UNIX 网络性能管理的流量监测技术研究*

苏金树　刘雪青

（国防科技大学计算机学院　长沙　410073）

摘　要　网络性能管理的主要目标是高性能和高吞吐量。标准 UNIX 系统缺乏对网络的流量的有效管理。本文研究了在标准 UNIX 网络实现性能管理中的网络流量监测、主机网络/系统状况动态检测等技术，依据检测结果用户可以进行网络动态配置或重新分配网络资源，从而改善 UNIX 系统的网络性能及吞吐量等技术。结果表明所提供的技术手段是有效的。

关键词　网络，unix 系统，管理

分类号　TP393

On Network Traffic Management of Unix System

Su Jinshu　Liu Xueqing

(Department of Computer, NUDT, Changsha, 410073)

Abstract　Efficient performance and high throughput are the major goals of the network performance management. Standard UNIX system is short of efficient management of network traffic. This research thesis implements a network traffic query utility for users to monitor the network traffic situations. The network users can also query the network/system status of their computer hosts. Having Realized the network traffic situation, manager and users can reconfigure the network configuration, or reallocate the network resources to improve the network performance and throughput. The results show that techniques presented work very well.

Key words　network, unix system, management

1　网络性能管理与流量监测技术

1.1　网络性能管理技术

性能管理包括两个不同的活动。一个是有效积极的监测网络主机以便检测故障和决定操作方法，目的是测定主机的可用性，并定位其中的瓶颈。因为瓶颈问题是网络管理者改善网络性能的焦点，同时高级管理者需要依据性能报告来确定计算机系统的费用及带来的好处。

另一个是有效的性能测试和吞吐量设计。许多计算机网络的分析模式是建立在一些关于网络流量负载的分配和服务率的假想基础上。因而，对于现实情况还必须通过实际测定来确定耗费高的因素。网络性能通常受专门应用的负载特性的影响，因而模拟是一种良好的测量通信量的方法。有一种通信模拟器可以产生实际的通信量并跟踪网络中的通信。网络性能管理必须提供现时的网络结点的名字和域信息、网络结点通信量的状态和网络统计信息等。

1.2　网络的流量监测技术

网络的流量监测是网络管理的一个基本方面。这里可实现两种流量监测：错误检测和基本监测。在网络通信中，必须检测出各种错误。在实际应用中各类错误也必须按时记录下来。而且，记录下来的错误率作为网络流量率的一个指标用来表明网络中的拥挤情况。在基本监测中，第一步是测定网络的

* 1997 年 5 月 20 日收稿
第一作者：苏金树，男，1962 年生，副教授。

使用性。因为在此测定过程，可能产生若干网络问题，因而可依此而采取改善措施。

先给出一种从源地到目的地的网络流量模型，从而说明本地主机与其它 Internet 网主机之间的通信状况，同时记录通信的吞吐量、类型和时间。另外，在各种延迟变化或在路径流量强度突然增加时也应注意，还要记录它们以表明网络中存在的某些问题。网络应用通常需要一个 Internet 地址以便开辟一条通信联系或发送一个报文。用户通常喜欢使用符号名而不是 IP 地址。同时，网络上的名字服务器必须提供 IP 地址和符号名之间映射的翻译表。

域名信息不仅仅是找到 Internet 地址，每个域名在数据库中均有一项，每项有大量定义不同特性的记录，如 Internet 地址、计算机类型以及一系列由网络主机提供的服务，用户可由此找到专门的信息或有关主机名字或别名的信息。

Internet 还有定义域服务器名字的操作和用来查询它们的协议。通过名字服务器和域名服务器由网络的应用程序来实现。通过查询名字或域名，用户易得到网络结点的信息，否则就要通过大量网络去获得这些信息，这必将增加网络的总负载而且也浪费时间，因此这样便可以减少 Internet 网的通信负载。对用户来说，也为他们提供一条有用的信息并节省时间。

1.3 网络结点的通信量测量手段

通过网络通信量的状态可了解到网络中拥挤的情况。通信量取决于网络本身的吞吐量和用户请求的应用，如果有大量的应用超过网络的吞吐量，则会发生网络拥挤现象。

如何测量 Internet 网上的流量呢？基本的方法是模拟网络的流量，由流量产生器通过网络发送一个专门的邮包，计算邮包到达一个与之有关专门的结点的循环时间，这个时间将表明网络流量的状况。

这里有两类通信量的监测，分别是实时的流量查询和长期流量趋势发展的分析。前者为用户提供测量和了解当前为用户服务的主机的通信量状况的设备。后者则搜集一段时间内的网络主机通信量的状况，它允许用户去发现将来通信量的发展趋势，并找到其中的瓶颈，通过长期通信量报告提供的信息，网络管理者可根据实际情况改善网络性能。

本文的研究提供两种测量通信负载的方法：查询专用的网络主机和查询网络配置 profile。前者提供一条简单的专用通信量负载链，另一方面，用户可能关心 Internet 网中某些过热点，他们可以将过热点的符号名或 IP 地址输进网络配置的 profile 中，然后通过查询 profile 测量管理站和几个过热点之间的网络状态。

1.4 网络信息量统计

主机性能直接影响网络性能的好坏。用户主机的统计信息应同时提供给用户和网络管理者。这个信息包括两个方面：网络的状态和主机的系统状态。网络的状态将表明各种相关的网络数据结构和有关信息的内容，包括（1）每个网络协议的一系列活动；（2）网络的数据结构；（3）网络的路由表；（4）累积的通信量统计数据等主要信息：通过这些状态信息，用户和网络管理者可以了解他们主机中当前网络的环境。

主机的系统状态能让用户知道当前的主机环境，并提供关于主机处理、虚拟存储、磁盘、陷阱和 CPU 活动的报告，磁盘的 I/O 活动及 CPU 的使用情况等情况。

2 软件系统实现概况

在上述理论分析基础上，实现了监测网络通信量和测量性能的软件，为用户、网络管理者和系统管理者提供一个获得各种网络流量状态和流量统计信息的工具。首先，这个程序将为用户提供一个交互式环境以便指定有关他们要求的监测方案。这其中有一些查询功能使用批处理方式，与在 TCP/IP 环境下使用的 SNMP (Simple Network Management Protocol) 相似，这些程序可以被当作 NMS (Network Management Station) 放置在任何主机中，也即是说，此程序是一种逻辑的监测程序可查询被管理的结点的星型拓扑结构中运行，用户面对的是中心结点，他可以直按监测或测量专用网络结点的流量状态。

这个应用软件的实现分为单个主机网络状态的查询、网络 profile 中流量状态的查询和网络 profile 的维护三个部分；前两个部分别用于查询网络信息和流量状态，通过查询可以了解网络的主要参数：结

点状态和流量状态。第三部分是维护网络 profile 文件并作为网络配置管理的工具。

2.1 单个主机网络流量状态的查询

该部分提供八个具有查询功能和为用户获得网络和主机系统信息的应用程序。

(1) 查询主机名和地址。此查询功能将给出用户关心的专用结点的名字信息。在此之前，用户需要输入网络结点的地址或名字，那么就可获得它的“名字”信息，包括正式的主要名，别名，地址类型，地址长度以及 IP 地址。

(2) 查询 Internet 域名服务器。交互地查询域名服务器，它涉及到名字服务器，要求有关专用结点的信息，并在一个专门域中打印一系列网络结点。在此查询中，用户可以选择服务器，查询类型以及其它查询域名服务器时的设置参数。为了以后的查看方便，用户也可以打印一些有用信息。此外，用户可以从域名服务器得到有关网络结点的详细信息，这样就可以节省查找时间，也减轻了 Internet 网上的负载量。

(3) ECHO 数据报文：查询程序提供两条主机网络信息：可达性测试和测量的循环时间。后者即为从用户需要的主机到专用的网络结点的时间和从该结点返回的时间。此信息将指明两结点间通信量状态。此功能使用了 ICMP 协议的命令式的 ECHO-REQUEST 的数据报文来得到一条从专用结点或网关发来的 ECHO-REPONSE。此 ECHO-REQUEST 数据报文有一个 IP 和 ICMP 报头，按下来是一个‘时变结构’和一个在填充包中以任意数目出现的字节。若从一个专用结点返回，它将指出到达另一个结点的可达性。当用户发送 ECHO-REQUEST 给目标结点时还可采用一些其它任选项如路由类型，冗余性和数据报文大小。当返回时，从目标结点发回 ECHO-REPONSE，那么就可知道两结点间的通信量状况。

(4) 显示网络状态的应用程序。此程序用于显示在各种选择格式中各种相关网络的数据结构的内容。这里有三种网络的状态信息：socket、网络数据结构和累积的流量统计。

(5) I/O 统计报告的应用程序

此程序用于记录终端和磁盘 I/O 的活动，收集的信息包括系统最近一次被引导起累积统计的信息。从内核中收集 I/O 信息。内核中设置几个计数器，每次磁盘寻找、数据传输的完成和大量词的传输都由计数器来维护。在每次时钟敲击时，计数器检查每个磁盘的活动状态，若有效则计数一次。内核还同时提供设备大致的传输率。

此程序允许用户以交互式或批处理方式执行。另外有关 I/O 活动的这些信息可被用来检测那些可能降低网络流量性能的不正常条件，故可以消除这些不利条件以提高网络的性能!

此外，还有虚拟存储统计报告的应用程序，UNIX 命令的应用程序、文件传输的测量工具等。

2.2 网络 profile 通信状态的查询

profile 是一个表示网络配置的正文文件。文件中每个记录表明一个结点和它所包含的符合名或 IP 地址。此函数可以在 SHELL 中以批处理方式使用专门的 profile 实现。执行主机测定由它到 profile 所有专用结点的通信状况。也即是说，任何一个被设置的结点可以象一个网络管理站那样执行任务。对于一个分布式系统，可以在几个结点中安装它并从远处启动它。

(1) 可达性测试

此测试使用网络 profile 中每个结点的符号名或 IP 地址以测试与另一结点通信的可能性。若主机收到 ECHO-REPONSE 的数据报文，则表明此结点存在，否则在主机和结点间无可达性。每个用户可依据所关心的网络配置情况去选择专门的网络 profile。这种可达性测试通常以交互式进行，网络结点间的可达性向用户提供网络的基本状态，故任何必须先进行可达性测试。

(2) 流量统计报告

主要用于收集从网络 profile 指定结点产生的流量统计信息，分为两类：实时流量统计和长期流量统计。实时方式下备选样机依照用户提供的 profile 交互地执行。用户认可 profile 后，必须指定被填充到 ICMP 数据报文中的数据长度和将被用于产生网络流量状况的统计数目。若用户想了解比较准确的流量状况，他们将选择较长的数据报文的较大的统计数目，但同时又会花费时间和增加网络负载。

(3) 网络结点流量状态的测定

每条通信链路可能有几种不同的流量状态，因此有不同的额定值，它有助于管理者决定网络应用的调度和诊断网络的故障。若一个或多个结点总有很差的额定值，那么就应注意这些结点。

结点间的链路通常有不同的长度和传播延迟。数据报文的传输时间不仅仅是用于测定结点流量状态的标准。好的流量状态意味着结点有稳定的传送数据的能力，而且传送时变化小。因而，按收包的百分比和循环时间的变化值将是判断结点状态的两个重要的标准。前者表明了在用户结点和被管理结点之间传送报文的能力，后者则指明了两个结点之间传输报文所用的循环时间的变化。另外，传送数据的能力是测定结点最重要的因素，即“接收包的百分比”常被用作测定的指标。其次是时间的变化意味着传输能力更稳定。

3 实际应用结果

最后，本文给出利用本系统在一个校园网上所得到“网络流量状态的测定”的结果。“网络流量状态的测定”主要是收集、测定以及记录流量状态。这里有两类记录：实时结点流量状态测定值的报告和长期的流量状态测定值的报告。前者允许用户选择网络 profile，同时要设定 ICMP 数据报文和大小和计数数目，如表 1 所示，报文长度：512；采样次数：5 次。

表 1 网络节点流量实时状态报告

	接收报文数	平均来回时间 (ms)	间隔变化	均差
1	100. 00	0. 000	0. 000	0. 000
2	100. 00	6. 000	30. 000	5. 477
3	100. 00	14. 00	230. 00	15. 166

当平均来回时间逐渐变大时，说明网络的负载较重，应考虑如何分解网段，当间隔变化变大时，说明网络的负载出现时大时小的非稳定状态，可能存在一些异常情况。

“长期结点流量状态的测定报告”同时还要设置整个测量期和采样期，并将产生的信息存储在用户利用专用参数指定的文件中，如表 2 所示。

这份长期报告是总结了许多实时流量状态的测定报告。时间可以唯一区别这两类报告，它指明采样每个结点通信状况额定值的时间。与此同时，此报告还可以用于专用结点的诊断工具，可以检测处于低额定值结点的性能。于是管理者可依据所得出的结论调整网络配置并重新分配网络资源，以提高网络的性能和专用结点的吞吐量。

表 2 长期结点流量状态的测定报告

	接收报文数	平均来回时间	间隔变化	均差
Wed Feb17 17：00：14 1996	100. 00	0	2	9
Wed Feb17 17：13：14 1996	100. 00	0	1	10
Wed Feb17 17：27：15 1996	100. 00	0	0	10

参考文献

1 John J. The Whole Internet User's Guide & Gateway. CA, O'Reiuy & Associates, Inc., 1992

2 Albert P. Tools for monitoring and TCP/IP internets and interconnected devices. 1990

3 苏金树，卢锡城. 高性能计算机网络技术. 北京：电子工业出版社，1996