实现服务定制,有较好的灵活性和扩展性等优势.目前最新的研究是关于一类基于代理的应用层选播结构,即选播的路由不再由路由器完成而是通过代理实现最近服务器的选择.但是目前的这类系统主要关注于某种特定的目标,如为了保证端到端的时延,对服务代理的带宽资

源进行优化或者对服务代理进行某种设置来保证特定系统的可靠性等,并没有提出一种效率高、通用性好的基于服务代理的选播结构的构造方法,来保证网络负载的均衡和网络的服务质量. 本文提出一种基于覆盖网的应用层选播 (Application Layer Anycast Overlay Network, ALAON) 构造方法,为选播的实现提供了良好的基础.

2 ALAON 体系结构

2.1 Overlay 技术

Overlay Network 技术是一种构造网络方法,它与特定技术、特定层次无关. 它可在原有物理网络的基础上,按照某种需要构造一个虚拟网络,以此来支持原网络没有或很难提供的功能,并能最大限度地保证与原有网络的兼容性. Overlay Network 无需物理构造,而只需要逻辑融合,即可提供所需的服务,并由此改善下层网络的一些相关属性,而且可以在与现有网络兼容的基础上,较为灵活地解决当前网络体系结构所存在的问题.

2.2 ALAON 的三层结构

在网络体系结构总思想的指导下,用 Overlay Network 的思想来解决选播所面临的问题.

我们把 ALAON 结构分为三层:基础网络层、覆盖网络层和选播应用层. 底层基础网络层为真实物理网络节点(或服务器)和链路,利用现有 TCP/IP 协议族在物理链路上提供基本的物理层传输、网络通信和网络层路由等功能. 中间层为覆盖网络层,其节点为服务代理节点及选播成员,它们之间用虚拟的覆盖网络链路进行连接,生成并维护 Overlay Network 网络拓扑结构,并按照一定的算法实现选播路由. 最上层是选播应用层,为端用户和选播应用程序提供应用接口. 因为本文的研究的内容与关键在于覆盖网络的构造.

本文将 ALAON 覆盖网的构造方案通过自治域内选播和自治域间的选播分层结构来实现:首先,把选播组成员分成若干个自治域,这里的自治域直接采用当前 Internet 为保证单播服务所划分的自治区域概念,在域内部署服务代理(Server Agent)节点,通过以它为根节点,选播成员为叶子节点构成一棵组播树的方式来实现对域内的选播资源的管理;其次,在每个自治域内选举较高处理能力及可靠性的 SA 节点来调度和管理域间资源,从而保证网络的负载均衡和选播服务的质量. ALAON 体系结构拓扑图见图 1.

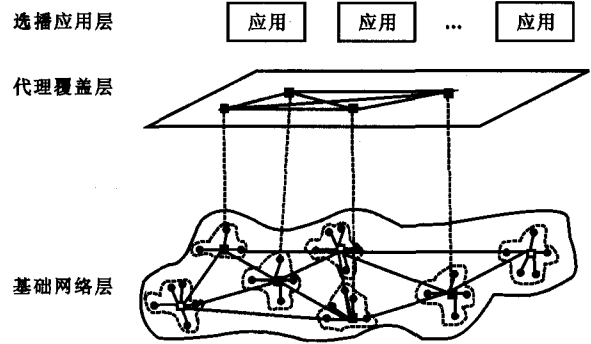


图1 ALAON 体系结构拓扑图

3 ALAON 节点的部署

根据上面提到的 ALAON 的体系结构,其部署重点在于代理节点 P 的数量和位置. 真实的物理网络可用无相图 $G(V, E)$ 来表示,定义客户端节点的集合 $C \subset V$ 及它们的需求 d_c ; 服务器节点的集合 $S \subset V$ 及它们的服务能力 c_s ; 边的权重 $w_e, e \in E$. 实际上就是构造这样一个虚拟代理覆盖拓扑 $G'(P, E')$, 其中 P 为代理节点的集合, E' 为虚拟链路的集合. 由此可见 ALAON 节点部署问题实际上就是一个优化问题,应考虑平衡网络资源利用代价,代理服务器的部署可分解为以下 2 个优化目标:(1) 找到合适的代理服务节点的位置及数量;(2) 优化链路的总流量. 条件 2 不能决定代理节点的位置和维数,但是可以对条件 1 部署的代理节点的效率进行检验.

对任意一个选播系统,设总共需要部署 $M(0 \leq M \leq N)$ 个代理服务节点, P_i 表示节点 i 是否被选为代理节点,对于 $\forall i \in V, P_i \in \{0, 1\}$. $a_i (i \leq N)$ 表示与路由器 R_i 链接的选播组的数量,而两个路由器 R_i 与 R_j 之间的最短距离 $w_{ij} (0 \leq i, j \leq N)$, 从而得到 ALAON 节点部署的目标函数如下.

$$\min \sum_{j=1}^N \left(\sum_{i=1}^N w_{ij} \times p_{ij} \right) \times a_i, \text{ 其约束条件如下:}$$

- (1) 每个代理至少与一个路由器相连, $P_i \leq \sum_{j=1}^N p_{ij}$;
- (2) 为了控制覆盖网的规模,最多部署 M 个代理节点, $\sum_{i=1}^N p_i \leq M$;
- (3) 每一个路由器 R_j 只能与一个代理相连,有 $\sum_{i \in V} p_{ij} = 1$.

由此可见,从目标函数和约束条件看,此问题是整数划问题(integers linear program). 随着网络规模的变化,这种组合优化 NP 难问题,很难采用传统的优

化算法进行求解,这里采用启发式算法进行求解。

4 ALAON 链路的选取

在进行虚拟链路选取时,由于虚拟链路所连接的是端用户,而端用户的不稳定性必然对虚拟链路的可靠性有所影响,因此在进行虚拟链路的选取时首先要考虑链路的可靠性;另一方面,每条虚拟链路都有路由代价,因此在进行虚拟链路的选取时还要考虑到全局网络的路由代价问题。所以,覆盖网上虚拟链路的选取设计关键在于使网络的可靠性要高,并且使网络的路由代价最低。因此,在给定覆盖拓扑维护开销的条件下,选择合适的覆盖链路,对构造高性能的覆盖网络拓扑具有重要意义。设网络流量总费用为 $f(x)$, x_e 代表流经边 e 的流量, w_e 代表边 e 的权重,链路选取的目标函数:

$$\min f(x) = \sum_{e \in E} w_e x_e, \text{约束条件为}$$

$$\sum_{e \in \text{out}(n_i)} x_e - \sum_{e \in \text{in}(n_i)} x_e = b_i, \forall n_i \in V, x_e \geq 0$$

式中, b_i 表示流过节点 n_i 的相对剩余流量,当节点 n_i 代理节点时, b_i 为非零实数;当 n_i 为普通节点时, b_i 为零。在实际选取链路时,我们假设所有的边都是连通的,如若不然,则必能找到两条连通的边来代替不连通的两点。这样链路选取问题可以看作固定费用网络流的问题(Fixed Charge Network Flow Problem),并被证明是 NP 难问题。这里用动态偏坡优化算法(Dynamic Slope Scaling Procedure (DSSP))进行求解。

5 仿真实验及结果分析

本文的实验目标主要是验证两个方面:一是,代理的数量随着代理节点费用的改变的变化情况;二是,代理节点的费用随着链路跳数的增加而变化的情况。利用 GT-ITM 网络生成工具根据 Waxman 模型随机生成节点为 250 的拓扑图作为底层物理网络,覆盖网络的最大代理节点数设为 80。为了验证本文提出的选播代理覆盖结构的优越性,从代理节点数目和链路跳数两方面对此算法和随机生成算法进行仿真,比较两者的性能。

从图 2~3 中,可以看出,随着代理节点的数量以及链路跳数的变化,随机算法和 ALAON 算法的变化有相同的趋势;从图 2 可以得到随着代理节点数量的增多,覆盖层的维护代价增大,但是 ALAON 算法较随机算法有更好的性能;从图 3 分析可得随着网络规模增大,选播服务的代理费用随着链路跳

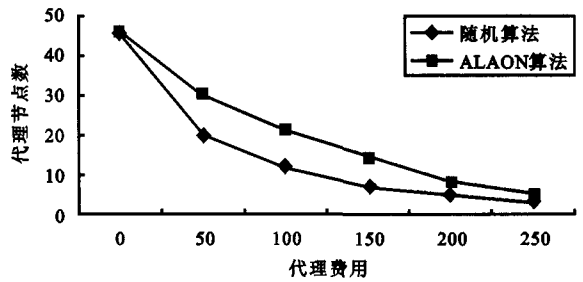


图2 覆盖代理数目代价

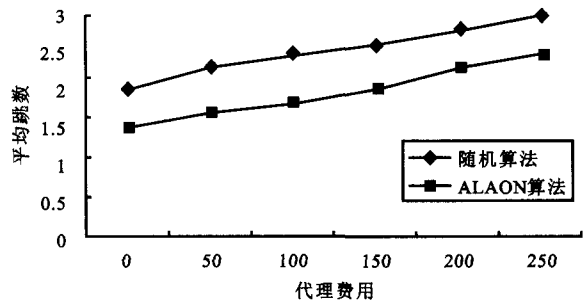


图3 覆盖代理链路代价

数的增加而增加,在链路选取的方法中将多余的边用已有的边代替,这样代理的费用一般会降低。同时,在逐渐增加网络规模过程中,随机算法易受网络规模的影响,而 ALAON 算法的平均跳数变化则相对平稳。

6 结束语

提出了一种新的基于覆盖代理的选播服务模式,解决了以往难以解决的可扩展性、无状态服务等问题。在深入分析选播特点的基础上,将此覆盖代理结构分解开来,从代理节点的部署和链路的选取两方面进行构造,并给出了相应的求解方法,通过实验验证了此结构的性能,为进一步研究选播服务提供了新的思路。

参考文献:

- [1] Animesh Nandi. LALA: location aware load aware overlay anycast[J]. Master thesis of Rice university, 2004 (8): 19-22.
- [2] Stevens T, De Leenheer M, Develder C, et al. ASTAS: architecture for scalable and transparent anycast services [J]. Journal of Communications and Networks, 2007, 9 (4): 457-465.
- [3] Yu Shui, Zhou Wanlei. An efficient reliable architecture for application layer anycast service[C]//6th International Conference on Algorithms and Architectures for Parallel Processing (ICA3PP). Melbourne, Australia, 2005, LNCS V371 9: 376-385.

(下转第 78 页)

参考文献:

- [1] Ando T, Fowler A B, Stern F. Electronic properties of two-dimensional systems [J]. Rev Mod Phys, 1982, 54(2):437-672.
- [2] Sune J, Olivo P, Ricco B. Quantum-mechanical modeling of accumulation layers in MOS structure [J]. IEEE Trans Electron, 1992, 39(7): 1732-1739.
- [3] Pregaldiny F, Lallement C, Van Langevelde R, et al. An advanced explicit surface potential model physically accounting for the quantization effects in deep-submicron MOSFETs[J]. Solid-State Electronics, 2004, 48(3): 427-435.
- [4] Van Langevelde R, Klaassen F M. An explicit surface-potential-based MOSFET model for circuit simulation [J]. Solid-State Electronics, 2000, 44(3):409-418.
- [5] 张衡, 张武. 二维 Poisson 方程边值问题的块三对角可扩展并行算法 [J]. 微电子学与计算机, 2008, 25(10): 117-120.
- [6] 张大伟, 章浩, 朱广平, 等. 亚 100nm 体硅 MOSFET 集约 I-V 模型 [J]. 半导体学报, 2005, 26(3): 554-561.
- [7] 高珊, 孟坚, 陈军宁, 等. OISD MOSFET 阈值电压建模与结构设计[J]. 微电子学与计算机, 2010, 27(7): 242-249.

作者简介:

李亦清 男, (1986-), 硕士研究生. 研究方向为新一代纳米半导体器件建模、非易失性存储器.

王子欧 男, (1968-), 博士, 副研究员. 研究方向为微电子工艺、电路设计以及表征分析等应用技术.

(上接第 74 页)

- [4] Wu Chijen, Hwang Renhung, Ho Janming. A scalable overlay framework for internet anycasting service[C]// The 22nd Annual ACM Symposium on Applied Computing (SAC'07). Korea; Seoul, 2007.
- [5] 裴胜玉, 周永权, 罗淇方. 求解约束优化问题的混合粒子群算法[J]. 微电子学与计算机, 2010, 27(4): 5-8.
- [6] Stevens T, De Turck F, Dhoedt B, et al. Achieving network efficient stateful anycast communications[C]// Proceedings of the 21st International Conference on Information Networking (ICOIN2007). Estoril, Portugal, 2007.
- [7] Ballani H, Francis P. Towards a global IP anycast service[J]. ACM SIGCOMM Computer Communication Review, 2005, 35(4): 301-312.
- [8] 熊琴, 李陶深. 一种基于遗传粒子群算法的选播路由算法[J]. 微电子学与计算机, 2009, 26(9): 75-77.

作者简介:

张园园 女, (1983-), 博士研究生. 研究方向为网络体系结构、路由协议、选播.

余镇危 男, (1942-), 博士生导师. 研究方向为网络体系结构、覆盖网、下一代互联网.