

服务定制网络

刘韵洁, 黄韬, 张娇, 刘江, 姚海鹏, 谢人超

(北京邮电大学 网络与交换国家重点实验室, 北京 100876)

摘 要: 针对当前互联网中存在的 2 个亟待解决的问题: OTT (over the top) 流量占用大量网络带宽以及信息高度冗余, 提出了一种具有差异化服务能力的未来网络体系架构——服务定制网络 (SCN, service customized networking)。SCN 基于软件定义网络设计, 继承了其数据控制分离以及网络可编程的主要特点, 并针对当前互联网中的问题, 增加了网络虚拟化能力以及内容智能调度能力。通过构建具有不同等级服务能力的虚拟网络, SCN 可以避免 OTT 业务大量占用带宽。同时, 通过信息资源智能有序调度, SCN 可以消减信息冗余。在小规模真实网络平台上进行的实验结果表明, SCN 确实具备为不同用户提供差异化服务能力以及有效减少信息冗余的能力。

关键词: 未来网络体系架构; 服务定制; 差异化; 信息冗余; 网络虚拟化

中图分类号: TP393

文献标识码: A

文章编号: 1000-436X(2014)12-0001-09

Service customized networking

LIU Yun-jie, HUANG Tao, ZHANG Jiao, LIU Jiang, YAO Hai-peng, XIE Ren-chao

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

Abstract: Today's Internet faces many challenges, two of which need to be addressed urgently. One is that most bandwidth is taken by OTT services, and the other one is that the content in the Internet is highly redundant. To address these two problems, a novel future Internet architecture, named as service customized networking (SCN), is proposed to provide differentiated services. SCN is designed based on the software-defined networking (SDN) technology. It inherits the two advantages of SDN, that is, separation of control and data plane and programmable networking. Besides, SCN adds the function of network virtualization, and intelligent content scheduling. SCN avoids OTT services taking too much bandwidth through employing the network virtualization technology to provide differentiated services. Besides, SCN effectively eliminates content redundancy by intelligent content scheduling. The results of experiments conducted on a real small-scale testbed indicate that SCN indeed provides differentiated services for different users and largely reduces content redundancy.

Key words: future internet architecture; service customization; differentiation; information redundancy; network virtualization

1 引言

当前, 中国互联网规模以及用户数量已位列全球第一。据中国互联网信息中心 (CNNIC, China internet network information center) 公布的第 34 次中国互联网络发展状况统计报告^[1]显示, 截至 2014

年 6 月, 我国互联网用户数量达到 6.32 亿, 占总人口约 46.9%。另外, 知名风投公司 KPCB 指出, 截至 2013 年, 中国移动互联网活跃用户已经高达 5 亿, 全球占比最高。

然而, 随着互联网规模的飞速增长以及互联网应用的多样化, 现有网络也逐步显现出了各种问

收稿日期: 2014-08-20; 修回日期: 2014-10-18

基金项目: 国家重点基础研究发展计划 (“973” 计划) 资助项目 (2012CB315801); 国家高技术研究发展计划 (“863” 计划) 资助项目 (2015AA011906); 国家自然科学基金资助项目 (61300184, 61302089)

Foundation Items: The National Basic Research Program of China (973 Program) (2012CB315801); The National High Technology Research and Development Program of China (863 Program) (2015AA011906); The National Natural Science Foundation of China (61300184, 61302089)

题, 其中 2 个非常关键的挑战如下。

1) OTT 业务的飞速发展。近年来, 实时通信工具、在线视频等 OTT 业务蓬勃发展, 用户数量和流量快速增长。截至 2013 年 10 月, 微信用户数量已经突破 6 亿, Facebook 用户数量超过 10 亿。思科预计 2016 年, OTT 业务数据流量将占据移动数据流量的 71%。OTT 业务的飞速发展带来的数据流量增加必然会导致互联网带宽被大量占用。然而, 由于 OTT 业务服务商直接面向用户提供服务 and 计费, 使基础网络提供商变成了单纯的数据传输通道, 因此缺乏足够的动力进行网络基础设施的扩容。可以预测, 如果按照这一趋势, 某一天整个产业生态链将会崩塌。那么如何缓解流量快速增长这一问题, 一些专家提出对 OTT 业务流量收费的方式, 然而这一方法可能会影响到整个互联网生态圈的蓬勃发展, 事实上, 互联网业务的快速发展与其不按流量收费密切相关, 如果按照流量收费, 将会制约互联网业务的发展, 甚至可能会阻碍其创新。

另外一个可行的思路是对网络进行改造, 以提供差异性服务的能力, 进而引入经济模型, 引导用户高效合理地使用网络资源。事实上, 当前互联网应用提供商对于网络差异性服务能力的需求非常强烈, 例如, 亚马逊指出网络延迟每增加 0.1 s, 每年销售将减少 1%^[2], 谷歌指出用户搜索延迟每增加 0.5 s, 搜索量将会减少 20%^[3]。据调查, 85% 的内容提供商为了提高用户的业务体验, 愿意为更好的网络服务质量支付最高 25% 溢价, 但是由于当前网络体系架构不灵活, 并不能为特定用户提供差异性的服务质量。于是, Google 等大型内容提供商已经开始试图构建自己的基础网络来更灵活地使用网络资源, 从而满足业务需求^[4]。但是, 这仅仅是少数大型企业能够选择的技术路线, 同时也会造成基础设施资源的巨大浪费。

如果可以构建更加灵活并能愿意付费用户提供有服务质量保障的基础网络, 则互联网可以持续健康的发展。为了构建这样一个网络, 可以借鉴交通系统的差异化运输方式, 令网络具有提供高速公路、普通公路、普通火车、高铁等差异化服务的能力, 普通用户可以免费使用普通公路, 但是服务质量不能得到保障; 而对于服务质量有需求的用户, 可以选择付一定费用来得到更高的服务质量。这样一来, 一方面对网络有服务质量需求的内容提

供商可以得到服务保障, 而普通用户和小型创业公司等可以继续免费使用互联网, 促进互联网的蓬勃发展。另一方面, 基础网络提供者会有更高的动力去建设基础设施, 来更好地满足不同用户对基础网络的需求。

2) 互联网中信息存在大量冗余。当前互联网中的大部分内容分发应用存在大量重复传输问题, 尤其是视频流量, 例如优酷 2011 年播放量过亿次电视剧超过 30 部。而当前互联网中视频流量占据了大部分, 思科预计未来在线视频流量将会占据互联网流量的 90% 以上。这些冗余流量会浪费大量网络带宽, 因此消除信息冗余需求十分迫切。当前网络中主要是利用内容分发网络 (CDN, content delivery network) 技术来减少信息冗余。但是, 传统 CDN 的解决方案存在诸多不足, 一方面, 其采用重叠网 (overlay network) 的方式进行设计, 增加了成本, 造成了基础设施资源的浪费。另一方面, 通常只有较成熟的公司才有能力去购买 CDN 服务, 因此 CDN 技术只是减小了购买其服务的内容提供商的数据流量, 并不是一个整体解决方案。对于基础网络来说, 由于缺乏对内容资源的智能调度能力, 信息冗余仍然大量存在, 因此基础网络具备信息冗余消除能力十分必要。同时调查指出, 存储容量比带宽增长速度更快, 而且存储成本价格比带宽价格下降更快, 按照目前数据, 存储容量平均每 1.7 年翻倍, 而带宽平均每 2.4 年翻倍; 存储成本每年下降 40%, 而带宽成本每年只下降 26%^[5]。因此从发展趋势来看, 未来在基础网络中部署缓存来减小信息冗余也是一个大的方向。

除了以上 2 个挑战, 当今互联网业务种类飞速增加, 个性化需求不断出现, 尤其是工业和工业互联网、能源互联网、车联网以及更多的业务体验需要差异化服务。因此, 需要探索新的网络体系架构来从根本上解决当今互联网面临的这些问题, 满足用户的个性化需求。在本文中, 提出一种新的网络体系架构服务定制网络 (SCN), 试图构建一种开放可扩展、高效灵活调度信息资源的具有差异性服务能力的网络体系架构, 从而网络基础设施提供商可根据用户的个性化需求, 实现计算、存储、带宽等网络资源灵活调度, 构建具有不同服务能力的虚拟网络, 进而构建新的产业生态链。

SCN 基于软件定义网络思想设计, 软件定义网络的主要特点是数据控制相分离以及数据平面可

编程, 而数据平面只提供了基本的数据转发能力, 并不具有提供服务区分以及消除信息冗余的能力。但是, 这些能力在未来网络中是非常必要的。所以, 在 SCN 中, 增加了网络虚拟化能力、大数据测量感知能力以及内容智能调度能力。SCN 网络体系架构主要具有 3 个特点。

1) 基于网络控制(软件实现)与数据交换(硬件实现)分离解决网络管理复杂的问题, 同时具有灵活构建不同服务质量等级的虚拟网络功能, 从而可以为不同用户提供差异化服务。

2) 实现信息内容资源智能有序调度, 内容贴近用户部署, 解决信息重复传输问题, 从而可以更有效利用基础网络资源。

3) 基于大数据思想实现网络测量、感知与分析, 为网络智能调度、控制提供支撑。

2 相关工作

当前互联网面临的问题日益严峻, 近几年学术界对未来网络体系架构的研究很多, 侧重点各有不同, 主要从网络管理、信息冗余、移动性、安全性等几个方面进行展开。本节将对当前典型未来网络体系架构进行简要的总结。

2.1 灵活管控

A. Greenberg 等人于 2005 年提出了 4D^[6]网络体系架构, 旨在针对当前网络逻辑决策平面和分布式硬件设备结合过紧的问题将决策逻辑从底层协议中完全分离出来, 从而简化网络管理。4D 项目倡导 4 个主要平面: 数据平面、发现平面、分发平面和决策平面。其中, 数据平面负责数据转发功能; 发现平面负责发现网络中的物理组件, 并为决策平面提供构建网络视图的基本信息; 分发平面在决策平面和路由器之间建立可靠的通信通道; 决策平面通过逻辑上集中的控制器将网络功能转换成数据分组处理状态。在 4D 项目的基础上, 2007 年, 斯坦福大学研究人员 Martin Casado 等人提出的 Ethane^[7]将安全管理策略添加到网络管理当中, 扩充了中央控制器的管理功能, 实现了更细粒度的流表转发策略。在 Ethane 网络中, 中央控制器和 Ethane 交换机是 2 个主要部件, 其中中央控制器负责完成网络主机认证、IP 分配和产生交换机流表等基本功能, 是整个网络的控制决策层; Ethane 交换机则负责根据控制器部署的流表进行报文转发, 是一个简单的、哑的数据转发单元。在 Ethane

工作的基础上, 斯坦福大学的研究人员于 2008 年提出了 OpenFlow 技术, 并逐渐推广 SDN 技术。SDN 技术中的逻辑控制和数据转发分离架构的设计思想简化了网络管理和配置操作, 实现了高层控制逻辑的顽健性。另外, 其软件可编程的特性使得网络体系架构更加灵活, 有利于创新应用的部署以及网络架构的演进。

2.2 信息冗余

施乐帕洛阿尔托研究中心的 Van Jacobson 于 2006 年分析认为网络的应用呈现以信息为中心的特点, 提出了信息中心网络的架构。其核心思想是采用类似 URL 的层次化命名方式对信息命名, 从而方便用户进行信息搜索和信息获取。这种架构将 IP 地址为中心过渡到了以信息为中心, 网络节点通过缓存数据可以有效降低信息冗余、加快对终端用户的响应速度。

这类网络体系架构的代表工作包括 DONA^[8]、NDN^[9]、CCN^[10]、ICN^[11]等。它们主要有 3 个共性特性。1) 基于分布/订阅模式。发布者将内容公开, 订阅者需要时要请求即可。这种模式对发布者和订阅者在空间和时间上都进行了解耦, 即发布者和订阅者不需要知道对方的位置, 也不需要同时在线。2) 全网缓存。当一个网络节点收到一个内容请求时, 如果本地缓存中有此内容则直接回复, 否则向邻居节点请求此内容, 并在收到此内容时在本地进行缓存。这种缓存的全局性体现在 3 方面: 对所有协议的内容都进行缓存; 对所有用户的内容都进行缓存; 网络中所有节点而非部分节点缓存数据。3) 面向内容的安全模型。由于在此类网络体系架构中, 内容大都来自于网络节点而非源服务器, 因此安全模型不能基于源地址。所以, 内容发布者会对发布的内容进行签名, 然后, 订阅者只需要检验内容签名的有效性即可。

2.3 其他

除了以上 2 种未来网络体系架构外, 还有一些架构重点对移动性、安全性和可重构方面进行了考虑。在移动性方面的典型代表为美国 NSF 的 FIA 项目提出的对移动性支持良好的未来网络体系架构 MobilityFirst^[12]。MobilityFirst 以一个基于名字的服务层作为协议栈的细腰, 此服务层使用全局唯一标识(GUID)来区分不同的终端对象。使用全局名字解析服务(GNRS)来动态绑定名字和地址。在安全性方面的典型代表为美国 NSF 支持的 XIA

项目。XIA^[13]指出由于源端主机和网络设备的协议难以改动，所以当前互联网很难演进。XIA 对主机和内容分别进行了内置安全操作，主机标识是主机公网的散列值^[14]，内容标识是内容的散列值，这样保障了安全性。最后，在可重构方面，2013 年美国 NSF 在其 FIA 项目第二阶段启动了 ChoiceNet^[15]项目，该项目为互联网演进开启了一个新的架构设计思路，它注重运用经济原则推动网络核心创新持续进行。这种新的网络体系架构的核心思想是支持选择，使网络不再是黑盒子，通过选择和竞争生成新的应用和商业模型。以服务为中心，建立可选择的技术方案和经济之间的关联。

3 服务定制网络

上节中提到新的未来网络体系架构都有各自的特点和优势，然而，它们不能或者只能部分解决当前基础网络建设者面临的 OTT 业务飞速发展以及信息存在大量冗余等挑战。本节将会具体介绍为解决基础网络建设者当前面临的挑战而提出的服务定制网络。

3.1 设计目标

1) 提供差异化服务质量

SCN 可以为内容提供商等用户构建差异化服务质量的虚拟网络，用户可以根据特定的需求以及经济承受能力选择适合的服务等级，从而构建良好健康的互联网经济模式。通过限制 OTT 业务占用的网络资源，OTT 业务对网络的冲击也可以迎刃而解。

2) 减少信息冗余

对于信息冗余带来的带宽浪费问题，CDN 也可以部分解决，但它是一种覆盖网的方案。而 SCN 试图让网络架构本身具备避免信息冗余的能力，网络具有感知内容、网络状态的功能，然后基于大数据的智能数据挖掘与分析，实现全网内容资源和网络资源的智能调度，从而实现有效消减信息冗余，充分利用网络基础设施的能力。

3) 网络可演进

虽然现有网络存在很多问题，但是已经受住了各种考验，用一种革命式的未来网络体系架构替换现有网络是一个非常漫长的过程，也面临着市场等各方面的压力。因此，为了渐进式部署，使 SCN 有可能被运营商等基础网络建设者采纳，以较低的成本逐步部署到现网中。SCN 当前设计方案仍然兼容采用基于 IP 的数据分组格式，采用深度数据分组

检测 (DPI, deep packet inspection) 的方式进行内容解析和调度。然而，当未来出现新的内容命名标识体系 (如 NDN/CCN 等) 或更优的网络协议后，SCN 系统需要能够方便地过渡至新的命名体系，支持新的技术。因此，SCN 的设计目标之三是立足于解决现有的互联网问题，同时符合未来的发展方向，具有可演进性。

3.2 设计框架

SCN 是基于软件定义网络设计的，然而软件定义网络主要解决了物理网络资源的调度，并没有对信息资源进行挖掘和调度。因此，SCN 在软件定义网络基础上增加了一个信息平面，如图 1 所示，SCN 的设计框架主要包括物理基础设施平面、控制平面和信息平面 3 个平面以及相邻平面之间的接口。

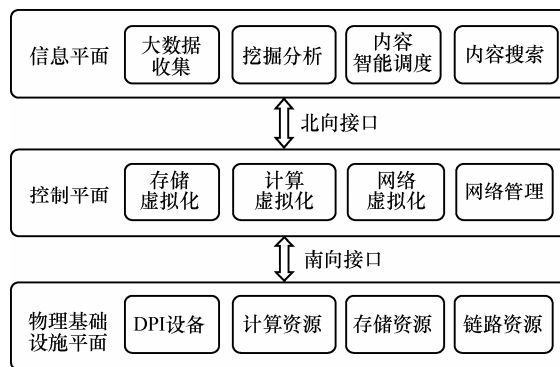


图 1 SCN 框架

信息平面的主要特点是开放和智能。主要负责对收集到的数据进行挖掘分析，从而指导内容的智能调度，实现信息贴近用户部署，降低信息冗余传输。

控制平面的主要特点是可管控、可扩展。主要负责全局网络视图创建与维护，虚拟子网的隔离等。从而构建不同服务等级虚拟网络，缓解 OTT 等业务的冲击。

物理基础设施平面需要高效、灵活可定义。例如，对于新协议，流表可以自定义；网络测量粒度可以动态定义；网络测量接口可方便配置等。

3.3 物理基础设施平面

SCN 中物理基础设施平面由交换设备、存储设备以及 DPI 设备等组成。控制平面通过南向接口协议可以对底层基础设施层设备的计算、存储、链路带宽资源进行统一的控制与调配，为不同等级的服务分配所需要的资源。同时，依赖控制平面下发的存储策略，基础设施层的存储设备可以对内容进行

缓存，为用户提供就近的访问，从而减少网络信息的冗余。此外，基础设施层的 DPI 设备将对分组进行检测分析，用于分析分组的特征与提取分组中的关键字段等，以支持大数据分析等，为上层做出决策。

3.4 控制平面

SCN 中的控制层除了具备基本的数据转发功能外，还具有网络虚拟化能力来为不同用户提供差异化的服务，解决 OTT 业务流量对网络的冲击问题。通过构建逻辑上隔离的子网，基础网络建设者可以给不同用户按照需求分配不同的网络、带宽等资源，同时，用户还可以根据自身的具体需求在网络内部使用不同的策略机制。另外，子网之间的隔离提高了不同子网间传输信息的机密性，有助于增加用户之间的信息安全。

为了实现网络虚拟化，首先 SCN 需要具有虚网映射和虚网隔离的功能。用户首先向网络提出虚网请求，这包括节点、链路、每个节点的 CPU、内存等需求以及每条链路对应的带宽等需求。然后控制器根据用户的请求信息在物理网络中找到恰当的节点和路径来实现一个逻辑上相当的虚拟子网，即虚网映射。最后，控制器通过一定的策略来实现虚网隔离。

当前对于虚网映射和虚网隔离都已经较多的研究^[16-24]，不过，虚网映射的研究大多偏重于理论研究，缺少工程实践的验证支持。虚网映射是 NP 难问题，已有工作提出了不同的启发式算法，这些算法均可以在 SCN 框架中使用。

虚网隔离除了具备基本的虚拟子网之间的流量隔离外，还应该具备：1) 可扩展性好，否则在大规模互联网中使用可能会成为性能瓶颈；2) 支持地址空间复用，即虚拟子网可以使用重叠的私网地址；3) 支持虚拟设备配置。从用户角度来看，操作虚拟网络中的设备应该如同操作物理设备一样，可以方便地对虚拟设备进行配置。

在 SCN 中，设计了一个新的网络虚拟化平台 CNVP^[25] (C-lab network virtualization platform) 来满足以上特性。CNVP 能够对底层物理网络进行抽象，并根据不同网络层协议定义虚拟网络，实现对网络基础设施的共享复用。CNVP 使用 MPLS 标签标识每个虚网（根据不同的需求此标签未来可以更换）。当虚网控制器向交换机添加流表信息时，此消息会被 CNVP 截获，然后 CNVP 根据虚网的配置

信息，向即将下发给交换机的流条目增加写 MPLS 标签的行为，以标识此流所属的虚网。当属于一个虚网的数据分组到达此虚网的边缘交换机时，边缘交换机为此数据分组添加 MPLS 标签，并在数据分组离开虚网时去除标签。CNVP 利用 MPLS 标签识别每个虚网，从而实现各虚网的流量隔离。另外，在 CNVP 中还实现了对物理资源的虚拟化，例如：1) 带宽，通过为虚网内的交换机配置端口队列来实现带宽隔离，从而为每个虚网分配独占的链路带宽资源；2) 虚拟化平台 CPU，通过为每个虚网设定处理队列，制定相应的处理策略来实现对虚网计算能力的分配和调度。

按照上述思路，图 2 给出了 CNVP 网络虚拟化平台的功能设计框图，各功能模块的详细说明如下。

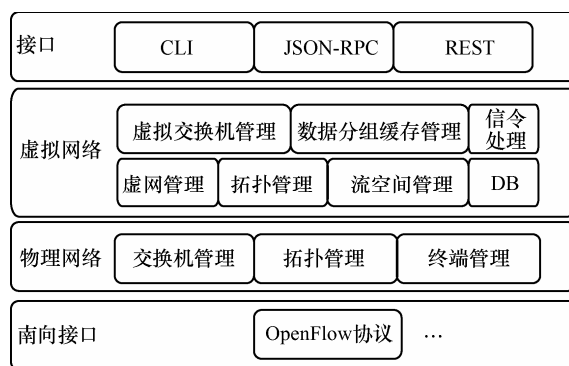


图 2 CNVP 的功能组成

1) 南向接口。CNVP 目前支持的南向接口有 OpenFlow 协议，具体版本包括 V1.0 和 V1.3 这 2 种，未来计划支持更多的南向接口协议，如 XMPP 等协议。

2) 物理网络。提供交换机管理、拓扑管理和终端管理的功能。交换机管理实现 CNVP 与数据平面转发设备通信、管理会话和端口的功能；拓扑管理通过定时发送 LLDP (link layer discovery protocol) 探测分组探测网络中的链路状态；终端管理提供对网络边缘节点进行管理的功能。

3) 虚拟网络。提供虚网管理、拓扑管理、流空间管理、虚拟交换机管理、数据分组缓存管理、信令处理和数据库管理的功能。虚网管理提供对虚网基本信息进行管理的操作以及虚网的创建、删除、启动和停止功能；拓扑管理提供对虚网拓扑的定义；流空间管理提供对流空间的配置和流空间冲突检测的功能；虚拟交换机管理提供与控制器交互的逻辑对象，维护与控制器之间的会话；数据分组缓存管理用于交换机和控制器之间通信时对负载数

据的缓存处理,以降低控制器和交换机之间的网络流量;信令处理提供上行消息的虚网分发和下行信令的整形处理;数据库用于保存虚网信息、拓扑定义以及流空间定义等信息。

4) 接口。CNVP 目前可以提供 CLI、JSON-RPC 接口和 REST 接口 3 种类型的接口,其中 JSON-RPC 接口实现了基于 Web 的网络虚拟化管理系统,可对物理网络、虚拟网络、拓扑管理等进行可视化操作,极大地方便了用户对网络虚拟化平台的了解和使用。

3.5 信息平面

SCN 中的信息平面主要通过用户对请求进行测量感知,然后对收集的信息进行大数据分析,根据分析结果指导信息资源在网络存储节点间的智能调度,使内容更加贴近用户,从而减小网络中的信息冗余,同时提高对用户请求的响应速度。

传统互联网中的缓存策略通常是针对具体协议的,例如 HTTP 协议。而 SCN 中的缓存能力是网络本身的一种属性:一方面,不再局限于特定的内容,即可以针对更加广义的内容进行缓存来去除信息冗余;另一方面,不再是针对特定内容提供商。所有网络使用者均可以享用网络缓存的这种能力来提高对用户的响应速度。

SCN 中的去除信息冗余功能主要是通过实时用户请求处理模块和定期内容调度模块实现的。为了实现对广义内容的缓存,SCN 中的接入网络设备需要具有 DPI 能力,控制器维护所有存储节点上的内容分布。用户对内容的请求与内容在不同服务器之间的调度均由控制器进行管理,2 个过程是独立的。

实时用户请求处理。当用户请求到达时,此模块负责对请求内容进行解析,并找到合适的内容提供者。具体流程如图 3 所示。

1) 用户的数据分组发向接入交换机。

2) 接入交换机利用 DPI 功能检测用户的数据分组是否是请求分组,如果是请求分组,则将整个请求数据分组发向控制器。

3) 控制器根据请求分组的请求内容,查看在存储节点上是否已经存在。如果不存在,则将用户此次请求的流到服务器的转发表下发到相应的转发节点上。如果存在,则找到一个离用户节点最近的存储节点的 IP 地址。将用户此次请求的流到相应存储节点的转发表下发到相应的转发节点上。

4) 如果用户请求内容不在存储网络中,则服务

器收到用户请求,并将内容传输给用户。

5) 如果用户请求内容已经存在于存储网络中,则相应的存储节点收到用户请求后,将内容传输给用户。

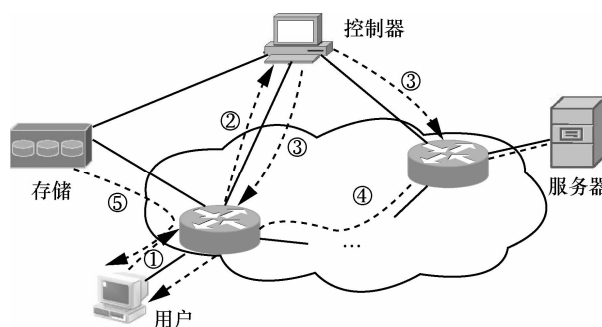


图 3 缓存策略

定期内容调度。控制器定期地根据用户请求的内容进行大数据分析,然后根据数据分析结果进行服务器到存储网络之间的内容调度。如果存储节点已经没有剩余存储空间,则可以根据某个缓存策略进行缓存替换,例如根据用户请求的热度对访问频率比较低的内容进行替换。

3.6 SCN 特性

3.6.1 提供差异化服务

SCN 中的网络虚拟化平台 CNVP,可以对不同的业务进行隔离,并为不同的业务进行不同的带宽资源分配。从而 OTT 业务无法过度占用带宽,侵占其他业务的资源。

3.6.2 减少信息冗余

SCN 中的缓存策略具有全网有效性和所有内容有效性。控制器通过对用户请求进行数据分析,并根据分析结果指导内容的智能调度,可以大大提高缓存的命中率,从而减少信息冗余。

4 SCN 实现和验证

4.1 小规模实现

为了验证 SCN 的基本性能,本文基于网络创新实验环境 C-LAB^[25],利用其计算和存储资源构建了一套小规模 SCN 实验平台。此实验平台共包括 30 多个交换节点,10 多个专用计算存储节点。整个实验平台的网络拓扑如图 4 所示。

该实验网络中的主要节点间采用了光纤直连以保证骨干网络的传输速度,交换节点采用了 Open vSwitch 软件交换机与 H3C、盛科、Pica8、xNet 等硬件交换机共存的模式,并且在各个节点做了数据平面的无线扩展,以方便用户接入。

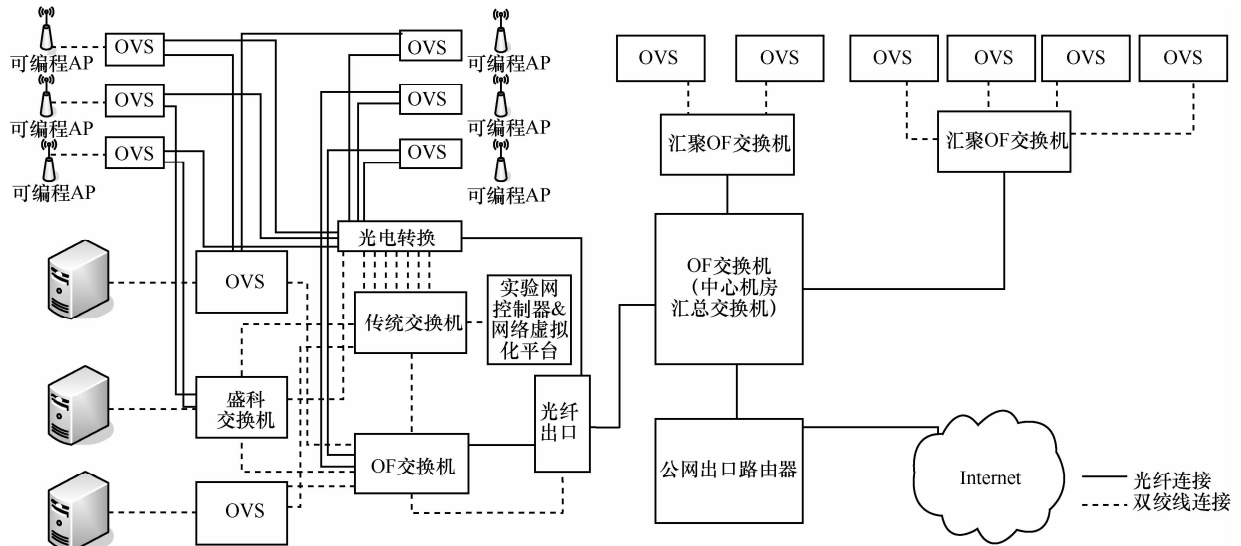


图 4 SCN 实验平台拓扑

4.2 性能验证

为了验证 SCN 网络提供的差异化服务能力和消除信息冗余能力，设计了如下 2 个实验。

实验 1 提供差异化服务。在这个实验中，用户向 SCN 网络提出 2 个虚拟子网请求，请求的带宽大小不同。记这 2 个子网分别为 A 和 B，拓扑如图 5 所示。子网 A 中所有的链路带宽请求值为 $C1$ Mbit/s，子网 B 中所有的链路带宽请求值为 $C2$ Mbit/s，其中 $C2=(100-C1)$ Mbit/s。在完成 2 个子网的映射之后，为了测试 SCN 网络确实按照子网的不同需求提供了不同的带宽值，在这 2 个子网中的 2 个终端间建立 TCP 连接，传输同样大小的文件，测试流的完成时间。在实验中，使用的文件大小为 10 MB。

如图 6 所示，流的完成时间与子网请求的带宽值是成正比的。例如，当给予子网 A 分配的带宽为 10 Mbit/s 时，传输完 10 MB 文件需要约 $10\text{ MB} \times 8/10\text{ Mbit/s} = 8\text{ s}$ 。这与图 6 中的实验数据 (8.2 s) 非常接近，说明 SCN 网络确实给予子网 A 保障了 10 Mbit/s 的带宽。而对于子网 B，当给予子网 A 分配的带宽为 10 Mbit/s 时，子网 B 得到的带宽为 90 Mbit/s，所以，传输完 10 MB 的文件需要约 $10\text{ MB} \times 8/90\text{ Mbit/s} = 0.89\text{ s}$ ，这与实验数据 (0.98 s) 也很接近。同样，对于图中的第 2 列数据 (横坐标为 20 Mbit/s)，这时给予子网 A 分配的带宽为 20 Mbit/s，所以，给予子网 B 分配的带宽为 $100-20=80\text{ Mbit/s}$ 。在子网 A 中，传输完一个 10 MB 的文件，需要的时间为 $10\text{ MB} \times 8/20\text{ Mbit/s} = 4\text{ s}$ ，这与图中左数第 2 个矩形标记所示的 4.1 s 非

常接近。在子网 B 中，传输完一个 10 MB 的文件，需要的时间为 $10\text{ MB} \times 8/80\text{ Mbit/s} = 1\text{ s}$ ，这与图中左数第 2 个圆形标记所示的 1.18 s 也非常接近。其他数据类似。因此，实验数据表明 SCN 网络确实可以按照子网的需求提供差异化的网络带宽资源。

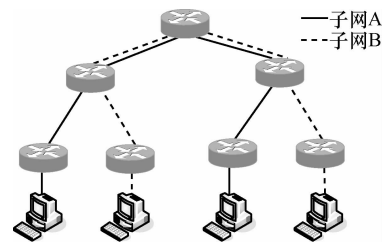


图 5 子网 A 和子网 B 拓扑

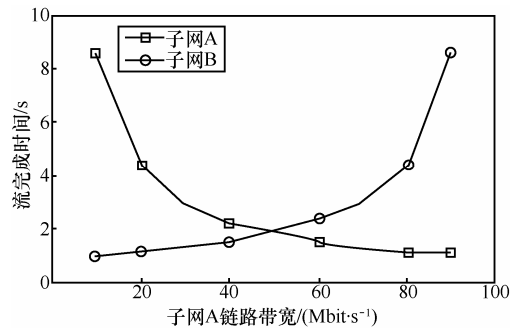


图 6 不同子网中的流完成时间

实验 2 减少信息冗余。此实验主要是为了测试 SCN 确实具有有效消减信息冗余的能力。为了对比，令内容提供商 A 使用 SCN 的缓存功能，而 B 不使用。然后，令同一个用户分别访问内容提供商 A 和 B 的内容 1 000 次，每次都记录

下响应时间。

图 7 显示了 2 种情况下访问时间的累积分布函数 (CDF, cumulative distribution function), CDF 曲线可以表明访问时间的分布情况。可以看到, 用户请求使用了缓存功能的内容提供商 A 的内容时, 纵坐标从 0 到 1 对应的横坐标 (访问时间) 全部集中在 2 s 左右, 这说明访问内容 A 的所有请求均可以在 2 s 左右得到响应。而访问不使用缓存功能的内容提供商 B 的内容时, 95% 的请求在 3 s 内得到响应, 而还有 5% 的请求响应时间介于 3 s 和 23 s 之间。由于缓存节点距离用户更近, 所以, 访问内容提供商 A 的时间分布说明大部分请求都在 2 s 左右的时间得到了响应, 即在缓存处得到访问的内容, 不需要再去访问内容源服务器, 这可以大大减少内容源服务器到缓存节点内容的冗余度。而访问内容提供商 B 的内容时, 有 95% 的请求在 3 s 左右才得到响应, 说明请求不能在缓存节点得到响应, 需要访问内容源服务器, 这样会造成访问内容在内容源服务器到用户的重复传输。此实验表明 SCN 的缓存功能使得向内容提供商 A 访问的绝大部分重复请求可以从缓存节点直接获取内容, 从而可以有效消减服务器到用户之间的冗余信息。

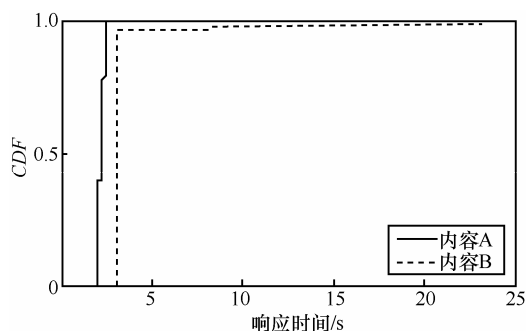


图 7 访问不同内容的响应时间分布

5 结束语

当前互联网中存在 2 个亟待解决的问题: OTT 业务飞速发展以及信息高度冗余。为了解决这 2 个问题, 提出了一种新的未来网络体系架构 SCN。SCN 主要包括基础设施层、控制层和信息层 3 个平面, 通过提供差分服务的能力能够有效避免 OTT 业务大量占用带宽, 同时通过提供缓存功能能够大大减少高度信息冗余对带宽的浪费。在小规模网络平台上进行的实验结果表明 SCN

确实具备为不同用户提供不同服务质量以及消减信息冗余的能力。

参考文献:

- [1] <http://www.cnnic.cn/hlwfzyj/hlwxzbg/hlwtjbg/201407/P020140721507223212132.pdf>[EB/OL]. 2014.
- [2] Amazon found every 100ms of latency cost them 1% in sales[EB/OL]. <http://blog.gigaspace.com/amazon-found-every-100ms-of-latency-cost-them-1-in-sales/>, August, 2008.
- [3] <http://glinden.blogspot.com/2006/11/marissa-mayer-at-web-20.html>[EB/OL]. 2006.
- [4] SUSHANT J, *et al.* B4: experience with a globally-deployed software defined WAN[A]. ACM SIGCOMM[C]. 2013.
- [5] MALIK O. The storage vs bandwidth debate[EB/OL]. <http://gigaom.com/2011/06/24/the-storage-vs-bandwidth-debate/>. 2011.
- [6] GREENBERG A, *et al.* A clean slate 4D approach to network control and management[A]. ACM SIGCOMM Computer Communication Review[C]. 2005.41-54.
- [7] CASADO M, MICHAEL J, *et al.* Ethane: taking control of the enterprise[J]. ACM SIGCOMM Computer Communication, 2007, 37(4): 1-12.
- [8] KOPONEN T, *et al.* A data-oriented (and beyond) network architecture[A]. ACM SIGCOMM Computer Communication Review[C]. 2007.181-192.
- [9] ZHANG L, ESTRIN D, BURKE J, *et al.* Named data networking project[A]. ACM SIGCOMM Computer Communication Review[C]. 2010.66-73.
- [10] JACOBSON V, MOSKO M, SMETTERS D, *et al.* Content-centric Networking[S]. Palo Alto Research Center, 2007. 2-4.
- [11] KUTSCHER D, *et al.* Information-centric networking[A]. Dagstuhl Seminar[C]. 2010.
- [12] DIPANKAR R, NAGARAJA K, *et al.* Mobilityfirst: a robust and trustworthy mobility-centric architecture for the future internet[J]. ACM SIGMOBILE Mobile Computing and Communications, 2012, 16(3):2-13.
- [13] ASHOK A, DOGAR F, *et al.* XIA: An architecture for an evolvable and trustworthy internet[A]. Proceedings of the 10th ACM Workshop on Hot Topics in Networks[C]. 2011.
- [14] ANDERSEN D G, BALAKRISHNAN H, FEAMSTER N, *et al.* Accountable internet protocol (AIP)[A]. ACM SIGCOMM Computer Communication Review[C]. 2008.339-350.
- [15] GEORGE N, *et al.* Optical network design and modeling (ONDM)[A]. Proceedings of the Optical Network Design and Modeling(ONDM)[C]. 2013.1-6.

- [16] GONG L, WEN Y G, ZHU Z Q, *et al.* Toward profit-seeking virtual network embedding algorithm via global resource capacity[A]. Proceedings IEEE INFOCOM[C]. 2014.1-9.
- [17] ZHU Y, AMMAR M. Algorithms for assigning substrate network resources to virtual network components[A]. Proc IEEE INFOCOM[C]. 2006. 1-12.
- [18] LU J, TURNER J. Efficient Mapping of Virtual Networks Onto a Shared Substrate[R]. Washington University in St Louis, 2006.
- [19] YU M, YI Y, *et al.* Rethinking virtual network embedding: substrate support for path splitting and migration[A]. ACM SIGCOMM Comput Communication Review[C]. 2008.17-29.
- [20] RAZZAQ A, RATHORE M. An approach towards resource efficient virtual network embedding[A]. Proc INTERNET[C]. 2010.68-73.
- [21] FAJJARI I, AITSAADI N, PUJOLLE G, *et al.* VNE-AC: virtual network embedding algorithm based on ant colony metaheuristic[A]. Proceedings IEEE ICC[C]. 2011.1-6.
- [22] CHENG X, *et al.* Virtual network embedding through topology-aware node ranking[J]. SIGCOMM Comput Commun, 2011, 41(4): 39-47.
- [23] ZHANG S, QIAN Z, WU J, *et al.* An opportunistic resource sharing and topology-aware mapping framework for virtual networks[A]. Proc IEEE INFOCOM[C]. 2012. 2408-2416.
- [24] ROB S, *et al.* Can the production network be the testbed [A]. OSDI[C]. 2010. 1-6.
- [25] [http://www.fnlab.org/\[EB/OL\]](http://www.fnlab.org/[EB/OL]).

作者简介:



刘韵洁 (1943-), 男, 山东烟台人, 中国工程院院士, 北京邮电大学教授、博士生导师, 主要研究方向为未来网络体系架构。

黄韬 (1980-), 男, 重庆人, 博士, 北京邮电大学副教授, 主要研究方向为未来网络体系架构、软件定义网络、内容中心网络。

张娇 (1986-), 女, 河北保定人, 博士, 北京邮电大学讲师, 主要研究方向为未来网络体系架构、软件定义网络、数据中心网络。

刘江 (1983-), 男, 河南郑州人, 博士, 北京邮电大学讲师, 主要研究方向为未来网络体系架构、软件定义网络、内容中心网络。

姚海鹏 (1983-), 男, 河北张家口人, 博士, 北京邮电大学讲师, 主要研究方向为未来网络体系架构、无线网络、大数据。

谢人超 (1984-), 男, 福建南平人, 博士, 北京邮电大学讲师, 主要研究方向为未来网络体系架构、内容中心网络、无线网络。