

对等网络自组织中间件模型的研究与设计^①

Design of Self-Organization P2P Network Middleware

汪婷婷 (安阳工学院 计算机科学与信息工程系 河南 安阳 455000)

周佳骏 (安庆师范学院图书馆 安徽 安庆 246003)

摘要: 提出一种自组织式分层对等网络中间件模型, 以实现应用进程间自适应通信。模型基于遗传算法封装结构化对等网络基本功能, 提供一套对外服务原语, 支持自组织结构化对等协议仿真。在 DVE-IW 平台上的实验表明模型能够按照小世界现象自组织对等进程并实现进程间可靠通信。

关键词: 对等网络 中间件 自组织 遗传算法

随着对等网络的飞速发展, 广泛应用于组播, 文件存储和资源定位等领域的对等网络中间件技术因其对上层应用屏蔽底层物理网络的复杂性和异构性, 进而大大提高系统的稳定性和健壮性等优点而成为对等网络研究的核心理论之一。当前典型的对等网络中间件研究包括 Gong^[1]提出 JATA 中间件, 定义了三层模型(核心层, 服务层和应用层), 基于 XML 的通信协议, 对上层应用提供了统一的 API 接口; Dabek et al^[2]提出一种封装 CAN, Chord, Pastry 和 Tapestry 等不同结构化对等网络的中间件模型, 提供对外统一接口; P2PSIM^[3]模拟了多种协议, 设计良好, 易于扩展; PLP2P^[4]是一个包括 PeerAPP 层, PeerAgent 层和 SocketAPP 层的仿真中间件模型; Karl Aberer^[5]提出了一种适合结构化和非结构化对等网络的中间件模型, 该模型分三层: 应用层, p2p 存储层和 p2p 基础功能层。但这些研究都因其领域局限性, 未考虑消息转发模块自适应网络动荡特征。

基于上述分析, 引入一种自组织对等网络中间件, 封装对等网络包括自组织寻径在内的所有模块, 提供对外统一服务原语, 因而支持自组织对等网络协议的仿真。本文首先定义、分析对等网络及其基本模块, 并在此基础上, 将其封装为三层中间件模型, 然后设计该中间件模型, 最后, DVE-IW 平台上实验数据证明模型是正确和有效的。

1 对等网络功能描述

对等网络是为实现通信目的使用相同的协议互连的一组对等体的集合^[6,7]。形式可描述为:

其中, **peer** 表示对等体, 指具有特定标识, 提供同等服务的进程。**protocol** 表示对等协议, 是一套用于对等体间信息交换的语用, 语法和语义规则集合。

下面基于上述定义, 分析对等网络的基本模块及其关系。

1.1 对等网络的基本功能模块

在文献[5]归纳的对等网络重要功能模块的基础上, 抽取对等网络重要功能模块包括以下几个部分:

(1) 标识符空间的设计及其组织(对等体)

主要包括两个参数: 标识符自身组成和对等体间距离函数 **d**, 具体实例如:

CAN 使用一种虚拟标识符的 Euclidean 空间, 该虚拟标识符配到该空间中。距离函数 **d** 是 Euclidean 距离。

P-Gird 的标识符是位串, 距离 **d** 可定义为

$$d(a,b)=\min(|\sum_{i=1}^k a_i 2^{-i} - \sum_{i=1}^l b_i 2^{-i}|, |\sum_{i=1}^k a_i 2^{-i} - \sum_{i=1}^l b_i 2^{-i}|)$$

Chord 和 DKS 的标识符空间是 **N** 个节点的集合, 且 $d(x,y)=(y-x)\bmod N$

(2) 标识符映射算法(对等体), 该映射函数 $F_p: P \rightarrow I$, 将对等体和 **I** 中唯一虚拟标识联系起来,

^① 收稿时间:2008-08-27

例结构化对等网络可采用哈希函数实现这种一一映射。

(3)对等体的结构及其管理策略(拓扑),又可理解为寻径表的结构及其维护,拓扑的结构包括环,树,超立方体等结构,其参数为邻居节点表,邻居关系 $N:P \rightarrow 2^P$,使得对每个给定的对等体 p , $N(p)$ 是一个对等体节点集,该集合由 P 维护着一种连接。拓扑的管理主要功能是在网络动荡的情况下,维护重叠网络的结构的完整性。

(4)基于拓扑的寻径协议及其消息转发算法(协议),这是对等网络构成后对外提供的唯一服务,其建模为异步消息传送模式: $route(p,i,m)$ 转发消息 m 到负责 i 的对等体 p 上去。

基于上述分析,对等网络的基本功能因子构成如下:

$$O = (I, d, F_p, C_{F_p}, M, C_M, N, C_N)$$

其中, I 表示标识符空间, d 表示标识符空间的距离函数, F_p 表示对等体映射到标识符空间的映射函数, C_{F_p} 表示映射函数的属性, M 表示标识符空间的划分函数, C_M 表示空间划分函数的属性, N 表示对等体结构的邻居关系表示, C_N 表示邻居节点关系的属性。

1.2 功能模块间关系

功能模块间关系可描述为如图 1 所示

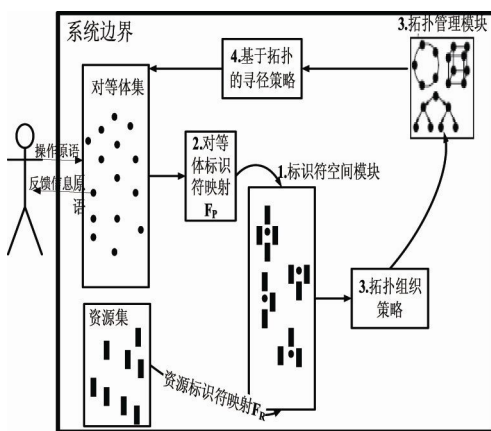


图 1 对等网络基本设计策略

图中, ●表示对等体, ■表示资源, 模块间执行关系可描述为: 首先, 预先定义标识符空间及其距离参数 d ; 其次, 对等体集或资源的原始集经过标识符映射算法映射到标识符空间; 再次, 标识符空间中对等体节点又依据拓扑组织策略(对等体间邻近关系)形

成拓扑(邻居表或映射表); 最后, 在邻居表和对等体标识的基础上, 执行消息寻径操作。

2 对等网络中间件的分层模型

传统中间件将对等网络基本功能封装为一层, 所有应用建立在此基础上, 难以扩展, 考虑到对等进程间自组织可靠通信的实际需求, 我们将确保通信可靠性的模块, 进程的管理模块, 自组织算法引入其中, 参考 ISO/OSI 模型分层, 在对等网络概念模型理论上, 将对等网络基本功能模块按其性质不同, 分别封装至各层, 使层次间任务接口清晰, 易于模块化和可扩展。

参照 ISO/OSI 七层模型, 将中间件模型分为三层: 虚拟寻径层, 虚拟传输层和管理控制层。如图 2 所示:

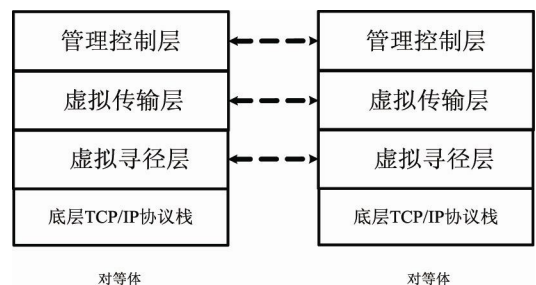


图 2 中间件模型

其中, 管理控制层, 对上层用户提供统一通信和进程管理原语, 其中通信原语包括发送(send)和交付(deliver); 而进程管理原语包括进程创建, 撤销, 运行, 就绪, 等待等原语; 本层负责组织进程间的会话, 负责进程组织和事件的触发, 其核心部件为事件控制模块。

虚拟传输层, 对上层提供连接(connect), 监听(listen), 消息发送(send), 接收(receive)和断开(disconnect)等操作。本层类似于七层模型的传输层, 负责控制消息有序, 可靠的传输, 主要包括进入消息模块和出去消息模块, 消息的格式为消息主体, 源节点标识, 目的节点标识, 消息序号, 消息属性。

虚拟网络层, 是模型的核心部分, 它对上层提供的接口包括: 其一, 对等体加入, 离开, 失效等操作; 其二, 对等体状态操作, 是否活跃; 其三, 对等体消息寻径操作: 寻径, 发送, 接收等操作; 其四, 映射表操作: 更新, 邻居节点查询, 本地节点查询。本层类似七层协议的网络层, 负责消息点

到点的传输,封装底层寻径细节,主要包括嵌入拓扑的标识符空间模块,寻径表模块以及基于寻径表的协议模块三部分组成。

分层模型细化对等网络功能模块,将它们分层封装,各有分工,大大提高了其功能可扩展性。

3 对等网络中间件模型的设计

基于分层模型,设计出分层中间件的结构框图,并详细设计每层的具体功能结构或关键算法。中间件的结构设计结构如图 3 所示,其接口包括:其一,管理控制层和应用进程间接口.主要是指管理控制层对上层所提供的 API 服务原语,包括进程管理原语(创建,撤销,运行,就绪和等待)和通信原语(发送,接收消息)。其二,虚拟网络层与 TCP/IP 的传输层接口,主要涉及到对等体标识与对等体实际物理地址端口映射(id,ip:port),以及 socket 调用。下面详细阐述各层的设计思想。

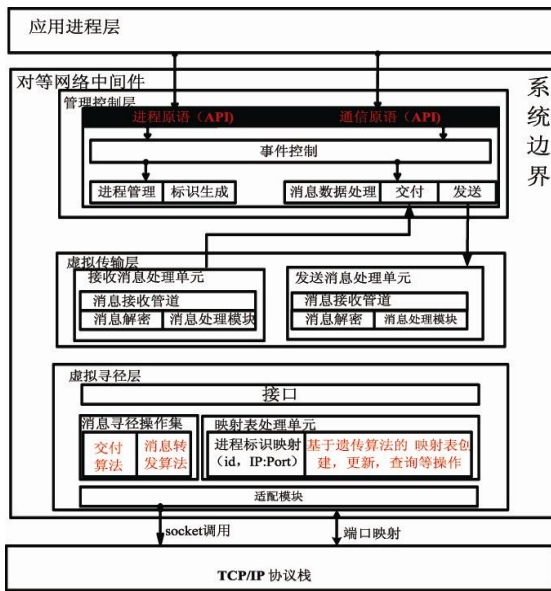


图 3 对等网络中间件的结构框图

3.1 管理控制层

管理控制层包括事件控制模块和通信创建模块,其中事件控制模块负责处理进程的加入,离开,创建等事件的调度和触发,通信创建模块负责生成消息头,并交付消息传输层来执行。该层封装了标识符映射算法重要模块,事件控制模块是该层的核心,其结构图如图 4 所示,事件执行模块最后负责将事件交付到其下层处理单元执行,事件主要包括:(1)进程管理和通

信相关的基础事件;(2)由用户生成的控制事件。

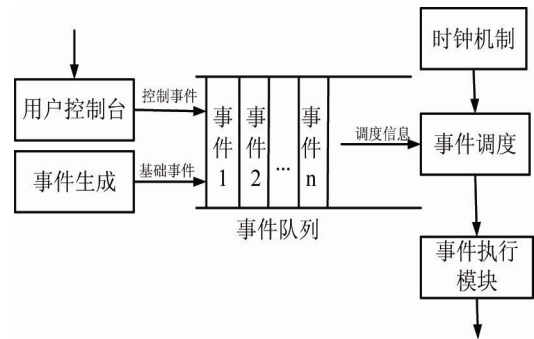


图 4 事件控制结构图

3.2 虚拟传输层

包括消息发送通道和消息接收通道两个模块。消息的执行过程,如图 5 所示。

消息发送时,管理控制层将消息内容(主体)交付给发送消息通道;经过对等协议模块处理消息,加入消息头;经过消息加密后,消息送入消息队列并等待被调度;被调度的消息交付给虚拟网络层传输。

消息接收时,虚拟网络层将消息入消息队列,消息解密;对等协议模块处理消息,去掉消息头;送给管理控制层,有管理控制层将最终消息内容交给应用进程。

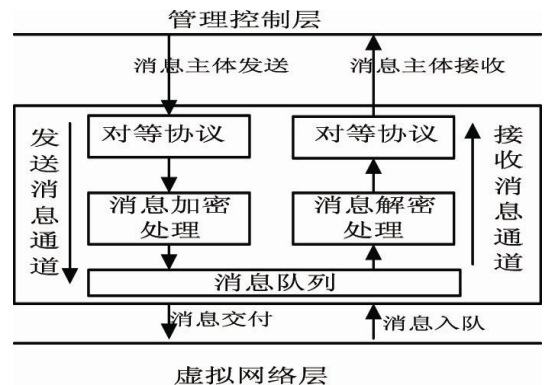


图 5 虚拟传输层消息处理流程图

3.3 虚拟消息寻径层

消息寻径层是整个中间件结构的核心部分,其主要功能模块包括映射表处理模块,和基于上述模块的消息寻径操作模块;它具有如下特点:

- (1) 对等体标识本身非简单哈希计算,引入了遗传算法,从父体经过遗传操作得到;
- (2) 对等体的标识符空间也是染色体空间,具备了生物界优胜劣汰的特征;

(3) 每个对等体的寻径表是其相邻对等体三元组(节点活动能力, 节点标识, 节点 IP 地址)构成, 由于三元组是遗传算法中构成染色体的基本单位基因利用遗传算法不断更新每个节点的路由表, 这同时解决了传统分布式路由算法中动态更新问题;

(4) 寻径中消息转发下一跳基于遗传算法。

基于遗传的消息寻径算法伪代码如下:

```

Begin
  While(x最优) do
    Begin
      select(x);
      if(x最优)
        then return x;
      else
        //在x上选择两个父体y, z;
        select(y,z)
        //对y, z执行交叉, 变异操作生成m;
        m=crossover(y,z);
        if(m大于变异率)
          m=mutation(m);
          xm;
    End
  End
End

```

图 6 基于遗传的消息寻径算法伪代码

4 验证

基于 DVE-IW 平台上对 50-100 个对等体进行验证, 仿真时间在 500m 内的得出如图 6 所示实验数据,

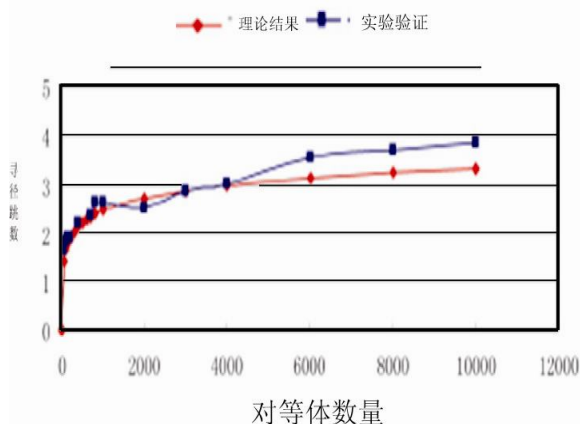


图 7 实验仿真数据

自组织寻径算法的在某个对等体失效后, 其可恢复性理论与仿真对比数据。

实验结果表明, 对等网络自组织中间件模型是正确而有效率的。

5 结束语

本文在定义、分析和抽象对等网络基本模块的基础上, 建立自组织的对等网络中间件模型, 同时给出对外统一服务原语, 为对等网络中间件提供了一种建模依据, 下一步尚需解决基于遗传的寻径算法的优化问题。

参考文献

- Gong L. JXTA: A Network Programming Environment. IEEE Internet Computing, 2001,5(3):88-95.
- Dabek F, Zhao B, Druschel P, Kubiawicz J, Stoica I. Towards a Common API for Structured Peer-to-Peer Overlays. IPTPS, 2003.
- P2psim <http://www.pdos.lcs.mit.edu/p2psim>
- He Q, Ammar M, Riley G, Raj H, Fujimoto R. Mapping Peer Behavior to Packet-level Details: A Framework for Packet-level Simulation of Peer-to-Peer Systems. Proceedings of MASCOTS 2003, Orlando, FL, 2003.
- Aberer K, Alima LO. The essence of P2P: A reference architecture for overlay networks. Proceedings of the Fifth IEEE International Conference on Peer-to-Peer Computing, 2005.
- Schollmeier R. A definition of peer-to-peer networking for the classification of peer-to-peer architectures and applications. Peer-to-Peer Computing, 2001.
- IRTF. Except from the Chapter of PTP Research. <http://www.irtf.org/charter?gtype=rg&group=p2prg>. Jun2003. up.