

基于多属性信息的数据中心间数据传输调度方法

李阳阳, 王洪波, 张鹏, 董健康, 程时端

(北京邮电大学 网络与交换技术国家重点实验室, 北京 100876)

摘 要: 给出一种基于多属性信息综合评价的数据中心间数据传输调度方法。首先利用层次分析法, 对多个属性之间的从属关系进行分析, 然后根据属性值的分布差异由信息熵计算属性的相对权重, 给出对于不同备选中转数据中心的评价。通过建立时间扩展图模型将基于多属性信息的数据中心间数据传输调度问题形式化为最小代价流问题。通过设置不同的参数, 对一般存储转发方法、基于单属性信息的方法以及基于多属性的方法在不同属性上的性能优化差异做了实验对比。结果表明, 相比于其他方法, 该方法能够综合考虑多种属性信息, 为数据中心间的数据传输选择综合评价最优的路径。

关键词: 云计算; 数据中心; 多属性决策; 数据传输调度; 路径选择

中图分类号: TP301

文献标识码: A

文章编号: 1000-436X(2012)Z1-0121-11

Multi-attribute aware scheduling for inter-datacenter bulk transfers

LI Yang-yang, WANG Hong-bo, ZHANG Peng, DONG Jian-kang, CHENG Shi-duan

(State Key Laboratory of Networking and Switching Technology, Beijing University of Posts and Telecommunications, Beijing 100876, China)

Abstract: A multi-attribute decision making method to schedule inter-datacenter bulk data was proposed. First, hierarchical analysis was used to analyze the dominant relationship among attributes. Based on the relative weights of attributes which were calculated via information entropy, the evaluation to each intermediate data center was provided. Then the multi-attribute aware scheduling problem was modeled on a time expanded graph and formulated as a minimal cost flow problem. Finally, the scheme with general store-and-forward mechanism and single attribute based approach under several different simulation settings was compared. The results demonstrate that our strategy can take multiple attributes into consideration, and can find the bulk transfer route with the best integrated evaluation.

Key words: cloud computing; inter-datacenter; bulk transfers; multi-attribute decision making; route choice

1 引言

随着云计算和互联网在线业务的迅速发展, 互联网巨头如微软^[1]、谷歌^[2]、亚马逊^[3]在全球各地兴

建了多个大规模的数据中心, 将用户的数据复制到多个地理位置以提高用户服务质量。一些内容分发网络提供商如 Akamai^[4]、ChinaCache^[5]等也在不同地理位置建立互联网数据中心来加快用户请求的

收稿日期: 2012-07-10

基金项目: 国家自然科学基金资助项目 (61002011); 软件开发环境国家重点实验室开放课题计划基金资助项目 (SKLSDE-2009KF-2-08); 国家重点基础研究发展计划 (“973” 计划) 基金资助项目 (2009CB320505); 国家高技术研究发展计划 (“863” 计划) 基金资助项目 (2011AA01A102)

Foundation Items: The National Natural Science Foundation of China (61002011); The Open Fund of the State Key Laboratory of Software Development Environment (SKLSDE-2009KF-2-08); The National Basic Research Program of China (973 Program) (2009CB320505); The National High Technology Research and Development Program of China (863 Program) (2011AA01A102)

响应速度。这些分布式的数据中心之间需要传输大量的数据。测量研究表明,数据中心之间的业务流量已经占据了接近一半的数据中心出口带宽^[6]。为满足应用的峰值带宽需求,数据中心管理者通常向运营商超额地订购接入带宽^[7,8],但是由于数据中心处于忙闲不均的使用情况,导致数据中心之间的链路并没有得到有效利用。因此,如何对数据中心间的数据进行传输调度是当前亟待解决的问题。

由于数据中心的接入带宽使用情况是随着时间而变化的,传统的路由选择技术并不能解决现有问题。例如当数据需要传输时,如果目的数据中心的下行链路处于过载状态,那么源、目的数据中心间就不存在带宽充裕的路径。LAOUTARIS 等人^[9]提出,虽然当前不存在这样的路径,但是有的数据中心却处于低负载的状态。因此,可通过使用接入带宽空闲的数据中心进行中转,待中转数据中心具备到目的数据中心的带宽富裕路径时再将数据进行转发。利用这样的存储转发机制,当前研究者们分别提出以提高数据中心接入带宽利用率^[9]、降低数据中心间流量传输成本^[10,11]、减少域间传输流量^[12]为目标的路径选择优化机制。然而在实际的数据传输路径选择过程中,经常会遇到管理者不仅关心数据中心接入带宽使用情况,而且对于链路费用、服务质量也有要求;或者管理者对其中的某一项属性要求较高而对其他属性要求不是很高的情况。所以仅考虑单属性或单一优化目标的路径选择是不全面的。

为克服仅仅依靠单属性信息进行数据中心间数据传输路径选择的不足,本文提出了一种基于多属性决策理论的数据中心间数据传输调度方法。考虑到管理者在进行路径选择时面临多属性或者多目标,并且不同属性还可能存在从属关系。本文利用层次化的方法分析不同属性间的从属关系,并使用信息熵理论确定不同属性的权重,从而给出多属性信息下中转数据中心的综合评价。通过建立时间扩展图模型将基于多属性信息的数据中心间数据传输调度问题形式化为最小代价流问题,并给出了求解该问题的算法。

本文的贡献总结如下。

1) 首次从实际情况出发,考虑多属性信息条件下的数据中心间数据传输调度问题。

2) 利用基于层次分析的多属性决策方法给出对于不同中转数据中心的综合评价。

3) 给出基于多属性信息的数据传输调度方法,通过大量仿真实验验证,相比于已有方法,能够给出综合评价最优的路径选择。

本文后续安排如下:第 2 节介绍相关研究工作;第 3 节对问题进行建模;第 4 节给出基于多属性信息的数据中心间数据调度方法;第 5 节进行仿真实验评价;第 6 节总结全文。

2 相关工作

数据中心网络是当前云计算领域的一个重要研究方向,目前已有很多工作研究数据中心内部的网络结构、路由及流量优化等问题。近一两年来,对于数据中心之间的网络问题得到了科研人员的广泛关注,数据中心间的数据传输调度更是最近的研究热点,研究者们从不同角度和优化目标出发,提出了一些方法。

在文献[9]中,作者发现在同一时刻,分布在不同时区的数据中心具有不同的带宽使用情况。他们将整个地理分散的数据中心构造成一个存储节点网络,利用中转数据中心对数据存储转发来充分利用未使用的带宽。Feng 等人^[10]提出连接数据中心的链路是由不同运营商提供的,而这些运营商索取的链路价格存在差异,因此他们设计了一个存储转发机制来最小化数据中心间流量传输的成本。在文献[11]中,作者也提出了一个减少带宽成本的方法,通过合理的作业调度降低对于峰值带宽的消耗。在文献[12]中,本文作者提出了一种运营商友好的以最小化域间链路流量为目标的数据中心间路径选择方法,根据作者的实验,域内链路时延通常是域间链路时延的 1/5~1/3,因此通过优先选择时延较低的链路来尽量使用域内链路,从而减少域间流量。

上述工作都是从单属性或单目标的角度出发,然而实际中,数据中心的管理者需要综合考虑多种属性信息。就作者所知,目前还没有研究工作从多属性角度出发,研究数据中心间的数据传输调度问题。

3 问题描述及建模

3.1 基于层次化的属性权重分析

在考虑多属性信息下数据中心间数据传输路径选择时,面临的一个重要问题是,多个属性之间存在一定的相互制约关系。例如,有的管理者需要考虑数据中心间链路的服务质量,而对该属性的评

价,实际是由带宽、网络时延、可靠性等属性决定的。因此,本文借鉴了美国运筹学家 Saaty 提出的层次分析法(AHP, analytic hierarchy process)^[13],使用层次化的结构来构造多属性决策问题。因为数据中心通常具有高效的监测管理工具,关键的属性值都可以通过监测定量获取。所以,相比于经典的层次分析法,不需要通过 1~9 标度法构造判断矩阵来分析定性的属性信息,避免了由于心理等主观因素带来的人为误差。

图 1 显示了基于多属性信息数据中心间数据调度问题的属性层次化结构。围绕最终综合评价目标,将多个属性信息根据其依赖关系划分成不同的层次。其中第 0 层是对多个属性信息的最终总体评价。第 1 层是管理者需要关心的核心属性,如中转数据中心的接入带宽成本、接入链路的服务质量、数据中心能源使用效率(PUE, power usage effectiveness)等。第 2 层是对第 1 层核心属性的分解,如服务质量可以被分解成带宽和时延 2 个子属性。属性的分解可以根据属性的分类及管理经验的日常管理经验获取,比如,链路的服务质量与几个基本的要素如链路的带宽及时延等有关,而内容的流行度同用户所在的地理位置以及所在位置的人口规模等有关。如果第 2 层属性还有子属性,可以进一步构造第 3 层、第 4 层结构直至第 $k-1$ 层。第 k 层是各个备选方案,代表上述属性取值的来源,即各个备选中转数据中心。

下面基于属性的层次结构,构造基于多属性信息的决策矩阵来评价不同备选数据中心的偏好次序。首先给出原始决策矩阵如式(1)所示。

$$\mathbf{D}_{\text{origin}} = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1n} \\ x_{21} & x_{22} & \cdots & x_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ x_{m1} & x_{m2} & \cdots & x_{mn} \end{bmatrix} \quad (1)$$

矩阵的行代表有 m 个决策备选方案 $\{A_i\}$,在本文研究的问题里为 m 个备选中转数据中心。列代表 n 个不同的决策属性 $\{X_j\}$,包括带宽成本、带宽、往返时延、数据中心能源使用效率等等。矩阵的每个元素 x_{ij} 代表数据中心 i 的第 j 项属性值。根据上文提出的层次化结构模型以及属性之间的从属关系,本文将原始的决策矩阵 $\mathbf{D}_{\text{origin}}$ 分解成若干个子决策矩阵。每一个子决策矩阵可以自上而下的逐层构造。例如,在本文的图示的结构模型中,第 1 层的关于服务质量 QoS 和内容流行度的决策矩阵可以由第 2 层的子属性构造造成如式(2)和式(3)所示。

$$\mathbf{D}_{\text{QoS}} = \begin{bmatrix} \text{QoS}_{11} & \text{QoS}_{12} \\ \text{QoS}_{21} & \text{QoS}_{22} \\ \vdots & \vdots \\ \text{QoS}_{m1} & \text{QoS}_{m2} \end{bmatrix} \quad (2)$$

$$\mathbf{D}_{\text{pop}} = \begin{bmatrix} \text{Pop}_{11} & \cdots & \text{Pop}_{1n'} \\ \text{Pop}_{21} & \cdots & \text{Pop}_{2n'} \\ \vdots & \ddots & \vdots \\ \text{Pop}_{m1} & \cdots & \text{Pop}_{mn'} \end{bmatrix} \quad (3)$$

服务质量决策矩阵 $\mathbf{D}_{\text{QoS}}$ 和内容流行度决策矩阵 $\mathbf{D}_{\text{pop}}$ 的行仍由 m 个备选数据中心构成,但决策属性分别由服务质量属性的 2 个子属性:带宽、时延以及内容流行度的 n' 个子属性:人口、地理位置等构成。对每个决策矩阵进行归一化,并计算每个决

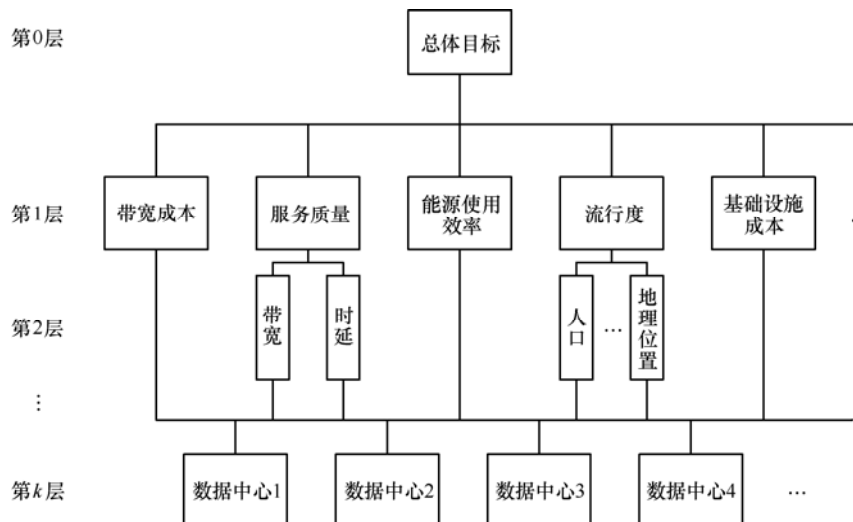


图 1 基于多属性信息数据调度问题属性层次结构

策子属性的权重, 可以生成对当前属性的评价结果, 标记为当前属性的值, 如根据 D_{QoS} 和 D_{pop} 可以生成不同备选决策方案对于服务质量属性和内容流行度属性的属性值。当获得该值后可以进一步构造第 0 层的决策矩阵, 即最终的决策矩阵为

$$D_{final} = \begin{bmatrix} X_{11} & X_{12} & \cdots & X_{1n} \\ X_{21} & X_{22} & \cdots & X_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ X_{m1} & X_{m2} & \cdots & X_{mn} \end{bmatrix} \quad (4)$$

其中, 决策属性的个数 n 只由第 1 层的 n 个核心属性决定。对该矩阵再次进行归一化处理和权重设置, 可以得到最终的决策评价结果。同理, 在具有更多层次的属性结构下, 也能通过子决策矩阵, 向上递归的构造最终综合评价结果。

3.2 基于信息熵的权重设置

上一小节讨论了利用层次分析的方法, 递归计算每个中转数据中心最终评价结果的方法。本节将讨论在通过子属性计算上层属性的评价价值(属性值)时, 如何决定不同子属性之间的权重。本文利用信息熵理论设置属性的相对权重, 并给出评价值的计算公式。

本文假设数据中心管理者对不同的属性并没有偏好, 属性的权重完全由不同数据中心的各属性值分布差异决定。其基本思想是: 当不同备选决策方案对于同一个属性的取值差异较大时, 该属性对于整个决策评价的影响就越大, 而当差异较小时, 该属性对于决策评价所起的作用就很小甚至可以忽略。信息熵由香农首次提出并应用到信息论中^[14], 根据信息熵的极值性, 熵的大小恰好能够反映出不同备选方案在同一个属性值上的接近情况, 属性值越接近, 熵值越大, 因此, 本文采用信息熵值来设置不同属性权重。在描述权重的设置前, 对上一小节构造的决策矩阵进行归一化。首先, 将所有类型的属性转化为成本型属性。在本文的模型中, 只有 2 类属性, 一类是效益型如带宽, 流行度等, 其属性值越高对其评价越好; 另一类属于成本型如带宽价格、时延等, 其属性值越低评价结果越好。本文通过用 $1/x_{ij}$ 取代 x_{ij} 将效益型的属性转化为成本型。然后, 将备选方案 A_i 关于属性 X_j 的属性值归一化为评价值

$$X_{ij} = \frac{x_{ij}}{\sum_{i=1}^m x_{ij}} \quad (5)$$

可以用 X_{ij} 逐一取代上一小节决策矩阵中的 x_{ij} 从而构造出评价矩阵

$$P = (X_{ij})_{mn} \quad (6)$$

根据香农公式, 计算属性 X_j 的信息熵

$$E_j = -k \sum_{i=1}^m X_{ij} \ln X_{ij}, \forall j \quad (7)$$

其中, k 是常数, 为使 $0 \leq E_j \leq 1$, 令 $k = \frac{1}{\ln m}$ 。

信息的偏差度定义为

$$d_j = 1 - E_j \quad (8)$$

本文假设数据中心管理者没有属性偏好。一种简单的设置不同属性权重的方法是令

$$w_j = \frac{d_j}{\sum_{j=1}^n d_j} \quad (9)$$

本文使用算术加权平均算子对备选数据中心的评价值进行计算, 其最终评价值可以用层次分析法递归求解为

$$U_i^{final}(X_{i1}, X_{i2}, \dots, X_{in}) = \sum_{j=1}^n w_j X_{ij}, \forall i \quad (10)$$

由于所有的属性都被转化为成本类型, U_i^{final} 的值越低, 对备选数据中心 A_i 的评价结果越好。

3.3 基于时间扩展图的数据中心互联网络建模

高效的数据中心通常部署许多强大的管理工具来监控数据中心的运行状态。例如, 带宽的使用情况可以采用检测工具如亚马逊的 CloudWatch^[3] 进行细粒度的周期性采集; 网络链路的服务质量可以通过网络测量工具采集(如 netperf^[15]); 能源使用效率也可以用一些精密的部件如 IBM 的 WebSphere sMash^[16] 获取。通过对这些工具获取的信息进行分析, 可以运用有效的预测算法对今后一段时间内的属性取值进行合理的预测。本文假设所有随时间变化的属性信息可以以一种可预测的方式获取。

同文献[9,10,12]类似, 将随时间变化的数据中心互联网络建立在时间扩展图上^[17]。将一个时间段划分成若干个时间间隔, 根据时间间隔的数量, 对每个数据中心引入多份“虚拟”拷贝, 即在每个间隔将所有节点复制一份表示数据中心在当前时间间隔的状态。因此时间维度以一种时间扩展的方式展现出来。

在本文的模型中,数据中心以及连接它们的覆盖层链路一起构成了一个数据中心互联网络。该网络可以被建模成一张时间扩展图 $G=(N,E,T,x_i^0(t),\dots,x_i^k(t)),i=1,2,\dots,M,k=0,1,\dots,K$,这是一个完全有向图, N 代表数据中心集合, E 代表连接数据中心的覆盖层链路集合。尽管大多数的数据传输是时延容忍的,但是这些数据仍然需要在一个相对不是很严格的截止时间内被传输,比如10个小时、2天等等。因此本文令 T 代表可以接受的最大传输时间,由时间间隔的数量表示,而每个时间间隔的长度可以根据需要进行设置,如可以设置为属性值采集周期的最小公倍数。 $x_i^k(t)$ 代表在 t 时刻数据中心 i 的属性 k 的取值,需要强调的是, i 的数量 M 代表的是图中数据中心虚拟拷贝的总的个数,由 $M=N\times T$ 计算得到,所有的属性都是随着时间变化的。当前在边 e_{ij} 上的流量可以用 $f_{ij}(t)$ 来表示,它也是随着时间而变化的。

图2给出了一个基于时间扩展的数据中心互联网络示意图,图2(a)是静态模型,表示4个数据中心彼此互联,有200单位的数据需要由数据中心1传输到数据中心4。图2(b)是与之相应的的时间扩展图模型,其中假设数据在3个时间间隔内完成传输,因此数据中心在图中的表示被拷贝成3组12个不同的虚拟节点,其属性值也取决于当前的时刻 t 。根据已提出的^[9~12]和本文将要介绍的数据传输调度方法,数据被临时存储或被分成多块进行传输。

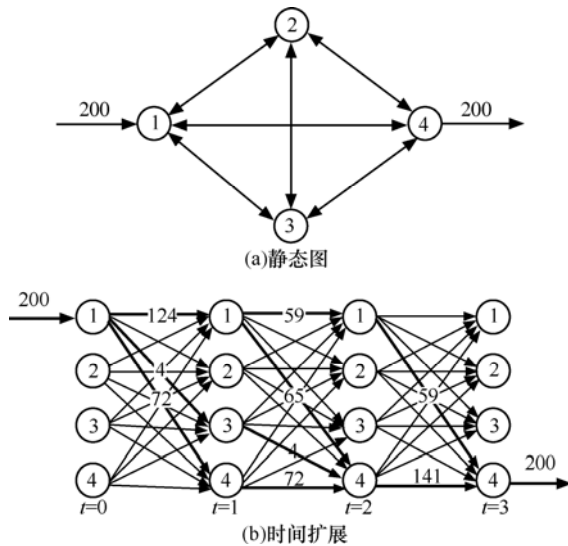


图2 时间扩展的数据中心互联网络实例

本文的一些重要概念在表1中做了总结。

表1 系统模型中一些重要概念

符号	描述
i	第 i 个数据中心
e_{ij}	连接数据中心 i 和数据中心 j 的覆盖层链路
T	最大期待传输时间
t	数据传输时刻
$f_{ij}(t)$	t 时刻,在链路 e_{ij} 上的当前流大小
$X_i^k(t)$	t 时刻,数据中心 i 的第 k 个属性取值

4 基于多属性信息的数据中心间数据传输调度方法

本节首先介绍一般的存储转发方法,接着以D4D方法为例介绍考虑单属性信息的数据传输调度方法,最后给出本文提出的基于多属性信息的数据中心间数据传输调度方法。

4.1 原始的存储转发方法

在原始方法中,每条链路将多个备选数据中心看作是等价的,除接入带宽约束外不考虑任何其他属性。其路径选择方法是只要到中转数据中心的路径具有空闲的带宽能够承载数据,那么该条链路就会被选择。因此,问题可以被形式化为

$$\text{minimize} \quad z = \sum_{t=0}^{T-1} \sum_{e_{ij} \in E} f_{ij}(t) \quad (11)$$

$$\text{s.t.} \quad \sum_{t=0}^{T-1} \sum_{e_{ij} \in E} f_{ij}(t) = F \quad (12)$$

$$\sum_{e_{ij} \in E} f_{ij}(t-1) - \sum_{e_{ji} \in E} f_{ji}(t) = 0 \quad i \neq 1, j \neq n \quad (13)$$

$$\sum_{t=0}^{T-1} \sum_{e_{jn} \in E} f_{jn}(t) = F \quad (14)$$

$$0 \leq f_{ij}(t) \leq u_{ij}(t), \forall e_{ij} \in E \quad (15)$$

在该形式化中,为了与下文描述一致,本文采用了式(11)所示的最小化表现形式,但由于每条链路是等价的,且数据块大小是给定的,因此,所有可行路径的目标值也是相等的。该问题下的求解结果实际是返回一条任意可行的路径。式(12)表述源数据中心发送的数据量之和等于数据总量 F 。式(13)描述了除源、目的数据中心外,其他中转数据中心的流量守恒。式(14)表示目的数据中心接收到

的数据量之和也等于数据的大小 F 。最后的不等式 (15) 表明每条流的大小满足链路容量约束。

4.2 基于单属性信息的方法

在文献[12]中,提出了一种运营商友好的数据中心间数据传输方法,记为 D4D 方法。相比于原始的存储转发方法,考虑了单属性的信息,即链路时延信息,提出减少数据中心间数据传输时产生的跨运营商的流量。为了实现这个目标,我们为每条链路添加了一个代价,如果 2 个数据中心接入同一个运营商的链路,那么该条链路的代价会设置的较低,反之则较高。选择一条具有最小代价和的路径来存储转发数据,尽可能的使数据在同一个运营域内进行转发,从而显著地减少域间的数据传输流量。式(11)可以被改写为

$$\text{minimize} \quad z = \sum_{t=0}^{T-1} \sum_{e_{ij} \in E} f_{ij}(t) \times c_{ij}(t) \quad (16)$$

$c_{ij}(t)$ 代表在 t 时刻数据中心 i 和数据中心 j 之间链路的代价。在本文的实验中^[12],发现域间链路的往返时延值是域内链路的 3~5 倍,因此,可以采纳往返时延值作为链路代价设置的参考值。因此 D4D 方法又可以被看作是解决以往往返时延值作为单属性考虑的数据中心间数据调度问题。

4.3 基于多属性信息的方法

基于多属性信息的数据中心间数据传输调度问题可以作以下描述。

当在 t 时刻准备传递数据时,需要同时考虑:选择接入带宽价格最低的数据中心中转数据以节省带宽成本;选择接入链路服务质量最好的数据中转数据以提高数据传输性能;选择能源使用效率最高数据中转数据以节省能源消耗等。

这就意味着面临一个多属性决策的问题,而数据中心的所有属性随着时间而变化使得该问题变得更加复杂。在第 3 节提出的系统模型中,使用层次化分析方法给出对每个中转数据中心的评价,并将随着时间变化的数据中心互联网络建模成一张时间扩展图。由于本文建立的最终评价函数 $U_i(t)$ 是一个成本型的函数,这就意味着函数值越低,得到的对备选方案 A_i 的评价结果越好。因此,基于之前的工作,一个很自然的想法就是用评价函数来代替 D4D 方法中的链路代价,从而将基于多属性信息的数据传输调度问题也形式化成一个最小代价流问题。需要做改进的是,函数 $U_i(t)$ 是对 t 时刻数据中

心 i 的综合评价,并不是对链路代价的描述,但是可以用接收端数据中心的综合评价来取代链路的代价。其现实含义是:在选择一条中转链路时,要综合考虑链路接收端数据中心的接入带宽、接入带宽价格、能源使用效率等因素。因而表达式(11)可以再次被改写为

$$\text{minimize} \quad z = \sum_{t=0}^{T-1} \sum_{e_{ij} \in E} f_{ij}(t) * U_j^{\text{final}}(t) \quad (17)$$

为了解决该最小代价流问题,本文使用了经典的负价环算法,基于多属性信息的数据中心间数据调度算法如算法 1 所示。该算法由数据中心的管理者根据数据中心监测工具采集的属性值信息在数据块需要调度时进行计算,为数据块的传输选择一条综合评价最优的路径。其基本思想是首先由 Edmonds-karp 算法计算出一条可行流(通常也是最大流),然后生成相应的残量网络,用 Ballman-ford 算法循环判断图中是否存在负价环,如有则消除负环并更新残量网络,直至图中不存在代价为负的环境,最终得到的流,即为最小代价流。

算法 1 基于多属性信息的数据中心间数据调度算法

输入: $G(N,E)$: 具有 N 的数据中心, E 条链路的数据中心互联图。

其中每条边 $e_{ij} \in E$ 有属性 $u_{ij}(t)$ //链路的容量大小

每个节点 $i \in N$ 有代价函数 $c_i(t)$ //链路代价函数,在不同方法中分别取不同函数

T : 最大期待传输时间

s : 源数据中心

d : 目的数据中心

输出: 由 s 到 d 的基于多属性信息的数据传输调度路径

1) 根据原始图 $G(N,E)$ 生成相应的时间扩展图 $G'(N',E')$

2) 在 $G'(N',E')$ 中建立一条可行流 x

3) $E_1 = e_{ij'} \in E' : f_{ij'} < u_{ij'}$

4) $E_2 = e_{ij'} \in E' : f_{ij'} > 0$

5) $G'(x) = \{N', E_1 \cup E_2\}$

6) while $G'(x)$ 包含一个负价环 do

7) 找到其中的一个负价环 W

8) $\delta := \min\{r_{ij'} : \forall e_{ij'} \in W\}$

9) 沿着环 W 的方向增加 δ 单位的流量,

并更新 $G'(x)$:

$$10) \quad f_{ij} = f_{ij'} + \delta, \forall e_{ij'} \in E_1 \cap W$$

$$11) \quad f_{ij} = f_{ij'} - \delta, \forall e_{ij'} \in E_2 \cap W$$

12) end while

13) return s 到 d 的路径

4.4 算法的时间复杂度分析

在寻找一条可行流时, Edmonds-karp 算法的时间复杂度是 $O(|V||E|^2)$ 。而寻找负价环需要进行至多 $|E|CQ$ 次迭代, 其中 C 是网络最大容量, Q 是网络最大代价。在每次迭代时, 使用 Ballman-ford 算法的时间复杂度是 $O(|V||E|)$, 更新残量网络的时间复杂度是 $O(|E|)$, 最后需要更新每条边的流量, 其时间复杂度也是 $O(|E|)$ 。因此最终算法的时间复杂度为

$$O(|V||E|^2) + |E|CQ(O(|E|) + O(|V||E|) + O(|E|)) \\ = O(|V||E|^2 CQ) \quad (18)$$

5 实验评价

本节将进行一系列的实验对原始存储转发、D4D 以及基于多属性信息等方法进行对比。在实验中, 设置了不同的最大期待传输时间, 比较 3 种方法分别在带宽成本、带宽大小、往返时延以及能源使用效率等属性上的性能差异改进。

5.1 参数设置

实验评价是在编写的 C++ 代码基础上实现的。模拟了一个云基础设施提供商拥有 20 个数据中心分布在多个地理位置。所有的数据中心由覆盖层链路两两互联, 构成了一个完全有向图。在本次实验中, 考虑带宽成本、服务质量、带宽、往返时延、能源使用效率等属性, 其中服务质量属性又由带宽和往返时延两个子属性构成。作者认为这样的实验设置足以验证基于多属性信息方法的性能。数据中心的接入带宽价格由 $[1, 10]$ 单位随机生成, 带宽大小由 $[1, 100]$ 单位随机生成。根据在文献[12]中的测量, 不同数据中心的往返时延介于 30~180ms 之间, 因此令往返时延由 $[30, 180]$ ms 随机生成, 能源使用效率由 $[1, 3]$ 随机生成。特别地, 数据有可能被临时地存储以等待今后某一时刻进行转发。在对应的时间扩展网络中, 标记同一个数据中心的“虚拟”副本之间的带宽成本为 0, 带宽为 INTMAX 单位, 在本文实验环境中为 16 位整型数最大值, 即 32 767 单位, 往返时延为 0, 由于存储仍需使用数据中心

的能源, 因此能源使用效率值仍随机生成。数据块的大小由 200 单位开始, 以 200 单位递增直至当前网络能够容纳的最大可行流。在本实验中, 令期待的最大传输时间分别为 $T = 3, 4, 5, 6$ 个时间间隔, 3 种方法采用相同的参数设置以作比较。

5.2 实验结果及分析

根据本文使用的随机数生成函数特征, 数据中心在能源使用效率上的属性值差异较大, 该属性的相对重要性也越明显, 因此权重也较大。在多次实验中, 权重大小分别以能源使用效率, 带宽成本, 服务质量依次递减, 其中服务质量属性的子属性往返时延的权重要高于带宽的权重。表 2 显示了在不同实验参数下, 由层次分析法获得的属性熵权。

表 2		权重信息				
T	PUE	COST	QoS			total
			bandwidth	RTT		
3	0.343 381	0.331 251	0.440 053	0.559 947		0.325 368
4	0.344 546	0.331 648	0.428 876	0.571 124		0.323 806
5	0.344 208	0.330 418	0.439 437	0.560 563		0.325 373
6	0.343 406	0.330 687	0.441 818	0.558 182		0.325 906

以下分别比较 3 种方法在能源使用效率、带宽成本、带宽、往返时延 4 个维度上的性能差异。

1) 能源使用效率评价

评价单位数据传输时选择的中转数据中心的平均能源使用效率值, 如图 3 所示。

PUE 值是数据中心总设备能耗和 IT 能耗设备的比值, 其基准是 2, 越接近 1 能源使用效率越高。可见在不同的最大传输时间情形下, 基于多属性信息的方法(MA, Multi-attribute)在单位数据传输时使用到的中转数据中心的平均能源使用效率要明显高于原始的方法(Origin)和 D4D 方法。由于数据中心能源使用效率权重对综合评价结果影响较大, 所以基于多属性信息的方法尽可能的寻找能源高效的数据中心进行数据的存储转发, 以节省能源开销, 而其他方法不考虑能源使用效率因素, 因此在此指标上明显劣于基于多属性信息的方法, 实验结果符合方法设定。

2) 带宽成本评价

评价在完成一次数据传输时, 需要花费的总的带宽成本, 如图 4 所示。

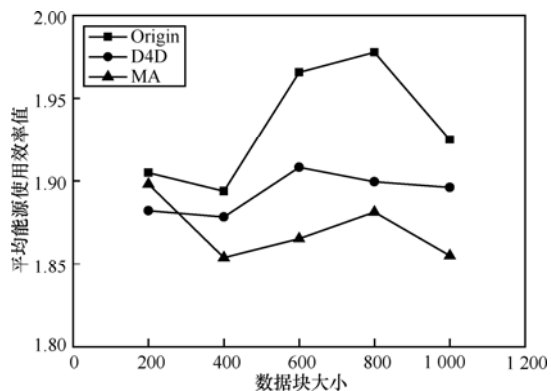
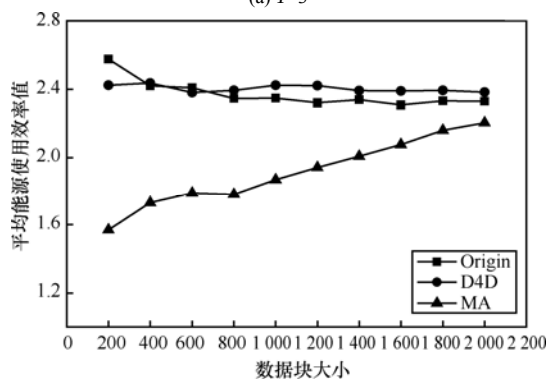
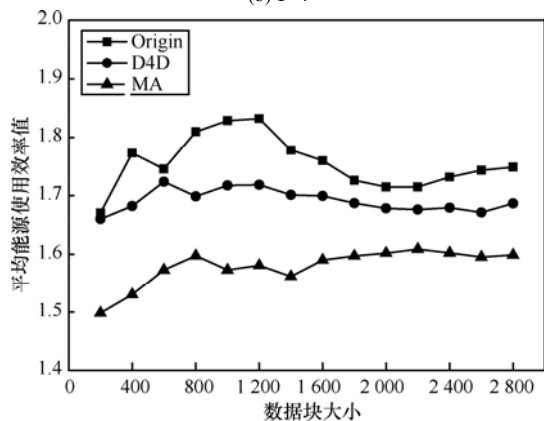
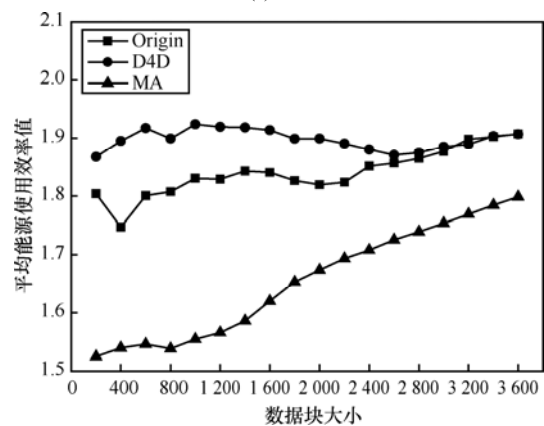
(a) $T=3$ (b) $T=4$ (c) $T=5$ (d) $T=6$

图3 平均能源使用效率评价

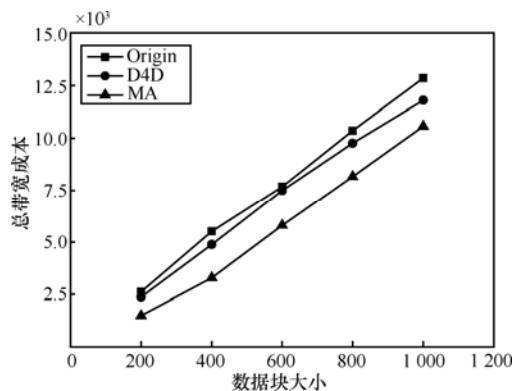
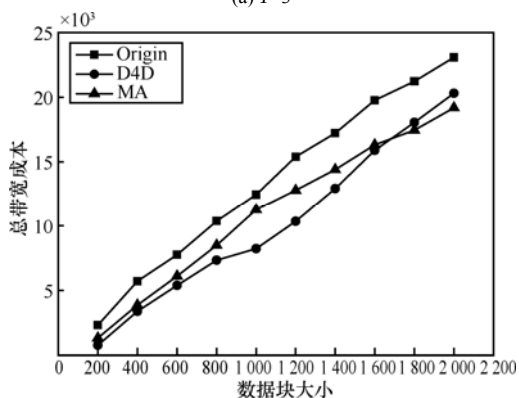
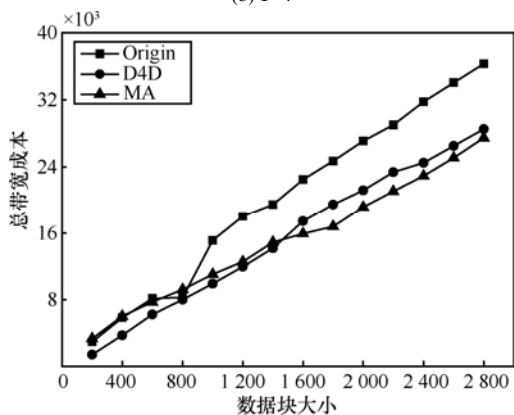
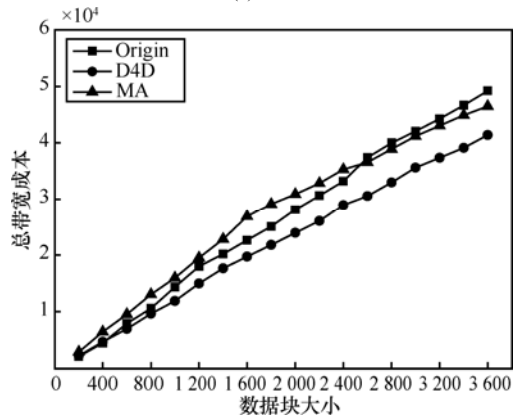
(a) $T=3$ (b) $T=4$ (c) $T=5$ (d) $T=6$

图4 总带宽成本评价

在该维度上, 总的带宽使用成本成为影响总体评价的次要因素, 因此基于多属性信息的方法在总的带宽使用成本上存在波动性, 在有的情况下优于 D4D 方法, 而有的情况下稍劣, 但绝大多数情况下都优于原始方法。

3) 带宽评价

评价在单位数据传输时所选择的中转数据中心的平均接入带宽大小, 如图 5 所示。

在该维度上, 由于 QoS 权重最低, 它对于总体评价的重要性相对影响最小, 作为其子属性的接入带宽的重要性就更小了。基于多属性信息的方法单位数据传输时使用的平均链路带宽容量小于原始方法和 D4D 方法。

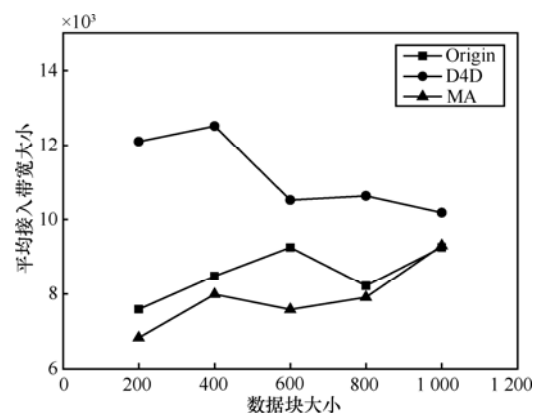
4) 往返时延评价

评价在单位数据传输时所选择的中转数据中心的接入链路的平均往返时延大小, 如图 6 所示。

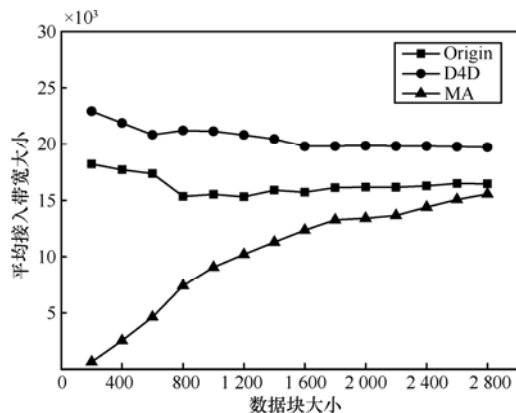
同样, 往返时延属于 QoS 的子属性, 该属性对总体评价的影响也很小。单位数据传输时基于多属性信息的方法使用到的链路平均时延要高于原始

方法和 D4D 方法。由于 D4D 方法的优化目标即在于通过最小化数据传输的时延来优先选择域内的链路, 因此在此维度上拥有最优的性能。

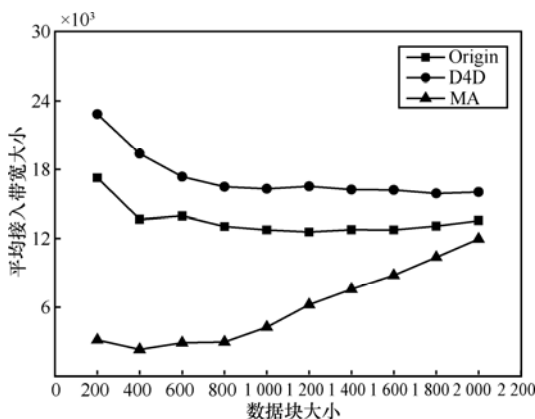
根据上述分析, 不难发现, 本文提出的基于多属性信息的数据中心间数据传输调度方法, 能够根据基于层次化分析得到的综合评价结果, 选择综合价值最优的路径进行数据的传输和调度。仿真表明, 方法会按照由层次分析法获得的权重, 根据权重高低优先选择对总体评价最优的路径, 实验结果符合方法设定。此外, D4D 方法在不同属性上也有较好的性能, 这是因为其优化目标是最小化传输链路的时延, 而对于将数据中心在本地进行存储稍后进行转发的情形, 其链路的带宽记为 INTMAX, 时延记为 0, 带宽成本也记为 0。因此优先选择时延值较低的链路也会偏向于选择成本较低、接入带宽较大的链路。使得在本实验场景下取得了较好的结果。但需要注意的是, 如果时延因素对管理者来说并不是最为看重的因素, 其最终评价会劣于本文提出的基于多属性信息的方法。



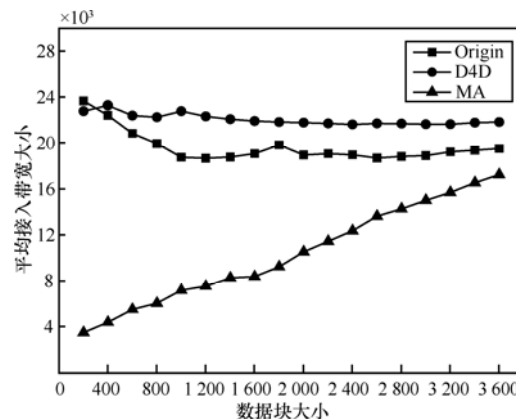
(a) $T=3$



(b) $T=4$



(c) $T=5$



(d) $T=6$

图 5 平均接入带宽大小评价

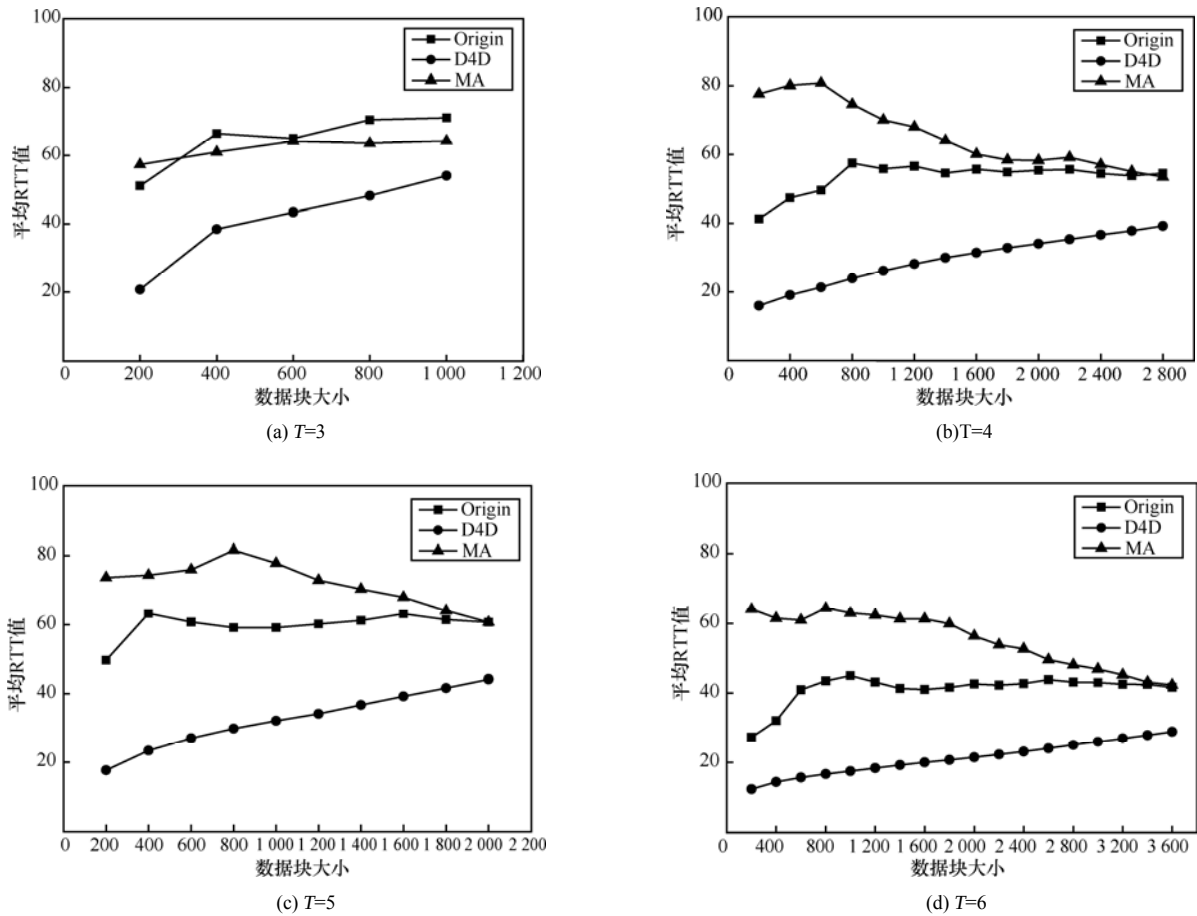


图6 平均往返时延评价

6 结束语

数据中心之间的数据传输调度是当前云计算及数据中心网络领域的热点研究问题。本文提出了一种考虑多属性信息综合评价的数据中心间数据传输调度方法。其基本思想是通过层次分析法分析不同层次属性对整体评价的影响,结合信息熵理论确定不同属性的权重,并生成最终的评价结果;最后,借助该评价结果,为数据选择一条综合评价结果最优的路径。仿真结果表明,根据获得的权重,基于多属性信息的方法在权重最高的维度上具有最高的优化性能。随着权重的降低,在其他维度的优化性能逐渐降低,直到在次要的维度上,劣于其他方法。仿真结果符合方法的预期。

本文主要在假设数据中心管理者对于多个属性没有明显偏好的情况下,研究基于多属性信息的数据中心间数据传输调度方法。但在有的场景下,管理者对于数据传输路径有特定的需求,比如由于应用的需要,需要在考虑多种因素时优先选择服务质量最好的

路径等。因此,存在属性偏好的情况下,研究多属性的数据中心间数据传输调度是下一阶段的工作。

参考文献:

- [1] Windows Azure[EB/OL]. <http://www.windowsazure.com/>, 2012.
- [2] Google App Engine[EB/OL]. <http://appengine.google.com/>, 2012.
- [3] Amazon web services[EB/OL]. <http://aws.amazon.com/>, 2012.
- [4] Akamai[EB/OL]. <http://www.akamai.com/>, 2012.
- [5] ChinaCache[EB/OL]. <http://www.chinacache.com/>, 2012.
- [6] CHEN Y Y, JAIN S, ADHIKARI V K, *et al.* A first look at inter-data center traffic characteristics via Yahoo! datasets[A]. Proceedings of IEEE INFOCOM 2011[C]. Shanghai, China, 2011, 1620-1628.
- [7] ARMBRUST M, FOX A, GRIFFITH R, *et al.* Above the Clouds: A Berkeley View of Cloud Computing[R]. EECS Department, University of California, Berkeley, 2009.
- [8] GREENBERG A, HAMILTON J, MALTZ D A, *et al.* The cost of a cloud: research problems in data center networks[J]. ACM SIGCOMM Computer Communication Review, 2008, 39(1): 68-73.
- [9] LAOUTARIS N, SIRIVIANOS M, YANG X Y, *et al.* Inter-datacenter bulk transfers with netstitcher[J]. SIGCOMM Computer Communication Review, 2011, 41(4): 74-85.
- [10] FENG Y, LI B C, LI B. Postcard: minimizing costs on inter-datacenter

traffic with store-and-forward[A]. Proceedings of the 2nd International Workshop on Data Center Performance, Conjunction with ICDCS[C]. Macau, China, 2012, 43-50

- [11] NANDAGOPAL T, PUTTASWAMY K P N. Lowering inter-datacenter bandwidth costs via bulk data scheduling[A]. Proceedings of the 12th IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing[C]. Ottawa, Canada, 2012, 244-251.
- [12] LI Y Y, WANG H B, ZHANG P, *et al.* D4D: Inter-datacenter bulk transfers with ISP friendliness[A]. Proceedings of 2012 IEEE International Conference on Cluster Computing[C]. Beijing, China, 2012, 597-600.
- [13] SAATY T L. Modeling unstructured decision problems-the theory of analytical hierarchies[J]. Mathematics and Computers in Simulation, 1978, 20(3):147-158.
- [14] Shannon C E. A mathematical theory of communication[J]. ACM SIGMOBILE Mobile Computing and Communications Review, 2001, 5(1):3-55.
- [15] The Netperf Homepage[EB/OL]. <http://www.netperf.org/netperf/>.
- [16] WebSphere sMash[EB/OL]. <http://www-01.ibm.com/software/webservers/smash/>.
- [17] FORD J L R, FULKERSON D R. Constructing maximal dynamic flows from static flows[J]. Operations Research, 1958:419-433.



王洪波（1975-），男，河北唐县人，博士，北京邮电大学副教授，主要研究方向为云计算及数据中心网络、互联网测量与管理、下一代互联网体系结构及新应用。

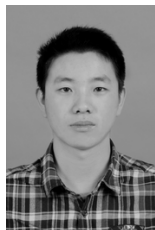


张鹏（1984-），男，河北邯郸人，北京邮电大学博士生，主要研究方向为数据中心网络。



董健康（1973-），男，甘肃陇南人，北京邮电大学博士生，主要研究方向为数据中心虚拟化技术和数据中心网络。

作者简介：



李阳阳（1987-），男，江苏扬州人，北京邮电大学博士生，主要研究方向为数据中心网络资源管理及调度。



程时端（1940-），女，上海人，北京邮电大学教授、博士生导师，主要研究方向为下一代互联网技术，互联网 QoS 技术，数据中心网络。