

网络测量与分析研究综述

陈晓霞 任勇毛 李俊 (中国科学院 计算机网络信息中心 北京 100190)

张潇丹 (中国科学院 计算技术研究所 北京 100190)

摘要: 网络测量与分析是人们认识、了解、更好地利用网络的重要手段,为互联网的科学管理和有效控制及互联网的发展与利用提供科学依据。介绍了网络测量与分析的主要研究内容,以及国内外相关领域的研究现状,并对网络测量的关键技术进行了分析,同时给出了网络测量的常用工具的性能比较。最后展示了中美俄环球科教网络 GLORIAD 的网络测量的应用案例。

关键词: 网络测量与分析;互联网实验平台;可用带宽;网络性能;科研网络

A Survey on Network Measurement and Analysis

CHEN Xiao-Xia¹, REN Yong-Mao¹, LI Jun¹, ZHANG Xiao-Dan²

(1.Computer Network Information Center, Chinese Academy of Sciences, Beijing 100190, China; 2.Institute of Computing Technowgy, Chinese Academy of Sciences, Beijing 100190, China)

Abstract: Network measurement and analysis provides vital means for people to know, understand and better utilize the internet. Also it offers the scientific basis for scientific management, effective control and development and utilization of internet. In this paper, the chief research of the network measurement and analysis and its research situation in the world are introduced. And the key techniques are also explored. Commonly used measurement and analysis tools are presented. Lastly, the application examples of the measurement of Global Ring Network for Advanced Applications Development (GLORIAD) are given.

Keywords: network measurement and analysis; internet experiment platform; available bandwidth ;network performance ; research network

据中国互联网信息中心《第 24 次中国互联网络发展状况统计报告》,截止 09 年 6 月 30 日,中国网民已达 3.38 亿,稳居世界第一。伴随着网络信息技术的快速发展,互联网的触角已经延伸到社会经济生活的方方面面,网络已经并正深刻地影响和改变着世界。在因特网发展早期,网络测量一直没有得到足够的重视。近年来,人们意识到有效的网络测量与分析有助于更好地理解网络结构及行为,网络测量开始得到国家政府、企业及学术界的普遍关注。然而随着因特网规模的迅速扩大、网络传输内容的多样化、恶意攻击行为的增多,以及对互联网的流量特征、性能特

征、可靠性及安全性特征及网络行为模型等缺乏精确描述,也严重影响到因特网发展及更有效的应用,这一切使得网络测量成为迫切需要解决的问题。

1 网络测量的基本描述

1.1 网络测量的性能指标

网络测量是按照一定的方法和技术,利用软件或硬件工具来测试网络的运行状态、表征网络特性的一系列活动的总和。

性能测量是网络测量领域中研究最多的部分,标准化组织 IETF(Internet Engineering Task Force,

基金项目:国家高技术研究发展计划(863)(2007AA01Z214)

收稿时间:2009-10-21;收到修改稿时间:2009-11-17

互联网工程任务组)^[1]和ITU-T(International Telecommunication Union, 国际电信联盟)^[2]分别对此作了定义。在参数的定义方法上,前者更注重对业务的支持,后者着重于网络描述IP包的时间传送效果;在参数的测量框架上,前者的研究较后者更为深入。二者在参数的选定上尽管有不同的表述方法,但基本含义是一致的。比如IETF的连通性参数对应ITU-T的服务可用性参数;IETF用选取函数代替ITU-T三种时延变化的定义方法。

不同的性能参数反映网络不同方面的性能。业务不同,对网络性能参数的选取也就不同。常用的分析指标有:单向时延、往返时间、丢包率、可连接性、吞吐率、瞬时分组延时差、大批传输容量、带宽利用率、瓶颈带宽估计、周期流网络性能。其中,最基本的三大指标为:时延、丢包率和带宽。

1.2 网络测量的分类

网络测量的分类的标准有很多种。按测量方式,可以分为主动测量和被动测量;按测量点的多少,可以分为单点测量和多点测量;按测量内容,可分为拓扑测量、性能测量和流量测量;按测量点所在层次,可分为网络层测量和应用层测量;按测量者是否知情,分为协作式和非协作式;按测量所采用的协议,分为基于BGP协议、基于TCP/IP协议和基于SNMP协议。

主动测量是指由测量用户主动发起测量,将探测分组注入网络,根据测量数据流的传送情况分析网络的性能。被动测量则通过在网络中的链路或设备(如路由器、交换机等)上借助包捕获数据的方式来记录网络流量,分析流量,获知网络的性能状况。

网络拓扑测量主要是了解网络拓扑结构,用于资源调度和流量分配。性能测量是通过监测网络端到端时延、抖动、丢包率等特性,了解网络的可达性、利用率以及网络负荷等。流量测量是对网络数据流的特性进行监测和分析,以掌握网络的流量特性,如协议的使用情况、应用的分布和用户的行为特征等。

比如,sniffer软件是被动测量,用于网络流量测量。而netperf软件作为网络性能测量工具,是主动测量,协作式,是基于TCP/IP协议的,测量点所在的层次为网络层。

2 网络测量的研究现状

网络测量不但具有广阔的工程应用前景,也有很

高的学术价值,国内外许多科研机构 and 学术组织都在进行网络测量方面的研究,很多公司也推出各具特色的网络测量工具。目前国内国际上涉及网络测量的研究及其相应的项目部署工作正以方兴未艾之势进行着。

2.1 国外情况

网络测量与分析的研究工作大致分为两类,一是网络测量与分析的参数研究;一是实验平台的研究。

自1996年始,IETF即开始对流量工程有关的领域进行跟踪,也成立了很多相关的工作组,如:ITRACE(ICMP Traceback, ICMP追踪),IPPM(IP Performance Metrics, IP性能测量),TEWG(Internet Traffic Engineering Working Group, 互联网流量工程工作组),ISMA(Internet Statistics and Metrics Analysis)小组等工作组,进行相关理论的研究。

下一代互联网的研究起源于实验平台建设,实验平台大致可分为三类:实验网络,具生产网络的特点,如Internet2^[3],Geant^[4],JANET^[5]等;实验系统,即模拟或仿真平台,不具备生产网络特点,在这个领域较为活跃的组织或团体包括CAIDA^[6]、PlanetLab^[7]、NS2^[8]、PerfSonar^[9],欧盟的EVERGROW下的ETOMIC^[10],日本的MAWI^[11]、英国的PPNCG^[12];集实验网络与实验系统双重特征为一体,如美国的GENI^[13],欧盟的FIRE^[14]等。另外,还有项目现已结束,但却在历史的发展中起了很大影响作用的项目,如美国的NLANR^[15]、Surveyor^[16],NIMI^[17]等。

其中,CAIDA重点关注测量、分析和可视化工具的开发,在全球都有测量节点,同时致力于有效部署和分析通过这些工具获得的数据。NLANR(1997—2006)率领下的两个项目AMP(主动测量)与PAM(被动测量)都为网络测量与分析提供了很多思路与经验。Paxson博士领导下NIMI项目也展示了在该领域取得的有关网络测量的研究成果。PlanetLab最明显的优势是研究及测度过程将保证研究人员体验到真正的网络行为,而不是虚拟的。另外,由于参与者众多,且来自不同领域,因此能够从多个不同角度出发来预测网络发展可能会出现的新服务。PerfSonar最大的亮点是解决跨多个网络的端到端的性能监视问题。

2.2 国内情况

在网络测量方面,国内起步虽晚,但是各大高校

及科研机构近些年来也做出了相当的工作。

CERNET(China Education and Research Network, 中国教育和科研计算机网):作为国内最大的研究性互联网络,很早就开展了以网络测量为基础的诸多研究工作,并将研究成果应用于流量计费、性能测量与管理以及网络安全检测等众多的网管分析课题中。DRAGON-Lab^[18]的全球分布式性能测量联邦GPERF目前在全球已有近千个测量点。

CSTNET(China Science and Technology Network, 中国科技网):中国主干网之一。2002年,中美俄三个国家共同出资建设环绕北半球的支持先进科研应用的高速长距离环状光网络(Global Ring Network for Advanced Applications Development, GLORIAD)。GLORIAD在高能物理、天文等领域支持实时大量数据的传输。我们在基于GLORIAD这一典型的高速长距离光网络的测量和实验中^[19],发现带宽利用率存在瓶颈,在分析高速长距离光网络的特点及TCP的机制后提出了TCP协议改进思路。

与此同时,各大高校网络实验室都在不同层面对网络测量展开了相关科学研究工作,比如:中科院计算所进行了端到端可用带宽的测量相关科研工作,北京航空航天大学,北京邮电大学都有对IP网络端到端性能测量技术的相关研究。值得一提的是,在实际网络环境不够充分的情形下,东北大学成为CAIDA中国第一结点,利用CAIDA数据进行相关研究。

3 网络测量的关键技术

(1) 拓扑测量:Internet网络拓扑研究分为四类:拓扑数据的获取、拓扑特征的发现、拓扑建模及拓扑内部关系。在网络上的网络拓扑研究一般分为IP级和自治系统AS级。不管是IP级还是AS级,研究人员为获得更加完整的拓扑数据通常采用多种方法混合的方式,但是由于网络规模的庞大,结构的复杂,想要获取完整的IP级或AS级网络拓扑数据仍然存在很多需要克服的困难。网络拓扑特征的研究在某种程度上是衡量和刻画Internet网络试题的提出和引入过程,一些新的度量也被不断的发现和提出,并用数学语言进行了描述。拓扑模型的建立在很大程度上依赖于拓扑特征的发现。用数学语言对一种或多种拓扑特征进行刻画并用算法生成该模型。当然也有拓扑生成器软件。如何理解和获取AS间的逻辑关系^[20-22],由于缺

乏统一的标准的方法,究竟应该如何将这些拓扑进行层次的划分以便于更容易的理解整个网络的拓扑结构还是一个未能解决的问题。另外,拓扑数据收集完整性方面,仍然有很多工作要做。对于Internet网络中幂律的形成、分布及正确性也面临诸多挑战。

(2) 带宽测量:主要集中在两个方面:端到端带宽测量和逐跳带宽测量。而端到端带宽测量又包括可用带宽和瓶颈带宽的测量。逐跳带宽测量则主要是测量网络路径上各段链路的带宽参数,着重于测量逐跳链路的带宽值。带宽测量算法分为三种模型:探测报文间隔模型(PGM)、探测报文速率模型(PRM)和基于网络流量模型的测量技术。PGM存在较大的测量误差。PRM的测量过程本身会影响网络状态和已有流量特征,导致网络不稳定和服务质量下降。最后一种测量结果与实际情况之间的差别较大。因此,好的带宽测量模型的提出还需要更多的工作。

(3) 流量测量:目前主流的三种方法有:基于SNMP统计数据来获取流量信息,但是只能提供粗粒度的流量信息;通过采集流经链路的报文(packet)来进行流量测量,因研究相对平等地分析每个报文,导致对报文间关系及其更高层次信息分析的缺失;流级别的测量,既提供详细的流量信息,又具备一定的可扩展性,因而受到了广泛的关注。流量抽样测量技术成为高速网络流量测量的重点,抽取流量子集来估计流量信息,由系统抽样,发展到随机抽样,后来出现了自适应抽样。Bloom filter^[23]是逐渐受到关注的高效的哈希结构。

(4) 延迟测量:单向时延的测量最主要的问题是要解决两个待测节点之间本地时钟的不同步,因此产生了各种不同精度的时钟同步方法和授时方法。而在已有的单向时延测量实验中,大多数是借助于GPS(Global Positioning System)接收机或NTP(Network Time Protocol)来实现同步。但是前者一般较贵且与接收环境有关,而后的精度又太低。因此,许多科学家研究出了几种经典时钟同步算法。例如V.Paxson的消除初始时间差与频率差的方法^[24];线性规划算法^[25]O.Gurewitz等人的算法^[26];Wei-xuan Gu等人的算法^[27],如何在前人的基础上提出更优的算法,在降低算法复杂度并提高更好的性能。

(5) 测量系统:网络测量系统中亟待解决的问题

是：网络测量系统的同一性结构、多种测量工具的动态协作、可扩展性、测量信息的建模和综合、自动化监测方法、网络测量的执行特性分析、资源授权和分配机制、最优化测量任务调度、指标量化问题。

(6) 可视化：测量的可视化的目标是直观地展示出全面而客观的网络测量分析信息。最常见的可视化技术是描出节点与节点间的链接^[28]。对于小型稀疏型网络，由节点和链接可以有效地表示出。但是对于大型网络，存在三个问题--显示杂乱、节点定位、可感知的紧耦合度。目前，解决这些问题主要将研究工作放在：聚焦动态参数，正确定位节点和 3D 布局。

(7) 不同的应用哪些性能指标是最为关键的，如何根据不同的业务自动地选取最为有效的网络性能参数。比如，与数据传输相关的业务，关注点在数据的可达性，丢包率是测量的重要指标；实时通信业务，比较关注消息传送的及时性，此时时延、抖动及丢包率是敏感指标；而在高速度长距离传输中，什么因素又是影响性能的主要方面，同时必须考虑网络测量的不确定性及误差。

4 常用测量工具

测量工具分为两类，一种是由公司开发的，有免费也有付费的；一种是由科研机构或组织开发的。以下主要从网络流量监控工具、网络性能测量工具展开描述。

网络流量监控工具主流的有：Sniffer Pro，Ethereal，PRTG，MRTG 及基于 Netflow 数据的软件等。Sniffer Pro 不支持 Linux，后三者同进支持 windows 和 Linux 系统。Sniffer Pro 和 Ethereal 是对网络的数据包进行监控，并可对数据包中的内容进行分析，功能上后者较前者更为强大，因为有很多的扩展插件。而 PRTG 和 MRTG 只是对流量信息的多少进行监视，并根据数据大小绘制图表，反复对比后发现异常现象，前者较后者操作更简单，监控内容更为丰富。

网络性能测量工具：netperf, iperf, tcptrace 等。Netperf 可确定大多数网络类型 TCP 和 UDP 端到端的性能，iperf 微调网络应用程序和服务器的 TCP 参数，可确定 TCP 窗口大小，生成该网络条件下的最佳吞吐量，用于优化应用程序和主机参数；而 TcpTrace 本身不执行网络测试，对由 tcpdump, wireshark 等

程序捕获的数据包进行分析，可使用 xplot, jplot 输出图形。

每一款工具的功能导向不同，各自有自己的特点及优势。因为每款工具都是以与现有工具较量的姿态面世的，故而在工具的选择上，根据需要，可综合选择几种工具。而像 PerfSONAR 也提供了 PerfSONAR Toolkit, PerfSONAR UI 等软件包，可以进行主动、被动测量解决跨多个网络的端到端的性能监测。CAIDA 组织也有提供了诸如主动测量工具 Scamper、可视化工具 cuttlefish 等。EVERGROW 组织提供了 DIMES 等测量工具。

5 GLORIAD网络测量应用

Gloriad(Global Ring Network for Advanced Applications Development, 中美俄环球科教网络)是一条新的环绕全球的高速环状光网络，用以支持先进科学应用。随着科教应用的发展，信道速率由目前的 2.5Gbps 进一步提升到 10Gbps。光网络是从一台计算机到另一台计算机间的一条直接的网络路径，采用永久的或暂时性的光路连接，不需要使用路由器。传统 IP 网络，是基于三层路由的“尽力而为”的传输方式，不能满足科研数据传输的要求。而光网络传输数据包不需要路由中转，端到端的光网络具有高容量、高 QoS、安全、透明等优势，更适合大量科研数据的传输。网络测量在保障数据的可靠有效传输方面起着举足轻重的作用。

5.1 针对天文 e-VLBI 数据传输应用的网络测量

e-VLBI 是指利用高速互联网将 VLBI 观测数据，实时高速送往数据处理中心，该技术也首次应用到“嫦娥一号”绕月探测工程首次飞行任务的精确测轨中。

为满足上海天文台同国家天文台、南山天文台、昆明天文台同欧洲甚长观测数据处理中心 JIVE、澳大利亚国家天文学院、日本国立天文台等机构之间的传输要求，中国科技网做了多次的网络测量及优化工作。在最初的端到端测试中，使用的 TCP 协议传输只有不足 50Mbps，丢包率达 4.6%，远远达不到 e-VLBI 实时观测实验要求的 512Mbps 及丢包率小于 1%的要求，对于 622M 的线路来说带宽利用率过低。

就此问题，采用主动式分段测量，在上海畚山、北京 CSTNET 机房、香港、JIVE 分别布置四台测试服务器(如图 1 所示)，使用 Traceroute、iperf、

TcpDump、TcpTrace 等开源网络测量工具定位网络故障，根据测量结果，优化网络路由及传输参数，并用 TCP 和 UDP 协议在三层以太网网络及光路上进行比较，有效地提高地网络性能及网络利用率，并成功地支持了 APAN 会议及多次 e-VLB 天文观测。

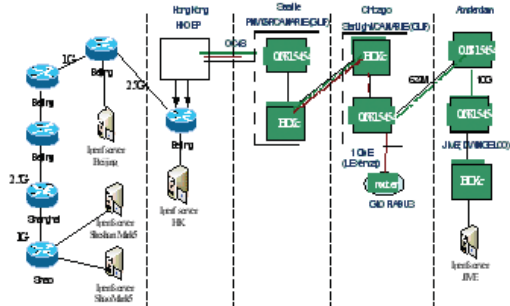


图 1 网络测试环境

5.2 针对高能物理数据传输应用的网络测量

06 年 6 月，中科院高能物理研究所与法国里昂之间进行 LCG(大型强子对撞机计算网格)的传输测试，使用 Grid-FTP 传输软件进行，测量结果显示单 TCP 连接传输速率为 875KB/sec，不能满足 LCG 本阶段应用 20MB/sec 的最低要求。针对此问题，对网络性能进行了测试。

对从香港到荷兰的链路性能进行测试：在中国科技网的香港节点部署了 MSTP 光设备，在对端荷兰节点部署同样的设备，中间的连接经过 GLORIAD 光网络上多台光传输和光交换设备，是一条端到端的直接的光通路，带宽为 OC-12(622Mbps)，在两台 MSTP 设备接两台测试服务器，此连接的瓶颈带宽是中间 WAN 的连接带宽为 622Mbps(如图 2 所示)。

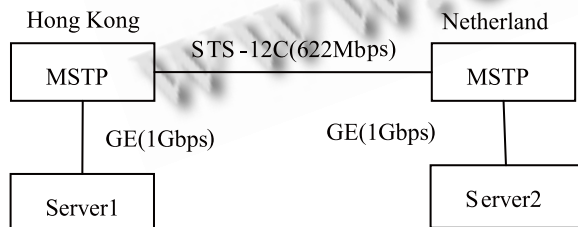


图 2 测试连接示意图

Ping 测试结果显示，链路往返延迟 RTT 约为 300ms。我们采用 Iperf 软件来测量吞吐量，并修改 TCP window size，MSS(Maximum Segment Size) 进行不同的实验测试。结果显示，默认设置下，单 TCP

连接吞吐率为 875Kb/sec，在 Window size 为 2MB，70 个并行流(Parallel Streams)时，得到最大实验结果吞吐率为 435Mbps。此吞吐率值约为瓶颈带宽 622Mbps 的 70%。

6 总结与展望

网络的设计和配置通常都遵循这样一句格言“网络出现拥塞时就增加带宽”，只要每条链路的可用带宽远远大于平均的通信负载，人们大概都会感到满意。可现实情况是，现有的带宽利用率并不高，即便带宽增加，对网络拥塞问题也没有得到实质解决。在没有找到真正制约因素前，增加带宽往往也是徒劳。根据从链路层到应用层的不同网络性能参数，如延迟、抖动、丢包率、传输协议、应用层协议、终端设备等对不同的应用环境的影响差异，建立一个能够全面反映网络性能的框架是下一步研究工作的重点方向。

目前，互联网上很多知名的实验平台都是由众多机构合作建设与运行的，但这些实验平台都是独立运行的，彼此缺乏相互合作，相互补充，即使是同类平台也缺乏合作机制。在国内，几乎所有大学的计算机系都会有研究人员选取网络测量与分析的某个方面进行相关研究，但是这些研究成果较为松散，如何有效地整合这些科研成果进而转化为工程应用，提升网络环境，这些都是今后努力的方向。

另外，在进行天文、高能物理等领域大数据传输提供网络保障过程中，目前中国科技网所提供的网络测量与分析，面临如下问题：所采用的测量方法单一、零散；测量效率不高、周期较长、重复工作多；测量结果不够精确；网络使用分析不够深入，难以快速高效定位网络传输瓶颈带宽并予以解决。因此着眼于高速长距离光网络可用带宽测量，设计并实施一个高效定位网络传输问题的网络测量分析系统是下一步要展开的工作。

参考文献

- 1 IETF.<http://www.ietf.org>
- 2 ITU-T.<http://www.itu.int/ITU-T/>
- 3 Internet2.<http://www.internet2.edu/>
- 4 Geant.<http://www.geant.edu/>
- 5 JANET.<http://www.ja.net/>

- 6 Brownlee N, Claffy K. Internet Measurement. In *Internet Computing*, IEEE. 2004,8(5):30 - 33.
- 7 Chun B, Culler D, Roscoe T, et al. PlanetLab: an overlay testbed for broad-coverage services. *ACM SIGCOMM Computer Communication Review*, 2003,33(3):3 - 12.
- 8 Breslau L, Estrin D, Fall K, et al. *Advances in Network Simulation*. IEEE Computer Society Press, 2000, 33(5):59 - 67.
- 9 Hanemann A, Jeliaskov V, Kvittem O, et al. Complementary visualization of perfSONAR network performance measurements. *Proc. of the International Conference on Internet Surveillance and Protection*. 2006.
- 10 Magana E, Morato D, Izal M, et al. The European traffic observatory measurement infrastructure (ETOMIC). *IP Operations and Management*. 2004:165 - 169.
- 11 MAWI:<http://tracer.csl.sony.co.jp/mawi>
- 12 PPNGC: <http://icfamon.dl.ac.uk/>
- 13 GENI Planning Group. GENI Design Principles. *Computer*, 2006,39(9):102 - 105.
- 14 FIRE.<http://www.future-internet.eu/>
- 15 NLANR.<http://www.nlanr.net/>
- 16 Kalidindi S, Zekauskas M. Surveyor: An infrastructure for internet performance measurements. *Proc. of the INET'99*. San Jose, 1999. http://www.isoc.org/inet99/proceedings/4h/4h_2.html
- 17 Paxson V, Adams A, Mathis M. Experiences with NIMI. IEEE Computer Society. 2002. <http://nm-group.tsinghua.edu.cn/classes/cl3/%5BPax99%5DExperiences%20with%20NIMI.pdf>
- 18 Wang JL, Li ZH, Lu GH, et al. DRAGON-Lab—Next generation internet technology experiment platform. *Science in China Series*, 2008, 51 (11):1908 - 1918.
- 19 Ren YM, Qin G, Tang HN, et al. Performance Analysis of Transport Protocol over Fast Long Distance Optical Network. *Journal of Computers*, 2008, 31(10):1679 - 1681.
- 20 Gao L. On inferring autonomous system relationships in the internet. *IEEE/ACM Transactions on Networking*, 2001,9(6):733 - 745.
- 21 Subramanian L, Agarwal S, Rexford J, et al. Characterizing the internet hierarchy from multiple vantage points. *Proceedings IEEE INFOCOM*, 2002, 2:618 - 627.
- 22 Battista G, Erlebach T, Hall A, et al. Computing the Types of the Relationships between Autonomous Systems. *IEEE/ACM Transactions on Networking*, 2007,15(2):267 - 280.
- 23 Broder A, Motzenmacher M. Network Applications of Bloom Filters: A Survey. *Internet Mathematics*, 2004,1(4):485 - 509.
- 24 Paxson V. On calibrating measurements of packet transit times. *ACM SIGMETRICS Performance Evaluation Review*, 1998,126(1):11 - 21.
- 25 Moon S, Skelly P, Towsley D. Estimation and removal of clock skew from network delay measurements. *IEEE INFOCOM*,1999,1:227 - 234.
- 26 Gurewitz O, Cidon I, Sidi M. One-way delay estimation using network-wide measurements. *IEEE/ACM Transactions*, 2006,14(SI):2710 - 2724.
- 27 Gu WX, Yu SZ. Inference of one-way queuing delay distribution using packet-pair probes without clock synchronization. *IFTP International Conferences*. 2007:169 - 175.
- 28 Eick S. Aspects of Network Visualization. *Computer Graphics and Applications*, 1996,16 (2): 69 - 72.