

IPv6 Overlay 网络多路径流量均衡实现机制研究

朱尚明, 储赞, 王雅兰

(华东政法大学, 上海 201620)

摘要: 针对基于 IPv6 的 Overlay 网络, 分析了其多重连接和多路径路由两项关键技术, 根据逻辑功能, 提出了 Overlay 网络节点的内部体系结构, 重点对 Overlay 网络多路径流量均衡机制进行了研究, 提出了一种按路径带宽进行流量均衡的简化方案并进行了设计实现。实验结果表明, 多路径流量均衡算法传输数据的效率明显优于单路径算法, 且随着数据分组的增大, 其优势更加明显。

关键词: Overlay 网络; 多重连接; 多路径路由; 流量均衡

中国法分类号: TP393

文献标识码: A

doi: 10.11959/j.issn.2096-109x.2018045

Research on the implementation mechanism of multipath traffic proportion of IPv6 Overlay network

ZHU Shangming, CHU Yun, WANG Yalan

East China University of Political Science and Law, Shanghai 201620, China

Abstract: Based on IPv6 overlay network, two key technologies, multihoming and multipath routing were analyzed. According to logic functions, architectures of overlay nodes were presented. The mechanism of multipath traffic proportion of IPv6 overlay network was discussed on emphasis. A simplified scheme of traffic proportion according to path bandwidth was proposed, and the design and implementation of the scheme were carried out. Experimental results show that the efficiency of the algorithm of multipath traffic proportion is much higher than that of single path, and the advantage is more obvious with increasing data packets.

Key words: Overlay network, multihoming, multipath routing, traffic proportion

1 引言

近年来, Internet 已发展成为一个具有高度路由冗余的网络^[1]。但是, 当前的 Internet 底层路由协议, 并不能完全利用这些路由冗余来提高网络的性能和可用性。当网络链路或路由器发生故障时, 底层路由协议无法快速地做出反应并恢复。另外, 由于当前底层路由协议的灵活性有限, 很

难识别和避开发生了拥塞的瓶颈链路。因此, 网络链路一旦出现故障或拥塞, 端到端的网络应用将会受到很大影响。

Overlay 网络是架构在 Internet 之上的一个虚拟的应用网络, 通常实施在网络体系结构的应用层, 并独立于底层的 IP 网络。Overlay 网络是一种能够利用路由冗余提高 Internet 性能的有效机制, 构成一个 Overlay 网络的终端主机和路由器称作

收稿日期: 2018-04-15; 修回日期: 2018-05-20

通信作者: 朱尚明, zhushm@ecupl.edu.cn

基金项目: 赛尔网络下一代互联网技术创新基金资助项目 (No.NGII20160102)

Foundation Item: CERNET Innovation Project (No.NGII20160102)

Overlay 网络节点, Overlay 网络节点间的一条 IP 网络路径称作 Overlay 链路。Overlay 网络技术也称为一种能根据数据内容进行处理的高层节点构成的网络技术。这些 Overlay 节点通过遵循特有的协议和现有的互联网进行连接。由于构建一个 Overlay 网络并不需要改变底层 IP 网络的体系结构,但能够提升网络性能,如提高网络的可靠性、保证网络的 QoS (服务质量)、平衡网络负载等。因此目前在内容分发、文件共享、远程视频会议、远程教育、远程医疗、VoIP 以及网络电视等多媒体通信领域有着广泛的应用^[2-4]。随着 IPv6 下一代互联网技术的普及应用,研究 IPv6 环境下 Overlay 网络的体系结构和多路径路由技术具有重要的意义和应用价值。

2 IPv6 Overlay 网络的关键技术

2.1 多重连接技术

多重连接是指一个网络站点或主机连接到2个以上 Internet 服务商,是 Internet 上的一个基本组件服务。多重连接的主机或网络站点通过连接到2个以上的互联网服务商 (ISP),提供了一种维持 Internet 连接的机制,提高了连接的可靠性,并可以将负载路由到多个 Internet 连接上。当其中一个 ISP 连接失败时,Internet 连接可以通过另外的 ISP 继续维持。另外,将负载路由到多个 Internet 连接上,可以均衡网络流量,平衡网络负载。一个简单的多重连接站点如图1所示。

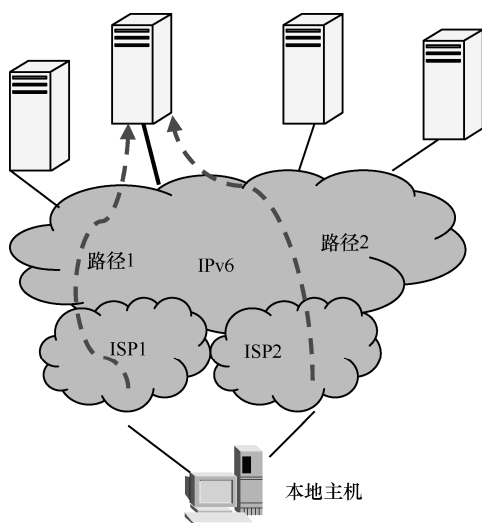


图1 多重连接示意

IPv6 由于支持多地址,具有动态监测和获得多个 IP 地址的能力,因此相比 IPv4,更适合部署多重连接。尽管如此,IPv6 并不具备一套解决多重连接的完整机制,根据 RFC4177 定义的体系结构,基于 IPv6 的多重连接解决方案可分为如下5种方式:路由方式、移动方式、插入新协议子层方式、修改协议子层方式、修改边界路由器和本地主机之间的交互方式。

根据解决方案的工作层次,基于 IPv6 的多重连接还可以分为如下3种机制:主机层次 (host-level) 的多重连接、站点层次 (site-level) 的多重连接、服务器层次 (provider-level) 的多重连接,其中,主机层次的多重连接对域内路由结构的影响最小,每个站点的地址前缀可以聚合在一个给定的提供者的地址前缀范围内。但需要解决以下4个问题。

- 1) 目的地址选择问题。
- 2) 源地址选择问题。
- 3) 边界路由器流入分组的过滤问题。
- 4) 快速响应网络拓扑的变化问题。

到目前为止,还没有一种单一的机制可以满足多重连接的所有目标和功能,实际应用中往往采用几种机制相结合的方式来实现所要求的目标和功能。例如,针对小型站点,可以使用基于主机的外部路径选择和基于源地址的路由机制。

2.2 多路径路由技术

多路径路由技术是 Overlay 网络构建的另一项关键技术^[5]。对于进入 Overlay 网络的业务流,传统的路由算法,如 Dijkstra 最短路径算法和 Bellman-Ford 距离矢量算法,都是在源-目的节点间计算出一条最佳路径,所有业务量都通过这条最佳路径传送,容易造成网络拥塞,而且没有明确的时间和可靠性保障。最短最宽路径 (SWP, shortest widest path) 算法和最宽最短路径 (WSP, widest shortest path) 算法等虽然能保证一定的服务质量,但是仍然使用单一路径进行路由。最佳路径路由算法可以获得最佳的传输路径,但这些被选路径由于常常被选上而负载过重,致使网络性能下降。多路径 QoS 路由技术能克服传统最佳路径算法的不足,为业务流选择多个候选路径,并在这些路径中均衡通信流量,从而提高网络的可靠性,降低网络拥塞,减少传输时延和提高链

路利用率。目前虽然已提出了多种多路径 QoS 路由解决方案^[6],但这些方案都是基于 IPv4 网络的。由于 IPv6 Overlay 网络是架构在 IPv6 网络之上的一个虚拟的应用网络,其多路径选择和路由技术与 IPv4 网络相比有很多特殊之处。针对 IPv6 Overlay 网络,研究其多路径 QoS 路由技术,应从以下几个方面展开。

1) 底层 IPv6 网络路径的性能探测技术

IPv6 Overlay 网络中的一条链路映射为底层 IPv6 网络的一条路径,IPv6 网络的路径决定了 Overlay 网络中链路的服务质量以及是否可能与其他链路物理上相交。而 IPv6 网络路径的性能(如时延、可用带宽和分组丢失率)是随着网络负载动态变化的。以较低的开销对这些参数进行动态实时地测量是比较困难的。目前通过 Ping 或 TraceRoute 工具可以获得 IPv6 网络节点之间的传输时延和分组丢失率,但要获得当前一条路径的可用带宽,往往需要较大的开销。

2) 多路径 QoS 路由算法

通过采用多路径 QoS 路由技术可以减少源-目的节点之间多媒体通信的时延和分组丢失率,提高端到端的传输质量和可靠性。如何针对时延和带宽约束,动态计算满足 QoS 条件的最优 IPv6 多路径集合,且使所计算的多个物理路径在瓶颈链路上互不相交,是学者研究的热点。多路径 QoS 路由算法可以通过弥补 Internet 传输质量的不足,减少对底层 IP 网络和 ISP 的依赖,来提高端到端的服务质量。

3) 多路径流量均衡技术

由于多路径集合中每条路径的服务质量不尽相同,如何根据 Overlay 网络多路径的服务质量和当前流量负载,来确定多路径集合中每个路径的流量比例系数,以使整个网络的性能达到最优,这是一个典型的最优化问题,并且其求解会随着路径性能和流量负载的动态变化而变化。流量均衡技术就是研究如何根据路径的时延、带宽和分组丢失率来计算各个路径上的流量比例系数,以使网络负载平衡,且使端到端的时延最低。

国内外学者在多路径流量均衡方面已有不少研究^[7,8],但这些研究方案都不能直接应用到 IPv6 Overlay 网络上,因此本文针对基于 IPv6 的

Overlay 网络,对其多路径流量均衡机制进行探索,提出一种按路径带宽进行流量均衡的简化方案并进行设计实现。

3 IPv6 Overlay 网络的体系结构

3.1 系统结构

IPv6 Overlay 网络的体系结构构建在 Overlay 网络节点的应用层。根据逻辑功能将 Overlay 节点分为发送节点、中继节点和接收节点 3 种类型,如图 2 所示。发送节点和接收节点是端节点,可采用多重连接机制连接到 Internet 上;中继节点是服务节点,部署在发送节点和接收节点之间。

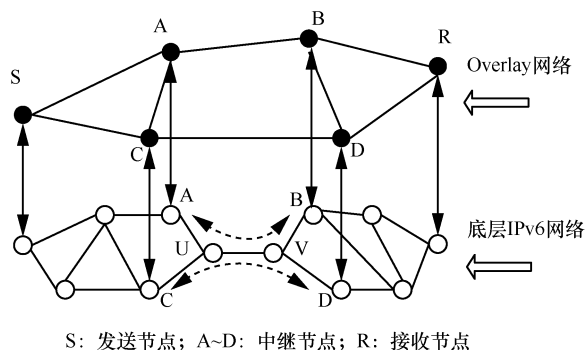


图2 Overlay网络示意

3.2 节点内部体系结构

Overlay 网络节点的内部体系结构如图 3 所示。发送节点用于发送数据,由流分类器、IPv6 网络性能探测、链路状态交换、不相交多路径选择、流量均衡分配、前向纠错编码和分组发送几个功能模块构成;中继节点用来转发 Overlay 网络中的数据分组,由包接收、IPv6 网络性能探测、链路状态交换、不相交多路径选择、流量均衡分配和分组转发功能模块构成;接收节点用于接收多媒体数据,由分组接收、前向纠错解码、流重组构成。

4 多路径流量均衡实现机制

4.1 设计思想

多路径流量均衡技术的核心思想是根据路径的性能,计算各个路径的流量比例系数,这是一个典型的最优化问题。可采用路径的时延、可用带宽和分组丢失率来简化流量比例系数的计算,对于一个路径,其流量比例系数和当前可用带宽成正比,与时延和分组丢失率成反比。简单起见,

本文以 2 条路径为例, 并仅采用路径带宽作为流量均衡计算的依据。

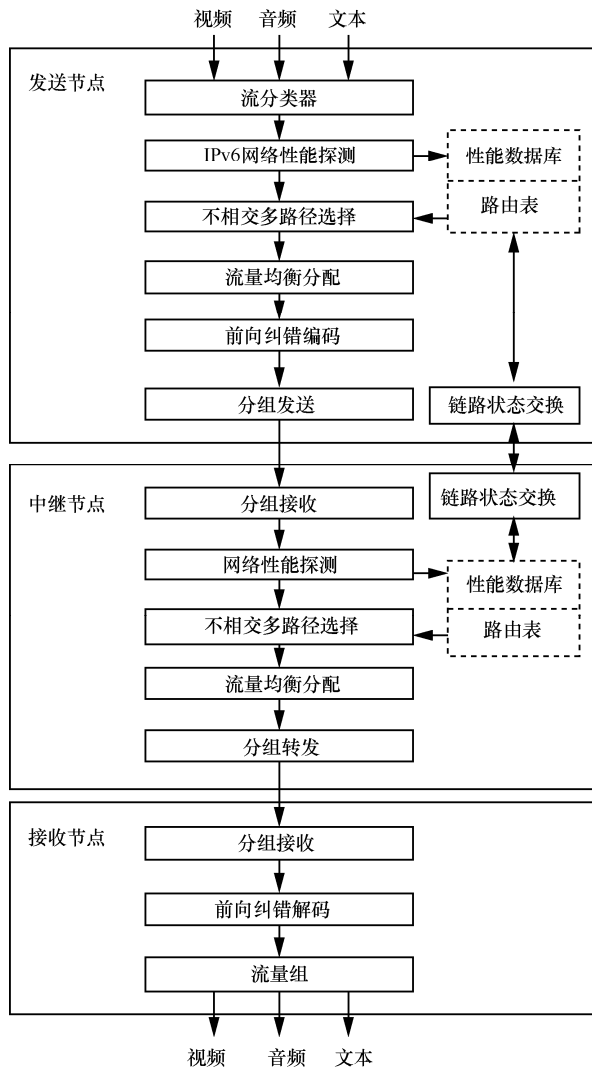


图 3 Overlay 节点内部体系结构

假设 $B1$ 和 $B2$ 分别为路径 1 和路径 2 的带宽, $p1$ 和 $p2$ 分别为路径 1 和路径 2 的流量比例系数, 则计算可以简化为

$$p1 = \frac{B1}{B1 + B2} \quad (1)$$

$$p2 = 1 - p1 \quad (2)$$

在发送端, 对数据流进行分块, 并打上块号标记, 然后根据数据块数量和比例系数 $p1$ 、 $p2$, 分别在路径 1、路径 2 这 2 条路径上按比例均衡分发; 在接收端, 对路径 1 和路径 2 这 2 条路径上的数据块进行接收, 并根据块号进行信息重组, 恢复为原来的数据流。

4.2 实现机制

本文通过套接字 (Socket) 利用 C# 在 IPv6 环境下对多路径流量均衡进行技术实现和模拟验证。

4.2.1 发送端

发送端 (Client) 作为客户端, 套接字初始化如图 4 所示, 首先设置好 IPv6 地址和端口, 然后设置套接字和发送缓冲区大小, 并向接收端发起连接请求。

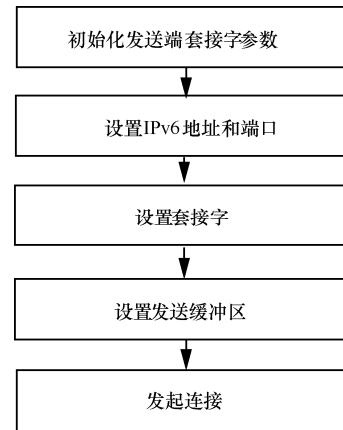


图 4 初始化发送套接字

发送数据前, 发送端首先根据单个数据块的长度 (sendLength) 计算出总的 数据块数量 (blockNumber), 然后根据路径 1 和路径 2 的带宽值计算相应的流量 $p1$ 和 $p2$ 。对每个已读入发送缓冲区 (buffer) 的当前数据块, 发送时都在数据块前面加上一个 5 byte 的数组数据, 其中 4 个字节用来存放数据块的块号, 1 个字节用来存放数据块结束标志。

发送数据时, 本文设计了 2 种流量均衡方式, 一种是通过产生一个随机数 p 进行流量均衡, 随机发送算法如图 5 所示。

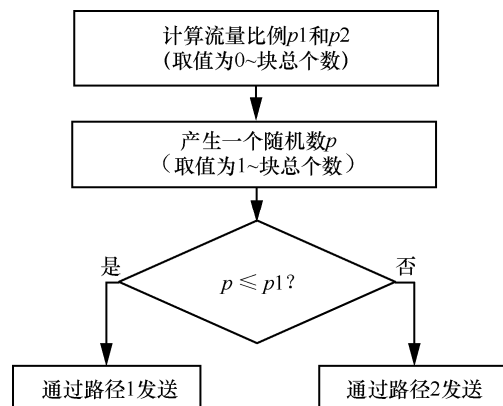


图 5 随机发送算法

首先计算出路径 1 和路径 2 对应的流量比例系数 p_1 和 p_2 , 然后系统产生一个在 1 和文件块总数之间的随机数 p , 根据随机数的大小选择路径 1 ($p \leq p_1$ 时) 或路径 2 ($p > p_1$ 时) 进行发送, 且该过程循环进行, 直至所有文件块都发送完。

另一种方式是按数据块块号的个位数进行按比例顺序发送, 顺序发送算法如图 6 所示。

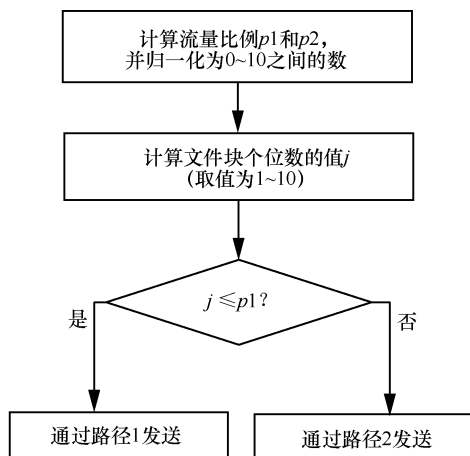


图6 顺序发送算法

首先计算出路径 1 和路径 2 对应的流量比例系数 p_1 和 p_2 , 并归一化为 0~10 的一个数, 然后计算文件块个位数的值 j , 根据该值的大小选择路径 1 ($j \leq p_1$ 时) 或路径 2 ($j > p_1$ 时) 进行发送, 且该过程循环进行, 直至所有文件块都发送完。

4.2.2 接收端

接收端 (Server) 作为服务器端, 套接字初始化及接收流程如图 7 所示。接收端套接字初始化方法同发送端类似, 但建立套接字后需开启侦听模式, 启动接收线程, 然后接收并重组数据。

接收时, 首先定义一个消息数组 `message`, 用来存放接收到的数据块, 这个消息数组的成员变量 `Num` 用来存放接收到的数据块号, `Flag` 用来存放数据是否结束的标志, `content` 用来存放数据块内容。

2 个路径的数据分组都接收完毕后, 还需对数据块进行合并、排序和重组, 恢复为原来的数据流。最后创建文件并将重组后的数据块写入文件。

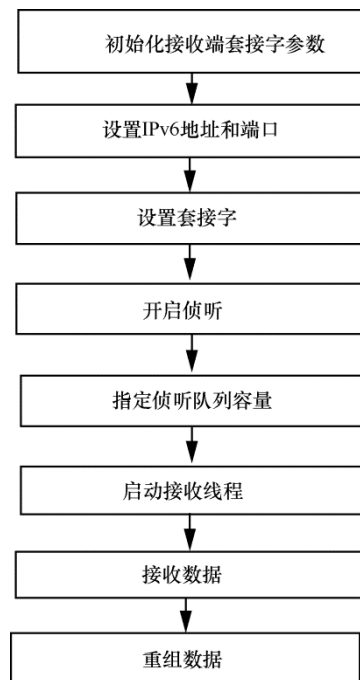


图7 数据接收流程

5 实验及结果分析

本文对上述多路径流量均衡的发送和接收算法在 IPv6 环境下进行了测试验证, 并和单路径传输进行对比分析。

5.1 实验环境

发送端为华东政法大学长宁校区一台主机, IPv6 地址为 2001:da8:8020:9b:515d:1e8:945c:2678; 接收端为松江校区一台主机, 配置 2 块网卡通过多重连接技术接入 IPv6 网络, 从而产生 2 条相对独立的 IPv6 路径, 2 块网卡的 IPv6 地址分别为 2001:da8:8020:3:48f6:5daa:eade:e120 和 2001:da8:8020:3:5792:236c:5866:157, 其中, 后缀为 e120 的 IPv6 地址同时作为单路径地址进行对比测试。

由于 2 块网卡的性能基本相似, 可以假定 2 条路径的带宽基本相同, 即 $B_1=B_2$ (每个路径的流量比例系数均为 0.5), 文件块大小设置为 32 kB (即 $32 \times 1024 = 32768$ byte)。

为了和单路径进行性能对比, 本文采用传输时间作为度量参数。为消除发送端和接收端的主机时钟差别, 在发送端和接收端分别计算多路径和单路径情况下的传输时间差, 具体设定如下。

1) 单路径情况下, 假设发送一个文件的时刻为 t_1 , 接收端接收完该文件的时刻为 t_3 , 那么其

传输时间为 $t_3 - t_1 \pm \Delta$ (Δ 为发、收端主机的时钟之差)。

2) 多路径情况下, 假设发送一个文件的时刻为 t_2 , 接收端接收完该文件的时刻为 t_4 , 那么其传输时间为 $t_4 - t_2 \pm \Delta$ (Δ 为发、收端主机的时钟之差)。

3) 单路径与多路径传输的耗时之差为 $T = (t_3 - t_1 \pm \Delta) - (t_4 - t_2 \pm \Delta) = (t_2 - t_1) - (t_4 - t_3)$ 。

即发送端时间差 $t_2 - t_1$ 减去接收端时间差 $t_4 - t_3$, 就是在传输同一个文件时单路径比多路径所多耗费的时间差值, 这样就消去了发送、接收主机之间时钟之差的影响。

5.2 实验结果

为进行测试对比, 对本文提及的通过产生随机数实现流量均衡的随机多路径发送算法和按数据块块号的个位数进行按比例发送的顺序多路径发送算法都进行了实际测试。

1) 随机多路径均衡

对随机多路径发送, 通过产生随机数来实现流量均衡和单路径传输的实验对比结果如表 1 所示。从发送端依次传输 10 个 jpg 图片文件到接收端, 10 个传输文件大小从 100 kB 到 13 MB 逐渐增大。

表 1 随机多路径和单路径传输耗时对比

文件序号	大小/块	$t_2 - t_1$	$t_4 - t_3$	T/ms
1	4	6 507	6 360	147
2	27	23 007	22 339	668
3	39	24 820	23 812	1 008
4	70	23 599	21 693	1 906
5	113	64 633	61 272	3 361
6	132	31 900	27 994	3 906
7	163	32 932	27 900	5 032
8	249	56 155	48 614	7 541
9	324	55 606	46 046	9 560
10	414	43 576	30 918	12 658

2) 顺序多路径均衡

对顺序多路径发送, 通过数据块块号的个位数进行按比例发送和单路径传输的实验对比结果如表 2 所示。从发送端依次传输 10 个 jpg 图片文件到接收端, 10 个传输文件大小从 100 kB 到 13 MB 逐渐增大。

表 2 顺序多路径和单路径传输耗时对比

文件序号	大小/块	$t_2 - t_1$	$t_4 - t_3$	T/ms
1	4	15 401	15 372	29
2	27	15 710	15 076	634
3	39	93 753	92 815	938
4	70	21 409	19 673	1 736
5	113	135 663	132 943	2 720
6	132	26 160	22 957	3 203
7	163	20 700	16 745	3 955
8	249	31 821	25 855	5 966
9	324	119 323	111 551	7 772
10	414	44 261	34 312	9 949

从表 1、表 2 中可以看出, 2 种算法下多路径耗时都明显比单路径耗时减少, 且随着文件大小的增大, 其节省的时间也线性增大, 采用流量均衡后的多路径传输明显优于单路径传输。

3) 随机多路径和顺序多路径耗时对比

随机发送 (T_1) 和顺序发送 (T_2) 2 种流量均衡算法耗时之差对比如表 3 所示。

表 3 随机发送和顺序发送多路径传输耗时对比

文件序号	大小/块	T_1/ms	T_2/ms	$T_1 - T_2/\text{ms}$
1	4	147	29	118
2	27	668	634	34
3	39	1 008	938	70
4	70	1 906	1 736	170
5	113	3 361	2 720	641
6	132	3 906	3 203	703
7	163	5 032	3 955	1 077
8	249	7 541	5 966	1 575
9	324	9 560	7 772	1 788
10	414	12 658	9 949	2 709

以文件大小 (块数量) 为 x 轴, 耗时差值 T 为 y 轴, 2 种算法下, 单路径和多路径的耗时差值曲线如图 8 所示。

从表 3 和图 8 中可以看出, 随机多路径均衡和顺序多路径均衡算法传输文件时比单路径算法所节省的时间随文件的增大而线性增加, 即其传输所耗费的时间更少, 传输效率更高, 验证了 2 种算法均优于单路径传输机制。且从图 8 中还可以

看出,对于多路径发送,通过产生随机数来实现流量均衡比通过按数据块块号的个位数进行顺序发送所节省的时间更多,算法效果更佳。

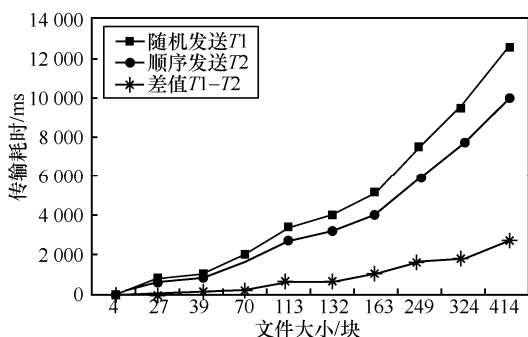


图8 随机多路径和顺序多路径传输耗时对比

6 结束语

本文探讨了 Overlay 节点的体系结构和流量均衡的具体实现机制,由于 Overlay 网络能够充分利用底层 IP 网络的冗余性,在提高数据传输的可靠性、提高链路带宽利用率以及平衡网络负载方面具有明显的技术优势,随着 IPv6 下一代互联网技术的深入应用,基于 IPv6 的多路径传输技术能够减少时延和提高音频流、视频流的服务质量,在文件传输、内容分发和多媒体通信和相关领域有着广阔的应用前景。

参考文献:

- [1] ARJONA-VILLICANA P, CONSTANTINOU C, STEPANENKO A. The Internet's unexploited path diversity[J]. IEEE Communications Letters, 2010, 14(5): 474-476.
- [2] BAI H, SHABAN K, KHODEIR M, et al. Overlay network

scheduling design[J]. Computer Communications, 2016(82): 28-38.

- [3] GRANDA J, NUÑO P, MOLLEDA J, et al. Resilient overlay network for real-time interactive multimedia sessions in corporate networks[J]. Multimedia Systems, 2016, 22(5): 543-557.
- [4] RIDHAWI Y A, KARMOUCH A. QoS-based composition of service specific overlay networks[J]. IEEE Transactions on Computers, 2015, 64(3): 832-846.
- [5] CHEN Y R, RADHAKRISHNAN S, DHALL S, et al. The service overlay network design problem for interactive internet applications[J]. Computers and Operations Research, 2015, 57: 73-82.
- [6] SAHHAF S, TAVERNIER W, COLLE D, et al. Adaptive and reliable multipath provisioning for media transfer in SDN-based overlay networks[J]. Computer Communications, 2017, 106: 107-116.
- [7] HESMANS B, TRAN-VIET H, SADRE R, et al. A first look at real multipath TCP traffic[C]//The 7th International Workshop on Traffic Monitoring and Analysis. 2015: 233-246.
- [8] GAO C, LING Z H, YUSN Y F. Packet reordering analysis for concurrent multipath transfer[J]. International Journal of Communication Systems, 2014, 27(12): 4510-4526.

[作者简介]



朱尚明 (1969-), 男, 河南虞城人, 博士, 华东政法大学研究员, 主要研究方向为计算机网络、计算机应用、信息安全。

储赞 (1981-), 女, 江苏常州人, 华东政法大学工程师, 主要研究方向为计算机网络、计算机应用技术。

王雅兰 (1993-), 女, 山西临汾人, 华东政法大学硕士生, 主要研究方向为网络安全、计算机取证、司法鉴定。