

# 遥科学网络环境下的可靠多播文件传输<sup>\*</sup>

石 军 鲁士文

(中科院计算技术研究所 北京100080)

**摘 要** 本文根据遥科学地基网络拓扑结构和试验数据分发的特点,在应用层设计了一个可靠多播文件传输协议。本协议具有简单、高效和安全性较高的特点,不仅适合遥科学试验数据的分发,也可用于需要安全、可靠保证的其它应用中的小规模多播文件传送。

**关键词** 遥科学,可靠多播,文件传送,差错控制,ARQ

## Reliable Multicast File Transfer in Telescience Network

SHI Jun LU Shi-Wen

(Institute of Computing Technology, Chinese Academy of Sciences, Beijing 100080)

**Abstract** Based on the characteristics of the architecture of telescience network, and for the purpose of transmitting important experiment data in the network, this paper designs a reliable multicast file transfer protocol. The protocol is simple, efficient and secure. It is suit not only for transmitting experiment data of telescience, but also for those small scale multicast transport applications that need security and reliability.

**Keywords** Telescience, Reliable multicast, File transfer, Error control, ARQ

## 1 引言

遥科学<sup>[1]</sup>是遥现(telepresence)技术和遥操作(teleoperation)技术的结合。遥现技术是人类感知范围的延伸,遥操作技术则是人类行为能力的扩展。这两种技术的结合使得研究人员可以在远离空间科学实验现场的地面实验室内,对空间科学实验进行具有交互性的操作。遥科学网络使得在空间试验室与地面设备之间以及宇航员与载荷专家之间可以建立实时联系,并允许载荷专家对其实验进行必要的遥操作。

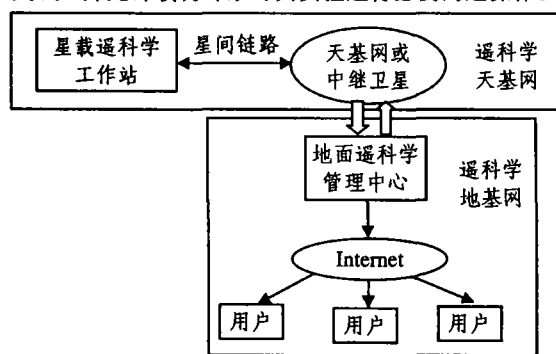


图1 遥科学网络拓扑结构

如图1所示,遥科学网络系统<sup>[2]</sup>分为天基网和地基网两个部分。遥科学地基网络在逻辑上是一个以地面遥科学管理中心为主节点的星状结构。通常情况下,地面遥科学管理中心工作站接收来自空间试验室返回的试验数据,然后再通过遥科学地基网络发送给在 Internet 上的各个用户(研究机构)。

目前,遥科学地基网使用 TCP/IP 进行可靠数据传输。如果遥科学工作站要向 N 个不同用户传输数据,它必须与每个用户单独建立一个连接,并且要为每个用户发送一次数据单元。因此,在传输数据时要同时维持 N 个不同连接,每个数据

单元也要发送 N 次。而如果使用多播技术传输数据,遥科学工作站不需要建立连接,并且任何数据单元在任何一条链路上仅出现一个备份,中心工作站也只要为所有用户发送一次数据单元。显然,多播技术可以节约网络带宽和路由器及信息服务器处理时间,但是 Stephen Deering 所提出的 IP 多播<sup>[3]</sup>(IP Multicast)模型是一种开放式的、只提供尽力而为服务的模型。这种模型的主要特点是:1)发送方不需要了解组成员关系和接收节点的信息,只要知道多播地址就可以向某个组发送数据;2)接收节点可以随时加入或离开某个组,而无需任何权威主体同意;3)IP 多播与 IP 单播一样只提供尽力而为的服务。而且 IP 多播的上层协议是 UDP 协议,因此一个使用 IP 多播的应用只能得到尽力而为的服务,除非在应用层或网络层提供可靠服务。

本文针对遥科学地基网络的星型拓扑结构和可靠性、安全性要求较高的特点,提出一个可靠多播文件传输协议,称为简单可靠多播文件传输协议(Simple Multicast File Transfer Protocol, SMFTP)。它运行在 UDP/ IP Multicast 之上,试图在应用层为多播数据传输提供可靠保证。所谓简单就是没有着重考虑协议的可扩展性。没有着重考虑协议的可扩展性是基于以下两点原因:1)遥科学试验数据传输必须考虑安全性,而在协议中增加安全保证必然影响可扩展性;2)遥科学数据传输中接收节点相对来说比较少(通常最多也就是数十个),可以不考虑可扩展性以便尽可能使协议简单、高效。

## 2 相关工作

到目前为止,人们已提出了许多可靠多播协议,但还没有一个可靠多播协议成为标准。按照接收节点的组织形式,可靠多播协议大致可分为三类:云状可靠多播协议、环状可靠多播协议和树状可靠多播协议。

1)云状可靠多播协议 直接利用 IP 多播协议所建立的

<sup>\*</sup> )本项目获国家863计划资助项目(编号2002AA742052)资助。石 军 研究生,主要研究方向为多播技术。鲁士文 教授,博导,主要从事网络协议研究,网络性能分析。

多播树传输数据,而其自身没有对接收节点做任何结构性组织。发送方在发送数据单元时加上序列号,接收节点收到正确的数据单元后就发送 ACK/NACK 消息给发送方,发送方收到重传请求后重传数据单元。如: LBRM<sup>[4]</sup> (Log-based Receiver reliable Multicast), MFTP<sup>[5]</sup> (starburst Multicast File Transfer Protocol), RAMP<sup>[6]</sup> (Reliable Adaptive Multicast Protocol)等。这类协议管理比较简单,但应用于大规模通信应用可能会出现信息爆炸现象。

2) 环状可靠多播协议 仿照令牌环网的设计方案,利用令牌来保证成功发送。特点是可靠性高、效率高。缺点是管理复杂,不支持接收节点数量较大的情况,只适合200个节点以下的网络数据传输。环状可靠多播协议有: TRP<sup>[7]</sup> (Token Ring Protocol), RMP<sup>[8]</sup> (Reliable Multicast Protocol)等。

3) 树状可靠多播协议 将所有的接收节点组成层次式的树状结构,由树状顶层向下一层发送,依此类推,并负责确认下一层是否收到。这样就可以支持大量的多播组成员,但是管理所用的额外开销比较大。这类协议通常用于大规模网络通信。TMTP<sup>[9]</sup> (Tree-based Multicast Protocol), RMTP<sup>[10]</sup> (Reliable Multicast Transport Protocol), TRAM<sup>[11]</sup> (Tree-based Reliable Multicast Protocol)等属于这类协议。

这些协议没有一个能很好满足所有应用的需求,它们都只能适合特定的应用。事实上,不同的应用对可靠多播的需求相差很大,人们很难找到一种适合大多数应用的可靠多播协议。在这些协议中,只有 RAMP 是针对小规模网络通信的,但是它是针对全光纤网络的通信应用而设计的。本文借鉴这些协议中的先进思想,提出适合小规模通信应用而且具有安全保障机制的可靠多播文件传输协议。

### 3 简单可靠多播文件传输协议

SMFTP 是基于应用层分帧原则<sup>[12]</sup> (application level framing, ALF),它运行在 UDP 之上,利用 UPD socket 接口来定义发送端的发送服务和接收端的接收功能。SMFTP 只用多播发送正常数据单元,使用单播发送控制信息和重传数据。

#### 3.1 会话阶段

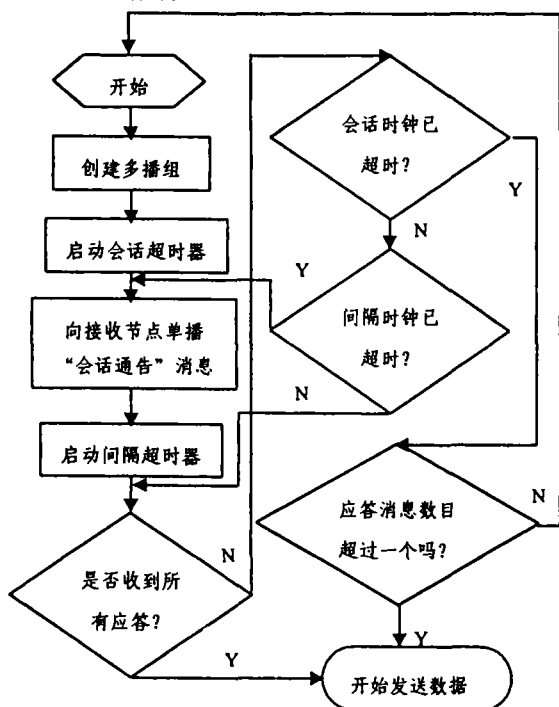


图2 服务器端会话流程

在 SMFTP 中,数据发送端是会话发起者,会话的主要任务是: 1) 利用 IP 多播协议创建一个多播组; 2) 通知各接收节点准备接收数据。由于遥科学数据传输安全性要求比较高, SMFTP 不像一般的可靠多播协议那样使用多播方式发送“会话通告”消息,而是使用单播方式将“会话通告”消息发送给各个合法节点。其中,“会话通告”消息含有本次通信的多播组地址和发送方主机地址等信息。接收节点收到“会话通告”消息后,也要先判断数据发送方是否合法。这样可以防止非法节点接收和发送数据,提高数据传输的安全性。

如图2所示,在会话期间发送方周期性地发送“会话通告”消息,这样做的好处是: 1) 确保所有接收节点收到此消息; 2) 可以防止接收节点所发送的应答消息在返回途中丢失。

#### 3.2 数据传输阶段

会话结束后进入数据传输阶段。为了保证数据传输的可靠性,任何可靠多播协议在数据传输阶段除了传输数据外,还要提供差错发现机制和差错恢复机制。为了防止出现网络拥塞和信息爆炸现象, SMFTP 采用分块、分趟的方式传输数据文件,采用基于接收节点的差错发现机制。与一般的可靠多播协议不同的是, SMFTP 只有接收完一块数据单元,才向发送方发送一个消息来报告这块数据的接收状态。发送端和接收端数据传输的流程如图3、图4所示。

• 差错发现机制和差错报告机制 SMFTP 采用基于接收节点的差错发现机制,并利用自动重传请求(ARQ)机制来报告接收状态信息。自动重发请求机制有两种应答方式: 肯定应答(ACK)和否定应答(NACK)。这两种应答方式在单播传输协议中使用都能很好的工作,如: TCP 使用 ACK, 而 NETBLT 使用 NACK。但是在可靠多播协议中使用 ACK 将带来新的问题: 1) 发送方必须计算出它到各个接收者的往返延时,从而根据延时来设置重传时钟,但是对于具有多个异构接收节点的多播通信来说,很难准确地估算出往返延时; 2) 过多的 ACK 消息可能导致网络拥塞。而使用 NACK 的缺点是,发送方不知道已发送的数据到底要保存多久。

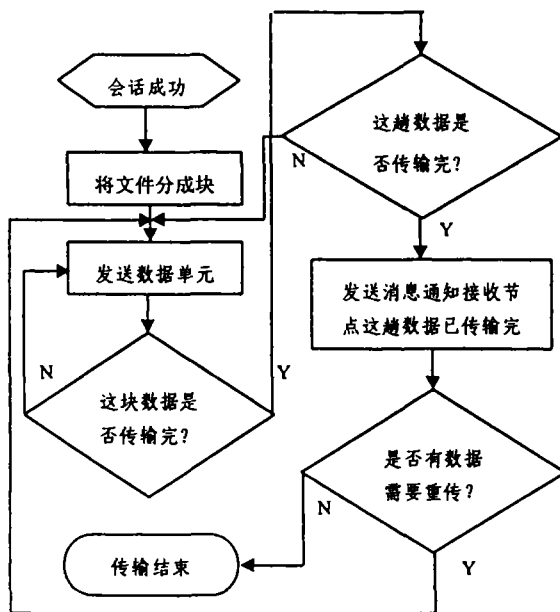


图3 发送端数据发送流程

SMFTP 为了克服两者的不足,使用 ACK+NACK 方式报告接收状态信息,并用单播方式发送消息给发送方。ACK 用于标识已传输成功的数据单元,通知发送方释放已传输成

功的数据单元所占内存;而 NACK 用于标识传输出错的数据单元,通知发送方要重传哪些数据单元。

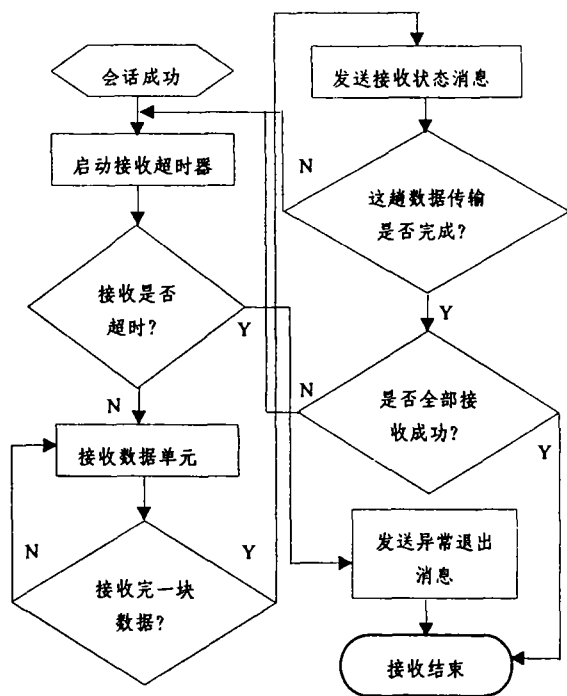


图4 接收端接收数据流程

为了适当减少报告消息的数量,发送方将文件分成 N 块相同大小的数据块,每块数据又等分为 M 个数据单元,当一块数据传输结束或等待一定时间还没收到数据后,接收节点就向发送方返回一个“报告消息”,报告这块数据的每个数据单元的接收状态;发送方收到所有接收节点返回的“报告消息”后,释放那些已传输成功的数据单元所占的内存。

•数据传输和差错恢复机制 其发送方采用分趟方式传输数据文件,其中第一趟以多播方式传输整个数据文件,不重传数据,以后第 N 趟重传第 N-1 趟中没有传输成功的数据单元。发送方可以采用多播方式或单播方式重传数据。通常针对大规模网络通信的可靠多播协议采用多播方式重传数据,而针对小规模网络通信的可靠多播协议大多采用单播方式重传数据。遥科学地基网络通信规模相对来说比较小,而且接收节点分布比较分散,多个接收节点同时在接收同一个数据单元时发生数据传输错误的情况很少。如果采用多播方式重传数据将会造成大量的带宽浪费,而且可能会导致网络拥塞。因此,SMFTP 采用单播方式重传数据。

### 3.3 关闭会话阶段

数据文件传输结束后,接收节点可以立即退出会话。发送方在退出前必须先向所有接收节点发送“DONE”消息,通知接收节点此次数据传输已完成,而且只能在下面两个条件有一个为真的情况下才能退出:1)收到所有接收节点的应答消息;2)等待应答超时。

## 4 性能分析

考虑到协议的数据传输效率对协议的最大吞吐量、带宽的有效使用和服务器处理时间都有很大影响,我们将从数据传输效率方面对 SMFTP 和 FTP 进行比较。

我们设一个文件可分为 N 个数据单元,文件要发送给 R 个节点,每个接收节点的包丢失率都为 p。另外,我们还假设

两个条件:1)各接收节点的丢包事件是不相关的;2)应答包不丢失。这两个条件在实际环境中不可能成立,因此会影响性能分析的精确度,但基本能反映实际情况,而且这种假设有利于 FTP,因为 FTP 应答包比 SMFTP 多,而且所要求的网络带宽也比 SMFTP 多。

若用 FTP 传输,要使每个节点都成功地收到整个文件,发送方总共要发送数据包的次数为:

$$A = R * (N + N * p + N * p^2 + \dots + N * p^{n-1} + \dots)$$

$$= R * N * \sum_{i=0}^{\infty} p^i = \frac{R * N}{1 - p}$$

因此,FTP 的数据传输效率为:

$$W = \frac{N}{A} = \frac{1 - p}{R}$$

用 SMFTP 传输文件,要使每个节点都成功地收到整个文件,发送方总共要发送数据包的次数为:

$$A' = N + R * (N * p + N * p^2 + \dots + N * p^{n-1} + \dots)$$

$$= N + R * N * \sum_{i=1}^{\infty} p^i = N * (\frac{1 - p + R * p}{1 - p})$$

因此,SMFTP 的数据传输效率为:

$$W' = \frac{N}{A'} = \frac{1 - p}{1 - p + R * p}$$

综合以上分析,SMFTP 与 FTP 的传输效率之比为:

$$S = \frac{W'}{W} = \frac{R}{1 - p + R * p}$$

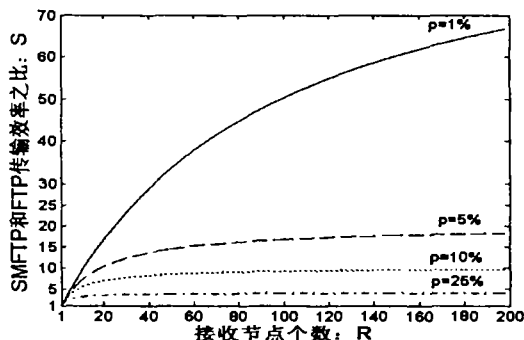


图5 SMFTP 和 FTP 的数据传输效率比较

图5显示了在丢失率为1%、5%、10%、25%时,SMFTP 和 FTP 的数据传输效率之比 S 随接收节点数目 R 的变化情况。从图中我们可以得出如下结论:1)SMFTP 和 FTP 的数据传输效率之比 S 大于1,而且 S 与 R 成正比,这说明 SMFTP 数据传输效率明显比 FTP 高,而且接收节点越多,SMFTP 优越性越能体现出来;2)当接收节点达到一定数目后 S 增加比较缓慢,这说明 SMFTP 不适合大规模网络通信应用;3)接收节点的包丢失率越小,SMFTP 和 FTP 的数据传输效率之比越高。从2)、3)我们可以看出,如果通信规模和包丢失率太大,则会降低 SMFTP 的数据传输效率。这是 SMFTP 为了节约网络带宽而采用单播方式重传数据的原故。在实际环境中,接收节点的包丢失率很低,一般低于5%,因此 SMFTP 特别适合于小规模、一对多点的文件传输应用。

结束语 本文结合我们承担的国家863遥科学课题针对遥科学地基网络试验数据传输的特点提出了一个简单可靠多播文件传输协议(SMFTP),它有如下几个特点:1)提供封闭式的组成员管理,来保证遥科学试验数据传输的安全;2)分趟传输数据;3)使用单播进行数据重传;4)针对一对多文件传输应用设计。SMFTP 的目标是为遥科学地基网络试验数据传

(下转第48页)

建筑物内位置和在这些位置上观察的(或者估计的)来自各 AP 的信号强度的数据库。例如,图上的某个出口可能记录成  $(x, y, z, ss, (i=1 \cdots n))$  的形式,这里,  $(x, y, z)$  是记录信号点的物理地址,  $ss_i$  是来自 AP<sub>i</sub> 的信号强度。

(2) 移动站测量区域内每个 AP 的信号强度。然后,它在电波地图数据库中搜索,以找出与所测信号强度最为匹配的强度值矢量。系统将与该最佳匹配的强度值对应的位置作为移动站的定位值。一种匹配方法是计算地图中每一强度矢量  $(ss_i, i=1, \cdots, n)$  与测得值  $(ss_m, i=1, \cdots, n)$  这两个信号空间之间的欧几里得距离。选取使该距离最小的 SS 矢量作为用户定位的物理坐标估计值。

### 3.3 室外射频定位系统

第一个典型的户外基于射频的地理定位系统是发展用于广域蜂窝移动通信的,最近,有若干种定位系统被提了出来<sup>[11]</sup>。用于蜂窝电话定位的可供选择的技术包括测量信号衰减、DOA、和/或到达时间差(TDOA)。研究表明,这些系统在户外环境下是有前途的。

基于 GPS 技术的系统对于 WLAN 用户的户外地理定位是非常有用的<sup>[11,12]</sup>。而那些无法接收到多于四颗 GPS 卫星信号的 WLAN 用户(如处于建筑物/大树阴影下甚至地铁里)的地理定位问题,就可以借助于所谓的伪 GPS 信号,当然,前提是该区域里已经为此目的建立了某些站来发射类 GPS 信号。由于未来 WLAN 的大部分增长点是在健康设备方面的应用,因此最近提出了一种称作“转发式 GPS”的方案<sup>[13]</sup>,为某些 WLAN 用户的精确地理定位提供支持,这些用户的保健医师需要监视他们的准确位置及其生物数据。

**结束语** 以前关于室内外定位系统方面的研究已经在专用硬件和技术上有了许多发明。除了上述地理定位方案之外,还有其它一些方案,尽管也很有趣和具有创造性,但没有在此列出。这些系统经常是为了单一性的用户定位目的而构建的,即它们不提供任何数据网络服务。

更不幸的是,现有的 WLAN 用户地理定位系统和技术,甚至某些发展中的技术并没有设计成适合基于 802.11a/b MAC 的 WLAN,因为它们没有考虑 802.11a/b 标准的特性,

尤其是在速度、带宽、频率以及编码方案等方面。由于这些原因,在此领域有许多研究工作有待于完成。这样,值得进一步研究基于 802.11a/b MAC 的 WLAN 地理定位的可行性问题。

### 参考文献

- 1 IEEE 802.11b Wireless LANs -- Wireless Freedom at Ethernet Speeds:[Technical Paper of 3Com Corporation]
- 2 Brenner P. A Technical Tutorial on the IEEE802.11 Protocol. Breezecom Wireless Communications
- 3 Crow B, Widjaja I, Kim J, Sakai P. IEEE 802.11 Wireless Local Area Networks. IEEE Communications Magazine, Sept. 1997. 116~126
- 4 IEEE P802.11, Working Group for Wireless Local Area Networks, 1999
- 5 Kurose J F, Ross K W. Computer Networking. Addison Wesley, 2000
- 6 Want R, Hopper A, Falcao V, Gibbons J. The Active Badge Location System. ACM Transactions on Information Systems, 1992, 10(1): 91~102
- 7 Ward A, Jones A, Hopper A. A New Location Technique for the Active Office. IEEE Personal Communications, 1997, 4(5): 42~47.
- 8 Azuma R. Tracking Requirements for Augmented Reality. Communications of the ACM, 1993, 36(7): 50~51
- 9 Hodes T D, Katz R H, Schreiber E S, Rowe L. Composable Ad Hoc Mobile Services for Universal Interaction. MobiCom'97 Proceedings, Sep. 1997. 1~12
- 10 Bahl P, Padmanabhan V N, Balachandran A. Enhancements to the RADAR User Location and Tracking System; [Technical Report (MSR-TR-2000-12) of Microsoft Corporation]
- 11 Tekinay S. Wireless Geolocation Systems and Services. Special Issue of the IEEE Communications Magazine, April 1998
- 12 Enge P, Misra P. Special Issue on GPS: The Global Positioning System. In: Proc. of the IEEE, Jan. 1999. 3~172
- 13 Chen Yongguang, Kobayashi H. A Repeating GPS System for Geolocation. In: Proc. of the 2001 Conf. on Information Sciences and Systems, The Johns Hopkins University, Baltimore, Maryland, 2001
- 14 protocol. In: Fifteenth Annual Joint Conf. of the IEEE Computer and Communication Societies, San Francisco, CA, March 1996
- 15 Chang J M, Maxemchuk N. Reliable broadcast protocol. ACM Trans. on Computer Systems, 1984, 2(3)
- 16 Whetten B, Montgomery T, Kaplan S. A high performance totally ordered multicast protocol. Theory and Practice in Distributed Systems, International Workshop, Lecture Notes in Computer Science, Sep. 1994
- 17 Yavatkar R, Griffioen J, Sudan M. A reliable dissemination protocol for interactive collaborative applications. In: Proc. of ACM multimedia, Nov. 1995
- 18 Lin J, Paul S. RMTP: A reliable multicast transport protocol. In: Proc. of the IEEE Infocom'96, Nov. 1995
- 19 Chiu D M, Hurst S, Kadansky M, et al. TRAM: A tree based reliable multicast protocol. <http://research.sun.com/research/techrep/1998/sml-tr-98-66.pdf>, 1998
- 20 Floyd S, Jacobson V, Liu Ching-Gung, et al. a reliable multicast framework for light weight sessions and application level framing. ACM Sigcomm Conf. 1995
- 21 张珩,李庚田. 遥科学的概念、应用与发展. 中国航天, 1997(11)
- 22 鲁士文. 天地一体化的遥科学系统和网络技术. 计算机工程与设计, 2002(1)
- 23 Deering S, Hinden R. Multicast Routing in datagram inter-network and extended LANs. ACM Transactions on Computer Systems, May 1990
- 24 Holbrook H W, Singhal S K, Cheriton D R. Log-based receiver reliable multicast for distributed interactive simulation. In: Proc. ding of ACM SIGCOMM, Aug. 1995
- 25 Miller K, Robertson K, Tweedly A, et al. Startburst multicast file transfer protocol specification. <http://www.globecom.net/ietf/draft/draft-miller-mftp-spec-03.html>, 1998
- 26 Koifman A, Zabele. RAMP: a reliable adaptive multicast

(上接第30页)

输提供一种高效、可靠的传输方法。但是, SMFTP 不局限于遥科学地基网络中的应用,它同样适合于其他小规模、一点对多点的文件传输应用。

### 参考文献

- 1 张珩,李庚田. 遥科学的概念、应用与发展. 中国航天, 1997(11)
- 2 鲁士文. 天地一体化的遥科学系统和网络技术. 计算机工程与设计, 2002(1)
- 3 Deering S, Hinden R. Multicast Routing in datagram inter-network and extended LANs. ACM Transactions on Computer Systems, May 1990
- 4 Holbrook H W, Singhal S K, Cheriton D R. Log-based receiver reliable multicast for distributed interactive simulation. In: Proc. ding of ACM SIGCOMM, Aug. 1995
- 5 Miller K, Robertson K, Tweedly A, et al. Startburst multicast file transfer protocol specification. <http://www.globecom.net/ietf/draft/draft-miller-mftp-spec-03.html>, 1998
- 6 Koifman A, Zabele. RAMP: a reliable adaptive multicast