高能效互联网传输技术研究

张国强[1,2,6]，林森[2]，刘真[3]，林涛[4]，张国清[2]，李幼平[5]

(1. 南京师范大学 计算机科学与技术学院，江苏 南京 210012； 2. 中国科学院 计算技术研究所，北京 100190;
3. 北京交通大学 计算机与信息技术学院，北京 100044；4. 中国科学院 声学研究所，北京 100190；
5. 中国工程物理研究院，北京 100083; 6. 苏州大学 江苏省计算机信息处理技术重点实验室，江苏 苏州 215000）

摘 要：从实现网络传输过程中的能耗比例计算理念以及降低网络中数据的重复传输 2 个角度综述了降低网络能耗的方法。实现能耗比例计算理念的技术包括边缘网络的网络存在性代理技术、以太网节能技术和核心网络的节能路由技术。人类对数据访问的异步性需求以及对数据访问呈现重尾分布的规律从宏观上为减少数据的重复传输提供了前提。比较了互联网上现有的和处于实验阶段的多种内容分发方式，包括 CDN、P2P、CCN 和双结构互联网，探讨了它们对提高网络传输能效的作用。

关键词：绿色互联网；能耗比例计算；网络存在性代理；节能路由；未来互联网

中图分类号：TP301 **文献标识码：**A **文章编号：**1000-436X(2012)05-0158-11

Energy efficient data transmission on the internet

ZHANG Guo-qiang[1,2,6], LIN Sen[2], LIU Zhen[3], LIN Tao[4], ZHANG Guo-qing[2], LI You-ping[5]

(1. School of Computer Science and Technology, Nanjing Normal University, Nanjing 210012, China;
2. Institute of Computing Technology, Chinese Academy of Sciences, Beijing 100190, China;
3. School of Computer and Information Technology, Beijing Jiaotong University, Beijing 100044, China;
4. Institute of Acoustics, Chinese Academy of Sciences, Beijing 100190, China; 5. China Academy of Engineering Physics, Beijing 100083, China;
6. Provincial Key Laboratory for Computer Information Processing Technology, Soochow University, Suzhou 215000, China)

Abstract: Approaches for reducing network energy consumption were surveyed from two aspects: realizing energy-proportional computing in data transmission, and reducing duplicated data transmission in the network. Approaches endeavor to realize energy-proportional computing include network presence proxy in edge networks, energy-efficient Ethernet and energy-aware routing techniques in core network. The asynchronous data access nature and the heavy-tailed frequency distribution for data access offer opportunities to reduce duplicated data transmission. Several existing and experimental content distribution mechanisms (e.g. ,CDN, P2P, CCN and dually structured internet) and their effects in enhancing energy efficiency in data transmission, were compared and investigated.

Key words: green internet; energy proportional computing; network presence proxy; energy efficient routing; future internet

收稿日期：2010-12-31；**修回日期：**2011-05-20

基金项目：国家自然科学基金资助项目（61100178, 61174152**）；南京师范大学科研启动基金资助项目（**2011119XGQ0248**）；北京市自然科学基金资助项目（**4112057**）**

Foundation Items: The National Natural Science Foundation of China (61100178, 61174152); The Startup Foundation of Nanjing Normal University (2011119XGQ0248); The Natural Science Foundation of Beijing (4112057)

1　引言

随着互联网用户的增长、高带宽需求的业务增加以及物联网应用的普及，互联网的联网设备数和所产生的数据量均呈现爆炸性增长的趋势，导致其能耗在过去若干年也急剧增长。统计表明，美国互联网消耗的能源已经占了所有能耗的 2%~10%[1,2]。能耗的急剧增长对网络的运营和互联网的持续发展都构成了严峻的挑战。一方面，能源开支在网络运营商和内容提供商的总体拥有成本 (TCO) 中所占的比例越来越重，降低能耗可以节省大量的开支。另一方面，由于传统的散热和冷却技术正遭遇瓶颈，能耗已经成为制约互联网进一步发展的主要因素之一。此外，对温室气体和全球变暖的担忧也对绿色通信提出了要求，有统计表明互联网的能耗已经超越了整个航空业的能耗[2]。中国在哥本哈根会议后提出了在 2020 年使单位 GDP 产值的二氧化碳排放量较 2005 年降低 40%~45%的宏伟目标，为了实现该目标，需要各行业的共同努力。

目前，针对具体硬件设备和部件的节能技术得到了广泛研究。在设备和系统级别，设备厂商提供了多种电源管理模式。高级配置和电源接口 (ACPI)[3]对系统不同的电源状态进行了规定并提供了相应的软件管理接口。美国能源署(EPA)于 1992 年启动了“能源之星”计划[4]，为符合其标准的产品贴上能源之星的标签。中国环境保护局也有类似的标准，如为计算机采购制定的环保标准[5]。但是，在网络层面提高数据传输的能效并没有得到太多的实际应用。一方面，网络边缘节点对联网的需求以及核心网络为应对峰值负载和网络顽健性的需求使得现有的设备级电源管理功能未得到有效利用。而实际上，终端节点、网络设备和网络链路的利用率通常都很低，致使现有的互联网能效低下，未能实现能耗比例计算的理念。另一方面，由于用户对数据对象的请求呈现重尾分布规律，基于端到端的单播传输方式造成了大量的重复数据传输，使得传输效率低下。

本文正是基于此背景，从两方面来探讨互联网的高能效传输技术。一方面，在假定现有网络流量需求不变的前提下，以能耗比例计算为轴，综述实现该理念的技术，具体包括边缘网络的网络存在性代理技术、以太网节能技术和核心网络的节能路由技术。另一方面，从降低网络流量的角度出发，探讨现有的及正处于实验阶段的内容分发架构对提高网络传输能效的作用。

2　网络节能研究现状

美国劳伦斯国家实验室的统计表明：2008 年，联网设备的能耗占所有能源消耗的 4%左右，而其中，路由器、交换机等网络设备占了 0.5%左右(见图 1)。在中国，2009 年三大运营商耗电 28.9TWh，较前一年同比增长 26%，耗电总量占全社会用电量的 0.8%，而且现在每年还在快速增长。据预计，至 2020 年将有 1/7 的电力被 ICT 产业所消耗[6]，而其中 PC 和网络设备将占据 1/3。但是，被消耗的能源中一大部分都被浪费了。虽然终端节点提供了多种电源状态，如待机、休眠等，但是随着网络业务的渗透，各种“反电源管理”因素正在不断增加，如 P2P 应用、远程访问、远程管理等。这些因素促使用户禁用电源管理功能，而将终端节点置于 7×24h 开机模式。而对网络设备来说，电源管理基本没有得到应用。即使在负载很低的时候，网络设备也必须保持工作状态，以应对可能出现的链路故障等不确定因素。这些现实浪费了大量的能源，而如果采取有效措施，则可以最大程度予以避免。目前，学术界已经启动了绿色互联网的相关研究计划，如南佛罗里达州的绿色互联网研究计划[7]和加州伯克利劳伦斯国家实验室的高能效数字网络研究计划[8]。工业界也于 2007 年 6 月由英特尔公司和谷歌公司倡议并携手超过 25 家企业和环保机构在美国共同发起了“电脑产业拯救气候行动计划”，次年，中国电子学会也加入了该计划。

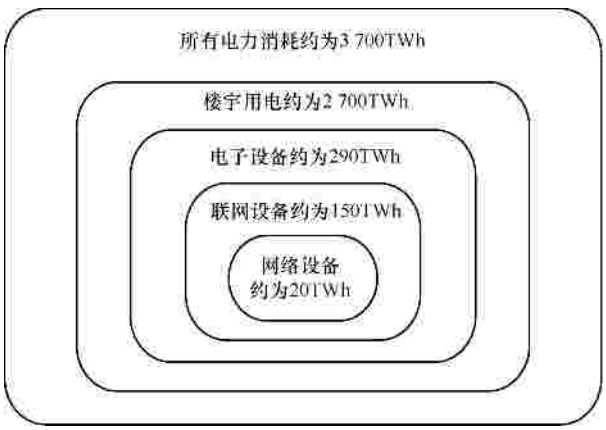

图 1　美国 2008 年的电力消耗

3 基于能效比例计算的网络传输节能技术

3.1 能效比例计算的理念

能耗比例计算(energy proportional computing)是实现绿色计算的关键理念[9]。这一理念将成为未来系统设计的目标之一，即系统的能耗应该与工作负载成正比，而不是与最大处理能力成正比。图 2 给出了能耗比例计算。在理想情况下， 当工作负载为零时，能耗也应趋近于零。而目前的绝大部分现实是，系统在工作负载为零时的基准能耗一般都超过了峰值负载能耗的 50%。若将整个网络看成一个系统，并将这一理念延伸至网络传输，则意味着网络的能耗应该与网络的负载正相关。理想情况下，网络为传输 1bit 信息量所耗费的能量应为常数。

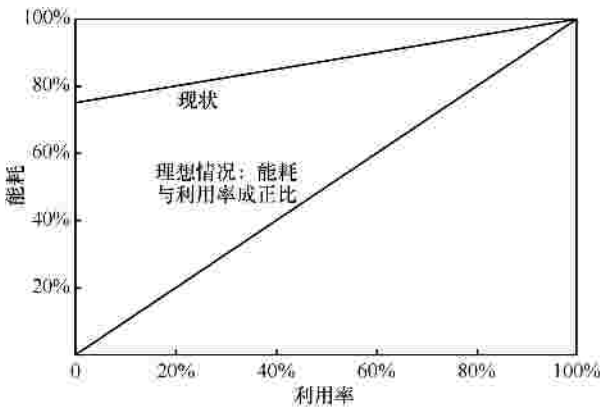

图 2 能耗比例计算

在硬件和系统层面提供多种电源状态是实现能耗比例计算的基础。目前，这方面的技术已经日趋成熟。例如，在 CPU 层面，动态电压频率调整(DVFS)等技术的应用已使 CPU 愈发接近了能耗比例计算的目标；在系统层面，高级配置和电源接口(ACPI)[3]也定义了多种休眠模式(S 状态)和性能模式(P 状态)，以支持系统级的能耗比例计算。然而电源管理与高能效计算存在本质区别：前者仅提供了硬件支持，而后者是一个全局优化问题[10]，需要高能效算法的支持[11]。传统的系统一般都是以最大化性能为设计目标。一般来说，高性能和高能效并不互斥。但事实上，要求系统的所有硬件都工作在最大能力的场景并不多见。文献[10]概括了在给定计算问题下实现高能效计算所必备的三要素。

1) 电源模型。详细规定了每个电源状态的能耗、状态转换的延迟以及状态转换的能耗。

2) 约束确定和性能评价模块。用于给出计算问题的应用层性能约束条件，如完成时间。

3) 能源优化器。基于电源模型和现有资源的工作负载调度器，满足应用性能约束条件，同时最小化所耗费的能源。

下面将分别对边缘网络和核心网络的节能技术进行评述，对构成高能效计算的三要素进行抽象。

3.2 边缘网络节能技术

3.2.1 网络存在性代理

网络存在性代理技术[1,12~16]是一种边缘网络的节能解决方案。目前，为时刻保持网络存在性已经成为了网络边缘节点禁用电源管理的主要因素，如远程登录、远程管理以及 P2P 等资源共享应用。但实际上，有效工作时间仅占总开机时间的一小部分，绝大部分无人值守时间内节点均处于空闲状态。在现有的网络应用模式下，若节点进入休眠，则同时也失去了对外的网络存在性。网络存在性代理技术能有效解决上述问题。它允许被代理节点进入休眠状态，而同时对外保持其网络存在性。一个节点的网络存在性是由它的对外行为表现的，为了维持节点的网络存在性，需要能对特定的外部请求消息加以响应，维持对外连接。

图 3 给出了网络存在性代理的工作方式示意：1) 初始时，节点处于活跃状态，它与其他节点之间的数据传输直接发生在它们之间；2) 当节点空闲了一段时间希望进入休眠时，它将该意图通知网络存在性代理，节点可能需要和代理之间进行一定的状态传输，使代理能获知节点的当前状态，之后，节点进入休眠状态；3) 在节点休眠期间，网络存在性代理截获发往被代理节点的网络报文，并予以处理；4) 当某些报文无法处理而必须唤醒被代理节点时，网络存在性代理将唤醒节点，并将报文转交给节点，同时清理被代理节点的状态，之后，节点恢复与目标节点的直接通信。

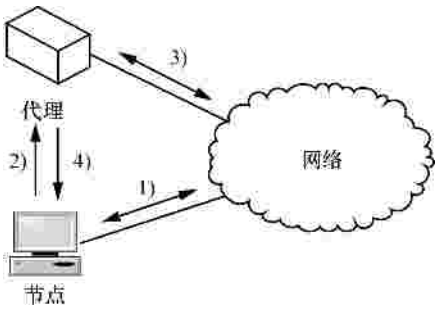

图 3 网络存在性代理的工作流程

上述场景中，对应的高能效计算问题中的三要素为：每个被代理节点活跃状态和休眠状态的能耗、状态转换的能耗和转换时间、以及网络存在性代理的能耗构成了电源模型；每个节点 i 都有一组在休眠时需要支持的网络应用需求 $APP(i)$，所有应用的约束构成了整个计算问题的约束集，通常表现为处理延迟的约束；而节点进入休眠的策略和网络存在性代理的处理逻辑则构成了能源优化器。

网络存在性代理的处理逻辑是该问题的关键，其目的是在代理设计的复杂性、节能效果和应用层性能之间寻求有效折中。代理所支持的行为集合决定了节点何时需要被唤醒。在最简单的模式下，代理不做任何处理，即对任何报文都唤醒原节点，这种方式称为 WoP(wake on packet)，实际上该模式下完全不需要代理。WoP 的有效性与报文到达的间隔时间密切相关。研究表明，在办公网络环境，大量的周期性广播和多播报文的存在使得 WoP 方法的有效性基本为零；而在家庭网络中，报文到达间隔时间更具重尾分布特性，因此，WoP 方法具有一定的效果[16]。

WoP 方法的低效说明：为实现更佳的节能效果，网络存在性代理需要支持更复杂的处理逻辑。一些典型的处理逻辑可以包括：1) 忽略某些报文（如广播报文或其他无关紧要的报文），而对其他报文则唤醒被代理节点；2) 代为应答一些简单的网络报文，如 ARP、PING 等，而对其他报文则唤醒被代理节点；3) 代理部分应用程序的功能，而对不支持的应用报文则唤醒被代理节点；4) 在复杂性和支持的应用种类之间进行折中，例如，仅对某些应用唤醒被代理节点，而对其他应用的报文则丢弃，这些应用在被代理节点休眠时将不再得到支持。

目前，ECMA (european computer manufacturers association)正在着手制定网络存在性代理需要支持的处理逻辑及其行为的标准[12]。按照不同的网络层次将网络存在性代理所支持的处理逻辑进行简单分类如下。

1) 网络层：应支持 IPv4 的 ARP 协议和 IPv6 的邻居发现协议；若使用 DHCP 协议，则还应支持 DHCP 以维持 IP 地址；若被代理节点参加了多播组，则还需要支持 IGMP。支持这些协议确保了被代理节点在休眠时的网络可达性，使发往被代理节点的报文能被正确寻址。

2) 传输层：需要维持应用可达性和 TCP 连接，如响应 TCP SYN 和 TCP 定时消息。

3) 应用层：对每种被代理的应用，需要具备处理简单的应用层请求和心跳消息的功能。例如，文献[17]提出了一种能用于处理 Gnutella 查询消息的代理，仅当需要传输文件时才唤醒节点。

网络存在性代理的具体表现形态（或部署位置）可以是网络中间件（如已有的防火墙和 NAT 等设备）、同一局域网上的另一个网络节点或自身的网卡。其中前两者称为外部代理，而后者称为内部代理。如果使用外部代理，则应支持网络报文唤醒功能。网络报文唤醒可以是基于魔分组（magic packet）的 WOL 技术[18]，也可以是基于特定报文模式匹配的技术，如 TCP SYN 或更细粒度的自定义报文。而使用内部代理则需要扩展现有网卡的功能。Yuvraj Agarwal 和 Steve Hodges 等人设计了一种基于 USB 接口的智能网卡[13]。它在网卡中内嵌了一款低能耗的处理器，运行嵌入式操作系统，在节点休眠时代理休眠节点的角色。

网络存在性代理的优点在于无需对现有的基础设施做任何改变。近期来看，外部代理可以有效利用现有网络设备短期内实现节能的目标[15]；而长远来看，内部代理和被代理操作系统之间的协同更为简单，具有更广泛的应用前景。

3.2.2　以太网节能技术

在链路层面，以太网技术已成为局域网组网的主要技术。调研表明，仅在美国，PC 机和其他网络设备的以太网卡在 2005 年消耗的电量为 5.3TWh[19]。传统的以太网标准中，即便没有数据传输，接收端和发送端也工作在最高能耗模式。实际上以太网的利用率通常都很低，一般在 1%左右[2]，但是以太网链路的能耗却与工作频率（而非负载）正相关（如图 4 所示）。为解决这一矛盾，研究人员提出了 2 种方案：负载自适应以太网变频技术（ALR）和以太网休眠技术。

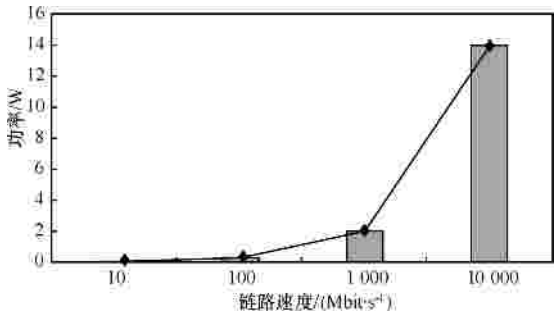

图 4　以太网的链路速率与能耗关系

3.2.2.1 基于变频的以太网节能技术

在以太网变频问题中，构成高效能计算的三要素为：以太网不同链路速率的能耗和链路速率切换时间共同构成电源模型；报文延迟则是约束集考虑的主要因素；切换策略则对应于能源优化器。

以太网变频的研究重点在于链路速率切换机制和切换策略。其中，切换机制的目标是降低切换延迟，而切换策略的目标是最大化链路处于低频的工作时间[20]。在链路建立阶段可采用现有的 IEEE 802.3 自动协商机制来进行能力的协商，包括是否支持 ALR 以及双方支持的传输速率。在链路速率切换时，ALR 提出了两阶段握手的机制，由请求方发送 ALR REQUEST 的 MAC 帧，而接收方通过 ALR ACK/NACK 确认帧来接受/拒绝切换请求。如果接收方接受切换请求，则开始链路重新同步。在切换策略方面，ALR 提出了双门限、利用率门限和超时门限 3 种策略，用于智能地对工作频率做出切换决策。

3.2.2.2 基于休眠的高能效以太网

目前，IEEE 802.3az 标准工作组提出了基于休眠机制的以太网节能方案[21]，称为高能效以太网(EEE), 相关标准已于 2010 年 9 月制定完成。其基本思路是在无数据传输时让链路进入低能耗的休眠模式，而在新数据分组到达时快速将其唤醒。图 5 给出了高能效以太网的工作示意。其中，T_s 表示进入休眠所需要的时间，T_w 表示唤醒链路所需的时间，T_r 表示刷新时间，用于定期刷新接收器的状态，以保证接收器单元与信道环境保持一致。空闲时，以太网就进入低能耗状态，低能耗模式下的能耗通常为正常模式的很小一部分(约 10%)。

与以太网变频技术不同，基于休眠的以太网节能方案不能作为高能效计算的实例，因为这里缺乏了三要素中的能源优化器的概念，因此基于休眠的以太网节能只能被视为提供了链路级的电源管理支持。

3.3 核心网节能路由技术

核心网中的节点和链路平均利用率不高，但网络是为应对峰值负载和可能出现的故障而设计的，因此即便利用率很低，目前的网络设备也维持峰值时的工作状态。而实际上，流量按天呈现周期性[22~24]，因此可以在负载较低时让部分节点/链路进入低能耗状态（如让节点/链路进入低频工作模式或休眠模式），从而达到节能的目的，使网络作为一个整体接近能耗比例计算的理念。

将网络看成一个系统，则节能路由问题对应的高效能计算的三要素为：电源模型由所有路由器和链路的能耗模型、不同电源状态之间转换延迟和转换能耗开销构成；约束和评价模型则主要从平均延迟、分组丢失率、网络吞吐能力及链路利用率等进行考虑；最后，能源优化器通常被形式化建模为能耗感知的流量工程问题，需要在全网范围内选择能耗最小的路由，确定链路的频率或可以关闭的节点或链路，但这通常是一个 NP-hard 问题。

文献[25]开创性地提出了在网络层面进行节能的 3 种方法，其中 2 点是针对全网而言的。

1) 网络层面，可以在低负载时改变路由，在更少的路径上聚合流量，从而允许空闲路径的设备进入休眠状态。这要求对第 2 层和第 3 层协议的工作方式进行修改。

2) 可以改变互联网拓扑结构，使网络结构依据网络负载的不同而动态调整（通过聚合和休眠）。换言之，应设计能源敏感的网络拓扑，使网络的活跃设备数和网络负载之间具备良好的相关性。当负载增加时，更多的设备被唤醒，而当负载下降时，更多的设备可以休眠。

针对全网节能优化的方法也可以分为链路变频技术和休眠技术。依据节点是否进行协同，可以分为自主方式和协同方式。其中，自主方式是设备依据自身感知的网络状态，独立地做出变频或休眠决策；而协同方式则依据全局的流量需求、能耗函数等，在满足链路利用率限制或网络吞吐能力限制的前提下，从优化全局能耗的角度决定路由策略。因此，针对全网的节能优化技术可分为自主变频技

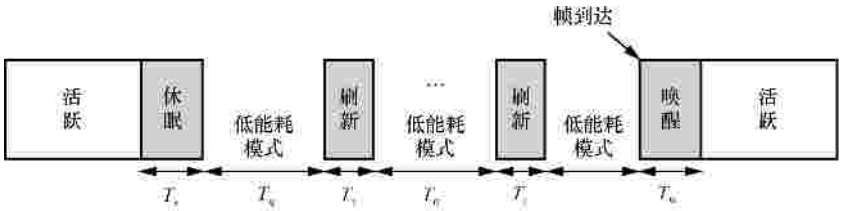

图 5 EEE 中活跃和低能耗状态之间的转换

术、协同变频技术、自主休眠技术和协同休眠技术这 4 种。

3.3.1　自主休眠

文献[26]提出了几种自主休眠的策略。第 1 种是定时休眠。在该策略下，休眠阶段到达的报文将全部丢失。第 2 种是 WoA(wake-on-arrival)，即将网卡处于供电状态，在主机休眠时侦听线路，一旦有报文到达即唤醒主机。然而在核心网络上，报文到达间隔通常很小(为微秒数量级)，因此 WoA 在这类链路上基本无效。为克服这一缺陷，可以利用流量缓存的方法在网络边缘对流量进行缓存和整形，人为地使网络流量呈现突发态势，从而增加空闲时间的长度[注1]。为保证入口处创建的突发流量在穿越网络时得以保持，入口路由器将对报文进行重新排列，使得去往同一出口路由器的报文在突发中相邻。此外，可以增加入口路由器的协同性，使到达中间路由器的报文在时间上更为接近，从而中间路由器可以获得更多的休眠机会。但完全协同并非总是可行的，对协同性的要求也大大增加了策略的复杂性。

3.3.2　自主变频

文献[26]还提出了自主变频技术，即依据网络流量自适应调整网络的工作频率。假设每个网络链路都支持 N 个速率 $r_1, r_2, \cdots, r_n$(其中，$r_i<r_{i+1}$,$r_n=r_{\max}$ 为默认最大链路速率)，不同的性能状态之间的转移时延为 d，设 $\hat{r}_f$ 表示链路依据历史报文到达速率对未来报文到达速率的预测值[注2]，d 表示每个报文排队时延的上限，则可依据下列准则来进行链路频率切换：

1) 假设链路速率为 r_i，队列长度为 q，则当且仅当 $\frac{q}{r_i} > d$ 或 $\frac{d\hat{r}_f + q}{r_{i+1}} > d - d$ 时将链路速率切换到 r_{i+1}。

2) 假设链路速率为 r_i，队列长度为 q，则当且仅当 q=0 且 $\hat{r}_f < r_{i-1}$ 时，将链路速率切换到 r_{i-1}。

3.3.3　协同休眠

协同休眠是一个典型的高能效计算问题。在性能约束方面，通常使用链路利用率[24,27]或网络吞吐能力[28]来衡量网络的传输性能。在电源模型方面，需要分别建立路由器和链路的能耗模型[24,29,30]。协同休眠研究的核心则在于能源优化器，即如何实现最优的全局决策以最大化节能效果。这类决策策略通常都是依据网络拓扑结构知识或流量需求在路由协议层面对流量进行汇聚，允许不参与路由的节点或链路进入休眠以达到节能目的。

文献[31]提出了一种 EAR(energy aware routing)算法，可以看成是网络负载较轻时的 OSPF 版本。例如，路由器在晚间可自动切换到 EAR，而在白天则采用 OSPF 路由。与 OSPF 相比，EAR 利用了更少的链路进行路由，因此不参与路由的链路可以进入休眠状态。

EAR 并没有显式地考虑流量需求，在 EAR 和 OSPF 之间的切换时机需要依赖于经验。文献[27]在选择路由时不仅考虑节能目标，同时还增加了对流量需求矩阵与服务质量约束的考虑。该研究同时考虑了路由器和链路的能耗，并将其建模为一个具有容量限制的多物资最小费用流问题(CMCF)。由于 CMCF 问题是 NP-hard 的问题，文章提出了几种关闭节点的启发式策略：随机策略、节点度最小最优先、最小流最优先、以及 opt-edge。基于同样的建模方法，文献[24]又提出了一种简单的贪婪算法用于关闭节点和链路，保证网络连通性和最大链路负载率。

笔者在文献[28]中提出了基于网络宏观拓扑结构知识选择性关闭链路的方法，实验表明，在 BA 网络上采用一种混合策略能取得较优的节能效果，即先关闭边介数(edge betweenness)较大的链路，然后切换到随机策略关闭链路。该研究也是文献[25]中所提出的第 3 种网络节能思路的一种体现，即 BA 网络比 ER 网络具备更好的“活跃设备—网络负载”相关性。

此外，路由器虚拟化技术也为节能路由提供了基础。若路由器支持虚拟化，则可以通过调整逻辑路由器和物理路由器的映射关系来提高传输能效。例如，文献[32]提出了 VROOM(virtual router on the move)，将逻辑路由器和物理路由器分离，依据流量负载动态地增加或减少物理路由器的数目。在晚间流量较低时，虚拟路由器可以迁移到更少的物理路由器上，不需要的路由器就可以关闭或进入休眠状态。应用 VROOM 的优点在于 IP 层的拓扑在迁移过程中保持不变。

3.3.4　协同变频

文献[33]提出了能耗最优化的路由模型，根据

注 1　这很反传统，在传统的拥塞避免中，一般都是避免突发流量，而不是创造突发流量。

注 2　可基于指数加权移动平均(EWMA)来获得。

能耗和负载的不同函数关系及流量需求，给出路由算法及每条链路的最佳工作频率设计方案。给定无向图 G 和一组流量需求，每个需求 i 在源节点 s_i 和目标节点 t_i 之间要求 d_i 整数单元的带宽。每一条链路 e 被赋予一个费用函数 $f_e(s)$，表示负载为 s 单元的数据时的能耗。设 0-1 函数 $y_{i,e}$ 表示流量需求 i 是否经过链路 e，x_e 表示链路 e 上的总负载。路由最优化问题（P_1）就表示为

$$(P_1)\quad \min \sum_e f_e(x_e)$$

满足：

$$x_e = \sum_i y_{i,e} d_i \quad \forall e$$

$$y_{i,e} \in \{0,1\} \quad \forall i,e$$

$$y_{i,e} \text{ ：} flow\ \ conservation$$

上述问题求解的复杂度依赖于费用函数，具体有以下结论：

1) 如果费用函数 $f_e(\cdot)$ 满足 $f_e(x_1+x_2)=f_e(x_1)+f_e(x_2)$，则最短路径路由算法是最优的；

2) 如果费用函数 $f_e(\cdot)$ 满足 $f_e(x_1+x_2)< f_e(x_1)+f_e(x_2)$，则对应于 buy-at-bulk 网络设计问题[34]；

3) 如果费用函数 $f_e(\cdot)$ 满足 $f_e(x_1+x_2)>f_e(x_1)+f_e(x_2)$，则一般不存在一个具有上界的近似多项式算法。

类似地，文献[35]在假设链路支持多个工作频率的基础上，提出了能耗感知的流量工程方法。其基本思路是如果网络中有些链路的流量负载稍高于某个工作频率，则可以将高出工作频率部分的流量通过流量工程的方法转移到一些其他链路，同时保证这些链路的流量不超过现有工作频率的上限，从而可以将一部分链路切换到低频工作状态。

3.3.5 小结

具体采用哪种技术方案依赖于网络硬件设备和系统的能耗属性。路由器的能耗一般可以表示为 $P(R)=B_R+f(v)$，其中，B_R 为基准能耗，$f(v)$ 为依据负载的可变能耗。在当前的硬件环境下，一般 B_R 占主导，基准能耗与峰值负载能耗之比一般超过 50%。而当设备实现了能耗比例计算后，基准能耗与峰值负载能耗之比将大幅降低。链路能耗一般可以表示为 $P(L)=h(L)+w(C, x)$，其中，$h(L)$ 表示与链路长度 L 相关的能耗，$w(C, x)$ 表示工作频率为 C, 负载为 x 的能耗。一般来说，$h(L)$ 可以视作常数且忽略不计，链路的能耗主要由其工作频率 C 决定。当链路不支持变频时，仅存在活跃和休眠 2 个状态；而当链路支持多个工作频率时，则不同的工作频率具有不同的能耗。

原则上来说，当网络设备的基准能耗占峰值能耗的比例较高时，采用基于休眠的方案能有效地实现网络节能；而当网络设备基本实现能耗比例计算后，则变频技术对网络节能的贡献将更为显著。

4 降低网络流量的方法

所有基于休眠和协同的策略都是在假设网络流量需求相同的前提下，通过优化资源的使用实现节能降耗。但是如果网络设备能实现能耗比例计算，那么减少网络流量是降低能耗的最根本出发点。减少网络流量对降低能耗的作用体现于两方面：一方面，网络流量的降低缓解了网络设备升级的压力，延长了设备的使用寿命；另一方面，网络流量的降低减少了为转发这些流量所需要的运行能耗。

降低网络流量的可行性源自下述 2 个基本现实：一方面，人们对数据的访问通常呈现重尾分布的特征，如 Web 页面的访问[36]和 P2P 对象的访问[37,38]；另一方面，现有的网络传输模式基本是单播模式。这两者使得现有的内容分发机制产生了大量的重复数据传输。这些数据对象的重复传输不仅耗费了大量的带宽，迫使运营商不断进行扩容，同时也耗费了大量的数据传输能耗。

用户对数据对象的访问呈现重尾分布这一规律以及互联网用户对数据访问的异步性需求为解决网络中数据重复传输提供了依据。一种行之有效的方式是利用缓存技术。缓存系统不断地积累用户的访问行为信息，通过缓存替换算法来动态地优化缓存，逐步隐式地形成对用户访问规律的认知。无论是 Web、CDN 或 P2P，都大量地使用了缓存系统以降低数据对象的重复传输、提高用户的体验。另一种方式是改变网络的传输模式和优化数据传输的路由。为了将某个数据对象从源节点发送到一组接收者，基于单播的传输模式将产生大量的重复传输；而网络层多播则能避免数据在同一链路上的重复传输，但目前的网络层多播仅支持数据的同步推送。介于两者之间则是基于应用层多播的折中方案。但是，如果不对应用层多播的拓扑和路由加以优化，可能产生比单播传输更多的流量。优化的应用层多播产生的数据流量能介于单播模式和多播模式之间。而如果存在广播媒介，则能最大程度地降低数据传输的次数。下面对几种现有的和正处于

实验阶段的内容分发系统加以介绍，着重剖析其为减少网络数据传输所做的设计选择。

1) CDN

CDN 最初被用于缓解中心服务器的负载。例如，从 abc.com 下载的网页可能包含图片、视频、音频，或其他高带宽的多媒体文件。Abc.com 的中心服务器可以只提供最基本的 web 页面，而将浏览器重定向到 CDN 来获取页面中内嵌的多媒体内容。CDN 在网络部署成本和分发性能之间存在折中。早期的 CDN 是一种中心化的模式，即仅在核心的运营商网络部署服务器。然而统计显示，只有 50%的流量是源自互联网最大的 35 个网络运营商的[39]，这表明大量的流量依然要穿越互联网。据统计，互联网端到端的路径平均要经过 15 个路由器[30]，因而中心化的模式只能部分缓解流量压力。高度分布化的 CDN 则大幅提高了服务器覆盖的广度，将热门资源通过 CDN 网络直接推送到边缘网络，实现一次传输、存储，多次访问，避免了大量的网络传输。当然，高度分布化的 CDN 也对系统的安全性、管理、扩展性和内容同步带来了挑战。

2) P2P

早期的 P2P 技术着重于降低内容提供商的分发代价，并不关心对运营商网络带来的流量冲击。作为一种应用层多播，P2P 网络拓扑通常与底层物理网络失配，使得 P2P 流量通常以非优化的方式传输，产生了大量的 P2P 流量[40~42]。但通过构建位置感知的 P2P 拓扑（使应用层拓扑与底层物理网络拓扑相匹配）、在网络的边缘缓存 P2P 流量、改进数据调度算法等手段，可以大幅提高 P2P 节点从本地获取内容的概率，从而大幅降低 P2P 引发的网络流量[40~42]。目前，中国通信标准化协会的 TC1 的 WG4 工作组和 IETF 的 ALTO 工作组都在制定基于承载网感知的 P2P 流量优化标准[43,44]，旨在提高应用层拓扑与物理网络拓扑的匹配度，大幅降低 P2P 传输对核心网的流量冲击。IETF 于 2010 年还成立了 DECADE[45] 工作组，试图在网络中提供公共的内置缓存，将 P2P 应用的控制层和数据层分开，通过开放的协议允许应用可以自主使用和管理缓存，形成一个可管可控的内容分发平台，解决 P2P 缓存的可扩展性问题。

3) CCN/NDN

以内容为中心的网络[46]则将缓存从应用层拓展到了网络层，对内容进行命名，并依据内容标识进行寻址和路由。目前，基于这一思路对互联网进行改革的方案已经于 2010 年得到了美国自然科学基金委的支持。该方案将内容的解析和路由 2 个逻辑上独立的概念在物理上合为一体，在解析的同时完成路由。从单次传输来看，NDN 实现了网络层的任播(anycast)。内容获取的路径最优性由路由层保证，而不像现在的 CDN 系统或 P2P 系统一样依赖于中间的解析体系来优化定位。更重要的是，由于目前的网络层多播缺乏路由器缓存内容的支持，只能实现同步多播，无法满足用户对内容的异步访问需求，而由于在路由器引入了缓存，从群体效应来看，NDN 实现了网络层的异步多播。这是网络体系结构在网络传输模式上的一大革新。通过设计合理的缓存协同机制和替换算法，可以大幅降低网络的流量传输。

4) 双结构互联网

李幼平院士提出的双结构互联网认为，为了有效利用人类对内容访问呈现重尾分布规律这一特征，网络需要引入内容存储库，用于缓存内容。库可以存在于终端节点、边缘网络以及核心网路由器。双结构互联网引入了广播的传输模式。与缓存对内容访问规律的利用方式不同，广播显式地利用了重尾分布这一统计知识，将用户经常访问的资源经卫星通过广播媒质定期推送到用户终端。终端通过对用户历史访问行为挖掘得到的本体代码对广播内容进行个性化过滤。与 CCN 类似，双结构互联网也支持网络层的缓存，对于终端未缓存的内容，通过正常的方式请求，一旦路径上的某个中间节点能够满足请求，则通过该节点来服务该请求。但与 CCN 不同，双结构互联网建议不改变现有的 IP 网络基础设施，而是利用 IP 协议的选项字段用于标识用户所需内容的本体代码。双结构互联网依赖于现有的解析体系，如 DNS、ALTO[43]等，来完成资源的优化定位，不一定能保证单个请求的路由最优性。但由于引入了缓存，从群体效应上看，双结构互联网也实现了异步多播的功能。因此，双结构互联网分别通过显式和隐式 2 种方式利用了人类对内容访问呈现重尾分布规律这一特征，并充分利用了广播和多播的传输模式，最大程度地降低了热门内容分发的数据流量。

5　结束语

5.1　其他网络节能技术

数据中心是许多大规模网络应用的后台支撑系统，也是互联网的一大能耗产业。针对数据中心

目前提出了许多节能解决方案，其中较为著名的是Google的绿色倡议[47]，它为绿色数据中心提供了一整套的解决方案。针对数据中心的节能技术主要包括利用虚拟化实现服务器按需配置、利用数据调度算法降低数据传输量、以及复用现有系统来实现分布化的数据中心等[48~56]。此外，对数据中心的运营者来说，降低总的能源开支是它要追求的一个目标。对那些在不同地区建有多个数据中心的运营者，可以依据电价的时空波动性设计高效的实时调度算法，最小化电源费用开支[23,57]。

对传统的传输层和应用层的网络协议进行改进，使其具备能源感知功能，是实现节能的又一研究方向。在TCP层，已经提出了多种节能策略，如功能转移、机会性休眠、改变TCP定时器的粒度、减少重传次数等[17]。在应用层，也可以让协议增加能源感知功能。例如，文献[58]对BitTorrent协议进行了修改，使会话参与节点在没有数据传输发生时能进入休眠，同时又不从邻居节点的邻居列表中删除。

5.2 互联网节能的技术途径总结

为了实现互联网的节能所采用的技术途径大致可分为以下4个方面。

1) 在硬件和设备层面，需要支持变频、休眠和远程唤醒机制。

2) 在传输协议层面，通过流量感知，自动实现链路变频、流量聚合和节点/链路休眠。

3) 在应用层面，可以通过网络存在性代理、控制数据通信量或流量特征、提高应用的位置感知能力从而降低路由开销等方式来降低网络能耗或为节能创造前提条件。

4) 在宏观层面，可以通过减少网络流量来降低网络能耗。一方面，可以显式或隐式地利用人类对内容访问呈现重尾分布的规律，利用主动推送或缓存等机制来降低数据的重复传输；另一方面可以改变单播的传输方式，以多播或广播等传输方式来降低数据传输量。

5.3 研究方向展望

降低互联网的能耗已经成为了研究界和工业界的热点。网络节能涉及硬件设备、传输协议和网络应用等多个层面。目前研究界已经提出了多种针对网络的节能技术，但这些技术距离实用还有漫长的路要走。

展望未来，网络节能将是未来几年的热点研究领域，但也面临着诸多挑战，主要包括以下几方面。

1) 需要大范围改造现有的网络硬件设备，使其在设备和系统层面逼近能耗比例计算的理念，为协议层和应用层的能源优化创造前提条件。

2) 不同网络环境的高能效计算问题的建模还有待进一步细化，在电源模型、应用层性能约束和能源优化器3方面都有广阔的研究空间。例如，网络环境下的能效最优化经常是NP-hard问题，如何设计高效的启发式算法依然有很广泛的研究前景；此外，现有的协同休眠机制大都依赖于集中式能源优化器，而在网络环境下，如何实现分布式的能源优化器更具现实价值。

3) 网络环境下的节能需要大量的标准化工作。例如，需要为不同厂商设备的电源状态提供标准化的访问和管理能力，如提供合适的MIB[59]；网络存在性服务器需要对代理的行为和操作进行标准化，以支持代理的广泛部署。

4) 改变现有的网络协议/应用程序与能源使用脱离的现状是实现网络节能的一个主要途径，但这是一项极为艰巨的任务，因为对现有协议的改造都可能导致意想不到的负面效应。如何有选择地改造网络协议/应用程序，在实现有效节能的同时尽量避免负面效应，是需要深入研究的问题。

5) 探索降低网络流量的新机制的可行性。互联网用户对内容的访问呈现异步多播的特征，内容分发机制是未来互联网的一大研究方向。需要深入比较网络层缓存和应用层缓存方案，更有效地支持异步多播的需求。

参考文献：

[1] CHRISTENSEN K, GUNARATNE C, NORDMAN B, *et al*. The next frontier for communications networks: power management[J]. Computer Communications, 2004, 27(18): 1758-1770.

[2] CHRISTENSEN K. Green networks: opportunities and challenges[EB/OL]. http://www.ieeelcn.org/prior/LCN34/2009_Keynote_Christensen.pdf, 2009.

[3] Advanced configuration & power interface(ACPI)[EB/OL]. http://www. acpi.info/spec.htm,2011.

[4] ENERGY STAR version 5.0 specification for computers[EB/OL]. http://www.energystar.gov/ia/partners/prod_development/revisions/downloads/computer/Version5.0_Computer_Spec.pdf/,2008.

[5] 中华人民共和国环境保护行业标准——环境标志产品技术要求微型计算机、显示器[S]. HJ/T 313-2006.
The Technical Requirement for Environmental Labeling Products-Microcomputers and Displays[S]. HJ/T 313-2006.

[6] DEMEESTER P, PICKAVET M. An inconvenient truth: energy-efficient future Internet[EB/OL]. http://www.future-internet.eu/ fileadmin/documents/consultation_meeting_31_Jan_08/Demeester_ Internet_of_the_Future_31_jan_08.pdf, 2008.

[7] The energy efficient internet project[EB/OL]. http://www.csee.usf.edu/~ christen/energy/mail.html,2012.

[8] The energy efficient digital networks project[EB/OL]. http://efficientnetworks.lbl.gov/,2012.

[9] BARROSO L A, HÖLZLE U. The case for energy-proportional computing[J]. IEEE Computer, 2007, 40(12): 33-37.

[10] BROWN D J, REAMS C. Toward energy efficient computing[J]. Communications of the ACM, 2010, 53(3): 50-58.

[11] ALBERS S. Energy-efficient algorithms[J]. Communications of the ACM, 2010, 53(5): 86-96.

[12] ECMA international standard organization TC38-TG4 (formerly TC32-TG21)[EB/OL]. http://www.ecma-international.org/ memento/ TC38-TG4.htm,2012.

[13] AGARWAL Y, HODGES S, CHANDRA R, *et al*. Somniloquy: augmenting network interfaces to reduce PC energy usage[A]. Proceedings of the USENIX NSDI 2009[C]. Boston, USA, 2009.

[14] JIMENO M, CHRISTENSEN K, NORDMAN B. A network connection proxy to enable hosts to sleep and save energy[A]. IEEE International Performance Computing and Communications Conference(IPCCC)[C]. Austin, USA, 2008.

[15] NORDMAN B, CHRISTENSEN K. Proxying: the next step in reducing IT energy use[J]. IEEE Computer, 2010, 43(1): 91-93.

[16] NEDEVSCHI S, CHANDRASHEKAR J, LIU J, *et al*. Skilled in the art of being idle: reducing energy waste in networked Proceedings of the USENIX NSDI[C]. Boston, USA, 2009.

[17] AKELLA S A, BALAN R K, BANSAL N. Protocols for Low-power[R]. Project Report, Camegie Mellon University, 2001.

[18] LIEBERMAN P. Wake-on-LAN technology[EB/OL]. http://www.liebsoft.com/pdfs/Wake_On_LAN.pdf, 2006.

[19] NORDMAN B, CHRISTENSEN K. Reducing the energy consumption of network devices[EB/OL]. http://www.csee.usf.edu/~christen/energy/pubs.html, 2005.

[20] GUNARATNE C, CHRISTENSEN K, NORDMAN B, *et al*. Reducing the energy consumption of Ethernet with adaptive link rate(ALR)[J]. IEEE Transactions on Computers, 2008, 57(4): 448-461.

[21] IEEE std 802.3az-2010[EB/OL]. http://www.ieee802.org/3/purchase/index.html, 2010.

[22] ROUGHAN M, GREENBERG A, KALMANEK C, *et al*. Experience in measuring backbone traffic variability: models, metrics, measurements and meaning[A]. ACM SIGCOMM Internet Measurement Workshop[C]. Marseille, France, 2002.

[23] QURESHI A, WEBER R, BALAKRISHNAN H, *et al*. Cutting the electric bill for internet-scale systems[A]. Proceedings of the ACM SIGCOMM[C]. Barcelona, Spain, 2009.

[24] CHIARAVIGLIO L, MELLIA M, NERI F. Energy-aware backbone networks: a case study[A]. Proceedings of the IEEE ICC 2009[C]. Dresden, Germany, 2009.

[25] GUPTA M, SINGH S. Greening of the internet[A]. Proceedings of the ACM SIGCOMM 2003[C]. Karlsruhe, Germany, 2003.

[26] NEDEVSCHI S, POPA L, IANNACCONE G, *et al*. Reducing network energy consumption via sleeping and rate-adaptation[A]. Proceedings of the USENIX NSDI 2008[C]. San Francisco, USA, 2008.

[27] CHIARAVIGLIO L, MELLIA M, NERI F. Reducing power consumption in backbone networks[A]. Proceedings of the IEEE C 2009[C]. Dresden, Germany, 2009.

[28] ZHANG G Q. Link power coordination for energy conservation in complex communication networks[J]. Europhysics Letters, 2010. 92,28001.

[29] CHABAREK J, SOMMERS J, BARFORD P, *et al*. Power awareness in network design and routing[A]. Proceedings of the IEEE INFOCOM 2008[C]. Phoenix, USA, 2008.

[30] TUCKER R. Modeling energy consumption in IP networks[EB/OL]. http://www.cisco.com/web/about/ac50/ac207/crc_new/events/assets/cgrs_energy_consumption_ip.pdf, 2008.

[31] CIANFRANI A, ERAMO V, LISTANTI M, *et al*. An energy saving routing algorithm for a green OSPF protocol[A]. Proceedings the IEEE INFOCOM 2010[C]. San Diego, CA, USA, 2010.

[32] WANG Y, KELLER E, BISKEBORN B, *et al*. Virtual routers on the move: live routers migration as a network management primitive[A]. Proceedings of the ACM SIGCOMM 2008[C]. Seattle, USA, 2008.

[33] ANDREWS M, ANTA A F, ZHANG L, *et al*. Routing for energy minimization in the speed scaling model[A]. Proceedings of the IEEE INFOCOM 2010[C]. San Diego, CA, USA, 2010.

[34] AWERBUCH B, AZAR Y. Buy-at-bulk network design[A]. Proceedings of the 38th Annual Symposium on Foundations of Computer Science[C]. Miami, 1997.

[35] VASIC N, KOSTIC D. Energy-aware traffic engineering[A]. Proceedings of 1st International Conference on Energy-efficient Computing and Networking[C]. Passau, Germany, 2010.

[36] BRESLAU L, CAO P, FAN L, *et al*. Web caching and zipf-like distributions: evidence and implications[A]. Proceedings of IEEE INFOCOM 1999[C]. New York, USA, 1999.

[37] HEFEEDA M, SALEH O. Traffic modeling and proportional partial caching for peer-to-peer systems[J]. IEEE/ACM Transactions on Networking, 2008, 16(6): 1447-1460.

[38] SALEH O, HEFEEDA M. Modeling and caching of peer-to-peer traffic[A]. Proceedings of International Conference on Network Protocols(ICNP'06)[C]. Santa Barbara, USA, 2006. 249-258.

[39] LEIGHTON T. Improving performance in the internet[J]. ACM Queue, 2008, 6 (6): 20-29.

[40] **唐明董**, **张国清**, **杨景等**. P2P **流量优化技术综述**[J]. **电信网技术**, 2009, (1): 1-7.
TANG M D, ZHANG G Q, YANG J, *et al*. A survey of P2P traffic

optimization techniques[J]. Telecommunications Network Technology, 2009, (1): 1-7.

[41] **张国强，唐明董，程苏琦等**. P2P **流量优化**[J]. **中国科学**, 2012, 42(1):1-19.
ZHANG G Q, TANG M D, CHENG S Q, *et al*. P2P traffic optimization[J]. Science in China series F, 2012, 42(1): 1-19.

[42] ZHANG G Q, CHENG S Q, ZHANG G Q. LANC: locality-aware network coding for better P2P traffic localization[J]. Computer Networks, 2011, 55: 1242-1256.

[43] ALIMI R, PENNO R, YANG Y. ALTO Protocol[S]. Draft-ietf-alto-protocol- 06.txt, 2010

[44] YD/T 2146-2010. **基于承载网感知的** P2P **流量优化技术总体技术要求**[S].
YD/T 2146-2010. Technical Framework for Carrier Network-aware P2P Traffic Optimization[S].

[45] DECADE working group[EB/OL]. http://datatracker.ietf.org/wg/decade, 2010

[46] JACOBSON V, SMETTERS D K, THORNTON J D, *et al*. Networking named content[A]. Proceedings of CoNEXT'09[C]. Rome, Italy, 2009.

[47] Google's green initiatives[EB/OL]. http://www.google.com/green/, 2011.

[48] KANT K. Data center evolution: a tutorial on state of art, issues, and challenges[J]. Computer Networks, 2009, 53(17): 2939-2965.

[49] VALANCIUS V, LAOUTARIS N, MASSOULIE L, *et al*. Greening the internet with nano data centers[A]. Proceedings of the International Conference on Emerging Networking Experiments and Technologies(CoNEXT'09)[C]. Rome, Italy, 2009.

[50] LIU L, WANG H, LIU X, *et al*. Greencloud: a new architecture for green data center[A]. International Conference on Autonomic Computing[C]. Barcelona, Spain, 2009.

[51] 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, 2009,39(1):68-73.

[52] MAHADEVAN P, SHARMA P, BANERJEE S, *et al*. Energy aware network operations[A]. IEEE Global Internet Symposium[C]. Riode Janeiro Brasil, 2009.

[53] VASIC N, BARISITS M, SALZGEBER V. Making cluster applications energy-aware[A]. International Conference on Autonomic Computing[C]. Barcelona, Spain, 2009.

[54] LEVERICH J, KOZYRAKIS C. On the energy efficiency of hadoop clusters[J]. ACM SIGOPS Operating Systems Review, 2010, 44(1): 61-65.

[55] FAN X, WEBER W D, BARROSO L A. Power provisioning for warehouse-sized computer[A]. Proceedings of the ACM International Symposium on Computer Architecture[C].San Diego, CA, USA, 2007.

[56] **邓莉，吴松，金海**. **虚拟化—绿化数据中心的有效途径**[J]. **中国计算机学会通讯**, 2010,6(3): 14-17.
DENG L, WU S, JIN H. Virtualization—an efficient approach to greening data centers[J]. Communications of the CCF, 2010,6(3): 14-17.

[57] RAO L, LIU X, XIE L, *et al*. Minimizing electricity cost: optimization of distributed internet data centers in a multi-electricity-market environment[A]. Proceedings of the IEEE INFOCOM 2010[C]. San Diego, CA, USA, 2010.

[58] BLACKBURN J, CHRISTENSEN K. A simulation study of a new green BitTorrent[A]. First International Workshop on Green Communications[C]. Dresden, Germany, 2009.

[59] BLANQUICET F, CHRISTENSEN K. Managing energy use in a network with a new SNMP power state MIB[A]. Proceedings of the IEEE Conference on Local Computer Networks(LCN2008)[C]. Montreal, Canada, 2008.

作者简介：

张国强（1980-），**男，江苏常州人，博士，南京师范大学副教授，主要研究方向为网络科学和未来网络。**

林森（1985-），**男，江苏盐城人，中国科学院计算技术研究所硕士生，主要研究方向为计算机网络。**

刘真（1977-），**女，江西南昌人，博士，北京交通大学计算机与信息技术学院副教授，主要研究方向为数据中心虚拟化和分布式系统。**

林涛（1978-），**男，河南焦作人，博士，中国科学院声学研究所副研究员，主要研究方向为互联网体系结构和未来网络。**

张国清（1965-），**男，浙江浦江人，中国科学院计算技术研究所硕士生导师，主要研究方向为网络科学。**

李幼平（1935-），**男，福建泉州人，中国工程院院士，主要研究方向为三网融合。**