

内容分发网络研究

李 乔,何 慧,张宏莉

(哈尔滨工业大学网络与信息安全研究中心,黑龙江哈尔滨 150001)

摘 要: 内容分发网络是一种以降低互联网访问时延为目的,在网络边缘或核心交换区域部署内容代理服务,通过全局负载调度机制进行内容分发的新型覆盖网络体系.随着多媒体网络流及实时交互技术的兴起,现今内容分发网络已成为互联网的核心应用之一.本文介绍了内容分发网络的演变过程与体系结构,深入剖析了内容分发网络的核心功能组件与工作机制,对当前主流内容分发网络进行分类,并讨论了内容分发网络性能度量,最后探讨了内容分发网络的未来挑战与进一步研究方向.

关键词: 覆盖网络; 内容分发网络; 内容路由; 负载均衡; 性能测量

中图分类号: TP393 **文献标识码:** A **文章编号:** 0372-2112 (2013) 08-1560-09

电子学报 URL: <http://www.ejournal.org.cn> **DOI:** 10.3969/j.issn.0372-2112.2013.08.017

Research on Content Delivery Networks

LI Qiao, HE Hui, ZHANG Hong-Li

(Research Center of Network and Information Security, Harbin Institute of Technology, Harbin, Heilongjiang 150001, China)

Abstract: CDN (Content Delivery Network) is a novel overlay network, which deploys surrogates on the edge or backbone of network in order to reduce the access delay; it improves the efficiency of content delivery on the global load balance component. In current status, CDN has been a core Internet application as the development of multimedia and network technologies. In this paper, the evolution of CDN and its architecture are introduced; the key component of CDN and mechanism are presented; the taxonomy of CDNs and the performance measured strategies are reviewed. The developing trends of CDN and further topics in this area are also analyzed.

Key words: overlay network; content delivery network; content routing; load balance; performance measure

1 引言

随着高带宽消耗的 Internet 服务不断涌现,单纯依赖高性能数据中心的网络结构已经无法为广泛分布于全球的用户提供可靠的服务质量,同时日益增长的流量以及服务器与客户端能力与资源的不对称性促使内容分发技术的产生.上个世纪 90 年代后期,网络拥塞成为当时制约互联网性能的最主要因素,因此研究人员提出内容分发网络(Content Delivery Networks)技术以提升互联网边缘的用户的访问体验^[1].

在当前 Internet 上构建一层内容分发覆盖网络对改进互联网性能有着诸多优势:(1)减少链路中的重复流量,缩短传输路径,从而缓解互联网的流量压力;(2)部署多个内容副本服务器(surrogate),降低源服务器负载;(3)提升内容提供商的服务质量与服务可靠性等.内容

分发网络将 Internet 用户与 Internet 资源提供商之间的通讯分割为两个部分:(1)用户与内容分发网络中的内容副本服务器的交互;(2)内容副本服务器与源内容服务器的交互.这种分离将源服务商面向用户的服务交付与内容分发网络,显著降低源服务商的操作成本、部署难度及管理复杂性,同时催生出巨大内容分发网络市场.

新兴的内容网络分发技术已引起国内外研究机构,IT 业界的广泛重视.自 2002 年以来, SIGCOMM, USENIX, WWW 等顶级学术会议开始关注内容分发网络^[2~6].同时,美国大型电信运营商 AT&T 于 2002 年推出 CDN 流媒体内容分发架构.2006 年,国际电信联盟 ITU(International Telecommunication Union)将 CDN 纳入 IPTV 标准化文档体系内^[7].IETF(Internet Engineering Task Force)也制定了多项涉及内容分发网络技术的标准化文档,如 RFC

3040,3466,3835,3866^[8~11].Spagna 等人在文献[12]中分别从缓存服务器位置部署、路由请求机制、内容副本放置策略及内容定位四个方面提出设计高分布式内容分发网络的关键因素.国内对内容分发网络研究的关注和相关工作还较少.清华大学尹浩等人与中国最大的内容分发网络提供商 ChinaCache 建立联合实验室,对 CDN 进行长期的研究^[13,14],并在文献[15]中对内容分发网络进行了简要的综述.北京邮电大学杨戈等人在文献[16]中针对内容分发网络中的流媒体技术进行了综述.鉴于内容分发网络在当前互联网所占比日趋增大,为深入理解内容分发网络的功能结构与部署机制,对其研究方向进行总体把握,并促进国内在该方向上的研究,对内容分发网络进行综述具有重大意义.

本文首先阐述了内容分发网络的演变过程,并对内容分发网络进行定义,同时重点分析了内容分发网络的体系结构,然后对内容分发网络中的部署策略、内容路由、内容管理、性能测量的各个方面的主要研究工作进行了总体介绍,最后对内容分发网络的发展趋势进行了展望.

2 内容分发网络演化进程

Internet 流量指数性增长以及用户对网络速度的需求已促使网络资源分发的准确性、可用性、可靠性等成为互联网技术的关键问题.在上个世纪 90 年代,由于网络接入带宽的限制,研究人员提出采用代理缓存技术提升网络性能.缓存技术具有显著的优势:通过缩短资源在网络中的传输距离降低带宽消耗与网络拥塞,进而增强资源可用性与服务可靠性.然而代理技术具有较强的本地局部性,对于整个互联网的效率提升作用有限.当前的代理缓存技术主要以层次结构缓存为主,以改进服务性能与减少带宽消耗^[17].继代理缓存技术之后,研究人员提出一种服务场概念^[18],但是服务场依旧只是在源服务器附近部署服务机群与 4~7 层交换机等对网络请求进行分发.

随着 Web 技术的不断推进,静态页面在整个互联网流量中的比重持续下降,动态交互数据与多媒体共享资源的迅猛增长导致代理缓存技术很难满足这些内容分发的需求.更关键的在于近源的部署策略并不能很好的解决类似于 flash crowd^[19]与 slashdot^[20]现象.1998 年,麻省理工大学 Leighton 等人为解决在整个互联网中长距离传输的时延与网络拥塞问题,首次提出内容分发网络这种新型的覆盖网络结构^[21,22],并依托该项新技术建立了全球最大的内容分发网络服务提供商 Akamai^[23].在经历 10 余年的技术改进,内容分发网络已经发展到第 4 代,表 1 给出它的演化过程.

表 1 内容分发网络演化进程

| | 时间 | 业务对象 | 简要描述 |
|------------------------------------|-------------|------------|------------------------------|
| 第一代 CDN (Web-based) | 1998 - 2002 | 静态/动态网络数据. | 主要针对静态/动态 web 页面进行分发. |
| 第二代 CDN (Video on demand-based) | 2002 - 2006 | 多媒体数据. | 主要针对大数据量的媒体流进行分发. |
| 第三代 CDN (P2P-assisted) | 2006 - 至今 | 共享类数据. | 通过与 P2P 网络融合,降低服务端负载等. |
| 第四代 CDN (cloud-based) | 2009 - 至今 | 整合型数据. | 通过与云计算平台融合,解决资源整合,服务统一管理等问题. |

内容分发网络的迅猛发展促使国内外研究机构、业界厂商与国际相关组织开始进行制定标准化文档.互联网工程任务组 IETF 早在 2001 年就将内容分发网络内容写进 RFC 规范中^[12,20],国际电信联盟 ITU(International Telecom union),互联网流媒体联盟 ISMA(Internet Streaming Media Alliance),欧洲电信标准化协会 ETSI(European Telecommunications Standards Institute)均随后在其标准化文件中对内容分发网络进行阐述.2009 年中国电信研究院报告指出全球内容分发网络市场年均复合增长率 CAGR(Compound annual growth rate)达到 44.6%,并预测到 2013 年其市场规模将达到 45 亿美元^[24].

3 内容分发网络定义与体系结构

3.1 内容分发网络定义

内容分发网络作为运行于整个互联网上的覆盖网,主要致力于解决互联网中的服务质量优化.一般而言,一个典型的内容分发网络将多个内容存储服务器部署于不同的地理位置的 ISP(Internet Service Provider)域中,包括网络接入点(Point of Presence)与骨干交换中心域.这些内容存储服务器统一由全局内容管理服务器管辖,并依据用户的访问体验动态调整内容路由策略,以优化服务负载,流量压力等网络性能.IETF 在 RFC3466 中将内容分发网络定义为内容网络中的一种:一个典型的内容分发网络由请求路由(Request-Routing)系统,内容存储代理服务器(Surrogate),内容分布(Distribution)系统与审计(Accounting)系统构成.Vakali 等人定义内容分发网络应当至少包括四个部分:内容代理服务器,内容路由转发系统,内容分布设施与审计机制^[25].本文在综合前述内容与文献[26]的基础上,对内容分发网络进行如下定义:以改善网络服务质量为目的,在网络边界或核心交换区域部署内容代理服务并

通过内容路由、全局负载调度、分布式存储与系统审计进行管理的覆盖网络。

文献[27]将内容分发网络分为四个功能结构:(1)内容外包组件(content outsourcing unit):将数据从源服务端推送到代理服务端;(2)内容分发组件(content delivery unit):将数据从代理服务端分发至用户端;(3)请求路由组件(request routing unit):通过特定的算法将用户请求转发到合适的代理服务端;(4)管理组件(management unit):对上述三个组件的信息数据进行收集、分析并对系统进行自优化。

典型的内容分发网络工作流程如图1所示,互联网用户首先向资源定位服务器发送内容请求,定位服务器收到请求后根据当前的内容路由策略将其转发至最低延迟低负载的内容代理服务器,最后内容代理服务器将所请求的内容传输至用户。从图1可看出,内容分发网络在传统的 C/S(Client/Server)模式中增加了数据分发层与全局服务管理层,其中数据分发层将内容提供商的资源进一步推向距离用户更近的区域,有效缩短

了互联网的传输距离和访问时延,同时减少了互联网中的重复流量,大幅降低源服务器的负载压力。在管理服务层中,内容管理服务负责对源内容进行分类与整合,把复杂的源数据分割为可缓存与不可缓存,以优化内容的存储机制;分布式数据管理负责对源数据进行分布式存储与更新;全局负载服务通过实时收集分析内容代理服务器的访问情况动态调节资源定位算法;数据分析审计负责对整体网络日志进行分析以促进整个内容分发网络的性能优化以及作为商业化内容分发网络的计费数据来源。内容分发网络在当前已有成熟的商业化运营模式,影响其定价的主要因素有:带宽消耗率、流量分布的种类、内容副本的空间大小、内容代理服务器数量及相关的稳定性可靠性与安全性^[28]。文献[29]指出商业化内容分发网络的核心竞争优势在于扩展性、安全性、可靠性与服务性能。

本文在文献[25,28,29]的基础上,结合当前的内容分发网络运营商的功能,提出如图2所示的内容分发网络体系结构。内容分发网络体系可以分为硬件设备

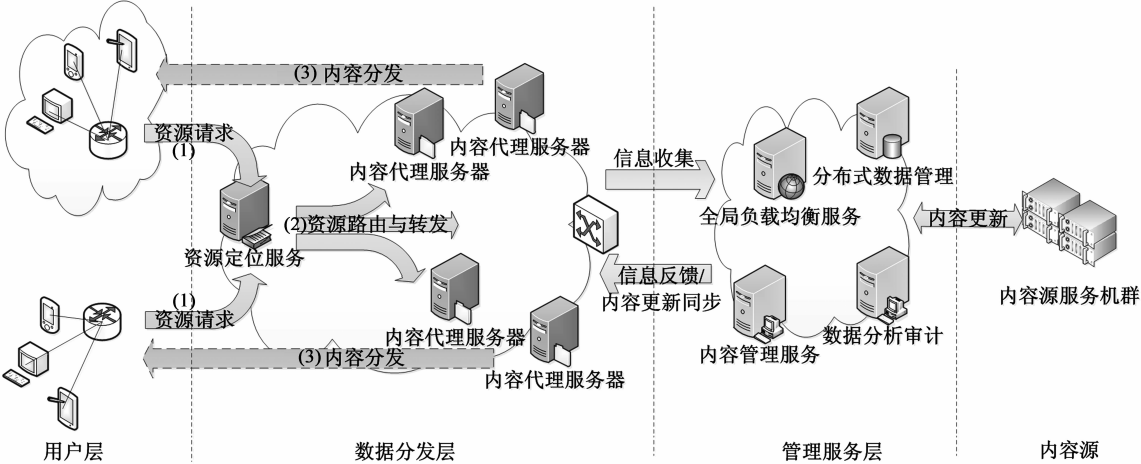


图1 内容分发网络工作机制

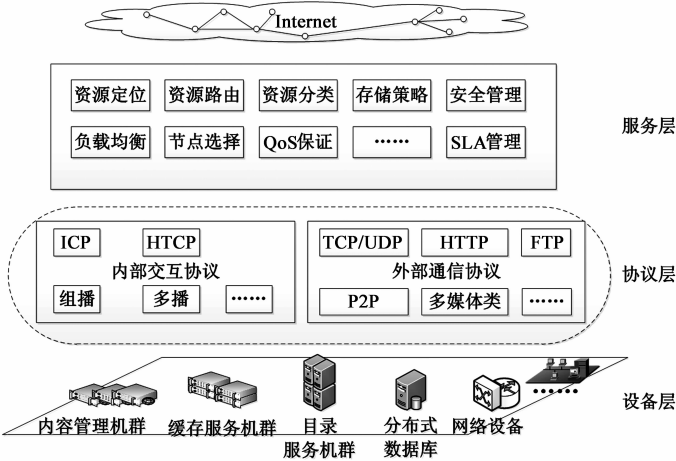


图2 内容分发网络体系

层,交互协议层与应用服务层三个层次:硬件设备层包括各种内容分发服务管理机群、分布式数据库与网络转发设备;交互协议层分为内部交互与外部交互两个部分,服务管理层是内容分发网络的核心部分,包括资源路由与定位、全局负载均衡、资源分发、存储模式与安全管理等。

3.2 部署机制

内容分发网络的部署策略主要分为两类:边缘化部署与核心域部署。边缘化部署的思想是将内容代理服务器部署于多个 ISP 内,从而降低跨域网络传输的时延,这种部署策略的代表是 Akamai,其在全球范围内 78 个国家的 1,000 个网络中部署超过 105,000 台服务器。核心交换区域部署的思想是将大规模数据中心部署于高速光纤骨干网络,通过高带宽、高处理能力降低用户访问时延,这种部署策略的代表是 Limelight^[30],其在全球范围 72 个光纤交换区域部署数据中心,与超过 578 个 ISP 进行光纤直连。

边缘化部署机制是内容分发网络的主流部署方式,主要是考虑到全球性网络传输距离过长,为减少数据在网络内的“最后一公里”(last mile)传输时延,将资源前置于离用户更近的区域以期有效改善访问体验。为保证内容分发的服务质量,内容分发网络管理者需要维护大量分布式服务器数据的同步和更新。总体而言,该问题等价于在 N 个不同的位置部署 M 个服务器($N > M$)以达到最小代价消耗,即 K-median 问题^[25],而最小化 K-center 问题已被证明是 NP 完全的^[31]。

Ni 等人在文献[32]中认为基于层次化结构的内容分发服务能够有效适应当前的层次化 IP 体系,并提出一种基于簇的协作化 cache 策略应用于多媒体内容分发网络。随着 P2P(peer to peer)技术在网络上的普及,一些多媒体内容分发技术将内容分发与 P2P 相结合,如 Jiang 等人在文献[33]中提出一种面向 P2P-CDN 的部署策略。总之,该类方法主要在逻辑拓扑上改进部署机制,通常都关注于整体的结构优化。

内容代理服务器的位置部署策略是内容分发网络中的一项关键技术,其直接影响分发服务质量、网络流量负载以及成本代价。合适的服务器部署策略能够降低 ISP 域间流量结算成本与域内重复流量。Vakali 等人在文献[25]中指出内容分发网络管理者通常依据用户所在的网络类型选择部署策略,主要分为 Single-ISP 与 Multi-ISP 两类。对于 Single-ISP,至少需要部署 40 个内容代理服务器才能保证在地理位置上覆盖整个 ISP 边缘网络。Multi-ISP 的最大优势在于能够将请求访问流量控制在本域内,降低了 ISP 的域间流量。

综上所述,当前内容分发网络的部署策略研究如图 3 所示,主要以层次化结构为主,包括基于树、基于簇

或基于图的算法。由于代价度量的多样化,如平均响应时间、访问距离/跳数、带宽利用率、服务负载等,而且互联网的实时变化也导致部署策略随之改变,因此并没有通用化的部署机制。总而言之,层次化混合结构的部署方式是当前主流的服务器部署体系,随着云计算技术的兴起,对于新型基于云的内容分发网络部署策略将是一个重要的研究方向。

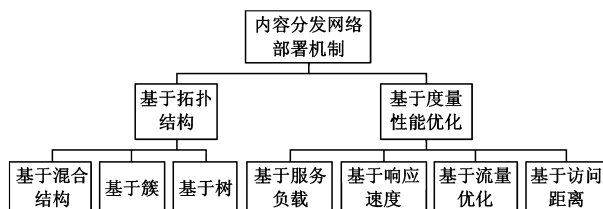


图3 内容分发网络部署机制分类

3.3 内容路由

内容分发网络作为内容网络的一种,内容路由策略的优劣直接影响整个内容分发网络的性能。内容路由主要负责将用户请求通过一定的路由算法重定向至最靠近用户的内容代理服务器上,并通过给定的资源选择策略从而达到内容的快速分发。内容路由主要分为两个部分:(1)资源路由算法,即通过监测当前网络各项性能指标与服务负载压力,选择最佳的内容代理服务器进行响应;(2)请求转发机制,即依据哪种策略对服务请求进行转发。典型的内容分发网络资源路由流程如图 4 所示:

(1)用户向请求解析服务器进行资源请求;(2)全局性能监测服务机群收集分析分布于互联网中的内容代理服务机群的服务质量以及网络链路各项性能参数;(3)节点选择服务器根据实时性能监视服务的反馈信息,动态调整选择算法与资源定位策略;(4)解析服务器通过资源定位结果,将用户请求转发至最佳的内容代理服务器;(5)内容代理服务器将数据发送至用户。

3.3.1 资源路由算法

请求路由属于应用层路由策略,其思想是根据应用层服务需求,对资源请求进行相应的转发。路由算法主要分为静态与动态两类:静态路由是指当内容分发网络部署后,对于用户的请求都根据给定的路由转发策略重定向至内容代理服务器,该算法不随网络状态与服务负载发生变化;动态路由是指实时监视网络参数与服务端负载,动态修改转发策略,尽可能保证当前服务维持在最佳状态。

静态路由策略通常依据用户的访问距离或跳数作为节点选择度量。文献[34]采用传输距离与服务器历史负载信息作为服务节点选择的依据,对请求进行转发。这种方法在性能上优于轮询,然而由于没有对链路进

行监测,对用户的服务质量并不高.

动态路由策略主要通过被动收集或主动探测方式对网络状态与用户访问质量进行监测,自适应修正路由转发机制. Pierre 等人开发的研究型内容分发网络 Globule 采用网络邻近性及周期性更新的传输距离动态

改变路由算法^[35]. 但由于网络探测的准确性较低且没有考虑流量对延迟的影响导致整体性能较低. 文献[36]使用改进的 BGP 协议将内容索引加入协议首部以提升查找效率.

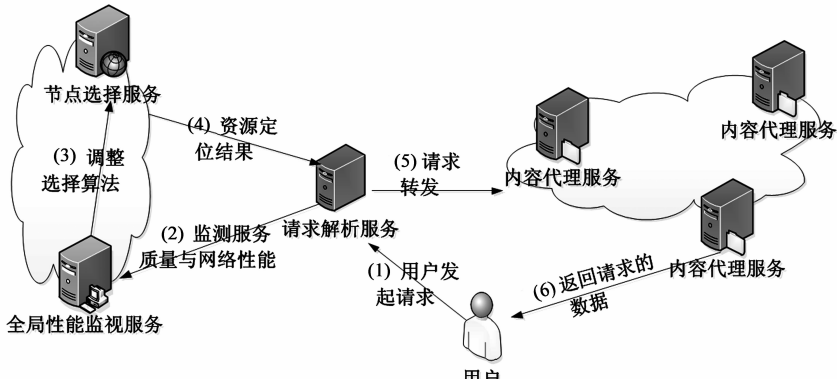


图4 内容路由流程

由于内容分发网络的商业化,工业界对资源路由算法也较为重视. 思科公司的分布式重定向器^[21] (Distributed Director) 把 AS (Autonomous System) 域内距离, AS 域间距离与端到端延迟作为动态转发算法的计算权值. Akamai 公司在多个层面对转发算法进行优化,包括内容代理服务器负载、传输路径上的带宽占用率、当前服务质量(可靠性,延迟等).

资源路由算法的优劣直接影响内容分发网络的服务质量,静态路由策略不具备网络感知能力,在性能上远低于动态策略. 动态策略由于其对各项度量指标进行实时监测,在线更新选择算法,大幅度提升内容分发网络的性能,但其开销与预测准确性还有待提升. 如何降低监测代价及提高对于网络状态的预测能力将是一个具有挑战性的研究课题.

3.3.2 请求转发机制

路由转发是资源路由中的另一项重要组成,其描述的主要是请求转发的实现机制. 本文在文献[37]的基础上,对请求转发机制进行了总结. 如图5所示,主要可分为基于协议转发、基于URL转发、基于泛播转发与基于P2P索引四类. 其中基于P2P索引的转发机制主要由P2P混合内容分发网络使用,单纯分发Web数据的内容分发网络并不采用该类方法.

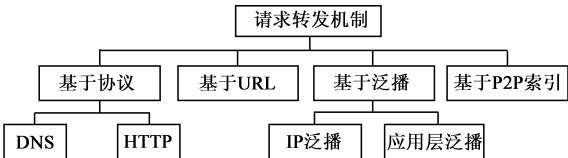


图5 内容分发网络转发机制分类

基于 DNS 的转发机制是当前应用最广泛的请求转发策略,在典型的 DNS 转发机制中,用户将所请求资源的 URL 发送至域名服务器,域名服务器根据当前资源路由由算法结果将请求转发至内容代理服务器. 互联网域名服务系统的层次结构也促使该策略成为内容分发网络商的首选. 文献[20]指出当前的内容分发网络商在分发网内部署多级域名解析服务器,通过层次结构的分流降低网络消耗与访问延迟. 然而,基于 DNS 转发机制由于 DNS 本身的安全威胁而可能导致服务故障. 基于 HTTP 的转发策略是使用 HTTP 协议首部字段对请求进行转发. 与 DNS 转发机制类似,HTTP 协议的通用性意味着该机制的易用性. 这种策略主要用于 Web 机群内的请求处理,如 HTTP 代理等.

基于 URL 的转发机制主要用于源服务器将请求转发至内容代理服务器的过程,该方法主要是将 Web 页面上的元素进行分类,将大量静态数据转发至内容代理服务器,动态交互数据则由源服务器返回给用户.

泛播转发策略可以分为 IP 层泛播与应用层泛播两类. Alzoubi 等人在文献[6]中重新审视了 IP 泛播技术在内容分发网络中的应用,指出 IP 泛播技术由于缺乏外界感知能力而无法提供较好的服务质量,并提出一种基于负载感知的 IP 泛播内容分发网络体系. 泛播技术的优势在于可以利用有限的 IP 地址空间获得良好的分发性能,然而随着内容代理服务器数量的增多,泛播路由表的变动开销迅速增大,使得维护代价过高. 文献[38]使用具备一定存储能力的路由器作为资源缓存节点,并采用多播技术提升网内内容分发效率.

P2P 混合的内容分发网络主要采用基于 P2P 索引转发策略对请求进行重定向. Rodriguez 等人在文献[3]

中认为互联网的研究重心在于如何更好的分发网络数据,并对 P2P 网络过去十年的进展进行简述,指出内容分发网络与 P2P 相结合是未来的重要研究课题.文献[39,40]对 P2P 混合的内容分发网络进行性能评价,并指出这种混合型内容分发网络可能由于 ISP 的干预而影响实际的分发质量. P2P 混合的内容分发网络的转发机制是由 P2P 的索引机制决定,主要分为中心化(Centralized directory)、分布式哈希表(Distribute Hash Table)与广播式(Broadcast)三类.在混合内容分发网络中,用户请求转发策略依赖于 P2P 网络,而且由于用户端性能远低于服务机群,这种转发策略大多作为商业化内容分发网络的辅助机制.

总体而言,请求转发策略是根据内容分发网络的结构决定,对于云内容分发网络体系,如何更加有效结合云的性能优势与 P2P 网络的连通性还有待进一步研究.

3.4 内容管理

内容管理是内容分发网络数据分发的基础组件,包括内容重组与缓存管理.内容重组是指将互联网数据分为可缓存与不可缓存两类,进而通过转发机制将请求分别发送至内容代理服务机群与源服务机群.缓存管理包括内容查询与内容更新两个方面.

内容重组依赖于该数据本身内在的特性.内容分发网络首先需要对内容进行分类,从而决定内容的路由策略.对于可以全缓存的数据,内容重组组件直接将其推送至内容代理服务器;对于部分缓存数据,则根据其缓存性进行分离,并按照数据类型分别存储于内容代理服务器.在源内容服务器视角上可将推送机制分为主动推(push-based)与被动拉(pulled-based)两种.主动推是指当源内容服务器上的数据更新时,同时将新数据推至内容代理服务器,被动拉是内容代理服务器根据缓存数据的生命期或当缓存未命中时向源服务器拉取数据.通常而言,网络内容提供商与内容分发网络运营商之间的交互主要采用主动推模式,而 ISP 建立的内容分发网络则是使用被动拉方式进行数据更新.

缓存管理是数据管理的传统研究课题.针对内容分发网络的缓存管理,可分为缓存更新策略与缓存定位算法两个重要部分.缓存更新策略主要可分成三类:(1)基于局部性,以 LRU(Least Recently Used)或其相似类为主;(2)基于流行度,以 LFU(Least Frequently Used)或其衍生类算法为主;(3)基于函数类,以对象空间大小,对象价值度等其他度量指标,构造对象评价函数,以 GDSF^[41,42]为主.在多媒体内容分发网络中,LFU 策略所达到的命中率优于 LRU,协作式缓存结构优于层次型缓存^[43].文献[44]以对象名称属性的概念分层为基础,通过模糊聚类找出同类对象修改的共同规律,提出了

一种以复制服务器中对象请求记录为依据的一致性算法.为了进一步适应互联网的异构性与数据的复杂性,内容分发网络的缓存更新机制应当更加着重研究多维度融合及动态适应的替换策略,提高内容分发的效率.

当前的数据定位算法主要分为基于查询和基于目录索引两种.前者还可细分为基于广播式查询^[45]与基于摘要式查询^[46],广播式查询的显著缺陷在于泛洪流量,而摘要式查询虽然避免了查询流量开销,但频繁的数据更新依旧可能造成性能瓶颈. Ni 等人在文献[47,48]中提出一种层次型半哈希(semi-hashing)定位机制,其考虑到哈希定位方式缺乏局部特性,将缓存空间分割为两部分:一部分数据组成一致性哈希空间,另一部分则由本地高频数据构成,从而提升本地命中率,且没有增加额外的通信开销.文献[49]提出一种资源分布树机制,对互联网中符合重尾分布(long-tail)的内容进行存储,并动态修改内容缓存服务器在分布树中的位置以降低资源副本部署代价.

4 内容分发网络性能评价

在标准的内容分发网络体系中,审计模块是影响其性能的关键之一,考虑到性能评价的复杂性以及相对独立性,本文将性能评价机制作为单独的一节进行叙述.由于内容分发网络在商业化上的巨大成功,性能评价机制也直接影响到用户的选择.而且在标准内容分发网络中,内容分发网络管理者需要对资源路由、数据分发、缓存状态进行收集分析,进而预测未来的网络状态及用户的访问趋势,对整个分发系统进行动态调整^[8]. Krishnamurthy 等人在文献[21]中对当前流行的内容分发网络进行了简要的对比.总体而言,内容分发网络的评价指标主要包括五个方面^[12,21,25]:(1)缓存命中率:在内容分发网络中,该命中率主要指对象数量的命中率而不是字节命中率,命中率越高表示当前的缓存机制越有效;(2)带宽节约率:指部署内容分发网络后,降低源服务端负载的比例;(3)平均时延:指用户通过内容分发网络获取到资源的平均延迟;(4)内容代理服务负载度:通过监视内容代理机群的服务质量(如 CPU 占用率、I/O 使用率与并发连接数等),判断服务器位置部署策略的优劣;(5)服务可靠性:主要指对丢包率的测量,在内容分发网络中,丢包率的高低可表示当前的网络状态以及服务机群的负载度.

内容分发网络性能指标主要可分为网络性能与服务端性能两个方面.内容分发网络管理者通过在服务端以及网络设备上收集分析系统日志,继而对当前的服务质量与运行机制进行评价.然而商业化的内容分发网络并没有公开对其系统的监测细节,因此本文主要对第三方对其网络性能评价指标与方法进行总结.

Su^[50]等人指出内容分发网络对互联网的流量以及传输路径的重定向在一定程度上改变了互联网的逻辑结构,并从 PlanetLab 中的 140 个节点上对 Akamai 进行长达两个月的测量,发现采用 Akamai 推荐的路径访问目标的延时低于实际路由的情况超过 50%,并且 Akamai 采用多维度方式定位边缘服务器,从而避免本地访问过热.

Huang^[51]等人分析了两种不同部署模式的内容分发网络 Akamai 与 LimeLight,将访问时延划分为内部 DNS 解析延时与 RTT(Round Trip Time)延时,发现 LimeLight 的在这两部分的延时均高于 Akamai,并提出通过相互学习部署策略以提升服务性能的方法.

Mao 等人在文献[52]中对内容分发网络的用户与内部 DNS 服务器的距离进行了细致分析,发现大多数用户并不邻近于内部 DNS 服务器(包括 AS 邻近,地理邻近,网络跳数少),从而认为单纯依赖 DNS 进行请求转发的策略对降低访问延迟的提升有限.

还有一部分文献通过构建研究性内容分发网络^[53,54]或使用网络模拟器^[55,56]对内容分发网络进行模拟分析.虽然通过模拟方法可以通过导入真实的网络流量进行分析,但是由于缺乏实际环境中不可控噪声对系统的影响,从而无法真实刻画内容分发网络系统的实际性能.

综上,当前对内容分发网络的性能评价方法主要分为实际网络探测与网络模拟两种,前者主要关注商业化内容分发网络的拓扑构成,边缘服务器部署位置,路由转发策略,试图为内容分发网络使用方提供系统性的指导,同时为其设计者提供改进参考.后者在模拟环境下对内容分发的策略进行评价,从而为分发机制的改进提供理论依据.当前对于内容分发网络的评价研究还处于起步阶段,如何更加全面的探测其网络构成以及建立完备的性能评价体系是研究内容分发网络性能的重要课题.

5 未来挑战

互联网的开放性与尽力而为(best-effort-based)的原则是其迅速发展的根本因素.然而基于简单的连接控制保证已经无法满足当前人们对互联网服务的可靠性、可用性、安全性与高可扩展性的需求.虽然 P2P 网络的高共享特性在大幅提高了互联网数据的可用性,但内容分发网络进一步提升互联网的高可靠性与高分发能力.

内容分发网络技术经过十余年的发展,无论在网络基础设施还是在网络流量上已经在很大程度上改变了早期的互联网结构.鉴于前述的分析,本文预期内容分发网络的未来发展趋势如下:(1)在拓扑结构上进一

步与 P2P 网络相结合,更有效利用端到端的带宽资源,提高多媒体流量的传输能力;(2)增强与 ISP 的合作关系,利用其拓扑优势以及协助其改善域内流量工程,降低域间流量结算成本,同时提高 ISP 域内用户的服务质量;(3)将内容管理服务与云计算平台进行融合,利用云平台的计算资源与弹性管理机制,降低服务能耗,提高内容分发网络的运营灵活性;(4)由于内容分发网络的复杂异构性,需要建立内容分发网络的安全传输标准与运营规范,提升整体的鲁棒性与安全性.

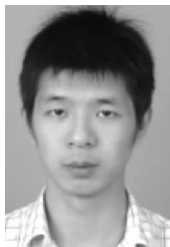
参考文献

- [1] Pallis G, Vakali A. Insight and perspectives for content delivery networks[J]. Communications of the ACM, 2006, 49(1): 101 - 106.
- [2] Karagiannis T, Rodriguez P, Papagiannaki K. Should internet service providers fear peer-assisted content distribution? [A]. Proceedings of the 5th ACM SIGCOMM conference on Internet Measurement[C]. Berkeley, USA, 2005. 6 - 6.
- [3] Rodriguez P, Tan S M, Gkantsidis C. On the feasibility of commercial, legal p2p content distribution [J]. ACM Sigcomm Computer Communication Review, 2006, 36(1): 75 - 78.
- [4] Bavier A, Feamster N, Huang M, et al. In VINI Veritas: Realistic and controlled network experimentation[A]. Proceedings of ACM Sigcomm[C]. Pisa, Italy, 2006. 3 - 14.
- [5] Wang L, Park K S, Pang R, et al. Reliability and security in the CoDeeN content distribution network[A]. Proceedings of the USENIX Annual Technical Conference [C]. Boston, USA, 2004. 14 - 14.
- [6] Alzoubi H A, Lee S, Rabinovich M, et al. Anycast CDNS revisited[A]. Proceeding of the 17th International Conference on World Wide Web[C]. Beijing, China, 2008. 277 - 286.
- [7] Jones S T. Iptv Delivery architecture[A]. ITU-T IPTV Global Technical Workshop[C]. Seoul, Korea, 2006. 1 - 1.
- [8] Day M, Cain B, Tomlinson G, et al. RFC 3466 (A Model for Content Internet Working) [S]. Internet Engineering Task Force, February 2003. www.ietf.org/rfc/rfc3466.txt.
- [9] Barbir A, Penno R, Chen R, et al. RFC3835 (An Architecture for Open Pluggable Edge Services (OPES))[S]. Internet Engineering Task Force, August 2004. www.ietf.org/rfc/rfc3835.txt.
- [10] Cooper I, Melve I, Tomlinson G. RFC3040 (Internet Web Replication and Caching Taxonomy)[S]. Internet Engineering Task Force RFC 3040, January 2001. http://www.ietf.org/rfc/rfc3040.txt.
- [11] Barbir A, Batuner O, Beck A, et al. RFC3838 (Policy, Authorization, and Enforcement Requirements of the Open Pluggable Edge Services (OPES))[S]. Internet Engineering Task Force RFC 3838, August 2004. www.ietf.org/rfc/rfc3838.txt.

- [12] Spagna S, Liebsch M, Baldessari R, et al. Design principles of an operator-owned highly distributed content delivery network[J]. IEEE Communications Magazine, 2013, 51(4): 132 – 140.
- [13] Yin H, Lin C, Zhang Q, et al. TrustStream: A secure and scalable architecture for large-scale internet media streaming[J]. IEEE Transaction on Circuits and Systems for Video Technology, 2008, 18(12): 1692 – 1702.
- [14] Yin H, Lin C, Qiu F, et al. CASM: A content-aware protocol for secure video multicast[J]. IEEE Transaction on Multimedia, April 2006, 8(2): 270 – 277.
- [15] 尹浩, 詹同宇, 林闯. 多媒体网络: 从内容分发网络到未来互联网[J]. 计算机学报. 2012, 35(6): 1120 – 1131.
Yin H, Zhang T Y, Lin C. Multimedia networking: from content delivery networks to future internet[J]. Chinese Journal of Computers, 2012, 35(6): 1120 – 1131. (in Chinese)
- [16] 杨戈, 廖建新, 朱晓民, 樊秀梅. 流媒体分发系统关键技术综述[J]. 电子学报, 2009, 37(1): 137 – 145.
Yang G, Liao J X, Zhu X M, Fan X M. Survey of key technologies of the distribution system for streaming media[J]. Acta Electronica Sinica. 2009, 37(1): 137 – 145. (in Chinese)
- [17] Gao L, Dahlin M, Nayate A, et al. Improving availability and performance with application-specific data replication [J]. IEEE Transactions on Knowledge and Data Engineering, 2005, 17(1): 106 – 120.
- [18] Day M, Cisco, Cain B, et al. RFC3466(A Model for Content Internet Working (CDI)) [S]. Internet Engineering Task Force RFC, 2003. <http://tools.ietf.org/html/rfc3466>.
- [19] Jung Y, Krishnamurthy B, Rabinovich M. Flash crowds and denial of service attacks: characterization and implications for cdns and web sites[A]. Proceedings of the 11th International World Wide Web Conference[C]. Hawaii, USA, 2002. 293 – 304.
- [20] Lampe C, Resnick P. Slash(dot) and Burn: Distributed moderation in a large online conversation space[A]. Proceedings of the SIGCHI Conference on Human Factors in Computing Systems[C]. Vienna, Austria, 2004. 543 – 550.
- [21] Krishnamurthy B, Wills C Zhang Y. On the use and performance of content distribution networks[A]. Proceedings of the 1st ACM SIGCOMM Workshop on Internet Measurement [C]. San Francisco, USA, 2001. 169 – 182.
- [22] Lazar I Terill W. Exploring content delivery network[J]. IEEE IT Professional, 2001, 3(4): 47 – 49.
- [23] Akamai Technologies Inc. <http://www.akamai.com>. [EB/OL]. 2012.
- [24] CDN 发展现状与趋势[R]. 中国工信部电信研究院. 2009.
- [25] Vakali A, Pallis G. Content delivery networks: status and trends[J]. IEEE Internet Computing, 2003, 7(6): 68 – 74.
- [26] Plagemann T, Goebela V, Mauthe A, et al. From Content distribution networks to content networks-issues and challenges [J]. Computer Communications, 2006, 29(5): 551 – 562.
- [27] Yin H, Liu X N, Min G Y, et al. Content delivery networks: a bridge between emerging applications and future ip networks [J]. IEEE Network, 2010, 24(4): 52 – 56.
- [28] Pathan A M K, Buyya R. A taxonomy and survey of content delivery networks[R]. Technical Report, GRIDS-TR-2007-4. The University of Melbourne, Australia, Feb 2007.
- [29] Saroiu S, Gummadi K P, Dunn R J, et al. An analysis of internet content delivery systems [J]. ACM SIGOPS Operating Systems Review, 2002, 36(SI): 315 – 328.
- [30] LimeLight Networks[EB/OL]. <http://www.limelight.com/> 2012.
- [31] Garey M R, Johnson D S. Computers and Intractability: A Guide to the Theory of NP-Completeness[M]. W H Freeman and Co, New York, USA, 1979.
- [32] Ni J, Tsang D H K. Large-scale cooperative caching and application-level multicast in multimedia content delivery networks[J]. IEEE Communications Magazine, 2005, 43(5): 98 – 105.
- [33] Jiang H, Wang Z, Wong A K, et al. A replica placement algorithm for hybrid cdn-p2p architecture[A]. Proceedings of 15th International Conference on Parallel and Distributed Systems (ICPADS)[C]. Shenzhen, China, 2009. 758 – 763.
- [34] Sivasubramanian S, Szymaniak M, Pierre G, et al. Replication of Web hosting systems[J]. ACM Computing Surveys, 2004, 36(3): 291 – 334.
- [35] Pierre G, Vansteen M. Globule: A collaborative content delivery network[J]. IEEE Communications, 2006, 44(8): 127 – 133.
- [36] Field B, Doorn J V, Hall J. Integrating routing with content delivery networks[A]. IEEE Conference on Computer Communications Workshops (INFOCOM 2012 WORKSHOPS) [C]. Orlando, Florida, USA, 2012. 292 – 297
- [37] Podlipnig S, Böszörményi L. A survey of Web cache replacement strategies[J]. ACM Computing Surveys, 2003, 35(4): 374 – 398.
- [38] Kim J, Lee G, Choi J. Efficient multicast schemes using in-network caching for optimal content delivery[J]. IEEE Communications Letters, 2013, 99: 1 – 4
- [39] Lu Z H, Gao X H, Huang S J, et al. Scalable and reliable live streaming service through coordinating CDN and P2P[A]. 2011 IEEE 17th International Conference on Parallel and Distributed Systems[C]. Tainan, Taiwan, 2011. 581 – 588.
- [40] Mori T, Kamiyama N, Harada S, et al. Improving deployability of peer-assisted cdn platform with incentive [A]. IEEE GLOBECOM[C]. Hawaii, USA, 2009. 1 – 7.
- [41] Cherkasova L, Ciardo G. Role of aging, frequency and size in web caching replacement strategies[A]. Proceedings of the 9th International Conference on High-Performance Computing and

- Networking[C]. Amsterdam, The Netherlands, 2001. 114 – 123.
- [42] Patil J B, Pawar B V. Trace driven simulation of GDSF# and existing caching algorithms for web proxy servers[A]. Proceedings of the 9th WSEAS International Conference on Data Networks, Communications, Computers[C]. Faro, Portugal, 2009. 378 – 384.
- [43] Chejara U, Chai H K, Cho H. Performance comparison of different cache-replacement policies for video distribution in CDN[J]. Lecture Notes in Computer Science, 2004, 3079: 921 – 931.
- [44] 陈妍, 李增智, 王云岚, 廖志刚. CDN 网络中对象一致性算法研究[J]. 电子学报, 2005, 33(8): 1480 – 1483.
Chen Y, Li Z Z, Wang Y L, Liao Z G. Study of object consistency algorithm in CDN[J]. Acta Electronica Sinica, 2005, 33(8): 1480 – 1483. (in Chinese)
- [45] Wessels D, Claffy K. RFC2186 (Internet Cache Protocol (ICP) version 2) [S]. Internet Engineering Task Force RFC 2186, September 1997. <http://www.ietf.org/rfc/rfc2186.txt>.
- [46] Rousskov A, Wessels D. Cache Digests[J]. Computer Networks and ISDN Systems, 1998, 30(22 – 3): 2155 – 2168.
- [47] Ni J, Tsang D H K. Large scale cooperative caching and application-level multicast in multimedia content delivery networks[J]. IEEE Communications, 2005, 43(5): 98 – 105.
- [48] Ni J, Tsang D H K, Yeung I S H, et al. Hierarchical content routing in large-scale multimedia content delivery network [A]. Proceedings of IEEE International Conference on Communications[C]. Anchorage, Alaska, USA, 2003. 854 – 859.
- [49] Yasuhiro M. An Overlay Architecture of Global Inter-Data Center Networking for Fast Content Delivery[A]. IEEE International Conference on Communications (ICC) [C]. Kyoto, Japan, 2011: 1 – 6.
- [50] Su A J, Choffnes D R, Kuzmanovic A. Drafting behind aka-mai: inferring network conditions based on cdn redirections [J]. IEEE/ACM Transactions on Networking, 2009, 17(6): 1752 – 1765.
- [51] Huang C, A. Wang, J Li, et. al. Measuring and evaluating large-scale CDNs[A]. Internet Measurement Conference[C]. Vouliagmeni, Greece, 2008. 15 – 29.
- [52] Mao Z M, Cranor C, Douglass F, et al. A precise and efficient evaluation of the proximity between web clients and their local dns servers[A]. Proceedings of the General Track of the annual conference on USENIX Annual Technical Conference [C]. Monterey, California, USA, 2002. 229 – 242.
- [53] Globule: the Open-Source Content Distribution Network[EB/OL]. <http://www.globule.org/index.html>. 2012.
- [54] CoDeeN: A CDN on PlanetLab[EB/OL]. <http://codeen.cs.princeton.edu/>. 2012.
- [55] Network Simulators[EB/OL]. <http://ee.lbl.gov/kfall/net-sims.html>. 2012.
- [56] CDNSim: A Content Distribution Network Simulator [EB/OL]. <http://oswinds.csd.auth.gr/~cdnsim/>. 2012.

作者简介



李 乔(通信作者) 男, 1984 年出生, 福建建瓯人, 哈尔滨工业大学网络与信息安全研究中心博士生, 主要研究方向为分布式计算, 网络安全.

E-mail: liqiao84@hit.edu.cn



何 慧 女, 博士, 哈尔滨工业大学网络与信息安全研究中心副教授, 主要研究方向为网络计算, 网络安全.