

Overlay 网络上的服务部署问题的研究 (一)¹

余镇危, 冯烟利, 潘耘
中国矿业大学(北京), (1000083)
zwyu@cumtb.edu.cn

摘要: 本文提出了 Overlay 网络上的服务部署问题, 给出形式化描述, 并给出了一对一单因子服务部署问题的数学模型和求解该问题的启发式算法, 通过数值模拟表明了算法的有效性。

关键字: Overlay 网络, 服务部署, 单因子服务

1. 引言

Overlay 网络是建立在一个或者多个已存在网络之上的网络, 通过增加额外的、间接的、虚拟的层, 来改善下层网络部分领域中的一些属性, 提高网络的性能。Overlay 网络有以下特性: 第一扩展性好, 一个 Overlay 网络不需要改变已有的网络结构, 不需整个网络范围内的网络组件的支持, 这一思想的实施将会加快新型的网络功能和服务的部署; 第二灵活性强, Overlay 网络所建立的是一个抽象的连接, 这些连接可根据应用程序的需要, 进行优化; 第三健壮性, 一个 Overlay 网络比下层的网络设施更加健壮。因此, Overlay 网络概念一经提出, 就引起了人们的重视。特别是近几年以来, 在现有的 Internet 传输网络之上构建的完全位于应用层的 Overlay 网络系统发展很快。例如提出了监测路径失效的弹性 Overlay 网络 (Resilient Overlay Networks, RON)^[1], 为现有的 Internet 提供端对端服务质量保证的服务 Overlay 网络 (Service Overlay Networks, SON)^[2]等。

Overlay 网络具有广阔的应用前景, 为许多在传统网络中难于解决的问题给出了一个可行的过渡解决方案。为了在 Overlay 网络上支持用户对多样服务的需求, 需要选取合理的节点部署所需的服务 (如组播、压缩、缓存、重新编码等, 用 proxy 表示), 这是在 Overlay 网络上进行路由的基础。

目前, 人们对 cache/mirror/CDN proxy/replica 部署问题进行了大量的研究^[4-6], 通常将它们抽象为两类著名的问题: 设备定位问题 (the facility location problem)^[7]和 k-median^[8]问题来研究。它们主要考虑服务节点与客户节点之间的通信代价, 而不去考虑客户节点之间在网络中收到某种服务支持的问题。

本文主要关心对端系统来说, 服务如何部署才能有利于端系统之间的通信。为了方便对服务部署问题的研究, 把服务进行了分类: 一对一的服务和一对多的服务。在 Overlay 网络上不只可放一类服务, 还可进行服务组合产生新的服务, 为此, 我们将部署服务的问题可以分成两类, 一类是部署单一的服务, 即端系统之间进行通信时, 需要网络提供单一的服务即可完成, 我们称为单因子服务; 另一类是必须几个服务的组合才能完成一次通信, 也就是说该服务需要多个因子的共同作用才能完成任务, 我们称为多因子服务。然后分别针对多中组合方式对 Overlay 网络上的服务部署进行了研究, 给出了相应的数学模型, 提出了求解这些问题的算法, 并进行了大量的数值模拟。

¹基金项目: 高等学校博士学科点专项科研基金资助课题 (20030290003)

2. 服务分类

Overlay 网络内部节点支持的服务，可分为两类：一类是从源端来的数据流接受该服务后流向唯一的节点，另一类是接受该服务后流向多个目的节点。

2.1 一对一服务

一对一服务（如图 1）是指数据流接受该项服务后，仍然流向确定的唯一目的节点，只是数据流的流量，或者负荷中的数据内容发生了改变。如压缩服务改变了数据流的流量（如图 2），加密解密和其他编码服务改变了负荷中的数据内容等，所以它们都可看作为一对一的服务。



图 1 一对一的服务

Fig. 2. one-to-one service

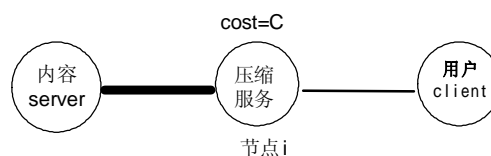


图 2 压缩服务

Fig. 2. compress service

2.2 一对多服务

一对多服务是指数据流接受该项服务后，复制数据流向多个目的节点(图 3)。组播通信服务就可归为一对多的服务(图 4)，组播节点对数据流进行复制，发向同组的多个目的节点。

我们可以把一对一的服务与一对多的服务进行组合，来完成更复杂的服务。例如加密服务与组播服务进行组合，这样每个组播组的目的节点都可获得加密服务。

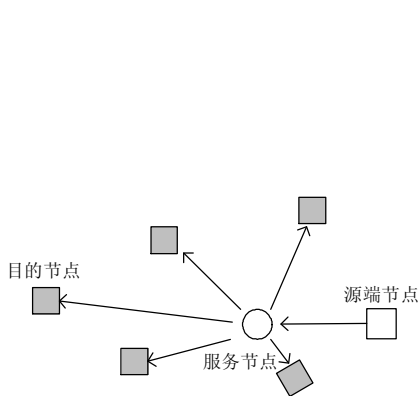


图 3 一对多服务

Fig. 3 one-to-many service

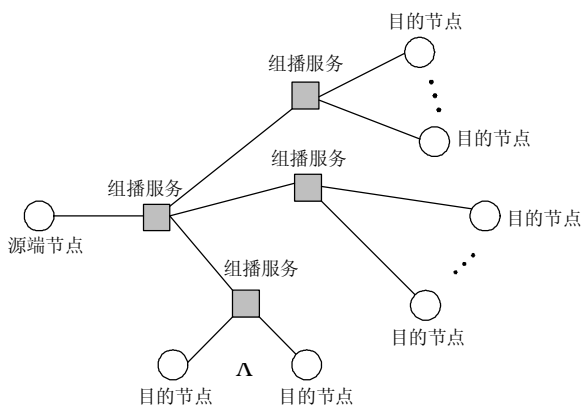


图 4 组播服务

Fig. 4 multicast service

3. 服务因子

在研究服务部署问题时，还需要考虑服务因子的影响。因为有些端系统之间的通信需要接受中间节点提供的一次服务，有些通信需要组合多个相关的服务才能完成。所以 Overlay

网络上的服务部署问题分为：单因子和多因子的服务部署问题。为了说明问题，我们借用了会话的概念，会话是指两个端系统通过 Overlay 网络进行通信的过程。

单因子的服务部署问题是指如何为多个需要同一服务的会话对选取合适的主动节点来部署适量数目的该类服务。多因子的服务部署问题是指如何部署多个相关联的服务因子才能满足多个会话组按不同次序来组合服务的需要。图 5 给出了两个服务因子的组合服务。

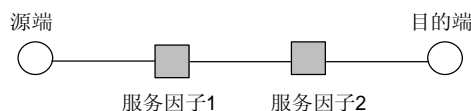


图 5 多因子服务

Fig. 5 multiple factors service

4. 单因子一对一服务部署问题

根据一对一服务的定义，在研究其服务部署问题时，我们只需要考虑源端到服务节点、服务节点到唯一目的节点这些路径上的问题。根据单因子服务的定义，数据流只获取一次服务，我们需要考虑一次部署服务的代价和该服务的成本代价。

4.1 模型的形式化描述

给定的网络拓扑结构 $G=(V,E,A)$ ， A 集合表示主动节点的集合。为了给出该模型的形式化描述，我们给出如下定义：

定义 1 主动服务标志函数 (active service tag)，用来表示该主动节点上是否部署服务。

$$x : A \rightarrow [0,1], x(i) = \begin{cases} 1, & \text{节点 } i \text{ 上放置服务} \\ 0, & \text{节点 } i \text{ 没有放置服务} \end{cases} \quad (1)$$

定义 2 主动服务代价函数 (the function of service cost)

$$f : A \rightarrow [0,\infty), f(i) = \text{节点 } i \text{ 建造服务的代价和提供的服务代价} \quad (2)$$

定义 3 上行服务节点

对于 s 到 d 的路径 $path(s,d)=s...j...d$ ，在 j 上部署服务为 s 到 d 的数据流服务，定义目的节点 d 的上行服务节点为路径 $path(s,d)$ 上距离 d 最近的且承担了服务的节点 j 。用函数 p 来表示 $p(d)=j$ ，这里 $x(j)=1$ 。显然，在只有一个会话的一对一单因子服务模型中， $path(s,d)$ 中唯一承担服务的节点就是 d 的上行服务节点。



图 6 上行服务节点

Fig. 6 The precedence service node

定义 4 一对一服务的 proxy 部署问题(UniService Proxy Placement Problem,简称 UPP)

给定 L 组源端节点和目的节点对 $\{< s_1, d_1 >, < s_2, d_2 >, \Lambda, < s_L, d_L >\}$ ，在集合 A 上部署服务，使得每一对会话都能获得单因子服务 C ，求使得代价最小的服务节点的部署策略。这

里还需要分两种情况来讨论：一是可以部署服务的节点数目无限制，只要是 A 中的节点，均可在其上部署服务 C ；二是由于资源等条件的限制，要求在 A 中不超过 $k(≤|A|)$ 个节点上部署服务 C 。这两种情形下的 UPP 问题均可形式化表示为如下优化模型（当 $k=|A|$ 时约束(c)可以忽略）：

$$C_U = \min_{x \in \{0,1\}^{|A|}} \sum_{i=1}^L c_G(s_i, d_i) + \sum_{i \in A} x(i) \cdot f(i) \quad (a)$$

$$s.t. \quad x(p(d_i)) = 1 \quad (i = 1, 2, \dots, L) \quad (b) \quad (3)$$

$$\sum_{i \in A} x(i) \leq k \quad (c)$$

其中：约束条件(b)表示每对话都获取了服务 C ，约束(c)表示部署服务 C 的主动节点个数不能超过 k ， $C_G(s, d)$ 表示图 G 中从 s 到 d 的最短路径长度(类似地，算法描述中的 $C_{G'}(s, d)$ 表示分层图 G' 中从 s 到 d 的最短路径长度，它已经包含了该路径上主动节点的建造代价和服务代价)。

4.2 求解 UPP 问题的启发式算法

UPP 问题是多于一个参数的最短路径问题，是 NP 完全问题。为了方便求解问题，把节点费用拆成链路费用，这样有助于采用动态规划^[9]的思想来解决问题。

首先，复制一份原始拓扑图，形成两层，如图 7 所示。对于在原图中每个主动节点 $r \in A$ ，其复制的图中对应的节点为 r' 。在两层中增加 r 指向 r' 的边 (r, r') ，其权值等于节点 r 的部署服务的建设代价和服务的成本代价的综合值 $f(r)$ 。这样构成的求解算法的新图 G' （原图称为第 0 层，复制的新层称为第 1 层）。

我们就可在图 G' 求解源点到目的点的获取服务 C 的最短路径。源点为第 0 层的 s_i 节点，目的节点为第 1 层的 d_i' 节点。这样可以使用一般的最短路径算法，计算好的这条路径将跨越两层，合并之后即是原图的最小代价的路径。

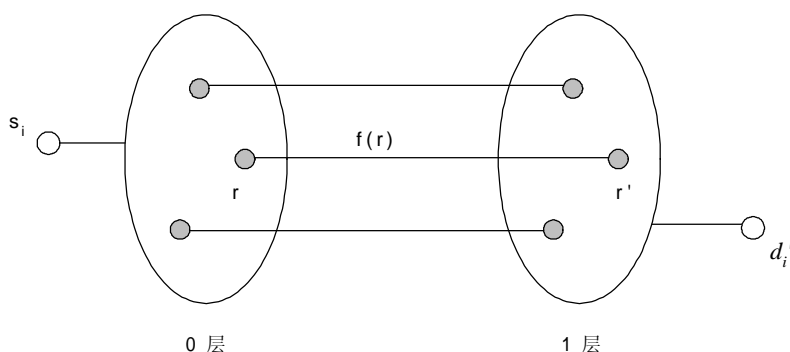


图 7 带约束的最短路径

Fig. 7 The constrained shortest path

在求得每对话的最短路径后，统计经过每个主动节点的会话数，若会话数为零，一定不会在这里部署服务节点，从 A 集合中去除，再统计不为零的主动节点个数，若小于规定的数目，说明这些节点上都可部署服务；若大于规定数目 k ，从 A 集合去除会话数最小的节点 i 。再次求经过节点 i 的会话的最短路径，继续统计重复以上工作，直到小于等于规定

的数目 k 。此时，仍然需要考虑是否可以进一步减少部署服务 C 的节点个数以改进优化目标函数值，即对 k 值的进一步寻优——这一步我们采用聚类分析中的层次法，即考虑合并 A 中距离最近的两个主动节点并重新估算目标函数值是否有“显著改进”，若有则继续寻找新的可以合并的节点，否则停止计算。此处“合并”的含义仅指承担服务功能上的合并，而非网络节点的真正合并，即不改变 Overlay 网络的拓扑结构。最后集合 A 中的节点就是需要部署服务的地点。

Algorithm1 UPP(G, A, k, SU)

Input: G is a network graph; A is a set of active service nodes; k denotes the total number; SU is a set of unicast session group $\{ \langle s_1, d_1 \rangle, \langle s_2, d_2 \rangle, \dots, \langle s_L, d_L \rangle \}$

Output: A' is a placement policy

Begin

Build layered model of graph G' according network G and active service nodes set A ;

Let $C_1 = \sum_{i=1}^L c_{G'}(s_i, d_i)$;

For $\forall j \in A$, define set $B(j) = \Phi$ and let $L' = L$;

L1: For $\forall i \in L'$, find the shortest path, π_i , from s_i to d_i using Dijkstra Shortest Path algorithm,

Meanwhile, for each node j in π_i , do

if j is also in A , then let $B(j) = B(j) \cup \{i\}$ and $p(d_i) = j$;

Let $A' = \{j \in A : |B(j)| > 0\}$;

If $|A'| \leq k$ then goto L3;

Find r , such that $|B(r)| = \min_{j \in A'} |B(j)|$;

L2: Delete edge $\langle r, r' \rangle$ from G' , and let $L' = B(r)$;

Let $B(r) = \Phi$ and goto L1;

L3: Let $C_0 = C_1$ and $A'' = A'$;

Let $C_1 = \sum_{i=1}^L c_{G'}(s_i, d_i)$;

If $C_0 - C_1 \leq M$ or $|A''| = 1$ then goto L4;

Find two node r_1 and r_2 in A'' , such that $dist(r_1, r_2) = \min_{\forall j_1, j_2 \in A''} dist(j_1, j_2)$;

If $\sum_{i \in B(r_1)} c_{G'}(s_i, d_i) < \sum_{i \in B(r_2)} c_{G'}(s_i, d_i)$, let $r = r_1$; otherwise let $r = r_2$;

Goto L2;

L4: A'=A";

Output A'

End

计算复杂性分析：求解每对会话的源端到目的端之间的最短路径采用 Dijkstra 算法(针对分层图 G)，其复杂度为 $O(|V| \log |V| + |E| + |A|)$ ，故 L 个会话的总的计算复杂度为 $O(L \cdot (|V| \log |V| + |E| + |A|))$ ；在对 k 寻优阶段，“合并”及删除某些主动边重复计算的次数不超过第一次计算得到的 $|A| - 1$ ，而每次重复时平均对 $L/|A|$ 个会话重新计算最短路径，考虑到 $\sum_{i=2}^{|A|-1} \frac{L}{i} \leq \sum_{i=2}^{|A|-1} \frac{L}{i} = O(L \cdot \log |A|)$ ，则对 k 寻优阶段的计算代价可表示为 $O(L \cdot \log |A| \cdot (|V| \log |V| + |E| + |A|))$ 。

4.3 仿真

为了测试算法 UPP 的性能，我们仍采用 BRITE 工具^[10]来生成网络的拓扑图，拓扑生成基于 Waxman 模型^[11]的拓扑生成算法。在我们的实验中，网络有如下属性：a. 边的费用在 [10,200] 上均匀分布。b. 主动节点上部署服务节点的费用在 [1,10] 上均匀分布。c. 平均每个节点连接边数 $m=5$ ， $\alpha=0.15$ ， $\beta=0.2$ 。以随机部署算法为对比对象，会话对随机生成。图 8 给出了网络节点为 40 时，随机选取 3 对会话，随机算法与 UPP 算法计算的代价与部署服务的个数之间的变化关系。可以看出 UPP 算法优于随机算法，随着部署服务节点的上界 K 增大，代价在减小，到一定程度代价将趋于平衡。这说明部署服务为 4-5 为最佳。

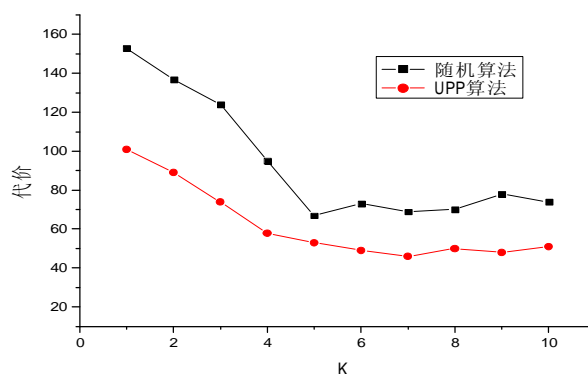


图 8 目标函数随服务节点个数的变化

Fig.8 The variation of objective function value with the number of service nodes

5. 结论

本文首先讨论了Overlay网络上的服务部署问题，进而把服务分成两类，还引入了服务因子的概念。对一对一的单因子服务问题展开了详细的讨论，在适当抽象的基础上，建立了

数学优化模型，讨论并实现了解决模型的算法，最后分析了这些算法的复杂度并进行了数值仿真实验。

参考文献

- [1] David G. Andersen, Hari Balakrishnan, M. Frans Kaashoek, and Robert Morris. Resilient Overlay Networks. In Proceedings of SOSP 2001, October 2001.
- [2] Yang hua Chu, Sanjay Rao, and Hui Zhang. A Case for End System Multicast. In Proceedings of ACM Sigmetrics, June 2000.
- [3] John Jannotti, David K. Gifford, Kirk L. Johnson, M. Frans Kaashoek, and Jr. James W. O'Toole. Overcast: Reliable Multicasting with an Overlay Network. In Proceedings of Operating Systems Design and Implementation (OSDI), October 2000
- [4] Xiaohua Jia, Deying Li, Xiaodong Hu, and Ding-Zhu Du: Optimal Placement of Web Proxies for Replicated Web Servers in the Internet. The Computer Journal 44(5): 329-339 (2001)
- [5] B. Li, M. J. Golin, G. F. Italiano, X. Deng, and K. Sohrawy, On the optimal placement of Web proxies in the Internet, in Proceedings of IEEE INFOCOM'99, Mar. 1999.
- [6] S. Jamin, C. Jin, A. Kurc, D. Raz, and Y. Shavitt, Constrained mirror placement on the Internet, in Proceedings of IEEE INFOCOM'2001
- [7] D.B. Shmoys, ' E. Tardos, and K. Aardal. Approximation algorithms for facility location problems. In 29th ACM Symposium on Theory of Computing, pages 265--274, 1997
- [8] Sven de Vries, Marc E. Posner, Rakesh V. Vohra, The K-median Problem on a Tree.
www.kellogg.nwu.edu/faculty/vohra/ftp/Kmpap.pdf
- [9] 卢开澄编著，单目标、多目标与整数规划，北京：清华大学出版社，1999.
- [10] Alberto Medina, Anukool Lakhina, Ibrahim Matta, John Byers. BRITE: An Approach to Universal Topology Generation. In Proceedings of MASCOTS '01, August 2001.
- [11] B. Waxman. Routing of Multipoint Connections. IEEE J. on Selected Areas in Comm., 6:1617-1622, December 1988.

Service Placement Problem in Overlay Networks (1)

Yu Zhenwei, Feng Yanli , Pan Yun

China University of Mining and Technology-Beijing, Beijing, 100083

zwyu@cumtb.edu.cn

Abstract

The service placement problem in Overlay networks and its formal description are proposed in this paper. The model and the algorithms for *one-to-one* single factor service placement problem have been given. Numerical experimental simulations have supported the feasibility of the model and the algorithms.

Key words: *Overlay networks, service placement problem, single factor service*