

多媒体网络:从内容分发网络到未来互联网

尹 浩^{1),2)} 詹同宇¹⁾ 林 闯¹⁾

¹⁾(清华大学计算机科学与技术系 北京 100084)

²⁾(清华大学信息科学与技术国家实验室(筹) 北京 100084)

摘 要 近年来,随着网络技术的飞速发展,互联网流量视频化的趋势日益明显,视频应用使互联网的可扩展性、可演进性、服务质量和网络安全面临诸多挑战.内容分发网络通过与应用与网络间构建覆盖网,为解决网络视频应用面临的分发难题提供了有力的支持,也成为当前网络视频产业中的关键技术.但是随着视频应用在移动性、交互性、异构性以及可扩展性方面需求的不断增长,现有内容分发网络技术面临重大挑战,学术界与工业界都试图从互联网体系架构入手解决网络视频应用中的问题.文中系统总结了内容分发技术面临的挑战以及现有互联网体系结构研究的思路与核心技术,提出了面向未来网络的多媒体网络设计思路.

关键词 互联网体系架构;内容分发网络;多媒体网络;下一代互联网

中图法分类号 TP393 DOI号: 10.3724/SP.J.1016.2012.01120

Multimedia Networking: From Content Delivery Networks to Future Internet

YIN Hao^{1),2)} ZHAN Tong-Yu¹⁾ LIN Chuang¹⁾

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

²⁾(National Laboratory for Information Science and Technology (TNList), Tsinghua University, Beijing 100084)

Abstract As network technologies develop rapidly in recent years, more and more network traffic are taken by video applications, which introduces challenges to the scalability, evolutionary, QoS and security of existing Internet. By constructing an overlay network between applications and networks, CDN provides a strong support to solve the distribution problem faced by video application, and becomes a key technology in current video industry. But with the growing demand of video applications in terms of mobility, interaction, heterogeneity and scalability, existing CDN faces significant challenges, and both academia and industry are trying to solve the challenges by redesigning Internet architecture. This paper summarizes the challenges that CDN faced, the ideas and core technologies of the existing Internet architecture, and then puts forward the design ideas of multimedia networking for future networks.

Keywords Internet architecture; content delivery networks; multimedia networking; next generation Internet

收稿日期:2012-03-21;最终修改稿收到日期:2012-05-23. 本课题得到国家“九七三”重点基础研究发展规划项目基金(2012CB315800, 2011CB302601)、国家自然科学基金(61170290,60736012)资助. 尹 浩,男,1974年生,博士,副研究员,主要研究方向为未来网络体系架构、内容分发网络与多媒体通信. E-mail: h-yin@mail. tsinghua. edu. cn. 詹同宇,男,1986年生,硕士研究生,主要研究方向为服务节点部署、未来网络体系架构. 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、系统性能评价、安全分析和随机 Petri 网.

1 引言

近年来,随着网络技术的飞速发展,Internet 服务的内容从以静态网页、小图片为主转变为以视频等大文件为主。2005 年,随着 Micromedia 公司被 Adobe 公司收购,互联网视频进入了网页视频阶段,从而迎来了网络视频产业高速发展的黄金时期。电影、电视剧、各类视频日益增多地通过互联网播放。同时,一些重要的事件也开始在互联网上进行直播。调查显示,互联网当中视频流量的增长速度是其它应用流量增长速度的 10 倍以上。Cisco 的报告指出,到 2013 年,互联网流量的 90% 以上都会被视频应用所占据^{①②}。另一方面,视频应用的发展越来越迅速,新的应用和服务不断涌现。从最初普通的标清视频,到逐渐兴起的高清视频、3D 视频等,这些应用在带给用户更完美体验的同时,也对互联网的体系架构提出了挑战,这些挑战可以归纳为如下几点:

可扩展性。是指网络适应服务规模、视频质量以及用户服务体验等变化的能力。网络的视频化对网络的可扩展性提出了更高的要求。以广播电视系统为例,在中国一个热播电视剧的收看人数是 5000 万人,按照 DVD 品质视频的码率 1Mbps 来计算,所需要的带宽就达到 50 Tb,而中国电信和联通刚刚将带宽扩容至 100 Gb^③。因此视频应用对互联网的服务规模和扩展能力提出了巨大的挑战。如何以成本可控的方式实现可扩展,就成为了视频应用面临的一大挑战。

可演进性。是指网络系统架构适应新应用与技术的能力。网络技术的不断发展以及新应用的不断涌现,使得网络架构与关键技术越来越复杂。复杂的网络会提高网络应用开发与升级的时间周期和成本,同时也对网络架构与核心技术的演进性提出了更高的要求。互联网的体系架构设计应借鉴操作系统的设计思想,对底层和复杂技术利用抽象的方式隐藏复杂技术的细节,为上层应用与服务提供简单的应用接口,从而能够灵活地满足不同种类应用对网络服务的不同需求。

异构性。随着手机、平板电脑、无线传感器等设备以及物联网、移动网的发展,互联网的终端与网络本身发生了很大的变化,主要表现在网络动态性的显著增加。网络节点动态性的增强导致数据

传输路径频繁变换,严重破坏了上层应用服务的连续性,影响了互联网用户享受视频应用的用户体验。传统体系架构无法有效地适应网络的高动态性。首先,IP 地址既表达身份又表征位置的特性不能很好地支持移动 IP 的思想;其次,TCP 端到端的通信模式将服务连接与流量管理交由终端完成,这对能量与能力都有较大局限的终端是个很大的挑战。

服务质量。多媒体应用对服务质量非常敏感。Dobrian 等人^[1]的研究指出,在一场 90 min 的足球比赛的直播当中,如果缓冲比例提高 1%,就将导致用户的平均观看时间缩短至少 3 min。现有的网络协议在长距离传输、吞吐率、域间传输等方面存在不足,使得视频应用的服务质量无法得到保证,如 TCP 协议在远距离传输时吞吐率将严重下降。另一方面,Internet 的规模大,自治域多,传输过程中要经过多个自治域,而域间传输效率却很低;此外网络中出现服务节点失效、链路故障、配置管理错误等都有可能造成网络的不可预测的错误,从而无法确保视频应用对服务质量的要求。

网络安全。互联网的前身 ARPANET 是美国国防部为进行研究而建立的网络,其设计基于可信实体之间的通信,因此安全性并未作为一项重要问题被考虑。但随着互联网的普及,网络安全成为发展过程当中非常重要的问题,涉及到访问安全、传输机密性和完整性等问题。如拒绝服务(Denial of Service, DoS)攻击会导致网络的不可用或者性能下降,而在目前的网络架构下无法有效防止 DoS 攻击^[2],急需在体系结构层面解决。

内容分发网络(Content Delivery Network, CDN)技术通过在与网络间构建覆盖网,弥补了 IP 网络的诸多缺陷,为视频应用提供灵活高效的内容分发服务,被广泛应用于网络视频应用中。在本文中,将首先对内容分发网络的发展和面临的挑战进行分析,同时综述了国内外关于未来网络的研究,进而指出,面对视频应用不断发展的需要,内容分发网络技术将逐步发展为面向未来网络的多媒体网络;最后提出了该网络的设计思路。

① Cisco Visual Networking Index: Forecast and Methodology, 2009-2014, 2010

② Schonfeld E. Cisco: By 2013 Video Will Be 90 Percent of All Consumer IP Traffic and 64 Percent of Mobile, 2009

③ http://www.cq.xinhuanet.com/2012-03/16/c_111660413.htm

2 内容分发网络的兴起与挑战

为克服传统 IP 网络在服务视频应用当中的不足,内容分发网络 CDN 被提出,并迅速成为互联网体系架构中的重要组成部分.根据 IETF RFC 3466 的定义,CDN 通过传统 IP 网络上部署服务节点,并利用应用层协议将这些服务节点联结构成应用层覆盖网络,为用户提供内容分发服务.在过去的十年里,内容分发网络技术不断发展,并逐渐融入对等网络和云计算等技术,以满足新的技术和应用对内容分发提出的需求.现在,内容分发网络已经成为了视频内容分发的重要手段.表 1 给出了一些主要的视频网站所采用的传输协议、编码技术和 CDN 供应商.从表中可以看出,这些视频网站虽然采用了不同的协议和编码

表 1 主要视频网站使用的传输协议、编码技术和 CDN 供应商

	传输协议	CODEC	CDN 提供商
ABC. go. com	Move Network	ON2 VP7	Limelight
www. NBC. com	Flash	ON2 VP6	Limelight
www. NFL. com	Flash	H. 263	nplayer
www. cbs. com	Flash	ON2 VP6	Akamai
FOX. VOD	Move Network	ON2 VP7	Limelight
www. hulu. com	Flash	ON2 VP6	Akamai
www. youtube. com	Flash	H. 263	Google
video. yahoo. com	Flash	H. 263	Yahoo
mypacstv. com	Flash	H. 263	Limelight
www. brightcove. com	Flash	ON2 VP6	Limelight
www. nba. com	RTSP	WMV9	Akamai
www. mlb. com	RTSP	WMV9	Akamai
Joost	P2P	H. 264	Level3
PPLive	P2P	H. 263	P2P

技术,但大都租用或自建了 CDN 来提供服务.而 CDN 的大量部署也使得互联网的拓扑结构和网络协议模型发生了变化.

2.1 互联网拓扑结构变化

传统的互联网拓扑,呈现的是层次化的树状结构.但随着网络的发展,网间传输瓶颈问题日益突出.在中国,运营商之间的竞争所导致的域间传输延迟过大问题尤为明显.图 1 为 2010 年 5 月份对中国主要省份各运营商之间的延迟状况测量结果,图中不同线段的颜色按照各省的不同运营商之间延迟累积分布函数的 85% 划分.从图中看出,各地区不同运营商之间的平均延迟大都在 90 ms 以上.为了解决网间传输瓶颈等问题,同时提高网络的灵活性与可扩展性.近年来,一些内容提供商如 Google、Microsoft、Facebook 等,都建立了自己的覆盖网.网络供应商开始提供 CDN 或云存储服务与应用,如 Google 在 2006 年和 2007 年分别投入了 19 亿和 24 亿资金用于部署数据中心,到 2008 年为止在全球各地部署了 36 个数据中心^①.由此带来的是网络流量趋势的转移.近几年大部分跨区域网络流量的增长直接发生在大型内容供应商、CDNs 以及用户接入网络之间,而不是传统的几大网络供应商之间,如表 2^[9].部分内容提供商和 CDN 所占据的流量已经与一级运营商相当,同时许多 CDN (如 Akamai、LimeLight 等)和内容提供商(如 Google、Microsoft、Facebook 等)的网络都直接与一级运营商互联,如图 2 所示,互联网的拓扑结构开始由层次化向扁平化转变,CDN、内容提供商与运营商之间的网络呈现彼此交织的状态.

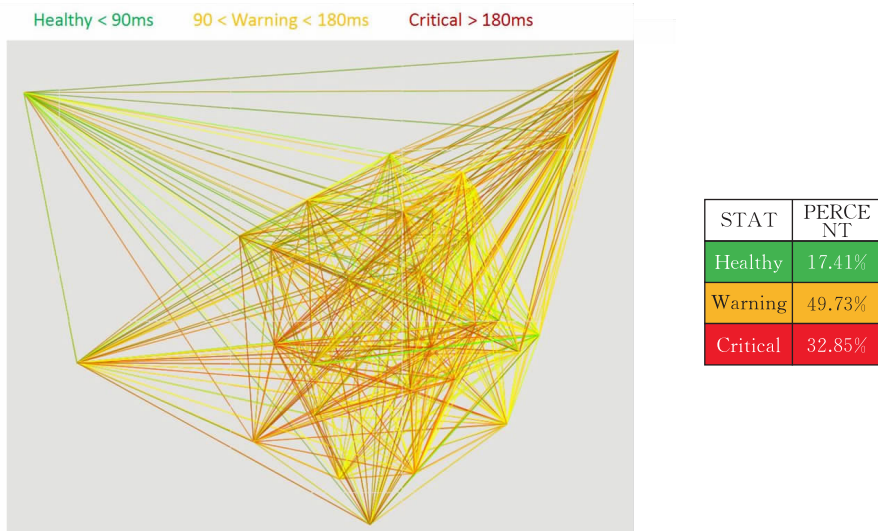


图 1 中国互联网运营商之间的平均延迟(2010 年 5 月)

① Map of all Google data center locations. <http://royal.pingdom.com/2008/04/11/map-of-all-google-data-center-locations/>

表 2 网络流量分布^[3]

排名	提供商	百分比/%
1	Google	5.03
2	ISP A	1.78
3	Limelight	1.52
4	Akamai	1.16
5	Microsoft	0.94
6	Carpathia Hosting	0.82
7	ISP G	0.77
8	Lease Web	0.74
9	ISP C	0.73
10	ISP B	0.70

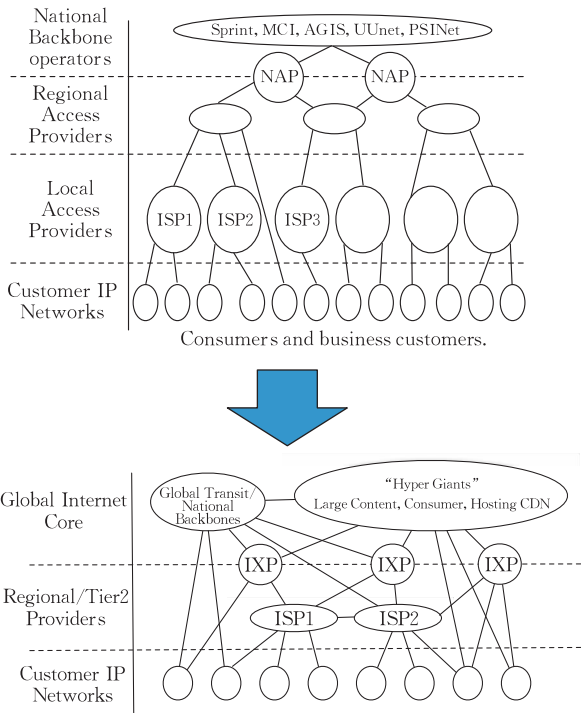


图 2 网络拓扑变化趋势^[3]

2.2 网络协议栈的变化

随着 CDN 的兴起与互联网拓扑结构的变化，互联网的协议栈也在发生改变。Akhshabi 等人^[4]对互联网协议栈进行了深入的研究，指出网络内部协议之间的竞争关系，使得互联网协议栈呈现两端粗、中间细的沙漏结构。论文通过对协议栈演化过程的模拟，研究协议栈各层的宽度，结果如图 3 所示。可以看出，在一个具有 10 层结构的协议栈中，沙漏的细腰往往出现在中间层，如第 5、6 层。论文同时对影响协议栈沙漏结构的因素进行了分析，并进一步指出，通过改变参数，可以使沙漏的腰向上移动，并且在腰的位置越往上就会变得越粗。另外，在某些情况下，沙漏甚至会出现两个或以上的腰。

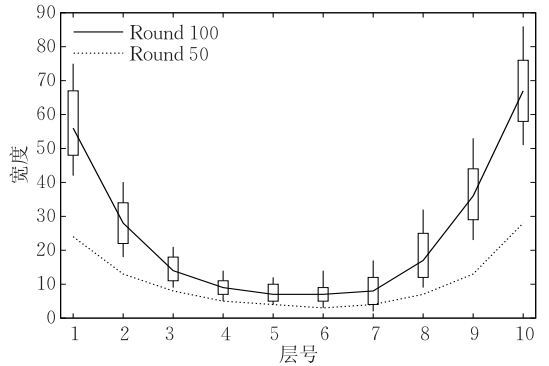


图 3 互联网协议栈沙漏结构各层宽度^[4]

在过去的十年里，多项针对互联网结构的研究表明^[5-8]，IP 网络的结构简洁而清晰，并且能够有效地支持上下层协议的发展，从而使得它在与其它协议的竞争中存活下来，成为了互联网协议栈的细腰。然而，随着 CDN 等技术的发展，大量的技术着手在应用层解决互联网的问题。这是由于现有的应用与技术大都依赖状态相关的技术来解决问题，如构建 P2P 的应用层组播关系等，而这些状态相关的技术如果应用在网络层的话，将大大增加网络层的工作负载，使得相关技术无法大规模应用，IP 层组播至今未得到大规模应用就是一个典型案例。因此在应用层解决某些问题比在网络层解决要更简单和有效，多项研究进一步指出，协议栈的“细腰”正在向上移动。

Cisco 的研究表明随着 NAT 的广泛使用和防火墙的设置，TCP 和 UDP 正在成为互联网新的细腰。NAT 的出现使得原有互联网结构中 IP 地址不足的问题得到了缓解，但同时也使得网络传输更加依赖 TCP 和 UDP。而另一方面，大多数的防火墙也配置为仅仅允许 TCP 和 UDP 的传输。互联网逐渐呈现出新的变化，如图 4 所示，从中可以看出，TCP 和 UDP 正在成为协议栈沙漏结构中新的细腰。

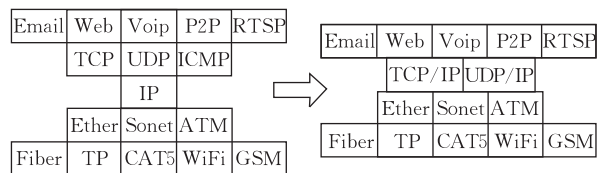


图 4 互联网沙漏结构变化

而加利福尼亚大学伯克利分校的 Popa 等人^[9]的研究指出，随着 CDN、Web Cache 等基于 HTTP 的基础设施被广泛应用，HTTP 已经成为了使用最广泛的互联网传输协议。同时，HTTP 的使用也扩展到了视频应用领域，目前大量的视频网站如 YouTube、

优酷等,都使用 HTTP 协议传输内容.相比于流式协议,分段的 HTTP 协议具有可扩展性强、可靠性高、适用性强等特点,更适用于大文件传输.同时,由于 HTTP 服务器的广泛部署,以及 HTTP 协议具有更好的兼容性,因此以分段 HTTP 传输为主的视频分发方式已经成为了主流. MoveNetworks 和 Swarmcast 等公司首先开始使用 HTTP Chunking 技术,保证了 HTTP 协议适用于流媒体的传输. 研究指出,HTTP 在互联网协议中占据着主导地位,并成为互联网协议栈新的腰,如图 5 所示.

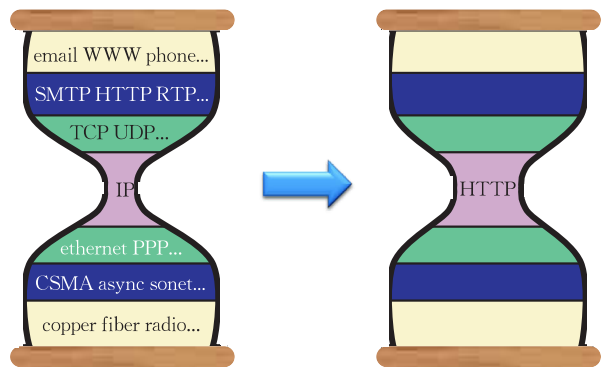


图 5 协议栈的细腰从 IP 变为 HTTP

作者认为,CDN 的广泛部署使得更多问题在应用层解决,而 HTTP 相比于其它应用层协议,具有部署成本低、可扩展性强、可靠性高等特点.同时,HTTP 是一个以内容为中心的协议,更加符合未来网络的发展趋势.因此,HTTP 在网络中将发挥越来越重要的作用,也很有可能成为协议栈沙漏结构的细腰.

2.3 CDN 面临的挑战

从互联网的发展趋势来看,由于互联网的传输内容发生改变,视频正逐渐成为互联网流量的主导,未来的互联网将被视频流量占据,互联网需要新的技术与架构来支撑视频应用. CDN 通过与应用与网络之间构建覆盖网络,为网络应用提供灵活高效的服务.但随着互联网流量的视频化日趋明显,CDN 在面对大规模视频服务时也存在诸多缺陷.总结起来有如下几点:(1)可扩展问题. CDN 技术为视频应用提供了有质量保证的服务,但 CDN 部署成本的增长与服务规模 and 用户数量的增长呈线性关系^[10],同时 CDN 之间很难有效合作,这使得服务的扩展能力受到很大限制.(2)智能化问题. 随着网络技术的发展,用户对服务体验和服务质量的要求越来越高,同时用户环境与应用模式也在不断更新,传统 CDN 的关键技术如请求路由技术^[11-12],只进行

一次定位,无法满足现有视频应用对服务质量的需求,未来网络需要更智能化的控制.同时,未来网络需要考虑多媒体应用的技术特点,如文献[13]就提出在多媒体应用中内容的自适应处理非常重要,这要求面向视频的多媒体网络除了需具备通信与缓存功能之外,还需要加入多媒体处理的计算功能.(3)抽象问题. 现有的网络体系架构对底层细节的抽象不够,导致无法有效应对网络和终端异构与快速开发的需要,未来网络应实现更好的抽象,屏蔽底层的异构网络和复杂的技术细节,为上层应用与服务提供简单的应用接口.

伴随着网络拓扑结构的改变,互联网协议栈本身也在发生改变,沙漏结构的细腰正在不断地向上移动.细腰的上提使得沙漏上下层的技术与协议有机会进一步创新,为网络的发展带来了更多机遇.

总体来说,现有网络解决挑战性问题,基本都是采取打补丁的方式,即依靠增加复杂的机制来弥补现有体系架构的不足.但是这种方法的弊端也日益凸显,使得网络结构更加臃肿,问题的解决与新应用的开发愈发复杂.借鉴编程语言的设计思想,编程语言通过不断的抽象,从而使我们摆脱了机器语言、汇编语言带来的对底层复杂系统的依赖性,让我们可以利用面向对象的方法去编写程序,更多地关注需要解决的实际问题,而不是花很多精力解决哪个数据应该存储在哪个寄存器上这些问题上,进而实现了对复杂性的屏蔽.因此,网络设计应该遵循简单化的原则,通过抽象将复杂技术和实现屏蔽,提高网络的灵活性和可演进性.为了解决 CDN 技术面临的问题,应从体系结构角度入手,对互联网进行反思与设计.

3 互联网体系架构

从 2005 年开始,国外主要发达国家都大力投入新一代网络体系架构的研究,如美国的 GENI 和 FIND 计划,欧洲的 4WARD、FIRE 计划,日本的 AKARI、JGN2+ 计划等.在工业界,Cisco、Juniper、Google、NEC、Microsoft、HP 等国际一流设备商,及 AT&T、法国电信等著名运营商都在该领域投入重要力量加强创新研究.美国自然科学基金于 2010 年提出 FIA 计划,希望在 FIND 计划的成果基础上,将网络架构研究汇聚成四大类架构进行深入研究.其中,NDN(Named Data Networking)是由加利福尼亚大学洛杉矶分校主要负责的研究项目^[14],该项

目基于 CCN(Content-Centric Network)的概念,认为互联网正从端到端的包传输转向内容为中心的模型. NDN 模型中,内容是以数据命名而非位置,从而将网络关注的焦点从“where”变为“what”. NDN 从加利福尼亚大学伯克利分校的 Scott Shenker 教授等提出的 DONA(Data-Oriented Network Architecture)体系结构^[7]出发,它采用名字路由,通过路由器来缓存内容,从而提高数据传输效率和内容的检索效率. NEBULA 项目^①由宾夕法尼亚大学等几所院校负责,针对云系统在可用性、安全性和一致性方面存在的缺陷进行研究. NEBULA 的架构中包含一个可靠、高速的数据中心,连接数据中心和核心路由器的平行路径,以及一系列的安全访问、传输机制、链接认证机制,从而改善了云系统的安全性、可用性和一致性. MobilityFirst^[15]项目针对移动互联网进行研究. 该项目使用延时可容忍网络 DTN(Delay-Tolerant Networking)提供鲁棒性,再结合自认证公

钥的使用,解决移动网络的可信问题. 同时把移动作为第一属性,使得环境和位置感知服务自然地适合于该网络,该项目侧重在移动、可扩展性和公平使用网络资源之间的权衡,实现移动终端间的有效通信. 卡内基梅隆大学牵头的 XIA(eXpressive Internet Architecture)项目^[16]针对网络应用的多样化、可信通信的需求,拟设计一个更通用的架构来支持不同的通信模式,包括基于内容的通信或基于主机之间的通信等,并从网络架构的角度系统解决网络的安全性问题. 通过对 FIA 各个项目进行总结和对比,从互联网的几个主要挑战,如可扩展性、可演进性、服务质量和网络安全等方面,将它们的主要优化手段进行归纳,结果如表 3 所示. 此外,在 GENI 项目的基础上,最近工业界也推出了可编程网络(Software Defined Network, SDN)的概念和系统架构,通过对互联网的转发抽象、分布状态抽象和构建网络操作系统,实现了对网络的简化与抽象.

表 3 各研究项目主要优化方案

	可扩展性	可演进性	服务质量	网络安全
NDN	分布式存储、路由可扩展	名字与地址分离、分层命名	路由存储、内容命名	分布式架构、命名安全
MobilityFirst	分布式存储	名字与地址分离	路由存储、类 DTN 技术	自认证公钥
NEBULA	云系统	抽象与虚拟化	高速骨干网、云计算	访问与传输安全、连接建立认证
XIA	分布式存储	普适的细腰结构、回溯机制	类 DTN 技术	自认证标识、可信通信

以上这一切都表明,未来网络领域的研究与产业化推进工作已经紧锣密鼓地展开了,随着一系列新的架构和技术被提出,一轮新的网络技术升级换代和产业变革即将到来.

4 多媒体网络体系架构

互联网拓扑结构的扁平化,与协议栈细腰的上移,为网络发展带来了新的机遇. 参考国外关于未来网络体系架构的主要研究工作,在解决互联网面临的挑战时,主要的手段包括 3 个方面:(1)通过抽象的手段,如名字与地址分离、可编程路由等,实现网络的可演进性.(2)采用分布式的系统架构提高系统的可扩展性,为网络加入智能化控制,通过路由存储与存储感知的选路协议(类 DTN 技术),提高服务质量.(3)在接入设备中加入存储和视频处理等计算功能,提高数据的传输效率和面向终端用户的自适应能力. 借鉴这些设计经验,我们提出了新一代多媒体网络的设计思想与关键技术.

4.1 体系架构设计原则

多媒体网络体系架构的设计,包括以下几个关键思想.

(1) 简单化.

网络技术的不断发展使得网络架构与关键技术越来越复杂. 复杂的系统会给未来网络的设计带来多方面的问题. 因此,新一代多媒体网络的设计应遵循简单化原则.

实现网络的简单化主要有如下几种方法:①通过抽象,提高各层次的灵活性与可演进性. 可编程网络的相关研究^[17]提出,网络的控制层需要实现三种抽象:转发抽象、分布状态抽象和网络操作系统. 进行转发抽象的目的是构建灵活的转发模型,从而实现对不同网络之间内容传输和转发的兼容. 而状态分布抽象是将控制层与底层的状态分布屏蔽,使得控制程序无需应付频繁变化的分布状态. 网络操作系统是运行在网络的服务器和控制器的分布式系统,目的是为上层提供完整的网络视图,实现网络不同转发元素之间的通信. ②通过构建覆盖网络,屏蔽底层细节. 借鉴操作系统的设计思想,隐藏底层复杂技术的细节,为上层应用与服务提供简单的应用接口. ③松耦合. 现有互联网的分层结构实现了对网络的部分抽象,这种抽象构成了现在互联网的体

① NEBULA Project. <http://nebula.cis.upenn.edu>.

系架构。但是,互联网的各层次接口之间呈现紧耦合的关系,它们之间相互依赖的关系导致互联网的灵活性与可演进性变差。未来网络的设计应本着松耦合的思想,降低各组成部分之间的依赖性,提供更灵活和可演进的解决方案。

进行抽象与松耦合的重要前提是实现名字和地址分离。其中的两个重要问题是内容命名与内容定位。其中,内容命名部分的主要策略包括平面式命名技术^[18-19]和分级式命名技术^[14,20]两种。其中,分级式的命名技术将内容命名成类似 Web URL 式的标识符(如 `www.tsinghua.edu.cn/publish/th/index.html`)。这种命名方式能够兼容现有的基于 URL 的网络,同时层次化的特性能够通过对内容的聚合来缓解路由压力。但同时,一旦内容所有者发生改变,将导致命名与内容的不匹配。而另一种平面命名策略将所有内容统一命名为其公共密钥的哈希值,如 128 位 GUID(Globally Unique Identifier)^[15]。这种平面命名策略保证了内容命名的持久性,但同时这种命名方式无法表示具体的语义,因此需要提供额外的解析服务。Ghods 等人^[21]对这两种命名技术在安全性、灵活性和可扩展性方面的优劣进行了详细的分析,指出平面式命名方式能够更好地抵御 DoS 攻击,同时拥有更好的灵活性,但可扩展性稍差。

而在内容定位方面,也有基于哈希的分布式方式^[15]和基于服务抽象^[16]的集中方式。采用分布式的策略,是利用一致性哈希将存储内容的 GUID 和存储位置信息映射到路由器中。当用户请求内容时,首先从路由查找该内容的 GUID 和对应的哈希函数,查找该内容的存储位置,之后将这些位置中距离该用户最近的地址发给用户。如 YouTube^[22]采用唯一的视频 ID 和 DNS 服务器间的映射函数,将视频 ID 空间分为 192 个部分,再利用 DNS 服务器查找到距离用户位置较近的服务器。而集中式策略的核心思想是视频搜索服务和 DNS 服务的整合抽象。如 XIA 中,用户向 Web 服务器明确描述连接意图,通过 Web 服务器搜索可获得的内容标识符,然后以该内容标识符为目的地址建立连接并享受服务。

(2) 智能化。

视频应用对服务质量和用户体验提出了更高的要求,这就要求网络在路由选择、负载均衡和任务调度等方面更加智能。同时,网络应用的多样化以及新应用的不断涌现,对互联网的可演进性提出了更高的挑战,这就要求网络的体系架构应具备上下文感

知能力,以适应应用的发展与多样化。终端设备的多样性和网络的异构性对网络的自适应能力提出了较高的要求,需要在进行用户感知和网络感知的前提下,实现不同尺度网络资源的虚拟聚合与协同工作。

由以上几点可以得出,网络需要变得更加智能与高效。为了实现网络智能化,需要进行充分的感知,包括对用户的感知、对内容的感知和对网络的感知。对用户和内容的感知是网络提供自适应能力的前提,而网络感知是为实施控制层的负载均衡、动态路由以及实现网络虚拟化提供必要的网络信息的重要手段,同时也是为多尺度网络资源按需聚合提供依据的依据,并可以对网络状况和服务状况进行实时监测,提高系统的可靠性。

为了进行高效准确的网络感知,可以采用网络坐标以及数据拟合等技术,从而有效地解决网络感知中,获取全面的网络信息困难以及测量成本高等问题。网络坐标采用将无形的网络空间映射到欧几里得空间当中^[23],将网络映射到坐标系中的方法,从而可以通过在坐标系中进行简单的数学计算,得到复杂网络空间中的数据。网络坐标的生成算法有很多种,包括 GNP 算法^[23]、动态三角算法^[24]、Vivadi 算法^[25]等。此外,IPlane^[26-27]采用的方法是在进行网络测量的基础上,借助预测与拟合等方式,来得到较完整的信息。

网络与应用接口的设计是实现网络智能化的重要手段。P4P/ALTO 等提出了全新的网络与应用接口的概念^[28],来提高对网络资源的使用效率与服务质量。

(3) 安全性。

安全性是指网络通信应具备有效性、完整性、真实性和机密性。在这些问题中,完整性、真实性和机密性等可以依靠加密与认证等手段实现。而有效性却面临着 DoS 攻击的挑战,在传统网络体系架构中难以解决。为了解决 DoS 的攻击,文献^[29-30]提出了一种称作“shut-up-message”(SUM)的思路,使得被攻击的服务器可以通过这个信息令攻击服务器停止向自己发包,进而抵御 DoS 攻击。同时,可以借鉴内容中心网络的安全机制,采用分布式的缓存机制,以及实现名字和地址分离的命名策略,从根本上解决 DoS 攻击的问题。

此外,NEBULA 也提出了自己的架构,来解决云系统中网络安全的问题,通过实现多种访问机制和复制技术,以及策略为基础的访问控制和连接的安全认证,来提高网络通信的机密性。

(4) 策略与政策.

随着网络连接趋于复杂,网络资源的协商机制也愈发复杂^[31]. Clark 等人^[32]分析互联网连接趋势后认为:网络之间的 Peer 连接将不再仅仅是收益中立的,网络使用相关的付费模式将占主导,按内容收费的模式或将出现.这使得多媒体网络中资源有效管理必不可少.由于多媒体内容的传输对成本控制要求非常高,新一代多媒体网络的设计应该遵循成本最低的管理原则,通过高效的、带成本控制的资源管理和协商方法,使得网络中各个利益相关者达到

成本最小化、利益最大化.

4.2 体系架构设计

基于以上考虑,借鉴操作系统的设计思想,设计新一代多媒体网络的体系架构.如图 6 所示,类比操作系统的三层空间,新一代多媒体网络包含三层环境、三种感知和三层空间.三层环境包括服务环境、内容中心网络环境、网络环境;三种感知包括用户上下文感知、内容感知和网络感知;三层空间分别是数据空间、名字空间和服务空间.

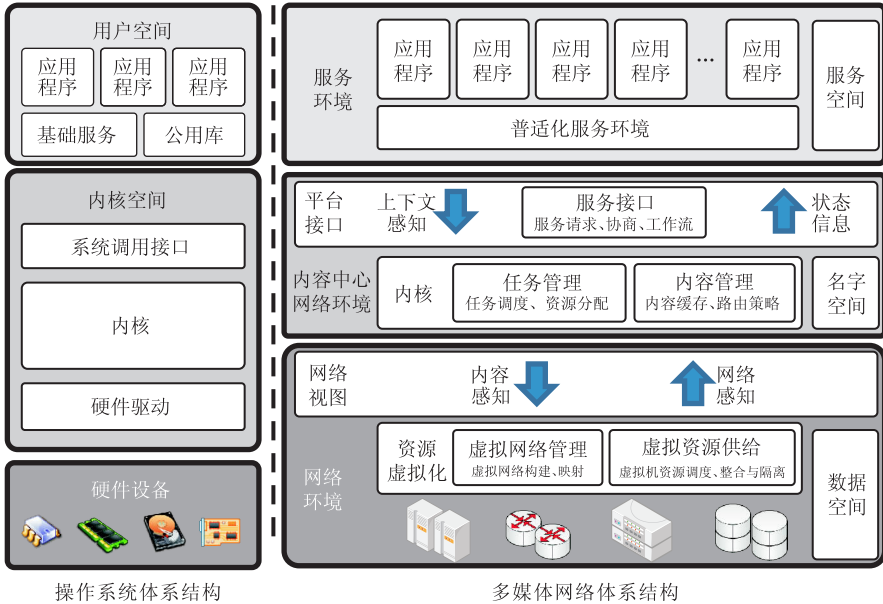


图 6 多媒体网络设计

4.2.1 数据空间

数据空间主要解决通信网络的异构问题与可扩展能力问题. 视频应用需要解决异构网络环境中海量内容的处理与分发问题. 为了解决这个问题,行之有效的办法是通过实现网络的虚拟化,并为网络加入可编程能力,提高网络的可扩展性与应对异构网络的能力. 如文献^[33]提供了一种新的解决方案,使用小路(Pathlet)和虚拟节点(Vnode)来建立独立于物理拓扑图上的虚拟拓扑图,通过抽象的手段,把路径地址与路径标识相分离,使得原来提供单一路径转发信息功能的数据平面,能够通过可编程的方式根据 AS(Autonomous System)自身的需求来动态地组合路径,以达到路径动态配置的目的. 在 Pathlet 中,每个 AS 拥有本地的策略约束,这些策略可以使用少量的小路表示,这使得转发表较小,而可以使用的路径较多,保证了网络的可扩展性. 同时,每个 AS 可以根据自己的情况使用不同类型的传输策

略,解决了异构网络的通信问题.

而延迟容忍网络技术^[34-35],则采用网络内存储和重传的消息交换、名字迟绑定与容忍网络分割路由等技术,构建了具有构建长延时、间歇性连接、不对称数据速率、低信噪比等特点的覆盖网,为不同类型网络,如空间网络、海底网络等性能差异较大的网络提供互操作通信.

在网络感知方面,可以借鉴 Onix^[17]的设计思想,如图 7. Onix 采用网络信息库(Network Information Base, NIB)来记录网络拓扑与性能信息,为上层应用提供网络视图. 在 NIB 中,所有的网络实体采用统一的命名格式,并进行了封装,从而为上层的控制逻辑提供统一的接口. 在网络测量中,可以采用网络坐标的方法来获取更加全面的网络信息,也可以借鉴 iPlane 平台的技术,使用数据拟合的方式,可以从有限的直接测量结果中得到更加完整的网络拓扑、时延信息、丢包率、链路能力等. 此外,还

可以采用被动测量的方法,从网络设备上实时地静态收集各种流量信息和统计日志,并进一步对收集到的日志信息进行深入挖掘和分析,可以获得网络更加全面的属性,如流量矩阵、带宽占用等性能影响因素。

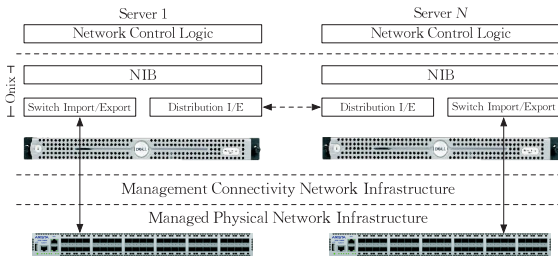


图 7 Onix 体系架构^[17]

4.2.2 名字空间

名字空间通过抽象手段,解决了网络移动性问题和终端的异构性问题。传统的互联网架构基于通信目标设计,关注的是主机与主机间的通信,而现在在互联网的主要需求已经上升到关注内容的订阅与获取为主,因此互联网应从通信网向信息中心网络转变。参考编程语言的发展,在早期汇编语言中,程序员需要自行对寄存器进行操作,而在高级语言中,程序员无需关心变量的具体存储位置,可以通过变量名和指针等方式直接对变量进行操作。因此,名字空间需实现名字与地址分离,令使用者无需关心内容存放的具体位置,实现更灵活和高效的内容定位与分发。

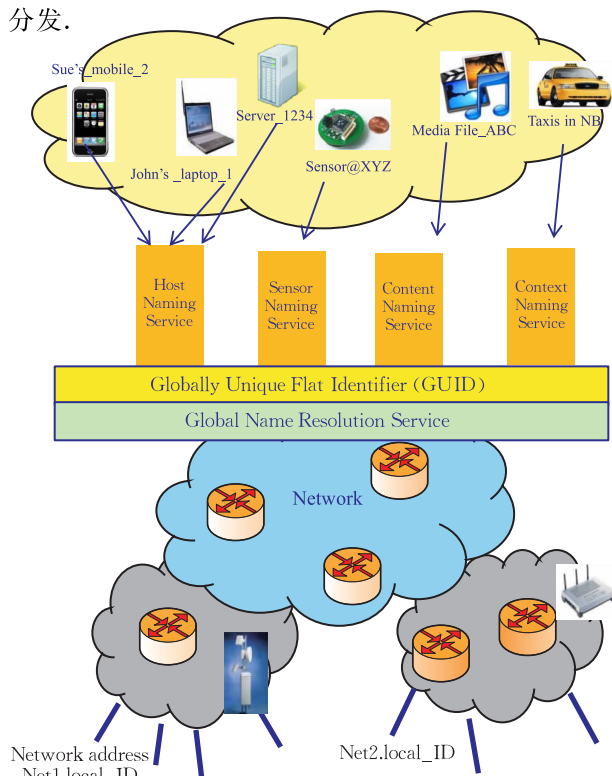


图 8 MobilityFirst 架构^[15]

名字空间的设计可以借鉴 MobilityFirst^[15] 的设计策略,采用平面化的内容命名与分层的内容定位技术,使用全局统一的 GUID 来命名每一个主机或者网络。GUID 作为主机或网络的全局唯一标识,可以自动生成,无须中心认证。GNRS(Global Name Resolution Service, GNRS)系统储存了每个 GUID 对应主机或网络的位置,而路由设备从 GNRS 系统获取这些位置并制定相应的路由策略,如图 8。在此基础上,通过不同的名字服务策略,解决终端异构性的问题。同时这种结构结合数据空间的小路和虚拟节点等技术,实现灵活的域间路由策略,解决了移动性的问题。

4.2.3 服务空间

服务空间解决智能网络中的感知与自适应问题。视频应用对服务质量要求的提升、网络新应用对互联网可演进性的要求以及终端设备的多样性,都对网络的智能化提出了要求。为了实现智能化网络,可以借鉴 Alicante^① 的设计思想,使用接入设备(如 Home Box)来实现终端用户和服务提供商之间的交互。接入设备由终端用户、服务提供商共同管理,通过对接入设备的配置和授权,可以使得终端用户在内容提供者、内容消费者和内容管理者等角色间进行切换,从而实现了服务层的灵活管理,提供高效的内容分发服务。同时,接入设备从终端用户获取上下文信息,如终端设备的类型、接入网络以及应用需求等,并提供给名字空间,为控制平面进行任务调度和内容管理提供依据。从名字空间获取状态信息,生成进一步抽象的网络视图,通过接入设备之间的交互协同和自适应,匹配用户的不同需求,以达到最佳的服务体验。接入设备从用户环境和内容中心网络环境获取用户上下文信息和网络状态信息,实现用户和网络间的双向感知与面向终端的自适应。通过智能化的终端设备管理与配置,实现灵活的服务模式,构建普适化的服务环境^[36]。同时,可在接入设备中加入图像处理单元(Graphics Processing Unit, GPU),提高传输效率,提供面向终端用户的自适应能力。

5 总结与展望

网络流量的视频化与新应用的不断推陈出新,

① ALICANTE Consortium, "D2. 2: Service/Content Adaptation Definition and Specification" and "D2. 3: Cross-Layer Monitoring Definition and Specification", FP7 ICT ALICANTE Project, <http://www.ict-alicante.eu/>

对现有互联网的体系架构提出了新的挑战。CDN 通过传统 IP 网络上部署服务节点, 并利用应用层协议将这些服务节点联结构成应用层覆盖网络, 为用户提供内容分发服务, 有力地支持了视频技术的发展。同时, CDN 技术的广泛使用也使互联网结构发生了新的变化, 如拓扑结构扁平化与协议栈沙漏结构细腰的上移, 这些变化为网络发展带来了新的机遇。但随着新应用的不断涌现与用户对服务质量的要求越来越高, CDN 在大规模视频分发服务时也存在诸多不足, 如可扩展问题、智能化问题和管理问题等。为了解决 CDN 技术面临的挑战, 应从体系结构的视角入手, 对互联网进行反思与设计。本文借鉴国外一些关于未来网络体系架构研究的思想, 提出了面向未来网络的多媒体网络体系结构与关键技术。多媒体网络的设计原则包括简单化、智能化、安全性、管理与策略。根据这几项设计原则, 借鉴操作系统的设计思想, 提出了多媒体网络的体系架构包含三层环境、三种感知和三层空间, 并重点对三种空间的设计要点进行了阐述。

参 考 文 献

- [1] Dobrian F, Awan A, Stoica I, Sekar V, Ganjam A, Joseph D, Zhan J, Zhang H. Understanding the impact of video quality on user engagement. *ACM SIGCOMM Computer Communication Review*, 2011, 41(4): 362-373
- [2] Koponen T, Shenker S, Balakrishnan H, Feamster N, Ganichev I, Ghodsi A, Godfrey P B, McKeown N, Parulkari G, Raghavan B, Rexford J, Arianfar S, Kuptso D. Architecting for innovation. *ACM SIGCOMM Computer Communication Review*, 2011, 41(3): 24-36
- [3] Labovitz C, Johnson S, McPherson D, Oberheide J, Jahanian F. Internet inter-domain traffic//*Proceedings of the ACM SIGCOMM 2010 Conference*. New Delhi, India, 2010: 75-86
- [4] Akhshabi S, Dovrolis C. The evolution of layered protocol stacks leads to an hourglass-shaped architecture. *ACM SIGCOMM Computer Communication Review*, 2011, 41(4): 206-217
- [5] Balakrishnan H, Lakshminarayanan K et al. A layered naming architecture for the Internet. *ACM SIGCOMM Computer Communication Review*, 2004, 34(4): 343-352
- [6] Guha S, Francis P. An end-middle-end approach to connection establishment. *ACM SIGCOMM Computer Communication Review*, 2007, 37(4): 193-204
- [7] Koponen T, Chawla M, Chun B-G et al. A data-oriented (and beyond) network architecture. *ACM SIGCOMM Computer Communication Review*, 2007, 37(4): 181-192
- [8] Stoica I, Adkins D, Zhuang S, Shenker S, Surana S. Internet indirection infrastructure. *IEEE/ACM Transactions on Networking*, 2002, 12(2): 205-218
- [9] Popa L, Ghodsi A, Stoica I. HTTP as the narrow waist of the future Internet//*Proceedings of the 9th ACM SIGCOMM Workshop on Hot Topics in Network*. Monterey, Canada, 2010
- [10] Yin H, Liu X, Zhan T, Sekar V, Qiu F, Lin C, Zhang H, Li B. Design and deployment of a hybrid CDN-P2P system for live video streaming: Experiences with LiveSky//*Proceedings of the 17th ACM International Conference on Multimedia (ACM MM)*. Beijing, China, 2009: 25-34
- [11] Krishnamurthy B, Wills C et al. On the use and performance of content distribution networks//*Proceedings of the 1st ACM SIGCOMM Workshop on Internet Measurement*. San Francisco, USA, 2001: 169-182
- [12] Krishnamurthy B, Wills C. Analyzing factors that influence end to end web performance. *Computer Networks*, 2000, 33(1): 17-32
- [13] Zhu W, Luo C, Wang J, Li S. Multimedia cloud computing. *IEEE Signal Processing Magazine*, 2011, 28(3): 59-69
- [14] Jacobson V, Smetters D K, Thornton J D, Plass M F, Briggs N H, Braynard R L. Networking named content//*Proceedings of the 5th International Conference on Emerging Networking Experiments and Technologies*. Rome, Italy, 2009: 1-12
- [15] Nelson S C, Bhanage G et al. GSTAR: Generalized storage-aware routing for MobilityFirst in the future mobile Internet//*Proceedings of the 6th International Workshop on MobiArch*. Bethesda, USA, 2011: 19-24
- [16] Anand A, Dogar F, Han D, Li B, Lim H, Machado M, Wu W, Akella A, Andersen D, Byers J, Seshan S, Steenkiste P. XIA: An architecture for an evolvable and trustworthy Internet//*Proceedings of the 10th ACM Workshop on Hot Topics in Networks*. Cambridge, England, 2011
- [17] Koponen T, Casado M, Gude N, Stribling J, Poutievski L, Zhu M, Ramanathan R, Hama T, Shenker S. Onix: A distributed control platform for large-scale production networks//*Proceedings of the 9th USENIX Symposium on Operating Systems Design and Implementation*. Vancouver, Canada, 2010
- [18] Koponen T, Chawla M et al. A data-oriented (and Beyond) network architecture. *ACM SIGCOMM Computer Communication Review*, 2007, 37(4): 181-92
- [19] Visala K, Lagutin D et al. LANES: An inter-domain data-oriented routing architecture//*Proceedings of the 2009 workshop on Re-Architecting the Internet*. Rome, Italy, 2009: 55-60
- [20] Gritter M, Cheriton D R. An architecture for content routing support in the Internet//*Proceedings of the 3rd Conference on USENIX Symposium on Internet Technologies and Systems*. San Francisco, USA, 2001: 4-4
- [21] Ghodsi A, Teemu K et al. Naming in CON architectures//*Proceeding of SIGCOMM Workshop on Information-Centric Networking*. Toronto, Canada, 2011: 1-6
- [22] Adhikari V, Jain S, Chen Y, Zhang Z. Vivisecting youtube: An active measurement study//*Proceedings of the 31st Annual IEEE International Conference on Computer Communications*. Orlando, USA, 2012: 2521-2525

- [23] Ng T E, Zhang H. Predicting Internet network distance with coordinates-based approaches//Proceedings of the 21st Annual IEEE International Conference on Computer Communications. New York, USA, 2002; 170-179
- [24] Wen Z, Rabinovich M. Network distance estimation with dynamic landmark triangles//Proceedings of the 2008 International Conference on Measurement and Modeling of Computer Systems. Annapolis, USA, 2008; 433-434
- [25] Dabek F, Cox R, Kaahoeok F, Morris R. Vivadi: A decentralized network coordinate system. ACM SIGCOMM Computer Communication Review, 2004, 34(4): 15-26
- [26] Madhyastha H, Katz-Bassett E, Anderson T, Krishnamurthy A, Venkataramani A. iPlane Nano: Path prediction for peer-to-peer applications//Proceedings of the 6th USENIX Symposium on Networked Systems Design and Implementation. Boston, USA, 2009; 137-152
- [27] Madhyastha H, Anderson T, Krishnamurthy A, Spring N, Venkataramani A. A structural approach to latency prediction//Proceedings of the 6th ACM SIGCOMM Conference on Internet Measurement. Rio de Janeiro, Brazil, 2006; 99-104
- [28] Xie H, Yang Y, Krishnamurthy A, Liu Y, Silberschatz A. P4P: Provider portal for applications//Proceedings of the ACM SIGCOMM 2008 Conference on Data Communication. Seattle, USA, 2008; 351-362
- [29] Andersen D G, Balakrishnan H, Feamster N, Koponen T, Moon D, Shenker S. Accountable Internet protocol(AIP)//Proceedings of the SIGCOMM. Seattle, USA, 2008; 339-350
- [30] Shaw M. Leveraging good intentions to reduce unwanted network traffic//Proceedings of the 2nd Workshop on Steps to Reducing Unwanted Traffic on the Internet. San Jose, USA, 2006; 9-9
- [31] Faratin P, Clark D, Gilmore P, Bauer S, Berger A, Lehr W. Complexity of internet interconnections: Technology, incentives and implications for policy//Proceedings of the 35th Research Conference on Communication, Information and Internet Policy (TPRC). Arlington, USA, 2007; 1-31
- [32] Clark D, Lehr W, Bauer S. Interconnection in the Internet: The policy challenge (August 9, 2011)//Proceedings of the 39th Research Conference on Communication, Information and Internet Policy (TPRC). Arlington, USA, 2011; 1-24
- [33] Godfrey P, Ganichev I, Shenker S, Stoica I. Pathlet routing//Proceedings of the ACM SIGCOMM 2009 Conference on Data Communication. Barcelona, Spain, 2009; 111-122
- [34] Cerf V et al. Delay-tolerant network architecture. IETF RFC 4838, Informational, 2007
- [35] Fall K. A delay-tolerant network architecture for challenged internets//Proceedings of the 2003 Conference on Applications, Technologies, Architectures, and Protocols for Computer Communications. Karlsruhe, Germany, 2003; 27-34
- [36] Satyanarayanan M. Pervasive computing: Vision and challenges. IEEE Personal Communication, 2001, 8(4): 10-17



YIN Hao, born in 1974, Ph. D., associate professor. His research interests include future Internet architecture, content delivery networks, and multimedia communication.

ZHAN Tong-Yu, born in 1986, M. S. candidate. His research interests include server placement, and future Internet architecture.

LIN Chuang, born in 1948, Ph. D., professor, Ph. D. supervisor. His research interests include computer networks, performance evaluation, network security analysis, and Petri net theory.

Background

With the rapid development of Internet applications, the traditional network architecture faces numerous challenges. The emerging demands for scalability, mobility, security, etc. are hard to be met by incremental changes for the Internet. There has been significant amount of effort on clean-slate architecture design for future Internet. This paper summarizes the existing challenges of current Internet and the research projects about future Internet architecture design, discusses the design philosophy of future multimedia network, proposes a new Multimedia Network architecture, and discusses the key technologies of the network. This work is partly

supported by the National Basic Research Program (973 Program) of China (No. 2012CB315800, No. 2011CB302601), National Natural Science Foundation of China (No. 60736012, No. 61170290). These projects aim to provide better performance in computer networks and information systems. Our group mainly focuses on Content Delivery Network, multimedia communications and Internet architecture. Our group has published over 100 papers in refereed journals and conferences and is a holder of more than 13 Chinese invention patents.