

基于虚拟组织的网格文件资源共享模型

吴雄奇, 曾文华

(厦门大学软件学院, 福建 厦门 361005)

摘要: 网格计算主要关注大规模的资源共享, 且这种共享是高度可控的。为解决网格环境下文件资源共享与管理的问题, 提出了一个网格文件资源共享模型 FsvGrid。该模型引入注册通知机制, 并采用确定性算法与非确定性算法相结合的消息传递机制, 使得网格中的各个节点之间能够高效协作; 采用分层结构, 屏蔽了文件资源的多样性; 增加了共享的安全性, 可以对共享进行控制; 提出了一种依靠虚拟组织来对文件资源进行管理的方式, 解决分布式资源难以管理的问题。

关键词: 网格计算; 资源共享; FsvGrid; 虚拟组织

中图分类号: TP393

文献标识码: A

文章编号: 1673-629X(2007)03-0001-04

Grid File Sharing Model Based on Virtual Organizations

WU Xiong-qi, ZENG Wen-hua

(Software School of Xiamen University, Xiamen 361005, China)

Abstract: Grid computing focuses on large-scale resource sharing, and this sharing is highly controlled. In order to solve the resource sharing and managing problem, a grid file sharing model is proposed in this article. This model has the following features: having register-notification mechanism and high performance message passing mechanism with both fix and randomized algorithms to ensure the nodes in the grid can well cooperate with each other; using layered structure to hide the heterogeneous resources; improving security and being able to control the sharing; proposing a file managing mechanism based on virtual organizations so as to manage distributed resource.

Key words: grid computing; resource sharing; FsvGrid; virtual organization

0 引言

网格计算相比于传统的分布式计算, 更关注大规模的资源共享, 而且这种共享是高度可控的。在目前的因特网环境下, 存在许多文件资源共享的方式, 主要可以分为单一的服务器端与客户端的共享方式和分布式的共享方式两类。

以传统 HTTP 服务器或 FTP 服务器等方式来提供文件资源下载, 从而对文件资源共享, 即为典型单一的服务器端与客户端的资源共享方式。这种共享方式缺点是随着服务的客户数目增多, 服务器端所需要占用的资源也随之增加, 最终制约了客户数目的增长。

P2P 资源共享方式是一种新型的分布式资源共享方式。在这种方式下, 资源共享者能够同时扮演客户端和服务器端两种角色, 这个方案最小化工作负荷而

最大化了全部网络的性能。现在比较常用基于 P2P 的共享方式有 Bittorrent^[1] 和 Emule^[2]。不过目前 P2P 的资源共享方式也有不足的地方: 比如难以对共享进行控制; 缺乏有效的对用户合法性进行验证的手段; 不同 P2P 的实现方式之间接口不统一; 没有对资源进行统一有效的管理, 资源生命周期很短等。

关于网格体系结构的研究国外比较著名的有 GLOBUS^[3], 国内的有织女星网格^[4], 分别应用在不同的领域。为了解决上文所述的两类文件资源共享方式的不足, 文中提出的一个网格体系结构: 基于虚拟组织^[5]的网格文件资源共享模型 FsvGrid。

1 FsvGrid 模型介绍

FsvGrid 是基于虚拟组织的文件共享模型, 模型目的是建立一个网格环境下文件共享的平台。如图 1 所示, FsvGrid 模型采用分层结构, 它由资源索引服务器层、虚拟组织索引服务器层、虚拟组织协调服务器层以及 Peer 层四个部分组成。

FsvGrid 既可以利用互联网上各种现存的文件共享方式, 也可以定制新型的共享方式来满足网格环境

收稿日期: 2006-06-05

基金项目: “985 工程”智能化国防信息安全技术科技创新平台项目 (0000-X07204)

作者简介: 吴雄奇 (1981-), 男, 福建福州人, 硕士研究生, 研究方向为网格计算等; 曾文华, 博士, 教授, 研究方向为人工智能、神经网络、进化计算、网格计算、智能控制等。

下新的需求,通过虚拟组织协调服务器之间以及虚拟组织协调服务器层与 Peer 层之间的交互,能够为用户找到尽可能多的共享资源,而且该模型屏蔽了文件资源异构性。

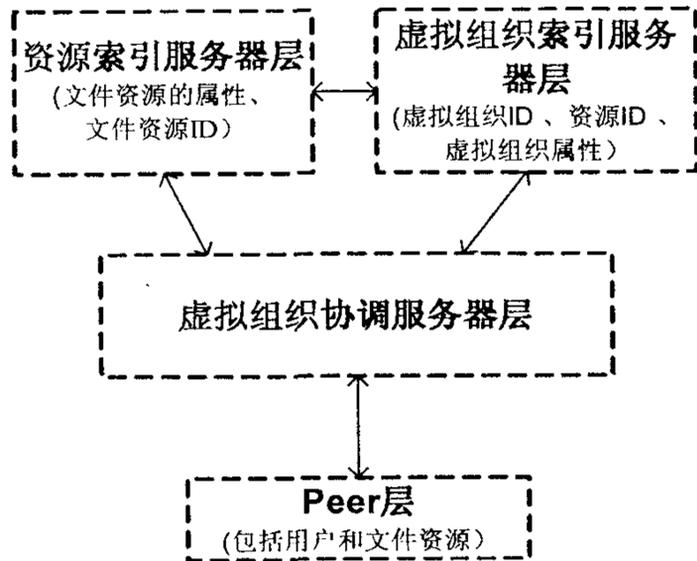


图 1 FsvGrid 模型

不同于传统的基于异步自治的访问控制策略^[6]以及所有者为中心的网格文件共享方式^[7],FsvGrid 采用安全虚拟组织方式,引入公钥、私钥增加共享安全性,使得安全模型层次清晰而且更加有效。P2P 的资源发现机制是目前比较高效的一种资源发现机制^[8]。Fsv-Grid 在 P2P 机制基础上,依靠虚拟组织来对资源进行发现,并且可以对资源进行管理,解决了分布式资源难以管理的问题。

2 FsvGrid 模型的体系结构

在 FsvGrid 模型中,虚拟组织协调服务器层是模型的核心部分,它从资源索引服务器层获得文件资源的属性,从虚拟组织索引服务器层得到虚拟组织的属性,并按照规则建立和维护虚拟组织。而 Peer 层中的 Peer 通过虚拟组织协调服务器层进入到虚拟组织,然后 Peer 可以从虚拟组织中或者从其它的 Peer 得到文件资源的信息,从而进行文件资源共享。

2.1 资源索引服务器层

资源索引服务器层由资源索引服务器组成。资源索引服务器存储了文件资源的属性和文件资源 ID 之间的对应关系,可以根据文件属性(包括文件名称、文件大小、文件描述等)从资源索引服务器中获得文件资源 ID。网格中每一个文件资源,都用一个唯一的文件资源 ID 来标识。文件资源 ID 作为一个文件的标识用在文件共享的虚拟组织以及文件资源维护的虚拟组织中。

2.2 虚拟组织索引服务器层

虚拟组织索引服务器层同样是由虚拟组织索引服务器构成。虚拟组织索引服务器是用来存储虚拟组织 ID、资源 ID、虚拟组织属性(包括虚拟组织类型和负责

该虚拟组织的虚拟组织协调服务器)之间的对应关系。虚拟组织索引服务器的作用主要是让虚拟组织协调服务器能够根据虚拟组织 ID 得到相应的资源 ID 以及虚拟组织属性,反之也可以根据资源 ID 以及虚拟组织属性来得到相应的虚拟组织 ID。

2.3 虚拟组织协调服务器层

虚拟组织协调服务器层是 FsvGrid 模型的核心部分。虚拟组织协调服务器分布存在于网格系统中。

虚拟组织协调服务器层的功能主要是通过制定规则维护一个虚拟组织正常运行,包括虚拟组织建立与撤销、虚拟组织成员的加入与离开、成员状态的更新、文件资源发现与维护、安全权限控制,并负责虚拟组织内部消息传递以及与其它虚拟组织的通信。

FsvGrid 模型在虚拟组织协调服务器层引入了注册通知机制以及邀请机制,以提高模型的扩展性。并采用确定性和非确定性的两种消息传递机制。

2.3.1 建立虚拟组织

虚拟组织是通过虚拟组织协调服务器来建立的。新用户通过向虚拟组织协调服务器提出申请来加入某个虚拟组织,虚拟组织协调服务器提供用户该虚拟组织的信息。用户加入到虚拟组织以后,可以通过和虚拟组织协调服务器以及虚拟组织中其它用户之间的交互,来更新虚拟组织的信息。

2.3.2 注册通知机制

引入注册通知机制的目的是让能满足用户请求的虚拟组织协调服务器或选择一个较为空闲的虚拟组织协调服务器来响应用户的请求。

如图 2 所示,用户 A 向虚拟组织协调服务器 B 提出请求,B 本身不能满足用户 A 的要求或 B 正处于繁忙状态,B 就向虚拟组织协调服务器 C 注册用户 A 的请求,由虚拟组织协调服务器 C 来响应用户 A 的要求,这就是所谓的注册通知机制。

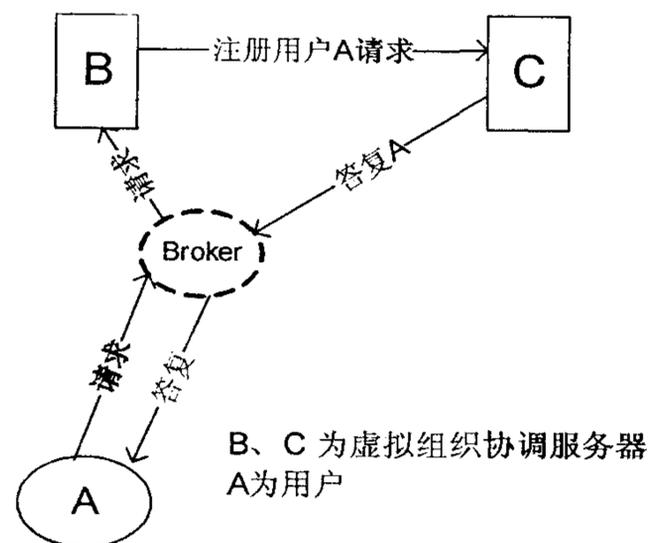


图 2 注册通知机制

通过注册通知机制,使虚拟组织协调服务器之间能够更好进行分工,并有助于提高 FsvGrid 模型的负

载平衡。

2.3.3 文件资源的发现

文件资源的发现是通过建立一个资源发现虚拟组织,文件资源服务器可以向资源发现虚拟组织注册自己,用户也可以向资源发现虚拟组织来注册已知的文件资源。

对一个虚拟组织来说,每个虚拟组织协调服务器存储该虚拟组织的一部分信息,它们之间通过交互来丰富自己的信息。这样加强了 FsvGrid 模型的存储能力,通过提供统一的对外接口,在外界用户看来, FsvGrid 是一个存储了大量信息的系统。

2.3.4 FsvGrid 模型安全性

除了普通的虚拟组织构成形式外,为了满足私有资源共享的目的, FsvGrid 提供了一种安全虚拟组织的形式。

在 FsvGrid 模型中,一个安全的虚拟组织存在授权虚拟组织协调服务器、认证虚拟组织协调服务器、认证用户等不同角色。

每个安全的虚拟组织,都有一个或多个认证虚拟组织协调服务器来进行维护。对于一个未经过认证的虚拟组织协调服务器,只有通过授权虚拟组织协调服务器的认证,才能参与到安全虚拟组织的管理中来。而用户只有经过认证虚拟组织协调服务器的授权,成为认证用户后才能参与共享安全虚拟组织的资源。

FsvGrid 模型采用公钥、私钥的方式来对安全虚拟组织中的角色进行授权。

在一个安全的虚拟组织中,存在用以授权的公钥 A、私钥 A 和用以接入虚拟组织的公钥 B、私钥 B。其中公钥 A、公钥 B 是共享的,而授权虚拟组织拥有私钥 A 和私钥 B,认证虚拟组织仅拥有私钥 B。也就是说用私钥 A 来对虚拟组织进行认证,用私钥 B 来对用户进行认证。

通过该安全模型,可以判断用户和虚拟组织协调服务器是否经过认证,这样就可以在虚拟组织中对资源下载进行控制,对虚拟组织成员的行为进行控制等。

2.3.5 邀请机制

邀请机制是指一个虚拟组织协调服务器可以邀请另一个虚拟组织协调服务器来提供某种服务或参与某个组织中。邀请机制使虚拟组织协调服务器具有一定被动性来参加虚拟组织的管理,加强了 FsvGrid 模型的扩展能力,并通过定义良好的规则在很大程度上减少人工的参与。

2.3.6 消息传递机制

为了保证虚拟组织协调服务器之间消息传递的准确性和有效性,采用了两种不同的消息传递算法:确定

性算法和非确定性算法。

①确定性算法。

为每个虚拟组织协调服务器节点,确定它的父节点和子节点。一个节点的父节点最多只能有一个,子节点可以有多个。父节点和子节点并不是经过确定后就不变了,它会根据网格中节点的情况来改变。

确立完父节点和子节点后,在消息传递过程中,每个节点在接到消息后,检查该消息是否以前接到过,如果以前接到过则不做任何处理;如果没有接收过,则对消息进行处理,并把消息转发到它的父节点和子节点。

如果父节点和子节点定义不好的话,会存在这样一个问题:仅仅由父节点和子节点的关系形成的图是不连通的。

如图 3 所示, A, B, C, D, E 为虚拟组织协调服务器节点,箭头表示的是父节点与子节点的关系,箭头从父节点指向子节点。图 3 所示的情况,在 A, B, C 之间传递的消息是不会传递到 D, E 之间的。这样就要求重新构建父子节点的关系。由于网格系统往往有许多节点,而且各个节点的状态随时都有可能改变,所以采用传统的检测图连通算法,所花费的代价太高,而且正确性因为网格节点变化也得不到保证。

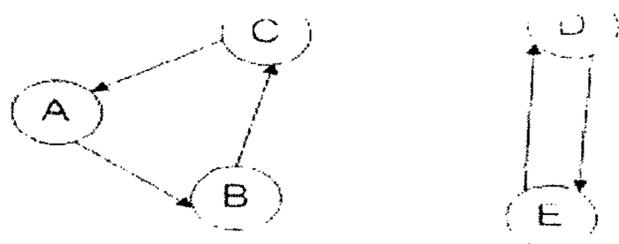


图3 节点关系图

这里提出一种检测算法,因为节点 A, B, C 和 D, E 在实际的网络中是连通的,这样节点 D 如果在一定时间内没有收到虚拟组织的消息,或者在经过一定的时间间隔后,就主动向其它非父节点,比如说 C 节点,来获取虚拟组织的消息,若发现有些消息以前并没有收到,则考虑重构父子关系, D 就向 E 要求撤销 E, D 之间的父子关系,从而建立 C, D 之间的父子关系。

②非确定性算法。

考虑到确定算法在进行消息传递时是通过父子关系向外传播的,这样对某些与消息发出节点之间父子关系比较远的节点来说,可能消息到达的时间比较长。为解决该问题,提出了非确定性算法。非确定性算法就是在节点向外传递消息时,除了传给父节点和子节点,还随机选择一些节点来进行消息传递,消息传递方式从简单单一父子关系的发散,变成随机的多点发散。这样一方面使得消息传到每个节点所需的时间更加平均,另一方面增加了消息传递的可靠性,因为在这种情况下,节点除了依赖于子节点和父节点给自己传递消息

外,还可能随机从其它节点得到消息。

2.4 Peer 层

Peer 层包括用户和文件资源。用户和文件资源都作为一个 Peer 存在,它们是对等的实体,从而丰富了 Peer 的内容。任何文件资源,不论采取何种提供方式都作为 Peer 存在。Peer 有不同的种类和属性,这样在与 Peer 进行交互时,就可以根据不同的种类和属性,采取相应的方法。

在 Peer 层当中,用户可以既是资源的下载者,也是资源的提供者。资源的类型可以多种多样,可以是 FTP 服务器、HTTP 服务器,或者是其它 P2P 网络中一个用户,区别只是它们之间提供的下载上传方法不一致,可以通过资源的属性来确定资源交互所用的方法。

3 在 FsvGrid 模型中建立虚拟组织

FsvGrid 模型提供了两种类型的虚拟组织定义,即共享文件资源的虚拟组织和维护文件资源的虚拟组织。

3.1 共享文件资源的虚拟组织

建立共享文件资源的虚拟组织目的是为用户提供一个共享文件资源的网格平台,来对资源进行高效方便的共享,并且尽可能最大化利用网络资源来提高下载速度。

如图 4 所示,用户 A 参与某个共享文件资源的虚拟组织。用户 A 从虚拟组织协调服务器得到资源信息,然后可以同时从不同的资源提供者下载需要的资源。用户下载时根据不同的资源提供者属性,采用相应的下载方式。

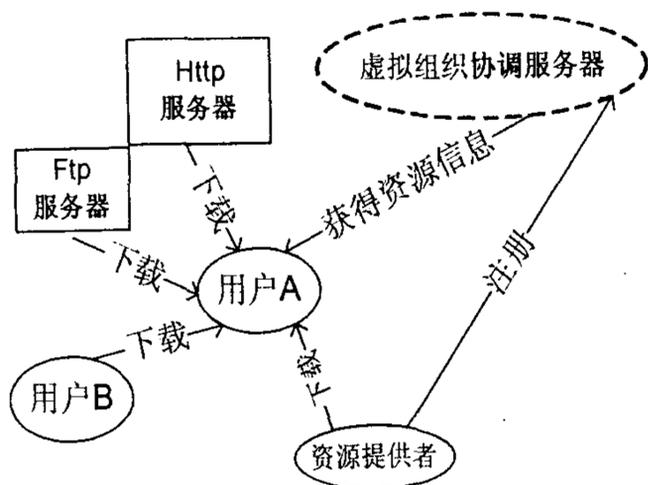


图 4 用户下载方式

3.2 维护文件资源的虚拟组织

FsvGrid 通过建立维护文件资源的虚拟组织,来管理网格中的文件资源,为网格环境下资源管理提出一个可行的解决方案。

在维护文件资源的虚拟组织中,不同的文件资源,可以由不同的虚拟组织来维护。一个用户可以同时是

资源的提供者和资源的维护者。

在文件资源众多的网格系统中,要对文件资源进行维护是十分困难的,不仅是因为文件种类的繁多,而且因为不同人维护文件资源的方式是不同的,增加了网格系统的复杂性。建立维护文件资源的虚拟组织的好处是可以提供一个统一的接口,使得对资源的管理更加合理高效。而且可以充分利用网格计算的能力,由网格中不同的节点来共同对资源进行维护,满足了大量文件资源分布式存在的客观事实。

3.3 FsvGrid 模型的扩展性

FsvGrid 模型关注的是如何对虚拟组织进行组织和管理,其中用到的各种机制,都是为了保证虚拟组织能够高效稳定地运行,这样就可以通过定制不同的规则来建立特定的虚拟组织以满足特殊的需求。

例如要建立一个即时通讯平台,可以用一个特定的文件资源 ID 来代表即时通讯,通过建立共享该文件资源的虚拟组织,使得用户加入虚拟组织就相当于登陆即时通讯平台,用户可以从虚拟组织中得到其它用户的信息,这样用户之间就可以进行消息通讯。

4 结束语

文中提出了网格文件资源共享模型 FsvGrid,它采取分层结构,通过虚拟组织协调服务器层屏蔽了资源多样性,通过虚拟组织协调服务器之间共同协作来对文件共享进行组织。在模型中采用注册通知机制以及确定和非确定的消息传递机制,保证了模型的高效协作性。并提出了在 FsvGrid 模型中建立共享文件资源的虚拟组织和维护文件资源的虚拟组织,从而搭建网格环境下的文件资源管理与共享的平台。FsvGrid 模型还具有良好的扩展性,这样在网格环境下,用户对文件资源的获取以及企业的软件产品发布都可以通过 FsvGrid 的平台来进行。

参考文献:

- [1] Bittorrent Inc. Bittorrent project [EB/OL]. 2001. <http://www.bittorrent.com>.
- [2] Emule Inc. Emule project [EB/OL]. 2002. <http://www.emule-project.net>.
- [3] The Globus Alliance. GLOBUS Project [EB/OL]. 1998. <http://www.globus.org/>.
- [4] 徐志伟,李晓林,游赣梅. 织女星信息网格的体系结构研究 [J]. 计算机研究与发展, 2002, 39(8): 948 - 951.
- [5] Foster I, Kesselman C, Tuecke S. The Anatomy of the Grid: Enabling Scalable Virtual Organizations [J]. International Journal of Supercomputer Applications, 2001, 15(3): 1 - 10.

(下转第 9 页)

方法传输 8M 大小的文件所花费的时间是只用质量好的路径的 4 倍!

当使用多路径同时传输时,这种由于路径质量差异而引起的接收端缓冲区阻塞问题肯定会出现,除非缓冲区不再由多个路径共享。但是研究表明^[5],重传时尽量选择质量好的路径可以有效降低接收端缓冲区阻塞问题。

3.6 重传时的路径选择技术

由于一个关联中可以存在多个路径,SCTP 的发送端在发送重传数据时就会面临多个选择。因为当前 SCTP 只允许数据从主路径发送,所以主路径上的拥塞控制信息(可用带宽、丢失率、RTT 等)比备选路径上的更加准确。正是基于这样的原因,通过原来的主路径发送重传数据将比使用备选路径更具性能优势^[6]。这种重传机制称为 RTX-SAME,即重传的数据始终通过原先的路径来发送。

但是在多路径同时传输时,由于同时使用了所有可用路径来传送数据,所以每条路径上的拥塞控制信息都能够比较准确地反映当前路径状态,上面的结论^[6]不再适用,研究表明^[7]RTX-SAME 在多路径同时传输时,效果最差。

在文献[7]中,作者研究了 5 种重传策略:

1) RTX-SAME:使用报文原先发送的路径,除非路径失效。

2) RTX-ASAP:随机选择一条可用路径。

3) RTX-CWND:使用具有最大 cwnd 值的路径。

4) RTX-SSTHRESH:使用具有最大 ssthresh 值的路径。

5) RTX-LOSSTATE:使用具有最低报文丢失率的路径。

在这 5 种重传策略中,RTX-SAME 最简单;RTX-ASAP(as soon as possible)希望重传的数据能够尽快发送出去,这是一种随机策略;RTX-CWND,RTX-SSTHRESH,RTX-LOSSRATE 都倾向于将重传的数据发送到具有最低报文丢失率的路径上。研究表明^[7],在多路径同时传输时,倾向于使用最低丢失率的三个重传策略的性能最好,而且这三种策略能够有效减轻上面提到的接收端缓冲区阻塞问题;而 RTX-

SAME 的表现最糟糕;RTX-ASAP 的表现一般。

4 总结

随着多宿主机在普通消费者中的流行,如何利用多宿进行负载分配正成为当前的研究热点。传输层是第一个端到端的协议层,它拥有详细的端到端的路径信息,所以传输层的多路径同时传输将比其它协议层在负载平衡上做的更好,同时在传输层还可以进行拥塞控制、丢失检测与错误恢复,而且还有利于程序移植。传统的 TCP,UDP 协议都不支持多宿;作为 IETF 推荐的传输层协议,SCTP 具有 UDP 和 TCP 的许多优点,更重要的是 SCTP 内在地支持多宿。

文中介绍了基于 SCTP 多宿特性的多路径同时传输;并且总结了多路径同时传输的关键技术,包括快重传触发技术、拥塞控制窗口增长技术、延迟应答技术、减轻接收端缓冲区阻塞问题的技术以及重传时的路径选择技术。虽然这个领域的研究刚刚开始,但是许多研究结果^[4,6]表明,基于 SCTP 多宿特性的多路径同时传输能够获得极大的吞吐量提升。

参考文献:

- [1] RFC2960-2000. Stream Control Transmission Protocol[S]. 2000.
- [2] STEWART R, ONG L. Stream Control Transmission Protocol Implementer's Guide[R]. IETF Internet Draft, Work in Progress, 2005.
- [3] RFC2581-1999. TCP Congestion Control[S]. 1999.
- [4] LYENGAR J, SHAH K, AMER P, et al. Concurrent Multipath Transfer Using SCTP Multihoming[C]// In SPECTS 2004. San Jose, California:[s.n.],2004.
- [5] IYENGAR J, AMER P, STEWART R. Receive Buffer Blocking In Concurrent Multipath Transport[R]. USA: CIS Dept, University of Delaware,2005.
- [6] CARO A, AMER P, STEWART R. Transport Layer Multihoming for Fault Tolerance in FCS Networks[C]// In MILCOM 2003. Boston, MA:[s.n.],2003.
- [7] LYENGAR J, AMER P, STEWART R. Retransmission Policies for Concurrent Multipath Transfer Using SCTP Multihoming[C]// In IEEE ICON. Singapore:[s.n.],2004.
- [8] Iaminitchi A, Foster I. On Fully Decentralized Resource Discovery in Grid Environments[C]//Proceedings of the Second International Workshop on Grid Computing. Heidelberg: Springer-Verlag,2001:51-63.

(上接第 4 页)

- [6] XU ZhiWei, BU GuanYing. A theorem on grid access Control [J]. J. Comput. Sci & Technol, 2003, 718(4): 515-522.
- [7] 梁建民, 李伟, 徐志伟. 所有者为中心的网格文件共享研究[J]. 计算机研究与发展, 2003, 40(12): 1781-1786.