

基于模拟的网络流量自相似现象分析^{*}

Simulation-based Research on Self-similar Phenomena

赵佳宁 李忠诚

(中科院计算技术研究所信息网络研究室 北京100080)

Abstract The notion of self-similarity has been the focus of recent research on traffic engineering. This paper addresses the cause of the self-similar network traffic. Based on a simulation with "Network Simulation-Version 2", we demonstrate that large number of ON/OFF sources with heavy tailed distributed random ON and/or OFF periods are not necessary to explain self-similarity, while TCP flow-control mechanism can be a deterministic factor of self-similarity.

Keywords Self-similarity, Heavy-tail, TCP congestion control, High variability

一、引言

自从1993年文[2]发表以来,网络流量的自相似特征日益受到重视,使传统的基于泊松和马尔可夫行为的流量模型受到了质疑^[1],但是关于自相似现象的成因以及这一现象对网络流量建模的影响一直没有确定性的答案,文[3]中提出了高可变性(High Variability)是网络流量自相似现象的可能原因。但在文[1]中,作者基于 Network Simulation-Version 2(以下简称 ns2)网络模拟器,模拟产生了单个 TCP 对话流量的自相似现象,从而提出高可变性未必是出现自相似现象的必要前提,而 TCP 流量控制机制本身就是一个能产生自相似现象的确定性过程。然而在文[1]中,作者只模拟出了链路上单个 TCP 对话流量的自相似现象,但是在现实网络观测到自相似现象都是对于链路上总的流量来说的,而不是针对单个对话流量来说的。因此,本文设计了新的网络模拟方案,以探讨网络链路总流量出现自相似现象的条件,从模拟结果可以看出同一瓶颈链路上的多个(但不是全部)自相似 TCP 对话流相累加,其流量和仍旧是自相似流,并通过对模拟过程的统计分析,说明该过程中不存在高可变性,从而更有力地说明:TCP 阻塞控制机制本身就可以成为出现流量自相似的决定因素,高可变性不是网络出现自相似现象的唯一原因。

本文介绍了网络自相似的一些重要概念;讲述关于自相似现象原因的研究情况;介绍我们采用的网络模拟方案;并通过对模拟结果的统计分析,说明了

TCP 过程本身就可以成为自相似现象的决定成因;最后是总结。

二、自相似的一些基本概念

2.1 自相似的定义及自相似参数的算法

自相似的定义不尽相同,本文采用的定义是文[2]对自相似的数学描述:

设 $X = (X_t; t=0, 1, 2, \dots)$ 是一个协方差稳定(或是广义上的稳定)的随机过程,即: X 有常均值 $\mu = E[X_t]$, 有限方差 $\sigma^2 = E[(X_t - \mu)^2]$, 并且 X 的自相关函数 $r(k) = E[(X_t - \mu)(X_{t+k} - \mu)] / E[(X_t - \mu)^2]$ 只依赖于 k , 设 X 的自相关函数还具有以下特征:

$$r(k) \sim a_1 k^{-\beta}, k \rightarrow \infty$$

其中 $0 < \beta < 1$, a_1 以及下文出现的 a_2, a_3 表示固定正常数。对每一个 $m (m=1, 2, 3, \dots)$, 令 $X^{(m)}$ 表示由 X 生成的新的序列, 其中的元素是从 X 中依次取连续的 m 个 X 元素平均值, 并且各个“ m 元素块”不重叠, 即: $X_k^{(m)} = 1/m(X_{k-m+1} + \dots + X_{k+m})$, ($k \geq 1$)。注意: 对每个 m 来说, $X^{(m)}$ 定义了一个协方差恒定的过程。令 $r^{(m)}$ 表示 $X^{(m)}$ 相应的自相关函数。如果 $X^{(m)}$ 的自相关函数与 X 的自相关函数有相同的结构, 例如: 对所有的 m , 有 $r^{(m)} = r(k)$, 则称过程 X 严格(二阶)自相似, 自相似参数为 $H = 1 - \beta/2$, 换句话说, 如果 X 与 $X^{(m)}$ 在二阶统计数据上相同的话, 则称 X 为严格自相似过程, 如果 X 满足 $r^{(m)}$ 与 $r(k)$ 渐进相符(例如当 m 值越来越大时有 $r^{(m)} \sim r(k)$), 则称 X 渐进(二阶)自相似。自相似参数 H 又称为 Hurst 参数, 并且 $0.5 < H < 1$ 。

*) 本工作得到国家自然科学基金项目 69703001 支持, 赵佳宁 硕士研究生, 主要研究方向为计算机网络流量工程和可编程序网络, 李忠诚 研究员, 博士生导师, 主要研究方向为计算机通信和网络综合评测。

从感性的角度看,自相似过程的特点是它的聚合过程(aggregate processes), $X^{(m)}$ 的相关结构在 m 趋于无穷大时是非退化的,即 $\sum_j r^{(m)}(k) = \infty$,而传统的随机模型产生的序列 X ,相对应的 $r^{(m)}$ 当 m 趋于无穷大时, $r^{(m)}(k)$ 趋于0,序列的Hurst 参数为 $H=0.5$ 。如果说某一网络流量具有自相似特征,说明该网络流量具有很强的突发性(burstiness)。

自相似过程还具有以下特点:

1 样本均值的方差的减小速度慢于样本大小的倒数,即:当 $m \rightarrow \infty$ 时

$$\text{var}(X^{(m)}) \sim am^{-2}, 0 < \beta < 1 \quad (1)$$

2 自相似过程的频率分布 $f(\lambda)$ 在 $\lambda \rightarrow 0$ 时,有

$$f(\lambda) \sim a_2 \lambda^{-1-\beta} \quad (2)$$

以上特点为判断某一过程是否自相似以及计算自相似参数提供了依据。

计算自相似参数 H 有很多种方法^[1],本文采用的方法主要是方差法、R/S 法、频谱图法以及小波分析法,以下简单介绍小波分析算法,文[7]提出的小波分析方法不仅能计算自相似参数 H ,而且还为进一步分析时序序列的单相似分形(monofractal)和多相似分形(multifractal)提供了工具^[11,12]。

下面我们介绍一下本文在计算参数时用到的关于小波分析的一些基本概念。

假定用一时间序列 $X_{n,k}$, $k=0,1,2,\dots$,来表示在监测中所取的最小时间间隔内所测到的字节数, n 代表最高级别的分辨率。可以采用以下算法将分辨率降一级:

$$X_{n-1,i} = \frac{1}{\sqrt{2}}(X_{n,2i} + X_{n,2i+1}) \quad (3)$$

而得到 X_{n-1} ,而这两个序列的差别可以用细节信号(detail signal):

$$D_{n-1,i} = \frac{1}{\sqrt{2}}(X_{n,2i} - X_{n,2i+1}) \quad (4)$$

来表示。我们可以将 X_n 用“模糊”一级的 X_{n-1} 与细节信号 D_{n-1} 来表示: $X_n = 2^{-1/2}(X_{n-1} + D_{n-1})$ 。不断重复这个过程,可以得到 $X_n = 2^{-n/2}X_0 + 2^{-n/2}D_0 + \dots + 2^{-1/2}D_{n-1}$ 。我们称 $D_{j,k}$ 为离散小波变换系数,可用 $d_{j,k}$ 表示。以上(3)、(4)式便构成了一个最简单的 Haar 离散小波变换,而 $|d(j,k)|^2$ 能表示被分析信号在 $2^j k$ 时刻和 $2^{-j} \nu_0$ 频率的能量,其中 ν_0 是所选小波类型(如上例中的 Haar 小波)决定的基准频率。根据自相似过程的特点,利用式(2),经过计算可得如下公式:

$$\log_2\left(\frac{1}{n_j} \sum_k |d(j,k)|^2\right) = (2H-1)j + c \quad (5)$$

其中 H 为对子相似参数 H 的一个估计值, j 表示分辨率下降的级数, n_j 表示分辨率下降 j 级时 $d(j,k)$ 的个数。根据上式,我们不仅可以估算小波参数 H ,并且可

以从不同频率中信号平均能量的大小分析信号分形(fractal)特征的变化。

三、关于网络流量自相似现象的解释

虽然局域网和广域网的自相似流量特征已经被接受,但对于它的成因并没有确定性的说法。通过实际监测、统计,文[5]提出广域网的流量呈自相似特征的同时,也针对这一现象提出了解释:网络上传输的文件大小、用户行为(例如两个请求之间的思考时间)都符合或近似符合“重尾”(heavy-tailed)分布,而这可能是出现自相似流量的原因,如果某一统计概率分布满足:

$$\text{当 } x \rightarrow \infty \text{ 时, } P[X > x] \sim x^{-\alpha}, 1 < \alpha < 2 \quad (6)$$

则称此分布为“重尾”分布,文[4]采用文[9]提出的“包行列”(packet train)流量模型,将局域网上不同的源-目对之间的流量的“有”、“无”定义为“开”、“关”,并在数学上证明:如果每个开关数据源(ON/OFF sources)的开时段和关时段互相交替,并且开时段时长和关时段时长两者至少有一个具有高可变性,那么多个互相独立的开/关数据源相叠加能产生具有自相似特征的网络流量。其中所谓高可变性,等同于统计上的“无限方差”,在数学上,可以用“重尾”分布来描述。同时,文[4]还将实际观测到的几组出现自相似现象的局域网数据进行统计,对这几组被观测网络数据不同源-目对用开/关模型进行分析,发现不同源-目对的开时段时长分布或关时段时长分布是符合“重尾”分布的。因此“互相独立的重尾分布的开/关源相叠加”成为对自相似现象一个可能的解释。然而文[1]通过 ns2 上一个简单的模拟配置模拟出了单个 TCP 流量的自相似现象,从而提出 TCP 的阻塞控制机制有可能是出现自相似现象的确定性因素,而不需要符合重尾分布多个开/关数据源。本文设计了新的网络流量模拟方案,在 ns2 上通过更符合网络实际运行情况的模拟设置,模拟出链路总流量的自相似现象,有力地说明 TCP 阻塞控制机制在一定情况下可以是自相似现象的决定性因素,而多个符合重尾分布的开/关数据源叠加未必是出现自相似的必要前提。

四、网络流量模拟方案的设计

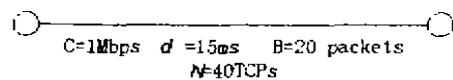


图 1

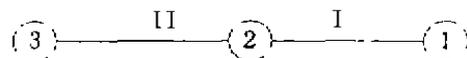


图 2

ns2是在网络研究中广泛使用的一种系统模拟软件,对网络流量工程更是一个重要的辅助工具。文[1]中在 ns2上模拟出了单个 TCP 流量的自相似现象,其模拟方案如图1所示,图中 C 为链路速率,d 为链路延迟,B 为缓冲区大小,N 为链路上的 TCP 通话数。根据文[8],B/N 在这个设置下控制着系统的行为,当 B=20、N=40 个 TCP 时,链路成了瓶颈,其丢包率达 16%,而此时单个 TCP 对话的字节流量呈自相似特征,其自相似参数约为 0.8。由此文[1]认为,尽管多个符合重尾分布的开/关源叠加可以产生自相似数据流,却未必是自相似现象的必要原因,而是 TCP 阻塞控制机制和符合重尾分布的开/关源共同作用的结果。但是文[1]只模拟出了链路上单个 TCP 对话流量的自相似特征,而在现实网络观测到的自相似现象都是针对网络链路总流量来说的。文[1]提出并使用数学手段证明:瓶颈缓冲区的单个的自相似流进入另外非瓶颈的缓冲区,多个互相独立的自相似数据流叠加可以产生新的自相似流,这样就得到了自相似的链路总流量。

虽然文[1]在数学上证明来自不同瓶颈缓冲区的单个自相似流量进入非瓶颈缓冲区,其流量和是自相似流量,但并没有做进一步的模拟,而且我们认为“来自不同瓶颈缓冲区的单个自相似流量进入非瓶颈缓冲区”在现实网络中并不经常出现。虽然在图1所示的网络模拟中,40个 TCP 流组成的链路总流量接近链路线

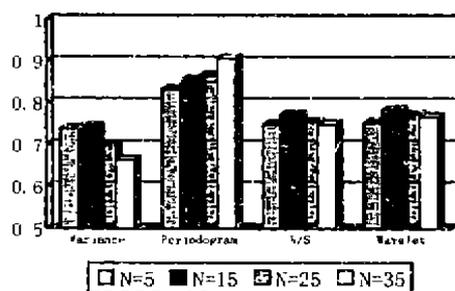


图3

采用四种方法计算出的链路 II 上的总流量的 H 虽然互有不同,但均明显大于 0.5。这样,可以认为,链路 II 上的总流量是具有自相似特征的,证实了我们的预测,即部分瓶颈链路的自相似 TCP 流进入非瓶颈链路后总流量是自相似的。但在这种情况下,我们并不认为可以下结论:实际网络自相似过程可能不需要多个重尾分布的开/关过程叠加。这是因为:(1)模拟中的连续 5 个小时的 TCP 对话与网络的实际情况有较大差距,现实中的网络很少会出现这样的情况;(2)根据文[4]的观点,多个重尾分布的开/关数据源叠加能产生

速,不是自相似流,但其中一部分 TCP 流的叠加和有可能是自相似的。基于以上考虑,我们提出一个更符合现实网络的模拟方案,目的是模拟出自相似的链路总流量,以此研究出现链路总流量自相似的条件。该模拟方案如图2所示,图中链路 I 与图1中的配置一致:C=1Mbps, d=15ms, B=20 packets;而链路 II 的设置: C=5Mb, d=5ms, B=120。同时,分别设置 M 个 TCP 对话在节点①、②之间,其中①为源节点,②为目的节点;再设置 N 个 TCP 对话在节点①、③之间,其中①为源节点,③为目的节点,这 N 个 TCP 对话要经由节点②。令 M+N=40,则可以看见,链路 I 的设置以及其上来往的 TCP 对话数与图1的配置是完全相同的,由此我们可以判断,在链路 I 上的单个 TCP 对话应该如文[1]中所说,具有自相似特征。而在链路 II 上,有 N 个 TCP 对话,来自同一链路 I,由于链路 II 的速率大于 I 速率,所以链路 II 不会成为“瓶颈”,在这样的情况下,当 N<40,尤其是 N<<40 时,根据我们的预测,链路 II 是有可能出现总流量自相似现象的。在这种配置下,我们进行了 5 个小时连续的 TCP 流模拟,以 0.1 秒为单位对链路 II 上的总流量进行监测,再分别采用方差法、频率法、R/S、小波分析四种方法计算 H 值,结果如图3所示。

五、模拟结果分析

从图3中可以看出,在 N=5,15,25,35 的情况下,

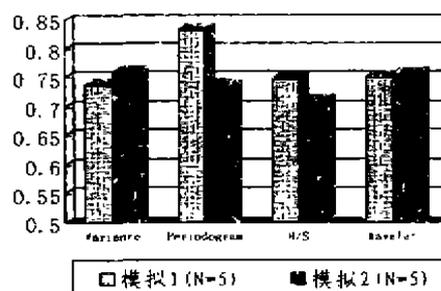


图4

自相似流,而如果我们把模拟中单个 TCP 流的“有”“无”定义为“开”“关”,不同的 TCP 流看成不同的源-目的对话,TCP 流的开关时间也可能具有重尾特征。基于以上考虑,我们在同一配置下,将模拟过程进行了修正以使其更符合实际网络的运行情况,重新进行了模拟。在新的模拟中,不再使用 5 个小时连续的 TCP 流,而是改为此伏彼起的 TCP 流。在模拟开始时,链路 I 上有 40 个 TCP 对话,每个 TCP 对话持续的时间采用 [2,1800s] 之间的随机数;在任一个 TCP 对话停止后,在等待 [2,50s] 上的一个随机时间后,会有一个新的

TCP 对话启动,而这个新的 TCP 流的持续时间依然是 $[2, 1800s]$ 上的一个随机数,如此此伏彼起,直到5个小时的模拟结束,持续时间和等待时间采用平均分布是为了尽量与重尾分布不同。值得注意的是,为了保证一定的阻塞程度,模拟中 TCP 对话持续的时间可能远大于此伏彼起的两个 TCP 对话之间等待的时间。在新的模拟过程下,我们重新对链路 II 上的流量进行了分析,如图4所示的是 $N=5$ 时的 H 值,其中“模拟1”代表连续5个小时 TCP 流的模拟,“模拟2”代表修改后此伏彼起的 TCP 流的模拟,可以看出前后两次模拟中链路 II 上的流量均呈自相似特征,H 参数相差不大。

同时,我们还考察了两种模拟设置之下的单个 TCP 对话流。定义单个 TCP 对话流流量的“有”“无”为“开”“关”,对“开”和“关”的持续时间分别考察分析,在 $N=5$ 时,共取得1050个“开”时间样本和1050个“关”时间样本,对样本采用文[4]提到的 qq 方法和 Hill 方法分析。qq 方法和 Hill 方法的结果采用图表判断,典型的重尾分布如 Pareto 分布在 qq 图上应该有一段明显的线性区,而在 Hill 图上应该有对应的一段趋于平稳的部分,如图5所示,而以我们模拟中的 $N=5$ 时的“开”时间样本为例,如图6所示,可以看出,模拟中链路

上的单个 TCP 流的开时段都没有重尾特征,分析表明,在历次模拟中的单个 TCP 流都不符合重尾分布,

通过以上分析可以看出,多个重尾分布的开/关源叠加并不是对自相似过程的唯一解释,TCP 的阻塞控制机制可以对产生自相似流起决定性作用。

在按照图2配置进行的模拟中,我们注意到,即使在 $N=35$ 的情况下,非瓶颈链路 II 上仍出现自相似现象,这说明此时链路 II 上的流量仍有突发性。而根据式(5)做出的小波分析图(如图7所示)已经可以看出,在 $N=35$ 时,曲线的线性区已经缩小在 $j \in [4, 12]$ 上。在根据式(1)对 $N=5, N=35$ 所做的方差图(如图8所示)上,在 m 接近最大值时,曲线的线性度下降,而 $N=35$ 时尤为明显。这是由于在链路 I 上总共40个 TCP 中的35个 TCP 对话流量进入链路 II,它们的平均流速和为 $945kb/s$,接近链路 I 的速率上限,所以此时在大的时间规格(scale)上(大于 $2^{12} \times 0.1s = 409.6s$),流速已经不再具有自相似性。尽管如此,当时间规格在 409.6秒内,在 $N=35$ 时,虽然流速已接近链路 I 的上限,链路 II 上的总流量仍具有自相似特征,这表明即使瓶颈链路上的大部分 TCP 对话流入非瓶颈链路时,非瓶颈链路的总流量仍存在自相似特征。

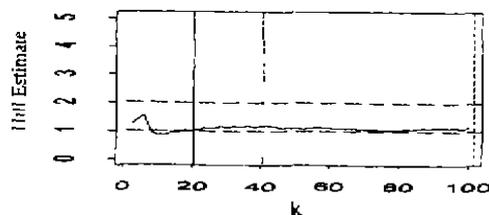
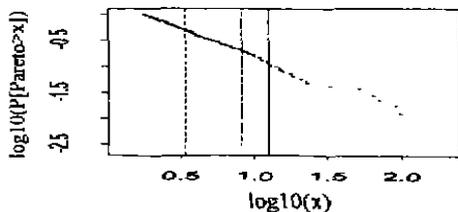


图5

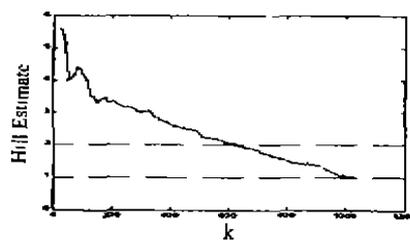
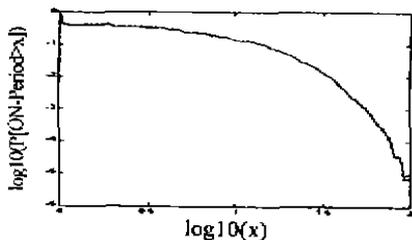


图6

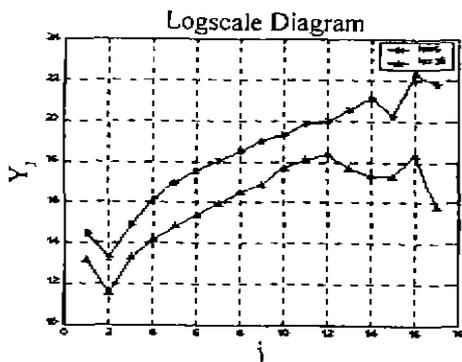


图7

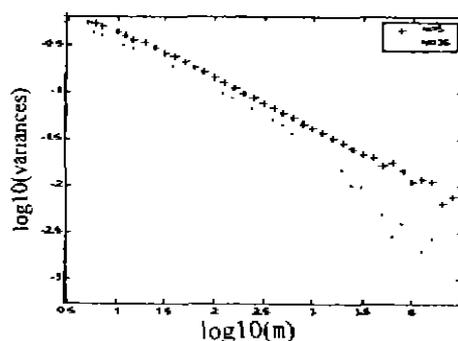


图8

结束语 通过在 ns2 上的模拟,我们研究了出现网络链路总流量自相似的条件,说明 TCP 的阻塞控制机制可以成为网络自相似现象的决定性因素,而不必要高可变性。而由于模拟配置很简单,可以很方便地改变其中流速、延迟、排队策略等因素,考察这些因素对自相似现象的影响,这也是我们今后工作的方向。

参考文献

- 1 Veres A, Boda M. The Chaotic Nature of TCP Congestion Control. IEEE Infocom'2000, Tel Aviv, Israel, Mar 2000
- 2 Leland W E, et al. On the Self-Similar Nature of Ethernet Traffic. ACM SIGCOMM'93, San Francisco, CA, USA, Sep. 1993
- 3 Paxson V, Floyd S. Wide Area Traffic: The Failure of Poisson Modeling. IEEE/ACM Transactions on Networking, 1995, 3: 226~244
- 4 Willinger W, et al. Self-similarity through high-variability: Statistical analysis of Ethernet LAN traffic at the source level. IEEE/ACM Transactions on Networking, 1997, 5(1): 71~86
- 5 Crovella M E, Bestavros A. Self-similarity in world wide

- web traffic - evidence and possible causes. In: Proc. of ACM Sigmetrics'96, 1996: 160~169
- 6 Taqqu M S, Teverovsky V, Willinger W. Estimators for long-range dependence: an empirical study. Fractals, 1995, 3(4): 785~798
- 7 Abry P, Veitch D. Wavelet Analysis of Long-Range-Dependent Traffic. IEEE Transactions on Information Theory, 1998, 44(1)
- 8 Morris R. TCP Behavior with Many Flows. IEEE International Conference on Network Protocols, Oct. 1997
- 9 Clafly K C, Braun H W, Polyzos G C. A parameterizable methodology for internet traffic flow profiling. IEEE Journal on Selected Areas in Communications, 1995, 13: 1482~1494
- 10 Floyd S, Jacobson V. Random Early Detection Gateways for Congestion Avoidance. IEEE/ACM Transactions on Networking, Aug. 1993
- 11 Feldmann A, Gilbert A G, Willinger W. Data networks as cascades: Investigating the multifractal nature of Internet WAN traffic. Computer Communication Review, 1998, 28(4)
- 12 Feldmann A, et al. Dynamics of IP traffic: A study of the role of variability and the impact of control. In: Proc of the ACM/SIGCOMM'99, Cambridge, MA

(上接第68页)

修改路由器端口生存周期为查询间隔。最后,将该包转发给 ID=1 VLAN 对应的所有端口(主机)。

当接收到主机组成员关系报告包时,IGMP 守护进程首先判断该多播组是否存在,如果不存在,为该多播组(地址)创建新的 VLAN,并将新的 VLAN 写到所有 GT48310A 的 VLAN 表;如果存在,检查该端口是否在 VLAN 中,如果不在,加入 VLAN 中,修改所有 GT48310A 的 VLAN 表;然后,设置/修改该端口的生存周期为组成员资格间隔,并将该包转发给“224. 0. 0. 2” VLAN 对应的所有端口(路由器组)。

当接收到主机离开某个多播组包时,IGMP 守护进程转发该包给“224. 0. 0. 2” VLAN 对应的所有端口(路由器组)。

当接收到路由交换设备发送的特定组成员查询包时,IGMP 守护进程将该包转发给特定组对应 VLAN 的所有端口(主机),修改该特定组 VLAN 的所有端口生存周期为离开查询时间间隔。

当接收到其它 IGMP 类型包时,IGMP 守护进程转发该包给除进入端口外的其它所有端口。

IGMP 守护进程设置一个定时器用于更改组播表项的每个端口生存周期。当接收到一个 IGMP 包或定时器到期时,修改表项中每个端口的生存周期。

这样,可以为每个电视频道(或 VOD 点播视频

流)建立一个多播组,每个用户选择电视频道(或 VOD 点播视频流)实际就是动态加入/离开一个多播组。我们通过实验室的两个卫星接收器,成功进行了计算机终端选择卫星电视频道的测试。

结束语 VLAN 的标准正在逐渐地统一和完善,它的种类比较多,各厂家的实现也互不相同,但大多数厂家支持 IEEE802. 1Q 标准。我们在 Galileo 公司提供硬件支持的基础上,在宽带接入交换设备中实现了基于端口的 VLAN、基于协议和端口混合的 VLAN 以及基于 IP 组播的 VLAN。在我们正在研发的宽带接入路由交换设备中,可以支持更高层次的 VLAN、IP 组播协议的实现更为简单。

参考文献

- 1 Draft Standard P802. 1Q/D11 IEEE Standards for Local and Metropolitan Area Networks. Virtual Bridged Local Area Networks July 30, 1998
- 2 Fenner W. RFC2236 Internet Group Management Protocol, Version 2. Nov. 1997
- 3 Parkhaust W R. 东京工作室 译. Cisco 组播路由与交换技术. 机械工业出版社, 1999
- 4 Rubini A. Linux Device Drivers. O'Reilly & Associates, Inc. 1998
- 5 Galileo Inc. GT-48310/GT-5. 311/GT-48312 Switched Ethernet Control for 100BASE-X Revision 1. 2 April 03, 2000
- 6 Galileo Inc. GT-48300 GalNet-II Crossbar Switch Revision 1. 1. June 16, 1999