

天津工业大学

毕业设计（论文）

基于虚拟机放置策略的数据中心网络节能算法研究

姓 名 朱昌健
学 院 计算机科学与技术学院
专 业 计算机科学与技术
指导教师 陈 志
职 称 讲 师

2015 年 6 月 1 日

天津工业大学毕业设计（论文）任务书

题目	基于虚拟机放置策略的数据中心网络节能算法研究				
学生姓名	朱昌健	学院名称	计算机科学与软件	专业班级	计算机 1104
课题类型	实际课题				
课题意义	<p>近些年来，随着经济的发展，绿色环保问题越来越受到人们的关注。然而数据中心日益增大的能量消耗对环境造成了巨大的破坏，大量二氧化碳的排放也加剧了温室效应。当前的数据中心拥有成千上万个交换机，所以设计合理的算法可以关闭多余的交换机和服务器，并减轻了冷却装置负担，对环境的绿色发展有重大意义。</p> <p>数据中心的能耗主要来自物理机和交换机。因此本课题目标就是通过合理的虚拟机放置，将不同需求的虚拟机放置在尽可能少的物理机资源上，并使得流量路径需要流经的交换机显著减少，从而达到关闭多余的物理机和交换机实现节能减排。</p>				
任务与进度要求	<p>03.09~03.29: 准备毕业论文任务书和毕业论文开题报告。</p> <p>03.30~04.09: 了解谱聚类, Mincut, Gomory hu tree 的算法。了解 Fat-Tree 拓扑结构。设计可行的数据结构和算法。</p> <p>04.10~05.03: 调试算法中的错误，验证各类算法在不同的模型中的正确性。得到在对于不同模型下，消耗的能量总和。</p> <p>05.04~05.10: 图形化数据，通过对比和分析，找到本算法的优势和缺陷。</p> <p>05.11~05.24: 撰写论文，提交论文提纲和论文初稿。</p> <p>05.25~05.31: 进一步修改论文，完成论文定稿。提交论文定稿。</p> <p>06.01~06.07: 打印论文，制作 PPT，准备答辩。提交论文打印稿。</p>				
主要参考文献	<p>[1] J. Dong, X. Jin, H. Wang, Y. Li, P. Zhang, and S. Cheng, "Energy-Saving Virtual Machine Placement in Cloud Data Centers," in Cluster, Cloud and Grid Computing (CCGrid), 2013 13th IEEE/ACM International Symposium on, 2013, pp. 618–624.</p> <p>[2] B. Heller, S. Seetharaman, P. Mahadevan, Y. Yakoumis, P. Sharma, S. Banerjee, and N. McKeown, "ElasticTree: saving energy in data center networks," in Proceedings of the 7th USENIX conference on Networked systems design and implementation, 2010, pp. 17–17.</p>				

	<p>[3] B. Zhang, Z. Qian, W. Huang, X. Li, and S. Lu, "Minimizing Communication traffic in Data Centers with Power-aware VM Placement," in Innovative Mobile and Internet Services in Ubiquitous Computing (IMIS), 2012 Sixth International Conference on, 2012, pp. 280–285.</p> <p>[4] O. Biran, A. Corradi, M. Fanelli, L. Foschini, A. Nus, D. Raz, and E. Silvera, "A Stable Network-Aware VM Placement for Cloud Systems," in 2012 12th IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing (CCGrid), 2012, pp. 498–506.</p> <p>[5] W. Fang, X. Liang, S. Li, L. Chiaraviglio, and N. Xiong, "VMPlanner: Optimizing virtual machine placement and traffic flow routing to reduce network power costs in cloud data centers," Computer Networks, 2012.</p> <p>[6] X. Meng, V. Pappas, and L. Zhang, "Improving the scalability of data center networks with traffic-aware virtual machine placement," in INFOCOM, 2010 Proceedings IEEE, 2010, pp. 1–9.</p>
起止日期	2015 年 3 月 09 日至 2015 年 6 月 8 日
备注	

院长_____

教研室主任_____

指导教师_____

毕业设计(论文)开题报告表

2015 年 3 月 24 日

姓名	朱昌健	学院	计算机科学与软件	专业	计算机	班级	计算机 1104
题目	基于虚拟机放置策略的数据中心网络节能算法研究					指导教师	陈志

一、与本课题有关的国内外研究情况、课题研究的主要内容、目的和意义：

1、国内外研究情况：

近十年来，随着云计算的不断发展，它为用户提供了灵活，可靠和低成本的服务。这些服务的基础就是云数据中心。然而随着数据中心规模的增大，服务器的不断增加，造成了一系列的问题。

首先是传统的树形结构的数据中心已经无法满足当前云计算的需求。因此，学术界提出了虚拟数据中心进行改进。它主要有两种网络拓扑结构：以交换机为核心和以服务器为核心的两种拓扑结构。例如 Fat-Tree 就属于前者，它设置了足够多的交换机来保证每个服务器都能一对一无阻塞通讯，三层的拓扑结构也将服务器划分为了不同的集合。由于这些优点，本文也主要基于这种成熟的拓扑结构进行研究。

数据中心规模的增大，造成的另外一个严重的问题就是能耗的大量浪费。当前主流的解决方案就是设计算法，提高物理机资源的利用率来关闭多余的服务器或交换机。这些的基础都是虚拟化技术（例如将存储、网络 and 桌面虚拟化），一个物理机可以放置多个虚拟机，每个虚拟机需要消耗物理机一定的资源，比如 CPU，内存和带宽。一组虚拟机还可以放置到不同的物理机中，各个虚拟机之间需要相互交流，并维持应用的性能。虚拟化大大降低了成本，并提供了灵活的资源管理和低能耗的解决方案。

所以在虚拟化数据中心的，虚拟机和物理机不同的映射关系将会对资源的利用率造成很大的影响，如何去放置虚拟机已经变成云供应商最关注的问题之一。在当前的研究中，因为数据中心的能耗主要来自于流量的传输，所以可以利用流量和距离（两个物理机交换信息需要经过的节点个数）的乘积总和作为总能耗的衡量标准。通常将虚拟机放置问题可以抽象成组合装箱问题和二次分配问题，这是一个经典的 NP 难问题和组合优化问题。例如文献[1]中就提出了一种从虚拟机中随机选取两点，利用迭代的 min cut 对虚拟机排序，然后用层次化的 best fit 将虚拟机放置到尽可能少的物理机中。但是这种单一的 min cut 算法有一定的缺陷，会造成分割后的两个虚拟机聚类大小极不平衡。文献[6]中，Meng 提出了使用 Gomory-hu-tree 的解决方案，这种方法虽然解决了 min cut 的局部性，但忽略了节点间的大流量。本文希望提出一个新的算法，实现对虚拟机的合理聚类，并进行层次性的放置。

2、课题研究的主要内容:

本课题主要研究基于虚拟机放置策略的数据中心网络节能算法研究, 希望提供了一个虚拟机安置方案, 用来适应多种物理资源情况和网络流量从而改进资源利用, 减少物理机服务数量和网络基础, 达到节能的效果。为了实现这个目的, 本课题从图的划分算法得到启发, 因为可以将虚拟机的流量矩阵看作一个有权无向图, 这样就可以基于图论谱聚类对虚拟机进行聚类, 使得分割的子图规模差不多(避免了 min cut 的缺陷), 并使得割边权重之和最小(子图之间流量最少, 将流量大的节点聚集在一起, 从而达到节点节能的效果)。并且这种聚类算法还有另一个优势, 在放置虚拟机到物理机的时候, 可以根据不同的物理机模型实现结构化的放置。在放置虚拟机时, 还考虑了资源的限制, 对每个虚拟机模拟各自占用的 CPU 和内存, 实现二维的 Bin-packing。

3、目的和意义:

虚拟化数据中心的能耗主要来自物理机和交换机。因此本课题的目的就是通过合理的虚拟机放置, 将不同需求的虚拟机放置在尽可能少的物理机上, 使得流量路径流经的交换机显著减少, 实现节能减排。

近些年来, 随着经济的发展, 绿色环保问题越来越受到人们的关注。然而数据中心日益增大的能量消耗对环境造成了巨大的破坏, 大量二氧化碳的排放也加剧了温室效应。当先的数据中心拥有成千上万个交换机, 所以设计合理的算法可以关闭多余的交换机和服务, 并减轻了冷却装置的负担, 对环境的绿色发展有很大意义。

二、进度及预期结果:

起止日期	主要内容	预期结果
03.09~03.29:	准备毕业论文任务书和毕业论文开题报告。	完成毕业论文任务书和毕业论文开题报告。
03.30~04.09:	了解谱聚类、Min cut、Gomory hu tree 的算法。 了解 Fat-Tree 拓扑结构。设计可行的算法。	构造数据结构, 并实现各类算法。
04.10~05.03:	调试算法中的错误, 验证各类算法在不同的模型中的正确性。	得到在对于不同模型下, 消耗的能量总和。
05.04~05.10:	图形化数据, 并进行分析	通过对比和分析, 找到本算法的优势和缺陷。
05.11~05.24:	撰写论文, 提交论文提纲和论文初稿。	完成提纲和初稿

05.25~05.31:	进一步修改论文，完成论文定稿。提交论文定稿。	完成修改
06.01~06.07:	打印论文，制作 PPT，准备答辩。提交论文打印稿。	完成答辩
完成课题的现有条件	编程环境: Eclipse, MATLAB, CodeBlocks 运行环境: Linux 或 windows 文档处理: Microsoft office 2010 硬件条件: 笔记本一台	
审查意见	指导教师: _____ 年__月__日	
学院意见	主管领导: _____ 年__月__日	

天津工业大学本科毕业设计（论文）评阅表
(论文类)

[illegible]

天津工业大学毕业设计（论文）成绩考核表

学生姓名	朱昌健	学院名称	计算机科学与软件	专业班级	计算机 1104
题目	基于虚拟机放置策略的数据中心网络节能算法研究				
1. 毕业设计（论文）指导教师评语及成绩：					
<div>成绩：</div> <div>指导教师签字：_____ 年 月 日</div>					
2. 毕业设计（论文）答辩委员会评语及成绩：					
<div>成绩：</div> <div>答辩主席（或组长）签字：_____ 年 月 日</div>					
3. 毕业设计（论文）总成绩：					
a.指导教师 给定成绩	a.指导教师 给定成绩	a.指导教师 给定成绩	a.指导教师 给定成绩	a.指导教师 给定成绩	

摘 要

近些年来,随着云计算的不断发展,虽然它为用户提供了灵活,可靠和低成本的服务(例如百度的搜索服务和新浪的社交网络服务),但对高性能需求的同时也促使了数据中心规模的增大。大规模数据中心的低效除了增加运营的成本也消耗了巨大的能源,加剧了环境的破坏。本文从绿色节能出发,提出了一个新的虚拟机放置算法,用来适应多种物理资源情况和网络流量,从而改进资源的利用率,减少网络设备的负载和物理机数量,最终达到节能的效果。

虚拟机放置问题可以抽象成组合装箱问题和二次分配问题。本文以网络流量大小和流量途径交换机个数的乘积作为优化目标,此目标代表了网络负载,进而反应了能耗。为了实现节能的目标,本文实现的放置策略主要分为三个步骤:(1)从图的划分算法得到启发,基于图论谱聚类对虚拟机进行聚类。此方法有两个优势,一是使得分割的子图规模大小相同(避免 mincut 算法的缺点),并使得割边权重之和最小;另一个优势是可以对聚类结果生成一颗二叉树,根据不同的物理机模型实现结构化的放置。(2)利用此二叉树得到虚拟机序列后,将它们用最优匹配或最先匹配放置到物理机中。(3)对放置好的虚拟机进行禁忌搜索,进一步减少流量的规模。

最后通过大量的仿真实验,验证了本文提出算法的有效性。

关键字: 云数据中心; 虚拟机放置; 谱聚类; 禁忌搜索

ABSTRACT

In these years, cloud computing is developing rapidly, it provide people with flexible, reliable and low-cost services (such as Baidu search service and Sina social service), However, the demand for high performance also prompts increasing the size of the data center. Inefficient large-scale data centers not only increase the operating costs, they also consume enormous energy and cause huge damage to the environment. This paper provides a VM placement scheme meeting multiple resource constraints, such as the physical server size (CPU, memory) and network link capacity to improve resource utilization and reduce both the number of active physical servers and network elements so as to finally reduce energy consumption.

Virtual machine placement problem can be abstracted into a combined packing problem and quadratic assignment problem. In this paper, the product of network flow and the number of switch is the optimization goal that represents the network load, and then reflecting the energy consumption. In order to achieve energy efficiency goals, placement schedule implemented in this article is divided into three steps: (1) motivated from spectral clustering algorithm to cluster virtual machines based on graph theory. This method has two advantages, one is dividing graph to sub-graphs in similar size (avoiding the min cut defects), the other advantage is that we can generate a binary tree and hierarchically put VMs in physical machine. (2) We use best fit or first fit to place vms to physical machines after getting the vms order. (3) We use Tabu search to get a better result, to further reduce the size of traffic.

Keywords: cloud data center; vm placement; special clustering; Tabu search

目 录

第一章 绪论.....	1
1.1 本文目的.....	1
1.2 本文意义.....	1
1.3 相关工作.....	1
1.4 本文组织结构.....	1
第二章 相关研究工作	3
2.1 云计算数据中心.....	3
2.2 数据中心的能耗分析.....	6
2.3 国内外对数据中心节能研究.....	6
2.4 网络节能相关算法介绍.....	7
2.5 本章小结.....	11
第三章 云计算数据中心谱聚类节能算法	12
3.1 谱聚类算法介绍.....	12
3.2 问题描述和定义.....	12
3.3 基于谱聚类的虚拟机放置算法.....	15
3.5 本章小结.....	18
第四章 仿真实验与结果分析	19
4.1 实验环境设定.....	19
4.2 最优匹配和最先匹配.....	19
4.3 聚类算法对于节能影响.....	22
4.4 禁忌搜索算法仿真结果分析.....	27
4.5 不同拓扑结构图.....	28
4.6 本章小结.....	29
第五章 总结与展望	30
5.1 工作的总结.....	30
5.2 未来工作的方向.....	30
参考文献.....	33
附录.....	32

I 英文原文	32
II 中文翻译	37
致谢.....	37

第一章 绪论

1.1 本文目的

云计算数据中心的能耗主要来自物理机和交换机节点，而虚拟机和物理机对应关系决定了资源的利用率。因此本文目的就是通过合理的虚拟机放置方案，将不同需求的虚拟机放置在尽可能少的物理资源上，使得流量路径显著减少，从而达到关闭多余的物理机和交换机实现节能减排。

1.2 本文意义

近些年来，随着云计算的不断发展，虽然它为用户提供了灵活，可靠和低成本的服务（例如百度的搜索服务和新浪的社交网络服务），但对高性能需求的同时也促使了数据中心规模的增大。能源消耗很大程度增加了一个云计算中心的成本，有调查显示直接的能源消耗成本占了整个运营成本的 19%，而配电和制冷基础设施占了总成本的 23%。数据中心中的服务器的不断增加还造成了一系列的问题，巨大的能耗对环境造成了巨大的破坏，大量二氧化碳的排放也加剧了温室效应。能耗的大量浪费引起了工业界和学术界的巨大关注。

最近的研究发现和过度配置的服务器相比，数据中心网络大部分时间的利用率却很低（甚至高达 40% 的链接是未使用的和空闲的），网络容量通常远超出流量负载。

所以设计合理的虚拟机放置算法，可以非常有效地提高数据中心的可扩展性，关闭多余的交换机和服务器，并减轻冷却装置的负担，能大大减轻能量的消耗，对绿色可持续发展和对成本的节省有很大意义。

1.3 相关工作

本文主要研究基于虚拟机放置策略的数据中心网络节能算法，主要由以下三部分组成：

1. 系统地描述云计算数据中心的概念，说明了数据中心中能耗的组成部分和节能的意义。用合理的数据结构模拟虚拟机，物理机，流量和物理机的拓扑结构等，为下一步的算法做准备。
2. 实现数据中心 Gomory-hu-tree, MC-BT 和谱聚类算法，对虚拟机进行排序并放置到物理机中。
3. 根据实现的算法进行大量试验，利用试验结果分析此算法的优势和缺陷。

1.4 本文组织结构

本文的内容如下。

第一章是绪论。阐述了这篇论文的目的和意义，从现实出发，阐述了网络资源的浪费，合理的虚拟机放置对网络节能的重要性。还简单介绍了本文的主要工作和结构。

第二章是相关工作的介绍。具体阐述了云计算数据中心的概念和已有的解决方案。并介绍了两个虚拟机放置的算法。

第三章具体解释了虚拟机放置和节能的关系，并详细介绍了由谱聚类，装箱问题和禁忌搜索对虚拟机放置问题的优化。

第四章是实验结果的展示。通过对比各个指标例如流量和距离的乘积，每一层的流量总和，流量的平均跳距（average hop）等，分析本文算法的优点。

第五章是总结与展望。对现有谱聚类算法的总结和未来工作的方向。

第二章 相关研究工作

2.1 云计算数据中心

2.1.1 虚拟化的概念

虚拟化技术是云计算的基础技术。虚拟化指的是对计算资源进行抽象，将操作系统（OS），硬件平台，存储设备，或计算机网络资源进行虚拟化。

这也就是说，因为有虚拟化技术，本文的算法中，一个物理机可以放置多个虚拟机，每个虚拟机需要消耗物理机一定的资源，比如 CPU，内存和带宽。一组虚拟机还可以放置到不同的物理机中，各个虚拟机之间需要相互交流，并维持应用的性能。虚拟化降低了成本，并提供了灵活的资源管理和低能耗的解决方案。

2.1.2 物理机距离的定义

在本文中，我们将两个虚拟机之间的距离定义为他们交换流量要经过的交换机个数。由图 2-1 可以简单给出胖树（Fat-Tree）网络的物理机距离矩阵。

例如物理机 1 号和物理机 2 号在一个 rack 中，他们交换流量只需经过一个交换机，所以两个物理机的距离为 1。以此类推，物理机 1 号和 3 号的距离为 3，1 号和 5 号物理机的距离为 5。

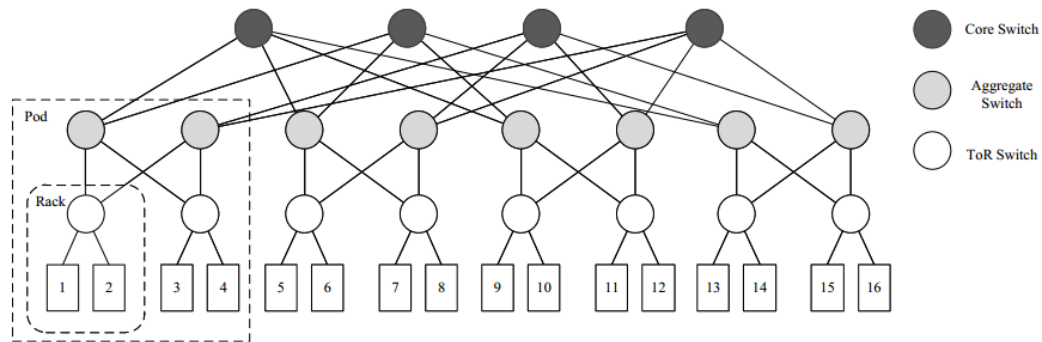


图 2-1 Fat-Tree 的拓扑图

表 2-1 主要的标记和他们的解释

符号	描述
C	两个物理机之间的距离（也代表两个物理机产生流量所需要流经的交换机）
I	物理机 I
J	物理机 J
k	一个 rack 中物理机个数（代表网络拓扑结构的规模）

Fat-tree 拓扑结构中物理机 I 和 J 的距离公式：

$$C_{IJ}^{Fat-tree} = \begin{cases} 0 & \text{if } I = J \\ 1 & \text{if } \lfloor \frac{2I}{k} \rfloor = \lfloor \frac{2J}{k} \rfloor \\ 3 & \text{if } \lfloor \frac{2I}{k} \rfloor \neq \lfloor \frac{2J}{k} \rfloor \wedge \lfloor \frac{4I}{k^2} \rfloor = \lfloor \frac{4J}{k^2} \rfloor \\ 5 & \text{if } \lfloor \frac{4I}{k^2} \rfloor \neq \lfloor \frac{4J}{k^2} \rfloor \end{cases}$$

2. 1. 3 网络拓扑结构

这一小节中呈现一些数据中心的网络模型，包括一个经典的树形结构，两种最新提出的拓扑结。它们的作用就是将巨大数目的服务器用网络设备（例如交换机）和高速链路连接起来。

当前大部分的网络模型都有一个相似的结构，被称作三层架构（three-tier architecture）。最下面的一层是接入层，每个服务器连接到一个（或者两个）访问节点。每个访问节点又连接到聚合层的一个或者两个交换机。最终，每个聚合层的节点连接到最上层核心层的多个交换机。

例如图 2-1 就是显示了胖树（fat-tree）拓扑结构的三种节点：核心层（Core Switch），聚合层（Aggregate Switch），访问层（ToR Switch）。

下文将简单介绍几种成熟的拓扑结构和它们的距离矩阵。

1. 树形结构是一种传统的网络结构。图 2-2 显示了一个 16 台服务器相连的树形网络和距离矩阵（标记为 Tree）。

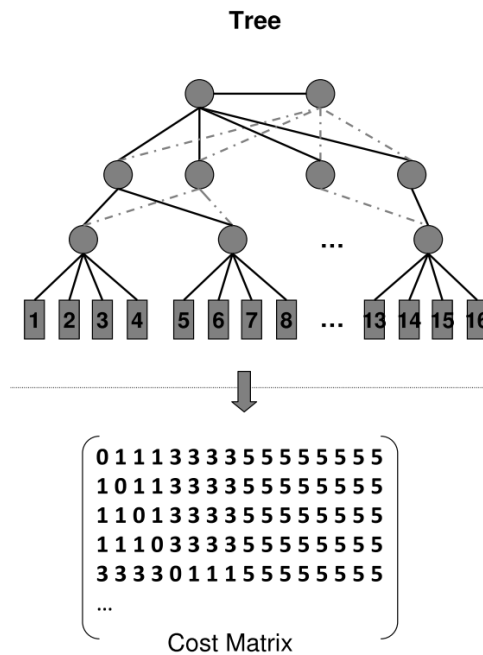


图 2-2 树形结构拓扑图和距离矩阵

2. BCube 是数据中心中一个新的多级网络架构，服务器是网络结构的一部分，某些服务器具有交换机的功能，他们能为其他服务器转发数据。

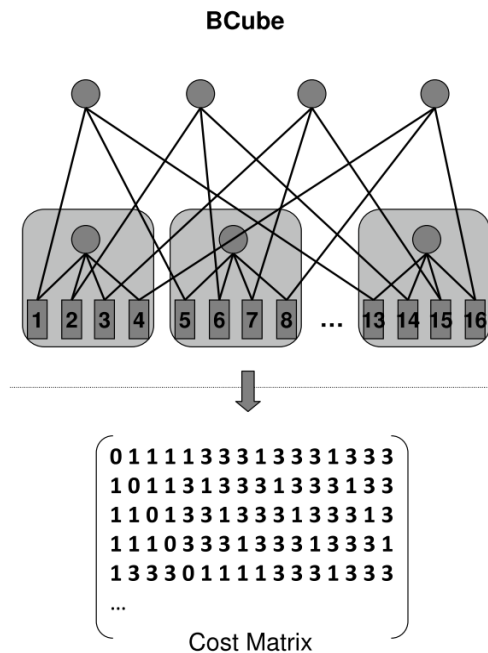


图 2-3 BCube 结构的网络拓扑图和距离矩阵

3. Fat-Tree 结构设置了足够多的交换机来保证每个服务器都能一对一无阻塞通讯，三层的拓扑结构也将服务器划分为了不同的集合。因为这些优点，本文也主要基于这种成熟的拓扑结构进行研究。所以下面详细介绍一下这种拓扑结构。

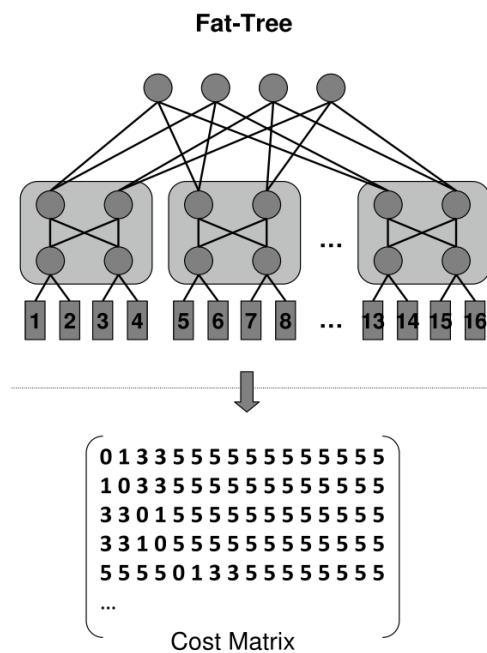


图 2-4 Fat-Tree 结构的网络拓扑图

2.2 数据中心的能耗分析

因为数据中心占全球能源使用的很大一部分，加剧了温室效应和能源危机，所以人们提出绿色数据中心，本小节通过分析能耗来寻找节能的突破口。

数据中心消耗的能量主要由两部分组成。

1. 主机资源：物理机（PMs）的集合。每个物理机由一个或多个处理器，存储器，网络接口和本地 I/O 组成，它们一起提供的云环境的计算能力。物理机通常已经部署了虚拟化软件，所以每个物理机可以放置多个虚拟机（VMs），虚拟机相互隔离，并且可以运行不同的操作系统，平台和应用程序。在之前的文献中，大多数研究人员对虚拟机和物理机建立模型，并约束他们的处理能力和内存。

2. 网络资源：数据中心中的物理机会被打包成几部分或聚类用于资源的分配。并且这些物理机必须用大带宽的网络相连，通常是千兆以太网。特别是那些通过并行计算的实现的云计算应用程序，好的网络流量分布对于性能的影响很大。

2.3 国内外对数据中心节能研究

2.3.1 总能耗的指标

在当前的研究中，最终的优化目标可以用虚拟机流量和距离（两个物理机交换信息需要经过的节点个数）的乘积表示，作为总能耗的衡量标准[6]。

2.3.2 主机节能研究

主机节能主要研究如何减少物理机（PM）的数量，将多余的虚拟机调整到睡眠状态来达到节能的效果。

文献[6]中 Meng 将这个问题模拟为装箱模型，利用最优匹配或者或最先匹配策略将虚拟机放置到尽可能少的物理机中。但大多数的研究都没有考虑虚拟机放置时，物理机实际的资源限制。本文将考虑物理机的计算负载和内存，对虚拟机进行多维的装箱。

具体算法将在下一章中介绍。

2.3.3 网络节能研究

网络节能可以通过对虚拟机聚类实现。因为应用程序通常部署在虚拟服务器上，而不是一对一直接放置到物理机上。通过优化虚拟机在主机服务器上的放置，虚拟机之间流量大小的分布可以更好地与它们之间的流量距离保持一致，也就是说，越大带宽的虚拟机应该放到越近的主机服务器上。

当前研究将虚拟机放置问题抽象为一个二次分配问题（QAP），而且是一个 NP 难问题。所以在当前超大规模的数据中心中，几乎不可能得到最优解，研究

者通常通过近似算法来找到一个近似解。

通常通过最小割(mincut)算法来解决这个问题,对虚拟机集合进行 K-mincut。但是单一的 mincut 算法有一定的缺陷,会造成分割后的两个虚拟机聚类大小极不平衡。文献[6]中, Meng 提出了使用 Gomory-hu-tree 的解决方案,这种方法虽然解决了 min cut 的局部性,但忽略了节点间的大流量。本文希望提出一个新的算法,实现对虚拟机的合理聚类,并进行层次性的放置。

基于数据中心拓扑结构的特点和流量模式的分析,文献[5]中,作者提出了一种新的方法,称为 VMPlanner,在基于虚拟化的数据中心网络中降低功耗。

VMPlanner 的基本思想就是同时优化虚拟机放置和路由流量,以便关闭许多不需要的网络设备尽可能的省电。文献[5]通过设计具有三个近似算法的 VMPlanner,逐步的优化地解决这个问题。

2.3.4 网络与主机节能相结合

网络与主机节能相结合的实现方法是先对虚拟机进行排序,例如文献[1]中就提出了 MC_BT 算法:从虚拟机中随机选取两点,利用迭代的 mincut (最小割算法)对虚拟机排序。生成虚拟机序列后,用最优匹配(best fit)算法将虚拟机放置到尽可能少的物理机中。这样的算法更加全面地解决了节能问题。

2.4 网络节能相关算法介绍

2.4.1 Gomory Hu 树算法

1. 算法介绍

Gomory hu tree 又称最小割树。也就是说求解出 Gomory-Hu tree,就可以简明地表示一个有权无向图中的两个节点间的最小割。

2. 算法的具体实现

输入: 一个有权无向图 $G = ((VG, EG), c)$.

输出: 最小割树 $T = (VT, ET)$.

这一小节是对最小割树算法的一个具体实现步骤。

图 2-5 为一个有权无向图,右侧为初始的 Gomory-hu tree,所有的点放置在一个集合之中。这时候树只有一个顶点,选择图 $X = VG = \{0, 1, 2, 3, 4, 5\}$ 。 $|X| = 6 \geq 2$,所以开始对图进行分割。

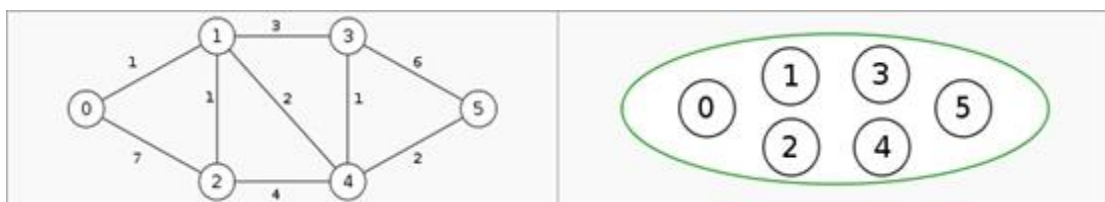


图 2-5 初始图

步骤一：选择 $s = 1$, $t = 5$, 所以 $\text{mincut } s-t(A', B')$ 就是 $(\{0, 1, 2, 4\}, \{3, 5\})$, mincut 的值为 6。也就是说将原图用 mincut 分割为两个部分了，但这两个集合中的结点个数大于等于 2，接下来步骤就是迭代对子图进行步骤一的处理。

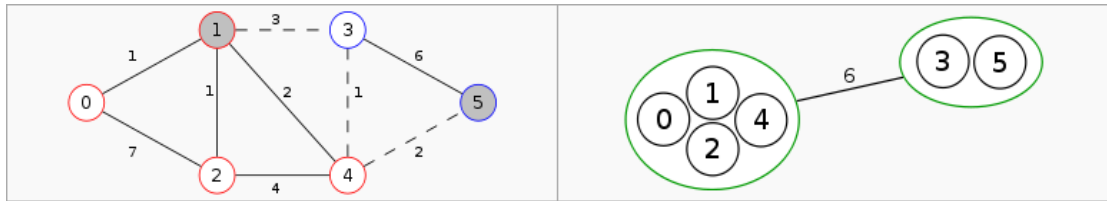


图 2-6 步骤一

步骤二：将边替换为切割的总权重

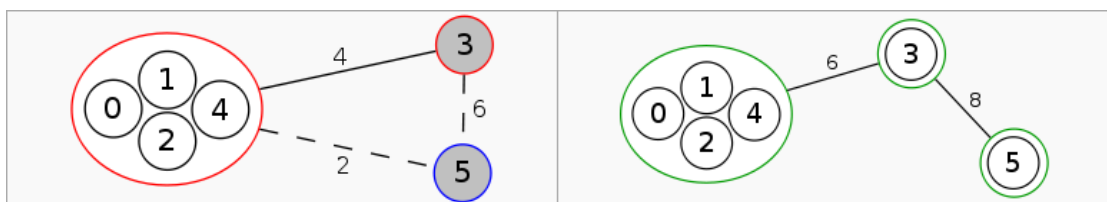


图 2-7 步骤二

步骤三：迭代步骤一的操作。

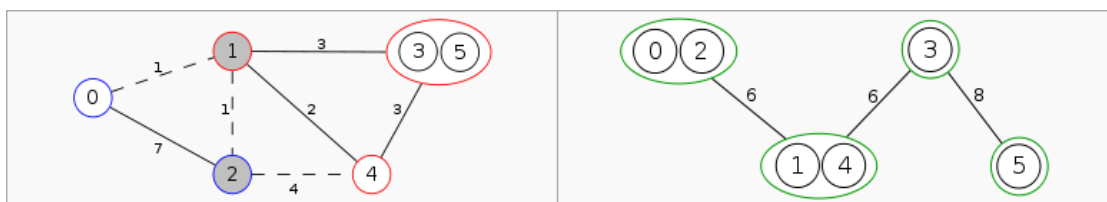


图 2-8 步骤三

最终每个集合 X 中的节点数都小于等于 2，就完成将原无向图划分为 Gomory-Hu tree，如图 2-9 所示：

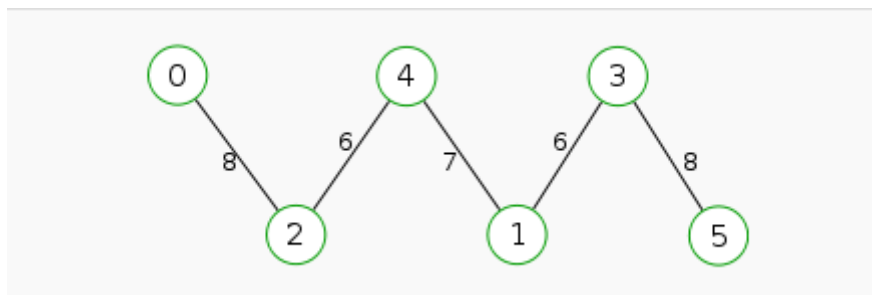


图 2-9 步骤四

模拟实验中对于随机生成的流量图进行 Gomory-hu tree 的生成。

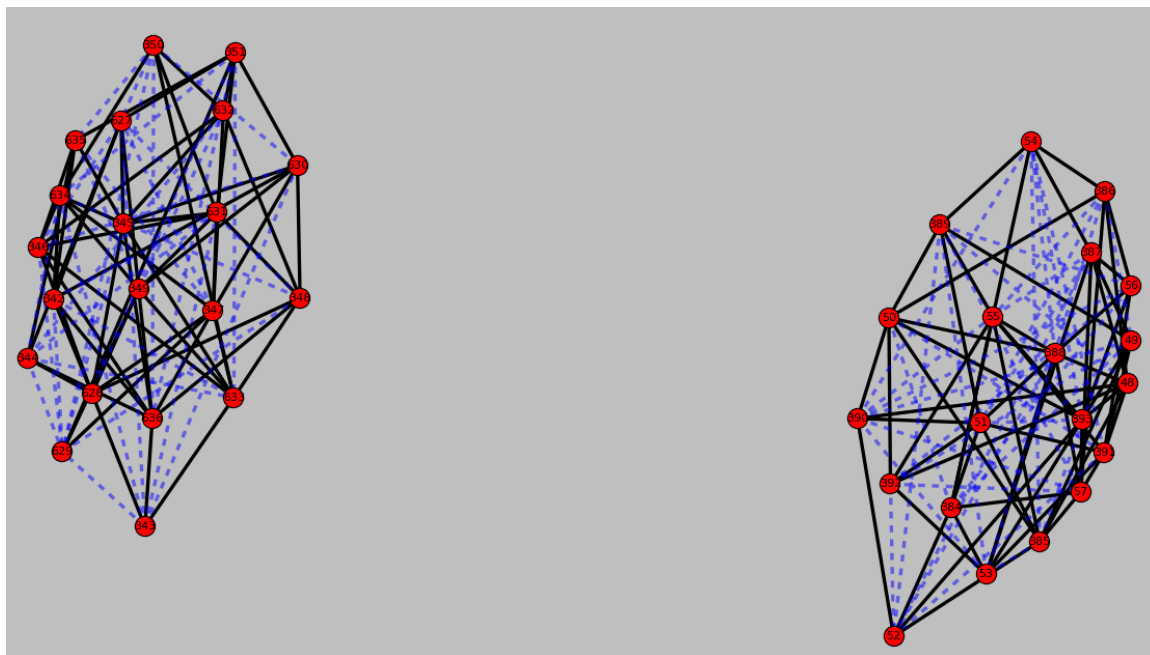


图 2-10 随机生成的流量图

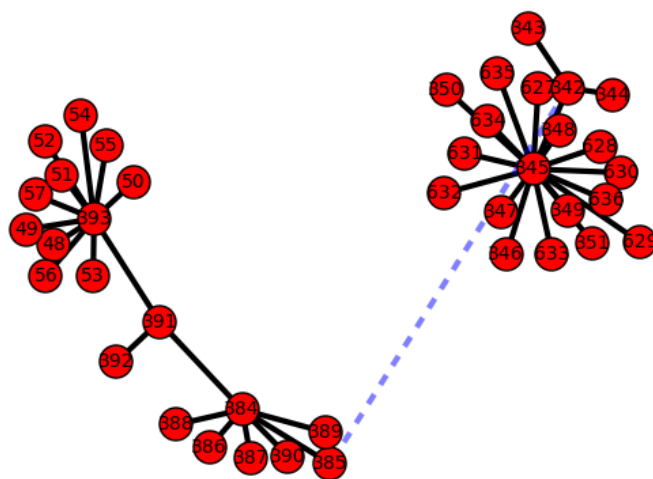


图 2-11 Gomory-hu tree 算法处理过的树

3.算法的改进

这个算法实际上是有一定缺陷，因为如果只根据最小割来对虚拟机排序，两

个虚拟机之间流量非常大的情况就可能被忽略，无法放在相近的位置，从而造成对网络资源的浪费。所以本文利用此算法对虚拟机排序时，在 Gomory-hu tree 的基础上做了改进。

改进的主要思想就是不再单单以 Gomory-hu tree 的权重作为虚拟机排序的标准，而是以原流量图每个节点的权重和树的权重的比值作为虚拟机排序的标准，这样就能体现出两个点之间的流量占 mincut 的比值，如果这个比值越大，说明这两个虚拟机应该放到越近的位置。

2.4.2 MC-BT 算法

1. 算法介绍

此算法是文献[1]提出的，实际上是利用最小割（mincut）的另一种改进，和上一节的 Gomory-hu tree 算法是异曲同工，但最后生成的是一棵二叉树。

2. 算法实现

此算法将一个有权无向图定义为 $G=(V, E)$ ， V 代表的是虚拟机的集合， E 代表各个虚拟机之间的流量。

首先随机选取两点 s 和 t ，对 VM 之间流量组成的图 G ，递归进行 s - t min cut，每次分两个子图，分别计算内部结点直接的流量之和，流量大的子图作为左子树，流量小的作为右子树，最终将所以结点形成一颗树 T 。

表 2-1 MC-BT 算法

算法 1 MC-BT 具体算法

Input: Graph $G=(V, E)$

Output: Binary tree $T(V)$

While $|G| \geq 2$ do

 Pick two nodes s and t randomly

 Compute s - t mincut on G

 if $W(G_s) > W(G_t)$

 Left subtree $\leftarrow G_s(V)$

 Right subtree $\leftarrow G_t(V)$

 Endif

 Replace G by G_s and G_t

Endwhile

Output $T(V)$

3. 算法缺陷分析

但是由于 mincut 本身的缺陷，在具体实验中发现，对于一些流量特别复杂的

虚拟机矩阵，会出现每次 mincut 只能割下一个节点的情况（如图 2-12 所示），最终的虚拟机排序就无法体现虚拟机两两之间的流量关系。

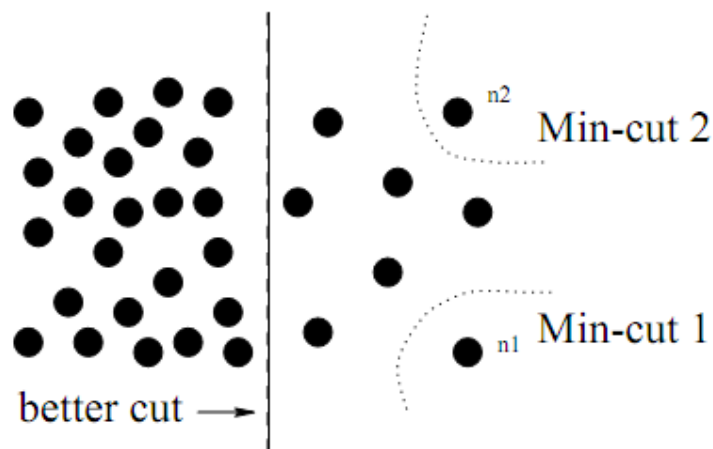


图 2-12 mincut 缺陷解释

2.5 本章小结

这一章主要介绍了云计算数据中心的概念和三种主流的拓扑图，并对数据中心的能耗进行了分析，由此提出了通过主机节能和网络节能的两种节能方法，并基于网络节能介绍了两种虚拟机排序的算法，两种都是基于 mincut 改进的算法。

为了避免最小割算法的一些缺陷，本文在下一章将尝试使用图论的谱聚类算法，对虚拟机矩阵进行聚类并排序，用来反应虚拟机间的流量关系。

第三章 云计算数据中心谱聚类节能算法

3.1 谱聚类算法介绍

3.1.1 传统聚类算法的缺点

近些年来，在机器学习和模式识别研究中，寻找好的聚类一直是研究者关注的焦点。传统的算法都有一定的缺陷，首先之前的算法需要预先输入一些密度参数，并做一些粗略的假设，例如 K-means 算法就有这些类似的缺陷。还有例如最小割算法在分割图形时，经常出现一直对单点进行分割的情况。

3.1.2 谱聚类算法与网络节能的关系

谱聚类（Spectral Clustering）和图划分密切相关，是一种基于图论的聚类方法，它利用矩阵的特征值对无向加权图进行聚类。此算法实现的效果就是将一个无向图平衡的分为两个大小相似的子图，并且使得割边的权重之和最小，如图 3-1 所示：

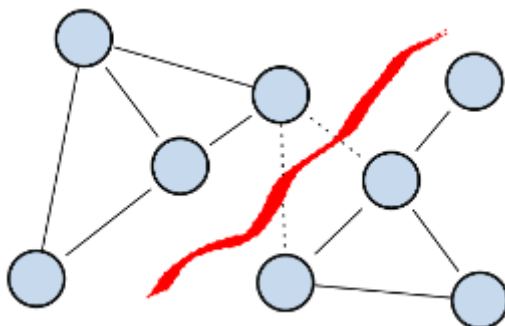


图 3-1 谱聚类示例

在云计算数据中心的，所有虚拟机之间的流量形成了一个流量矩阵，实际上也是一个有权无向图，受到图论算法的启发，本文将使用谱聚类算法对虚拟机进行聚类，此时两个子图之间的流量和最小，将流量大的节点聚合在一起。

经过迭代的谱聚类处理，所有的虚拟机被分割为单点，并生成考虑流量优化的虚拟机(VMs)列表，这时就转为多维装箱问题模型，再用最优匹配或最先匹配策略放置到物理机中。

经过这样的处理，流量大的虚拟机会聚合到一起，放在距离近的物理机或同一个物理机中，关闭一些网络设备最终达到节能的目的。

3.2 问题描述和定义

1. 问题描述

在 IaaS（云计算基础设施即服务）中，云服务提供商将资源租给有需求的用户，并且保证服务的性能。对于用户来说最关键的是服务的质量和应用的性能。而对于云服务提供商来说，最重要的是最大限度地提高资源的利用率并减少能耗，提升利润。所以云服务提供商最需要考虑的是设计合理的虚拟机放置策略，用来提高物理机的资源利用率，关闭多余的物理机和网络元件，这样就能减少硬件的投资和电能的消耗，最终达到减少数据中心开支的目的。

在云计算数据中心中，通常是用合理的虚拟机放置策略，尽可能减少交换机上的流量来减少能耗。这个问题可以抽象为二次分配问题(QAP)：给定 n 个虚拟机和 n 个物理机，每个物理机分配一个（或多个）虚拟机。很显然有 $n!$ 种不同的分配组合。每个虚拟机和它放置的物理机之间的代价函数被定义为二者间距离和流量的乘积，这个目标值越小代表网络中的总通讯量越小，也就意味着网络设施（交换机，链路等）的数量将会减少，一些空闲出来的网络设备就能进入睡眠状态，功率消耗最终将被降低。本文将使用这个目标值（距离和流量的乘积）作为衡量能耗的标准。

例如图 3-2 是通过虚拟机放置策略，对网络资源进行优化的一个实际情况。当 2 号虚拟机和 4 号虚拟机的位置交换时，就可以关闭一个多余的交换机，实现节能的效果。

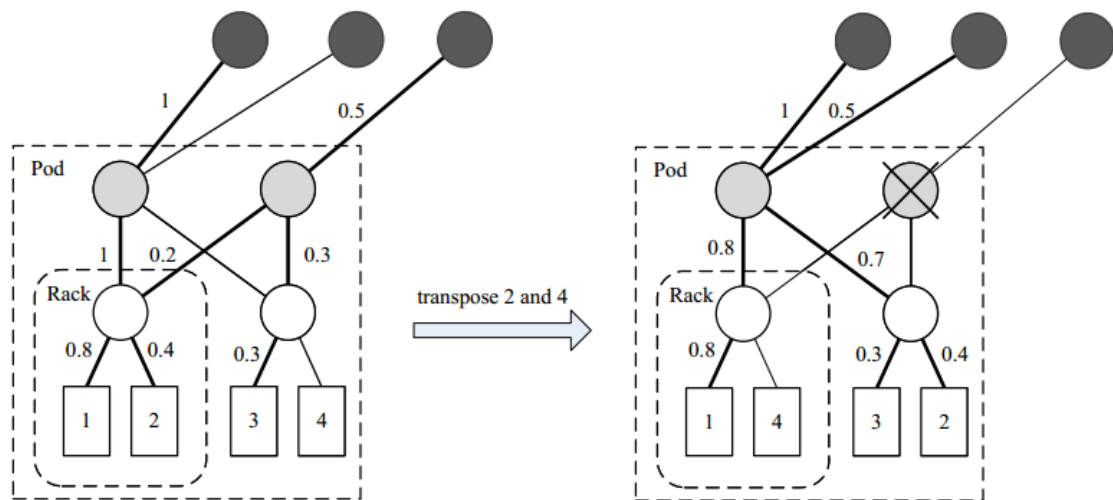


图 3-2 放置策略和节能的关系

问题定义。

表 3-1 各符号的表示

符号	描述
i, j	虚拟机 i 和 j
m, p	物理机 m 和 p
M	物理机的数量，从 1 增加到 M

N	虚拟机的数量，从 1 增加到 N
H_m	物理机 m 的能提供的资源
S_i	虚拟机 i 的消耗的资源
Y_m	物理机的状态的二维变量，1 表示物理机处于活动的状态；0 表示物理机处于睡眠的状态
$X_{i,m}$	虚拟机 i 是否放在物理机 m 上的二维变量，1 表示是，2 表示否
A	虚拟机的流量矩阵 $A=(a_{i,j})_{n \times n}$ ，表示虚拟机 i 和 j 之间的流量大小。
B	物理机的距离矩阵 $B=(b_{m,p})_{m \times m}$ ，代表从物理机 m 发送流量到物理机 p 需要经过的交换机个数。

本文的优化目标为： $\min Cost_{net} = \sum_{i,j=1}^M a_{i,j} b_{m,p}$

这个公式代表两两流量和距离的乘积，用它来代表总能耗。

约束条件为： $s, t. \sum_{i \in \pi_t} S_i^{(resource)} < H_t^{(resource)}$

这个公式代表每个物理机内所有虚拟机占用的资源之和必须小于物理机的最高负载，作为优化目标的约束条件。

2. 服务器资源的优化

对服务器资源的优化可以抽象为多维装箱问题，利用最优匹配或者或最先匹配策略将虚拟机放置到尽可能少的物理机中。目标是将不同规模的虚拟机放置到尽可能少的物理机，关闭多余的物理机大大降低能耗。

定义 $X_{i,m} = \begin{cases} 1 \\ 0 \end{cases}$ 1 表示虚拟机 i 放置在了物理机 m 上，反之为 0

F 表示所有的虚拟机： $F = \{X_{i,m} | X_{i,m} \in \{0,1\}, \sum_{m=1}^M X_{i,m} = 1, \forall i\}$

虚拟机放置问题的正式化表示如下：

$$\min Cost = \sum_{i=1}^N X_{i,m} \cdot \bar{S}_i \leq Y_m \cdot \bar{H}_m, \forall m$$

限制条件是： $\{X_{i,m}\} \in F$

二进制 $Y_m \in \{0,1\}$ 表示的是物理机 m 是在运行还是处于关闭状态将要被启动。限制条件就是要放置的虚拟机的资源必须要比虚拟机放置在的物理机上的资源数少。

上一小节由流量大小和距离推算出的总能耗目标公式，其实代表了总通信量的规模，从侧面反映了物理机的总数，所以本文暂时不考虑物理机的数量作为能耗的标准。

3.3 基于谱聚类的虚拟机放置算法

3.3.1 基于谱聚类对虚拟机进行聚类

1. 谱聚类最优切割方法

一个有权无向图定义：虚拟机集合 V ，流量矩阵 W

A 和 B 两点间的最小割可以定义为：

$$\text{cut}(A, B) = \sum_{i \in A, j \in B} \omega_{ij}$$

因为上一章提到 min cut 的一些缺点，于是有人对 mincut 做了改进，提出了 RatioCut，利用割边权重的总和和节点的权重的比值作为分割的标准：

$$\text{RatioCut}(A, B) = \frac{\text{cut}(A, \bar{A})}{|A|} + \frac{\text{cut}(A, \bar{A})}{|\bar{A}|}$$

还有 NCut(NormalizedCut)：

$$\text{NCut}(A, B) = \frac{\text{cut}(A, \bar{A})}{\text{vol}(A)} + \frac{\text{cut}(A, \bar{A})}{\text{vol}(\bar{A})}$$

RCC(ratio Cheeger cut)：

$$\text{RCC}(C, \bar{C}) = \frac{\text{cut}(C, \bar{C})}{\min\{|C|, |\bar{C}|\}}$$

2. 谱聚类的一般步骤。

- 1) 准备图的矩阵 $G(V, E)$ ；
- 2) 产生普拉斯矩阵；
- 3) 最后产生特征值和特征向量；
- 4) 对特征向量进行 kmeans 聚类

3. 谱聚类程序的输入输出。

1) 输入：

W ：稀疏的权重矩阵，必须是对称的

criterion：多种割图的标准，提供的方法有：

'ncut' - Normalized Cut

'ncc' - Normalized Cheeger Cut

'rcut' - Ratio Cut

'rcc' - Ratio Cheeger Cut

k: 最后聚类的个数

2) 输出:

clusters: 输出结果每一列中相同的数字的虚拟机表示在同一聚类中。

cuts: 输出结果每一列中数字与前一列不同的虚拟机表示相对上一次最后由输出的程序可以得到一颗关于每次分割过程的二叉树 $T(V)$ 。

3.3.2 最优匹配和最先匹配放置策略

基于上面的谱聚类算法我们得到了 $T(V)$ ，对这棵树进行前序排列就能得到虚拟机的顺序 (VMlist)，它是由所有 T 树的叶子节点组成的。因为上个谱聚类的算法，VMlist 中相邻近的虚拟机一般有比较大的流量，而距离比较远的虚拟机的流量相对比较小。

在这一小节，我们将 VMlist 中不同大小的虚拟机节点（每个虚拟机的 cpu 和内存需求不同）通过 best-fit 或 first-fit 放置到对应的物理机中。

Bset-Fit (最优匹配): 对于每一个虚拟机，我们都从第一个物理机开始搜索是否能放置。只有当所有活动的物理机都无法放置时，才会将这个虚拟机放置到一个新的物理机中。BF 的时间复杂度是 $O(n^2)$ ，空间复杂度是 $O(n)$ 。

算法 1 Bset-Fit 算法

Input: 物理机资源 PMlist，二叉树 $T(V)$

Output: VM 和 PM 之间的对应关系 VM_to_PM

initialize 二叉树 $T(V)$ 进行前序遍历得到虚拟机的序列

place_VM_in_PM(VMlist[0], 0)

for VM in VMlist:

if(VMlist.index(VM) >= 1):

for PM in PMlist:

if(PM_cpu_left >= 0 and PM_mem_left >= 0

and PM_sum_left < PM_best_capacity_left):

PM_best_capacity_left = PM_sum_left

VM_best_location = PMlist.index(PM)

place_VM_in_PM(VM, VM_best_location)

Output VM_to_PM

First-Fit(最先匹配): 从算法的名字可以看出，first-fit 和 best-fit 的区别在于放置虚拟机时，如果一个物理机无法放置，之后的虚拟机就不再考虑放置到这个物

理机了。这样的优点在于运行的速度更快，保留了虚拟机的顺序，但是对物理机资源的利用率没有 Best-fit 高。

算法 2 First-Fit 算法

Input: 物理机资源 PMlist, 二叉树 T (V)

Output: VM 和 PM 之间的对应关系 VM_to_PM

initialize 二叉树 T (V) 进行前序遍历得到虚拟机的序列

place_VM_in_PM(VMlist[0], 0)

pm_flag = 0

for VM in VMlist:

if(VMlist.index(VM) >= 1):

if(PM_cpu_left >= 0 and PM_mem_left >= 0):

place_VM_in_PM(VM, pm_flag)

else:

pm_flag += 1

place_VM_in_PM(VM, pm_flag)

Output VM_to_PM

3.3.3 基于禁忌搜索的放置优化

上面两步实现的其实就是初始放置，下面将利用禁忌搜索对虚拟机进行优化。

1.背景介绍

禁忌搜索 (Tabu search) 是一种启发式的搜索算法，可以用来解决优化组合问题 (例如获得最佳的排序)。总的思路是在循环中避免搜索已经搜索过的步骤，具体的途径是通过在下一次迭代中，不再访问之前的解空间 (因此叫做“禁忌”)。禁忌搜索从人类的行为受到启发，和人类对随机事物的操作行为一致。这保证了算法对于一个每次都会从一个新区域开始搜索以避免局部最小值，最终找到所需的解决方案。

2.算法基本实现

(1) 定义禁忌表，并使它初始值为空，利用算法随机产生初始解。

(2) 判断是否达到算法的终止条件？如果达到了，就结束算法，输出优化后的结果；如果不是就继续下面的步骤。

(3) 从当前解中的邻域函数，产生邻域解，并从中找出一些候选解。

(4) 判断对数据中心的总能耗是否有优化？如果总能耗小于之前的能耗，就对虚拟机的位置进行调整。

（5）跳转到步骤（2）。

3.5 本章小结

本章主要介绍了虚拟机放置问题对云服务供应商的重要性。以网络节能为目标，物理机能力为约束条件，将谱聚类算法与虚拟机放置问题相联系。解决了之前算法的一些不足之处。将虚拟机放置问题抽象为一个二次分配问题，从对总能耗进行优化。

第四章 仿真实验与结果分析

4.1 实验环境设定

4.1.1 实验程序

本文使用 Matlab 实现谱聚类算法，利用 python 模拟 MC-BT 和放置算法，用 c 语言模拟 Tabu 搜索的算法进行优化。最后用 python 编写脚本将各个步骤整合在一起进行仿真实验。

4.1.2 实验参数设定

本实验包括三个最基本的输入：虚拟机的流量矩阵（ 1024×1024 ），物理机的距离矩阵和每个虚拟机占用的资源。

1. 虚拟机的流量矩阵（ 1024×1024 ）：虚拟机数量为 1024，初始时两两虚拟机之间的流量都是 0。利用算法在 1024×1024 矩阵中随机生成局部流量。所以生成的“(m)partitions@(n)percent.data”的流量文件代表的就是这个流量矩阵分为 m 个不连通子图，每个部分的虚拟机个数占总虚拟机的 n%。下文将用[m, n]来表示流量矩阵的类型。

2. 物理机的距离矩阵：上文 2.1.2 介绍过，矩阵中的值代表的就是流量从物理机 I 到物理机 J 所要经过的交换机个数。本实验测试了传统的树形拓扑结构和 Fat-Tree 拓扑结构。

3. 每个虚拟机占用的资源：本文利用程序模拟了每个有流量虚拟机各自占用的 cpu 和内存。“Node1024_cpu(a)_men(b)_stdvar(c)”就是随机生成的 1024 个虚拟机各自消耗的资源，cpu 和内存的平均值为 0.5，标准差为 1。下文将用[a, b, c]来表示虚拟机资源的类型。

4.2 最优匹配和最先匹配

4.2.1 物理机资源利用率

这个实验分析了[1, 5%], [1, 15%], [1, 25%]和[1, 35%]的流量矩阵对物理机资源利用情况。模拟虚拟机的资源都为[0.2, 0.2, 1]。因为本实验只观察最先匹配和最优匹配的特点，所以虚拟机排列的顺序为随机。

FF: 最先匹配（First Fit），BF: 最优匹配（Best Fit）

纵坐标代表物理机资源利用情况，定义每个物理机的资源上限为 1。

横坐标代表物理机的序号

1. [1, 5%]的流量矩阵前 18 个物理机资源利用情况。

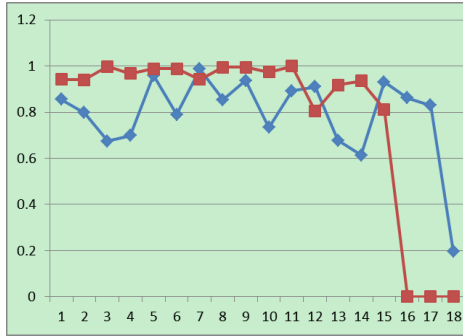


图4-1 [1,5%]的cpu匹配

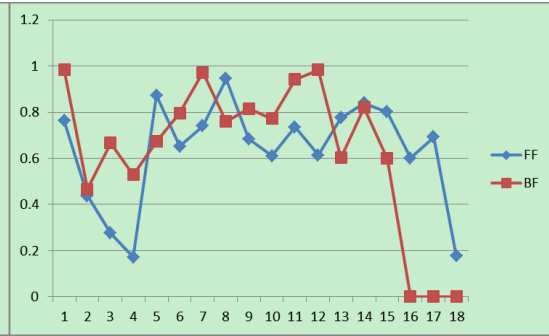


图4-2 [1,5%]的mem匹配

2. [1, 15%]的流量矩阵前 8 个物理机资源情况:

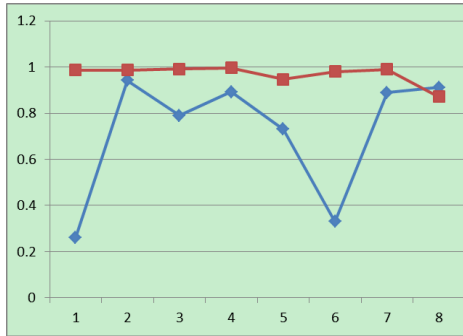


图4-3 [1,15%]的cpu匹配

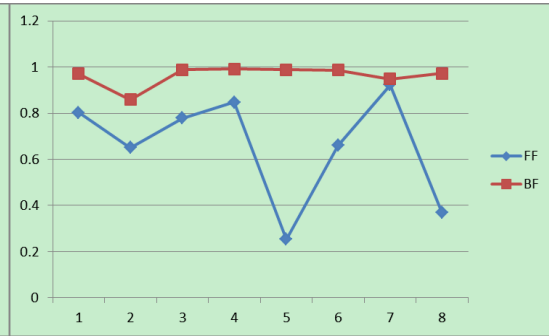


图4-4 [1,15%]的mem匹配

3. [1, 25%]的流量矩阵前 20 个物理机资源情况:

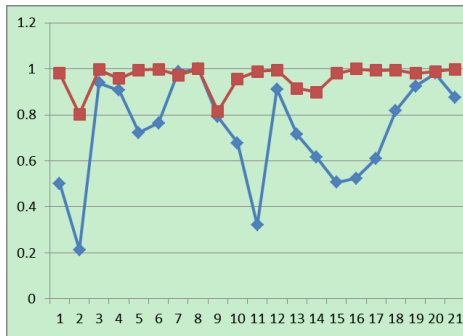


图4-5 [1,25%]的cpu匹配

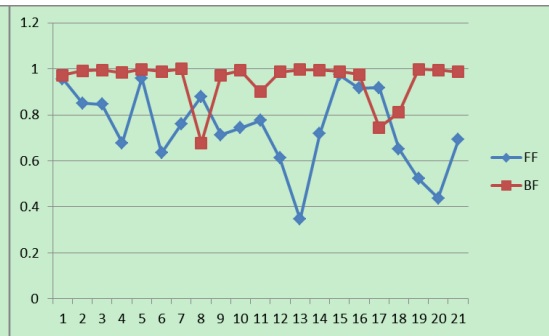


图4-6 [1,25%]的mem匹配

4. [1, 35%]的流量矩阵前 20 个物理机资源情况:

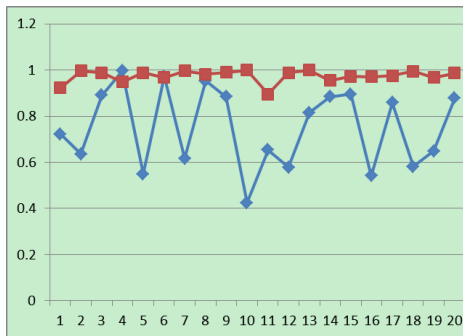


图4-7 [1,35%]的cpu匹配

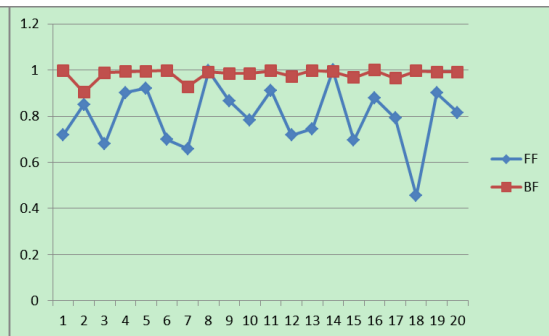


图4-8 [1,35%]的mem匹配

从各个流量矩阵的BF和FF可以看出：用BF放置虚拟机的时候，明显比FF占用更多的资源。例如图4-1和图4-2中，大部分BF后的物理机资源都接近1，而且大于FF，所以用BF进行放置相比于FF，节约了三个物理机的资源。

4.2.2 物理机开启数量

这个实验对[1, 15%]，[2, 15%]，[3, 15%]，[4, 15%]和[1, 25%]的流量矩阵对物理机数量进行分析（虚拟机的模拟资源都为[0.5, 0.5, 1]）。

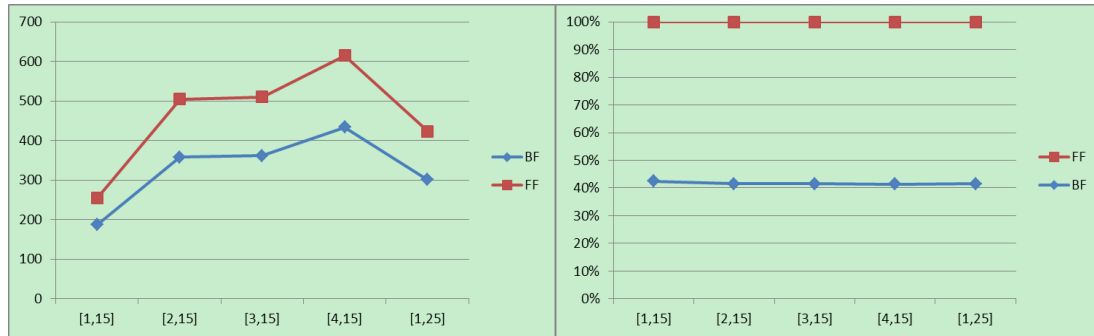


图 4-9 物理机开启的数量

从图 4-9 可以看出，和 FF 相比，BF 算法大大减少了物理机开启的数量。而且 BF 相对于 FF 算法的对物理机优化非常稳定，保持在 41% 左右。

4.2.3 平均跳步数

最先匹配和最优匹配对平均跳步数的影响：

纵坐标代表平均跳步数。

横坐标代表不同的虚拟机流量图

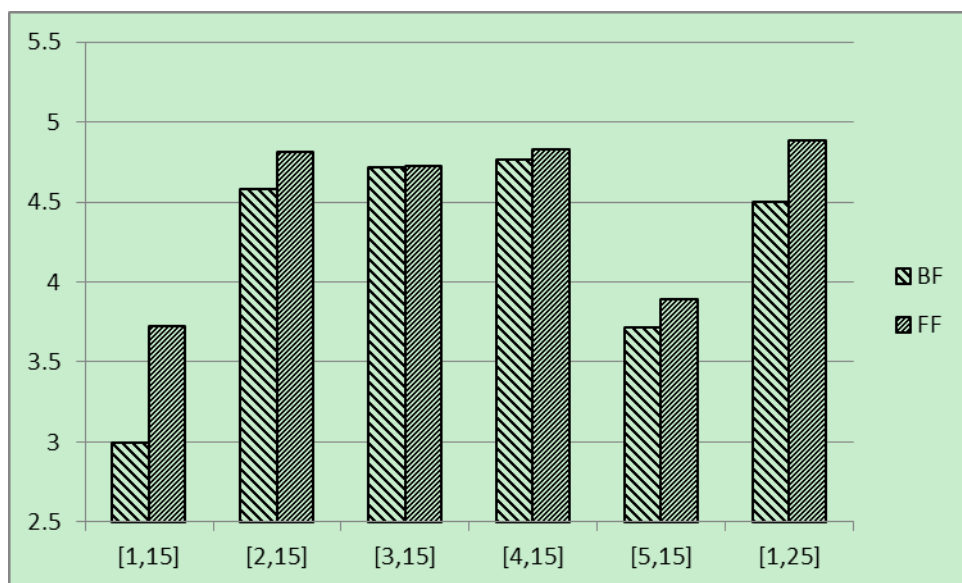


图 4-10 平均跳步数

4.2.4 结论

从物理机上 cpu 和内存的资源占用情况，平均跳步数和物理机开启数量可以很明显的看出，最优匹配（Best Fit）在放置虚拟机时能够更好的利用物理机上的资源，将虚拟机排列的更紧密。

所以下一章各个算法对虚拟机排序后，将都默认采用最优匹配进行放置。

4.3 聚类算法对于节能影响

4.3.1 优化目标的结果分析

图 4-11 到图 4-14 是各个算法对四个流量图进行处理的结果，模拟虚拟机消耗的资源为[0.5, 0.5, 1]。纵坐标表示对虚拟机进行排序后，最先匹配放置到物理机后的总能耗值。横坐标表示各个算法的缩写。

表 4-1 各个算法的简称和含义

简称	算法的含义
ran	vm 按序号排列，子图分开，代表虚拟机随机排序
mc_bt	MC_BT 算法对虚拟机进行排序
node	节点权重对虚拟机进行排序
hu	Gomory Hu Tree 对虚拟机进行排序
ncut	Normalized Cut 对虚拟机进行排序
ncc	Normalized Cheeger Cut 对虚拟机进行排序
rcut	Ratio Cut 对虚拟机进行排序
rcc	Ratio Cheeger Cut 对虚拟机进行排序

1) 对[1Partitions, 5percent]的流量图节能目标值的优化。

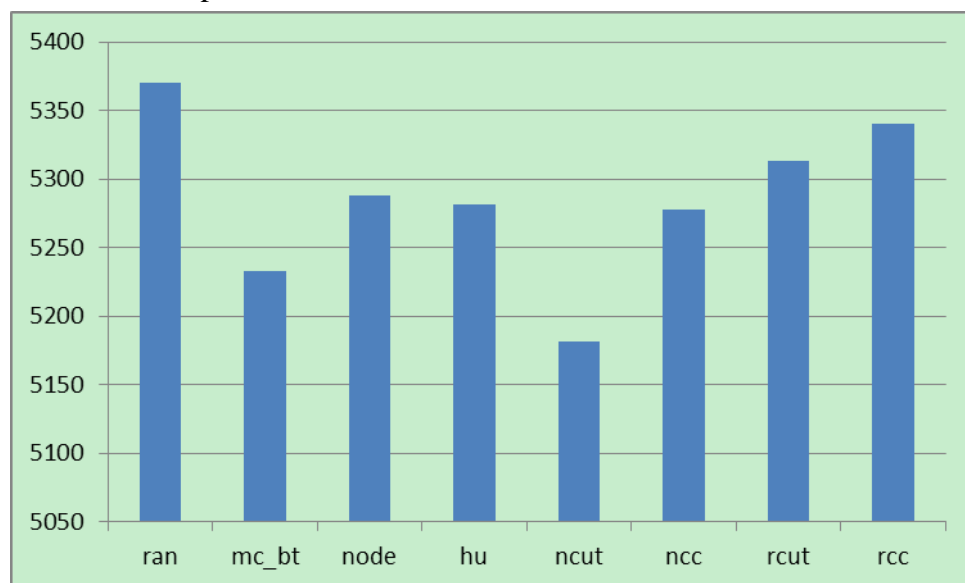


图 4-11 各算法对节能目标值的优化

2) 对[1Partitions, 15percent]的流量图节能目标值的优化

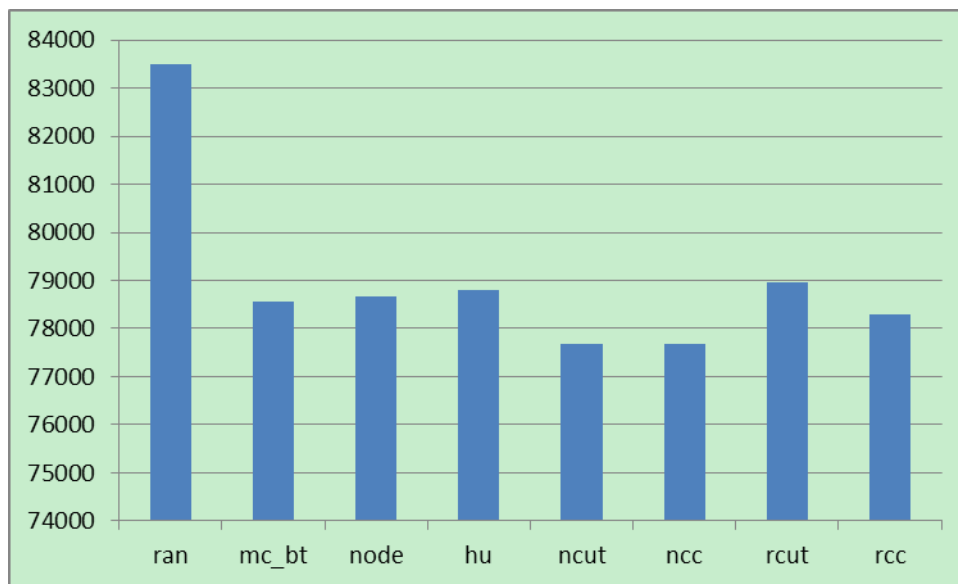


图 4-12 各算法对节能目标值的优化

3) 对[2Partitions, 15percent]的流量图节能目标值的优化

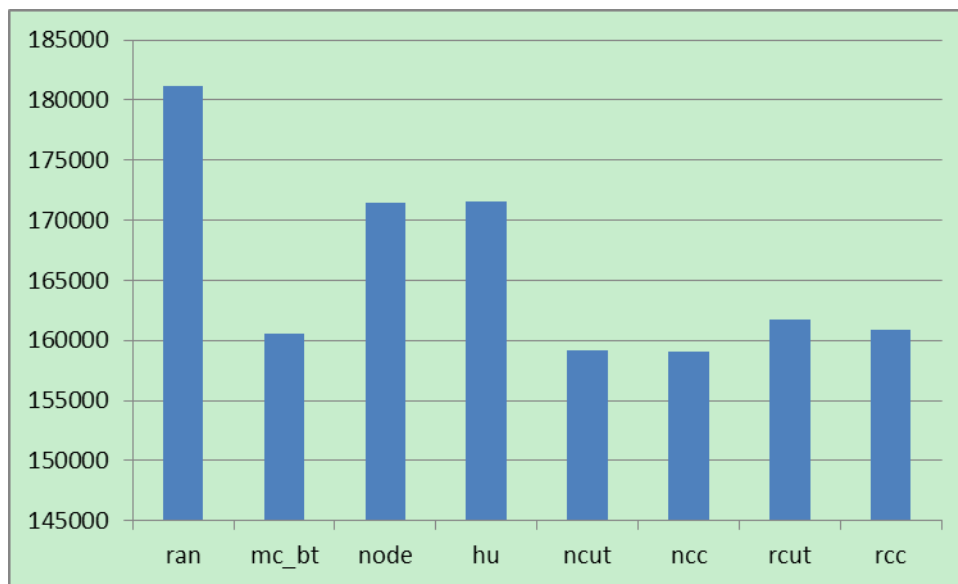


图 4-13 各算法节能目标值的优化

4) 对[5Partitions, 5percent]的流量图节能目标值的优化

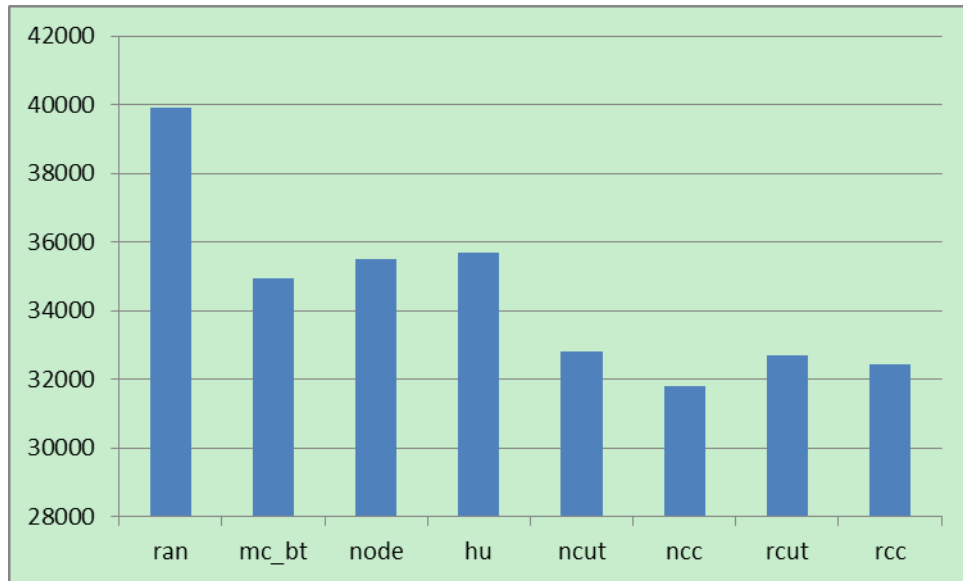


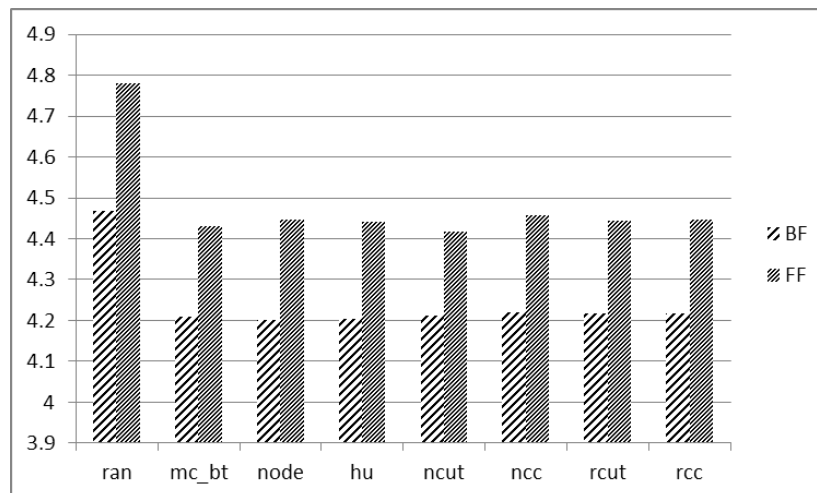
图 4-14 各算法节能目标值的优化

从图 4-11 到图 4-14 不难看出各个算法生成的虚拟机排列顺序，对目标值的影响还是很大的。虚拟机的排序为随机时，需要消耗的网络资源最大。而谱聚类算法大大优化了网络功耗。例如图 4-11 到图 4-14 中，ncc 谱聚类算法得到的网络能耗值就比随机的虚拟机序列分别优化了 3.6%，7%，12.2%，20.3%。

MC_BT 算法和 Gomory hu tree 算法也取得了不错的结果，例如图 4-13 中 MC_BT 算法和谱聚类优化效果近似。但谱聚类算法对网络流量的优化是普遍最大的。并且从优化的百分比看出，流量矩阵中的流量子图越多时，效果一般越好。

4.3.2 对平均跳步数（Average Hops per Flow）的分析

这一小节将对拓扑结构中平均跳步数的进行分析，平均跳步数的含义就是所有流量经过的交换机个数的平均值，这个指标也同时反应了总能耗。这个值越小就说明大的流量都汇集到一起，例如平均跳步数为零的话，所有的流量的有虚拟机就放在同个物理机中。



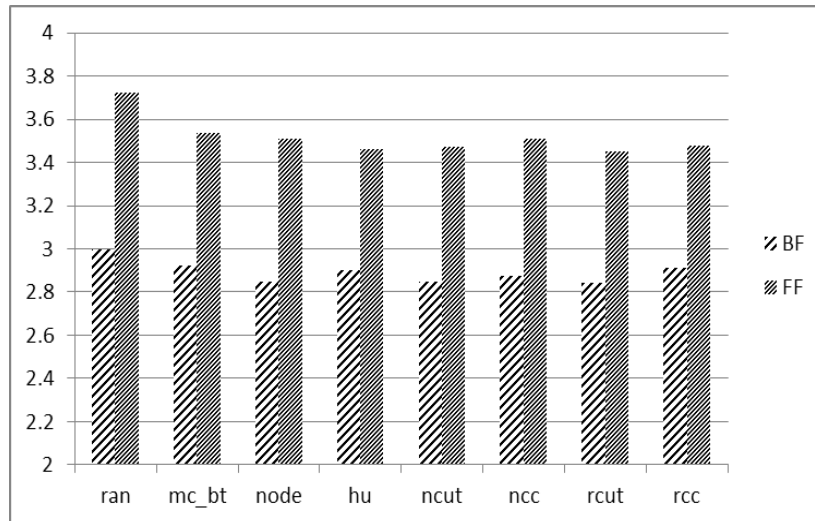


图 4-15 各算法的优化平均跳步数

从图 4-15 可以看出，各个算法对于平均跳步数有一定优化，但跳步数主要受放置算法的影响。

4.3.3 不同算法对于放置不同虚拟机资源的结果分析

这一小节将分析虚拟机模拟资源对各算法的影响，模拟的虚拟机资源分别为 $[0.5, 0.3, 1]$ ， $[0.5, 0.5, 1]$ ， $[0.2, 0.2, 1]$ 。

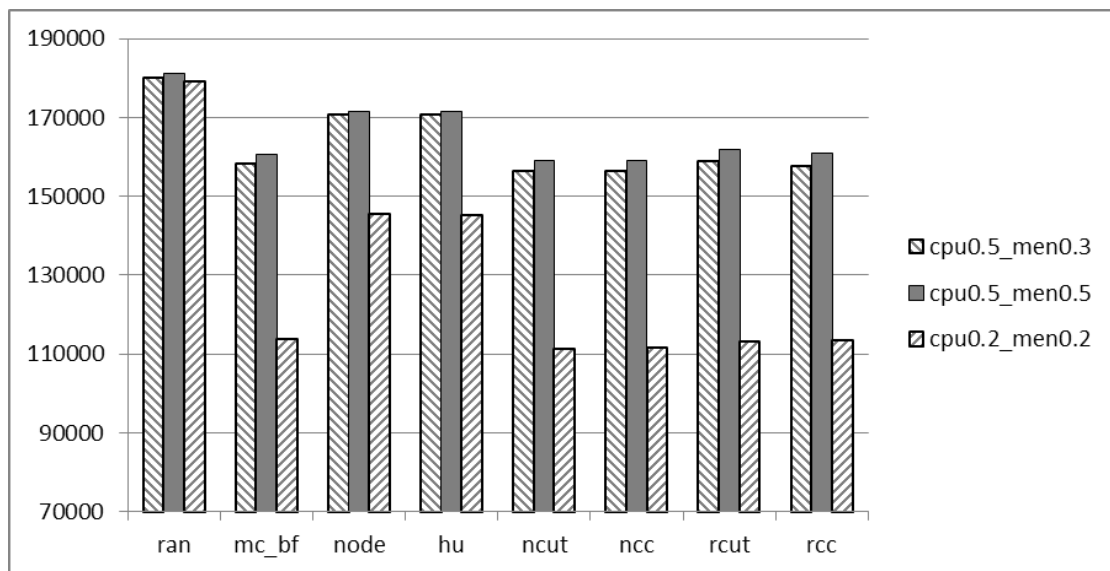


图 4-16 网络中的流量规模

从图 4-16 中可以看出，谱聚类算法对于各种虚拟机资源的情况都有一定的优化。但对于 $[0.2, 0.2, 1]$ 的虚拟机资源放置到物理机时，聚类算法的优化作用发挥的最大。

4.3.4 对于各层流量优化的分析

这一小节将对拓扑结构中每一层的流量总和进行分析，分别为核心层/聚合层/接入层（Core/Aggregate/ ToR）。

纵坐标表示虚拟机排序后，最先匹配放置到物理机后各层的流量总和。

横坐标表示各个虚拟机排序算法的缩写。

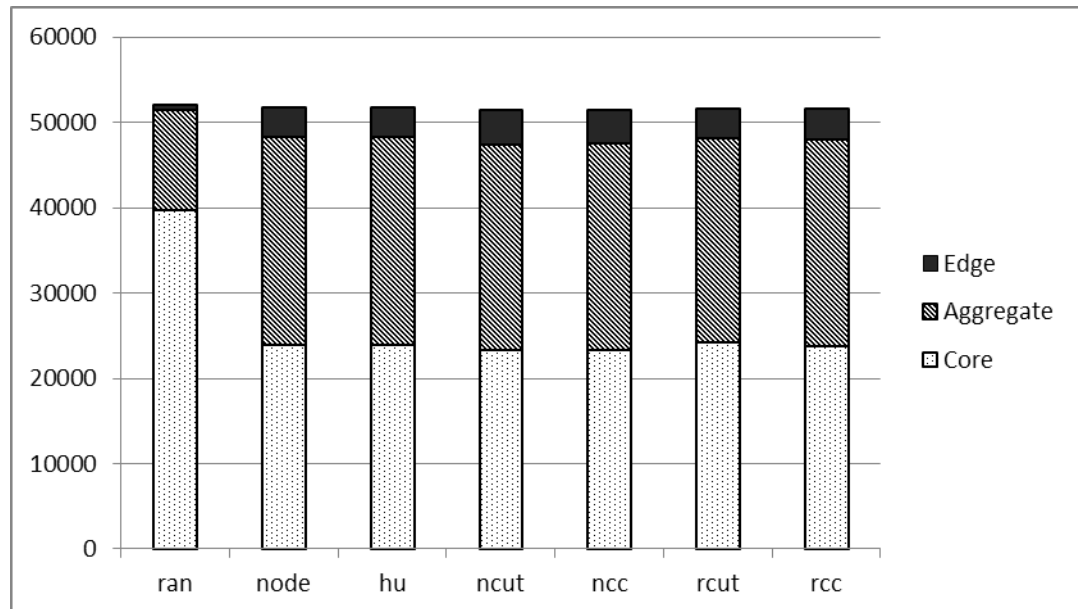


图 4-17 [1, 25]流量矩阵放置后各层流量

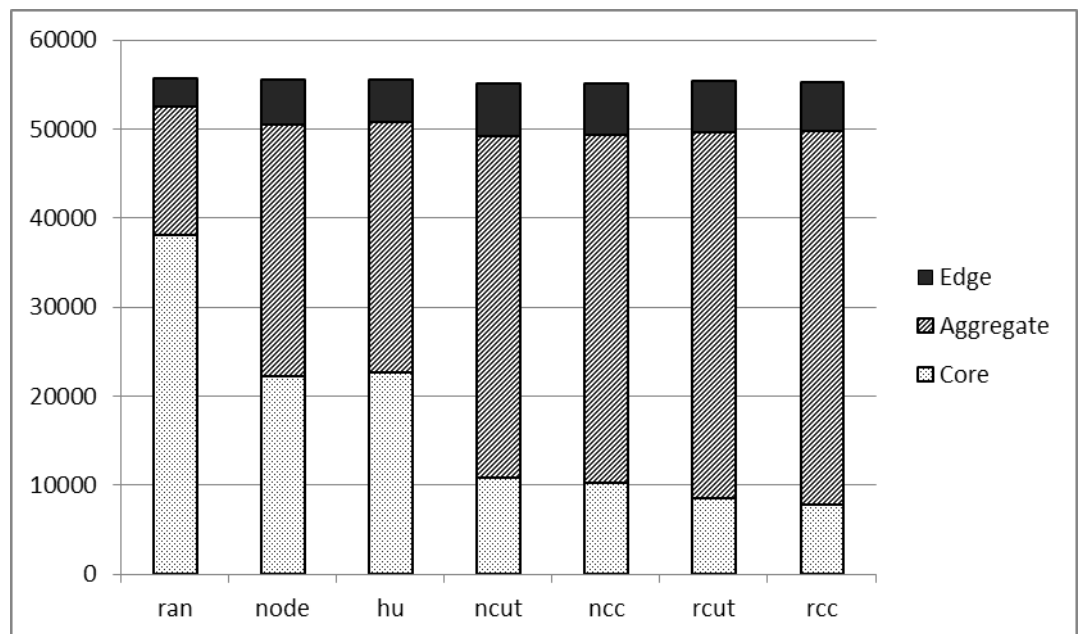


图 4-18 [3, 15]流量矩阵放置后各层流量

从图 4-18 和图 4-17 可以明显看出谱聚类算法大大降低了核心层中的流量。因为能耗的目标值为各层流量和距离的乘积，总流量是相同的，但核心层和聚合层的流量最多，所以核心层的流量越少，目标值也越小，就能节约更多的能耗。各层流量的差别显示了聚类算法对节能的优化。

4.3.5 结论

这一章从各个指标对谱聚类算法的节能效果进行分析，发现均取得了不错的结果。

4.4 禁忌搜索算法仿真结果分析

4.4.1 虚拟机资源对禁忌搜索的影响

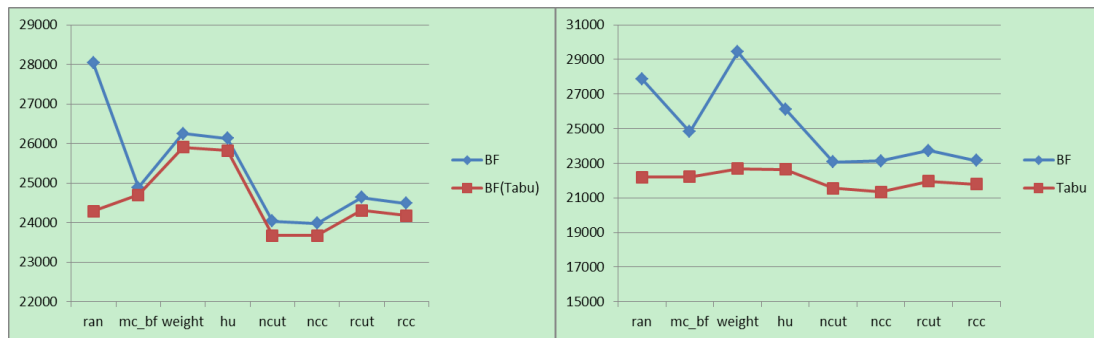


图 4-19 对[5, 5]目标值的优化
虚拟机资源为[0.5, 0.5, 1]

图 4-20 对[5, 5]目标值的优化
虚拟机资源为[0.2, 0.2, 1]

从上面图 4-19 和图 4-20 可以看出，当虚拟机资源的平均值更小时，禁忌搜索的优化越有效。原因是当虚拟机平均资源更小的时候，尝试移动虚拟机时，能更多的满足之前物理机的约束条件，达到优化网络规模的结果。

4.4.2 最先匹配和最优匹配对禁忌搜索的影响

对于最先匹配放置后的虚拟机进行搜索。

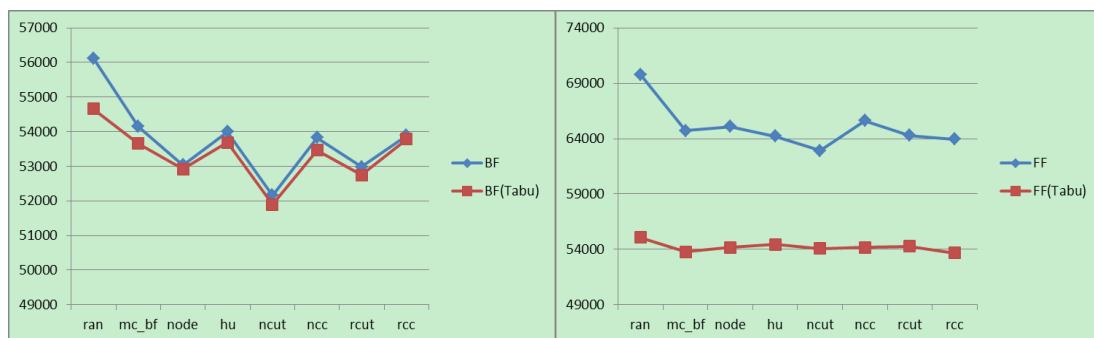


图 4-21 禁忌搜索对[1, 15]目标值的优化



图 4-22 禁忌搜索后的最终结果

从图 4-21 可以看出，当虚拟机的放置策略为最优匹配时，很难在对虚拟机进行禁忌搜索继续优化。因为最优匹配后，大部分的物理机资源都被占满，受限于约束条件，很难寻找合适虚拟机进行移动寻找更优解。

从图 4-22 可以发现虽然禁忌搜索对最先匹配优化远远超过最优匹配，但最终优化的结果还是无法超过最优匹配。

所以在对虚拟机进行放置时，还是应该选择最优放置策略。

4.5 不同拓扑结构图

这一节对传统的树形拓扑结构（Tree）拓扑结构进行分析，虚拟机模拟消耗资源为[0.5, 0.5, 1]。

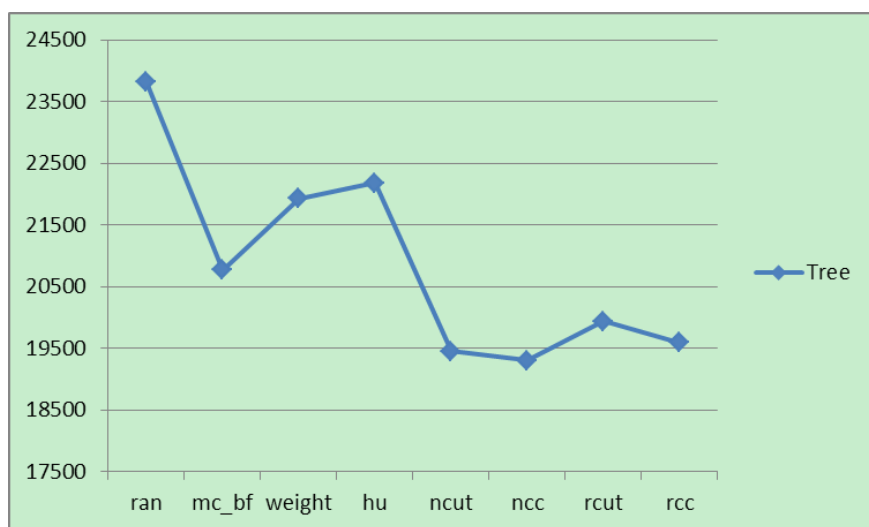


图 4-23 对[5, 5]流量矩阵的节能目标值优化

从图 4-23 可以看出谱聚类算法对传统的树形结构的优化和对胖树相似。说明了谱聚类算法对其它云计算数据中心拓扑结构也能有不错优化。

4.6 本章小结

本章通过制定不同变量，从优化目标，平均跳步数，各层流量大小，和不同拓扑结构进行分析，对比了各个算法对节能问题的优化。放置虚拟机后，对物理机资源利用情况进行分析，找出最优放置和最先放置各自的优势。

大量的实验验证了谱聚类算法，最优匹配放置算法和禁忌搜索算法对节能目标值优化的优势，大大降低了云计算数据中心的能耗。

第五章 总结与展望

5.1 工作的总结

本文将图论的谱聚类算法利用到虚拟机的放置策略中，在数据中心减少能耗的方面取得了不错的效果。在优化能耗的同时满足网络业务量需求。

本文以网络节能为目标，将物理机的能力作为约束条件优化问题。并将放置算法分解成三个子问题，一步一步解决它。本文最后实施和运行中的数据中心流量负载的模拟环境评估设计的算法。

在网络拓扑结构中，对模拟的虚拟机流量的评估表明，通过对虚拟机合理的放置，此算法比之前的一些算法相比，更高效地减少了网络能耗。

5.2 未来工作的方向

1. 当前的优化能耗目标值是流量和距离的乘积，虽然它反应了总能耗，但是实际的能耗并没有给出。未来将利用物理机和网络设备来计算出更准确的能耗总值，来衡量算法的优化效果。

2. 本文没有考虑交换机的优化问题，未来将需要交换机算法对交换机的数量进行优化。因为只要分配的每个链路上的总流量不超过链路的容量，不同的虚拟机之间的流量要共享尽可能多的交换机和链路。这样的优化目标未来可以通过基于路由算法实现。

3. 当前的算法只是静态的初始化放置，如果遇到虚拟机流量或者资源的变化，放置策略就不再适用。所以未来将考虑动态放置，对虚拟机随着时间进行动态的迁移。

参考文献

- [1] J. Dong, X. Jin, H. Wang, Y. Li, P. Zhang, and S. Cheng, "Energy-Saving Virtual Machine Placement in Cloud Data Centers," in *Cluster, Cloud and Grid Computing (CCGrid)*, 2013 13th IEEE/ACM International Symposium on, 2013, pp. 618–624.
- [2] B. Heller, S. Seetharaman, P. Mahadevan, Y. Yakoumis, P. Sharma, S. Banerjee, and N. McKeown, "ElasticTree: saving energy in data center networks," in *Proceedings of the 7th USENIX conference on Networked systems design and implementation*, 2010, pp. 17–17.
- [3] B. Zhang, Z. Qian, W. Huang, X. Li, and S. Lu, "Minimizing Communication traffic in Data Centers with Power-aware VM Placement," in *Innovative Mobile and Internet Services in Ubiquitous Computing (IMIS)*, 2012 Sixth International Conference on, 2012, pp. 280–285.
- [4] O. Biran, A. Corradi, M. Fanelli, L. Foschini, A. Nus, D. Raz, and E. Silvera, "A Stable Network-Aware VM Placement for Cloud Systems," in *2012 12th IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing (CCGrid)*, 2012, pp. 498–506.
- [5] W. Fang, X. Liang, S. Li, L. Chiaraviglio, and N. Xiong, "VMPlanner: Optimizing virtual machine placement and traffic flow routing to reduce network power costs in cloud data centers," *Computer Networks*, 2012.
- [6] X. Meng, V. Pappas, and L. Zhang, "Improving the scalability of data center networks with traffic-aware virtual machine placement," in *INFOCOM, 2010 Proceedings IEEE*, 2010, pp. 1–9.
- [7] On Heller Srinu Seetharaman Priya Mahadevan Yiannis Yakoumis Puneet Sharma Sujata Banerjee Nick Mck. Elastictree: Saving Energy In Data Center Networks[J]. Nsd, 2010, 51(11):32 - 38.
- [8] Zhang B, Qian Z, Huang W, et al. Minimizing Communication Traffic in Data Centers with Power-Aware VM Placement[C]// 2012 Sixth International Conference on Innovative Mobile and Internet Services in Ubiquitous Computing. IEEE Computer Society, 2012:280-285.
- [9] Jennings B, Stadler R. Resource Management in Clouds: Survey and Research Challenges[J]. *Journal of Network & Systems Management*, 2014:1-53.
- [10] 李丹, 陈贵海, 任丰原,等. 数据中心网络的研究进展与趋势[J]. *计算机学报*, 2014, 37(2):259-274. DOI:10.3724/SP.J.1016.2014.00259.

附录

I 英文原文

Amazon.com Inc., Microsoft Corp. and Google Inc. are warring over the future of corporate computing, and executives like Michael Simonsen are reaping the benefits.

Amazon's eight-year-old business, called Amazon Web Services, pioneered the notion of leasing computing power, sparing companies the costs of building their own computing backbone. So far, it has primarily appealed to small firms like Altos.

Microsoft and Google recently bolstered their own offerings, sparking a three-way price war. Within days last month, each company cut prices on various services by up to 85%.

That is changing the math for corporate executives who spend roughly \$140 billion a year to buy computers, Internet cables, software and other gear for corporate-technology nerve centers.

The tussle will determine how companies orchestrate the computing that runs their businesses, and it threatens the makers of traditional computing-center equipment such as International Business Machines Corp., Hewlett-Packard Co. and EMC Corp. A spokesman said IBM offers both outsourced computing services and 'higher-margin software.' H-P executive Bill Hilf said it is focused on helping organizations modernize their computing backbones to own, rent or a mix. EMC president Jeremy Burton said its sales are growing quickly to providers of Web-friendly computing.

The price war is already delaying the day of reckoning for companies that thought they would one day own their own computing centers.

Marketing-technology firm Krux Inc. has used Amazon's computers to personalize information or ads on customers' websites since it started four years ago. These days CEO Tom Chavez says the company sifts through 160,000 pieces of data a second. He figured Krux would need its own corporate-computing hubs once 750 million Web surfers encountered its technology each month. But it's now at more than double that level, and Mr. Chavez says he isn't planning a move anytime soon.

'These latest price wars among the big guys are sure to postpone it even further, which is terrific for my business,' Mr. Chavez says. He is also looking at Microsoft's computing-rental offering, called Azure. Mr. Chavez estimates Krux would spend five

or six times as much on people and equipment to operate its own computing infrastructure.

In all, researcher Gartner Inc. estimates companies will spend \$13.3 billion this year on pay-as-you-go computing horsepower from Amazon and others, up 45% from a year ago. That's still less than 10% of the total spending on corporate computing centers.

Amazon is by far the biggest player, with revenue last year of more than \$3 billion, up 85% from a year earlier, according to Bernstein Research, which calculates its figures differently than Gartner. Bernstein estimates that Microsoft and Google each pulled in several hundred million dollars from their offerings. None of the companies discloses the actual figure.

Behind the growth are big savings. A medium-sized website with about 50 million page views a month might spend about \$1,200 a month to buy two computer servers, hardware to push data to the Web and other gear, according to calculations by Simon Margolis at technology-consulting firm SADA Systems. The same company might pay roughly \$270 to \$530 to rent equivalent computing power from Amazon, Microsoft or Google, he says.

Consultants say such savings are tempting some bigger companies to rent, rather than own, more of their computing power. Already, about 87% of technology executives say they use an outsourced computing provider for at least one task, according to a recent survey by consultant RightScale Inc. But it's rare for a large company-- Netflix Inc. is the most cited example--to operate primarily this way.

'The cost models [clients have] done have all gone out the window' because of the price cuts, says Sebastian Stadil, founder of Scalr Inc., which helps companies manage their computing backbone. Mr. Stadil says an Internet firm that his company works with is considering moving some workloads to Google's Compute Engine because of the recent price cuts and a new policy that applies discounts without requiring a service contract.

There are obstacles to the rental model that may limit growth. It's often cheaper and more reliable for large companies, especially those with predictable computing needs, to own or control their own computing backbone. Companies also often have to rework many of their older software programs to run on another company's computers.

In highly regulated industries such as health care, companies often want to keep sensitive digital information on their own computers. Recent revelations about U.S.

government surveillance programs have also made customers of outsourced computing providers ask tougher questions about the data's security, according to technology vendors and customers.

Some companies, including Comcast Corp., try to blend the approaches, applying Amazon's techniques on their own computing centers. Forrester Research says that about a third of the big companies it surveyed last year had created an Amazon style of pay-per-use technology on computing hubs they owned.

In the rental business, the heightened competition is prompting some customers to switch vendors. CliQr Technologies, which helps companies calculate the costs of their computing needs, says more customers in the last year have jumped from one outsourcing provider to another, particularly after Google revamped its offerings.

II 中文翻译

译文:

亚马逊公司(Amazon.com Inc.)、微软(Microsoft Corp.)和谷歌(Google Inc.)正在就商用计算市场的未来厮杀,而像西蒙森(Michael Simonsen)这样的公司管理人员则从中获益。

亚马逊有八年历史的云计算服务 Amazon Web Services 可谓是这一领域的先驱,率先提出了出租计算能力,帮助公司节约自己打造计算基础设施的成本。到目前为止,该服务主要吸引了像 Altos 这样的小公司。

微软和谷歌近来也加大了它们自己的服务,引发了一场三方价格战。上个月,就在数天时间里,每家公司把不同服务的价格下调了至多 85%。

这正在改变企业高管的做事方式。这些高管每年花费 1,400 亿美元用于为企业技术中心购买电脑、互联网线路、软件和其他设备。

价格战将决定各公司如何使用负责经营它们业务的计算,同时也对国际商用机器公司(International Business Machines Corp., IBM)、惠普公司(Hewlett-Packard Co., HPQ)和 EMC Corp.等传统计算中心设备制造商构成威胁。一位发言人说,IBM 同时提供外包计算服务和更高利润率的软件服务。惠普管理人员希尔夫(Bill Hilf)称,该公司专注于帮助各种组织实现计算设施的现代化,具体方式可以是自有、租赁或二者兼有。EMC 总裁伯顿(Jeremy Burton)说,该公司面向网络计算供应商的销售正在快速增长。

对那些认为有一天会拥有自己计算中心的公司而言,上述价格战正在推迟它们实现这个愿望的时间。

自四年前创办以来,营销技术公司 Krux Inc.就使用亚马逊的计算机为客户网站上的信息及广告提供个性化服务。如今 Krux 首席执行官查维斯(Tom Chavez)称,该公司可在一秒内对 16 万条数据进行筛选。他原本认为,一旦网络用户对其技术的使用量达到每月 7.5 亿人次的水平, Krux 就需建立自己的企业计算中心。当前使用量已达这一数字的两倍多,但查维斯表示他近期没有这方面的打算。

查维斯说,目前行业巨头之间的价格战,势必会使 Krux 建立自有计算中心的计划进一步延后,这对他的业务非常有利。他还在关注微软提供的云计算平台服务 Azure。查维斯估计,若 Krux 选择自主运营计算基础设施,则所需支付的人力、物力将是现有支出的五到六倍。

研究公司 Gartner Inc.预测,今年各企业将支付 133 亿美元以使用亚马逊等公司的现付现用制的云计算服务,较一年前增长 45%。而这一数字尚不及企业计算

中心费用总额的 10%。

据 Bernstein Research（数据计算方法不同于 Gartner）称，亚马逊是目前云计算服务领域的最大企业，去年收入超过 30 亿美元，较上年同期增长 85%。据 Bernstein 估计，微软和谷歌各凭藉云计算服务创造了几亿美元的收入。上述三家公司均未公布确切数字。

云计算业务增长的背后是大笔费用的节省。据科技咨询公司 SADA Systems 的马戈利斯(Simon Margolis)计算，一个每月页面访问量约五千万次的中型网站，平均每月或需花费 1,200 美元购买两台计算机服务器及数据推送硬件等设备。而马戈利斯称，同样一家公司若从亚马逊、微软或谷歌租用同等计算能力的设备，只需支付大概 270 到 530 美元。

咨询人士称，如此划算的价格正促使一些规模更大的企业不再建设自己的计算中心，而是改之以租用的方式来增强计算能力。据咨询公司 RightScale Inc. 近期的一项调查，已有 87% 的科技企业高管称，至少已将一项计算任务外包给了计算服务供应商。不过大公司（例如经常被拿来举例的 Netflix Inc.）一般不会主要采取这种运营方式。

Scalr Inc. 的创始人 Sebastian Stadil 说，由于三大公司的价格战，客户已有的成本模型都被扔到了窗外。Scalr 是一家帮助企业管理计算基础设施的公司。Stadil 说，考虑到近期的减价潮，以及一项打折无需服务合同的新政策，与该公司有合作关系的一家互联网企业正在考虑将部分工作量转移至谷歌的计算引擎。

租赁模式面临的障碍可能限制该行业增长。对大公司，尤其是那些可以预测计算需求的公司来说，拥有或者控制自有计算基础设施通常更加廉价，也更加可靠。此外，这些公司往往还需要改进许多陈旧的软件程序，才能在另外一家公司的电脑上运行数据。

在医疗保健等被高度监管的行业，企业通常希望将敏感的数字信息保存在自己的电脑里。根据技术供应商和客户反映，美国政府监控计划近期被披露之后，外包计算提供商的客户围绕数据安全提出了更尖锐的问题。

包括康卡斯特(Comcast Corp.)在内的一些公司尝试采用一种混合的方法，即将亚马逊的技术应用于自有计算中心。Forrester Research 表示，在去年接受该公司调查的大型企业中，大约有三分之一已将亚马逊的技术融入了自有计算中心，按使用次数向亚马逊付费。

租赁业务的竞争加剧则促使一些客户更换供应商。CliQr Technologies 称，去年有更多的客户更换了外包供应商，尤其是在谷歌改进了云计算服务之后。CliQr Technologies 是一家帮助企业预测计算需求成本的公司。

致谢

经过一个学期的准备，大学本科毕业论文终于已经接近尾声了。在这里要感谢计算机学院老师孜孜不倦的教导和专业知识的授予，感谢学校给我们营造了一个良好的学习环境，让我们拥有一个和谐愉快和充满活力的大学四年。

本次毕业设计本是我的校内“师生合作”课题，此课题最后也成为了我的毕业设计。因为之前大量的准备工作，论文得以顺利的完成。首先感谢这个课题和这篇论文，让我将所学的专业知识运用到实践中，并慢慢了解到如何从事学术工作。感谢我的导师陈志老师，不管是学习还是生活，每次遇到问题，陈老师都能好朋友一样热心，耐心地对我进行指导，指出我的各种不足之处，给我点明人生的方向：)。

非常感谢陪伴我大学四年的所有朋友。四年的时光，因为有你们才变得更加精彩。可是这样的时光总是过得太快，快的感觉就像是昨天发生的那样。也许未来我们在天南海北，但是我相信我们依旧会心系彼此，我们的友谊会天长地久。当然也要感谢我的女朋友，一路走来她让我慢慢变得成熟，灿烂的微笑给了我活下去的动力和奋斗的目标。

与此同时，我也要非常感谢我的父母和家人，没有他们的养育，我不可能有现在的自己。风风雨雨二十载，他们一直是我坚强的后盾，成功时为我喝彩，失败时为我打气，没钱时就给我打钱。我相信在未来，凭我的坚持和努力，我一定能在学业和事业上闯出一片天，成为他们心中的骄傲。当然，我也希望他们能健康快乐，他们的健康快乐也是我努力奋斗的意义。

最后，要感谢评审本论文的老师。由于专业水平的不足，文章难免会有错误和缺陷，也希望你们多多指正，提出宝贵的意见和建议。