·博士论坛·

网络测量及行为学研究综述

程 光 龚 俭 丁 伟
(东南大学计算机科学与工程系,南京 210096)
(江苏省计算机网络技术重点实验室,南京 210096)
E-mail:gcheng@njnet.edu.cn

摘 要 随着互联网的发展,理解网络行为对于网络管理、规划和发展都有重要意义,网络流量测量是研究网络行为的基础。论文首先根据目前 Internet 发展的特点、网络管理的缺点和不足,分析网络行为学的研究难点和重点。将网络测量及行为学研究分成测量、流量行为分析、端至端行为分析等三个方面,并分别从这三个方面的国内外研究状况进行分析总结。论文最后定义网络行为学的概念,并提出网络行为学的研究方向。

关键词 计算机网络 网络行为学 综述 测量

文章编号 1002-8331-(2004)27-0001-08 文献标识码 A 中图分类号 TP393

Research on Network Measurement & Behavior

Cheng Guang Gong Jian Ding Wei
(Department of Computer Science & Engineering,Southeast University,Nanjing 210096)
(Key Lab of Computer Network Technology Jiangsu Province,Nanjing 210096)

Abstract: Due to the development of Internet,it is very important for network management,plan,and development to understand the network behavior,so the network traffic measurement has been becoming the basic of network behavior research.According to the development character of Internet and the existing problem,the paper analyzes the research difficulty and focus of network behavior.Then it divides the network measurement and behavior into three research directions including traffic measurement,traffic behavior analysis,and end-to-end behavior analysis.Thirdly,the research status in china and others countries are analyzed and summarized.Finally,the concept of network behavior is defined,and the research direction of network behavior is provided.

Keywords: computer metwork,network behavior,summarize,measurement

1 引言

互联网是上亿台计算机互联成的全球性网络,虽然相关的组网与管理技术在不断地完善,但人们对它在局部和整体范围内所体现出的行为特征依然没有一个正确和完整的认识。掌握 Internet 的行为是网络规划、网络管理和网络安全、新网络协议和网络应用设计等诸多研究工作的重要前提,近年来对大规模互联网络行为的研究成为该领域被关注的热点目标。网络行为测量和分析是网络行为学研究的基础,通过测量分析可以掌握网络行为的基本特征(base-line),有助于寻找网络行为变化的规律,构造并验证网络行为的数学模型。所以,针对网络行为的测量与分析方法展开系统性的研究将对 Internet 行为学方面的研究取得理论突破具有重要的意义。

目前在网络测量和行为学方面研究的主要组织有 CAIDA、MERIT、ITA、OC3 和 IETF 等,主要工作是对网络设备的数据采集,从历史数据中掌握网络运行状态,以预知可能出现的问题,合理安排网络负载;同时在充分了解设备和连接的使用情况下,对未来的网络规划做出指导。国外有很多网络组织对其自身提供的网络服务进行一定程度的测量分析并相互进行合作,以期获得对网络状态的全面认识,如 CAIDA 进行的网络数据流测量和分析工作。目前的研究大多是根据某些特定应用的要求而进行的,目标比较狭小,没有形成较完整的体系,已经出现的一些理论模型距实际运用还有一定的差距。如加利福利亚大学用一些简单的算法由历史数据对网络的延时、吞吐量进行短期的预报以支持负载动态调整;波士顿大学通过对信道的可用带宽的测试数据进行动态服务器的选择。总体来说,在这一领域现有的研究是直接根据测量值进行一些网络负载的即时优化,没有对数据的长期分析和综合运用。

论文首先讨论互联网发展特点、网络管理的缺点和不足,分析网络行为学的研究难点和方向:网络测量、网络流量行为分析、端至端行为分析等,并分别就此对国内外研究状况进行

基金项目:国家自然科学基金项目(编号:90104031);国家 973 基础科学研究发展计划项目(编号:2003CB314803);东南大学基金项目(编号:9209002157)资助

作者简介:程光(1973-),博士,讲师,研究方向:网络测量和网络行为学。龚俭(1957-),教授,博导,研究方向:网络安全,网络体系结构。丁伟(1962-),教授,研究方向:网络管理,网络测量。

分析。最后文章提出网络行为学的概念。

2 网络发展特点及存在的问题

2.1 Internet 发展的特点

互联网自 20 世纪 80 年代末以来，一直以指数的速度膨胀，互联网具有以下特点。首先，TCP/IP 技术将多种不同的网络技术和管理域统一成一个整体，但 TCP/IP 技术只是将各种多样性统一起来，并没有统一它们的行为成分。

第二个特点是 Internet 规模庞大，2002 年 7 月的估计有 162,128,493 台联网主机[1]，Internet 规模过于庞大会产生两方面的问题：第一个问题是网络异质性规模很大，表现出特殊行为的少数计算机在 Internet 中可能也会有数千台，因此就需要考虑这些计算机的行为。异质性具体表现为以下几方面：(1)网络拓扑结构难以准确刻划，规模大导致拓扑的一致性比较难维持，因此任意时刻实际的网络拓扑很难准确了解；其次 Internet 拓扑结构由许多不同机构的网络拓扑决定，但不是所有的单位都愿意提供它们网络的拓扑信息，虽然目前有些拓扑结构发现算法，但实际的拓扑很难准确了解。(2)网络中不同链路的属性差异很大，如：一条路由上的不同链路可能是高速光纤主干，也可能是低速以太链路。(3)Internet 路由动态变化，变化幅度在几秒到几天之间，这造成拓扑结构的不稳定。在大规模网络拓扑中路由可能是非对称的，从计算机 A 到计算机 B 的路由不同于从 B 到 A 的路由，这也说明网络规模的扩大可能产生一些不可预料的问题。(4)网络协议不同。理论上 Internet 使用统一的一套协议，但实际上不同的协议实现可能有不同特征，如：Internet 传输层核心协议 TCP 就有多个不同版本，这些不同的 TCP 协议表现出不同的行为。(5)基于相同传输层协议的网络应用也具有区别，这些差异也会造成端至端传输性能的变化。网络规模庞大引起的第二个问题是比例问题，有些网络协议和机制在小规模网络中能很好地工作，但当网络规模大了数倍以后可能就不能正常运行，大规模网络意味着一些罕见的事情将会经常出现在网络的某处。另外由于规模庞大，依靠人工干预以维持网络的正常运行相当困难。

Internet 的第三个特点是增长速度快。根据 Internet 软件联盟(Internet Software Consortium)ISC 的调查报告，1981 年 8 月联网主机仅有 213 台，到 2002 年 7 月有 162,128,493 台联网计算机。图 1 是从 1981 年 8 月到 2002 年 7 月主机增加的曲线图，从图中可以看出，主机增长速度呈现指数增长的趋势，图中的指数曲线和实际增长趋势非常接近，其增长拟合公式为：

$$y=1327.9e^{0.2883xx} \quad (1)$$

式中 x 表示需要计算主机数的时间距离 1981 年 8 月相隔多少半年，y 表示主机数，如果按公式(1)的趋势增长，十年以后联网主机将达到 150 亿台，几乎平均 1 人有两台联网主机，因此人们不知道几年以后 Internet 规模有多大，也不知道什么时候它的增长速度会放慢。

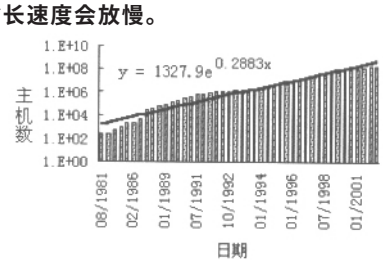

图 1 Internet 主机增长曲线图

网络中一些看似稳定不变的统计数据也可能随着时间的变化，如 1992 年 10 月在一个实验室中采集 FTP 流量的中间值是 4,500 字节，6 个月以后同样实验发现中间值仅为 2,100 个字节[2]，1998 年 3 月的实验得出的值是 10,900。另外，新应用会大大改变 Internet 的行为，如 WWW 应用自 1992 年出现，1993~1994 年每六个星期 WWW 流量翻一倍，近几年来 WWW 已经成为主要的应用，占总流量的 75%以上。

2.2 网络管理及其存在的问题

网络测量和行为学与网络管理密切相关，在讨论网络行为学之前，先研究网络管理的现状和存在的问题。网络管理分为 5 个功能域：失效管理、配置管理、安全管理、性能管理和计费管理。管理信息库 MIB 是对于通过网络管理协议可以访问信息的精确定义，MIB 的结构定义和内容是网络管理的核心。RFC1052 中定义用于 SNMP 和 CMIS/CMIP 的扩展 MIB；RFC1065 描述在管理 TCP/IP 网络 MIB 中可用的信息语法和类型；RFC1066 给出 TCP/IP 协议簇的第一个 MIB 版本，该版本目前成为 MIB-I，它精确地解释并定义监视和控制基于 TCP/IP 的互联网所需要的信息库；RFC1158 提出 TCP/IP 协议簇的第二种 MIB(MIB-II)，通过扩展定义在 MIB 中的对象集扩展定义在 MIB-I 中的信息库；RFC1271 定义的远程网络监视(RMON)对 SNMP 的增强，它定义的 MIB 补充 MIB-II，并且向网络管理人员提供有关互联网的关键信息。

利用 MIB-II，网络管理员可以获取单个设备本地信息。在具有一定数量设备的 LAN 中，这些设备均有一个 SNMP 代理。SNMP 管理者可以使用 MIB-II 获得出入每个设备的信息流量，但不易获知 LAN 上的整体流量。通常一个用于研究网络整体流量的设备被称为网络监视器、网络分析器或探测器。典型情况是监视器以监听方式在 LAN 上运行，监视 LAN 上的每一个包。监视器可以产生统计信息，包括错误统计和性能统计；监视器还可以存储全部或部分报文记录，并使用过滤器根据报文的类型或报文的其它特性以限制计数或捕获的报文个数。

为了在互联网环境中实现网络管理，每个子网需要一个监视器，这个监视器可以是一个独立设备，其唯一目的就是捕获并分析流量；也可以由具有其它任务的设备如路由器来执行。为了有效实现网络管理，监视器需要与一个中央网络管理站通信。下面分析 RMON MIB 的结构。

RMON MIB 分为 9 组：(1)统计：维护代理监视每一子网的基本使用和错误统计。(2)历史：记录从统计组可得到信息的周期性统计样本。(3)警报：允许管理控制台人员为 RMON 代理记录的任何计数或整数设置采样间隔和报警阀值。(4)主机：关于连接到子网上主机的各种流量计数。(5)最高 N 台主机：排序后的主机统计，该报告根据主机表中的一些参数生成列表，并从中选出最高 N 台主机。(6)矩阵：以矩阵形式显示出错和使用信息。(7)过滤：运行监测器观测与过滤器相匹配的数据包。(8)报文捕获：控制数据发往管理控制台的方式。(9)事件：关于由 RMON 代理产生的所有事件的表。

从以上网络管理机制分析来看，其存在以下几个问题：(1)虽然 RMON MIB 能够测量 LAN 内的流量信息，但随着网络应用和带宽的快速发展，RMON 不可能测量所有通过监测器的流量，即使测量也无法对所测量的流量实现在线实时分析。(2)由于 MIB 统计组中包含的流量测度偏少，难以满足各种应用需求。(3)由于资源的限制，不能将所有的历史数据资料永久保

留,RMON MIB 直接将最老的数据丢弃,这样超出容量空间的历史资料就被丢弃。(4)由于网络流量是由多种偶然因素构成,且不同的时间段变化很大,RMON MIB 简单定义阀值以判断流量行为的异常情况,这种固定的报警阀值无法满足网络流量的各种变化,如白天流量是晚上流量的 4 倍。(5)已有一些高级网管工具可帮助管理者预测未来流量行为、估计网络未来发展趋势,但预测方法只是简单采用线性预测模型或时间序列模型,没有根据网络流量产生机制建立网络流量行为模型。(6)随着网络实时多媒体应用的发展,如 IP 电话、视频会议等,这些应用的使用性能主要由单向延迟(delay)和丢包率(loss)等分布式测度来体现,但是网络管理采集流量的机制无法实现分布式测度值的分析。

2.3 网络行为学研究难点及其方向

由于网络技术和管理的多样性、规模庞大、增长迅速以及网络应用的变化等,所有这些因素对网络行为研究提出挑战,也使网络行为学从网络管理分支中分离出来,成为独立的研究方向。

研究网络行为的另一个难点是多种不同的外界和内部因素会引起网络的变化、发展,这些因素包括:首先,新的价格体系使得用户对流量类型和服务质量更加敏感。第二,网络路由器的分组调度算法将会由"先来先服务"转移到不同的链路稳定地共享资源。第三,目前 Internet 的组播大多是通过单播隧道传送,随着纯组播网的逐渐普及,链路传送组播的流量类型和数量将有很大的变化。第四,Internet 支持不同服务类型和服务质量的机制将产生不同的链接和性能。第五,网页缓冲将普遍存在,这将有助于缓解主干网的传输压力;最后是从流量比重看,出现新的应用有可能会取代 WWW 的地位,从而大大改变网络特征,如:多人协同网络游戏的应用。

综合前面网络发展的特点和网络管理技术中存在的问题,可以将网络测量及行为学分成三个具体研究发展方向:网络流量测量研究、单点网络流量统计行为研究和多点分布式网络行为研究,其中多点分布式网络行为研究主要包括:端至端网络性能研究和流量路由行为研究。这三个方向是统一整体,每个方向的进步都会给其它研究带来帮助,其中网络流量测量是其它研究基础和核心,下面将分别从网络测量、流量统计行为、端至端网络性能行为等三个方面来探讨目前国际上网络行为学相关领域的发展状况。

3 网络测量技术的发展

早在 1995 年,当 NSF 停止对整个 NSFnet 的资助以后,Internet 出现多元发展的趋势,试图完整地跟踪和监控 Internet 行为已经不可能。NSF 没有留下可以监测不同 ISP 之间的网络性能问题和安全事件的框架结构,同时 ISP 也没有重视收集或分析他们网络的数据,这样造成今天不但缺乏理想的测量和分析网络行为工具,同时也缺乏用于分析网络行为变化的测度数据。

尽管网络测量没有得到足够的重视,但学术界对其研究一直没有间断。因此现在有许多独立的网络端至端测量方法和相应的工具,这些测量方法主要是在端主机发送主动探测分组进入网络,然后记录分组返回的延迟。但是这种测量涉及到大量的难以独立建模的参数,同时测量结果的复杂性使得收集的数据难以比较或标准化。有些研究组织正试图发展技术和体系结构来支持标准化测量、性能评估、选择网络路由的可靠性等,但这些工作进行的很慢,还不能满足用户、研究者和 ISP 的需求。

网络测量可分为主动测量和被动测量两种方式,主动测量是通过主动产生流量直接测量网络的属性,可以使用直接分析的方法。但是主动测量流量会对被测网络流量的性能产生负面的影响,因此,主动测量系统开发需要仔细考虑对网络实际传输流量的影响。被动测量完全取决于被测网络中目前已有的流量,它的最大优点是在测量期间不影响被测网络的流量,但会引起测量、分析、存储等资源短缺的问题。论文将从抽样测量技术、测量测度和国际上著名的测量体系结构等几个方面来讨论网络测量问题。

3.1 主动测量技术

主动意味着测量过程中产生新的网络流量,主动流量可以引起网络部件的特殊响应(如:traceroute),也可以用于观测网络的性能(如:treno)。主动测量给网络增加潜在的荷载负担,特别是如果没有仔细设计使得该方法产生的流量最小,那么附加的流量会扰乱网络,影响分析结果。例如:如果为了测量在 IP 网络中瓶颈链路的带宽,定期地向网络发送巨大的 tcp 传输,由此产生的附加流量可能会引起 Heisenberg 效应,并使测量的网络吞吐量低于实际瓶颈链路的带宽。

其次,主动测量至少要多个网络部件以某种形式的参与,如:使用 ping 估计主机 A 到主机 B 的 RTT,需要主机 B 响应 ICMP ECHO 请求信息。不同机器之间主动测量的合作方式有被动合作和主动合作,如:响应 ICMP 请求、匿名 FTP 服务器允许主机和服务器之间实现吞吐量测量,这种合作称为被动合作。另一种合作方式是主动合作,需要具体的安排,如果要测量 A 至 B 之间单向延迟的对称性,从 B 到 A 和从 A 到 B 都需要进行延迟测量,因而 B 需要主动参与测量。

网络拓扑结构的变化需要使用主动测量技术,CAIDA[5]开发的 skitter 动态测量工具可用于动态发现和绘制全球 Internet 拓扑。另外主动测量技术可以探测网络的具体问题,如发现许多 Internet 端至端的延迟分布具有重尾特征。

3.2 被动测量技术

被动测量需要在网络中的一点收集流量信息,如使用路由器或交换机收集数据或者使用一个独立的设备被动地监测通过被测量网络链路的流量。被动测量的优点是可以完全取消附加流量和 Heisenberg 效应,但有些测度被动测量相当困难:如决定分组所采用的路由。当然如果用户关心的不是完整的网络路由,而是 AS 之间的路由,那么可以被动监测两个对等 BGP 之间的流量,因为流量包含全部的 AS 之间的路由信息。被动技术的另一个重要的问题是现在正在提出的要求确保隐私和安全的矛盾。

网络流量是由大小不一的分组构成的,收集到的数据可以进行各种流量分析,如:流量中各种应用的成分、分组的大小分布、分组的相隔到达时间、性能和路径长度等,这些能帮助设计下一代互联网设备和体系结构。网络运行者最感兴趣的是流量流矩阵:有多少流量从一个网络流向另一个网络,这个信息能有助于优化设计决定。不同的流量粒度矩阵有不同的用处,AS 粒度流量矩阵有助于优化拓扑结构;企业网或校园网的管理者为了解各部门之间流量交换的情况,可以建立系或工作组粒度的流矩阵;国家粒度的流量矩阵有助于了解各国的网络开放策略和国际商业前景,例如美国是世界 Internet 流量主要中转国,

71%的国际中转流量经过美国。

与荷载特性相关的一个重要特性是网络分组的大小。分组大小分布和到达时间的统计数据同设计网络路由交换设备相关，因为交换一个分组的代价有每个分组和每个字节的成分，因此，有关于典型 Internet 荷载的测度允许设计者优化路由器的硬件和软件体系结构。另一个重要的荷载分析是估计流量中各种协议类型的比重。

被动测量还有许多其它的应用，如：识别、刻划和跟踪网页缓冲和代理的优化配置；网络体系结构的安全性；拥塞控制算法的有效性；流量增长是由于用户数量的增加还是每个用户流量的增加；流行协议和应用使用的变化；新的技术和协议（如：组播和 IPv6）的渗透力和影响等等。这些被动测量的应用是 Internet 流量行为研究的重要内容。

有时为了能够从被动收集的数据中提取某些参数可能需要借助于主动测量。另外，被动测量是不应该丢失分组的，否则分析的数据将不精确。但随着流量速率的增加，保证不丢失数据变得几乎不可能，因此抽样网络流量测量技术成为解决这一矛盾的非常有效的方法和技术。

3.3 抽样测量技术

网络被动测量流量主要用于网络分析、规划和计费等，但由于网络带宽的高速发展，同时计算机测量、存储、分析资源的限制，难以直接测量全部的网络流量。一种解决方案是抽样测量部分流量，从统计角度获得对流量总体行为的认识。早在 1989 年，Paul 等人首先将抽样技术引入网络测量中。在 1993 年，Claffy 系统研究抽样技术在网络流量统计分析中的应用，分析系统抽样、简单随机抽样和分层抽样技术用于估计分组大小分布和分组到达时间分布[4]。在 1998 年 Cozzani 提出使用模式匹配技术抽取 ATM 位元。2000 年 Duffield 使用同样的抽样技术用于寻找网络中分组流的路径[5]。2001 年和 2002 年抽样测量技术出现新发展，Tanja 使用被动测量技术测量端至端 SLA 的单向延迟测度[6]。Duffield[7]和 Cristian 分别根据流大小建立不等概率模型。同时抽样测量技术也普遍应用于网络产品中，如 Cisco 的 Netflow 和 NetranMet 测量器。

2002 年 4 月，IETF 成立网络流量抽样测量工作组（PSAMP）[8]，PSAMP 的任务是建立被动报文抽样测量的标准，其标准需要能够可靠、详细、直接和实时支持已有和未来的应用，且在各种网卡、监测器等硬件设备中容易实现。因此 PSAMP 认为抽样测量需要具有以下条件：(1)提供分组级的测量；(2)测量过程的实时性；(3)测量数据处理的实时性，即：多个抽样测量点可以同时对同一个流或同一个报文进行处理。为了满足这些目标，抽样算法的选择非常重要。其次抽样测量的数据能够满足各种网络应用需求，如：流量工程、DoS 攻击检测、流量规划和计费、网络路由、端至端 SLA 等研究。Albert 讨论一些具体应用，如：拥塞控制、流量工程、容量规划、对等关系的管理等，认为抽样测量对高速流量分析的可行性、实时性具有重要作用。Trajectory Sampling 使用一个基于哈希的报文抽样模型来提供对网络行为实时和直接的观测。Albert 认为控制测量流量的负荷对于网络流量行为分析和网络性能行为分析是非常重要的，因此对于高速流量通过抽样机制控制输出的分析流量是非常有意义的。

从高速网络流量中测量一个流量子集主要有两种测量方法：时间驱动和事件驱动方法，这两种方法可以在各种粒度下随机和确定地选择模型，然后使用抽样样本估计总体流量数据集合的参数。当前研究的流量测度主要有报文长度分布和报文到达间隔时间分布；根据网络流矩阵研究网络应用和分布的地理位置分布。抽样估计不仅考虑抽样方法的类型，而且还考虑在一个子集内的抽样粒度范围和抽样子集。抽样方法的类型和粒度将对估计目标测度的置信区间具有影响，研究结果表明事件驱动的抽样（即每 n 个报文抽样一个报文）比时间驱动的抽样方案更有效（即每 n 毫秒抽样一个报文）。

在简单随机抽样中，总体中每个样本被抽中的机会都是相等的，称为等概率抽样，与此相反的抽样就是不等概率抽样。不等概率抽样是指在抽样过程中，总体中每个样本被抽中的概率值 P_i，这些概率值 P_i 可能是不相等的，也就是说，不同的样本被抽样的可能性是不相等的，有的概率大些，而有的概率小些。由于网络流量中流是服从重尾分布的，巨流占总流量的比率大，但是巨流的数目少，细流占总流量的比重小，但细流的数目大，如果将所有流按等概率抽样，那么一个巨流被抽样或不被抽样都会严重影响网络流量的统计估计。Duffield 根据流的大小建立抽样模型，巨流抽样概率大，细流抽样概率小，抽样模型定义为：

$$p_z(x)=\min\{1,x/z\} \tag{2}$$

式中 x 代表流的大小，z 是规定的阀值，如果流的长度大于 z，则抽样的概率为 1，否则流的抽样概率和流的长度成正比。

由于网络资源的限制，测量全部网络流量难以实现，目前国内外还没有人专门从网络资源限制角度研究流量测量。当从网络主干上测量分组，最终到达分析服务器涉及到的网络资源限制有：(1)测量器 CPU 速度，当分组还没有到达应用程序能够捕捉的时刻就可能已经被丢失；(2)测量器的内存，即在流量数据被传送到分析器之前需要考虑内存的存放空间问题；(3)从测量器到分析器之间的网络传输带宽；(4) 测量数据的硬盘存储空间，由于网络流量很大，不可能将所有的测量数据完整地、永久地保留下来；(5)另外还有分析器计算资源的限制，控制数据处理和查询时间的限制等。这些资源限制都可能会造成测量数据的丢失而必须采用抽样或粗粒度聚类的方法进行处理。粗粒度聚类能够从整体上无误差地对网络行为作统计上的认识，但缺乏对网络行为细节上描述的数据。抽样技术是根据统计数学理论提取部分细节数据，根据这些细节数据能够对网络总体流量行为作出在规定概率范围内的描述。因此可以规定抽样测量精度，提取部分细粒度数据，以实现既能反映流量总体信息，又能获得流量详细信息。

3.4 网络行为测度

Internet 取得巨大成功是开放性的一个胜利，但是 Internet 测量的开放性还不能满足网络快速增长的需求。如何测量沿着网络某个确定路径的吞吐量和延迟等基本问题都缺乏标准测量框架，于是对 SLA 的监控仍然是一个困难问题。在这种情况下，IETF 建立的 IP 性能测度工作组（IPPM）[9]发展了一套标准的测度，用以刻画网络数据传送服务的质量、性能和可靠性。设计的这些测度能被网络运行者、终端用户或独立的测试组使用，这些测度并不是好坏判断的标志，而是作为一种公正的性能数据测量。

“测度”定义 Internet 组成部分的不同属性[10]。“在 Internet 中有一些关于 Internet 性能和可靠性的参量，这些参量的值笔者是希望知道的，当其中一个参量被详细说明后，称这个参量

为一个测度”。

IPPM 目前已经制定 5 个 RFC：(1)IP 性能测度框架(RFC 2330)，它为 IETF IPPM 发展的具体测度定义一个通用的框架结构，包括几个测度的标准，定义相关的 Internet 术语、测度和测量方法的基本概念，以及处理定义合理测度和方法相关的几个问题。(2)测量连接性的 IPPM 测度(RFC 2678)，定义 Internet 主机对之间连接性的一组测度，包括：①瞬间单向连接性，定义在某一时刻一个方向的连接性；②瞬时双向连接性，定义在某一时刻两个方向的连接性；③相隔时间连接性，定义两台主机之间一段时间内两个方向的连接性。(3)单向延迟测度(RFC 2679)，从单个、采样和统计三个方面来分析该测度。(4)单向分组丢失测度(RFC 2680)，也是从单个、采样和统计三个方面来分析该测度。(5)往返延迟测度(RFC 2681)，采用同 RFC 2679 和 2680 同样的定义方法。

同时 IPPM 还制定几个测度草案：(1)瞬间分组延迟变化测度，定义通过 Internet 路径分组延迟变化的一个测度，这个测度是基于连续分组单向延迟差异的统计数据，它在两台主机之间时钟同步或不同步时都有效，在不同步时这个测度允许评估二者差异。(2)经验上定义的块传输能力测度的框架，块传输能力(BTC)是测量网络通过单个 TCP 链路传输大数据量的能力，BTC 的直觉定义是期望得到一条理想 TCP 链路的平均数据速率。然而 IETF 标准允许许多拥塞控制算法，传输算法的多样性造成了标准化测度的困难。这篇草案为标准化多种 BTC 测度定义一个框架结构。(3)单向丢失模式采样测度。丢失模式或丢失分布是决定用户观察性能的关键参数，对于相同的丢失率，两个不同丢失的分布可能产生对性能的感觉极其不同。丢失模式的影响对使用 TCP 协议的非实时应用相当重要，大量研究表明突发性丢失对声音和视频应用有较大影响。为了获得分组丢失模式，这篇草案中提出两个衍生测度：丢失距离和丢失周期。丢失周期测度测量丢失的频率和长度(突发性)，丢失距离测度测量丢失周期之间的间距。(4)周期流量网络性能测量。这篇草案讨论关于 IP 网络应用层性能测量的概念，最初的动机是交互式周期流(如，基于 IP 的多媒体会议)的 QoS，但应用层测量的思想也许有更广泛的应用。(5)单向延迟测量协议，建议使用 OWDP 服务器，它可以使对单向延迟的测量就象使用 ping 测量往返时间一样普通，另外 OWDP 还提供必要的安全功能和支持小测试分组。

目前 IPPM 工作组的工作重点主要是网络性能的平均测度和测度标准化，在评价网络流量统计测度和路由测度等方面工作很少。在路由测度方面，CAIDA 定义一套评价两 IP 之间网络长度的测度：通过 IP 的数目、通过路由器的个数、通过 AS 的数目、通过的 ISP 的数目、地理距离(公里)、RTT 中值(ms)等。在统计测度方面目前有 RMON MIB 中定义的流量报文总数、字节总数、平均报文长度等统计信息。

3.5 网络分布式测量平台

历史上一直没很好地对 Internet 进行测量，随着网络规模的扩大，这个问题变得越来越严重。国际组织试图通过发展网络分布式测量平台来解决网络测量问题，这些测量平台由分布在网络不同位置的专有工作站上的测量平台构成，这些工作站相互协作、相互交换测试流量来测量网络路径和网络云的属性。

分布式测量平台有以下几个用途：(1)通过从不同点探测网络的特性，诊断网络内部的性能问题；(2)测量广大范围网络路径的属性，来研究网络行为和进化；(3)测量从 ISP 网络的一个端点到网络另一端点的探测流量，评价不同 ISP 的性能；(4)综合 ISP 性能的评估，为 ISP 提出合理建议，促使他们优化自己的网络。

测量网络端至端性能需要分析一系列不同 ISP 的行为，没有一个 ISP 能完全依靠自己的网络完整地监控这一行为，因此，体系结构的各部分需要相互合作共同管理。网络的异质性决定测量体系结构不能只属于一个单位或组织，如果希望得到广泛地部署，ISP 必须能自己决定是否配置测量平台。这要求：(1)体系结构要有利于在 ISP 的网络或对等网络中监控和调试性能问题，使 ISP 们能更好地操纵自己的网络；(2)出于竞争的需要，ISP 和用户之间的‘服务水平约定’可提供在线测量，使用户能监测他们是否真正得到协议中规定的服务；(3)ISP 可以限制使用这个平台来测量自己的网络，而且还可以利用这个平台有效地测量竞争者的网络性能。ISP 可以使用‘测量政策’来限制公众利用这个平台测量流量的类型，测量政策指谁能用平台进行什么样的测量。

CAIDA 列出一些用于监测 Internet 流量的测量项目，其中有代表性的包括：(1)Paxson 在 Internet 的 50 个站点上部署了网络探测幽灵，详细研究了路由、延迟、丢包和 TCP 的动态性；(2)IPMA 研制了一个路由协议测量器，通过广域部署这些测量器，以获取网络路由的动态行为[11]；(3)Pinger 监测高能物理研究所(HEP)的站点，通过在一个固定的时间表发送一系列的 ICMP 请求报文以得到网络性能信息[12]；(4)WAND 进行单向延迟和丢包测度测量，使用 GPS 系统同步时钟，WAND 能被动测量 ATM 位元，另外 WAND 还开发了可以在 SONET 接口上获取 IP 分组的以太网卡；(5)RIPE 是 IETF IPPM 的单向延迟和丢包测度的一个应用，它的使用范围是同 RIPE 联系的欧洲网络；(6)NIMI 是一个网络测量通用平台，NIMI 原型可进行测量单向延迟、丢包和路由；(7)CAIDA 开发的 Skitter 工具测量从一个源点到数千个目标点的 IP 路由和 RTT，通过主动探测以获取和跟踪全球 Internet 拓扑。

4 流量统计行为的研究

到目前为止，Internet 主要经历了三个发展阶段：NSFNET 时代(1988~1995)、Web 时代(1995~2000)和现在的高速网络时代(2000 年以来)。特别是进入高速网络时代，计算机、路由器、光纤等硬件技术的发展，用户和应用的增加，以及网络拓扑结构的复杂化，使得描述网络应用和流量特性更加困难。为了进一步优化网络配置、掌握新的协议对目前已有网络流量的影响，需要对当前网络主干流量特性准确了解。

在 1988 年至 1995 年由 NSFNET 负责 Internet 主干网络期间，NSFNET 主干流量数据统计特性是研究重点。在 1993 年，K.Claffy[13]、K.D.Frazer 和 H.Heimlich 详细研究 Web 出现之前的主要应用流量的统计分布，报告 FTP、SMTP、NNTP、DNS 和 Telnet 等主要应用的统计数据。在此期间，还有一些研究广域网流量特性的工作，Caceres 1989 年研究 Bell 实验室的流量情况，表明 TCP 占了 90%以上的流量，其中 SMTP 占了约 50%的流量。Paxson 研究 SMTP、FTP 和 X11 等应用流量的长期发展情况。

在 NSFNET 结束后，Web 应用成为主要流量，占总流量的近 80%，网络行为特性有了很大的变化。J.Apisdorf(1997 年)、

K Claffy(1998 年)、K Thompson(1997 年)[14]等人使用 OC3MON 监测器从两个监测点分析 Internet MCI OC3 主干网络的流量特性。2000 年 McCreary 利用 NAI 项目从 AIX 主干网络测量的流量数据和 CAIDA 的 CoralReef 流量分析软件，对 IP 流量统计特性进行研究。NLANR/MOAT 的三维数据项目存储和分析由 Coral 监测器从 HPC 站点收集的数据，三维分别是指：数据的测量点、分析方法和数据收集的日期，MOAT 将数据放在三维数据结构的 Web 服务器中，用户可以查询任何维数据。

进入了 2000 年，由于高速网络的发展，大多 Internet 的主干网络升级到 OC-48(2.5Gbps)，有的升级到 OC-192，局域网也普遍使用高速千兆以太网技术。由于带宽的发展，出现新的多媒体应用，同时 FTP 应用又有新的发展，Web 应用在网络流量中的比重逐渐下降，回到 1995 年的水平，因此，网络流量行为也有了变化。同时，由于在 NSFNET 时代和 Web 时代，网络带宽较低，全报文测量器能够满足监测需求，而在高速网络时代，由于处理器、存储空间等限制，难以实现高速流量长期、实时地测量处理和存储，因此测量和分析中需要引入抽样技术。

根据采集网络流量时间粒度的不同，可分为细时间粒度的流量采集和粗时间粒度的流量采集。细时间粒度采集持续时间较短，一般为几个小时到几个天，但时间粒度在几百毫秒到几秒之间；而粗时间粒度采集持续时间较长，一般为几个月甚至几年。细时间粒度采集的流量数据主要是用于网络流量特性的研究，而粗时间粒度采集的流量数据主要是用于长期流量预测研究。另外流量特性研究根据研究流量数据的粒度不同可分为分组粒度流量特性研究和流粒度的流量特性研究。流量数据采集技术已在前面进行了评述，这里主要讨论流量分析的研究，包括网络流量以自相似为基础的队列模型和网络流量预测建模等。

4.1 网络流量自相似模型

为了能正确地设计和应用计算机网络和网络服务，理解网络流量的性质是相当重要的。计算机网络是基于分组交换的，这与传统电路交换网络具有本质区别，计算机网络主要传输的是数据流量和多媒体流量，而传统的电信网络传输的主要是语音流量。在电信网络中的网络设计和系统设计采用泊松过程或马尔可夫过程，但在计算机网络中泊松过程和马尔可夫过程难以模拟实际流量，主要原因是这些模型是无记忆、短相关的，而实际计算机网络流量则表现为长相关性。

然而 1994 年 Leland 等人对从 1989 到 1992 年测量的以太网数据进行分析时发现以太网 LAN 流量具有自相似性质[15]；其后 1995 年 V.Paxson 和 S.Floyd 分析从 1989 年到 1995 年之间 24 种不同广域网的流量数据发现 TCP 分组队列符合自相似规律；1995 年波士顿大学的 Mark E.Crovella 和 Azer Bestavros 通过对一个广域网的 WWW 日志文件的分析发现 WWW 流量具有自相似性；这三篇重要的学术论文奠定了网络流量自相似特性的理论基础。[16]列出 1996 年之前国外网络流量研究领域的数据分析、统计推理、数学建模、队列和性能分析等研究论文。国外掀起自相似网络流量的研究热潮，主要领域有：(1)进一步验证在不同网络环境中流量符合自相似规律；(2)研究网络流量自相似的建模方法；(3)从应用和协议角度解释网络流量具有自相似性的原因；(4)网络流量自相似特性对网络性能的影响。

由于网络流量的时间尺度变化范围很大，实测的网络流量不仅具有长相关性，同时还具有短相关性，因此目前很多研究采用多分形模型流量过程。由于多分形过程本身具有多比例特性，采用多分形模型描述网络流量时，不但能表现网络流量在大时间尺度的突发性，还能表现单分形模型不能表现的小时间尺度下的突发性。由于小波变换具有比例特性，因此小波可以用来分析网络流量多分形特性。

4.2 网络流量预测模型

预测未来网络流量行为可以监控和为未来到达的流量分配缓存和带宽。目前常用的预测方法有几种：基于均值的方法、基于中间值的方法和基于时间序列的方法，以及三种方法的综合。

文献[17]有效地利用时间序列模型来预测 Internet 流量。有许多因素对 Internet 流量有影响，大多数因素是不能测量和识别的。为了预计将来可能的流量，最好的方法是对以前观察的流量进行分析。流量是随时间变化的，应该可以采用时间序列模型来描述；由于 Internet 流量是非静态的，因此可以采用非静态的自回归滑动平均模型 ARIMA。

ARIMA 模型是对自回归模型(AR)、滑动平均模型(MA)和自回归滑动平均模型(ARMA)的发展。阶数 p 的自回归模型 $AR(p)$ 是由 p 个历史数据和一个白噪声构成的时间序列模型；阶数 q 的滑动平均模型 $MA(q)$ 是具有 q 个历史白噪声和一个随机的白噪声的时间序列模型；阶数为 (p,q) 的 ARMA 模型是由 p 个历史数据和 q 个历史白噪声的时间序列模型。ARMA 的优点是许多静态的时间序列问题可以建模成 p 值和 q 值为 0、1 或 2 的 ARMA 模型。

Nancy 和 George 成功地利用 ARIMA$(p,d,q)\times(P,D,Q)s$ 季节模型对 NSFNET 主干网络流量进行长期的预测，ARIMA 模型还可以用来预测 NSFNET 一个节点的流量。Sabyasachi 将时间序列模型用于网络流量预测，并建立校园网和以太网的预测模型，说明季节性 ARIMA 模型满足大多情况下的网络流量预测。[18]使用时间序列分析技术对 Web 流量建模和预测，建立精确的季节 ARIMA 模型预测客户机请求特性：请求数目和查询文档的大小，这个预测结果能作为判断规模大小的依据。

然而流量增长变化的有些趋势也可能限制 ARIMA 模型的使用。ARIMA 所有的预测完全是取决于以前的数据，这些模型不考虑同这些数据相关的一些外部因素的变化。来自于商业用户的行为和来自于大学用户的行为极其不同，另外新的应用、新的协议、新的技术、新的政府规章或国家经济的变化，所有这些都可能给网络流量造成很大的影响，这些方面使得季节 ARIMA 模型难以预测，因此需要考虑新的网络流量预测模型。

5 网络端至端性能行为的研究

网络端至端性能监测的关键在于统一监测内容，使网络的性能状态不仅纵向可比，而且具有横向的可比性，才能形成对网络使用状况的正确评价。但是，由目前的情况看来，这方面的研究还未形成统一的标准。IETF 的 IPPM 工作组定义了一系列描述网络性能的参量建议，如 one-way delay、one-way packet loss、round-trip delay 等，明确了测量对象的参数、单位、定义和方法。ANS 根据 IPPM 定义的测度建立 Current Surveyor Network Map，发送带时标的测试报文对美国 38 个主节点间的 one-way delay 和 packet loss 进行测量，这似乎是最早的一个根据已成形的定义对网络性能进行监测的网络状况报告。

目前的测量端至端性能测量主要使用主动测量方法,但会影响网络流量。被动测量由于在高速网络环境中容易丢失信息,且难以监测端至端网络性能,因此在端至端性能测量中很少采用。

5.1 网络端至端 SLA 研究

服务系统一般是由多个服务提供商提供,所以为了给用户提供端至端服务质量保证,就要求各服务提供商都能保证其相应部分的可用性和性能。用户不关心提供服务的各个组成部分,只关心整体效果。对 QoS 的期望使用户要求同他的服务提供商协商 QoS 级别,这通过 SLA(service level agreements)[19]来实现。SLA 是服务提供商和用户之间为指定期望的 QoS 级别而签订的一个服务期望行为和 QoS 参数的合同。

SLA 是服务提供商和用户之间的协议,用户希望 IP 服务提供商向他们提供与真实业务相关的 SLA,客户需要了解服务提供商是否能保证网络、服务和应用的性能和可靠性。由服务提供商实时地、不间断监测来验证 SLA。对于企业和服务提供商来说,主动的 SLA 监测已经成为购买和销售高附加值网络服务的关键。客户需要实时知道服务是否达到要求,如果服务提供商的服务能得到测量,并得到保证,那么用户将愿意花更多的费用购买基于高质量保证的高端服务。SLA 规定客户需要和预期的服务品质,而服务提供商要同意保证其性能水平。但除非服务提供商用真正的令人信服的数据来支持他们的客户之间的 SLA,否则含糊其词的 SLA 不过是一些不明确的保证,用户是难以接受的。服务提供商为了能有效地参与竞争,保证高利润的业务,需要将工作重点放在网络和服务的质量、性能和可靠性等的监测上,建立 SLA,并按新的预期值和可信度水平来核实它们。因此,服务提供商建立可信的并能够被监测的 SLA 体系,将会为他们带来更先进、价格更高的服务,而且还是一个直接关系涉及到服务提供商生存问题的话题。

传统的 SLA 管理工具要么不能收集性能参数,要么即使能够收集性能参数,也难协同参数之间的关系,来确定服务是否达到 SLA 协议所规定的质量水平。服务提供商目前的网络管理解决方案,不能实现端至端性能管理,因此,SLA 促使服务提供商从传统的基于 SNMP 的解决方案转向 SLA 管理工具。目前市场上有一些服务等级跟踪和报告的工具,这些工具正从传统的网络和系统管理转向服务管理,但管理系统真正能管理 ISP 的 SLA 还需要一定时间。这些工具包括:Cisco 公司的 CiscoWorks2000 SMS 工具,Response Networks 的 Pulsar xSP,Info Vista 公司的 Info Vista 系统,INS 的 VitalSuite,ConCord Communications 的 Network Health Reporter 和 Agilent Technologies 的 Firehunter 等。其中有的产品仅能监控网络服务,有的能监控终端用户的应用程序,但这些工具基本都是采用轮询 MIB 数据库和使用简单的 ping 工具,且只能测量低速小型网络的 SLA。

在 SLA 管理研究方面:Bhoj 提出基于 Web 的 SLA 的管理框架,证明服务提供商能给用户提供多种服务的 SLA 监控能力,包括 E-mail 和网络访问服务等。Park 支持使用 SLA 概念进行 QoS 管理,Park 提出将微观经济领域的效用模型用于管理和控制多媒体 Internet 服务,这种模型为多媒体 Internet 服务和 SLA 管理提供一个计算可行的解决方案。Mun 考虑大规模网络结点太多、流量太大,全部测量端至端 SLA 负担过重,因此,提出基于监测的聚合和细化算法(ARM)以减少信息交换量。

进入 90 年代以来,随着计算机互联网络建设在我国的兴起,国内在网络管理研究开发方面已做过不少一些工作,但通常主要是基于 SNMP 的网络管理方面,在基于 SLA 的网络管理方面目前的研究工作主要集中在 QoS 的研究,还没有进行 SLA 管理工具的开发工作。而且由于条件的限制这些工作都是面向 10M 以下低带宽的 IP 网络的。

5.2 单向延迟测度研究

单向延迟测度是网络性能评估中的一个重要测度。大多数网络应用,特别是多媒体应用受到单向延迟的限制,一旦报文延迟超过规定的最大阀值,则可能会导致网络应用不能正常使用。传统延迟测量采用往返延迟(RTT),但是网络存在以下几方面因素决定往返延迟测度难以正确评价网络性能:(1)在 Internet 中,从源 IP 到宿 IP 和从宿 IP 到源 IP 的可能是非对称路由,这导致往返延迟实际是两个不同路由单向延迟的累加和。(2)即使从源 IP 到宿 IP 的往返路由是相同路径,但可能会由于非对称队列造成性能差异。(3)网络应用的性能主要是由一个方向的性能决定的,如 TCP 协议的 FTP 应用在一个方向传输大量数据文件,而相反方向仅需要传输少量的应答流量。(4)基于 QoS 的网络,在不同的方向提供不同的 QoS 保证,需要独立测量不同方向的报文延迟。

Claffy 等人修改 ping 功能[20],修改后的 ping 能在 ICMP 请求报文中记录通过源主机的时戳 T_1 和宿主机的时戳 T_2,ICMP 应答报文中记录通过宿主机时间 T_3 和源主机时间 T_4,因此从源 IP 到宿 IP 单向延迟 $SDdelay$,和宿 IP 到源 IP 的单向延迟 $DSdelay$ 为公式(3)、(4):

$$SDdelay=T_2-T_1 \tag{3}$$

$$DSdelay=T_4-T_3 \tag{4}$$

他们的实验表明,往返延迟仅能测量估计近似的单向延迟,非对称路由和网络资源的动态变化会引起网络单向延迟的非对称性。

RFC2679 定义类型 P 报文在 T 时刻的从源 IP 到宿 IP 的传输延迟 dT 为:在 T 时刻源 IP 发送 P 类型报文的第一个比特,而宿 IP 在 $T+dT$ 时刻收到该报文的最后一个比特。RFC2679 认为实际测量单向延迟时需要考虑以下几个问题:(1)实际的测量值应该是正数,因此测量的结果为负数或者是 0 可以忽略不计,但是如果要研究一个流的单向延迟分布,则需要考虑负值和 0 值。(2)由于报文从源 IP 到宿 IP 的单向延迟范围一般在 0.1 毫秒到 10 毫秒之间,因此需要保证源主机和宿主机之间的时钟同步。GPS 全球定位系统能保证精度在 0.01 毫秒;NTP 时钟同步系统仅能保证时钟同步在 10 毫秒以内,其精度取决于 NTP 客户机和 NTP 服务器所处的网络环境。时钟的基本原理是时钟的精度有限,且随着时间的变化会出现时钟漂移现象,V.Paxson 详细讨论了基于延迟测量的时钟行为。(3)需要定义一个单向延迟的最大阀值,超出阀值则认为报文丢失,Mahdavi 和 Paxson 定义上限值为 255 秒。(4)如果报文重复到达,则使用第一个到达的报文计算单向延迟。(5)如果报文被分段,或者出现其它情况,则认为报文丢失。

6 结论

网络行为学是研究网络发展、进化规律的科学,网络行为学研究可以将整个研究网络看成是一个人体或人类社会,而网

络路由器和主机看成人类社会中的个体或人体的器官,从社会学或医学角度系统地研究网络行为,以便发现网络发展规律和未来发展方向,进而迅速发现和处理网络行为异常情况,以维护网络的健康运行。研究网络行为学涉及到网络流量行为、网络硬件设备行为、网络用户行为、网络管理行为、网络安全行为、社会经济行为和政府宏观调控行为等方面,所有这些行为因素的变化和发展都会对网络行为产生正面或负面的影响。由于影响网络的因素很多,目前主要是研究网络流量行为变化而反映的网络行为学行为规律,其研究手段主要是以主动方式或被动方式测量通过测量网络的报文,采用概率论、数理统计等数学工具分析测量的网络流量,从而发现一些规律性网络行为特征。下面给出基于网络流量行为的网络行为学定义。

定义 1 网络行为学:是根据事先定义的网络测度,测量主动流量或被动流量通过被研究网络的测度值,进而研究相应测度值的变化规律。

定义 1 将网络行为学分为三个研究方向:测度定义研究、网络流量测量研究和行为测度规律研究,测度行为规律研究可以分为网络流量行为、网络性能行为和网络路由行为,其中网络流量行为指通过网络测量点的流量特性,网络性能行为是指网络端至端的流量特性,而路由行为是指报文从起点到终点所经过的各个路由。传统上,网络流量行为是用简单测量通过测点的全部网络流量来描述,但是随着网络带宽的不断增加,网络应用的更加复杂,同时流量测量器和流量分析器资源的限制,全流量测量、分析已经是不可能。网络性能行为和路由行为则主要是通过主动流量测量端至端性能来描述。主动测量会影响网络性能,同时由于差分服务,使得不同类型的流量会有不同的传输性能,因此主动流量的测量结果并不能真正反映网络的真实性能。从前面综述来看,网络行为学研究还处于初步阶段,还没有形成一个系统的理论体系。

目前网络行为学发展存在以下问题:(1)没有提出一个既能进行网络流量行为研究,又能进行网络性能行为研究的测量体系结构模型。由于没有一个合理的测量模型,在路由器中无法测量到基于网络行为的管理所需要的流量数据。(2)网络行为学的概念缺乏系统性的定义,研究内容缺乏统一,从各方面对网络行为学的研究还都是简单片面的。(3)在统计分析领域,缺乏对测量流量进行实时统计分析处理的数学理论和模型,也没有提出对大规模网络流量的建模模型,目前看来自相似流量模型对于高速大规模宏观网络模型也不适合。(4)从网络性能行为和路由行为研究来看,主要还是基于主动测量方式,但主动方式会影响网络性能,同时由于差分服务将会普及,测试流量并不能真正反映网络性能的真实情况。

进一步进行网络测量与行为学研究需要依托 CERNET 等大规模网络环境,通过对高速网络环境中流量抽样理论与方法和大规模可扩展的协同测量体系结构的研究与实践,深入探讨网络行为数据工程问题,分析并寻找大规模网络流量的统计特征,从而发现流量宏观统计运行规律并建立网络流量行为统计模型。采用新的方法研究网络流量行为的运行机制,结合网络流量非平稳的特点,研究流量长期行为和短期行为预测的数学模型,从而推动大规模网络流量行为测量和分析的基础理论研究。(收稿日期:2004 年 5 月)

参考文献

1.Number of Hosts advertised in the DNS.Internet Domain Survey,http://www.isc.org/ds/WWW-200207/index.html,2002-07

2.Walter Willinger,Vern Paxson.Where Mathematics Meets the Internet[J].Notices of the AMS,1998;45(8):961~970

3.The CAIDA Web Site.http://www.caida.org/

4.K Claffy,G Polyzos,H Braun.Application of Sampling Methodologies to Network Traffic Characterization[C].In:Proceedings of ACM SIGCOMM '93,1993-05

5.Duffield N G,Grossglauser M.Trajectory sampling for direct traffic observation,Networking[J].IEEE/ACM Transactions on,2001;9(3):280~292

6.Tanja Zseby:Deployment of Sampling Methods for SLA Validation with Non-Intrusive Measurements[C].In:Proceedings of Passive and Active Measurement Workshop(PAM 2002),Fort Collins,CO,USA,2002:25~26

7.N G Duffield,C Lund,M Thorup.Charging from sampled network usage[C].In:ACM SIGCOMM Internet Measurement Workshop 2001,San Francisco,CA,2001

8.Packet Sampling(psamp).http://www.ietf.org/html.charters/psamp-charter.html,2002-12

9.IP Performance Metrics(ippm).http://www.ietf.org/html.charters/ippm-charter.html

10.V Paxson.Towards a Framework for Defining Internet Performance Metrics[C].In:Proceedings of INET 96,1996

11.C Labovitz et al.The Internet Performance and Analysis Project.http://www.merit.edu/ipma/

12.L Cottrell.PingER Tools.http://www.slac.stanford.edu/xorg/icfa/ntf/tool.html,1998-05

13.K Claffy,G C Polyzos,H W Braun.Traffic Characteristics of The T1 Nsfnet Backbone[C].In:proceedings IEEE INFOCOM'93,San Francisco,California,1993:885~892

14.K Thompson,G Miller,R Wilder.Wide Area Internet Traffic Patterns and Characteristics[J].IEEE Network,1997;11(6):10~23

15.W E Leland,M S Taqqu,W Willinger et al.On the Self-Similar Nature of Ethernet Traffic[J].IEEE/ACM Transaction on Networking,1994:2

16.Walter Willinger,Murad S Taqqu,Ashok Erramilli.A Bibliographical Guide to Self-Similar Traffic and Performance Modeling for Modeling for Modern High-Speed Networks,Stochastic Networks:Theory and Applications[M].Oxford University Press,1996:339~366

17.N Groschwitz,G Polyzos.A Time Series Model of Long-term Traffic on the NSFnet Backbone[C].In:Proceedings of the IEEE International Conference on Communications(ICC'94),1994-05

18.Jean chrysostome Bolot,Philipp Hoschka.Performance Engineering of the Web Wide Web:Application to Dimensioning and Cache Design[C].In:Fifth International World Wide Web Conference,1996

19.Service-Level Management,Definition and Overview.http://www.iec.org/online/tutorials/service_level/index.html,2002-12

20.K C Claffy,G C Polyzos,H W Braun.Measurement considerations for assessing unidirectional latencies[J].Internetworking:Research and Experience,1993;4(3):121~132