

内容网络服务节点部署理论综述

尹 浩¹⁾ 袁小群²⁾ 林 闯¹⁾ 张 法³⁾ 庞善臣¹⁾ 刘志勇³⁾

¹⁾(清华大学计算机科学与技术系 北京 100084)

²⁾(华中科技大学电子与信息工程系 武汉 430074)

³⁾(中国科学院计算技术研究所 北京 100190)

摘 要 内容网络通过在 Internet 上部署由多个服务器组成的服务节点来构成覆盖网,有效解决了新应用对网络的多样性需求与 Internet 尽力而为服务之间的矛盾,一直是学术界和工业界研究的热点. 服务节点部署是内容网络研究的一个核心问题,部署好坏将直接影响到覆盖网络的服务质量和运行效率. 随着 Internet 和网络应用的发展,内容网络所处的网络环境、服务质量的要求和与网络运营商之间的关系都出现了新的变化,传统服务节点部署理论面临着众多的新挑战. 近年来,如何改善现有的服务节点部署理论已成为一个热点问题. 文中系统总结了现有内容网络服务节点部署理论与应用,及其面临的挑战,针对挑战提出了新的研究方向与解决方案.

关键词 内容网络;节点部署;基于确定信息的选址模型;基于概率的选址模型;基于博弈论的选址模型

中图法分类号 TP393 **DOI 号:** 10.3724/SP.J.1016.2010.01611

The Survey of Service Nodes Placement Theories for Content Networks

YIN Hao¹⁾ YUAN Xiao-Qun²⁾ LIN Chuang¹⁾ ZHANG Fa³⁾ PANG Shan-Chen¹⁾ LIU Zhi-Yong³⁾

¹⁾(Department of Computer Science and Technology, Tsinghua University, Beijing 100084)

²⁾(Department of Electronics and Information Engineering, Huazhong University of Science and Technology, Wuhan 430074)

³⁾(Institute of Computing Technology, Chinese Academy of Sciences, Beijing 100190)

Abstract By deploying service nodes, which composed by multiple servers in diverse geographical locations, to form the overlay network, Content Networks can effectively reduce the gap between the demand of ever-emerging application for network diversity and the property of the best effort for Internet. It has been the research focus in academy and industry. The deployment of service nodes, one of key problems in content network, affects directly the quality of service (QoS) and operational efficiency of the overlay network. However, with the development of Internet and network application, the environment of content networks, QoS requirements and the relationship with network operators have changed, the traditional deployment theories face many new challenges. Recently, how to improve the existing service nodes deployment theory has become a hot issue. In this paper, we systematically survey the existing service node deployment theories and their applications, address the challenges faced. For the challenges, we also present the research tendencies and the potential solutions in final.

Keywords content networks; service-nodes deployment; graph-based location model; stochastic location model; game theory location model

收稿日期:2010-04-20;最终修改稿收到日期:2010-08-05. 本课题得到国家自然科学基金项目“新一代内容分发网络基础理论研究”(60873254)、“Petri 网可重写理论及在服务组合中的应用”(60970001)和国家“九七三”重点基础研究发展规划项目基金(2011CD302600)资助. 尹 浩,男,1974 年生,博士,副教授,主要研究方向为内容分发网络、分布式系统中资源管理与安全控制以及性能评价. E-mail: h-yin@mail. tsinghua. edu. cn. 袁小群,男,1976 年生,博士研究生,主要研究方向为内容分发网络、P2P、适时流媒体分发和资源优化等. 林 闯,男,1948 年生,博士,教授,博士生导师,主要研究领域为计算机网络和系统性能评价. 张 法,男,1974 年生,博士,副研究员,研究方向为计算机算法、高性能计算、网络降耗算法. 庞善臣,男,博士,研究方向为 Petri 网理论与应用、分布式并行计算分析和认证. 刘志勇,男,1946 年生,博士,研究员,主要研究领域包括计算机算法、高性能计算、网络降耗算法.

1 引言

近年来,随着网络和通信技术的发展,Internet 已成为人们日常生活不可缺少的信息交流平台. Internet 服务的内容也从当初以静态网页、小图片为主转变到现在以大文件、视频为主,并进一步朝交互式信息共享、协同工作等方向发展,如视频共享和社交网络服务(Social Networking Sites, SNS)等. 这些新的互联网应用需要 Internet 能够提供高扩展性、大容量、高服务质量保证、强交互性和安全性等方面的保障. 但现有 Internet 的基础架构采用端到端的设计原则,单纯实现数据包的存储转发^[1]. 虽然这种基础架构扩展性强且实现简单,但难以满足新应用的需求. 为此,学者们分别从改变现有网络体系和改进现有网络体系两种思路来解决上述新应用需求与尽力而为网络基础架构之间的差距问题. 前者主要通过设计新的基础网络架构,使之能够满足当前和今后一段时间的网络发展需求,如 Clean-slate design^[2],该方案由于对现有系统的改动过大,一直未得到大规模应用. 而后者则以现有基础设施为基础,通过引入新的技术与服务层来对现有网络性能加以改进,以达到满足应用需要的目的,如内容网络. 根据 RFC 定义,内容网络通过在 Internet 部署服务节点,并通过应用层协议将这些服务节点组织形成一个构建在 IP 网络之上的覆盖层,为网络应用提供灵活高效的服务^①. 其典型应用有内容分发网络(Content Delivery Network, CDN)、对等网络(Peer to Peer, P2P)、Cloud Computing^②等.

使用内容网络来为用户提供服务,能够增加网络灵活性、提高用户体验并降低骨干网络流量压力. 但这些优势的取得是通过在现有网络上部署服务节点来实现的. 因此,服务节点部署的好坏,直接影响到内容网络的服务质量和使用效率,对内容网络服务性能产生重大影响.

服务节点部署一直以来都是学术界和工业界关注的热点和难点,产生了众多有意义的成果. 近年来,计算机技术、网络技术和网络应用的飞速发展,使得内容网络的发展规模、所处的 Internet 环境、需要服务的应用类型以及与其它系统之间的关系都发生了较大改变. 传统服务节点部署理论在应对新的网络环境和新的应用需求时,需要解决如何应对复杂网络环境、如何克服传统服务节点部署理

论的缺陷以及如何对构建的部署模型加以实证分析等问题.

本文首先总结了现有服务节点部署理论与这些理论在内容网络中的应用. 在此基础上,文章结合当前内容网络面临的复杂环境以及内容网络需要服务的各种应用,找出当前内容网络服务节点部署面临的挑战及存在的研究方向和可能的解决方案.

2 节点部署理论研究

服务节点部署在理论上属于选址问题^[3],一直都是学术界研究的热点,包括部署模型的构建和搜索算法设计两部分. 其中,模型构建决定最终系统部署性能和搜索算法的总体性能,是服务节点部署问题的核心,也是本文研究的重点. 从数学建模来看,根据所需获得信息的不同,服务节点选址模型可以分为基于确定信息的选址模型、基于概率模型的选址模型和基于博弈论的选址模型.

2.1 基于确定信息的选址模型

基于确定信息的选址模型是最常用的一类选址模型. 这主要是由于在现实中,设计者通常会事先知道网络拓扑信息或者假设已知网络拓扑信息等,以方便模型简化工作. 根据部署模型的实现过程,选址模型可以分为单步选址和多步选址两类. 其中,单步选址技术的核心思想是通过一次输入所需信息,然后一次求解所需的部署策略;而多步策略则指部署模型的求解需要根据输入信息得到一些中间信息,再通过中间信息获得最后求解策略. 无论是单步选址技术,还是多步选址技术,都包含 Facility Location 和 K-Median 两种模式^[4].

2.1.1 单步选址模型

单步无容量约束 Facility Location 可以被描述成最小化部署成本与服务性能折中后的混合标准:

$$\min \sum_{k \in K} \sum_{j \in J} c_{kj} z_{kj} + \sum_{j \in J} f_j s_j \quad (1)$$

其中, c_{kj} 表示备选用户 k 到备选节点 j 间的距离权值, $z_{kj} = 1$ 表示用户 k 由节点 j 服务, f_j 表示在备选节点 j 部署单位服务器的成本, s_j 表示分配给备选节点 j 的容量.

① Day M, Cain B et al. A Model for Content Internetworking (CDD). <http://www.ietf.org/rfc/rfc3466.txt>

② Armbrust M, Fox A et al. Above the Clouds: A Berkeley View of Cloud Computing, Technical Report, 2009. <http://www.eecs.berkeley.edu/Pubs/TechRpts/2009/EECS-2009-28.pdf>

上述模型可以用 LP-relaxation 来分析,但该解法只能获取弱下界。Korkel 等人在对偶自同构的基础上设计出 Branch-and-bound 算法来获取更强的解法^[5], Goldengorin 进一步对 Branch-and-bound 算法进行了改进^[3]。

通过对每个备选节点的容量增加约束 $p_L \leq \sum_{j \in J} s_j \leq p_U$, 可以获得固定 p 值的单步 K -Median 模型。其中 p_L 表示容量下界, p_U 表示容量上界。

单步选址问题包括聚合容量选址^[6]和节点容量选址问题^[7]。前者对单个备选节点容量无限制,但对总需求容量进行限制,后者则对潜在的备选节点有一定的容量限制。显然,对于单步聚合容量选址问题,只需在上述无容量限制模型上增加聚合容量限制条件 $\sum_{j \in J} S_j y_j \geq d(K)$ 即可。其中, $S_j > 0$ 表示备选节点 j 的最大容量, $y_j \in \{0, 1\}$ 表示备选节点 j 是否被选择, $d(K) = \sum_{k \in K} d_k$ 表示总的容量需求, d_k 表示用户 k 的服务需求量。

而单步容量限制选址模型则需要在单步无容量限制模型基础上对每个潜在的备选节点增加容量约束条件 $\sum_{k \in K} d_k z_{kj} \leq s_j y_j, j \in J$ 。其中, $z_{kj} \in \{0, 1\}$ 表示备选节点 j 是否服务用户 k , $s_j y_j$ 为节点服务用户 j 的容量限。聚合容量模型通常可以通过抽取算法来获得部署方案, Ryu^[6] 等人在这方面做了相应的工作。而容量限制模型则通常采用 Lagrangean 方法加上一些其它的不等式约束来获取模型的下界,但很难使用精确抽取法来获取解^[8]。此外,一些基于 Primal-and-primal 对偶分解算法也被提出来求解容量限制模型^[9]。

2.1.2 多步选址模型

多步选址模型主要应用于分层结构的内容网络中。在这类内容网络中,位于高层的节点在条件满足的情况下(如高层节点被分配的容量应大于低层节点被分配的容量)可以独立选择其低层节点。以两层为例,在上述条件下,假定 x_{ij} 表示下层节点对上层节点 i 的需求量, t_{ij} 表示上层节点 i 服务来自下层节点 j 的单位请求的代价。此时,两层带容量约束的 Facility Location 模型可以通过同在模型(1)的基础上增加上层节点到下层节点服务代价函数 $\sum_{i \in I} \sum_{j \in J} t_{ij} x_{ij}$ 得到。

对于两层选址技术, Tcha^[10] 等人讨论了网络拓扑变化的两层 Facility Location 问题的建模,

Geoffrion^[11] 等人讨论了两层带容量约束的 Facility Location 问题,而 Tcha 等人则讨论了两层无容量 Facility Location 问题。

同理,多步 K -Median 模型可以通过在模型(1)基础上增加多层服务过程中的总代价函数 $\sum_{i \in I} \sum_{j \in J} \sum_{k \in K} q_{ijk} \omega_{ijk}$ 来得到。其中, ω_{ijk} 表示节点 k 通过 $i \rightarrow j \rightarrow k$ 服务过程的服务请求占总服务请求的比值, q_{ijk} 指通过 $i \rightarrow j \rightarrow k$ 服务过程的服务成本。上述模型需要对每一个客户的请求量加以约束,同时对服务种类加以约束。多步 K -Median 问题是一个研究热点, Geoffrion^[11] 等人都在这一方面做了相关研究。

2.2 基于概率模型的选址模型

基于概率模型的选址模型是主要针对系统设计者在未知网络状况或服务用户状况情况下来设计内容网络时而构建的一种选址模型。这类模型通常具有很好的适应性,能够在恶劣的环境下有较好的系统性能表现,其核心思想是假设服务节点或用户节点依某种概率分布,然后采用优化理论进行建模。典型建模方法是将网络节点假设为均匀分布,然后以服务距离和最小为优化目标来构建模型:

$$\min_{(p)} \sum_{k \in K} \omega_k d_k(p), \quad d_k(p) = |p - a_k|_l,$$

其中, ω_k 表示备选节点到用户间距离的权值, p 为备选节点, $\| \cdot \|_l$ 表示备选节点 p 与用户 a_k 之间的 l 测度,即 l 测度下的距离。

Weiszfeld 等人针对这种模型提出一种类似梯度的搜索算法^[12], Miehle^[13] 等人对这一方法进行了改进。对上述模型进行改进,可以用于 K -Median 问题。它是一个 NP-hard 问题,可以采用非线性综合法来进行建模:

$$\min \sum_{k \in K} \sum_{j=1}^p (\omega_k d_k(p)) z_{kj} \quad (2)$$

其中, $z_{kj} = 1$ 表示用户 j 由节点 k 服务, $d_k(p)$ 表示用户到备选节点的距离, ω_k 表示备选节点到用户间距离的权值。

模型(2)很难获得其精确解。因此, Rosing 采用线性规划松弛法来对上述模型重新建模,然后得到准确解^[14]。但文献[15]设计的搜索算法复杂度很高,为此, Brimberg 等人提出一些快速的启发式算法^[16]。

除了上述基于概率模型的选址模型外,还有一些其它情况也需要以概率方式来建模,如当需求和边的权值为随机变量时, Mirchandani 考察了这一情况下的 K -Median 问题^[15]。他假设图具有有限个状

态,且每一状态可能存在的概率已知.在这种情况下,Mirchandani 将选址问题描述成总的边权重和最小:

$$\min \sum_{i \in I} \sum_{k \in K} \sum_{j \in J} \pi_i c_{ikj} z_{ikj} \tag{3}$$

其中, π_i 表示节点*i*的需求量, c_{ikj} 表示在状态*i*下节点*k*到节点*j*间的权值距离, z_{ikj} 表示在*i*状态下被选节点的选择,它考虑了在状态*i*下对被选节点的容量分配.这类随机选址问题通常采用 Branch-and-cut 算法来获取选址方案,如 Laporte 采用 Branch-and-cut 算法来获取随机需求的选址方案^[17].

2.3 基于博弈论的选址模型

基于博弈论的选址模型从描述内容网络与其它系统之间存在的关系出发,采用博弈论的相关理论来构建相关模型,以求得全局最优部署方案.

Blanchini 讨论了在已知网络拓扑的网络流和网络设计过程中,当需求满足均匀分布时的服务节点选择问题.他将这个问题描述成一个两人合作博弈的过程.网络管理者为第 1 个博弈者,他的目标是最大化网络流量.第 2 个博弈者为用户请求,它的目标是尽量获得最大的服务.模型最终目标是获得一

个双赢的服务节点部署策略^[18].

Chun 等人讨论了 P2P 系统中 cache 节点的部署问题.他们将 P2P 系统中每一个 cache 节点描述成单纯追求本节点服务性能最优化,不同节点间缺乏相互协作.在此基础上,Chun 设计出一种非协作式博弈模型.实验和分析结果表明,该模型能够在一定程度上优化 P2P 系统整体服务性能^[19].Chekuri^[20]等人也用博弈论讨论内容网络的服务节点部署问题,但他们关注于非协作式组播游戏系统中的服务节点部署问题,并设计出两种博弈节点部署模型:单连接路径模型和多连接路径模型.此外,Cardinal 等人考察了非协作式的服务节点部署问题,并使用非协作式博弈模型将其抽象化为一个覆盖优化问题.在此基础上,他们提出一种近似稳定的次优算法来获取纳什均衡点^[21].

2.4 小 结

服务节点部署问题属于优化问题,其设计过程是一个选择优化目标,然后根据所得到的相关信息,构建部署模型,最后求解的过程如图 1 所示.

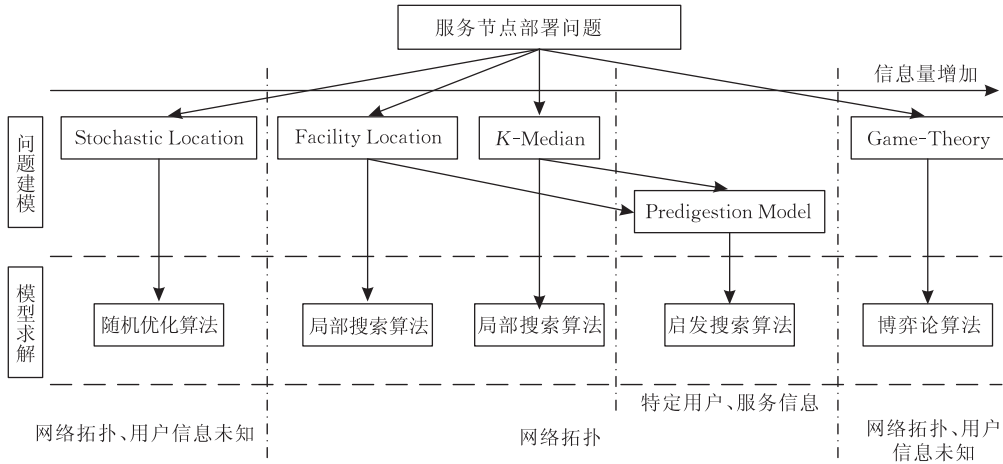


图 1 节点选址问题解决方案

(1) 模型构建. 根据获取信息的不同,节点部署模型可分为基于确定信息的选址模型、基于概率模型的选址模型和基于博弈论的选址模型等三类.基于概率模型的选址模型主要是指对先验知识缺乏了解的情况下,假设网络拓扑或用户相关信息(如用户分布)依某种概率分布,然后根据上述假设加以模型化.基于确定信息的选址模型是建立在已知网络拓扑的情况下,通过图论的相关知识来构建节点部署模型,分为 Facility Location 和 K-Median 两种.在实际问题建模的过程中,根据获取的实际信息,如特定拓扑信息、用户流量信息等,可以对这两种模型进

行简化,构建带约束的简化模型,便于近似求解,如贪婪节点部署模型^[22]、基于拓扑信息的节点部署模型^[23]等等.上述两种解决方案主要以内容网络本身为分析对象,寻找其优化目标,定义代价函数,构建部署模型.而基于博弈论的选址模型则从描述内容网络与其它系统之间关系入手,构建一个全局优化的节点选择模型.

(2) 优化求解. 对于基于概率模型的选址模型,根据对网络拓扑或需求假设的概率不同,有类梯度解法和最大似然解法等两种搜索算法.对于 K-Median 或 Facility Location 模型,可由局部搜索得到优化解.这种解法的优点是解的精确度高,但算法复杂度

高,当节点过多时,存在搜索效率低的问题.对于带约束的简化模型,多采用启发式算法,其优点是算法简单,效率较高,缺点是解法针对性强,普适性差.对于基于博弈论的选址模型,多数采用相应的求解方法.

3 服务节点部署应用研究

服务节点部署问题是内容网络的基础问题,其部署的好坏对内容网络应用性能有极大影响,是内容网络的研究热点与难点,产生了众多成果.根据组成内容网络服务节点的不同,可将这些成果归纳为基于 Cache 服务的节点部署技术、基于代理(proxy)服务的节点部署技术和基于种子服务的节点部署技术.

3.1 基于 Cache 服务的节点部署

Cache 是指一种程序信息存储设备或子系统,它能够在本地存储、恢复和删除信息.Cache 通过预先存储内容来减小终端用户响应时间和带宽消耗,提高系统服务性能^①.文中的 Cache 服务节点由一组能够在本地存储内容并服务终端用户的服务器组成.这类节点一般服务静态、小容量、频繁请求的内容,如静态 Web 页面、文件下载和静态流媒体等.

Li 等人较早关注 Cache 节点的部署问题.他们从用户体验角度讨论这一问题,提出一种基于结构树的节点部署模型,并设计了一种复杂度为 $O(N^2M)$ 的搜索算法.其中, M 为被选节点, N 为备选节点.该模型具有较好的适应性,但优化目标只考虑了节点与用户之间的流量和延迟,较为简单^[24];在 Li 等人的研究基础上, Qiu 等人进一步从服务节点部署对内容网络服务性能影响角度来考察服务节点部署问题.他们从 AS(自治域)网络拓扑的层面,对几种不同启发式部署算法的服务性能进行比较.研究结果显示:(1) 服务节点部署的好坏对内容网络服务性能好坏具有重要影响;(2) 服务节点部署性能受服务节点负载、用户分布和网络拓扑等信息的变化影响较小^[25]. Jamin 等人也从 AS 拓扑的层面研究服务节点部署问题.但他们以减小请求响应时间和服务器负载为目标,将节点部署问题描述成 K -median 问题,并提出了一种启发性部署算法.仿真结果表明:对于内容网络,通过适当方法,只需少量增加镜像节点数,便可明显降低系统服务响应时间和服务节点负载^[26].由于 AS 拓扑只能粗略反映 Internet 拓扑,为了获取更精确的网络拓扑信息, Radpslavov 从路由拓扑层面研究服务节点部署问

题.他们以用户体验和网络容量为优化目标并提出一种优化搜索算法.实验结果显示该算法不依赖于用户分布,且能达到贪婪算法的 0.8~1 倍的性能效果^[27].

上述节点部署问题研究的性能分析结果通常是建立在对简单拓扑和简单准则的仿真基础上的,与之不同, Oppenheimer 在真实系统中对服务节点部署算法进行考察,其核心目标是找出影响服务节点部署性能的主要因素.实测数据分析显示:在特定应用中,(1) CPU 和带宽是影响节点部署性能的重要参量;(2) CPU 和带宽起伏较为频繁.因此,有必要对服务节点进行动态部署^[28].

一些网络服务提供商(Internet Service Provider, ISP)也通过在合适的位置部署 Cache 节点来优化相应 Internet 流量. Cha 等人从 ISP 角度,以最小路由跳数为优化目标来考虑如何在 ISP(运营商)内部通过部署合适节点,以增加 ISP 内连通线路的灵活性^[29]. Cohen 等人也研究 ISP 内部节点部署问题,但他们以网络流量最大化为优化目标,从流量工程角度来考察网关设备的部署问题^[30].

此外, Guha 等人从分层网络设计出发,以部署成本和层间路由成本最小化为优化目标,提出一种分层 Cache 部署模型.在模型中,每层的 Cache 服务对应比例的请求^[31]. Krishnan 等人研究单服务器的选址问题,提出一种广义 Cache 选址模型,其核心目标是 minimized 网络流量和平均响应时间.在此基础上,他们提出一种基于线性和环形网络的最佳搜索算法^[32]. Buchholz 等人从内容网络自适应性角度考察了服务节点的部署问题^[33]. Laoutaris 等人研究了大规模 CDN 网络的服务节点部署问题,提出了分布式的无容量约束 K -Median 和无容量约束节点选址等两种模型.这两种分布式部署模型只需候选节点周围网络拓扑环境信息便可搜索出较优解^[34]. Cardinal 用非合作博弈的方式,将节点部署问题模型化为一个优化覆盖问题,并加以研究^[23].

3.2 基于代理服务的节点部署

内容网络除了提供静态网页服务外,还需要为大量不能被本地节点存储的内容(如动态网页、适时流媒体等)提供应用服务.这些应用通常采用部署代理服务节点来提高其服务性能. RFC2616 将代理定义为既能充当服务器,又能充当客户的中间媒介.因

① Fielding R, Getty J et al. Hypertext Transfer Protocol — HTTP/1.1. <http://www.w3.org/Protocols/HTTP/1.1/rfc2616.pdf>

此,当终端用户请求到来后,服务节点能够迅速响应服务或将请求转发到其它服务节点去.

Zhou 从流媒体分发网络中的可扩展性角度考察服务节点部署问题,并将这一问题描述成优化目标为平均响应跳数的优化问题,提出相应的流媒体复制算法.实验结果显示:与传统算法相比,算法能够将平均响应跳数减小 50%,具有较好的环境适应性^[35]. Almeida 等人将流媒体分布系统中的内容复制、请求路由和组播路由等问题转换为服务器部署问题,并构建了一个优化模型.与传统单播相比,该模型能将网络效率提高 50%~150%.此外,文章还提出 6 种启发式算法,并传统抽取算法进行性能比较,结果显示它们能在提高搜索效率的前提下获得近似最优解^[36]. Cahill 则试图解决电视流媒体服务的服务器部署问题,并将该问题模型化为一个优化目标为用户到服务器间的连接成本和服务器的存储成本之间融合的优化问题.在此基础上,文章提出一种优化搜索算法.实验结果表明:与“最近”部署方法相比,该算法能明显提高用户响应时间^[37].

此外,Ahuja 则从描述流媒体服务角度来研究服务器部署问题.他们以保证系统服务品质为优化目标构建服务节点部署模型,并提出一种有效搜索算法^[38]. Chekuri 将组播游戏模型化为非合作博弈问题来加以考察^[20].

3.3 基于种子服务的节点部署

与上面两类服务节点直接服务每一个终端用户请求不同,基于种子类服务系统的服务节点通过记录终端用户拥有的资源,然后通过分配终端用户间的资源相互传输来实现系统服务,不需要直接服务每一个终端用户请求,典型的系统有 P2P 网络.

Shi 等人从提高 P2P 网络可扩展性角度入手,将服务节点部署问题转化为覆盖问题,并分别采用线性规划和启发式贪婪算法来实现.该模型以网络距离为优化准则,其目标是寻找最少的服务节点数量^[39]. Roy 等人则从提高 P2P 网络端到端延迟和系统可靠性角度入手来研究 P2P 网络的服务节点部署问题.他们将服务节点部署问题转换成一个优化目标为可靠性和 TCP 吞吐量的图论问题.文章通过模拟实验,在比较了 4 种搜索算法的性能优劣基础上,指出随机搜索和贪婪搜索相结合的混合搜索算法能获得最佳搜索性能^[40].

除此外,Allen 考察了有线 P2P 流媒体服务器部署问题,并在中国电信网络中部署服务节点,实测数据分析显示通过在有线 P2P 网络中部署服务器,

可明显地减小 VOD 中心服务器的服务负载.而且,系统服务性能随 Peer 邻居规模和订阅流程序度的增大而提高.因此,通过部署 Cache 服务器来提高 P2P 系统服务效率在实际 P2P 系统中是切实可行的^[41]. Chun 等人将 P2P 服务节点部署问题模型化为一个非合作博弈问题^[19].

4 面临的挑战和趋势

4.1 面临的挑战和问题

传统的服务节点部署理论研究通常采用图论作为主要的数学工具,构建基于确定信息的部署模型,然后设计有效的搜索算法,最后求解部署策略.这种部署理论在一定条件下,能取得好的效果,但随着网络规模扩大,各种新型网络应用如移动视频、网络社区和网络游戏等应运而生并得到快速发展.上述新应用对服务需求各异,使得内容网络结构更加复杂.如移动视频应用需要内容网络提供高质量跨异构网络的透明服务,网络社区需要内容网络提供强交互性服务,而网络游戏则对内容网络服务的延时提出高的要求.在网络规模扩大,新应用快速发展和内容网络结构复杂化的情况下,传统服务节点部署理论面临着新的挑战,有不少难题亟需解决,总结起来,主要有如下几类:

(1) 如何应对复杂的网络环境. 各种新型网络终端和接入方法出现,计算机技术的发展和 Internet 的普及,使得内容网络面临的网络环境更复杂.同时,各类新应用的出现及快速发展,推动着内容网络规模朝超大规模发展,内容网络的体系结构和实现技术随之变得更复杂.复杂的外部环境和内部结构以及系统规模的扩大给内容网络节点部署模型的构建带来新的挑战,主要表现在:(a) 除了需要考虑传统服务节点部署中涉及的服务性能与部署成本等因素外,还需要考虑系统的运行和管理成本、网络异构性等新的影响因素;(b) 在复杂网络环境下,由于测量技术的限制,设计者设计节点部署方案时会面临不同程度的信息缺乏,甚至有时会面对无法获取建模需要的先验知识的情况.因此,如何从众多影响因素中选择出合适的影响因子,以此获取构建部署模型时需要的优化指标以及如何解决先验知识不足的问题,是当前服务节点部署问题面临的挑战.

(2) 传统理论的局限性. 由于 Internet 规模和内容网络规模以及新应用的飞速发展,更多新的影响因素出现,如系统的运行和管理成本、网络异构

性、网络服务提供商(ISP)的流量调度策略以及内容提供商在网络中的位置等等。这些新的影响因素在很大程度上影响到了内容网络的服务性能。同时,这些影响因素之间的关系变得错综复杂。以内容网络与 ISP 为例,随着内容网络规模扩大(如 Akamai),内容网络与 ISP 之间不再仅仅是传统的客服关系,而存在对等协作的关系。而传统服务节点部署理论的研究目标集中于内容网络本身,通过对内容网络某些性能进行优化来构建节点部署模型,不适合描述这些新影响因素和复杂利益关系。因此,如何很好描述各种影响因素及其相互间的复杂关系,是当前节点部署理论亟需解决的问题。另一方面,传统的服务节点部署策略存在搜索算法的效率随样本空间增加快速下降的缺点,而现有的内容网络规模都在快速增大,并正朝超大规模方向发展,这也给当前服务节点部署理论带来新的挑战。

(3)实证分析。传统服务器部署理论研究多采用模拟或仿真的方式来验证其模型构建或部署策略的有效性。这些验证方式限制了网络环境的变化,不能真实地反映现实网络环境,而实证分析是评估模型和算法有效性,提高并完善模型和算法的有效途径。因此,如何解决传统服务节点部署问题研究的局限性,构建灵活且与实际系统接近的实验平台,如何在构建的实验平台上再现和模拟真实网络环境,从而更好地评价服务节点的理论模型与相关部署策略,是新环境下服务节点部署问题研究面临的挑战。

4.2 研究方向及解决方案

通常,在设计服务节点部署方案时,首先需要确定部署优化目标,然后根据设计目标,构建节点部署模型,设计部署搜索算法,最后根据搜索结果来确定部署方案并验证该方案的有效性。

4.2.1 研究方向

为了确定优化目标,设计者需要确定能够影响部署性能的重要因素,然后根据这些因素来确定合适的优化目标。但在新环境下的内容网络节点部署设计中,有些因素在传统节点部署模型中缺乏研究但却对节点部署性能产生重大影响,如:(1)服务节点部署规模。服务节点部署规模是指服务节点部署规模的大小和数量的多少,即每个节点拥有的资源和设备数量的规模以及节点数量。当前,服务节点部署规模有类似于 Google 的大节点但少数量的部署方式来构建服务节点和类似于 Akamai 的小节点但多数量的部署方式等两种。虽然这两种部署方式各有利弊,但都对内容网络的服务性能有影响,还对

构建系统时的设备投入以及后续的系统管理成本有影响,已成为近年来内容网络的一个研究热点。Triukose 等人的实验结果表明:对于文件下载,过多数据中心并不能明显改善服务性能,但过少数据中心可能导致服务性能明显下降^[42]。Leighton 的研究则表明在流媒体应用下,需要部署更多的数据中心来保证终端用户服务质量^[43]。显然,上述问题并没有很好解决。因此,在规划部署服务节点时,如何决定服务节点的大小、分布的集中与分散与否,是一个亟需解决的问题。(2)服务节点的部署方式。服务节点的部署方式是指如何在被选择的位置部署节点,通常有深度部署和引入部署两类。前者是指将服务节点直接部署到运营商的网络中,后者则是在已选定的位置构建服务节点,然后通过私有光纤将该节点与运营商的网络直接相连。显然,这两种不同部署方式反映着内容网络与运营商之间的不同关系,对内容网络的构建、系统的服务性能有重要影响。但传统服务器部署模型关注于服务节点位置的选择,忽略服务节点在 Internet 中的部署方式问题。如何处理这种与运营商的关系目前也没有很好的模型支持。

新环境下的内容网络与 Internet 关系更紧密,所处的网络环境更复杂,所涉及的影响因素和关系更多更复杂,传统的节点部署模型已不能应对服务节点部署的新挑战,亟需新模型来分析和描述新的部署问题。Agarwal 等人从云计算服务系统服务节点部署问题的角度出发,试图找出在新网络环境和新应用环境下能够很好描述各种复杂关系的服务节点部署方案^[44]。同时,传统服务节点部署的搜索算法的效率会随内容网络规模扩大而急剧下降,因此,面对内容网络规模的扩大,设计出高效的搜索算法,也是需要重点研究的内容。

传统服务节点部署理论研究多采用模拟或仿真来验证其模型构建或部署策略的有效性。但这些验证方式较为简单,限制了网络环境的变化,不能很好地反映真实的网络环境,而实证分析是评估模型和算法有效性,提高并完善模型和算法的有效途径。因此,如何构建灵活且与实际系统接近的实验平台,如何在构建的实验平台上再现和模拟真实网络环境,从而很好地评价服务节点部署的理论模型和部署策略,是需要重点研究的内容。

4.2.2 解决方案

根据面临的挑战和研究内容,我们提出了一种系统的解决方案,如图 2 所示。该方案的核心思想是

对当前 Internet 网络和典型内容网络进行测量,以获取设计节点部署策略时所需信息.通过对这些信息的分析,找出影响节点部署的关键因素.在此基础上,分析已有节点部署模型的优缺点,引入博弈论,构建服务节点部署模型,并设计出有效的搜索算法,找出合适的部署方案.最后,通过实证分析的方法,验证和分析理论模型的正确性和部署策略的有效性.上述方案包括如下 3 个核心部分:

(1) 影响因子的确定. 影响因子的确定是指抽取众多影响内容网络服务节点部署性能因素中的核心因素,有利于简化并构建节点部署模型.一种有效的解决方法是通过对典型内容网络内部结构进行深入的分析,找出影响服务节点部署的内部影响因子;通过对内容网络外部环境及与其他系统的关系深入分析,找出影响内容网络服务节点部署的外部影响因子.最后,找出关键影响因子.

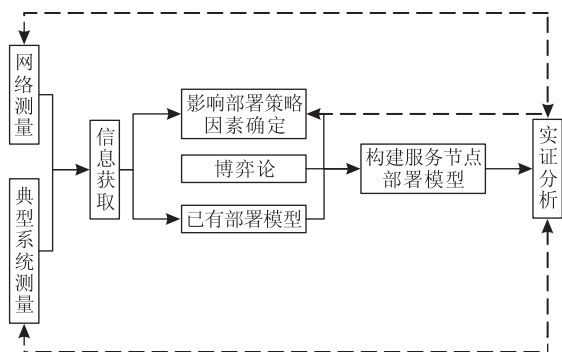


图 2 解决方案

(2) 部署模型构建. 新环境下的内容网络与 Internet 关系更紧密,但传统服务节点部署模型仅以内容网络本身为考察对象,不能很好描述内容网络与其它系统之间的复杂关系.为此,解决方案引入博弈论,通过已知关键影响因子,找出合适的优化目标.然后,通过博弈论相关理论来描述内容网络与其它系统间的关系,构建基于博弈论的服务节点部署模型,并设计相应的求解方法,以获得能实现全局优化的服务节点部署策略.

(3) 实证分析. 实证分析一直是服务节点部署问题研究面临的挑战之一.有效的解决方案就是通过在现有教育网或一些开放的试验网络(如 Coral^①、DipZoom^②等)上搭建接近真实网络环境的测试平台,测试确定的影响因素部署模型的正确性和部署方案的有效性.

上述三个核心步骤分别有效解决了传统服务节点部署理论在当前网络环境下面临的挑战,能适应

当前网络环境和内容网络的发展.

5 总 结

节点部署问题是内容网络的基础和核心问题,是学术界和工业界研究的热点,产生了众多成果.但随着网络技术、通信技术的发展,Internet 规模急剧扩大,各种新应用随之产生.这使得传统服务节点部署理论不能适应新的应用和新的网络环境,面临新的挑战.本文从数学模型和服务节点部署理论在内容网络中的应用等角度详细分析了现有的服务节点部署理论以及这些理论在当前网络环境和新的应用条件下面临的新问题和存在的挑战.在此基础上,文章总结出在当前网络环境和新应用条件下内容网络服务节点部署研究可能存在的趋势及解决方案.

参 考 文 献

- [1] Blumenthal M S, Clark D D. Rethinking the design of the Internet: The end to end arguments vs. the brave new world. *ACM Transactions on Internet Technology*, 2001, 1(1): 70-109
- [2] Feldmann A. Internet clean-slate design: What and why? *ACM SIGCOMM Computer Communications Review*, 2007, 37(3): 59-64
- [3] Goldengorin B, Ghosh D et al. Branch and peg algorithms for the simple plant location problem. *Computers & Operations Research*, 2004, 31(2): 241-255
- [4] Drezner Z, Hamacher H W. *Facility Location: Applications and Theory*. Springer, 2004: 132-141
- [5] Korkel M. On the exact solution of large-scale simple plant location problems. *European Journal of Operational Research*, 1989, 39(2): 157-173
- [6] Ryu C, Guignard M. An efficient algorithm for the capacitated plant location problem. Working Paper 92-11-02, Decision Sciences Department, University of Pennsylvania, The Wharton School, 1992
- [7] Harkness J, ReVelle C. Facility location with increasing production costs. *European Journal of Operational Research*, 2003, 145(1): 1-13
- [8] Geoffrion A M, McBride R. Lagrangean relaxation to capacitated facility location problems. *AIIE Transactions*, 1978, 10(1): 40-47
- [9] Van Roy T J. A cross decomposition algorithm for capacitated facility location. *Operations Research*, 1986, 34(1): 145-163

① Coral-The Coral Content Distribution Network, <http://www.coralcdn.org/>

② DipZoom-Deep Internet Performance Zoom, <http://dipzoom.case.edu/>

- [10] Tcha D, Lee B. A branch and bound algorithm for the multi-level uncapacitated facility location problem. *European Journal of Operational Research*, 1984, 18(1): 35-43
- [11] Geoffrion A M, Graves G W. Multicommodity distribution system design by Benders decomposition. *Management Science*, 1974, 20(5): 822-844
- [12] Weiszfeld E. Sur le point pour lequel la somme des distances de n points donnees est minimum. *Tohoku Mathematical Journal*, 1937, 43(190): 355-386
- [13] Miehe W. Link-length minimization in networks. *Operations Research*, 1958, 6(2): 232-243
- [14] Rosing K E. An optimal method for solving the (generalized) multi-Weber problem. *European Journal of Operational Research*, 1992, 58(3): 414-426
- [15] Mirchandani P B, Oudjit A et al. Multidimensional extensions and a nested dual approach for the m -median problem. *European Journal of Operational Research*, 1985, 21(1): 121-137
- [16] Brimberg J, Hansen P et al. Improvements and comparison of heuristics for solving the multisource Weber problem. *Operations Research*, 2000, 48(3): 444-460
- [17] Laporte G, Louveaux F et al. Exact solution to a location problem with stochastic demands. *Transportation Science*, 1994, 28(2): 95-103
- [18] Blanchini F, Rinaldi F et al. A network design problem for a distribution system with uncertain demands. *SIAM Journal on Optimization*, 1997, 7(2): 560-578
- [19] Chun B-G, Chaudhuri K et al. Selfish caching in distributed systems: A game-theoretic analysis//*Proceedings of the Twenty-Third Annual ACM Symposium on Principles of Distributed Computing*. St. John's, Newfoundland, Canada, 2004: 21-30
- [20] Chekuri C, Chuzhoy J et al. Non-cooperative multicast and facility location games//*Proceedings of the 7th ACM Conference on Electronic Commerce*. Ann Arbor, Michigan, USA, 2006: 72-81
- [21] Cardinal J, Hoefer M. Non-cooperative facility location and covering games. *Theoretical Computer Science*, 2010, 411(16-18): 1855-1876
- [22] Krishnan P, Raz D et al. The cache location problem. *IEEE/ACM Transactions on Networking*, 2000, 8(5): 568-582
- [23] Radoslavov P, Govindan R et al. Topology-informed Internet replica placement//*Proceedings of the WCW'01: Web Caching and Content Distribution Workshop*. Boston, MA, 2001
- [24] Li B, Deng X, Go L-M et al. On the optimal placement of Web proxies in the Internet; The linear topology//*Proceedings of the 8th IFIP Conference on High Performance Networking*. Vienna, 1998: 485-495
- [25] Qiu L, Padmanabhan N V et al. On the placement of Web server replicas//*Proceedings of IEEE INFOCOM*. Anchorage, Alaska, USA, 2001: 1587-1596
- [26] Jamin S, Jin C et al. Constrained mirror placement on the Internet//*Proceedings of IEEE INFOCOM*. Anchorage, Alaska, USA, 2001: 31-40
- [27] Radoslavov P, Govindan R et al. Topology-informed Internet replica placement//*Proceedings of the WCW'01: Web Caching and Content Distribution Workshop*. Boston, MA, 2001
- [28] Oppenheimer D, Chun B et al. Service placement in a shared wide-area platform//*Proceedings of the Annual Conference on USENIX '06 Annual Technical Conference (ATEC'06)*. Berkeley, CA, USA, 2006: 26-26
- [29] Cha M, Moon S et al. Placing relay nodes for intra-domain path diversity//*Proceedings of the IEEE INFOCOM*. Barcelona, Spain, 2006: 1-12
- [30] Cohen R, Nakibly G. A traffic engineering approach for placement and selection of network services//*Proceedings of the IEEE INFOCOM*. Anchorage, Alaska, USA, 2007: 1793-1801
- [31] Guha S, Khuller S. Greedy strikes back: Improved facility location algorithms//*Proceedings of the 9th Annual ACM-SIAM Symposium on Discrete Algorithms*. San Francisco, California, USA, 1998: 649-657
- [32] Krishnan P, Raz K D et al. The cache location problem. *IEEE/ACM Transactions on Networking*, 2000, 8(5): 568-582
- [33] Buchholz S, Buchholz T. Replica placement in adaptive content distribution networks//*Proceedings of the 2004 ACM Symposium on Applied Computing*. New York, NY, USA, 2004: 1705-1710
- [34] Laoutaris N, Smaragdakis G et al. Distributed placement of service facilities in large-scale networks//*Proceedings of the IEEE INFOCOM*. Anchorage, Alaska, USA, 2007: 2144-2152
- [35] Zhou S, Katto Jiro et al. Replication algorithms to retrieve scalable streaming media over content delivery networks//*Proceedings of the 5th ACM SIGMM*. Berkeley, California, USA, 2003: 255-261
- [36] Almeida J M, Eager D L. Minimizing delivery cost in scalable streaming content distribution systems. *IEEE Transactions on Multimedia*, 2004, 6(2): 356-365
- [37] Cahill A J, Sreenan C J. An efficient CDN placement algorithm for the delivery of high quality TV content//*Proceedings of the 12th Annual ACM International Conference on Multimedia*. New York, USA, 2004: 975-976
- [38] Ahuja S, Krunz M. Algorithms for server placement in multiple-description-based media streaming. *IEEE Transactions on Multimedia*, 2008, 10(7): 1382-1392
- [39] Shi S, Turner J. Placing servers in overlay networks//*Proceedings of the International Symposium on Performance Evaluation of Computer and Telecommunication Systems (SPECTS) 2002*. San Diego, CA, 2002
- [40] Roy S, Pucha H et al. Overlay node placement: Analysis, algorithms and impact on applications//*Proceedings of the 27th International Conference on Distributed Computing Systems*. Washington, DC, USA, 2007: 53
- [41] Allen M S, Zhao B Y et al. Deploying video-on-demand services on cable networks//*Proceedings of the 27th International*

Conference on Distributed Computing Systems. Toronto, Ontario, Canada, 2007: 63-71

- [42] Triukose S, Wen Z et al. Content delivery networks: How big is big enough? //Proceedings of ACM SIGMETRICS Performance Evaluation Review, 2009, 37(2): 59-60
- [43] Leighton L. Improving performance on the Internet. Com-

munications of the ACM, 2009, 52(2): 44-51

- [44] Agarwal S, Dunagan J et al. Volley: Automated data placement for geo-distributed cloud services//Appear in 7th USENIX Symposium on Networked Systems Design and Implementation. San Jose, CA, 2010



YIN Hao, born in 1974, Ph. D., associate professor. His current research interests include multimedia, content delivery networks, Peer to Peer and performance evaluation.

YUAN Xiao-Qun, born in 1974, Ph. D. candidate. His current research interests include content delivery networks, Peer to Peer system, live streaming and source optimization.

LIN Chuang, born in 1948, Ph. D., professor, Ph. D. supervisor. His current research interests include computer networks, performance evaluation, logic reasoning, Petri net

theory and its applications.

ZHANG Fa, born in 1974, Ph. D., associate professor. His current research interests include computer algorithms, parallel processing and energy-efficient algorithm in network.

PANG Shan-Chen, Ph. D.. His current main research interests include Petri net theory and application, formal methods, distributed concurrent system analysis and verification.

LIU Zhi-Yong, born in 1946, Ph. D., professor. His research interests include computer algorithms, parallel processing and energy-efficient algorithm in network.

Background

With the increasing demand on Internet services, there has been a significant amount of effort on the development of placement strategies for resources including servers and bandwidth to further improve Internet services performance. Placement strategies are important because appropriate allocation of resources benefits content providers by reducing latency for their clients, and benefits the hosting platform providers by reducing bandwidth consumption. The essential challenge that the ASP (application service provider) faces is to provide content distribution to clients with good Quality of Service (QoS) while retaining efficient and balanced resource consumption of the underlying infrastructure. Central to addressing this challenge is the careful placement of server replicas and the dissemination of requests.

From this perspective of mathematics, service node placement can be transformed to an facility location problem, which can be modeled as an optimization problem, such as Stochastic Location Model, Graph-based Location Model and Game Theory Location Model. Among them, the Stochastic Location Model considers that there is few prior knowledge and designers design their placement strategies to satisfy the service performance under worst network states; unlike this model, Graph_facility Location Model considers that the designers know the topology, so they apply Graph theory into this problem and formulate this problem as this problem

transformed the facility location as an optimization problem.

However, the Game Theory Location Model focuses on the Interaction between different systems, and is suitable for dynamic server placement.

The problem of replicated server placement has been studied extensively in the past in several papers. Li et al. proposed an algorithm for placing k Web proxies on the network by solving an optimization problem given the request pattern and performance constraints as input. The optimized function is usually the overall latency averaging across clients. Later work studied different placement strategies or with diverged constraint metrics under the same problem formulation. Some researchers tackle this problem from a different angle, i. e., minimizing the number of replicas while meeting clients' latency constraints and servers' capacity constraints. Intelligent data structures are proposed such as dissemination trees to improve efficiency. Besides the effort of innovating theoretical solutions, another group of work focuses on validating these proposals in reality.

In this paper, the authors systematically survey the existing service node deployment theories and their applications, address the challenges faced. For the challenges, the authors also present the research tendencies and the potential solutions in final.