

分布式虚拟环境 AIMNET 的关键技术概述^{*}

卢 威 陈继明 徐晓阳 潘金贵

(南京大学计算机软件新技术国家重点实验室, 南京大学计算机科学与技术系 南京 210093)

摘要 本文概要介绍南京大学自主研究开发的分布式虚拟环境 AIMNET 的关键技术。AIMNET 借鉴了兴趣管理的思想, 提出了双向-共享组播树的通讯结构, 从而大大降低了网络通讯量, 并保证了 AIMNET 系统良好的扩展性; AIMNET 中静态数据采用分布式存储结构, 同时为了实现场景的远程渲染而采用了负载均衡、数据压缩、累进式渲染、航位预测算法以及数据预取等技术。

关键词 分布式虚拟环境, 兴趣管理, 主动路由, 数据分布, 远程实时渲染

AIMNET: The Key Techniques in Distributed Virtual Environment

LU Wei CHEN Ji-Ming XU Xiao-Yang PAN Jin-Gui

(State Key Lab for Novel Software Technology, Department of Computer Science and Technology, Nanjing University, Nanjing 210093)

Abstract This paper introduces the key techniques in distributed virtual environment AIMNET developed by Nanjing University. Using the methods of Interest Management by reference, AIMNET advances the communication topology structure of bi-directional shared multicast tree, in order to reduce the number of messages and emphasize the scalability of the system. At the same time, the distribution of static data is presented in this paper, in company with the remote real-time rendering technology, which includes load balancing, data-compression, progress rendering, dead reckoning and data pre-fetching.

Keywords Distributed virtual environment, Interest management, Active routing, Data distribution, Remote real-time rendering

目前 DVE 系统需要关注的重要问题就是如何解决系统的扩展性问题, 以及如何减小大量用户之间的通讯量。我们自行设计开发的分布式虚拟环境系统 AIMNET (Active Interest Management NET) 在这方面做了有益的尝试。目前 AIMNET 的研究重点主要集中在两个方面: 首先是建立 AIMNET 中动态数据的通讯结构, 在此我们主要借鉴了兴趣区域管理的方法, 采用了树状拓扑结构, 大大地降低了用户之间的通讯量, 比较好地解决了系统的扩展性问题; 其次是静态数据的分布及其远程渲染, 我们主要采用分布式的服务器/客户端(C/S)模式, 采用这种模式的优点是易于维护数据的一致性, 但是对客户端的远程渲染提出了较高的要求。为了解决这一问题, 我们采用了累进式渲染 (Progressive Rendering) 的概念。

本文重点介绍动态数据通讯的拓扑结构和静态数据分布及其远程渲染。文章的最后给出了实验的结果以及对进一步工作的展望。

1 AIMNET 体系结构

1.1 相关概念

通讯扩展性问题是分布式虚拟环境需要解决的核心问题之一。DVE 的重要特点或约束条件是有大量对象参与并相互通讯, 且这些对象的状态和位置在持续进行动态变化, 这就给 DVE 的通讯扩展性提出了难题。解决这个问题的基本思

路是兴趣管理^[2], 该方法的基本思想是: 对象只与自己相关的对象进行通讯, 而其他不相关的对象和数据则被过滤, 因而降低整个虚拟环境的整体通讯量。但是兴趣管理方法发展至今仍然没有圆满地解决这个问题。

传统的体系结构有两种: Peer-to-Peer 结构和 C/S 结构。Peer-to-Peer 即对等模型, 这种结构的主要缺点就是通讯量大, 一致性难以维护, 系统难以扩展。而 C/S 模型的主要缺点就是服务器容易产生瓶颈, 也导致系统可扩展性差。这两种结构都不适合进行兴趣管理。

目前流行的是层次结构。层次结构将虚拟环境的场景数据分布在不同的服务器上, 服务器组织成层次树。主机从一个场景进入另一个场景, 对应从一个服务器过渡到另一个服务器。层次结构存在堆积问题 (clumping problem), 而且数据分布及服务器结构与具体应用有关。

在介绍 AIMNET 的拓扑结构之前, 先简单介绍几个概念:

(1) 兴趣区域管理 (Area of Interest Management, AOIM)

是指对象只与自己相关的对象进行通讯, 而其他不相关的对象和数据则被过滤, 其实质是通过信息过滤为对象定义兴趣范围。AOIM 的构成分为三个要素: 表示模型、过滤规则和体系结构, 主要应用在三个方面: 动态数据 (状态信息等) 过滤、静态数据 (几何数据等) 过滤和本地渲染数据过滤。

(2) 发布-订购模式^[7] 所有主机首先向系统发出订购消

^{*} 本课题受分布式虚拟环境中可扩展主动兴趣管理研究 (60473571) 和远程沉浸式虚拟奥运博物馆关键技术 (60533080) 国家自然科学基金基金资助。卢 威 硕士生, 研究方向为计算机图形学、分布式虚拟环境; 陈继明 博士生, 研究方向为分布式虚拟环境、agent 技术; 徐晓阳 硕士, 研究方向为分布式虚拟环境、实时渲染; 潘金贵 教授, 博导, 研究方向为多媒体信息处理技术、计算机图形学、远程网络教育。

息,表明自身感兴趣的事件和消息;兴趣管理系统维护这些兴趣信息。之后,当某个主机简单地向系统发送(发布)消息时,系统根据已有的发布区域与订购区域的匹配关系将消息转发给合适的接收者。

(3)表示模型^[3] 兴趣区域内待检测的其他对象、对象状态的表示方式、对象表达自身兴趣的方式(通常是条件集合)组成了表示模型。其中对象表达自身兴趣的方式本文称为“兴趣表达式”,也就是订购区域和发布区域的表示方法。

(4)过滤规则 对感兴趣的信息进行过滤的方式。可以是简单的值比较,也可以是复杂的逻辑规则。

(5)体系结构 描述了实现兴趣管理的通讯结构和软件配置。

(6)匹配 根据对象的表示方式以及兴趣表示方式来判断哪些对象通过检测的过程。匹配过程的核心是过滤规则的定义。

1.2 通讯结构

为了能够很好地解决 DVE 系统可扩展性的问题,我们摒弃了传统的通讯体系结构,采用了双向-共享组播树(bi-directional shared multicast tree)的拓扑结构,如图 1 所示。图中的路由器称为主动路由器(Active Router)。所谓主动路由器,简单地说就是负责转发数据包的路由器。在此,它们组织成一棵组播树^[8]进行兴趣管理。其中标记为 Wuhan 的路由器作为组播树的根结点,称为核心路由器(Core Router)。主机或场景服务器都连接在主动路由器上,构造一棵双向共享组播树,并在每个路由器上存储下游节点的订购区域兴趣表达式。主机发送的数据包到达核心路由器或者中途的主动路由器时,这些路由器将会根据发布和订购表达式进行匹配,决定是否向其他下游节点进行转发。

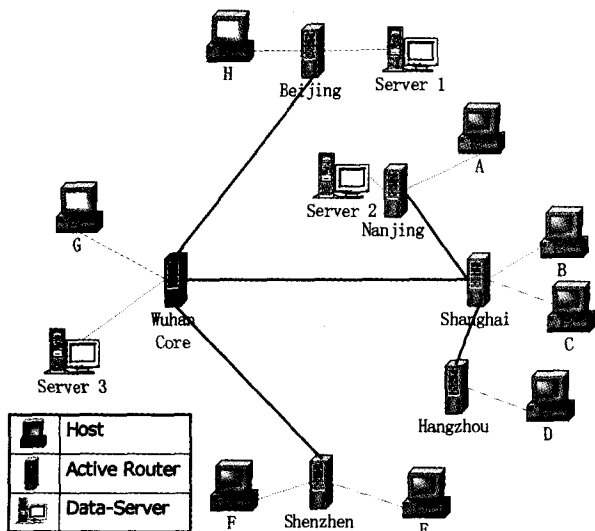


图 1 双向共享组播树结构

我们的主要工作如下:

(1)首次采用双向共享组播树作为兴趣管理的通讯拓扑结构,结合运用主动路由^[6]思想和发布-订购模式,根据数据包的兴趣区域进行主动过滤和转发。这种方法同时满足了减少存储量和通讯量的要求,并能快速动态地改变通讯关系。

(2)采取属性关系表达式作为兴趣管理的表示模型,提出了兴趣表达式的聚集方法,从而进一步降低了系统的存储量。特别借鉴了 IP 协议中超网/子网的构建方式,提出了属性值的聚集方法。这个方法不仅实现了属性值的聚集,还实现了

属性值的超集/子集(集合)关系。

(3)首次提出利用兴趣表达式进行流量控制的方法。这些方法不止是兴趣管理传统研究中的高级过滤规则,它们更类似于网络协议中的流量控制。这一策略从通讯底层保证了分布式虚拟环境的通讯量不超过给定阈值,从而在很大程度上提高了系统的可扩展性。

1.3 协议结构和配置

为了构造一个完整的系统,需要以下四个协议:

(1)订购路由协议(Subscription Routing Protocol, SRP)

用于建立组播树、传播、分布和维护兴趣表达式。

(2)实时数据包转发协议(Real-time datagram Delivery Protocol, RDP) 用于对订购表达式和发布表达式进行匹配并进行转发。

(3)流传输协议(Stream Transport Multicast Protocol, STMP) 提高数据传输的可靠性。

(4)几何传输协议(Geometry Transport Protocol, GTP) 用于传输场景几何数据。

其中 RDP 和 STMP 协议是对等协议,GTP 协议采取 Client/Server 方式。这些协议与 TCP 和 UDP 协议的关系如图 2。

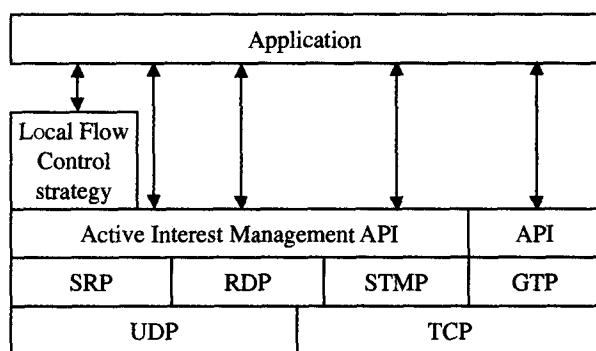


图 2 协议结构

主动路由器作为进行兴趣管理的路由器,与主机配置的协议并不相同。图 3 列出了上述协议在不同节点上的配置。其中路由器只需配置 SRP/RDP/STMP 协议,这些协议用于维护兴趣信息(也就是路由信息)和转发数据包。主机除了配置这三个协议外,还需要配置 GTP 的客户端,而场景几何数据服务器则需要配置 GTP 协议的服务器端。

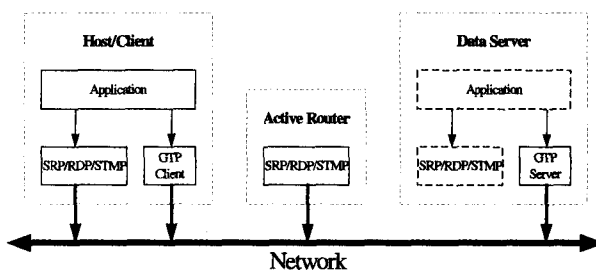


图 3 协议配置

1.4 主机端软件体系结构

分布式虚拟环境应用层配置与底层通讯结构是密不可分的,图 4 很好地反映了应用层的体系结构以及应用层与协议层之间的交互。

应用层由三个模块组成:仿真引擎(simulation engine)、渲染引擎(rendering engine)和输入处理(input manipulator)。

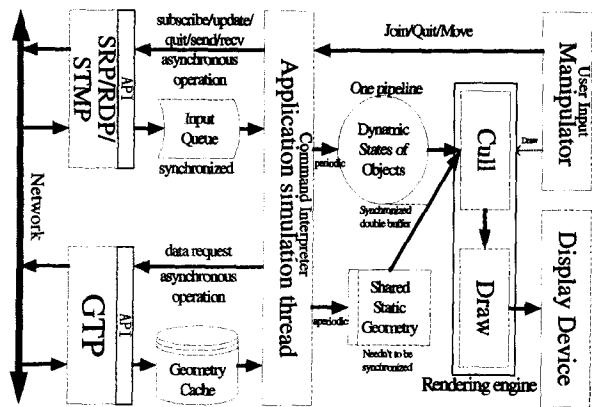


图4 应用层与协议层之间的数据流和控制流

输入处理模块负责将用户的所有键盘和鼠标操作转换为仿真引擎能够处理的命令,并将这些命令发送给仿真引擎,也就相当于“命令处理器”设计模式中的命令控制器。这些命令相对而言比较简单,只包括:加入虚拟环境、退出虚拟环境、移动化身这三个命令。

渲染引擎根据收到的几何数据和动态状态数据进行渲染,首先通过裁减(Cull)过程从这些数据集合中选择需要渲染的数据,然后由渲染(Draw)过程负责绘制。

仿真引擎是连接渲染引擎及输入处理模块与协议层之间的桥梁。对上游接口,它负责处理输入模块传来的命令并更新本地动态对象的状态,再把本地或者其他对象的这种状态更新信息通知给渲染引擎,进行即时绘制。对下游接口,它负责把本地用户的对象信息交给 SRP/RDP/STMP 协议来实现与其他用户的交互,同时通过 GTP 协议从远程数据服务器上下载几何数据并且通知渲染引擎进行渲染。

2 数据分布结构和实时渲染技术

在 AIMNET 中,我们改进了传统的客户端/服务器(Client/Server, C/S)模式,采用分布式服务器集群模式来存储视景数据。与客户端完全复制视景数据相比,这种分布式 C/S 结构的优点是显而易见的:结构灵活,易于维护数据的一致性。但是这种结构要解决的一个关键问题是:由于网络带宽不足而影响渲染速度。为了达到实时渲染,我们借鉴了各种技术,如负载均衡、数据压缩、累进式渲染、航位预测算法以及数据预取等。

2.1 几何数据的分布式组织

分布式虚拟环境中的静态数据分布在多台数据服务器上,我们采用了一种简单的虚拟空间的划分方法^[4]把三维空间分成多个长方体。具体实现如下:我们对划分做了一个简化,只对 X、Y 这两维进行划分,将整个虚拟空间划分为多个贯穿 Z 轴的子空间,每个数据服务器维护一个兴趣区域表达式表示的空间区域内的数据。如图 5 所示,将整个空间分为五部分,对应的静态数据分别存储在五台数据服务器上,数据服务器 1、3、4、5 维护的区域都可以用简单兴趣区域表达式来表示,而服务器 2 对应的区域则需要用复合兴趣区域表达式来表示。

2.2 实时渲染

采用分布式 C/S 模式需要面临的另一个重要挑战就是由于网络传输速度不足而影响实时渲染的效果。

我们的系统为了解决这个问题而采取的主要策略如下:

(1)均衡负载。随着用户数量的增加,场景数据库服务器的负担就会相应增加。当超过一定负荷的时候,系统的性能和资源的效率就会急速下降,这是传统的 C/S 模型经常会遇到的瓶颈问题。我们的解决方法就是采用分布式的服务器集群策略,底层的网络结构使得我们能够方便地实现多服务器架构:只要把视景数据根据空间大小分块存储在不同的数据服务器上,然后数据服务器向定位服务器注册空间信息,由定位服务器来组织管理数据服务器,这样就保证了客户端和数据服务器之间位置的透明性,也很好平衡了负载。

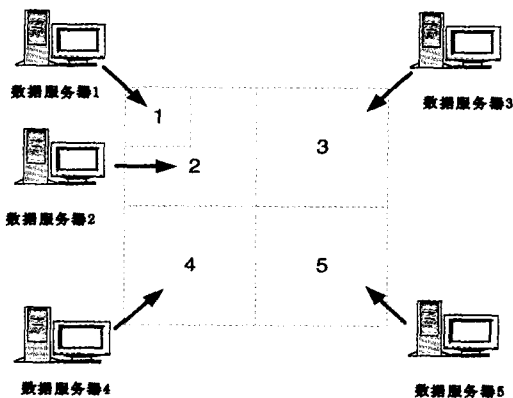


图5 空间划分在二维平面上的投影表示

(2)数据压缩。采用 C/S 模型就意味着视景数据需要通过网络进行传输,所以要减小网络传输量的一个主要方法就是采用压缩算法来压缩视景数据,再传送到客户端进行解压还原,这样可以极大地减小传输量。但是由此带来了一个缺点:增加了客户端的负担,不过目前客户端的计算能力是可以信赖的。

(3)累进式渲染。用树状结构组织场景数据,在远程实时渲染的时候采取由粗到精,逐渐细化的方式。当数据下载到本地之后,为了提高渲染速度,在客户端还采用了 LOD 算法和视锥裁减算法。

(4)航位预测算法和数据预取。航位预测算法(Dead Reckoning, DR)是根据用户的当前状态来判断用户下一步可能出现的状态的一种算法。根据 DR 算法,事先下载用户可能会到达的场景数据放入缓冲区,这种数据预取策略大大加快了渲染的速度,提高了连续性。

3 实验与评估

AIMNET 系统所涉及的这些协议都在网络协议层的应用层实现。主动路由器和主机端的协议及软件均建立在 Windows NT/2K 平台上。我们将主动兴趣管理系统选择在校园网内进行实验,用于实验的网络拓扑如图 6。图中用粗线条连接的是设置为主动路由器的计算机。核心路由器(Core Active Router)左边的计算机(Active Router 1)作为下游的一个主动路由器,该路由器与所在的局域网中的 7 台主机通过 IP 组播方式进行通讯。该局域网中的第 8 台主机直接与核心路由器通过 UDP 单播方式进行通讯。核心路由器跨越两个局域网。校园网内的另一个主机也通过单播方式连接到核心路由器上。在这种配置情况下,由于内部网络都是总线式以太网,单播的数据包也是以广播方式实现的,因此即使内部网络中的主机不接收数据包,主动路由器 1 向核心路由器发送的数据包也会被广播到内部局域网中。同样,核心

路由器发送给该主动路由器的数据也会被广播给内部网络中的所有主机,再由主动路由器转发给这些主机。由于实验条件的限制,本文采取这种方案,一方面可以尽量利用实验室的局域网,另一方面这种方案只增加了局域网的负担,并不对模拟实验的结果产生本质的影响。

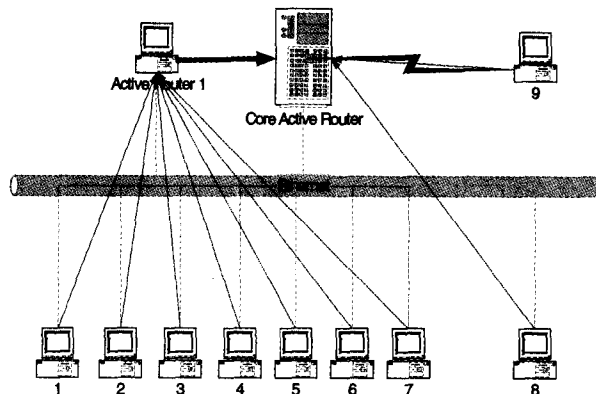


图6 实验网络结构

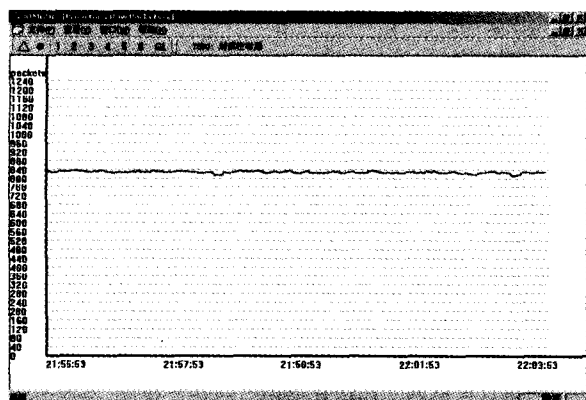


图7 核心路由器上的数据包接收率

在实验中,除了主动路由器之外,每个主机都对64个活动对象的行为进行了仿真(真正用于实际应用时,每个主机实际上只需要维护一个活动对象)。用于实验的虚拟环境空间大小固定,为 $360\text{m} \times 360\text{m} \times 200\text{m}$ 的空间范围,主机维护的活动对象在虚拟环境中呈均匀分布,并在各自均衡位置附近的垂直方向上做简谐振动。这些对象之间的间隔是10m,不同主机维护的对象边界之间的距离是20m。对象的最大订购区域是以各自位置为中心,尺度为 $110 \times 110 \times 100$ 的范围内,其中方框内表示的是5号主机所有活动对象订购区域的总和。每个对象每隔50ms更新一次状态,每次每个主机在一个数据包中发送两个对象的状态。为了避免占用局域网大量带宽,限制主机发送数据包的最大速率为300packets/s,其中每个数据包包括1个发布表达式和两个活动对象的状态信息,有200字节左右。为使主机发送数据包的速率保持大致稳定,没有采用航位预测算法(Dead Reckoning)来减少发送量。

核心路由器接收数据包的速率见图7。这是在没有应用流控制策略时,9个主机全部加入后核心路由器接收的数据包的速率(这是接收速率稳定后采样得到的数据)。核心路由器接收三个计算机发送的数据包,包括两个主机和一个主动路由器。这个数据包接收率相当于两到三个主机发送数据包时接收到的速率。如果没有裁剪区域,那么主动路由器1将

把所有下游主机发送的数据包转发给上游路由器,那时核心路由器的预计接收速率应该是现在的三倍,也就是2700packets/s左右(也就是说9个主机发送的数据包都被核心路由器接收并转发)。而采用了兴趣管理的区域裁减方法之后,通讯量明显下降,大约是840packets/s(如图7所示),因此裁剪区域可以有效地减少通讯量。

目前AIMNET中构建的静态场景比较简单,实现了对真实的地形、树木等的渲染,而对复杂的场景的渲染还在进一步扩充中。现阶段场景渲染效果如图8所示。



图8 AIMNET的场景渲染效果

结论和进一步的工作 AIMNET系统主要用VC++ 6.0和OpenGL编程实现,实验平台采用Windows NT/2K,在南京大学的校园网内取得了很好的试验效果。

DVE牵涉到了计算机技术的各个方面。现阶段AIMNET侧重对DVE的系统体系结构、场景数据分布和远程渲染等进行探讨,进一步的工作还在继续,比如把AIMNET系统改造成一个通用的基于内容的发布-订购服务,通过改造路由算法,为不同的应用构造不同的路由空间,以支持各种类型的应用;针对不同的应用采用不同的组播树,采取不同的核心路由器;拓扑结构不再静态配置,而是在初始状态下构成网格(Mesh),采用自动构造算法在网格上为不同应用构造不同的共享组播树,从而有效提高系统的健壮性和转发效率。又比如:把AIMNET系统建设成互连网上能够支持1000个以上的用户的网上虚拟社区等等。

最后,在此向林诚凯、李惠、孙元浩、龚震宇、崔业怡等表示感谢,谢谢他们对本文及在试验中作出的贡献。

参考文献

- 1 潘志庚,姜晓红,张明敏,等. 分布式虚拟环境综述[J]. 软件学报,2000,11(4):461~467
- 2 Macedonia M, Pratt D, Zyda M. NPSNET: A network software architecture for large scale virtual environments[D]:[PhD Thesis]. Naval Postgraduate School,1995
- 3 龚震宇. 分布式虚拟环境中的事件模型[D]:[硕士论文]. 南京大学,2003
- 4 李惠,翟磊,林诚凯,等. ROAM算法在超大规模地形渲染中的应用[J]. 中国图像图形学报,2003(8):526~529
- 5 孙元浩,龚震宇,李惠,等. 可扩展主动兴趣管理技术研究[J]. 中国图像图形学报,2003(8):771~775
- 6 Tennenhouse D L, Smith J M, Sincoskie W D, et al. A Survey of Active Network Research[J]. IEEE Communications Magazine, 1997, 35(1): 80~86
- 7 Aguilera M, Strom R, Sturman D, et al. Matching events in a content-based subscription system[A]. In: Eighteen ACM Symposium on Principles of Distributed Computing (PODC'99)[C], 1999. 53~61