

# 覆盖网体系结构及应用研究

杜丽娟<sup>1,2</sup>, 余镇危<sup>2</sup>

DU Li-juan<sup>1,2</sup>, YU Zhen-wei<sup>2</sup>

1. 华北科技学院 电子信息工程系, 北京 101601

2. 中国矿业大学 机电与信息工程学院, 北京 100083

1. Department of Electronic Information Engineering, North China Institute of Science and Technology, Beijing 101601, China

2. School of Mechanical, Electronic and Information Engineering, China University of Mining & Technology, Beijing 100083, China

E-mail: dulijuan1976@ncist.edu.cn

DU Li-juan, YU Zhen-wei. Survey of overlay network architecture and application. *Computer Engineering and Applications*, 2009, 45(28): 102-104.

**Abstract:** Overlay network, which is based on Internet, can enhance the performance of network and deploy new applications flexibly by constructing a virtual logic network. This paper summarizes two kinds of influential Overlay architectures, introduces several applications based on overlay network and their key techniques, hot research area, then explores the utilization of overlay network in mobile grid and its direction.

**Key words:** overlay network; P2P network; multicast; Content Delivery Network(CDN); mobile grid

**摘 要:** 覆盖网以 Internet 为基础, 通过构建虚拟的逻辑网络可以提高网络性能, 灵活地部署新应用, 是一种被广泛采用的网络技术。总结分析了目前两种有影响的覆盖网体系结构, 覆盖网的几种典型应用, 及各自的关键技术和研究热点; 最后提出了覆盖网思想在移动网格中应用的探索, 及未来的发展趋势。

**关键词:** 覆盖网络; 对等网络; 组播; 内容分发网络; 移动网络

**DOI:** 10.3778/j.issn.1002-8331.2009.28.030 **文章编号:** 1002-8331(2009)28-0102-03 **文献标识码:** A **中图分类号:** TP393

Overlay 网络是一种构造网络的方法, 与特定技术、特定层次无关, 是建立在一个或者多个已存在网络之上的网络, 通过增加额外的、间接的、虚拟的层, 来改善下层网络部分领域中的一些属性, 提高网络的性能。Overlay 网络上的应用扩展不需要全网范围内组件的支持, 增加新应用也无需网络内部节点的改变。这使得 Overlay 网络成为能让现实网络满足新型应用需求的过渡方案。因此对 Overlay 网络的研究成为近年来的研究热点。覆盖网络的出现实质是对网络资源利用在结构和分布上的调整, 但目前 Overlay 网络的研究多局限在对特定应用的支持上, 不同的应用需建立不同的 Overlay 网络, 缺乏服务定制的能力。

目前发展比较迅速和应用比较广泛的覆盖网络有对等网络、应用层组播和内容分发网络等。对等网络的目标主要是支持大量的节点相互通信; 应用层组播研究集中于视频会议系统、媒体流的分发系统(如视频广播)和订阅/分发系统等, 它主要用于实时的多媒体传输; CDN 偏重于非实时的、文件形式的传输。

## 1 覆盖网体系结构

### 1.1 弹性覆盖网络 RON

在目前的 Internet 中, 由于自治系统是独立管理和配置的,

路由域之间不会共享内部链路信息, 故障也是彼此独立的, 在发生路径失效或性能降低时, BGP 协议通常需要几分钟才能够检测并恢复。RON<sup>[1]</sup>可以使分布式应用检测到路径的失效和周期性的性能降低现象, 并能够在 20 s 之内从中恢复。具体来讲, RON 的设计目标包括以下几个方面: (1) 快速的失效检测和恢复; (2) 路由和路径选择可以和应用紧密集成; (3) 灵活的策略路由。

$N$  个节点的 RON 中, 每个节点都有  $N-1$  条虚拟链路, 因此节点形成网状覆盖网络。RON 通过主动探测和被动监听相结合的方式得到链路的度量参数, 通过分布式路由协议来交换各种链路度量信息, 从而每个节点都获得拓扑结构中全部虚拟链路信息, 并使用该信息进行路径选择。RON 体系结构模型如图 1 所示。

数据转发过程中, RON 客户端通过 Conduit 接口和 RON 交互。接收分组的第一个 RON 节点对分组分类以决定分组需要使用的路径的类型(如低延迟或高吞吐率), 并对分组进行标记以简化后续节点的转发, 然后从拓扑表中寻找一条路径转发分组。后续 RON 节点根据目的地址和标记决定下一跳。最后一个 RON 节点将分组交给客户程序。

RON 中, 每个节点要维护  $N*(N-1)$  条虚拟链路的度量信

基金项目: 教育部博士点基金项目(No.20030290003)。

作者简介: 杜丽娟(1976-), 女, 博士生, 讲师, 主要研究领域为移动网格; 余镇危(1942-), 男, 博士生导师, 教授, 主要研究领域为计算机体系结构。

收稿日期: 2009-06-04

修回日期: 2009-07-29

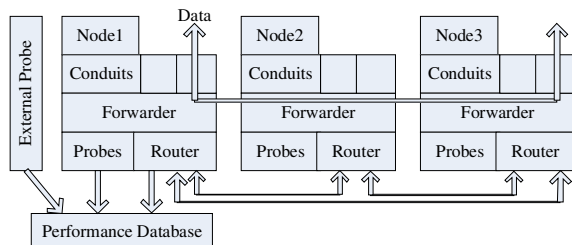


图1 RON 体系结构模型

息,而端主机的处理能力和存储容量有限,因此它的可扩展性非常差,只适用于非常小规模的应用。

## 1.2 服务覆盖网络 SON

目前 Internet 由多个 ISP 互联而成,每个 ISP 通常只关心自己维护的网络的性能并只对自己的用户提供服务保证,跨越多个 ISP 的端到端应用的性能难以得到保证。因此 SON<sup>[2]</sup>概念的提出用于在现有的 Internet 上提供端到端的服务质量保证。其主要思想是 SON 从私人网络那里通过双向服务级别合约购买具有一定 QoS 保证的带宽,从而在现有的数据传输网络之上建立逻辑的端到端服务传输结构。通过服务协定,用户直接向 SON 付费,从而使用 SON 提供的增值服务。SON 体系结构模型如图 2 所示。

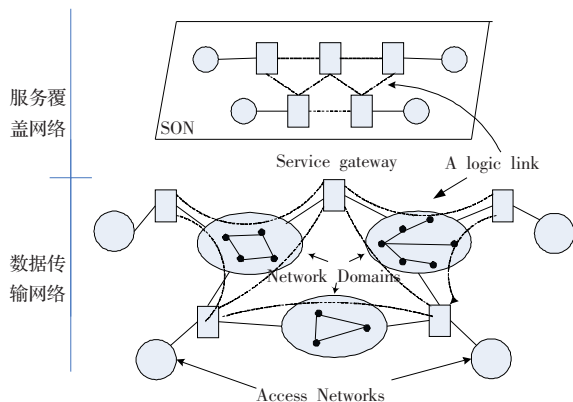


图2 服务覆盖网络体系结构模型

SON 通过服务网关相互连接,服务网关提供面向特定服务的数据转发和控制功能。两个服务网关之间的逻辑连接由底层的 ISP 提供并满足带宽和其他的服务质量保证,这些服务质量保证通过双向的 SLA 定义。SON 实际上把 RON 模型推广到更加通用的情况,但 SON 的实现也更加困难(复杂),它不但涉及到相关的网络技术问题,还涉及到网络供应商之间的协作问题。

## 2 典型覆盖网应用

### 2.1 对等网络

在对等网络体系结构中,位于网络边缘的众多计算机构成统一的服务器为客户提供文件下载等服务。每台计算机既是服务器又是客户机,对等网络本身构成一个覆盖网络,对等体需要自己进行服务器发现、选择到其他对等体的路由等。根据覆盖网拓扑结构的不同对等网络可分成三类:集中式对等网络,全分布无结构对等网络,结构化对等网络。

#### 2.1.1 集中式对等网络

集中式对等网络使用集中的服务器存放所有 P2P 节点共

享资源的信息;P2P 节点通过查询集中服务器找到存放所需资源的 P2P 节点位置信息,再直接到该节点去获取信息。集中式对等网络拓扑结构是星型结构,服务器是整个网络的核心,因此还不能称之为完全的覆盖网络,因为它还离不开服务器的支持,如路由、自组织、自适应性、容错性等要依靠服务器的监控,但服务器的作用相对于传统网络已经被大大弱化<sup>[3]</sup>。

#### 2.1.2 全分布无结构对等网络

全分布无结构对等网络没有固定、严格的拓扑结构,而是一个随机生成、松散组织的普通图,理论上这张图可以是任何形状的。但研究发现,虽然无结构的网络拓扑不严格遵循某种形状,但总是符合一定的规律——小世界模型或者幂律模型。

无结构对等网络的资源定位方法以洪泛法为基础,路由效率低并存在查询局部性。研究人员从覆盖网络构建的角度提出了不同的优化机制。Wouhaybi 提出协议,将对等节点构建成节点度数服从幂律分布的短直径覆盖网络拓扑 Phenix,思想是节点之间的距离越短,定位请求资源的延迟越小<sup>[4-5]</sup>。

#### 2.1.3 结构化对等网络

结构化对等网络是目前 P2P 领域的研究热点,其最大的特点在于都有严格的覆盖网拓扑结构,系统的其他机制如分布式哈希表、路由与定位、负载均衡、自适应、自组织、容错性等,都是基于特定的拓扑结构。比较有代表性的覆盖网拓扑结构包括带环环 chord、多维空间拓扑结构 CAN、超立方体结构 Tapestry、混合式结构 Pastry。

所有结构化 P2P 网络都使用分布式哈希表将节点、数据对象映射到覆盖网中,路由和定位方式取决于覆盖网拓扑结构和路由表结构,通常维护一个比较小的路由表,采用分布式、局部性的贪心路由算法,逐步缩小当前节点与目的节点之间的 ID 差异<sup>[6]</sup>。

从目前的研究发展来看,P2P 的研究重点已经从核心机制转向增强机制,目前的研究热点集中在语义模糊查询、容错性、拓扑意识与一致性、声誉和安全性问题。从更宽的范围来看,P2P 技术和其他技术的结合,P2P 领域和其他领域(如网格计算、无线网络)的融合,是一个必然的趋势。

## 2.2 应用层组播

应用层组播<sup>[7-8]</sup>体系结构包括两类:一是对等型。组成员节点都是平等的,通过自组织算法组织成控制网格和数据分发树,每个节点仅维护自身参与组的状态信息,组播服务功能以软件形态集成在参与会话的节点中;二是代理型。组播服务功能由增值服务提供商根据一定策略在互联网特定位置部署代理节点完成,端节点与自己拓扑相邻的代理节点相连,通过单播收发报文,代理节点类似组播路由器。

#### 2.2.1 应用层组播协议

根据构建分发树的策略分成集中式和分布式。集中式协议假定有一个服务器知道成员之间的拓扑关系,然后基于某种优化目标运行最小延迟、最大带宽、最小代价之类的算法来计算分发树,随后将计算结果发布给相应的节点,特点是简单、扩展性不好、单点故障;分布式协议中,每个组成员运行应用层组播协议,分布地构造分发树。为了适应系统的动态性、保证组播传输的效率,分布式协议引入控制拓扑和数据拓扑。控制拓扑(mesh)中节点之间存在多条路径,负责监视节点变化、维护网络连通性;数据拓扑指分发树。



### 2.2.2 覆盖网络路由算法

应用层组播算法是学术界的研究热点,研究人员提出的比较有影响的算法包括:Narada 是最早在应用层提供组播服务的方法之一;为了提高组播的可扩展性,Maryland 大学的 Suman Banerjee 提出了 NICE 应用层组播结构;针对 NICE 中处于中心的节点负担较大的问题,Florida 大学的 Tran 等人提出了 ZIGZAG 方法;Washington 大学的 Shi 等人提出了 ALMI 组播覆盖网络体系结构,讨论了覆盖网络路由目标和策略。

### 2.3 内容分发网络

CDN 由地理上分布的若干内容服务器、内容路由器、以及相应的管理软件构成,通过 Web 缓存、预取、复制、负载均衡和请求重定向等关键技术,基于 Internet 构筑一个广域覆盖(overlay)的内容传送网络,从带宽、响应速度、扩展性、可靠性等方面极大地改善了网络性能。CDN 技术已经在 Internet 上广泛应用并已成为网络利润的增长点。

CDN 的实现需要依赖多种网络技术支持,其关键技术主要有负载均衡、内容分发、分布式存储和内容管理技术等。(1)CDN 负载均衡系统实现 CDN 的内容路由功能,其作用是将用户的请求导向整个 CDN 网络中的最佳节点。负载均衡系统是整个 CDN 的核心,负载均衡的准确性和效率直接决定了整个 CDN 的效率和性能。(2)内容分发包含从内容源到 CDN 边缘的 POP 节点的过程,主要采用两种内容分发方式:PUSH 和 PULL。在实际的 CDN 系统中,PUSH 方式适合内容访问比较集中的情况,如热点的影视流媒体内容;PULL 方式适合内容访问较分散的情况。(3)分布式存储技术——对于 CDN 系统而言,需要考虑两类内容存储问题:内容源的存储,通常采用海量存储架构;内容在 POP 节点中的存储,需要考虑功能、性能两方面因素。(4)内容管理主要针对一个 POP 节点进行。内容管理的目标是提高内容服务的效率和本地节点的存储利用率。

## 3 覆盖网在移动网格中的应用探索

### 3.1 移动网格的发展背景

移动通信是近几年通信各领域中最发展最快的领域之一。截至 2008 年 3 月底,我国手机用户数超过 5.74 亿户。国际电信联盟(ITU)2008 年 9 月 27 日表示,到 2008 年年底全球手机用户总人数将达 40 亿。并且手持设备的配置也越来越高端化(处理器 400 MHz 以上,内存 90 M 以上)。以上数据足以说明移动设备资源无比巨大并呈上升趋势。

同时随着移动无线网络系统的快速发展,用户在任何地点、任何时间都可以访问全球网络资源。这意味着除了静态节点外,网格系统也应考虑把移动节点包含在内,这种结合所产生的技术就称为“移动网格计算”。它实际上是传统网格计算在无线计算环境中的扩展,它以无缝、透明、安全、有效的方式支持移动用户和资源,是无线技术与网格计算这两种新技术的融合<sup>[9]</sup>。移动网格要解决的问题是移动设备如何获取网格资源,以及自身的资源如何进行共享。

### 3.2 基于覆盖网的移动网格

移动网格是网格计算的进一步发展。网格起源于科学领域的高性能计算,已有的网格应用主要是将大型机聚合起来,节点数目少且节点之间的关系相对固定,直接基于现有的物理网络进行研究和构造,需要现有的网络层协议的支持,其兼容性不好,而且其扩展性和动态自适应性不高。

随着网络技术和计算技术的发展,人们对于网格技术的观点也发生改变——组成网格系统的资源不再局限于某几家科研机构的资源互联,而是面向广域分布的性能差别巨大的终端系统(各种移动设备、个人计算机、高性能服务器、巨型机、机群系统等),这些资源的异构性、分布性和动态性将变得更加复杂。尽管现有网格技术在一些重大科学领域取得了成功,但是在商用和民用的发展道路上步履缓慢,究其原因还是还没有恰当的网格应用能够把普通的移动设备(蜂窝电话、PDA、膝上型电脑)和个人计算机整合到网格环境中<sup>[10]</sup>。如果把网格节点扩展到普通计算机和移动设备,那么网格所连接的节点数量剧增,且这些节点加入或离开网格的频率更为频繁,尤其是移动设备的移动性会给网格技术发展提出挑战性的难题<sup>[11]</sup>。这些问题是网格要面对的新的考验,它必须能承受更大的规模,更不稳定的环境,与更沉重的管理负担。

基于覆盖网来发展传统网格技术是一种可行的方法。Overlay Network 擅长于隐藏底层物理网络的动态变化,提供一个虚拟的逻辑网络,满足高层应用的特定需求,例如服务注册与查找、任务调度、安全与管理等等。根据网格应用需求的特点,来构造 Overlay Network 的结构,建立一个网格节点交互的虚拟的通信基础,并且把底层的网格资源进一步地聚集,实现运行环境与物理资源的隔离,屏蔽物理网络的异构性,增强网格系统对物理网络的分布、自治、动态以及演化特性的适应能力。以 Overlay Network 为基础研究网格技术,使问题得到一定程度的简化,易于兼容现有的网络和系统,从而在动态异构的物理网络环境上实现大规模网格应用。

## 4 结束语

这几种应用由于出发点不同形成了各自的研究领域:对等网络的初衷是利用网络边缘的存储、计算等各种资源,主要是支持大量的节点相互通信,如文件共享、数据搜索等,研究的重点是如何提供更有效率的分布式哈希表查找机制;应用层组播研究集中于视频会议系统、媒体流的分发系统(如视频广播)和订阅/分发系统等,它主要用于实时的多媒体传输,研究的重点是针对不同应用的高效的组播协议;CDN 偏重于非实时的、文件形式的传输;移动网格和网格一样,目的是实现资源的有效聚合、广泛共享,相比而言,移动网格聚合了形式更多、地理范围更广的资源。

但是这几个领域的研究也不是完全相互独立、互不相交的。例如有些应用层组播方案也是建立在分布式哈希路由查找机制的基础上,从而如同结构化对等网络一样,具有良好的可扩展性;同样也存在着拓扑一致性问题需要解决。从这个角度讲,这种类型的应用层组播就是对等网络应用的一个子领域;在安全、覆盖网分割、容错性等问题研究上,两者都可以相互借鉴。内容分发网络也可以基于 P2P 技术构建,从而不需要在网络上额外部署服务节点;移动网格的研究正处于起步阶段,其关键技术的研究也能借鉴覆盖网较成熟的理论,如服务的注册和查找可以借鉴结构化对等网络的资源定位技术。

由于覆盖网具有扩展性好、灵活性、健壮性等特点,使得对等网络、内容分发网络、组播覆盖网等应用得到迅速发展;还有一些新的应用也基于覆盖网来研究,如移动网格。可以预见,基于 Internet 的覆盖网应用仍将体现强大的生命力。

(下转 108 页)