

分布交互仿真 数据传输关键技术

Key Technology of Distributed Interactive
Simulation Data Transmission

■ 刘 鹏 著



國防工業出版社
National Defense Industry Press

分布交互仿真数据传输 关键技术

Key Technology of Distributed Interactive
Simulation Data Transmission

刘 鹏 著

国防工业出版社

· 北京 ·

内 容 简 介

本书主要介绍分布交互仿真数据传输的基本理论和相关关键技术。本书分为基础篇、算法篇和应用篇。基础篇包括第1~2章,读者可以了解相关领域的研究现状,并且通过具体的实例了解建立一个分布交互仿真系统的数据传输平台的具体步骤,以及其中可能存在的问题与解决方法。算法篇主要介绍分布交互仿真数据传输系统扩展性技术,包括第3~5章。通过这几章,读者可以了解构建一个分布交互仿真数据传输系统中涉及的关键技术及主要算法,例如数据过滤算法、丢包恢复算法及拥塞控制算法等。应用篇为第6章,对分布交互仿真数据传输系统运行平台的实现方法和测试方法进行介绍,以实际的例子帮助读者迅速理解掌握分布交互仿真数据传输的关键技术方法。

本书以分布式仿真数据传输平台为主,同时介绍了相关理论、技术与应用,内容全面,循序渐进,可以作为分布式仿真研究与应用开发人员的参考手册,亦可作为高等院校计算机、仿真、自动控制相关专业学生的教学参考用书。

图书在版编目(CIP)数据

分布交互仿真数据传输关键技术/刘鹍著. —北京:
国防工业出版社, 2015. 12
ISBN 978-7-118-10591-9

I. ①分… II. ①刘… III. ①计算机仿真—数据
传输—研究 IV. ①TP391.9

中国版本图书馆 CIP 数据核字(2015)第 296907 号

※

国防工业出版社出版发行
(北京市海淀区紫竹院南路 23 号 邮政编码 100048)
北京嘉恒彩色印刷有限责任公司
新华书店经售

*

开本 880×1230 1/32 印张 4½ 字数 150 千字
2015 年 12 月第 1 版第 1 次印刷 印数 1—1500 册 定价 38.00 元

(本书如有印装错误,我社负责调换)

国防书店: (010)88540777
发行传真: (010)88540755

发行邮购: (010)88540776
发行业务: (010)88540717

前言

科学技术的发展,使我们在许多领域不断面临一些前所未有的复杂问题,这些问题很难进行现实环境下的实验论证,因此计算机仿真技术,特别是随着网络技术发展起来的分布交互仿真技术,越来越受到许多行业的重视。

分布交互仿真中实体间信息交换较为频繁,为了及时跟踪实体的状态和外部世界的变化,实体之间以及实体与外部世界(如地形数据库等)之间要及时交换数据(包括实体状态、交互事件、地形数据等),因此数据的传输与处理需要较高的网络带宽和处理能力。另外,在一些大型应用系统中可能存在大量的仿真实体,甚至要求能支持数万个仿真实体的实时交互。分布交互仿真中具有明显的多对多的组传输特性,各个主机节点都不停地产生仿真实体的数据,这些数据从不同的主机被发送到所有需要接收的那些主机。因此,组播技术对大规模分布交互仿真应用有着重要的意义。

基于组播技术实现分布交互仿真是分布式虚拟现实领域的重要研究内容,而跨广域网的分布交互仿真应用需求日益增加。与点对点、服务器转发或广播等通信机制相比,组播技术更适合于多对多的通信,具有高吞吐量、低开销的特点。但 IP 组播是不可靠的传输方式,且目前 Internet 上路由器对组播的支持不够普及,IP 组播不能跨广域网,影响了其应用发展,由此产生了应用层组播技术。相对于 IP 组播,应用层组播在传输效率上存在缺陷。针对分布交互仿真中多对多组通信、组成员动态变化、实时性要求高等特点,将二者恰当地结合起来,利用它们的特点实现分布交互仿真的广域网组播传输是本书所介

绍的核心内容。

本书内容还包括分布交互仿真数据传输系统中涉及的关键技术及主要算法,如基于发布/订购的域间数据过滤算法、基于 Sender-Group 的域内可靠组播丢包恢复方法、基于趋势分析的拥塞控制方法等,并介绍了平台实现和测试方法。可供读者开展相近的研究或应用时参考。

本书在出版过程中得到了国防工业出版社的鼎力支持,同时虚拟现实技术与系统国家重点实验室的吴威老师、周忠老师、蔡楠、唐昕、刘冬梅、王跃华、于庆等也为本书的撰写和顺利出版付出了心血,在此一并表示感谢!

作 者

2015 年 10 月

目录

第一章 绪论	1
1.1 重要概念	1
1.2 相关领域研究现状	2
1.2.1 IP 组播研究现状	2
1.2.2 应用层组播研究现状	5
1.2.3 可靠组播技术研究现状	7
第二章 基于广域网网关的多对多组播传输模型	12
2.1 引言	12
2.2 相关研究	12
2.2.1 客户—服务器模型	13
2.2.2 广播/组播模型	15
2.2.3 网关模型	17
2.3 基于广域网网关的多对多组播传输模型	21
2.3.1 MDM 模型总体	21
2.3.2 广域网网关节点	24
2.4 基于广域网网关的域间组播通信	26
2.4.1 问题建模	26
2.4.2 基于延迟带宽约束的多对多组播路由算法	27
2.5 基于可靠服务节点的域内可靠组播模型	33
2.5.1 域内可靠组播模型总体结构	34
2.5.2 可靠组播模型软件模块设计	36
2.6 小结	37

第三章 基于发布/订购的域间数据过滤算法	39
3.1 引言	39
3.2 兴趣过滤技术及其发布/订购机制	40
3.2.1 HLA 标准中的 DDM 服务	41
3.2.2 发布/订购区域匹配定义	42
3.2.3 基于发布/订购的数据过滤机制	43
3.3 基于发布/订购的域间数据过滤算法	45
3.3.1 符号定义	45
3.3.2 发布/订购区域的转化	46
3.3.3 网关发布/订购的表示	48
3.3.4 数据过滤关系的建立	51
3.4 过滤机制的代价权衡	53
3.5 小结	55
第四章 基于 Sender-Group 的域内可靠组播丢包恢复算法	56
4.1 引言	56
4.2 相关研究	57
4.2.1 TCP 协议中的丢包恢复方法	57
4.2.2 可靠组播协议中的丢包恢复方法	57
4.2.3 可靠组播协议中的丢包恢复方法总结	61
4.3 基于 Sender-Group 的报文编号及备份方法	62
4.3.1 面向多对多组播的报文编号方法	62
4.3.2 组播报文备份方法	63
4.4 多对多可靠组播丢包恢复算法 BH_RMER	65
4.4.1 面向动态组播组的丢包检测	65
4.4.2 基于报文传输属性的丢包恢复算法	67
4.5 算法分析	70
4.5.1 延迟分析	71
4.5.2 吞吐量分析	72
4.6 实验与分析	73

4.6.1	实验环境	73
4.6.2	丢包恢复延迟	74
4.6.3	平均往返延迟抖动	74
4.6.4	吞吐量对比	75
4.7	小结	77
第五章 基于趋势分析的拥塞控制算法		78
5.1	引言	78
5.2	相关研究	79
5.2.1	拥塞控制算法的分类	79
5.2.2	拥塞控制算法总结	82
5.3	多对多组播拥塞控制的主要问题	83
5.4	基于趋势分析的拥塞控制算法	85
5.4.1	仿真节点发送速率初始化	86
5.4.2	基于丢包率变化趋势信息的速率抑制	87
5.4.3	基于缓冲区占用量变化趋势的速率抑制	89
5.4.4	随机推迟的拥塞反馈抑制	92
5.5	算法分析	93
5.6	实验及数据分析	94
5.6.1	实验环境	94
5.6.2	组内公平性	95
5.6.3	吞吐量波动对比	96
5.6.4	平均往返延迟对比	97
5.7	小结	97
第六章 多对多组播传输系统集成与实验		99
6.1	概述	99
6.2	可靠组播模型的集成与测试	99
6.2.1	可靠组播模型系统 RMSP 的实现	99
6.2.2	实验环境	103
6.2.3	原始丢包率测试	104

»» 分布交互仿真数据传输关键技术

6.2.4	RMSP 吞吐量测试	105
6.2.5	平均往返延迟测试	108
6.3	BH RTI 2.3 集成可靠组播的设计	109
6.3.1	集成对比测试	112
6.3.2	与典型 RTI 对比测试	113
6.4	广域网网关的集成与测试	115
6.4.1	实验环境	117
6.4.2	功能测试	118
6.4.3	性能测试	123
6.5	小结	125
参考文献		126

绪 论

本章阐述了相关重要概念,并对 IP 组播、应用层组播和分布交互仿真中可靠组播等领域的研究现状进行了分析和展望。

1.1 重要概念

分布式虚拟现实(Distributed Virtual Reality, DVR)^[1,7]是在虚拟现实技术和网络技术的基础上发展起来的,又称为分布式虚拟环境、网络化虚拟现实环境、多用户虚拟环境等。它将地理分散的虚拟现实系统通过网络连接起来共享信息,多个用户在一个共享的虚拟环境中进行交互,协作完成任务^[8,9]。分布式虚拟环境的研究是虚拟现实研究中的一个重要方向^[10]。

分布式虚拟环境最早、最广泛的应用领域是分布式的仿真训练,并称为分布交互仿真(Distributed Interactive Simulation, DIS)或分布式仿真(Distributed Simulation)^[10]。分布交互仿真技术由于具有有效性、可重复性、经济性和安全性等特点,在国民经济和国防建设领域有着广泛的应用,尤其是在军事领域,包括我国在内的很多国家已经建设了大量的分布交互仿真系统,其中美国推出的相关标准和建立的系统得到了广泛关注和参考^[11]。

组播技术是分布交互仿真中的重要技术,美国所举行的历次大规模军事演习都是以组播技术为基础的。现在的 Internet 是建立在 TCP/IP 协议基础上,其通信以点对点单播通信为主,而对组播通信的

支持很少。随着社会和技术的发展,出现一些具有一对多或者多对多通信特点的应用,一对多和多对多的通信模式对 Internet 提出了更高的要求。组通信应用对 Internet 数据传输的要求可能包括带宽需求、延迟需求、可靠性需求、多发送方、扩展性、动态加入/退出等要素。在视频点播等应用领域,已经有以组播支持为重点的核心路由器推出,在一些企业网或小区网起着重要的作用。

单播通信时,发送者需要向各个接收者逐一发送数据报文,而组播通信中,发送者只需向指定的组播地址发送一个数据报文,该报文的副本则由具有组播功能的路由器或交换机完成,并最终发送给各个接收者。因此,组播通信可以节省大量的网络通信资源,提高通信效率。但 IP 组播存在如大规模组管理、服务质量、安全性等问题,现在尚未在 Internet 上广泛部署,限制了其应用范围。尽管 IPv6 中将组播作为一项重要改进内容,在技术规范、安全性等方面有所完善,但大规模的组播应用仍将是极其困难的,需要寻求新的技术方案以满足特定领域的组通信需求。

组播应用的领域很多,不同的应用对组播组规模、组播数据发送方数量、组成员动态性、带宽延迟要求等方面都存在很大的差异,所以组播应用在可预见的时期内实现像 TCP/IP 一样的通用协议是不太可能的。分布交互仿真对组播通信各项指标的要求较为严格,尤其是苛刻的实时性和多点对多点要求,组播对广域网的分布式交互仿真来说更是一个很大的挑战。

1.2 相关领域研究现状

本节介绍目前 IP 组播研究现状,应用层组播研究现状,总结分布交互仿真中组播及可靠组播问题的相关研究。

1.2.1 IP 组播研究现状

组播(Multicast)是相对于单播(Unicast)和广播(Broadcast)而言的,在发送者和每一接收者之间实现点对多点网络连接。如果一台发送者同时给多个接收者传输相同的数据,只需复制一份相同数据包。

它提高了数据传送效率,减少了骨干网络的数据交换量。这三种数据传输方式如图 1-1 所示。

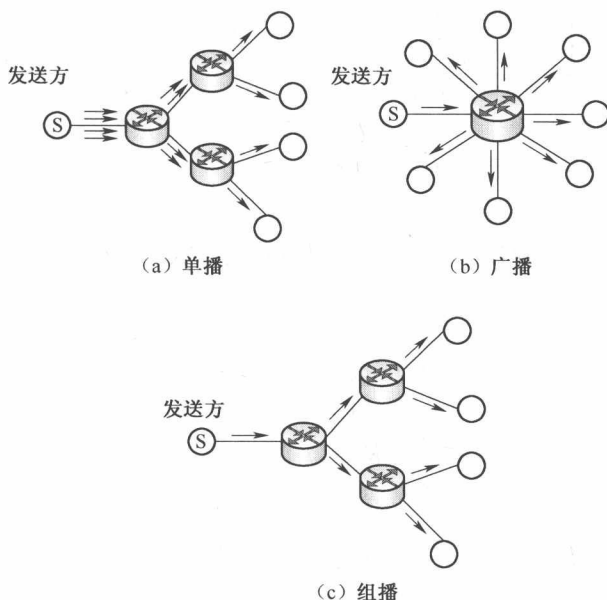


图 1-1 单播、广播、组播传输示意图

在图 1-1 中,单播传输的方法,需要发送方重复发送报文,占用了大量的带宽。广播传输的方法,则向局域网内的每一台主机发送消息,主机会接收到很多不需要的消息,造成了主机和网络资源的浪费。组播传输的方法由路由器根据组播组与主机的关系,判断链路上哪些主机对组播报文感兴趣,并完成报文复制和转发。

IP 组播的概念是在 20 世纪 80 年代中期由斯坦福大学的博士生 S. E. Deering 提出的。1988 年,D. Waltzman,C. Portridge,S. E. Deering 发表了介绍距离向量组播路由协议的文章。这篇文章提出了一个开放、动态无连接的组播服务模型,该模型使用户可以通过 D 类组播地址,在任何时间创建、加入和离开组播组,并通过组播组发送和接收数据。Internet 工程任务组 IETF 为组播研究成立了众多工作组,分别研究组播路由、组播安全、可靠组播等 IP 组播技术。随后构建了 IP

组播实验床 Mbone^[22]。1992 年的 IETF 大会上,首次通过 Mbone 实现了广域网上的 IP 组播音频传输^[23]。IP 组播的研究至今已经开展了 20 年,形成了大量的研究成果及协议。在 IP 组播路由协议方面,主流的组播路由协议有以下几种:距离矢量组播路由协议 DVMRP^[24](Distance Vector Multicast Routing Protocol)、组播扩展 MOSPF^[25](Multicast OSPF)、密集模式独立组播协议 PIM-DM^[26](Protocol Independent Multicast-Dense Mode)、CBT^[27](Core Based Trees)、稀疏模式独立组播协议 PIM-SM^[28](Protocol Independent Multicast-Sparse Mode)等。

PIM-DM 主要被设计用于组播局域网应用程序,PIM-DM 使用了和 DVMRP 及其他密集模式一样的溢出和修剪机制。DVMRP 和 PIM-DM 之间的主要不同在于 PIM-DM 主要引入协议独立的观念。PIM-DM 可以使用由任意底层单播路由协议产生的路由表执行反向路径转发(RPF)检查。MOSPF 是通过在 OSPF 链接状态通告中包含组播信息而工作的。每个 MOSPF 路由器都可以了解到哪个局域网(LAN)上的哪个组播组在活动。MOSPF 为每对源/组建立一个分配树并且为发送到组的活动源确定一个树。树的状态被缓存,并且当链接状态发生变化或高速缓存器超时的时候,必须重新确定树。PIM-SM 是一种能有效地路由到跨越大范围网络组播组的协议。PIM-SM 协议不依赖于任何特定的单播路由协议,主要被设计来支持稀疏组。它使用了基于接收初始化成员关系的 IP 组播模型,支持共享和最短路径树,此外它还使用了软状态机制,以适应不断变化的网络环境。它可以使用由任意路由协议输入到组播路由信息库(RIB)中的路由信息,这些路由协议包括单播协议如路由信息协议(RIP)和开放最短路径优先(OSPF),还包括能产生路由表的组播协议,如距离矢量组播路由协议(DVMRP)。

除了以上的组播路由协议之外,还有一些应用与域间路由或地址管理等方面的协议和方案也得到了广泛的研究,其中包括:组播协议边界网关协议 MBGP^[30](Multiprotocol Border Gateway Protocol)、组播源发现协议 MSDP^[31](Multicast Source Discovery Protocol)、组播地址集请求协议/边界网关组播传输协议 MASC/BGMP^[32]。

IP 组播的研究虽然已经取得了很多成果,但一直未能在 Internet 上全面部署并发挥作用,它本身也存在很多问题^[33]。IP 组播的扩展

性受到一定限制,路由器需要保存每个组播组以及每个组播节点的状态,随着组播规模的扩大,组播节点与组播组数量的增加,路由器的负载迅速增加,成为系统的瓶颈,扩展性受到限制;IP 组播采用尽力投递方式,不保证传输的可靠性。路由器或主机缓存溢出等问题可能引发网络或主机拥塞,造成数据包的丢失,因此不能保证所有的接收方都收到报文。传输的顺序性是另一个问题,由于传输路径的不同,及在路由器处排队队列的不同,造成了数据包被接收方收到的顺序与它们的发送顺序不一致;同时,目前现有的域间组播协议存在扩展性差、实现复杂度高的问题,阻碍了 IP 组播在广域网的部署。另外,IP 组播需要支持组播的路由器,而 Internet 现有的路由器只有少部分支持组播路由,这需要各路由器厂商对相关标准取得共识和支持,同时还需要确立组播地址分配中心的权威性和职责,才有可能全网部署。

因此,在现有的基础设施条件下,需要研究其替代方案,如在现有的 Internet 路由器基础上通过上层应用实现组播路由。因此,近年来应用层组播的研究逐渐成为组播领域新的研究热点。

1.2.2 应用层组播研究现状

IP 组播因为技术和市场等方面的问题难以在 Internet 广泛地部署,因此,结合 Internet 的性质和应用的特点,在 IP 组播模型、覆盖网(Overlay Network)和 Peer-to-Peer 等技术日新月异的基础上,发展出了应用层组播技术。

应用层组播将依赖于路由器支持的组播功能转移到终端系统上,通过终端进行数据的复制与缓存,而终端之间仍运用现有的单播传输技术,使得在现有的网络基础设施上可以支持组播,而不需要路由器进行数据的复制和组播组管理。图 1-2 为应用层覆盖网组播示意图。

应用层组播中,组播组状态在终端系统中维护,不需要路由器保持组播状态,改变了 IP 组播在广域网中对路由器的依赖,解决了组播扩展性的问题。应用层组播可以根据所运行的应用需求构建不同的覆盖网络,不需要网络设备升级和功能支持。建立在网络连接之上的应用层组播节点之间一般通过 TCP 或可靠 UDP 进行消息转发,可利用 TCP 的可靠性和拥塞控制实现简单的组播可靠性保证和拥塞控制。

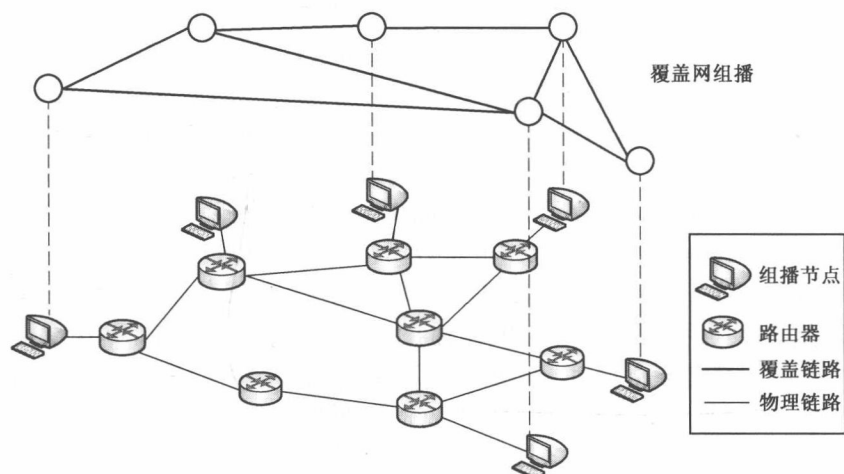


图 1-2 应用层覆盖网组播示意图

根据节点拓扑的不同,应用层组播的体系结构可分为对等型、代理型、服务器型三类。对等型体系结构中,每个组播节点都是对等的,节点之间通过一定的算法和协议形成自组织网络,组播的复制转发等功能均通过组播节点来完成,每个节点的地位和任务都是平等的。代理型体系结构中,通过部署一定的代理节点负责组播消息,组播节点通过接入距离最近的代理节点加入组播组,而这些代理节点的部署和数据传输路径均由服务提供商负责预先确定服务器型是介于对等型与代理型之间的一种应用层组播体系结构。组播数据的转发由具备较大负载的服务器负责,这些服务器可能是性能较高的普通组播节点。对于这些节点构建的组播转发树而言,结构和性能均比较稳定,通常可支持大规模的应用层组播服务。

针对不同的体系结构,出现了许多应用层组播协议。这些协议通常可分为集中式和分布式两类。集中式协议通过一个服务器负责构建所有组播节点的拓扑,由服务器进行延迟、带宽等指标的计算构造转发树,然后将这些关系发给相应的组播节点。这种方法的优点是实现简单,负载小。但其带来的问题是扩展性受到了单个服务器的限制。集中式应用层组播协议有 ALMI^[130], HBM^[131] 等。分布式应用

层组播协议中,每个组播节点分布的构建组播拓扑,数据的转发树也由组播节点动态生成。分布式协议中控制拓扑(mesh)维护了节点之间的多条路径,转发树(tree)则为组播数据转发的依据。在 mesh 基础上构建转发树的协议称为 mesh 优先,代表性的协议有 NICE^[132], Overcast^[133], Yoid, HMTTP, HostCast^[134]等;在 tree 的基础上引入额外的边从而构成 mesh 的协议称为 tree 优先,代表性的有 Chord^[135], Scribe^[136], Narada^[137]等。

但是,应用层组播主要是以单播为基础的,在稳定性和传输效率等方面也存在一些问题。通过应用层方式实现组播,增加了网络资源的需求,控制信息增加了数据传输的开销,引入了构造和维护组播路由的空间和时间代价。覆盖网的构建,多条逻辑链路的生成可能占用同一条物理链路,或者一条逻辑链路重复利用多条物理链路,影响了数据传输的性能。相对通过路由器实现的 IP 组播,应用层组播系统主机性能受到影响,稳定性较差,牺牲了 IP 组播服务的高效性。

图 1-3 为 IP 组播和应用层覆盖网组播的对比示意图。IP 组播和应用层组播在特性和局限性上具有互补性,IP 组播传输实时性好,延迟小,但是需要路由器维护组播组的状态信息,在 Internet 的部署上存在困难。应用层组播在性能上低于 IP 组播,但不需要路由器的支持,便于部署。如何将两者有效地结合起来以满足应用的需求是我们拟解决的关键问题。

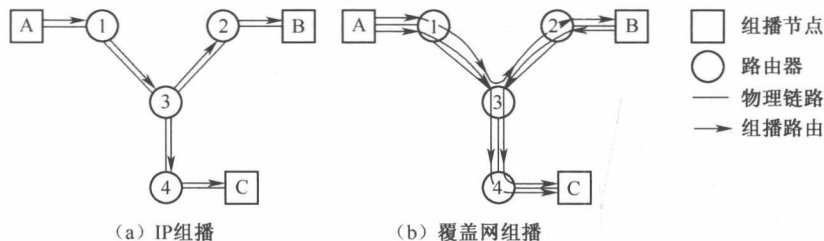


图 1-3 IP 组播和覆盖网组播对比图

1.2.3 可靠组播技术研究现状

由于不同的组播应用对可靠性的要求差异很大,对可靠组播的研

究要比 TCP 可靠传输机制困难得多,不存在一个能够满足所有可靠组播要求的传输模型。因此在可靠组播的研究中出现了大量的可靠组播协议,每一种协议往往都是针对一种类型的组播应用而设计的。大规模分布交互仿真应用需要使用大量的组播组通信,是一种典型的多对多组播通信应用,其中大量的动态组播组并发通信,且组播报文对传输的延迟非常敏感。现有的可靠组播研究主要还是面向基于端到端的一对多可靠组播,在多对多的可靠组播研究方面目前国际上主要的做法有基于 TCP 的可靠组通信和基于 IP 组播的可靠传输。

1. 基于 TCP 的可靠组通信 (TCP Exploder)

应用层的组播通信 TCP Exploder^[34,35]是多数分布交互仿真系统中实现可靠组播的一种方式,将组播数据复制多份,通过 TCP 依次发送给不同的接收方。美国军方早期使用的 RTI-s,在设计中其可靠消息服务就基于 TCP 实现了应用层组播,其体系结构如图 1-4 所示。

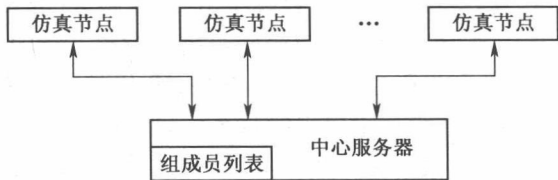


图 1-4 TCP Exploder 体系结构

从图 1-4 可以看出,TCP Exploder 技术采用客户端/服务器的模式,客户端仅需要维护一个中心服务器的连接,而服务器负责与所有客户端进行交互。中心服务器记录组成员列表,发送方将数据以 TCP 方式发送给中心服务器,由中心服务器遍历组成员列表,将该数据依次通过 TCP 发送至各组成员。在 20 世纪 90 年代,受美国国防部资助,MIT 林肯实验室基于 TCP Exploder 在应用层模拟实现了 STOW-RTIs^[31]的可靠组播传输,但在实际应用中发现,STOW 级别的战术演习中即使是可靠传输的数据都有可能引起拥塞而造成网络灾难性崩溃,所以美军进行军事演练时使用的 RTI-s 一直没有使用可靠传输方式。

由 Georgia 大学 Richard Fujimoto 教授主持开发的 FDK (Federated