

文章编号:1001-9081(2008)02-0394-03

## 一种可扩展的移动 Agent 容错组管理机制

王娟,黄永忠,武林,刘沙

(信息工程大学 信息工程学院,郑州 450002)

(guss@163.com)

**摘要:**把组通信技术引入到移动代理系统中,对建立可靠的移动代理系统有着重要意义。容错组管理机制是组通信机制的重要组成部分。首先提出了组通信系统框架,并在此框架的基础上提出了 Agent 容错组管理机制,这种容错组管理机制保证了组成员管理的灵活性、可扩展性,并且保证了故障检测完整性、精确性和网络负载低的要求。

**关键词:**容错;组成员管理;心跳检测;移动 Agent

**中图分类号:**TP391 **文献标志码:**A

## Scalable fault-tolerant mechanism of mobile Agent membership management

WANG Juan, HUANG Yong-zhong, WU Lin, LIU Sha

(Institute of Information Engineering, Information Engineering University, Zhengzhou Henan 450002, China)

**Abstract:** The group communication technology plays an important role in building a reliable mobile Agent system. The fault-tolerant membership management mechanism is a significant part of the group communication mechanism. After presenting a system framework of the group communication, based on this framework, a new mechanism of fault-tolerant membership management was brought forward, which can guarantee the flexibility and expansibility of the membership management. What's more, it can meet the needs of the integrity and accuracy of the fault detection and low net load.

**Key words:** fault-tolerant; membership management; heart-beating; mobile Agent

### 0 引言

移动 Agent 技术是近年来发展起来应用于分布式研究方向的一项新技术,移动 Agent 具有移动性、自主性、减少网络负担、延迟等特性<sup>[1]</sup>,目前越来越多的研究机构致力于移动 Agent 技术的研究。但是目前移动 Agent 系统中的通信机制大都是一对一的进程通信,对组通信的研究仍存在不足。组通信是保障分布式系统可靠性的有力工具,对于建立基于移动 Agent 的可靠分布式系统有着重要意义<sup>[2]</sup>。在通常情况下,一项任务是由多个 Agent 共同完成的,引入组通信机制,可以提高通信效率,减少通信开销。要实现移动 Agent 组通信机制,首先是保证消息传递的可靠性和原子性;其次是组成员管理,及时发现有故障的 Agent,并把它移出多播组。这就要求有一种高效的组管理机制,可以准确及时地发现问题 Agent,并把它移出多播组。

### 1 移动 Agent 组管理问题

移动 Agent 组管理机制要解决两方面的问题<sup>[3]</sup>,一是维护组成员视图,处理移动 Agent 加入和撤出组的请求,二是检测出故障 Agent,把它移出多播组。

当成员失效、成员的加入或者退出等引起视图的变化时,必须要更新视图,以便使 Agent 知道其他组成员 Agent 的存在。

传统的分布式故障检测中常使用 all-to-all 心跳 (Heart-Beating) 故障检测,每一个组成员定期 (心跳周期) 向其他的组成员发送一条心跳消息,目的是说明此组成员还有心跳,也即此成员没有发生故障。如果连续  $n$  个周期内非故障组成员

$i$  都没有收到组成员  $j$  的心跳消息,则可判断成员  $j$  出现故障。

对于 all-to-all 心跳检测,最简单的实现方式是每一个心跳消息都发送到其他成员,这种故障检测方法满足完整性要求;但是对精确性要求而言,它满足弱精确性,这种方法最主要的问题在于网络负载比较大。

为了解决网络负载过大的问题, Van Renesse 等人提出一种闲聊 (gossip) 风格的心跳传播协议,其算法思想是<sup>[2,5]</sup>: 每一个成员都维护它所能知道的所有成员的信息列表,每一个列表项为每个 Agent 的 ID 和心跳计数器,在每个时间周期  $T_{gossip}$  内,每个活动的成员在列表中增加其对应的心跳计数器,并随机选择几个成员发送此列表,一个成员接收到此列表,比较列表中的计数器,取最大的计数器并入自己的最新列表中,每个成员还为列表中的成员维护一个上次心跳计数器更新时间,如果在预先定义的时间周期  $T_{suspect}$  ( $T_{suspect} = n \times T_{gossip}$ ) 过后,不能收到使某个成员的心跳计数器增加的心跳消息,则此成员疑为失效。但是由于其选择发送心跳消息的成员是随机的,可能会出现某一时刻,几个 Agent 同时向一个 Agent 发送消息,而其他成员收不到任何心跳消息,这样造成了网络带宽被浪费的现象。

移动 Agent 具有移动性,使得移动 Agent 组通信不同于网络层的组播通信,网络层的组管理协议会把由于计算需要移动到其他主机上的 Agent 误以为 Agent 已经移出多播组,在 Agent 到达新的节点时重新加入多播组。因此,必须有一种适合移动 Agent 的容错组管理机制解决移动 Agent 的移动性问题,为了方便叙述容错组管理机制,引入了组通信系统的结构。

收稿日期:2007-08-28;修回日期:2007-11-13。 基金项目:军队重点基金资助项目(BF2000230)。

作者简介:王娟(1980-),女,山东德州人,硕士研究生,主要研究方向:分布式计算、软件工程; 黄永忠(1968-),男,湖北咸丰人,教授,主要研究方向:分布与并行处理、软件工程; 武林(1983-),男,北京人,硕士研究生,主要研究方向:网络计算、软件工程; 刘沙(1984-),男,江苏盐城人,硕士研究生,主要研究方向:分布式计算。

### 2 移动 Agent 组通信系统

一个移动 Agent 计算网络包括移动 Agent 及移动 Agent 平台 MAP,MAP 为 Agent 提供行环境,提供移动和通信支持,运行平台一般都包括以下基本服务:生命周期服务、事件服务、目录服务、安全服务、应用服务。为了支持组通信,把同一个组中的成员按照地理位置划分为几个域,每个域中的 Agent 位于同一个子网的 MAP 上,便于实现组通信。借鉴文献[4]中的思想,在每个域中引入一个新的模块,称之为监控模块。监控模块负责本域内的成员的消息通信和本域内成员的管理,包括处理其加入和退出组的请求。同时系统中还引入一个集合点 CP,CP 用来维护系统中 Agent 的组视图,组视图是一个三元组(组 ID,域 ID,组内 Agent 列表项),组内每一个 Agent 都知道 CP 的位置。

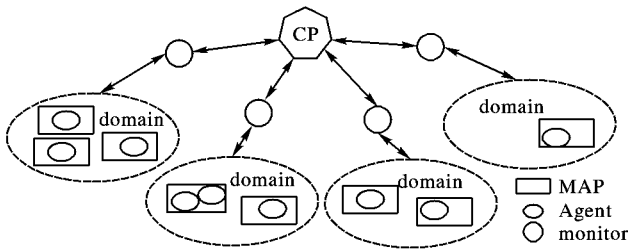


图 1 组通信系统结构

由图 1 可知,组通信系统自底向上由三层组成,依次是每个域中的 Agent、监控模块和 RP。这种管理方式可扩展性比较好,可以比较容易地支持容错。

在容错组管理机制中,主要关注崩溃型故障,如 MAP 崩溃或移动 Agent 崩溃,而不考虑灾难性故障,如一个组内所有移动 Agent 都崩溃。

### 3 一种可扩展的 Agent 容错组管理机制

组管理机制的主要操作为:

- 1) 组的创建。创建一个组。
- 2) Agent 的加入或撤出组。监控模块处理 Agent 的加入或者撤出请求,并把此信息报告给 CP,CP 更新组视图。
- 3) 故障检测。检测出现故障的 Agent,把它移出多播组。
- 4) 组视图更新。当组内成员发生变化时,更新组视图。

新成员 本地监控模块 CP 其他监控模块

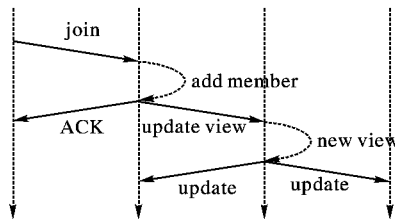


图 2 新成员加入

组管理算法如下:

```

GroupManager{
  Create( groupID );
  Join( AgentID, groupID, dominID );
  Quit( AgentID, groupID );
  ViewUpdate();
  Detector();
}

```

创建组、Agent 加入或者主动撤出组(如图 2、3 所示)的操作比较简单,在本文中,主要研究内容由两部分组成:故障

检测机制和组视图更新。

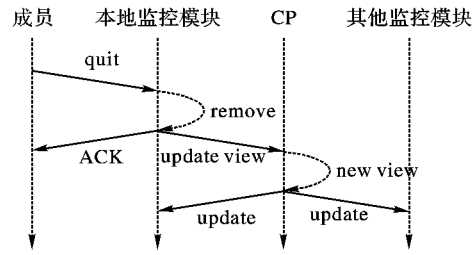


图 3 组成员撤出

#### 3.1 故障检测机制

##### 3.1.1 相关概念

故障检测是组成员管理的一部分,其目的就是检测出出现故障的 Agent,为了清楚地描述故障检测器,引入一个定义。

**定义** 每一个 Agent 成员列表项是一个五元组( domain, Agent ID, Counter, Time, Flag)。domain 是成员所在的域, Agent ID 是创建时分配给每一个 Agent 的唯一标识, Counter 是计数器的计数值, Time 是上次心跳消息发送的时间, Flag 是 Agent 是否失效的标志,可以及时把出现故障的 Agent 移出多播组。

在上述定义中, Flag = 0 表示此成员没有发生故障; Flag = 1 表示此成员在第一个预定义的时间周期内可能出现故障,也就是可疑的; Flag = 2 连续的两个周期内此成员都没有发送心跳消息,故障检测器判定此成员出现故障。系统默认值为 Flag = 0。

在大规模的计算网络中,其网络由几个子网组成,如果故障检测直接采用组内成员心跳检测,则网络负载会加大,也可能会出现网络拥塞;同时由于网络延迟等问题的存在,会对检测的完整性和精确性产生影响,从而加大检测难度。为了实现低负载、可扩展的故障检测器,Agent 的故障检测由监控模块监控本域内组成员 Agent,处理 Agent 加入和撤出组的请求,检测有故障的 Agent,同时把消息报告给 CP,更新组视图。

##### 3.1.2 域内故障检测

基于闲聊的心跳传播协议存在两个缺陷:1)可能会导致错误的检测;2)如果多个成员在某个周期内向同一个成员发送消息,那么网络带宽有可能被浪费。为解决上述问题,对协议进行了优化,主要有以下两方面:

(1)对任意一个 Agent 选择发送消息的目标 Agent 有一个限定,即在每一个回合内,每个组成员发送心跳消息的目的地由下述表达式决定:

$$\text{目的 Agent ID} = (\text{发送 Agent ID} + 2^{r-1}) \% N, \\ 1 \leq r < \log N, r \in \text{整数}$$

在上述表达式中,r 为发送消息的回合数,N 为组成员的个数。由此表达式可知,对其中一个成员来说,在每个时间周期内,其发送消息和接收消息都是均匀的,且不会出现一个成员 Agent 向某个 Agent 连续不断地发送心跳消息,从而使其检测精度提高。而且这种心跳消息的通信模式确保所有成员将在一个有限的 log N 回合内收到一个指定成员的“心跳”计数器更新值。

(2)设置标志位 Flag,提高了检测的可靠性。每个成员 Agent 列表项增加 Flag 标志位,通过两个周期检测,可以有效地提高检测精度。

基于上述改进的闲聊心跳传播协议,进行组管理机制中的域内故障检测。检测中保留了原有心跳传播协议中每一个成员都维护它所能知道的所有成员的信息列表,同时在每个时间周期  $T_{gossip}$  内,每个活动的成员在列表中增加其对应的心

跳计数器,并设定  $2 \times T_{\text{suspect}}$  为监控模块检查 Agent 的周期。具体故障检测过程如下:

a) 在一个预定的时间周期  $T_{\text{suspect}}$  过后,如果 Agent  $i$  检测到 Agent  $j$  可能出现故障,则将其列表项中对应 Agent 的 Flag 置为 1,则此 Agent 是可疑故障 Agent。

b) 第二个预定周期过后,如果 Agent  $j$  的计数值没有增加,则把此 Flag 位置为 2,即此 Agent 发生故障。如果 Agent  $j$  的计数值增加了,则把 Flag 位置为 0,即此 Agent 工作正常。

c) 在  $2 \times T_{\text{suspect}}$  过后,监控模块随机抽取一个 Agent,要求此 Agent 把其成员列表项发送给监控模块,监测模块检查成员列表项,如果发现某个 Agent 的 Flag 位为 2,则把此 Agent 撤出 Agent 组,同时通知 CP,更新组视图。

综上所述,每一个预定的周期过后,首先检查 Agent 对应的计数值,如果计数值没有变化,再修改其 Flag 位。

### 3.2 组视图更新及一致化

当组成员失效、加入或者撤出组时,会引起视图变化,因此必须要更新组视图。因为其中某个 Agent 向其他组成员发送组播消息时,必须先获得组成员视图,得到组内 Agent 的成员列表,避免错发消息,所以视图更新及一致化是组成员管理的一个重要部分。

#### 3.2.1 组视图更新

组成员 Agent  $i$  在运行某项计算任务时,必须首先得到系统的组视图,得到组内 Agent 的成员列表,才能向组成员 Agent 发送多播消息及计算数据。但是,系统的组视图可能时刻发生变化,一般情况下可以把改变的组视图在组内向各组成员多播出去,达到组视图一致的目的,保证任务的正常运行,但是这种方法会加大网络负载,并且如果这时某个域中的 Agent 不需要向其他成员发送消息,就不需要接收新的组成员视图,从而可以提高效率。

因此为了解决全部组成员多播带来的网络负载过大问题,在动态组管理机制中,设定在组成员视图改变后,CP 只是把改变后的视图组播到各监控模块,如果某个组成员 Agent 需要向组内成员发送多播消息,则它可以从监控模块处取得组视图。当成员的加入和撤出或者成员失效引起组视图发生改变时,其组视图更新过程可以由图 2 和图 3 表示出来。

#### 3.2.2 视图一致化

移动 Agent 组管理机制是分层的管理机制,每层的组件完成相应的任务。为了保证视图一致化,组管理机制由单独集中的模块管理组视图。同完全分布式组通信机制相比,该模块管理组成员的加入和退出,同时对视图进行更新,大大简化了组成员 Agent 的管理。

在本系统中,当组成员发生变化时,首先通知 CP,由 CP 更改组视图,然后组播到各监控模块,则各监控模块中保存的都是最新的组视图,保证了组视图的一致性。当某个 Agent 需要同组成员通信时,首先从监控模块中取得组视图,由于每个监控模块中都是最新的 Agent 的组视图,从而 Agent 可得到的组视图是最新的,可以保证计算任务的正常运行,同时在执行任务的过程中,Agent 本身不保存组视图,每次都由监控模块处取得最新的组视图,这种方式也保证了组视图的一致。

## 4 动态组管理机制性能分析

故障检测机制是移动代理系统中必不可少的组成部分,研究者从理论和实践的角度对分布式故障检测机制精确性和

可测量性等特性进行了研究和分析,这些特性主要包括<sup>[2]</sup>:

**完整性** 如果某个成员发生故障,它将最终被其他成员检测到。

**精确性** 避免出现判断错误,存在一个时刻  $t$ ,保证时刻  $t$  之后任何有效的移动 Agent 不会被其他有效的移动 Agent 怀疑为无效。

**网络负载**:故障检测产生的网络流量。

本文提出的故障检测器属于非可靠的故障检测器,现依据以上三个特性对故障检测机制进行分析。

1)完整性。基于改进的闲聊风格的心跳传播协议的故障检测器满足完整性的属性,这种心跳消息的通信模式确保所有成员将在一个有限的  $\log N$  回合内收到一个指定成员的“心跳”计数器更新值,因此如果 Agent  $i$  出现故障,它就会停止发送心跳消息。因此,在  $\log N \times T_{\text{gossip}}$  时间内,Agent  $i$  将会被其他活动 Agent 检测到,改进的心跳检测保证出现故障的 Agent 终会被其他活动 Agent 所检测到。

2)精确性。基于改进的闲聊风格的心跳传播协议的检测精确性比原算法有所提高,在 Agent 成员列表项中设置 Flag 位,使得检测精确性提高。即使某个活动 Agent 在第一个  $\log N \times T_{\text{gossip}}$  时间内被误判断为出现故障,在第二个  $\log N \times T_{\text{gossip}}$  时间内,如果收到了心跳消息,则可以纠正错误,对由于有限时间的网络延迟造成的心跳消息延缓接收问题有一定的缓解作用。

3)网络负载。相对于闲聊风格的心跳传播协议来说,网络负载比较小。基于闲聊的心跳传播协议对组内所有成员进行检测,这种方法适用于组成员比较少的情況。但是在组成员比较多的情况下,带来的负载问题同样不容忽视,同时对于不同子网中的组成员,这种检测方法的精确性会降低。为了解决这个问题,引入了监控模块,由监控模块处理子网中的 Agent 的请求,只在一个域中应用改进的心跳检测,这种心跳检测方法使得网络负载随着组成员的增多成对数级增加,网络负载相应也会减少。

## 5 结语

本文提出的动态组管理机制,可以实现 Agent 容错组成员的管理,处理组成员的加入和撤出组的要求,检测出现故障的 Agent,强制其退出多播组。采用三层的分析结构可以实现组管理的灵活性、扩展性,同时故障检测具有完整性、弱精确性、低负载的属性。但是在此管理机制中,当几个域的监控模块同时把视图改变的消息发送到 CP,可能出现网络拥塞现象。因此,必须提供可靠的消息通信机制,保证消息提交的次序性。

### 参考文献:

- [1] 朱森良,邱瑜.移动代理系统综述[J].计算机研究与发展,2001,38(1):16-25.
- [2] 徐伟.面向服务的移动 Agent 组通信研究[D].北京:中国科学院软件研究所,2005.
- [3] 卢朝霞,曾广周,杨公平.一种基于复制组管理的移动 Agent 容错模型[J].计算机工程与应用,2005,41(11):26-29.
- [4] STELLING P, FOSTER I. A fault detection service for wide area distributed computations [C]// Proceeding 7th IEEE Symposium on High Performance Distributed Computing. Redondo Beach: IEEE Computer Society, 1998: 268-278.
- [5] 刘畅,刘西洋,陈平.一种可靠可伸缩组通信系统设计与实现[J].计算机科学,2005,32(10):46-48.