

基于经济机制的资源管理与网格经济研究进展

黄智兴¹ 刘洪涛^{1,2} 邱玉辉¹

(西南大学计算机与信息科学学院 重庆 400715)¹

(重庆教育学院计算机与现代教育技术系 重庆 400607)²

摘要 本文通过回顾网格计算的特点,阐述了基于经济机制的资源分配在网格计算中的重要性和优势,分析了建立基于经济机制的网格系统所需的必要条件,通过比较有代表性的基于经济机制的资源管理系统和工作,以及参考相关的网格协议,如网格经济服务框架(GESA)等,指出了网格经济研究和实施中面临的问题与发展的方向。

关键词 网格经济,效用计算,多智能体,拍卖,网格经济服务框架

The Research Progress of Economic-based Reserouce Management and Grid Economics

HUANG Zhi-Xing¹ LIU Hong-Tao^{1,2} QIU Yu-Hui¹

(Faculty of Computer Science and Information, Southwest China Unversity, Chongqing 400715)¹

(Department of Computer and Modern Education Technology, Chongqing Education College, Chongqing 400067)²

Abstract In this paper, by firstly reviewing the characteristics of grid computing we address the importance and advantages of economic-based mechanisms for resource management in grid computing, and then we anylize the requirements for economic-based Grid system. After comparing a set of repretative systems and works related to this field with reference to the ongoing Grid protocols such as the GESA, we finally reveal problems to be faced in research and implementation of Grid economics, and its research orientations as well.

Keywords Grid economics, Utility computing, Multiagent, Auction, GESA

1 引言

网格计算起源于美国政府资助的高性能计算科研项目。网格计算的目的是把跨地区的多台高性能计算机、大型数据库、贵重科研设备、可视化设备和各种传感器等整合成一个巨大的超级计算机系统,支持科学计算和科学研究^[4]。

无论是简单的计算系统还是复杂的集群系统、并行系统、分布式系统,都存在着不同的资源,如内存、储存空间、带宽等都是资源。与传统的资源相比,网格资源具有如下特点。

多层次上的异构性:构成计算网络的计算资源和网络连接常常是高度异构的,这种异构特性表现在各个层次上,从硬件设备、系统软件到调度策略、安全策略、使用策略等等都具有异构性。这是由计算网络所面临的现实环境决定的^[2]。

可扩充性和可选择性:网格首先必须保证系统的可扩充性。网格中广域范围内的并行与分布式计算,拥有庞大的计算机资源,可扩充性是一个网格的基本特点。任何应用都可以根据自己在连接特性、开销、安全和可靠性等方面的要求选择适合的计算资源完成计算。局部的优先级规则、策略和预算方法,在网格环境下将不再具有良好的可扩展性。

动态性与不可预测性:传统上,高性能应用一般在单独一类系统上开发,很多特性是固定、可知的。然而,网格的应用要求执行在较广泛的环境中,这种环境是由适合资源动态构成的。这就导致了执行系统的结构和特性很难确定,动态变化,无法进行预测。同时,资源在地理上的分布式特性以及网络环境的复杂性更加剧了这一点。传统的高性能应用系统往往可以预测其对处理机和网络的访问,而在计算网格中,资源

特别是网络链路很可能是共享的,其结果是系统行为和性能随时间而变化。此外,大规模的计算网格也可能会遇到资源和网络的失效,以致无法保证基本的服务质量要求,这又会导致应用系统的行为无法预测。在本地化管理的组织中,这些需求随着时间的变化保持相对稳定,需求能够被比较容易地预测和测量。与此相反,在网格环境中,需求可能是零星的和不均匀的,因为异构的和分散的用户群,对资源的需求会受到非传统的和难以预计的因素影响。

自治性和管理的多重特性:网格应用所使用的资源通常不是只被一个组织拥有和管理,而是由多个组织的管理实体管理自身的计算资源^[7,8]。网格上的资源,首先是属于某一个组织或个人的,因此网格资源的拥有者对该资源具有最高级别的管理权限,网格应该允许资源拥有者对其资源有自主的管理能力,这是网格的自治性。但是,网格资源也必须接受网格的统一管理,否则不同的资源就无法建立相互之间的联系,无法实现共享和互操作,无法作为一个整体为更多的用户提供方便的服务,因此网格的管理具有多重性。不同组织的管理实体可能采用不同的管理机制、不同的验证机制、授权机制和访问机制,这使得本已令人头疼的网络安全问题更加复杂,这也要求计算网格要解决用户代码的远程执行问题、资源的协作分配问题以及分布和并发调度问题。传统的资源分配和收费方法已经不能满足需要。

2 网格资源的管理和调度

在网格环境中,生产者和消费者有不同的目标、策略和供求模式,资源和用户在地理上分布于不同的时区。管理这

黄智兴 博士生,主要研究方向:基于经济机制的资源分配与管理;刘洪涛 博士生,主要研究方向:网格计算、多智能体;邱玉辉 教授,博士生导师,主要研究方向:人工智能。

种复杂环境,最常用的方法可以分为:以系统为中心的和以用户为中心的策略。以系统为中心的策略是一种传统的资源管理方法,它的目的是优化系统范围内的整体性能,通常是用在单个管理区域的资源管理。而以用户为中心的策略是基于用户对服务质量请求,强调实现系统用户效用的最大化。例如,可以依据像用户作业完成期限这类重要的属性来保证达到某种程度的性能。

2.1 系统为中心的策略

以系统为中心的网格资源管理系统,如 Legion, Condor, AppLeS PST, NetSolve, PUNCH, XtremWeb 等,它们采用传统的策略,调度部件由系统参数决定哪些任务在哪些资源上执行。它们的目标是增强系统的吞吐量、效用和最短完成时间,而不是提高应用程序的效用。它们不考虑资源访问的价格因素,也就是说在任何时间处理应用程序的价值都是相等的。事实上,用户更愿意根据不同任务对资源的需求、完成任务的价值、任务的优先级和执行任务的预算等因素来协商执行不同任务的价格。

2.2 用户为中心的策略

在以用户为中心的方法中,调度决策可以根据用户的需求在运行时动态决定。用户根据其使用的软硬件资源而支付服务的费用。资源的定价策略是基于用户的需求和资源的供给,是由竞争的市场模型决定的。因此,在该模型中用户与用户竞争,资源的所有者与资源的所有者竞争。

2.3 基于经济机制的策略——计算经济(Computational Economy)

经济学的方法成功地管理了当前人类经济生活中存在的分布性和异构性。经济模型增强了网格的社会结构,因此保证了它的稳定和效率。由于网格环境的动态性,不可能预先定义一个资源的价格-性能-供求关系的数据表来实现网格资源的中介。借鉴人类社会的市场调节模型,需要在网格中建立一个市场模型,通过市场机制调节供求关系和资源价格,提供一个公共的解决方案。

虽然网格经济目前没有统一的定义,但是通常我们把基于经济机制的网格资源选择、分配、控制和支付的整个生态系统(Ecosystem)称为网格经济(Grid Economics),其相关研究领域包括:电子商务、网络经济、基于市场机制的控制(Market-Based Control)、经济工程学(Economics Engineering)^[1]等。

竞争的经济模型可提供一些算法、策略和工具,来解决网格系统的资源共享与分配问题。许多经济模式包括微观经济学和宏观经济学,其原则包括:商品市场模式、价格发布模式、议价模式、招投标模式、合同网模式、拍卖模式、基于标的比例(Proportional)的资源共享模式、协作交换模式、垄断与寡头模式等。

利用网格经济有助于建立大规模的网格,资源的拥有者提供了可以通过贡献其资源为他人所用而获得利益的激励。同时,对用户提供了经济的激励,通过减小优先级来产生较少的支出,因此鼓励重要的问题优先解决,有助于规范资源的供求。

网格经济提供了一个通用的原则来比较有冲突的用户需求,这些需求和目标通过货币的形式表达;提供了所有资源统一的处理,所有的资源都可以交易,包含计算能力、内部和外存储器、网络带宽、数据以及仪器设备等;可以有助于建立高可扩展性的系统,这种决策处理是分布式涵盖所有的用户和

资源的拥有者;建立了简单但同时有效的基础,为不同的应用在不同的时间提供不同的服务。

总的来说,网格经济模式将权力交到用户和资源的拥有者手中,用户和资源拥有者为实现效用最大化或收益最大化做出他们自己的决策。此外,对资源或服务使用,以及它们能实时提供的服务质量的谈判,将需要明确的公共契约来支持和保证不同的服务质量水平。这些与静态的、以文档为主的谈判有很大的不同。

2.4 建立基于经济机制的网格资源分配策略

基于经济机制的资源管理策略需要提供允许资源的消费者和资源的提供者表达他们的需求并且方便他们进行决策的机制和工具。这些机制和工具包括^[23,6]:

1) 价值的表达(Value Expression)。包含表达他们的价值和目标的方式。

2) 价值的转换(Value Translation)。资源调度策略将用户的价值表达转换成实际的资源分配。

3) 价值的实施(Value Enforcement)。用于强制选择和分配不同服务,以及执行时动态适应服务变化的机制。

在基于经济机制的网格计算环境中,资源管理系统需要为资源消费者(用户)和资源生产者(资源所有者)提供表达他们需求,以及能够方便地实现他们的目标的工具和机制。如:资源消费者需要一个效用模式,表达他们的偏好参数和对资源的需求;资源的中间代理(Broker)需要有支持资源的发现和具备能根据分布式资源的可用性、容量,以及与用户 QoS 需求相关的费用等参数动态调度应用程序的策略。资源的提供者需要工具和机制支持价格的指定和产生模式来增加系统的利用率,以及支持服务的发布、交易、清算的协议。同时,需要提供协作机制,使供求达到均衡,保证市场的竞争和健康。

网格计算的资源管理和调度系统需要根据资源的消费者和所有者的需求管理资源和应用程序的执行,并且它们必须不断地适应可用资源的变化。

3 代表性的工作

判断市场模型有效性的标准包括:社会福利、帕累托效率、个人理性、稳定性、通信效率等。基于经济机制的资源管理机制是分布式系统中研究的热点,这里我们仅列出比较有代表性的系统。其中有的系统并没有在网格环境中实现,而有的系统目前用于对等网中。我们选择的模型标准是其系统所采用的经济模型可以直接或经过改进移植到未来的网格经济模型之中。

Spawn^[19] 是支持并行任务执行的多 Agent 系统。Agent 要求在给定预算的条件下完成计算任务,当可能的时候一个任务被分解成多个并行的子任务。Agent 通过参加由每个工作站管理的第二价格的密封,拍卖竞争工作站 CPU 时间。Spawn 的关键是 Agent 如何把资金分配到不同的子任务之中,资金的分配实际上就是控制并发计算。

D'Agent^[11] 在 D'agent 系统中,移动 Agent 在网络中到处移动并执行它们的任务。当移动 Agent 的目的结点属于其他管理组织时,移动 Agent 根据这些结点发布的价格决定是否接受或拒绝这些结点。通过限制贪心用户的请求达到系统内部的稳定,实现用户之间资源分配的均衡。

Mariposa^[12] 提出了一种分布式的经济学方法,处理分布式数据库的查询执行和存储管理。用户应用程序的每个查询都被指派了一个预算,代理将负责将用户的查询分解成较

小的子查询,并根据查询响应时间和对资源使用的均衡计算每个子查询占预算的比例。服务器通过出售存储资源或购买数据对象实现利润最大化。Mariposa 采用了招投标和合同网方式组织和管理市场,用户代理通过获取黄页、广告、优惠券等信息选择合同对象。

Popcorn^[18]、JavaMarket^[20]和 JaWS^[21] 在这些方法中,资源的购买者用 Java 语言描述他们的应用程序,并把他们的应用程序注册到市场服务器(Market Server)结点中。另外一方面,资源的出售者利用注册在市场服务器上 URL 链接以 Java Applets 的方式执行这些应用程序。市场服务器管理应用程序的上传和下载,并匹配和保留资源的购买订单和出售订单。这些系统主要是采用简单的客户服务器模式,没有考虑市场服务器的可扩展性,或自动指定用户结点作为市场服务器结点。

Tycoon^[10] 是 HP Palo Alto 实验室的分布式的基于市场机制的资源分配系统,允许用户指出希望他们的应用程序在哪些主机上运行,并给出相应的出价。资源提供方采用比例共享(Proportional Share)的方法分配源。与一般的拍卖方法相比,用户不必为等待拍卖的结果而等待很长的时间。

Mirage^[14] 是美国 Harvard 大学 Intel Research Berkeley 实验室设计的基于组合拍卖的传感器网络实验平台系统,它建立了一个封闭式的虚拟货币环境,用户以投标(bid)的方式竞争资源。投标指定用户感兴趣的资源组合,并给出其出价。系统周期性地利用组合拍卖决定拍卖的胜者,并将资源分配给拍卖的胜者。由于在封闭系统中用户没有获取资金的外部资金来源,Mirage 将系统分配资源的所得收入,以及向资金超过一定数目的用户征收固定比例的税收的收入,根据用户账户当前资金的金额按比例分配给所有的用户。

G-Commerce^[15] 是 Rich Wolski 提出了一种网格资源管理框架。在网格中包含多种资源,每种资源的价格,G-Commerce 利用经济学中的一般均衡理论,根据每个用户对不同资源的需求,以及网格资源提供商对资源的供给,计算网格市场中全部商品的均衡价格。文[5]采用集中式价格调整算法,结合集中式同步调价算法速度快、分布式 WALRAS 算法可扩展性的优点,提出一种改进的分布分组调价算法。

OCEAN^[16] (Open Computation Exchange and Arbitration Network) 提出了一种基于对等网技术的可扩展的市场模型。为了协助用户找到合适的资源,OCEAN 中提出了高效的匹配算法,以及对等网的进化算法。科研人员在 Java 和 .NET 平台上实现了 OCEAN 系统。OCEAN 中提供了用户提交任务的图形界面,并用 XML 封装用户请求。通过匹配算法返回满足条件的服务,系统可以采用自动或人工的方式对价格等参数进行协商。

GridBus^[17]、GRACE^[23]和 Nimrod-G^[23] 是澳大利亚 Monash 大学 Buyya 领导的网格经济研究项目,Nimrod-G 是基于对市场机制在分布式系统中应用的早期研究成果;GRACE 是一个基于多种经济学模型之上的、分布式的、可计算的经济学体系框架;GridBus 是基于上述框架的开发源代码的网格工具包。其中 GRACE 网格市场环境有网格资源提供者(Grid Resource Provider, GRP)和网格资源中介者(Grid Resource Broker, GRB)两个关键角色。网格计算环境提供必要的基础设施,如安全通信、信息服务、对远程资源的透明访问及使上述两个角色发生联系与交互的功能。消费者与中介者交互,委托其管理和调度网格上的计算。GRP 通过运行

软件系统建立一个网格接口,同时提供网格交易服务,使资源的交易能够顺利进行,并在 GRB 的指导下执行消费者的请求。在资源交易期间,GRB 与 GRP 的交互通过网格市场目录(Grid Market Directory, GMD)作为中间媒介,它们使用各种经济模型和交互协议来确定资源使用价格。GridBus^[17]中除了上面提到的 GMD、GRB 外,还包含网格银行 GridBank、网格监控器 G-Monitor、网格模拟器 GridSim 等组件。

PeerMart^[24] 为了克服在 P2P 中资源共享时出现的搭便车现象(Free Riding),PeerMart 采用分布式的拍卖方式,实现信息的发布与资源的共享,其关键技术是利用分布式哈希表 DHT 实现价格的发现。

其他 为了克服网络资源不可存储的问题,IBM Zurich 研究中心的 Giorgos Cheliotis 和 Chris Kenyon 提出了利用期货价格(Future Price)来进行资源的预留等管理^[1]。M. Tamai^[26]研究了如何将分散的资源聚集起来进行交易的方法,同时研究了如何实现市场服务结点负载均衡的问题。P. Golle^[25]等提出了利用激励机制解决 P2P 环境中的搭便车的问题。

从上述系统我们可以发现拍卖机制(包括拍卖的各种形式:英式拍卖、荷兰式拍卖、双向拍卖、维克瑞拍卖、组合拍卖等)、一般均衡原理、激励兼容机制、期货理论是基于经济机制的资源管理系统中最常用的经济学理论。上述系统往往只采用单个经济模型,而单个模型往往只能解决特定环境下的资源分配和管理问题。可以预期的是,未来的网格经济是建立在多个经济模型共同存在的基础之上。从这个角度上看,上述系统中,GridBus 是目前网格经济研究项目中实现较为全面的系统。当然,研究人员仍需探索适合于网格环境的新的经济机制。

3.1 支持网络经济的协议

GESA 即网格经济服务框架(Grid Economic Services Architecture),是 Global Grid Forum(GGF)^[22]的一个工作组计划,目标是在 GGF 所提出的开放网格服务架构 OGSA 之上制定关于网格经济学服务架构的一些标准和规范。GESA 中包含可交易网格服务(Chargeable Grid Services, CGS)、网格银行服务(Grid Banking Services, GBS)、GBS 持有服务(GBSHold Service)3 大部分。

可交易网格服务 CGS 包括服务数据元素(Service Data Elements, SDE)、网格经济学服务接口(Grid Economic Services Interface, GESI),SDE 中描述了服务的定价机制、价格、使用方式、货币、证书等内容。GESA 目前所支持的定价机制只包括固定价格和(英式)拍卖方式,但是可以扩展其他机制;而价格根据所消费的服务资源、其使用方式以及该资源的定价机制来产生。GESA 中还指明了当服务方不能提供服务时用户可获得的补偿(Composition),或由于用户的原因任务不能成功执行时,服务方归还(Refund)用户部分费用的接口,这两个接口往往可以用百分比的形式描述。GESI 中定义了价格获取、价格参数、价格接受等接口,用户通过该接口服务最终的价格可能会经过多阶段的拍卖或协商得到。用户接受价格后,服务双方将会签署一个由 SDE 构成的、由 XML 格式描述的合同。GESI 中没有描述资源协同分配时的情况。

网格银行服务 GBS 中定义了一系列的与支付相关的接口,包括建立账户、删除账户、账户核对(creditCheck)、转账(transferIn, transferOut)和交易查询(getLastTransaction, getTransactionbyDate)等功能。当然,网格银行同样存在认

证等安全方面的考虑,主要通过网络安全基础结构 GSI 来实现。

在全球网络论坛 GGF 下与网络经济服务结构工作组相关的其它工作组包括^[21]:资源使用服务(RUS)、网络资源分配协议(GRAAP)、开放式网络服务结构(OGSA)、使用记录(UR)、预留(AR)、网络安全基础结构(GSI)等工作组。

4 网络经济研究中面临的问题及发展动向

网络计算走向商业化时,应当从经济领域吸取一些经验和教训^[1],并将它们集成到系统的开发之中。我们应该注意以下方面。

避免“公共地悲剧”现象 由于目前仍然缺乏可扩展的和动态的标准机制来限制系统使用,因此像网格一样基于共享的体系结构尤其容易导致负的外部性(Negative Externalities),即个人的理性活动对整个群体造成了负面影响。如何预防这种“公共地悲剧”现象的发生,是网格开发中值得认真考虑的问题。

发现网格资源的动态价值 资源的价值取决于用户对资源的需求和资源的稀缺性两方面的综合。与其他传统的资源(如汽油、电)一样,网格资源的价值也是动态变化的。发现资源的动态价值,将用户的需求映射为价格,拍卖是一种切实可行的、动态的价格形成机制。资源的动态价值应当从某个层次上(用户不必了解价格产生的细节)及时地传递给用户,而固定的价格机制只在某些有限的范围内使用。

采用真正的货币 在当前基于经济机制设计的系统中,一般是采用虚拟的网格货币(Grid Currency)。由于网格的构性和可扩展性特点,使得我们有必要建立标准的货币单位和汇率机制。但是虚拟货币不可避免存在信任、可兑换性、管理困难等问题。IT的管理者应当尽快选择真正的货币,跨过使用虚拟货币的中间阶段。采用真实的货币使得共享资源的购买、建立、出租、升级换代等决策变得简单,可以实现预算直接有效地分配。

实现资源预留服务 为了避免价格的剧烈波动以及能够实现风险管理,传统的不可存储的商品都建立了相应的期货市场,进行商品的预定。类似地,计算资源一般是不可存储的,也就是说今天的计算能力不能留到将来使用。切实可行的网格市场应该以预留服务为中心。值得关注 GGF 已经制定了关于预留服务的协议,为网格资源的商业化走出了非常重要的一步。

保证服务质量 目前大多数 IT 系统不能严格地保证服务质量(QoS),而是提供尽最大努力的(Best-effort)服务。这种不能保证服务质量的服务将影响用户对系统的评价。计算和应用程序虽然很复杂,但是它们都有明确的定义。高度详细的合同规范和严格地保证这些规范是树立用户使用系统信心的必要元素,它们可以使这些高度分布的、跨多个组织的用户得到他们所期望得到的。在某些时候,网格中可交易的资源必须用统计学的术语来描述。

建立激励机制 从大量不成功的 B2B 和 B2C 例子我们可以知道,切合实际的激励机制对于推动资源的共享是非常重要的,激励机制应当是网格系统设计中的基本组成部分。软件和系统可以支持激励机制,而激励应该在更高商业层上或策略层上设计。

包含信任机制 在商业环境中,信任是基于法律上和经济上的惩罚,不产生相应的后果的系统是没有办法产生信任的。在 IT 资源的共享环境中应当具备这样的功能:当出现用户违背信任现象时,会导致用户在时间上和数量上承担相应

比例的后果。

建立支持处理的工具 资源分配的处理工具必须使用户能够以一种简单的方式表达他们对资源的需求,同时包括伴随时间流逝对这些需求的不确定性。它们可以使用户获取现货市场和期货市场资源有效的价格动态,可以对资源进行交易,而用户不必关注实现的细节。要建立集成了用户预算和商业目标同时包含资源分配决策支持的工具看上去太困难,但例如多阶段随机优化工具^[1]已经成功应用于日常的金融管理系统之中,这些工具同样可以应用于网格这样的大规模资源系统。

与多智能体的结合 正如文^[13]指出的那样,网格与多 Agent 之间的关系,就像肌肉和大脑的关系一样。由于网格资源管理系统是一个天然的多 Agent 系统,因此基于经济学原理的多 Agent 研究的成果,如多 Agent 的竞争、协作、协商、机器学习等理论也是建立基于经济机制的网格系统系统时需考虑的重要因素。同时,基于多 Agent 的模拟是研究网格经济的一个重要手段。

结束语 本文介绍了网格经济的概念、优势、研究和应用现状、支持网格经济的协议以及还面临的问题。在下一代网格中,经济因素将成为一个关键的设计参数,现在许多国家和科研机构已经就这一点达成共识,例如欧洲的下一代网格 2005~2010 计划等。能否成功运用基于经济机制的资源管理方式对网格计算的环境进行管理,将在很大程度上影响网格计算的进一步推广与发展。

参 考 文 献

- 1 Kenyon C, Cheliotis G. Grid Resource Commercialization: Economic Engineering and Delivery Scenarios. *Grid Resource Management: State of the Art and Research Issues*, Kluwer, 2004
- 2 Kenyon C, Cheliotis G. Architecture Requirements for Commercializing Grid Resources. In: 11th IEEE International Symposium on High Performance and Distributed Computing (HPDC-11), Edinburgh, July 2002
- 3 Kenyon C, Cheliotis G, Buyya R. 10 Lessons from Finance for Commercial Sharing of IT Resources. *Peer-to-Peer Computing: The Evolution of a Disruptive Technology*. IRM Press, 2004
- 4 徐志伟,冯白明,李伟. 网格计算技术. 北京:电子工业出版社, 2004
- 5 翁楚良,陆鑫达. 一种基于市场机制的网格资源调价算法. *计算研究与发展*, 2004, 41(7)
- 6 Bredin J, Kotz D, Rus D. Market-based Resource Control for Mobile Agents. In: Proc. of the Second Int'l Conf. on Autonomous Agents, 1998, 197~204
- 7 Buttyan L, Hubaux J. Stimulating Cooperation in Self-Organizing Mobile Ad Hoc Networks. *ACM Mobile Networks and Applications*, 2003, 8(5)
- 8 Chen K, Nahrstedt K. iPass: an Incentive Compatible Auction Scheme to Enable Packet Forwarding Service in MANET. In: Proc. of the 24th Int'l Conf on Distributed Computing Systems, 2004
- 9 Ratnasamy S, Francis P, Handley M, et al. A Scalable Content-Addressable Network. In: Proc. of ACM SIGCOMM 2001, 2001, 161~172
- 10 Lai K, Rasmusson L, Adar E, et al. Tycoon: an Implementation of a Distributed Market-based Resource Allocation System, 2004
- 11 Bredin J, Kotz D, Rus D. Utility Driven Mobile-Agent Scheduling; [Tech Rep. CS-TR98-331]. Dartmouth College, Hanover, NH, 1998
- 12 Stonebraker M, Devine R, et al. An economic paradigm for query processing and data migration in mariposa. In: 3rd Int Conf. Parallel and Distributed Information Systems, Austin, TX, 1994
- 13 Foster I, Jennings N R, Kesselman C. Brain Meets Brawn: Why Grid and Agents Need Each Other. *aamas04*, 2004
- 14 Chun B N, Buonadonna P, et al. Mirage: A Microeconomic Resource Allocation System for SensorNet Testbed. In: 2nd IEEE Workshop on Embedded Networked Sensors, 2005
- 15 Wolski R, Plank J, Brevik J, et al. Analyzing marketbased resource allocation strategies for the computational grid. *Int J High-Performance Computing Applcat*, 2001, 15(3)
- 16 Padala P, Harrison C. OCEAN: The Open Computation Exchange and Arbitration Network, A Market Approach to Meta-computing. In: Proceedings of the ISPEC'03, 2003
- 17 Buyya R, Venugopal S. The gridbus toolkit for service oriented grid and utility computing; An overview and status report. *GECON 2004*, Seoul, Korea, Apr. 2004

18 Nisan N, London S, et al. Globally distributed computation over the internet: The POPCORN project. Int Conf Distributed Computing Systems, Netherlands, May 1998
 19 Waldspurger C, Hogg T, et al. Spawn: A distributed computational economy. IEEE Trans Softw Eng, 1992, 18(2)
 20 Amir Y, Awerbuch B, Borgstrom R S. A cost-benefit framework for online management of a meta computing system. In: 1st Int Conf. Information and Computational Economy, Charleston, SC, Oct. 1998
 21 Lalis S, Karipidis A. An open marketbased framework for distributed computing over the internet. In: 1st IEEE/ACM Int Workshop Grid Computing, Bangalore, India, Dec. 2000
 22 Global Grid Forum(GGF). <http://www.ggf.org>

23 Buyya R, Abramson D, Venugopal S. The Grid Economy, Special Issue on Grid Computing, 2005, 93(3)698~714
 24 Hausheer D, Stiller B. Decentralized Auction-based Pricing with PeerMart. In: The 9th IFIP/IEEE International Symposium on Integrated Network Management (IM 2005), Nice, France, May 2005
 25 Golle P, Leyton-Brown K, Mironov I, et al. Incentives for Sharing in Peer-to-Peer Networks. In: Proc. of the 2001 ACM Conference on Electronic Commerce, 2001
 26 Tamai M, Shibata N, Yasumoto K, et al. Distributed Market Broker Architecture for Resource Aggregation in Grid Computing Environments. In: Proc. of Cluster Computing and Grid 2005 (CC-Grid2005)

(上接第 51 页)

设系统中文件 f_i 的原始文件和副本总和为 T_i , 则 $\alpha 2^{n_i} \leq T_i \leq \alpha 2^{n_i} + \epsilon$, 故对于拥有文件 f_i (可能是 f_i 的原始文件, 也可能是 f_i 的副本) 的节点而言, 其平均收到的查询数为:

$$\bar{q}_i = \frac{\epsilon}{T_i} * q_{a_i} + (1 - \frac{\epsilon}{T_i}) * q_{r_i} \quad (5)$$

在系统中, 假设查询流的到来符合泊松分布, 则由 M/D/1/ ∞ 排队系统^[10]可知, 系统中节点对文件 f_i 的查询的平均响应时间为:

$$\bar{W}_i = (2 - \frac{\bar{q}_i}{\mu}) / 2\mu(1 - \frac{\bar{q}_i}{\mu}) \quad (6)$$

对所有文件查询的平均响应时间为:

$$\bar{W} = (\sum_{i=1}^M i * \bar{W}_i) / \sum_{i=1}^M i \quad (7)$$

文[6]的模型可以看成上述模型的特例, 令 $N=M, \alpha=1, \epsilon=1$ 就可以得到文[6]的模型。如果将 α 看成节点的在线率, 则本文提出的模型更具有一般性。

4 DynRM 模型分析

首先将文[6]中的模型(NO-DynRM)和本文的模型(DynRM)进行了对比, 如图 2 所示。从图中可以看出, 随着阈值 r 的增加, 两个模型的平均响应时间都随之而增加。当阈值 $r=0$ 时, NO-DynRM 模型的平均响应时间为 1.517641, 而 DynRM 的响应时间为 1.170684, 这个时候, 系统中的副本数量达到最大数, 故此时系统响应时间较快。由于 DynRM 系统中对受欢迎的副本索引缩短位数, 然后取反, 使得受欢迎的副本不但在本区域扩散, 而且扩散到地域较远的其它区域, 故 DynRM 模型的系统响应时间较 NO-DynRM 模型的系统响应时间快。当阈值 r 增加到一定程度时, 即 $r=0.5$ 时, 两个模型的平均响应

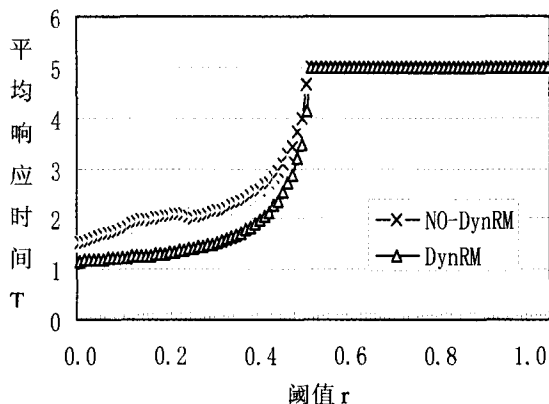


图 2 DynRM 和 NO-DynRM 对比图

时间接近 5, 为 4.975593。当阈值 r 大于 0.5 以后两个模型的系统平均响应时间处于稳定状态, 实际上此时两个模型中的文件都没有增加副本, 故响应时间没有变化。

结束语 在 P2P 分布式存储系统中, 副本能有效提高文件的可靠性, 降低访问延迟, 避免 Hot Spots 的产生和平衡负载。本文在分析了现有的分布式存储系统动态副本管理机制的基础上, 提出了一种动态副本管理机制——DynRM。在 DynRM 中, 当文件的受欢迎程度达到一定的阈值时, 系统就增加文件的副本, 在放置增加的副本时, 选择高性能的节点放置副本, 以提供更好的系统性能。通过增加文件副本, 一方面提高了文件的可靠性, 降低了访问延迟; 另一方面避免了系统热点的产生, 到达了负载平衡的目的。

参考文献

1 Ranganathan K, Iamnitchi A, Foster I. Improving Data Availability through Dynamic Model-Driven Replication in Large Peer-to-Peer Communities. In: The 2nd IEEE/ACM International Symposium on Cluster Computing and the Grid (CCGRID'02) Berlin, Germany, May 2002
 2 Cuenca-Acuna F M, Martin R P, Nguyen T D. Autonomous Replication for High Availability in Unstructured P2P Systems[C]. In: the 22nd IEEE Inte Symposium on Reliable Distributed Systems, 2003
 3 Gnutella: To the Bandwidth Barrier and Beyond. <http://lambda.cs.yale.edu/cs425/doc/gnutella.html>. 2001
 4 Backx P, Wauters T, Dhoedt B, et al. A comparison of peer-to-peer architectures. In Germany. <http://allserv.rug.ac.be/~pbackx/architectures.pdf>. 2002
 5 Clarke I, et al. Freenet: A Distributed Anonymous Information Storage and Retrieval System. In: ICSI Workshop on Design Issues in Anonymity and Unobservability, July 2000
 6 Cudre-Mauroux P, Aberer K. A Decentralized Architecture for Adaptive Media Dissemination
 7 Saroiu S, Gummadi P N, Gribble S D. A measurement study of peer-to-peer file sharing systems [C]. In: Proc. of Multimedia Computing and Networking (MMCN), Jan. 2002
 8 Xu Z, Hu Y. SBARC: A Supernode Based Peerto-Peer File Sharing System. in(ISCC), Kemer-Antalya, Turkey, June 2003
 9 侯孟书. 基于 P2P 的分布式存储研究; [博士论文]. 电子科技大学, 2005
 10 唐应辉, 唐小我著. 排队论—基础与应用. 电子科技大学出版社, 2000