

Grid 和 P2P 混合环境中一种基于信任的资源搜索机制^{*})

周金洋 杨寿保 郭磊涛 王莉苹

(中国科学技术大学计算机科学技术系 合肥 230026)

摘要 Grid 和 P2P 两种分布式计算模式中的资源搜索算法均假设节点提供可靠的资源,但 Grid 和 P2P 混合计算环境的动态、异构、自组织等特点使得一些节点存在冒名和提供虚假服务等行为。本文对基于经验和最好邻居搜索机制进行改进,引入信任因子,提出了基于信任的资源搜索机制。该机制有效抑制了欺骗行为,提高了资源搜索的可靠性和安全性。

关键词 资源搜索, 网格, 对等网络, 信任, 信任值

Trust-based Resource Searching Mechanism in Grid and P2P Integrated Environment

ZHOU Jin-Yang YANG Shou-Bao GUO Lei-Tao WANG Li-Ping

(Department of Computer Science, University of Science and Technology of China, Hefei 230026)

Abstract The resource searching algorithms in distributed computing environment of Grid and P2P are based on the assumption that reliable resources are provided. But the dynamic, heterogeneous and self-organizing features of the grid and P2P integrated computing environment have indulged some malicious behaviors, such as imposture and providing pseudo services. In this paper, the search mechanism based on experience and best neighbor is improved, the trust factor is introduced and the resource searching mechanism based on trust is put forward. The new searching mechanism can effectively improve the reliability and security of searching information through the experiences.

Keywords Resource searching, Grid, Peer-to-Peer, Trust, Trust value

1 引言

网格计算(Grid Computing)^[1]和对等网络计算(Peer-to-Peer Computing)^[2]是利用主机剩余资源进行协同计算和分布存储,以获取高计算性能和大存储容量的计算模式。在 Grid 和 P2P 混合计算环境中,节点的自组织性和匿名性不能保证所有节点都能提供诚实的服务,一些节点可能存在恶意欺诈行为,从而导致了所搜索到的资源的不确定性和非法性。因此,若能通过对资源搜索算法的改进尽量避免恶意节点的欺诈行为,就能有效提高资源搜索的可靠性和安全性,从而进行更为可靠的协同计算。根据这一基本事实,本文基于信任对经验和最好邻居搜索算法进行了改进,提出了基于信任的资源搜索机制,并进行了仿真实验,给出了实验结果。

本文第 2 节介绍了 Grid 和 P2P 混合计算环境以及资源搜索算法;第 3 节对信任关系及其一些基本的概念进行了描述;第 4 节提出在 Grid 和 P2P 混合计算环境中基于信任的一种资源搜索机制;第 5 节模拟仿真及其结果分析;最后是结束语。

2 相关工作

在描述 Grid 和 P2P 混合计算环境中基于信任的资源搜索机制之前,首先了解一下混合计算环境的基本架构、常见的资源搜索机制。

2.1 Grid 和 P2P 混合计算环境

Grid 和 P2P 两种分布式系统具有许多共同特性:分布、动态、异构、多样等。而二者在服务质量、安全机制等方面又有诸多不同。一个显著的区别就是 Grid 提供有 QoS 保证的服务,而 P2P 则不提供。目前,纯粹的 Grid 系统在某些情况下不能满足大规模计算要求,因而出现了 Grid 和 P2P 相互融合的趋势,我们称之为 Grid 与 P2P 混合计算环境,其结构如图 1 所示。图 1 上层表示网格节点 GN(Grid Node),中间层是超级节点 SN(Super Node),下层是对等网络节点 PN(P2P Node)。超级节点是由下层的 peer 节点所选举出来的信誉好、可靠性高的节点,作为和网格节点与对等节点交互的中间节点。上层的网格节点是资源和作业的调度者与管理者,既可以获取资源也可以提供资源。基于 Grid 架构特征,网格中主要考虑计算环境的动态主体特性,保证网格计算环境中不

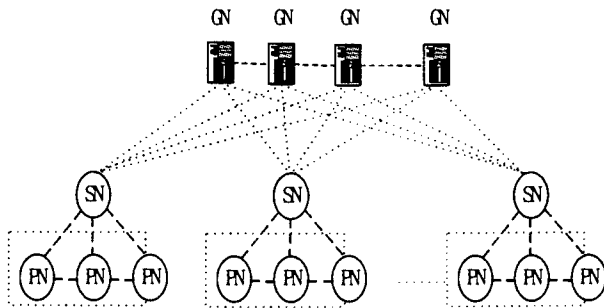


图 1 Grid 和 P2P 的混合计算环境

^{*})本文得到国家自然科学基金项目(编号 60273041)和国家 863 计划(编号 2002AA104560)的资助。周金洋 硕士研究生,主要研究方向:网格安全和 P2P 信任模型。杨寿保 教授,博士生导师,主要研究方向:网格计算、密码学和网络安全。郭磊涛 博士研究生,主要研究方向:网格安全和密码学。王莉苹 硕士研究生,主要研究方向:网格 P2P 安全。

同主体之间的相互鉴别和各主体之间通信的保密性和完整性。为此,网格中主要采取了安全认证、安全身份相互鉴别、通信加密、私钥保护以及单点登录等安全手段,从而使网格系统能够保证较好的服务质量。然而在自组织的下层 P2P 系统中,所有 P2P 节点自愿地提供计算服务,运行在其上的计算任务是不可再分的原子任务,适合于可分解的计算密集性任务。其节点可以自由地加入和退出,并且还能够根据网络带宽、节点数、负载等变化不断地做自适应式的调整。这就使得 P2P 网络在提供了系统良好的可扩展性的同时,也增加了欺诈,提供虚假数据,节点的恶意退出等不良行为的风险性,对 Grid 来讲 P2P 不能保证提供 QoS 保证的服务。因此,在 Grid 和 P2P 混合的计算环境中引入信任机制,作为获取可信资源的一种补充方式,从而保证混合计算环境中资源的安全性。在本文中采取的方法是对每个 Grid 节点和 P2P 节点评定信誉度,选择符合 Grid 节点预期信誉的 P2P 节点,允许这些 P2P 节点成为混合计算环境中可靠节点,同时选择多个可靠节点作为数据的冗余备份,使得 Grid 和 P2P 混合计算环境提供良好的服务质量。

2.2 资源搜索机制

在 Grid 和 P2P 混合计算环境中一个最基本的服务就是资源搜索。所谓的资源搜索机制主要是为了发现资源的结构和状态信息,给出期望资源的一个具体的描述,发现一系列和资源描述相匹配的资源。在目前 Grid 系统中资源发现定位主要采取了依据 MDS(元计算目录服务)的请求-转发策略。这里介绍基于经验(Experience)和最好邻居(Best Neighbor)的请求-转发策略^[3]:当收到请求时,如果没有所需资源,则根据以往的经验,将请求转发给曾经响应过类似请求的节点;若没有相关经验,则将请求转发给它的最好邻居^[3];如果既没有相关经验又没有最好邻居存在,则使用随机转发策略(flood-^[4]ing^[4]算法)进行转发。

关于 Grid 资源搜索机制效率的模拟在文^[3]中具有详尽的描述。Grid 能够提供 QoS 保证服务,而在 Grid 和 P2P 混合计算环境中,存在着一些恶意节点的欺诈行为,为了保证混合计算环境的可靠性,提供一个 QoS 保证的可靠网络必不可少。因此,本文在 Grid 和 P2P 混合计算环境中,对基于经验和最好邻居搜索机制进行改进,引入信任因子,提出了基于信任的资源搜索机制。

3 信任关系

本文我们使用 Diego Gambetta 给出的一个比较完整的信任定义:“信任(或相对来说的不信任)是一个 Agent 评价其他 Agent 或 Agent 团体实际行为的主观可能性程度,评价在对该行为进行监控(或根本不可能监控该行为)之前和与该行为对其自身行为产生影响的情况下进行”^[5]。在这里我们给出 Agent 的概念是为了和静态的实体相区分(如打印机等),在网络中所有的实体都可以成为提供资源的节点,但只有 Agent 可以成为请求-转发节点。信任具有以下几个重要特征:(1)主观性,不同的个体对同一事物的看法会受个体喜好等因素影响而有所不同;(2)可能性预期,信任由那些不能监视到的行为所影响;(3)内容相关,信任是对事物的某个方面(完成某项任务的能力)而言的。

当代理 A 确定相信代理 B 时,则将在二者之间建立一个信任关系(Trust Relationship)。信任关系可以分为两类:直接信任(Direct Trust)和推荐信任(Recommendation

Trust)^[6]。如果在两个节点之间存在直接交互行为,则建立直接信任关系(用实线来进行表示);如两节点之间的交互行为需要推荐节点,则建立推荐信任关系(用虚线来进行表示),即推荐信任是在给定的信任等级上由一系列的诚实推荐节点所得。特别注意,信任关系是单向不可逆的(即如果 A 在某信任类别的信任等级上信任 B,但并不代表着 B 也信任 A),如图 2 所示。

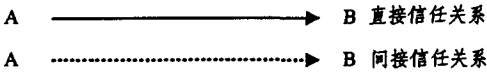


图 2 信任关系

本文采用信任类别(Trust Class)和信任等级(Trust Level)的概念来描述每一对实体之间的信任关系。信任类别是一个代理相对其他代理在某段时间对于某种行为的信任。信任等级是基于某种信任类型的对其他实体的信任程度。在本文中我们使用离散的数值来表示其直接信任以及推荐信任等级。其表示如表 1 所示。

表 1 信任等级表

信任值	含义
-1	不信任
0	不能判别是否有信任关系
1	最低信任级别
2	一般信任级别
3	较好的信任级别
4	完全信任级别

4 基于信任的资源搜索机制

4.1 基于信任的框架设计

Grid 和 P2P 是分布式的网络结构,没有固定的网络拓扑结构,其网络体系结构都是在节点的加入和退出中动态自发形成的。Grid 和 P2P 混合计算环境中具有种类繁多的各类计算资源:这里主要包括静态主机信息(操作系统及其版本号、CPU 类型、CPU 数目等)、动态直接信息(平均负载、运行的进程数等)、存储系统信息(可用磁盘空间、总磁盘空间等)以及可以获得的网络信息(网络带宽、网络延迟等)。这里把混合计算环境逻辑上划分为两个虚拟系统:Grid 系统和 P2P 系统,其间通过网络相连接。其结构如图 3 所示。

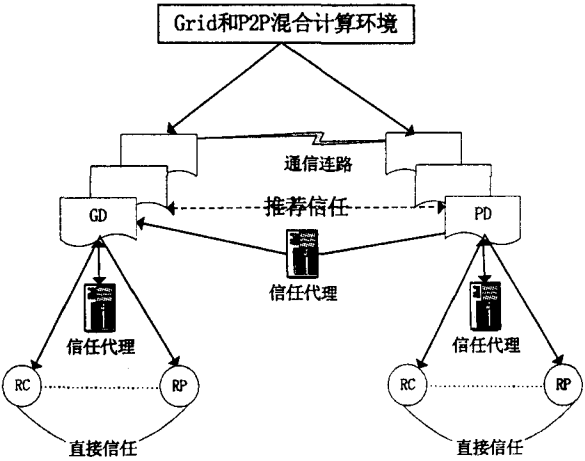


图 3 基于信任的混合计算环境

Grid 和 P2P 被分别划分为一系列的虚拟自治域,称为

Grid Domains(GDs)和 P2P Domains(PDs)。对每个 GD 和 PD 都由一系列 RP(Resource-Provider)和 RC(Resource-Consumer)自治管理节点组成。在系统中 RP 和 RC 节点并不严格区分,在某些情况下可以映射到相同的虚拟自治域,在 GDs 和 PDs 中存在一系列的信任代理(Trust Agent)。同样,虚拟的 Grid 系统和 P2P 系统之间也存在着信任代理。系统中的信任代理用来访问控制信任等级表(Trust Level Table)。对于虚拟域中的每个节点,建立对应于信任代理的信任值。如果事务处理所形成的新的信任关系的信任值与信任等级表中的信任等级不同,则代理更新信任值,并且同时复制更新其他虚拟域中的信任代理表的值。

在对基于经验和最好邻居的搜索策略中,系统中需要每个节点维护一个经验信息表来表示其经验和最好邻居。所谓的经验信息就是记录响应请求的节点标识和响应类型。所谓最好邻居,即响应相应类型请求次数最多的节点。其表结构如表 2 所示。

表 2 经验信息表

请求节点标识	响应节点标识	响应请求类型	响应次数
RC _i	RP ₁	CPU 负载	4
	RP ₂	内存负载	6
			
	RP _i	磁盘容量	2

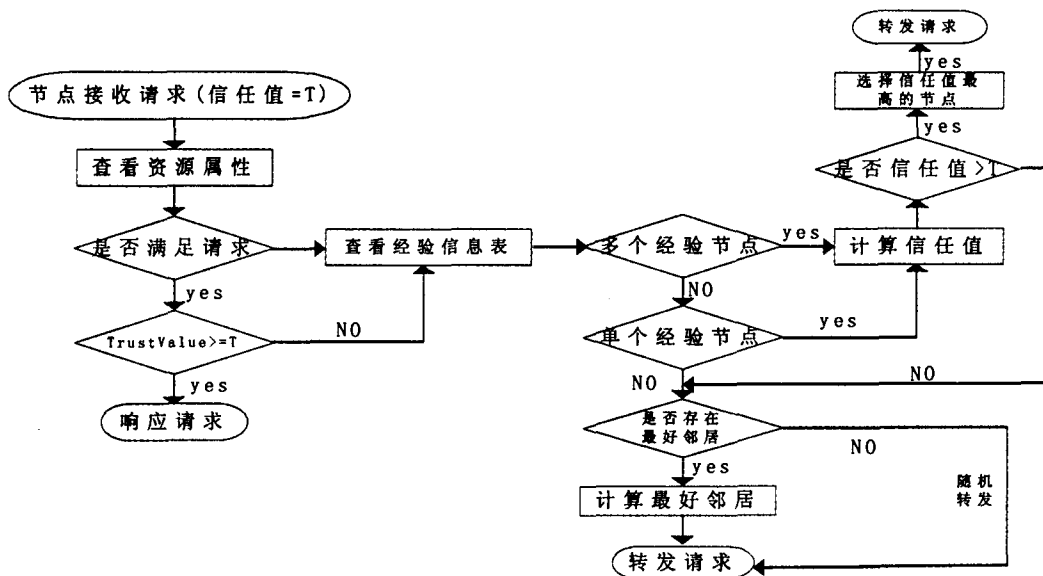


图 4 资源搜索流程图

4.3 信任值的计算

以上流程图其关键问题主要集中在对信任值的计算上,在此主要讨论有关信任值计算问题,由图 3 中的信任代理来管理。

(1) 域内两节点信任值的计算

对于单个 GD 域或者单个 PD 域内的两实体之间的信任值计算公式如下所示^[7]:

$$T_{ab}(b) = \begin{cases} t.dir_a^b(b) & (a) \\ t.rec_a^b(i_1)/4 \times t.rec_{i_1}^b(i_2)/4 \times \dots \times t.dir_n^b(b) & (b) \\ -1 & (c) \end{cases}$$

(a)直接信任关系;(b)间接信任关系;(c)不信任。

4.2 资源搜索机制流程

在此改进请求-转发机制:经验+最好邻居+信任。可从以下三个方面引入信任的因素:1)响应节点资源信任等级要等于或高于请求节点资源的信任等级;2)在推荐转发节点时,选择信任等级最高的节点;3)引入信任因子,改进最好邻居的定义(详述在 4.4 节)。下面给出改进后的请求-转发机制。节点接收到一个请求之后,按照如下步骤进行处理,其流程如图 4 所示。

(1)节点接收到一个请求,搜索自身看能否提供所需资源,存在且信任等级高于或等于请求资源信任等级则响应请求,并且同时在二者之间建立一个直接信任关系;如果不存在相匹配资源,则转到(2)。

(2)查看经验信息表,分下面三种情况:

a. 多个节点响应过类似请求:分别计算该节点对这些节点的信任值,选择信任值最高的节点转发。否则执行 b。

b. 单个节点响应过类似请求:验证该节点相应资源类型的信任值,若大于等于请求资源类型的信任值,则将请求转发给该节点;否则转向(3)。

c. 没有节点响应过类似请求,转向(3)。

(3)计算最好邻居,并将请求转发给最好邻居。如无最好邻居的信息,则转到(4)。

(4)采用随机转发策略,选择信任值最高的节点进行转发。

本文采用 $T_{ab}(b)$ 来表示同一虚拟域内两实体基于某种信任类别的信任值, $t.dir_a^b(b)$ 表示基于某种信任类别两个实体 a 和 b 之间的直接信任值, $t.rec_x^b(y)$ 表示基于信任类别实体 x 对实体 y 的推荐信任值,其中, $i_1, i_2, \dots, i_n$ 表示 a 和 b 之间的推荐实体。此公式是由一条推荐路径来进行信任值的计算,如果同时有多条推荐路径,我们则取多个推荐信任值的平均值来作为其最终的推荐信任值。

(2) 不同虚拟域间两节点信任值计算

假设 i 和 j 两个节点分别位于两个不同的虚拟域 M、N 当中,其虚拟域中的信任代理的标识分别记作:m,n,不同自治域中两个节点 i 对 j 的信任值 $T_{ij}(j)$ 的计算步骤如下:

1. 首先查找信任代理关系表,分别计算 i,j 对应的本虚

拟域中信任代理的信任值, $Td_f(m)$ 和 $Td_f(n)$ 的值。

2. 通过信任代理来计算两虚拟域间的信任值 $Tv_M(N)$: 查找信任关系表, 如果两个信任代理之间存在直接信任关系, 则其信任值为 $T.dir_M(N)$, 否则 M 域的信任代理向其他域发出信任评估请求, 各个虚拟域的信任代理通过推荐的方式来决定 M 和 N 两域之间的间接信任关系, 其值表示为,

$t.rec_M(N) = \frac{\sum_{i=1}^n t.dir_{x_i}(N)}{n}$, 其中 x 表示其他的评估域对于域 N 的直接信任评估, n 表示和域 N 有直接信任关系的域的个数。

3. 计算不同域中两节点的信任值, 公式如下: α, β, γ 分别表示其权重

$$Td_f(j) = \begin{cases} \alpha \times Td_f(m) \times \beta \times Td_f(n) \times \gamma \times t.dir_M(N) & (a) \\ \alpha \times Td_f(m) \times \beta \times Td_f(n) \times \gamma \times t.rec_M(N) & (b) \\ -1 & (c) \end{cases}$$

(a) 域间直接信任; (b) 域间间接信任; (c) 不信任。

4.4 最好邻居的计算

设节点 A 的经验信息表中记录了每个实体相应信任类别的响应情况, 对节点 A 相联系的节点分别依据信任类别的响应次数来计算其邻居值, 我们用来 $Neighbor_m^x(n)$ 表示节点 m 对节点 n 在某个信任类别上的最好邻居值:

$$Neighbor_m^x(n) = \alpha \times N_m^x(n) + \beta \times Tc_m^x(n)$$

其中 $N_m^x(n)$ 表示节点 m 的经验信息表中记录的基于某种信任类型的节点 n 的响应请求次数, α, β 为权值。计算出来的 $Neighbor_m^x(n)$ 值最大的节点就是最好邻居节点, 即在最好邻居的计算中考虑了信任的因素。

5 模拟分析

该仿真实验用来评测前面所描述的基于信任的资源搜索机制。在仿真实验中, 把基于各种类型的信任简化为对各个节点的信任, 即一个节点所拥有资源的信任值是一样的。并且在实验中我们假定认为所有网格节点自身为可信节点, 即其所有节点信任等级满足资源请求的信任等级。因此在这里我们只是模拟从混合网络中的 P2P 节点中寻找信任等级不符合资源请求的恶意节点以及符合要求的诚实节点的数目。本文假设网格节点为 10 个, P2P 节点为 1000 个, 对其进行模拟其实验结果如图 5 和图 6 所示。

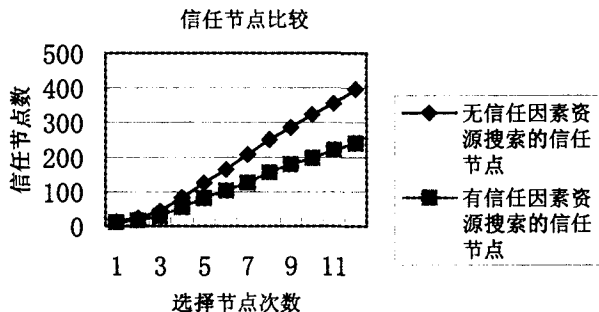


图5 两种资源搜索的信任节点比较

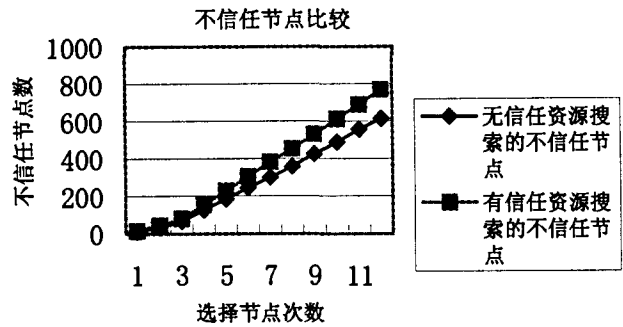


图6 两种资源搜索的不信任节点比较

图5表示了在一般的资源搜索算法和基于信任的资源搜索算法中, 分12次, 从1100个节点选取不同数目的节点, 所搜索到的诚实节点比较图。图6显示了对于同等数量的节点进行搜索时, 所搜索到恶意节点数量比较图。

由两图分析可以得出, 在引入信任因素的资源搜索机制中, 能够一定程度地发现更多的不信任、不可靠的节点的数目, 从而使得资源搜索时避免选择这些不可靠节点, 减少了恶意节点的欺诈行为, 提高了系统的可靠性。

结束语 在Grid和P2P混合计算环境中, 用户对于网络提供可靠的QoS服务具有的强烈的需求。而混合系统的动态性、异构性、不可靠性都可能导致不可预知因素的发生。本文针对传统请求转发过程中所出现的问题, 在基于经验和最好邻居资源搜索机制中加入了信任的因素, 提出了一种新的基于信任的资源搜索机制, 该算法适用于混合计算环境下的资源发现。它将传统的请求转策略同信任机制进行了有机的结合, 保证了资源发现的安全性, 减少了恶意节点的欺骗行为, 提高了混合计算环境的安全性和可靠性, 因此提供了一个良好的可靠计算环境。基于信任的搜索机制的算法实现、搜索效率分析和计算信誉值的开销都将是下一步的研究方向。

参考文献

- 1 Foster I, Kesselman C, Tuecke S. The Anatomy of the Grid; Enabling Scalable Virtual Organization. International Journal of Supercomputer Applications, 2001
- 2 Milojicic DS, Kalogeraki V, Lukose R. Peer-to-Peer Computing. HP Laboratories Palo Alto HPL-2002-57 March 8 th
- 3 Iamnitchi A, Foster I. On Fully Decentralized Resource Discovery in Grid Environments. In: Proc. of the Second Intl. Workshop on Grid Computing, 2001
- 4 Harchol-Balter M, Leighton T, Lewin D. Resource Discovery in Distributed Networks. In: 18th Annual ACM-SIGACT/SIGOPS Symposium on Principles of Distributed Computing, Atlanta, May 1999. 229~238
- 5 Gambetta D. Can We Trust? In: Trust: Making and Breaking Cooperative Relations. Gambetta D, ed. Basil Blackwell. Oxford, 1990. 213~237
- 6 Abdul-Rahman A, Hailes S. A distributed trust model. In: Proc. of the 1997 New security Paradigms Workshop, Cambrian, ACM Press, 1998
- 7 Beth T, Borchering M, Klein B. Valuation of Trust in Open Networks. In: Proc. 3rd European Symposium on Research in Computer Security-ESORICS '94