

支持普适服务的新一代网络传输层构架

杨 冬 李世勇 王 博 张宏科

(北京交通大学电子信息工程学院下一代互联网互联设备国家工程实验室 北京 100044)

摘 要 新一代网络没有准确定义,但普遍共识是提供全分组网络 and 普适服务.其中分组网络一般认为采用 IP 分组,但普适服务仍不清晰.3G、4G 中称“泛在服务”;NGI、NGN 中强调无缝融合电信网和互联网服务;计算机界类似的概念是“普适计算”;IEEE 普适服务会议汇聚各种思想,首次提出了“普适服务”的模糊概念.虽然上述研究对普适服务的定义并不统一,但普遍共识是:多样化的服务种类、多样化的网络接入.文中分析传输层在新一代网络设计中的重要地位,以提供普适服务为目标,通过多连接抽象表示服务多样化、多路径抽象表示接入多样化,给出符合新一代网络要求的传输层构架.基于这一构架,可以提高网络的 QoS 支持、安全可靠、移动性、可控可管性.文中通过理论分析和原型实现证明了新构架的正确性和可行性.

关键词 普适服务;传输层;多连接;多路径;映射

中图法分类号 TP393

DOI号: 10.3724/SP.J.1016.2009.00359

New Transport Layer Architecture for Pervasive Service

YANG Dong LI Shi-Yong WANG Bo ZHANG Hong-Ke

(National Engineering Laboratory for Next Generation Internet Interconnection Devices,
School of Electronics and Information Engineering, Beijing Jiaotong University, Beijing 100044)

Abstract There are no uniform definitions of the new generation network, and most researches tend to construct new packet network and provide pervasive service. New packet network may well be IP, but the target of pervasive service is not clear. 3G and 4G call it ubiquitous service; NGI and NGN emphasize service merging; Computer world uses the concept pervasive computing; IEEE International Conference on Pervasive Services brought forward pervasive service firstly. Thought there are many different researches about pervasive service, the most important goal is to provide different kinds of services and accesses. This paper introduces a new transport layer architecture based on multi-connection and multi-path mapping for pervasive service. The new architecture could improve QoS, reliable, security, mobility and controllable of new network. Theory models analysis and prototype implement prove its correctness and feasibility.

Keywords pervasive service; transport layer; multi-connection; multi-path; mapping

1 引 言

新一代网络研究是当前最热门的课题,尽管对

这一问题开展了很多不同的研究,如国外的 3G、4G、NGN、NGI,国内的可信网络^[1]、一体化网络^[2]等,但是普遍的观点认为:未来网络是基于分组的统一网络,该网络可以为用户提供普适服务.目前统一

的分组网络趋向于使用 IPv6,但“普适服务”目前还没有标准的定义. 4G 网络中称之为“泛在的服务”(ubiquitous service),强调多种接入;NGI、NGN 中强调无缝融合以话音业务为代表的电信网服务和以数据业务为代表的互联网服务;计算机界相关的概念是“普适计算”;2004 年,IEEE 创办了普适服务国际会议(ICPS),其内容除了普适计算等传统主题外,还新增了普适服务及应用. ICPS 对普适服务的定义如下:“普适服务和计算是新出现的计算范例,其基础架构和服务可以在任何时间、任何地点、通过任何格式无缝地接入和获取.它是近年来无线和传感器网络、分布式系统、移动和代理计算以及自治计算和服务等技术发展进步的结果”^①.从以上各种对“普适服务”的研究可以看出,虽然没有准确、统一的定义,但普适服务的主要内容是:多样化的服务种类、多样化的网络接入.

本文研究符合新一代网络普适服务要求的传输层构架,之所以选择传输层为切入点主要基于以下原因.

首先,传输层在网络分层中处于承上启下的重要位置.从网络分层设计的角度出发,新一代网络普适服务的两个研究内容:多样化的服务和网络接入,分别属于服务层和网络层.传输层连接服务层和网络层,是实现未来网络普适服务两个主要内容都无法回避的.

其次,传输层在新一代网络设计中的重要作用.从实现目标来看,4G、NGN 和 NGI 代表了通信技术的统一,即网络融合技术.随着这些技术的发展,将出现以下情况:上层应用程序对传输服务要求越来越高,而下层的接入网技术越来越复杂. TCP/IP 不仅实现了对底层异质传输技术的屏蔽和融合,同时对上层的语音和数据实现提供支持.因此, TCP/IP 的基本结构在未来网络中将发挥更大作用.其中全 IP 网络已成为各种新网络研究的趋势,然而 IP 作为一种尽力而为的网络肯定无法满足未来网络的要求,因此以 TCP 为代表的传输层协议将成为未来网络实现普适服务、一体接入的关键技术.

最后,新一代网络需要对 QoS、移动性、安全可靠、可控可管等有很好的支持,这些技术都与传输层密不可分.由于传输层为尽力而为的 IP 网提供了面向连接支持,因此传输层需要维护大量的网络协议参数,这些参数是实现端到端 QoS 的关键.主机移动性问题是互联网最主要的问题之一,目前在互联网的多个层中都有解决方案.链路层被强制要求提供

移动性支持,但是链路层移动性不能解决高层移动性问题;网络层移动性需要增加新的网络功能实体,实现成本较高;会话层由于没有被应用程序开发者广泛使用,其移动性的解决缺乏统一的开发方案.基于以上考虑,传输层移动性逐渐得到重视.传输层移动性可以为应用程序提供透明的支持,同时不需要增加新的网络功能实体.但是传输层移动性有自己的问题,就是需要修改现有的传输协议,这正是本文研究的内容.另外,传输协议对于数据传输的可靠性、安全性、可控可管性等设计具有重要的作用.

基于以上考虑,本文提出基于多连接多路径映射的传输层新构架.体现在应用程序具体实现上,多连接对应于多服务,多路径对应于多接入.通过映射将多连接和多路径进行有效的结合,为实现普适服务需要的多服务和多接入提供了一种有效的实现方式.

本文第 2 节给出新构架的基本设计思想;第 3 节基于排队模型建立新构架;第 4 节对各种排队模型进行数值分析;第 5~7 节是 QoS 支持、安全可靠性、移动性的支持;第 8 节简单介绍新构架原型系统的实现方法;第 9 节总结全文.

2 设计思想

2.1 多连接

本文所指的多连接包括 3 种情况:基于不同传输协议建立的多个连接,如 UDP 连接(UDP 是面向无连接的,这里仅指应用程序在两台终端之间通过 UDP 完成数据传输)、TCP 连接、SCTP^[3] 联盟;一种传输协议为同一个应用建立的多个连接,如多个 UDP 连接、多个 TCP 连接,或者多个 SCTP 联盟;一种传输协议的一个连接内的多个子连接,如一个 SCTP 联盟中多个流.上述多连接都与互联网最初设计有所不同,以下逐一分析 3 种情况的出现.

当前互联网使用最多的传输协议是 TCP 和 UDP,然而 UDP 的完全无序和 TCP 的完全有序决定它们无法满足未来网络服务的多样化需求,新一代分组网络的传输层将由多种不同类型的传输协议或者一种可以提供不同传输特性的协议组成,目前 SCTP、DCCP^[4] 等已经逐渐显示出许多优良的应用前景.以上情况导致了多连接的第 1 种情况.

多连接第 2 种情况的出现源于网络服务的日益

① ICPS Conference Scope. <http://icps2005.cs.ucr.edu/>

复杂.以 Web 浏览为例,随着网络内容的不断丰富,网页所承载的数据种类不断增加,除传统的文本、图片、视频外,各种动画、插件、实时交互命令等都依托网页页面开展.这一情况导致 HTTP 1.0 中,客户端根据解析网页 HTML 文件得到的 URIs 建立多个传输层连接.由于 HTTP 基于 TCP,多个 TCP 连接需要多次三步握手,降低了效率.作为 HTTP 1.0 的升级版本,目前广泛使用的 HTTP 1.1 支持持续的管道连接.持续管道减少了 TCP 连接的数目,但由于网页内容多样化,适当数目的多连接很有必要(Firefox 浏览器支持用户根据需要自己设定管道数目).有关这类多连接的研究较多,文献[5]设计实现了一个支持多连接的函数库(Psockets),可实现应用数据分割,再基于多个 TCP 连接传输,这一设计与单 TCP 连接相比可以明显提高传输效率;文献[6]和文献[7]都通过修改 TCP 实现了并行 TCP 套接口,并且性能获得了很大提升;文献[8]设计了并行 TCP 在卫星通信中的应用,将数据分割成 8KB 大小的数据块,然后基于多个套接口传输;Storage Resource Broker (SRB)^[9]使用多连接提高了数据传输吞吐量;I2-DSI^[10]使用多连接提高分布式数据缓存的性能.

HTTP 1.1 的管道连接在减少 TCP 握手次数的同时带来了头端阻塞问题,即网页中逻辑独立的对象之间相互影响,先到的对象如果丢失了数据包会影响后到完整对象向浏览器递交.造成这一问题的根源是多个传输对象共用一个连接.为此 SCTP 的多流设计用来消除这一问题,也就是第 3 种多连接的情况出现的原因之一.另外,传统语音服务需要将信令数据和语音数据分离传输,也需要使用多连接.

最后,作为 3G、4G 等研究的一个杀手级应用,服务整合将成为新一代网络设计必须考虑的问题.然而这种技术加剧了一次应用的复杂度,使得上述各种多连接的情况更加复杂.

2.2 多路径

本文的多路径等同于网络接入设备(或网卡).相对于多连接技术,传输层端到端多路径并行传输技术还比较新.端到端多路径技术的兴起得益于接入技术的多样化以及接入设备成本的降低.随着网络用户的不断增加,各种网络接入技术不断涌现.对于无线用户,可以选择 Globalstar 或 Iridium 进行卫星接入;通过 CDPD, GPRS, EDGE 或 3G 进行广域网接入;通过 Ricochet 进行城域网接入;通过 IEEE 802.11、HiperLAN、蓝牙、红外或 HomeRF

等进行局域网的接入;有线用户接入互联网也可以通过 ISDN 接入、DDN 接入、ADSL 接入、Cable Modem 接入、无线接入、光纤接入、电力线接入等多种技术.另一方面,随着芯片制造成本的降低,各种接入设备的价格逐渐下降,一台主机上有多块网络接入设备越来越平常.

随着接入技术的多样化以及终端上网络接入设备的增加,一个很自然的问题就是:能否让不同的接入设备同时为一次应用服务,最大程度地提高网络的效率?针对上面的情况,目前一个广泛应用的技术是端到端多路径并行传输技术. mTCP^[11] 基于修改 TCP 实现在覆盖网上多路径并行传输,它的主要贡献是解决了多路径带来的乱序问题;文献[12]和文献[13]通过修改 SCTP 的发送端机制实现多路径传输,主要思想是引入一个可以跨路径记录数据包顺序的序号;LS-SCTP^[14] 是基于 SCTP 的多路径实现,它通过修改 SCTP 数据包的格式引入一个记录路径内数据包顺序的序号,将 SCTP 中面向关联的拥塞控制扩展到面向路径.

2.3 支持普适服务的传输层构架思想

通过以上对多连接、端到端多路径的分析和技术回顾可以看出,现有多连接和多路径的实现单纯从连接和路径自身出发,缺乏两者之间的合理关联.另一方面,为了使用多连接和多路径设计,需要将应用层的服务数据进行分割,目前的分割比较简单:多连接一般是将服务层数据直接分割成固定大小的数据块,如文献[5,8]中将数据分割成 8KB 数据块;多路径中服务数据的分割没有任何策略,只是通过侦听机制将数据包发送到空闲的路径,保证每条路径尽量不会空闲即可.这些问题使得服务分割、连接、路径之间缺乏合理的联系和协同工作能力,不能充分利用下层网络设施,也不能为上层的网络应用提供多样化支持.

图 1 左边是现有网络终端的常见结构,网络服务较为单一,一次服务以建立单一连接为主;终端的接入方式和接入设备一般只有一个.在这种情况下,不用过多考虑连接建立的选择以及连接与接入设备的对应关系.随着现有各种网络无缝融合为图 1 右边所示的新一代网络,网络服务将日益多样化,而且每种服务将更加复杂化,一次服务不能再简单地通过一次传输层连接完成,需要分割服务并建立多个连接,这些连接可以基于一种传输协议,也可以选择不同的传输协议并行完成,而且新的传输协议支持

一次连接内再建立多个子连接. 另一方面, 未来网络接入方式的增加以及接入设备成本的降低, 一台终端上有多个接入设备的情况越来越普遍.

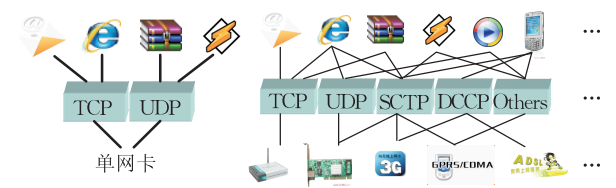


图 1 网络服务和接入变化趋势

基于上述变化, 本文设计传输层提供的多连接多路径映射构架. 设计核心思想是: 根据数据类型、用户需要等, 为多个服务以及复杂服务分割成的子服务灵活建立不同的连接; 多个连接再经过合理的映射对应到多个接入设备. 基于这种设计思想的传输层构架可以提高网络的 QoS 支持、安全可靠、移动支持、可控可管能力.

3 新传输层构架设计

新映射构架基于排队论建模, 服务器表示路径; 客户表示连接对应服务的数据包. 为了使模型更加完整, 除给出 3 种多路径映射模型外, 同时给出现有单路径下的两种映射模型. 实际上单路径是多路径的一个特例, 新构架下这两种类型将共存.

3.1 基于排队的映射模型

本节给出各队列模型代表的映射类型, 具体的排队论解释在 3.2 节. 图 2 为现有多连接单路径模型 M1; 图 3 为具有优先的多连接单路径模型 M2, 用于对 QoS 支持有要求的情况; 图 4 为简单多路径模型 M3, 该模型用于一般情况多路径传输, 特点是不会出现路径闲置, 整体传输效率高; 图 5 为流与路径绑定多路径 M4、M5 (M4 和 M5 的区别是绑定顺序不同), 该模型用于移动切换情况下多路径传输, 特点是数据流之间不受路径差异的影响; 图 6 为具有优先级多路径模型 M6, 该模型用于具有 QoS 要求的多路径传输, 特点是高优先级的数据可获得高的传输效率, 并且不降低整体效率.

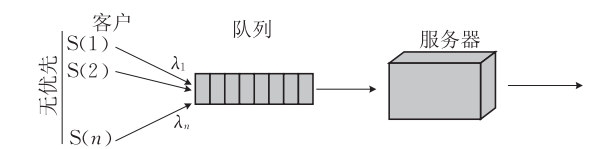


图 2 现有多连接单路径模型

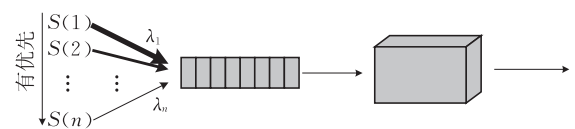


图 3 优先多连接单路径模型

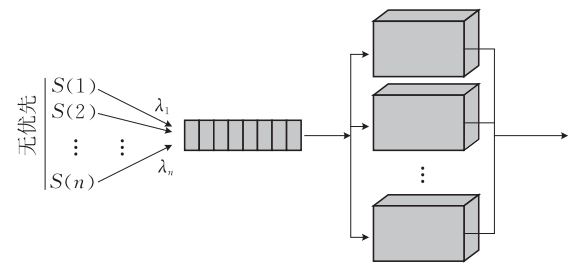


图 4 简单多连接多路径模型

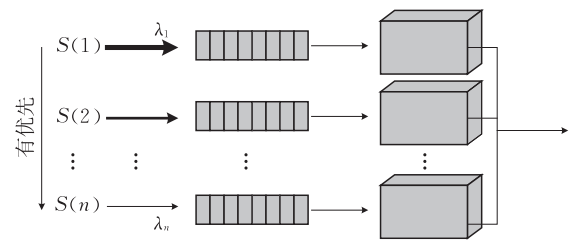


图 5 绑定多连接多路径模型

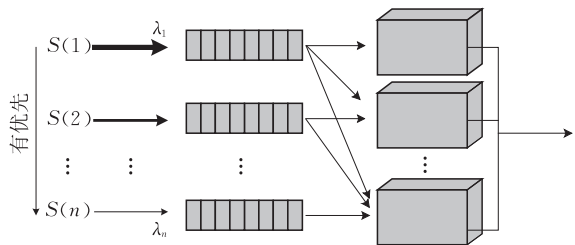


图 6 优先多连接多路径模型

3.2 模型分析

为说明各模型的意义, 使用排队论分析方法^[15]求解各模型. 通过分析可以获得服务数据包在各种模型下的平均服务时间和平均等待处理数目, 这些结果可用于模型间的比较和分析. 为简化分析, 采用 2 个连接和 2 条路径, 其中连接 1 和路径 1 的优先级高于连接 2 和路径 2. 服务数据包是队列的客户, 数据包的传输时间是客户的服务时间. λ_i 是连接 i 表示的服务数据的到达率; 路径 i 服从参数为 $1/\mu_i$ 的指数分布. 假设服务数据包的大小一样, 因此路径的能力只取决于参数 μ_i ($\mu_1 > \mu_2$); 再假设单路径模式下路径的能力是多路径模式下两条路径能力和的一半, 即 $\mu = (\mu_1 + \mu_2)/2$, 该假设使得模型间的比较更加公平. S_i 和 L_i 分别为连接中数据平均服务时间和平均等待处理的数据包数目.

(1) 现有多连接单路径模型(M1)

该模型(如图 2 所示)用于描述现有多连接单路径的实现,模型由 2 个连接(客户)、1 条路径(服务器)组成,2 个连接无优先级. 该模型是一个有两类客户的简单 M/M/1 队列,令 $\rho = \lambda_1 + \lambda_2 / \mu$, 则

$$E(S_1) = E(S_2) = \frac{1/\mu}{1-\rho}, \quad E(L_1) = E(L_2) = \frac{\rho}{1-\rho}.$$

(2) 优先多连接单路径模型(M2)

该模型(如图 3 所示)用于描述有优先级的多连接单路径映射. 模型由 2 个连接、1 条路径组成,连接 1 的优先级高,该模型是一个有两类不同优先级客户的 M/M/1 队列,令 $\rho_i = \lambda_i / \mu$, 则

$$E(S_1) = \frac{1/\mu}{1-\rho_1}, \quad E(L_1) = \frac{\rho_1}{1-\rho_1},$$

$$E(S_2) = \frac{1/\mu}{(1-\rho_1)(1-\rho_1-\rho_2)},$$

$$E(L_2) = \frac{\rho_2}{(1-\rho_1)(1-\rho_1-\rho_2)}.$$

(3) 简单多连接多路径模型(M3)

该模型(如图 4 所示)用于描述简单映射关系的多连接多路径. 模型由 2 个连接、2 条路径组成. 该模型是一个有两类客户的 M/M/2 队列. 令 $\rho = \lambda_1 + \lambda_2 / \mu_1 + \mu_2$, 则

$$E(S_1) = E(S_2) = \frac{1}{\mu_1} \frac{1}{(1-\rho)^2} \left(1 + \frac{\lambda_1 + \lambda_2}{\mu_1} \frac{1}{1-\rho} \right)^{-1},$$

$$E(L_1) = E(L_2) = \frac{\lambda_1 + \lambda_2}{\mu_1} \frac{1}{(1-\rho)^2} \left(1 + \frac{\lambda_1 + \lambda_2}{\mu_1} \frac{1}{1-\rho} \right)^{-1}.$$

(4) 绑定多连接多路径模型(M4、M5)

M4 模型(如图 5 所示)用于描述连接与路径的绑定映射. 模型由 2 个连接、2 条路径组成. 该模型可认为是两个 M/M/1 队列,路径 1 接收处理连接 1 的数据包,路径 2 处理连接 2 的数据包. 令 $\rho_i = \lambda_i / \mu_i$, 则

$$E(S_1) = \frac{1/\mu_1}{1-\rho_1}, \quad E(L_1) = \frac{\rho_1}{1-\rho_1},$$

$$E(S_2) = \frac{1/\mu_2}{1-\rho_2}, \quad E(L_2) = \frac{\rho_2}{1-\rho_2}.$$

该模型的另一种情况 M5 是路径 1 接收处理连接 2 的数据包,路径 2 接收处理连接 1 的数据包.

$$E(S_1) = \frac{1/\mu_1}{1-\rho_1}, \quad E(L_1) = \frac{\rho_1}{1-\rho_1},$$

$$E(S_2) = \frac{1/\mu_2}{1-\rho_2}, \quad E(L_2) = \frac{\rho_2}{1-\rho_2}.$$

(5) 优先多连接多路径模型(M6)

M4 基于绑定映射的模型可能造成路径的闲置,M6(如图 6 所示)描述具有优先级的多连接多路径实现,同时消除路径闲置. 模型由 2 个连接、2 条路径组成. 该模型是一个有两个优先级区分客户的

M/M/2 队列,其中具有高优先级连接 1 的数据包抢占优先低优先级连接 2 的数据包,令 $\rho_i = \lambda_i / \mu_i$.

图 7 的 2 维马尔可夫链用于描述这一模型,一维用于记录高优先级数据包数目,另一维记录低优先级数据包数目. 由于高优先级连接 1 的数据包的处理不依赖于低优先级的数据包,可得到以下结果:

$$E(S_1) = \frac{1}{\mu_1} \frac{1}{(1-\rho_1)^2} \left(1 + \frac{\lambda_1}{\mu_1} \frac{1}{1-\rho_1} \right)^{-1},$$

$$E(L_1) = \frac{\lambda_1}{\mu_1} \frac{1}{(1-\rho_1)^2} \left(1 + \frac{\lambda_1}{\mu_1} \frac{1}{1-\rho_1} \right)^{-1}.$$

低优先级的数据包数目依赖于高优先级数据包的数目,这类客户具有优先级的多服务器模型的分析非常困难. 为了定量分析以上模型,本文使用“Dimensionality Reduction”方法^[16]进行处理. 该方法可将一个多维的马尔可夫链降级为一个可以近似替代该多维马尔可夫链的一维的马尔可夫链.

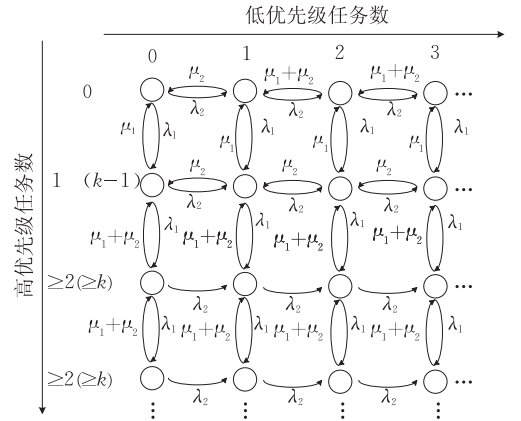


图 7 M6 模型的二维马尔可夫链表示

首先注意低优先级的数据包的处理仅由高优先级数据包数目(为 0, 1, 或 ≥ 2)决定,当系统中高优先级数据包的数目多于 2 个时,高优先级数据包的具体数目不再影响低优先级数据包的处理. 于是构造一维马尔可夫链,该链中高优先级数据包的数目仅为 0, 1, 或 ≥ 2 这 3 种情况,高优先级多于 2 个的状态逗留时间可使用一个两阶段的 Coxian+PH 分布来近似表示. Coxian+PH 分布的参数可通过“moment matching algorithm”^[17]生成. 经过以上处理可以得到近似模拟有两个具有优先级客户的 M/M/2 队列的一维马尔可夫链(见图 8). 该一维马链可借助矩阵分析“matrix analytic methods”^[18]进行处理,计算低优先级数据包的情况.

设高优先级无限状态过程为 B , 近似代替该过程的有限状态过程为 \tilde{B} . 通过使用 \tilde{B} 来替换 B . 通过以上分析可以得到的 \tilde{B} 的生成矩阵 \mathbf{Q}_B , 基于 \mathbf{Q}_B 又可以得到一维马尔可夫链的生成矩阵 \mathbf{Q} .

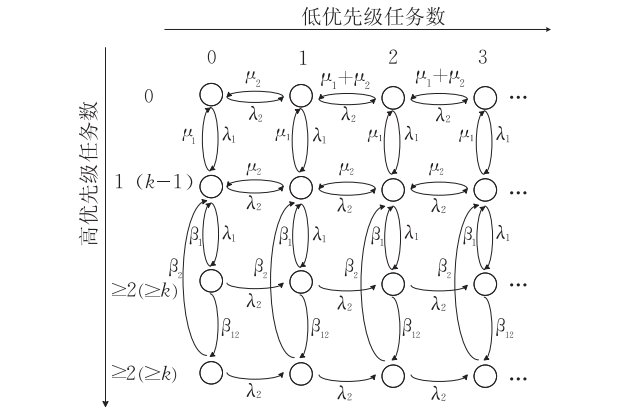


图 8 M6 模型维数减少后的马尔可夫链

$$Q_B = \begin{bmatrix} -\lambda_1 & \lambda_1 & & & \\ \mu_1 & -(\mu_1 + \lambda_1) & \ddots & & \\ & \ddots & \ddots & \ddots & \lambda_1 \\ & & \mu_1 & -(\mu_1 + \lambda_1) & \lambda_1 \tau \\ & & & t & T \end{bmatrix},$$

其中 $t = -T\mathbf{1} = \begin{bmatrix} (\beta_1 + \beta_{12}) & -\beta_{12} \\ 0 & \beta_2 \end{bmatrix} \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix}$, Q_B 是一个 $(k+2) \times (k+2)$ 阶的三角阵, 生成矩阵 Q 为

$$Q = \begin{bmatrix} L^{(0)} & F \\ B & L & F \\ & L & \ddots \\ & & \ddots & \ddots \end{bmatrix},$$

其中 F 和 B 是 $(k+2) \times (k+2)$ 阶三角阵, $L^{(0)} =$

$$Q_B - F, L = Q_B - F - B.$$
$$F = \begin{bmatrix} \lambda_2 & & & \\ & \lambda_2 & & \\ & & \lambda_2 & \\ & & & \lambda_2 \end{bmatrix}, \quad B = \begin{bmatrix} & \mu_2 & & \\ & & \mu_2 & \\ & & & \mu_2 \\ & & & & \mu_2 \end{bmatrix}.$$

有了 Q 的生成矩阵, 就可以分析低优先级过程。
以上的映射模型及相应的队列分析结果将用于本文后续分析. 简单映射 M3 用于安全可靠的高效传输, 优先映射 M2、M6 用于 QoS 支持, 绑定映射 M4、M5 用于解决移动切换问题.

4 模型数值分析

上节对各种模型进行了排队论求解, 为了更加直观地表现各种模型的区别, 本节通过数值分析对各种模型进行分析和比较, 说明其各自的优缺点。
我们对 $\lambda_1, \lambda_2, \mu_1$ 和 μ_2 取值, 进行直观的数值分析, 表 1 为 $\lambda_1 > \lambda_2$ 时的结果, 表 2 为 $\lambda_1 < \lambda_2$ 时的结果 (为增加比较的全面性, 增加了排队模型 M1++ 和 M2++, 它们与 M1 和 M2 的唯一区别是将单路径的处理能力提高到 M3 等多路径模型中两条路径的和, 即 M1++ 和 M2++ 的结论与 M1 和 M2 一致, 只是对参数 μ 的选择不同, 前者 $\mu = \mu_1 + \mu_2$, 后者 $2\mu = \mu_1 + \mu_2$). 其中 M6 模型中低优先级数据包的处理无公式解, 需要基于上节 M6 模型分析方法求解.

表 1 $\lambda_1 > \lambda_2, \lambda_1 = 0.3, \lambda_2 = 0.1, \mu_1 = 0.9, \mu_2 = 0.5$ 时的数值结果

	高优先级数据包				低优先级数据包			
	到达率	服务率	平均队列长度	平均服务时间	到达率	服务率	平均队列长度	平均服务时间
M1++	0.3	1.4	2/5 (0.4)	1	0.1	1.4	2/5 (0.4)	1
M2++	0.3	1.4	3/11 (0.27)	10/11 (0.91)	0.1	1.4	7/55 (0.13)	14/11 (1.27)
M1	0.3	0.7	4/3 (1.33)	10/3 (3.33)	0.1	0.7	4/3 (1.33)	10/3 (3.33)
M2	0.3	0.7	3/4 (0.75)	5/2 (2.5)	0.1	0.7	7/12 (0.58)	35/6 (5.83)
M3	0.3	0.7	0.54	1.34	0.1	0.7	0.54	1.34
M4	0.3	0.9	1/2 (0.5)	5/3 (1.67)	0.1	0.5	1/4 (0.25)	5/2 (2.5)
M5	0.3	0.9	3/2 (1.5)	25/9 (2.78)	0.1	0.5	1/8 (0.125)	9/4 (2.25)
M6	0.3	0.9	0.38	1.25	0.1	0.5	0.25	2.5

表 2 $\lambda_1 < \lambda_2, \lambda_1 = 0.1, \lambda_2 = 0.3, \mu_1 = 0.9, \mu_2 = 0.5$ 时的数值结果

	高优先级数据包				低优先级数据包			
	到达率	服务率	平均队列长度	平均服务时间	到达率	服务率	平均队列长度	平均服务时间
M1++	0.1	1.4	2/5 (0.4)	1	0.3	1.4	2/5 (0.4)	1
M2++	0.1	1.4	1/13 (0.77)	10/13 (0.77)	0.3	1.4	21/65 (0.32)	14/13 (1.08)
M1	0.1	0.7	4/3 (1.33)	10/3 (3.33)	0.3	0.7	4/3 (1.33)	10/3 (3.33)
M2	0.1	0.7	1/6 (0.17)	5/3 (1.67)	0.3	0.7	7/8 (0.875)	35/9 (3.89)
M3	0.1	0.7	0.54	1.34	0.3	0.7	0.54	1.34
M4	0.1	0.9	1/8 (0.125)	5/4 (1.25)	0.3	0.5	3/2 (1.5)	5
M5	0.1	0.9	1/4 (0.25)	25/18 (1.39)	0.3	0.5	1/2	3
M6	0.1	0.9	0.115	1.15	0.3	0.5	1.5	5

从以上结果可以看出: M1++ 和 M2++ 两种模式在各种情况下性能都是最优的, M1++ 和 M2++ 是将 M1 和 M2 中单路径容量扩展一倍, 这需要提高下层网络设施的硬件处理能力, 而本文的设计是在现有网络能力下进行; 除去 M1++ 和 M2++ 两种模式, 优化的优先映射模式 M6 中高优先级的数据包在各种情况下都可以获得最高的处理能力, 而且当 $\lambda_1 > \lambda_2$ 时, 即使对于低优先级的数据包, M6 模式仍能提供较高的处理能力; 对于高优先级数据包, 优先多连接 M2 模式明显强于无优先多连接模型 M1, 这说明具有优先的映射可以提高用户对高优先数据获取的满意度; 另外可以发现即使高优先级数据包的处理效率也低于 M3, 还存在 M4 和 M5 两种简单优先设计, 这是因为 M4 和 M5 模型相当于将连接与路径作了静态绑定, 其带来的结果是路径不能充分利用, 而 M3 模型中任何一个连接的数据包只要找到空闲路径都可以发送. M6 经过优化提高了路径利用率, 同时也发挥了高优先级路径的优势, 因此可获得最好的高优先级数据包处理效用.

5 新构架对 QoS 的支持

新构架在 QoS 支持方面的作用主要体现在具有优先的映射中, 用户越紧迫需要的应用得到优先级越高的映射优先级, 获得高的完成效率, 用户从而获得更好的 QoS 体验. 首先引入基本效用评估模型用来将用户体验的 QoS 引入分析框架; 之后对这一模型进行增强, 增强模型可以将之前排队论的模型引入进行共同分析.

5.1 基本效用评估模型

定义效用函数的评估模型: 设集合 $P(S)$ 为服务 S 所用的路径, $P(s_n)$ 为子服务 $s_n \in S$ 所用的路径; $S(p)$ 为使用路径 p 的服务, $S_N(p)$ 为使用路径 p 的子服务. 显然,

$$P(S) = \{P(s_1), P(s_2), \dots, P(s_n)\},$$

$$S(p) = \{S_1(p), S_2(p), \dots, S_N(p)\},$$

而且, $p \in P(s_n)$ 当且仅当 $s_n \in S_N(p)$, 即 p 是子服务 s_n 所用路径当且仅当 s_n 是使用路径 p 的子服务.

当某个子服务 s_n 获得的处理能力(速率)为 x_{s_n} 时, 其效用函数为 $U_{s_n}(x_{s_n})$, 效用函数是一个服务或者子服务对于其所获得的处理能力(速率)的满意度, 处理能力越大则满意度也就越高. 当子服务 s_n 获

得的处理能力(速率)为 x_{s_n} 时, 设该子服务在路径 p 上获得的处理能力为 y_{ps_n} , 其中 $p \in P(s_n)$. 对某个子服务, $(A \cdot H)y = x$ 表示在各路径的一种处理能力(速率)分配方案, 写成分量形式为 $\sum_{p: p \in P(s_n)} y_{ps_n} = x_{s_n}$, 即支持子服务 s_n 的各条路径上分配的处理能力总和等于 x_{s_n} . 同时, $y = (y_{ps_n}, p \in P, s_n \in S)$ 还应该满足: $y_{ps_n} \geq 0$, $\sum_{s_n: s_n \in S_N(p)} y_{ps_n} \leq C_p$, 其中 C_p 是路径 p 的最大处理能力, 即在路径 p 上为所有子服务提供的处理能力总和不超过路径 p 的最大处理能力 C_p . 网络分层优先映射的主要目标是如何合理有效地将路径的处理能力分配给各个服务的各个子服务, 从而使得使用服务的所有网络用户效用函数的和达到最大值. 使服务 S 所属的网络用户获得最大满意度, 可通过构造以下的最优化问题求解:

$$\begin{aligned} & \max \quad \sum_s \sum_{s_n} U_{s_n}(x_{s_n}) \\ \text{P0: } & \text{subject to } \sum_{p: p \in P(s_n)} y_{ps_n} = x_{s_n}, \quad \sum_{s_n: s_n \in S_N(p)} y_p \leq C_p \\ & \text{over } \quad y_{ps_n} \geq 0 \end{aligned}$$

5.2 可分析排队论的增强效用评估模型

为了能将排队论模型和效用模型结合, 需要选择一个在优化模型和排队模型中都使用的变量进行分析. 我们选用数据包的传送时间作为效用函数的自变量, 它同时也是排队模型中数据包的处理时间. 这样效用函数的自变量将受限于排队模型的数据包到达率(因为排队模型中数据包服务时间受限于数据包到达率), 这就是两个模型分析的结合点.

$$\begin{aligned} & \max \quad \sum_s \sum_{s_n} \omega_{s_n} U_{s_n}(x_{s_n}) \\ \text{subject to } & \omega_{s_n} = F(x_{s_n}), x_{s_n} = f(\lambda_{s_n}) \\ \text{P1: } & \sum_{p: p \in P(s_n)} y_{ps_n} = x_{s_n}, \quad \sum_{s_n: s_n \in S_N(p)} y_{ps_n} \leq \mu_p N \\ & \text{over } \quad y_{ps_n} \geq 0 \end{aligned}$$

其中 $\omega_{s_n} = F(x_{s_n})$ 是关于 x_{s_n} 的单调递增函数, 反映了该子服务 s_n 的优先权大小, 即根据排队论模型中得到的 x_{s_n} 较大, 则该子服务的优先权也就较大, 其对应的效用也较大(为效用函数引入优先级, 同时与排队模型挂钩). $x_{s_n} = f(\lambda_{s_n})$ 是排队模型与效用模型结合的关键, 函数形式即为排队模型分析得到的各种模型中数据平均处理时间与到达率的关系.

$\sum_{p: p \in P(s_n)} y_p = x_{s_n}$ 表明某个子服务 s_n 在各个路径上获

得的处理能力之和等于 x_{s_n} . $\sum_{s_n: s_n \in S_N(p)} y_p \leq \mu_p N$ 中, N 为一个表示数据包大小的常量, 则 $\mu_p N$ 表示路径的传输能力. 下面证明该增强模型存在最优解(即存在最优的映射矩阵), 首先构造与问题 P1 等价的问题 P2.

P2:

$$\begin{aligned} \max \quad & \sum_s \sum_{s_n} F(f(\lambda_{s_n})) U_{s_n}(f(\lambda_{s_n})) \\ \text{subject to} \quad & \sum_{p: p \in P(s_n)} y_{ps_n} = f(\lambda_{s_n}) \\ & \sum_{s_n: s_n \in S_N(p)} y_{ps_n} \leq \mu_p N \\ \text{over} \quad & y_{ps_n} \geq 0 \end{aligned}$$

对于问题 P2, 构造 Lagrange 函数

$$L(x, r, k, l, q, \lambda) = \sum_s \sum_{s_n} F(f(\lambda_{s_n})) U_{s_n}(f(\lambda_{s_n})) + r_{s_n} \left(\sum_{p: p \in P(s_n)} y_{ps_n} - f(\lambda_{s_n}) \right) + \sum_p k_p \left(\mu_p N - \sum_{s_n: s_n \in S_N(p)} y_{ps_n} - \epsilon_p^2 \right),$$

其中, $\lambda = (\lambda_{s_n}, s_n \in S)$, $r = (r_{s_n}, s_n \in S)$, $k = (k_p, p \in P(S))$. r_{s_n}, k_p 是 Lagrange 因子, ϵ_p 是松弛变量. 则可以得到:

$$\begin{aligned} \frac{\partial L}{\partial f} &= F'(f) U_{s_n}(f) + F(f) U'_{s_n}(f) - r_{s_n}, \\ \frac{\partial L}{\partial \lambda_{s_n}} &= (F'(f) U_{s_n}(f) + F(f) U'_{s_n}(f) - r_{s_n}) \frac{df}{d\lambda_{s_n}}, \\ \frac{\partial L}{\partial y_{ps_n}} &= r_{s_n} - \sum_{p: p \in P(s_n)} k_p, \\ \frac{\partial L}{\partial \epsilon_p} &= -2\epsilon_p k_p. \end{aligned}$$

由上面的式子可以得到在最优点 $\lambda^* = (\lambda_{s_n}^*, s_n \in S)$ 处, 满足 $\partial L / \partial \epsilon_p = -2\epsilon_p k_p = 0$, 得到 $\epsilon_p = 0$ 或者 $k_p = 0$. 若 $k_p = 0$, 则约束 $\sum_{s_n: s_n \in S_N(p)} y_{ps_n} \leq \mu_p N$ 是不积极约束, 可以省略; 若 $\epsilon_p = 0$, 则约束 $\sum_{s_n: s_n \in S_N(p)} y_{ps_n} \leq \mu_p N$ 是积极约束. 只要存在至少一个服务, 则在最优点处该不等式约束就是积极约束(服务使用带宽的最大值以实现其最大的效用), 因此在下面的分析中, 假设 $\epsilon_p = 0$. 在最优点处, 满足

$$\begin{aligned} (F'(f) U_{s_n}(f) + F(f) U'_{s_n}(f) - r_{s_n}) \frac{df}{d\lambda_{s_n}} &= 0, \\ r_{s_n} - \sum_{p: p \in P(s_n)} k_p &= 0. \end{aligned}$$

由上式即可得到最优点 $\lambda^* = (\lambda_{s_n}^*, s_n \in S)$, 但是最优处理能力的分配 $y^* = (y_{ps_n}^*, p \in P, s_n \in S)$ 不一定唯一.

5.3 数值分析

为说明新构架对 QoS 支持的作用, 基于增强的效用评估模型, 使用 3.2 节排队分析的结论, 给出两组具体的数值比较. 根据文献[19]中弹性服务的性质选择对数函数 $\log(x+1)$ 作为效用函数. 根据 3.2 节的队列模型, 考虑两个服务, 服务 1 对效用函数的贡献大于服务 2(服务 1 的优先级高), 以下对服务 1 的效用函数为服务 2 的 4 倍和 10 倍两种情况进行分析(即 P1 问题中的 w_{s_n}). 为增加比较的全面性, 增加了排队模型 M1++ 和 M2++, 它们与 M1 和 M2 的唯一区别是将单路径的处理能力提高到 M3 等多路径模型中两条路径的和.

5.3.1 效用函数 4 倍关系

两个服务的效用函数分别为 $U_1 = 4\log(x_1 + 1)$ 和 $U_2 = \log(x_2 + 1)$, 总效用为 $V = U_1 + U_2$. 为了使效用函数与数据包的传输成正比关系, 使用 x_i 的倒数作为传输时间. 令 $\lambda_2 = 0.1, \mu_1 = 0.9, \mu_2 = 0.5$, 根据 3.2 节得到的各模型的 $E(S_i)$ 和 $E(L_i)$ 值, x_i 成为 λ_1 的函数. U_1, U_2 和 V 的分布如图 9~图 11. 由图 9 可知, M6 模型中高优先级数据包效率最高; M2 模型由于引入了优先级, 因此高优先级数据包效率高于 M1. 从图 10 可以看出 M2、M6 模型中低优先级数据包效率低于 M1、M3, 说明 M2、M6 模型中高优先级数据包效率的提高是以牺牲低优先级数据包为代价. 图 11 是各种模型的总效用, 由图可知当两类效用函数的差别不大时(4 倍), 由于低优先级数据的影响, 优先模型 M2、M6 中虽然高优先级数据效率很高, 但总效用可能不如一般模型(另外注意到模型 M1++ 和 M2++ 的效率总是很高, 说明若单网卡的能力与多网卡的和一样, 则单网卡效率更高, 理论分析见文献[15]中单服务器与多服务器分析. 然而这种情况超出了本文多路径的讨论范畴).

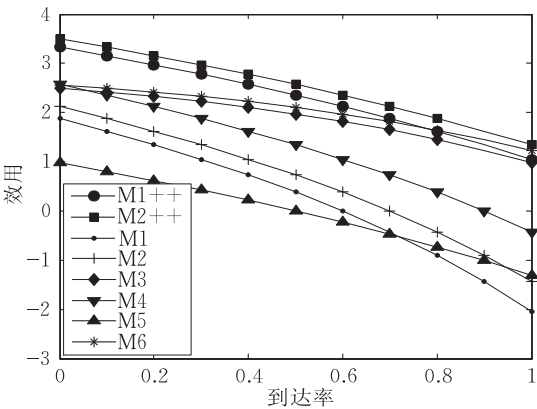


图 9 4 倍关系 U_1 分布

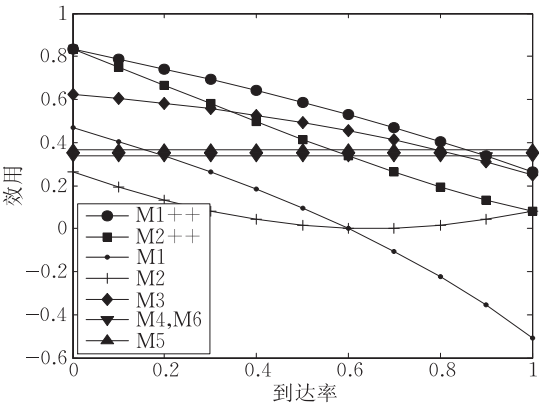


图 10 4 倍关系 U_2 分布

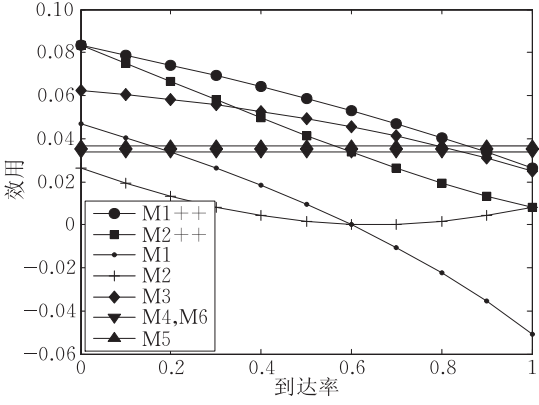


图 13 10 倍关系 U_2 分布

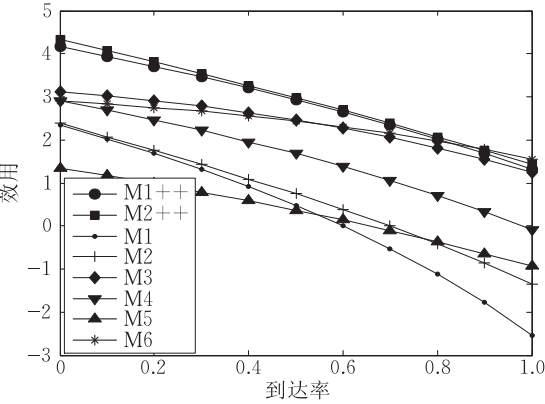


图 11 4 倍关系 V 分布

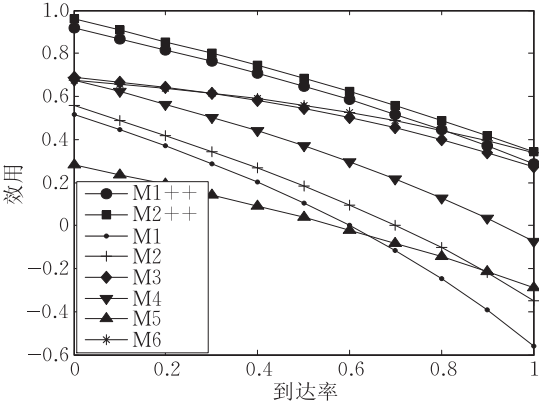


图 14 10 倍关系 V 分布

5.3.2 效用函数 10 倍关系

两类服务效用函数为 10 倍关系, $U_1 = \log(x_1 + 1)$ 和 $U_2 = 0.1 \times \log(x_2 + 1)$, 其它条件不变. U_1, U_2 和 V 的分布如图 12~14. 当高低优先级数据对用户的区别较大时(10 倍关系), 图 12 和图 13 的结果与 4 倍关系的分析一致, 即具有优先级的设计可以提高效用. 区别体现在图 14 中, 由于高优先级数据为总效用提升作了更大的贡献, 因此图 14 中 M6 模型的总效用超过了 M3 模型. 此时, 具有优先的设计为高优先级数据带来的效用和总效用都是最好的.

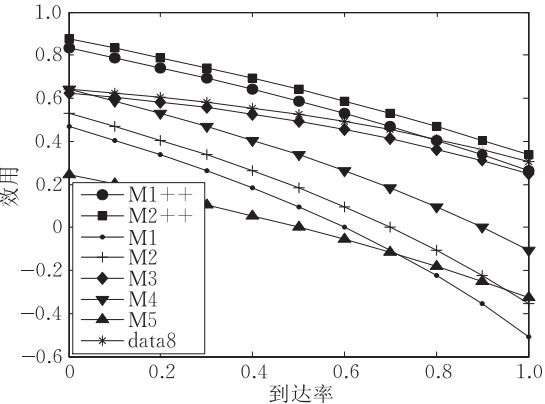


图 12 10 倍关系 U_1 分布

6 新构架在安全可靠方面的应用

6.1 新构架对安全性的支持

通信过程中的信息安全是新一代网络设计非常关注的问题. 基于多连接多路径映射的新传输层构架的安全性分析模型如下.

多连接多路径传输为服务的完成提供了冗余路径, 使得服务的完成具有较高的安全性和可靠性. 首先对于安全性, 网络中可能会有窃听者窃听路径上的数据, 以获得通信的数据内容, 多路径的传输使得安全性极大提高, 窃听者能够成功窃听并破解某个服务内容的几率下降.

假设网络中可以为服务提供数据传输的路径共有 n 条, 某个服务 s 使用多连接多路径技术进行数据传输, 路径数为 $m (\leq n)$ 条. 网络中的窃听者窃听了路径中的 $k (\leq m)$ 条, 想获得服务的相关信息, 除非窃听者窃听到了该服务所用到的所有路径的数据包, 否则该窃听者很难通过局部路径的数据包得到整个服务的机密信息. 因此, 该窃听者能够成功破解这个服务机密信息的概率为

$$P(n, k, m) = \begin{cases} C_k^m / C_n^m, & k \geq m \\ 0, & k < m \end{cases}$$

通过以上分析,当网络中可用的路径数较多时,或者某个服务 s 所使用的路径数较多时,窃听者想要窃听并获得该服务的一些信息内容的难度就会增大很多,所以使用多路径传输增强了数据传输的安全性。

6.2 新构架对可靠性的支持

当路径受到攻击(如 DoS 攻击)等破坏后就暂时不能再为服务提供数据传输,当攻击停止一段时间后,这条路径还能服务传输数据,即路径具有一定的可恢复性。以下我们假设在路径可恢复的前提下分析一下多路径传输的可靠性。

假设网络中可以为服务提供数据传输的路径共有 n 条,某个服务 s 使用多连接多路径技术进行数据传输,路径数为 $m(\leq n)$ 条。由于攻击(如 DoS 攻击)等破坏因素的存在,网络所能提供的 n 条路径中可能会有 $l(\leq n)$ 条断掉。对于某条路径,破坏持续一段时间 t_f 后,该路径就要断掉了,但是针对该路径的破坏或者故障停止一段时间 t_r 后,该路径又可以为服务提供数据传输,即该路径又恢复了可以传输数据的能力。假设 t_f 和 t_r 都是指数型随机变量,参数分别为 λ 和 μ ,即路径故障的到达过程和路径恢复的到达过程都是泊松过程。设因为攻击而正受到破坏的路径数为 $A_f(t)$,令 $\xi_j = Pr[A_f(t) = j]$,则

$$\xi_j = \frac{\frac{\rho^j}{j!(l-j)!}}{\sum_{k=0}^l \frac{\rho^k}{k!(l-k)!}},$$

其中 $\rho = \lambda/\mu$ 。

因此,当路径具有恢复能力时,利用多连接多路径传输协议传输数据时,某服务能够成功完成的概率为

$$Q = \sum_{j=0}^l \xi_j (1 - P(n+j-l, j, m)),$$

其中

$$P(n+j-l, j, m) = \begin{cases} C_j^m / C_{n+j-m}^m, & j \geq m \\ 0, & j < m \end{cases}$$

通过以上分析,若某个服务使用的路径数较多时,或者网络中断掉的路径数较少时,那么该服务能够成功完成的概率都是比较大的。因此,使用多路径传输提高了服务完成的可靠性。

7 新构架对移动性的支持

实现移动性支持的主要问题是切换过程中的数

据包丢失。在无线网络中,数据包丢失可以分成三类。第一类是拥塞造成的丢包,由无线节点的带宽和缓存受限引起溢出造成;第二类是差错造成的丢包,由噪声干扰、隐藏终端等造成;第三类是路径失效造成的丢包,由新旧路径切换过程中旧路径上的数据包丢失造成。在实际应用系统中,拥塞和差错造成的前两类丢包一般通过重复的 SACK 回应就可以发现;而路径失效造成的第三类丢包通过超时发现。根据传输协议丢包重传算法,超时出现是由于连续丢失数据包过多造成,而重复 SACK 回应是个别丢包造成。以上现象可以看出由于路径切换造成的数据包丢失通常是无线网络丢包重传的主要原因,好的切换技术可以极大地提高无线网络性能。

基于多连接多路径映射模型可以使移动终端在经过多个无线网络覆盖区域时,将处于活动状态的同一连接映射到多个路径,进行数据包的并行传输,这样将较好地避免切换过程中数据包的丢失。

8 原型系统

原型系统的传输协议除 TCP、UDP 外主要采用了 SCTP(使用开源项目 Lksctp^①),选择 SCTP 是基于其多流多宿主两个重要特性考虑,多流符合 2.1 节多连接的范围,多宿主经过扩展可以很容易地实现多路径并行传输。通过在 Lksctp 代码中添加多路径并行传输、流优先级、流与路径绑定 3 个功能模块,完成了 3.1 节提到的 3 种多路径映射在 SCTP 中的实现,这 3 种映射分别实现了安全可靠、QoS 支持和移动性。

在安全可靠方面,实现了基于简单多路径映射的文件下载,通过分割文件内容在不同路径传输降低文件被截获的可能。SCTP 本身并不支持多路径并行传输,我们首先扩展 SCTP 多宿主实现了简单映射的多路径并行传输。简单映射是指不同连接的数据到达后选择空闲的路径发送即可,从 5.3 节的数值分析结果可以看到这种模式效率很高(无路径闲置)。这种模式的主要问题是:不同路径间的传输延时差异使得接收端会产生接收乱序问题,导致接收端递交数据时多条路径相互制约。现有解决这一问题的研究主要有 SFR(Split Fast Retransmit)算法^[20],该算法将各个路径的缺失数据报告分离,使路径间的传输相互独立,避免了不必要的快速重传。

① Linux Kernel SCTP Project. <http://lksctp.sourceforge.net/>

但是由于各路径中数据的逻辑关系以及各路径使用同一接收缓存,仍无法很好地消除干扰。目前我们的改进是:不分离各路径的缺失数据报告,但增加协商功能,较好地消除相互干扰。

在 QoS 支持方面,实现了基于优先映射的 Web 浏览,提高了用户感兴趣数据的传输效率。首先通过添加用户协商模块和内核调度模块实现了 SCTP 的流优先功能。用户协商模块在结构体 `sctp_equal_streams` 和 `sctp_assocparams` 中添加了与优先级有关的成员,修改 SCTP_ASSOCINFO 套接口完成用户对流优先级设定的支持。内核调度模块通过扩展结构体 `sctp_asociation` 成员完成,主要包括多个优先调度队列,为了满足用户的不同需求,实现了两种不同的调度算法:优先级队列调度算法和加权循环队列调度算法。之后,我们将多流优先级特性应用到了 Web 浏览中。为了实现这一想法,首先将原先基于 TCP 的 Apache 和 Firefox 移植到 SCTP 上的,同时为了充分发挥 SCTP 的优势,改变 Apache 和 Firefox 的一些工作机制。最后将多流优先级引入到基于 SCTP 的 Apache 和 Firefox 中去。

在移动性方面,实现了基于绑定映射的 WLAN 与 CDMA 之间的移动切换。`mSCTP`^[21]是 SCTP 解决移动切换的版本,它基于 SCTP 多宿主特性和动态地址重配置扩展 DAR 技术^[22]实现。然而 `mSCTP` 工作在现有单路径下,切换时仍存在由于路径失效引起的丢包问题。为此我们首先实现了基于绑定的多路径映射,主要通过简单多路径映射基础上添加了 4 项新功能:路径独立接收缓存、完全面向路径的拥塞控制、限定流与路径的绑定关系(流与路径可以是多对一,但不允许一对多)、引入面向路径的数据包标识符 PSN(不再使用 SCTP 原有的 TSN、SSN)。在实现了绑定的多路径映射基础上,构造了一个 WLAN 与 CDMA 之间的移动切换的原型实现。

9 结束语

根据新一代网络对普适服务的要求,本文给出一个传输层新构架,可用来提高现有网络在 QoS 支持、安全可靠、移动性、可控可管等方面的能力,基于该构架的原型系统实现证明了新设计的正确性和可行性。但是该构架及原型还有大量需要深入研究的工作。首先,缺乏与传输协议中流量控制、拥塞避免等算法的结合分析;其次,原型系统还有一些问题,主要是简单映射多路径并行传输中不同路径间数据包的乱序问题,绑定映射不存在这一问题,优先映射

在路径差别较大时也不明显;再次,为使原型系统更加真实,使用了广域网模拟器 Nistnet,但仍缺乏在大规模真实网络中的测试;最后,原型系统在 SCTP 中实现了多连接路径映射,还需要将多连接跨越不同传输协议。

参 考 文 献

- [1] Lin Chuang, Peng Xuehai. Research on trustworthy networks. Chinese Journal of Computers, 2005, 28(5): 751-758(in Chinese)
(林闯, 彭雪海. 可信网络研究. 计算机学报, 2005, 28(5): 751-758)
- [2] Zhang Hong-Ke, Su Wei. Fundamental Research on the architecture of new network — Universal network and pervasive services. Acta Electronica Sinica, 2007, 35(4): 593-598 (in Chinese)
(张宏科, 苏伟. 新网络体系基础研究——一体化网络与普适服务. 电子学报, 2007, 35(4): 593-598)
- [3] Stewart R et al. Stream control transmission protocol. RFC 2960, Internet Engineering Task Force, Oct. 2000
- [4] Kohler E, Handley M, Floyd S. Designing DCCP: Congestion control without reliability//Proceedings of the SIGCOMM. Pisa, Italy, 2006: 27-38
- [5] Sivakumar H, Bailey S, Grossman R L. PSockets: The case for application-level network striping for data intensive applications using high speed wide area networks//Proceedings of the SC2000: High-performance network and computing conference. Dallas, TX, 2000: 38
- [6] Eggert L, Heidemann J, Touch J. Effects of ensemble-TCP. ACM Computer Communication Review, 2000, 30(1): 15-29
- [7] Balakrishnan H, Rahul H, Seshan S. An integrated congestion management architecture for Internet hosts//Proceedings of the ACM SIGCOMM. Cambridge, MA, USA, 1999: 175-187
- [8] Allman M, Kruse H, Ostermann S. An application-level solution to TCP's satellite inefficiencies//Proceedings of the 1st International Workshop on Satellite-based Information Services (WOSBIS). Rye, New York, 1996
- [9] Baru C, Moore R, Rajasekar A, Wan M. The SDSC storage resource broker//Proceedings of CASCON '98. Toronto, Canada, 1998: 5
- [10] Floyd S. Congestion Control Principles. RFC 2914, IETF, September 2000
- [11] Zhang M, Lai J, Krishnamurthy A, Peterson L, Wang R. A transport layer approach for improving end-to-end performance and robustness using redundant paths//Proceedings of the USENIX. Boston, MA, USA, 2004: 8
- [12] Iyengar J R, Shah K C, Amer P D, Stewart R. Concurrent multipath transfer using SCTP multihoming//Proceedings of the SPECTS 2004. San Jose, CA, USA, 2004

- [13] Iyengar J R, Amer P, Stewart R. Concurrent multipath transfer using SCTP multihoming over independent end-to-end paths. *IEEE/ACM Transactions on Networking*, 2006, 14(5): 951-964
- [14] Al A A E, Saadawi T, Lee M. A transport layer load-sharing mechanism for mobile wireless hosts//*Proceedings of the 2nd IEEE Annual Conference on Pervasive Computing and Communications Workshops*. Orlando, FL, USA, 2004: 87-91
- [15] Lin Chuang. *Computer Network and Computer System Performance Evaluation*. Beijing: Tsinghua University Press, 2001(in Chinese)
(林闯. 计算机网络和计算机系统的性能评价. 北京:清华大学出版社, 2001)
- [16] Osogami T, Wierman A, Harchol B M, Scheller W A. A recursive analysis technique for multi-dimensionally infinite Markov chains. *ACM SIGMETRICS Performance Evaluation Review*, 2004, 32(2): 3-5
- [17] Bobbio A, Horváth A, Telek M. Matching three moments with minimal acyclic phase type distributions. *Communications in Statistics-Stochastic Models*, 2005, 21(2): 303-326
- [18] Neuts M F. *Matrix-analytic methods in queueing theory//Advances in Queueing. Theory, Methods, and Open Problems*. CRC Press, 1995: 265-293
- [19] Shenker S. Fundamental design issues for the future Internet. *IEEE Journal on Selected Areas in Communications*, 1995, 13(7): 1176-1188
- [20] Iyengar J, Caro A, Amer P, Heinz G, Stewart R. Making SCTP more robust to changeover//*Proceedings of the SPEC-TS 2003*. Montreal, Canada, 2003
- [21] Stewart R, Xie Q et al. Stream control transmission protocol (SCTP) dynamic address reconfiguration. RFC 5061, IETF, September 2007
- [22] Koh S, Chang M J, Lee M. mSCTP for soft handover in transport layer. *IEEE Communications Letters*, 2004, 8(3): 189-191



YANG Dong, born in 1980, Ph.D.. His research interests include next generation internet research, pervasive services and peer-to-peer networks.

LI Shi-Yong, born in 1982, Ph.D. candidate. His re-

search interests include next generation internet research, peer-to-peer networks and queueing theory.

WANG Bo, born in 1982, Ph.D. candidate. His research interests include wireless Ad hoc network.

ZHANG Hong-Ke, born in 1957, Ph.D., professor, Ph.D. supervisor. His research interests include new generation network theory, wireless sensor networks and router theory.

Background

There are a lot of researches on the new generation network, such as 4G, NGN, NGI in the abroad and Trustworthy Networks, Universal Network in China. Most of these researches think that providing pervasive service is one main objective of the new generation network, and another is constructing new packet network. Through what is and how to implement pervasive is in the air, to provide different kinds of services and accesses is accepted as the main goal of pervasive service.

The project is supported by the National Basic Research Program (973 Program) of China under grant

No. 2007CB307100. The long-term goal is to build universal network and pervasive service. The main contribution of this paper is proposing a new transport layer architecture which is suitable for pervasive service. From the perspective of transport practice, different kinds of services are based on multiple connections, and different kinds of accesses are based on multiple paths (network access device). The new architecture is to build different relationship of connections and paths, called mapping. Based on the mapping technology, the new architecture could improve the QoS, reliable, security, mobility and controllable of new network.