

SDN/NFV架构下基于MPTCP协同SDN的空间网络资源调度技术

贾 敏, 郭 家, 徐媚琳, 郭 庆, 顾学迈

哈尔滨工业大学电子与信息工程学院, 黑龙江 哈尔滨 150006

摘 要: 空间网络主要由各种架构的卫星网络组成, 组成的网络节点众多, 相比于地面网络, 卫星网络中的节点难以观测且不易控制, 同时网络设备极其昂贵。用户量激增给网络带来极大的负担, 同时导致网络资源无法得到合理的利用, 将传统网络架构应用在大规模的卫星网络中导致其无法灵活地配置网络, 资源分配不均。针对以上问题, 提出基于软件定义网络(SDN)/网络功能虚拟化(NFV)架构的空间网络资源调度系统, 实现空间网络资源的实时监测, 同时将多路径TCP(MPTCP)和SDN相结合, 完成对空间网络的灵活配置以及其资源的合理利用, 有效提高了网络的吞吐率。

关键词: 空间网络; SDN; 资源调度; MPTCP

中图分类号: TN927.2

文献标识码: A

doi: 10.11959/j.issn.2096-8930.20200110

Spatial Network Resource Scheduling Based on Joint MPTCP with SDN for SDN/NFV Architecture

JIA Min, GUO Jia, XU Meilin, GUO Qing, GU Xuemai

School of Electronics and Information Engineering, Harbin Institute of Technology, Harbin 150006, China

Abstract: The space network is mainly composed of satellite networks of various architectures with many network nodes. Compared with the ground network, the nodes in the satellite network are difficult to observe and control, and the network equipment is extremely expensive. The rapid increase in the number of users has brought a great burden to the network, and at the same time caused the network resources to be unreasonably used. The application of traditional network architecture to large-scale satellite networks results in the inability to flexibly configure the network and uneven resource allocation. Aim to the above problems, a space network resource scheduling system based on SDN(software defined network)/NFV(network function virtualization) architecture was proposed to realized real-time monitoring of space network resources. At the same time, it combined MPTCP (multipath TCP) and SDN to completed the flexible configuration of the space network and the rational use of its resources. It improved the throughput of the network.

Key Words: space network; SDN, resource scheduling, MPTCP

1 引言

对于下一代空间卫星网络设计来说, NFV、SDN和软件定义无线电被视为最好的推动空间卫星网络发展的新型技术。目前在将SDN引入卫星网络相关技术方面已经进行了一些探索。2016年希腊雅典太空研发部的GARDIKIS等^[1]研究了NFV在卫星网络中的适用性, 并验证了其优势和提出了NFV应用在卫星网络中所要面临的挑战; 之后SHENG等^[2]研究了网络软件化, 即SDN和NFV技术在卫星

网络中的适用性, 其中还介绍了通过特定的用例将SDN/NFV集成到卫星基础设施中的好处, 并提出了基于SDN/NFV的分层架构卫星地面网络, 这为基于卫星的空间信息网络架构的搭建提供了很好的思路和技术途径; 2016年加泰罗尼亚理工大学FERRIS等^[3]进一步研究了SDN/NFV技术引入卫星地面部分带来的好处和技术挑战; 2016年希腊电信组织的AGAPIOU等^[4]详细地阐述了一种新的地面卫星架构, 将NFV带入卫星通信领域并利用基于SDN的新型架

收稿日期: 2020-08-15; 修回日期: 2020-09-11

基金项目: 国家自然科学基金资助项目(No.61671183, No.61771163)

Foundation Item: The National Natural Science Foundation of China(No.61671183, No.61771163)

构进行资源管理,首次提出了SDN的资源管理层面,为资源虚拟化技术提供了思路;FERRUS等^[5]进一步描述了基于SDN和NFV技术构建的卫星地面段的系统架构,基于这种架构,将有助于卫星通信集成到5G系统中。但以上的文献提出的技术方法,仅限于架构层面,没有对数据资源层面进行相应的研究和优化。

LEO系统中的传播时延可高达100 ms,并极大降低TCP吞吐量,现有的传输层解决方案包括中国科技大学HA等^[6]提出的高带宽时延TCP方法,但是其不能解决频繁切换的问题;REDA等^[7]提出的传统等价多路径转发算法ECMP被广泛地应用在大型数据中心网络中,当网络中存在大量的冗余链路时,该算法可以利用此资源来解决数据传输中的负载问题,然而研究表明,传统的ECMP算法对小流处理较为有效,对持续时间长且带宽敏感的大流会造成链路拥塞,因为其不考虑链路代价;美国加利福尼亚大学DU等^[8]提出了多路径TCP和多路径路由联合策略来解决频繁切换等问题,但是其在发现路径时出现额外的开销,同时无法将子流拆分成不相交路径,但是提供了改善空间信息网络通信方法的思路;文献[9]中BONFIM等对NFV和SDN的最新技术进行了相关的研究,并提出了相关的改进领域以及此架构将要面临的问题;文献[10]中QU等提出了基于NFV/SDN的框架,针对嵌入式服务的动态流迁移问题进行了研究,以满足随时间变化的流量的端到端(E2E)时延要求,提出了多目标混合整数优化,解决了负载均衡和重新配置开销之间的权衡问题;文献[11]中YONG等实现了利用SDN和NFV来按需修改网络配置,提出了一种选择性移动性管理方案,用于在用户设备切换的情况下实现服务连续性。数值结果表明,在5G核心网络上部署的云数量方面,不同的成本因素(即网络成本和服务资源成本)之间需要进行权衡;文献[12]ZHUANG等根据软件定义网络和网络功能虚拟化技术提供了潜在的解决方案,以实现具有成本效益的灵活和自动化的网络管理,验证了其增强数据通信、计算和缓存的能力;文献[13]中哈尔滨工业大学JIA等对大规模低轨卫星路由算法进行了研究,基于SDN技术,提升整个网络性能,为路由相关技术提供了新的思路。以上两个技术可应用在对异常流量监测的方法中。

2 基于SDN/NFV的空间网络架构研究

基于SDN/NFV的多层卫星星座网络架构如图1所示,主要包括地面的网络控制中心、卫星地面站,如网关和骨干网络、LEO卫星和GEO卫星。空间网络中的卫星按照一定的倾角和轨道组成,实现包括南北极的全天实时可见性,同时GEO与自己连接范围内的LEO卫星

组成SDN,并与地面控制中心相连^[14],由地面控制中心作为技术支持,对整个空间网络进行实时监测与灵活管控。将控制器放在GEO卫星和地面终端,相比于仅有一端的控制平面,可以实现集中式与分布式相结合的空间网络架构,即地面端可控制GEO卫星,GEO卫星再选择局部的LEO卫星进行管控,同时将控制中心安放在地面端,可方便操作人员进行实时操作。

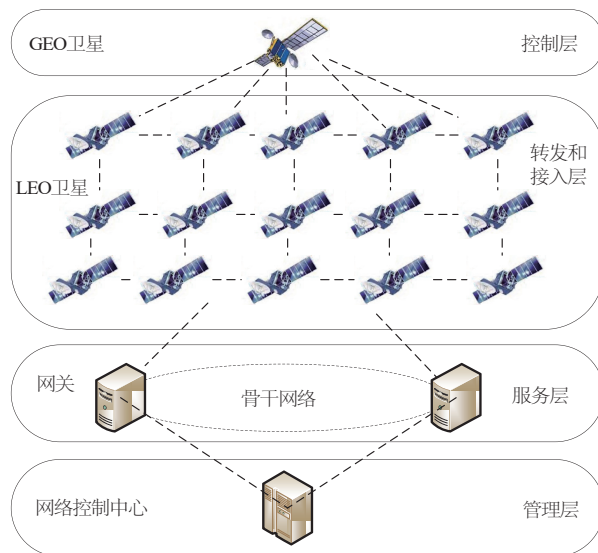


图1 基于SDN/NFV的多层卫星星座网络架构

本文参考基于SDN/NFV的空间网络传输体系架构^[15],该架构如图2所示。该控制器位于地球静止轨道(GEO)卫星中,并且这些交换机部署在SDSN中的中地球轨道(MEO)和低地球轨道(LEO)卫星上。当数据包到达转发卫星时,卫星在其内部流表中执行查找。除了流表未命中之外的流条目,卫星将以常规方式将分组转发到下一颗卫星,否则该包应该属于一个新的流程。在这种情况下,转发卫星通过发送封装到达分组的Packet_In消息向控制卫星请求指令,控制卫星确定相应的流量规则并将其安装到位于转发卫星中的流表中。之后,流表中的所有数据包都会正确转发到目标,而不会请求控制器。在图2中,卫星业务到达LEO1卫星并匹配流表中的条目。然后,流量将被转发到沿路径的下一个卫星。流量到达MEO卫星,而不匹配本地条目。因此,MEO卫星向GEO卫星中的控制器发送Packet_In消息以用于转发规则。

3 基于MPTCP-SDN协同方法的资源调度

3.1 MPTCP技术分析

MPTCP与传统的TCP不同,MPTCP传输模型如图3

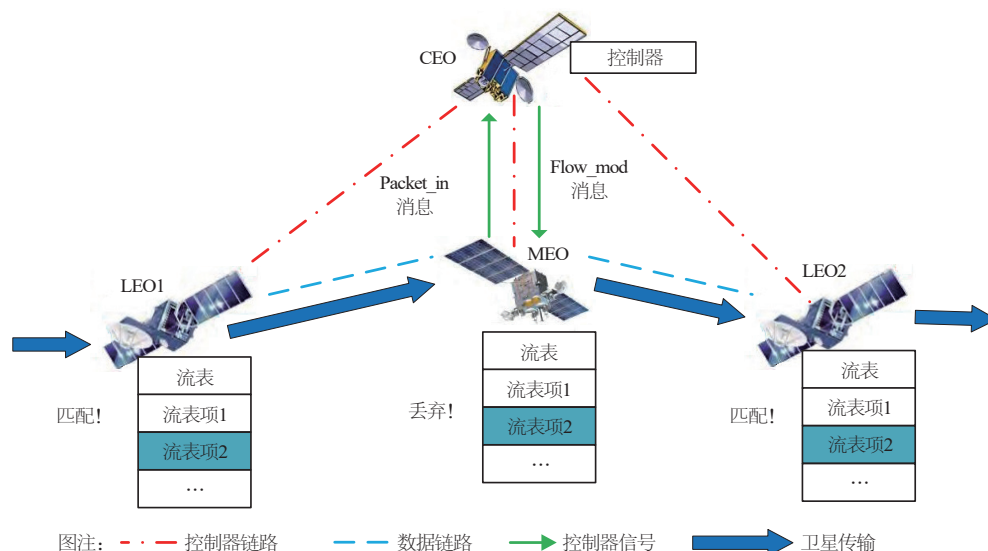


图2 基于SDN/NFV的空间网络传输体系架构

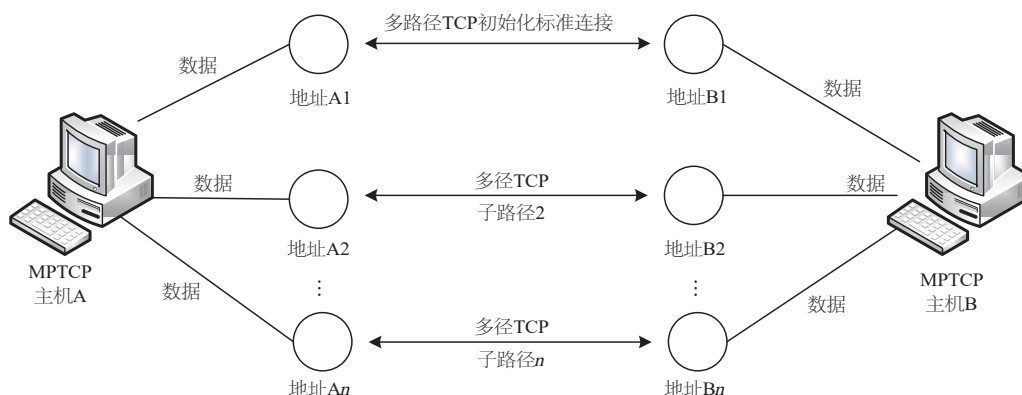


图3 MPTCP传输模型

所示，当数据包到达MPTCP主机A中时，首先被拆分为多个子流，即多个数据包，然后在选择的多个路径上集中实现数据的传输。

对于存在多条冗余链路的网络来说，将数据流拆分为多个子流通过多条路径传输可以最大化地利用网络中的冗余链路，将不再仅使用一条链路进行传输，进而提高了链路的利用率，实现吞吐量的增加。与TCP类似，MPTCP也是通过3次握手连接以及4次挥手断开连接，同时对于通信双方，若想使用MPTCP进行传输，必须要整主机A/B都支持MPTCP，同时进行MPTCP传输时，可以根据用户需求，选择不同的子路径进行传输，实现合理的资源调度。

在主机双方之间，传输层仍然是单路径传输，MPTCP的主要作用是将需要传输的数据先进行拆分，之后进行路径选择，使其在不同的路径上进行传输，实时监测通信两

端可用路径，传输层的MPTCP收到应用层发来的数据需要对其进行处理，再发送给各个子路径进行传输，数据在子路径上添加序列号和确认号发到网络层上，当所有被拆分的子流到达目的主机后，MPTCP则根据子流数据上的标识重组为一个数据包，发送给应用层。

3.2 MPTCP在空间网络拥塞控制中的应用

LEO卫星网络由极地轨道（飞机）组成，每个轨道都具有一定数量的卫星，这些卫星与地球的通信距离相比于其他卫星更近，形成了一个连接地球的网格状网络。此卫星星座通信系统与其他卫星网络相比，由于相对较短的时延和较低的发射功率要求，应首选LEO系统，但是，缺点是LEO卫星在天空中移动很快，每颗LEO卫星平均访问时间为10 min。因此，大多数语音和

视频应用程序会频繁切换,从而导致连接中断和性能下降。另外,LEO系统中的传播时延可能高达100 ms,并使得TCP吞吐量大幅度降低。相关文献提出了一种多路径TCP(MPTCP)和多路径路由联合策略来应对这些挑战。与以上的解决方案相比,该框架通过使用多个物理接口来很好地处理切换,并显著提高了吞吐量,但仍存在一些路由规划与相关负载等问题。因此,本文将MPTCP和SDN结合将更好地实现智能路由。其中GEO卫星层可作为网络中的控制层,对全网进行监控,大规模的LEO卫星作为网络中的转发和接入层,对流量数据进行转发以及和地面终端相连。地面终端作为服务层,可构建网络控制中心,进一步对整个空间网络的流量信息进行合理的监控及调度。

相比于ECMP算法,MPTCP扩展了多条路径来进行数据传输,主要设计目标是提高吞吐量,并且保持相对的公平性,因此网络的吞吐量受到传输到网络中的数据增加的影响。在使用多条路径传输数据时,相比于单条路径来说,将会产生路径拥塞的情况,严重的会导致传输中断。将收发双方的连接分成多条路径上的多个子流,每个子流都具有相同的性质。尽管互相通信的端点之间存在流量控制,但也只能对发送方和接收方的通信数据进行协调。若通信使用的路径中,存在能够被其他用户共享的路径,则当多个用户共享某条路径时,长时间将会造成拥塞链路,导致网络数据传输发生中断。其次,由于数据拆分为多个子流进行传输,如果子流使用相同的拥塞控制,同样会带来很多问题,那么若其中的子流同时使用了网络中的某一条路径,那么在该路径上,就占用了成倍的带宽资源。

大量子流的创建和维护会给最终主机带来额外的开销,这需要更大的缓冲区来应对重排序,并且由于MPTCP调度程序的使用量增加,CPU利用率更高。然而,更重要的是,在SDN环境中,大量子流通过在交换机上安装大量规则并增加SDN控制器的负载,给网络带来了额外的开销,因此,期望在不牺牲性能的情况下具有尽可能少的子流。给定一组路径,建议SDN控制器可以通过确定性地为其分配MPTCP子流,更好地利用这些路径。正如本文所介绍的,这可以形成最佳性能,而无须像基于随机方法那样需要那么多子流。

3.3 MPTCP协同SDN系统的设计与实现

本节对子流的建立过程进行介绍,如图4所示。

对于初始子流,主机使用包含随机生成的密钥的MP_CAPABLE选项执行握手,这些键用于令牌的计算(密钥的加密哈希)。为了建立其他子流,主机使用

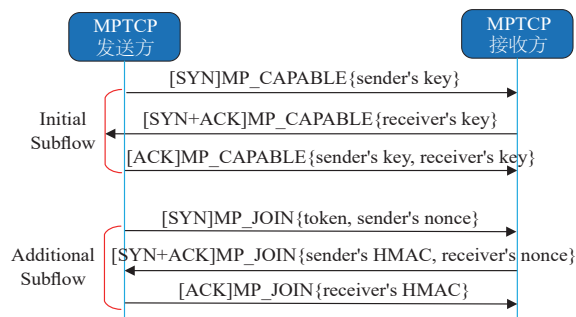


图4 MPTCP中子流的建立过程

MP_JOIN选项执行握手,如图4所示,发送方发送带有令牌和随机数SYN的数据包,当数据包到达接收方时,接收器作为反馈发送SYN+ACK数据包作为响应。该数据包包括自己携带的随机数和经过计算后的发送方随机数的HMAC代码。最后发送方发送一个ACK数据包,包含计算后的接收方随机数HMAC,其中HMAC的计算使用在MP_CAPABLE握手过程中交换的密钥和MP_JOIN握手过程中交换的随机数。

接下来考虑如何在Floodlight控制器中实现MPTCP协同SDN算法。图5为MPTCP协同SDN系统,其中系统的传输协议为MPTCP。系统将底层数据中心拓扑作为输入,每个交换机都通过控制平面连接到SDN控制器。控制器负责路径的计算,并为交换机安装适当的OpenFlow规则,当数据包到达交换机时,安装的流规则会对数据包进行相应的处理。当子流的数据包到达交换机,交换机中不存在相应的流规则时,该数据包将被重新定向到控制器,控制器计算子流的路径,并将规则安装到计算的路径中的所有交换机。相同子流的后续数据包将通过相同的路径,避免再次被定向到控制器中,减少控制器负载。

本文基于Floodlight控制器实现了此系统。此系统主要由两个组件组成:拓扑管理器(Topology Manager)和转发模块(Forwarding Moudle)。

(1) 拓扑管理器: 拓扑管理器可以获取最新的信息,具有最新的全局视图,这是通过SDN控制器和交换机之间交换的消息获得的。当转发模块接收到消息,需要一组源IP地址与目的IP地址之间的路径时,就会通过拓扑管理器查询获得。拓扑管理器首先使用深度优先搜索图遍历算法(DFS),查找长度不超过特定跃点计数阈值的所有可选择路径,然后路径管理器过滤路径集,以选取路径子集之一返回给转发模块:

- ①最短路径,包括与最短路径相同长度的所有路径;
- ② k 条最短路径,其中包括递增跳数顺序前的 k 条

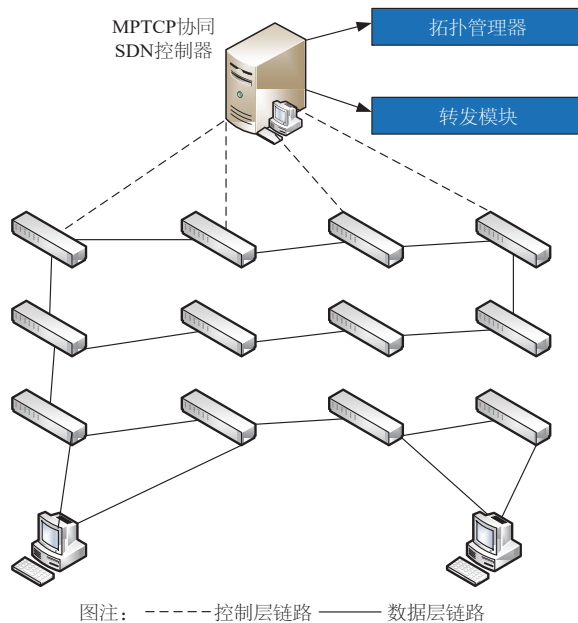


图5 MPTCP协同SDN系统

路径；

③ k 条边缘不相交最短路径，其中包括前 k 条路径，这些路径不共享任何边。

对于MPTCP传输协议来说，最重要的是路径集的选择，由于是多路径传输，因此在网络中应多利用冗余链路，提升链路利用率。在选定好路径集之后，进一步将路径分配给确定的子流，为了能够以最少的子流来最大化MPTCP性能。因此除减少终端主机的开销外，还需最小化子流数量。

(2) 转发模块：此模块负责将路径分配给确定性或随机性子流。针对随机性子流来说，转发模块所获取的路径集为子流的源目的IP地址对随机选择一条路径，并将此路径分配给该子流。当路径集为最短路径时，此方法等效于基于随机流的ECMP算法。在确定性方法中，每个子流都会分配不同的路径，此时方法为MPTCP协同SDN算法。此方法要求转发模块从子流设置数据包中提取MPTCP选项，使子流与现有的MPTCP连接匹配。此方法可以有效地避免ECMP所存在的问题，将同一连接的多个子流分配给同一个类路径。如果子流的数量大于可用路径的数量，则以统一的方式确定性地将子流分配给路径。

转发模块维护两个哈希表：pathCache和flows。pathCache哈希表用于缓存源目的IP地址之间的所有路径集。flows哈希表中包括每个MPTCP连接的条目，每个此类的条目都包含该连接的子流所使用的每对源目的IP地

址的子条目，此子条目成为IPentry，都会缓存源目的IP地址对所获取的路径集，并为该组路径中的子流进行分配。此条目的作用为使转发模块可以实时跟踪，为新的子流选择可用的路径。

4 仿真结果及分析

本文实现了MPTCP内核的安装及编译，并移植到Mininet建立的网络拓扑中，可使用MPTCP进行数据的传输。实验时在Ubuntu16.04系统上运行，依据铱星系统的参数进行仿真，实现所需网络拓扑搭建，同时创建50个子流实现网络的拥塞，验证全局感知多路径TCP算法的增益。拥塞控制使用链路增加算法，路径管理模块为全网格。通过Mininet建立网络拓扑，其拓扑形状为网格状，类似于LEO卫星通信系统，同时其链路参数采用铱星系统链路参数进行仿真，卫星与地面终端带宽为2 Mbit/s，星间链路带宽为9 Mbit/s，链路时延为14 ms。

根据以上参数以及建立的拓扑，首先对TCP协议进行仿真，采用Reno拥塞控制算法，对50个流其中的一个流进行吞吐量分析。可以看出，当数据流较多时，网络出现拥塞，数据传输的速率波动很大，并且没有很好地分配带宽资源。采用ECMP基于流的多路径传输策略来实现网络的负载均衡，通过寻找最短路径，将流分散到不同的最短路径上进行转发。与上述仿真参数设定相同，仿真对比图如图6所示。

首先对ECMP算法进行仿真。ECMP选择最短路径进行数据传输，同时使用随机散列。使用lia拥塞控制算法，选择所有最短路径对数据进行传输。随机选取4个数据流进行分析，仿真时间为200 s，传输速率仿真图如图7所示。

进一步对MPTCP协同SDN算法进行仿真，此算法选择4条不相交路径进行数据传输，仿真条件相同。同样对50个数据流进行吞吐量性能测试，仿真图如图8所示。

最后对原始的TCP协议与Dijkstra算法进行仿真分析。与上述仿真条件相同，同样对相同的流数进行仿真，使用Reno拥塞控制算法。同时选择所有最短路径进行数据传输，仿真图如图9所示。

根据以上对3种算法仿真性能的测试，分析不同数据流下，即不同拥塞程度情况下，数据流占比情况，如图10所示。其中5个数据流分析其吞吐量不低于0.1 Mbit/s，10个数据流分析其吞吐量不低于0.05 Mbit/s，15个数据流分析其吞吐量不低于0.04 Mbit/s，30个数据流分析其吞吐量不低于0.02 Mbit/s，40个数据流分析其吞吐量不低于0.018 Mbit/s，50个数据流分析其吞吐量不低于0.018 Mbit/s。

由以上分析可以看出，当需传输数据量增加，网络出现拥塞时，传统的TCP协议使得很多数据流传输速率

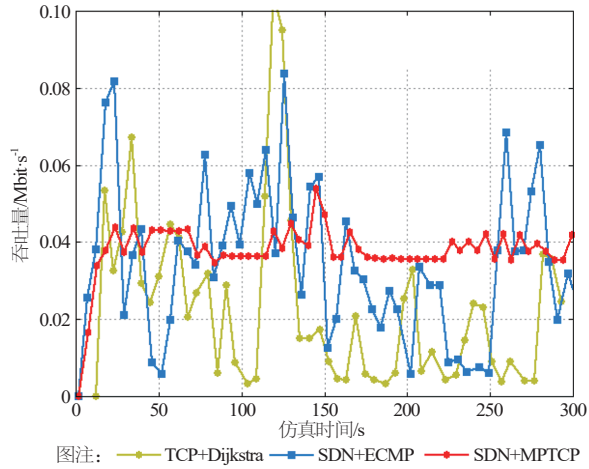


图6 算法仿真对比

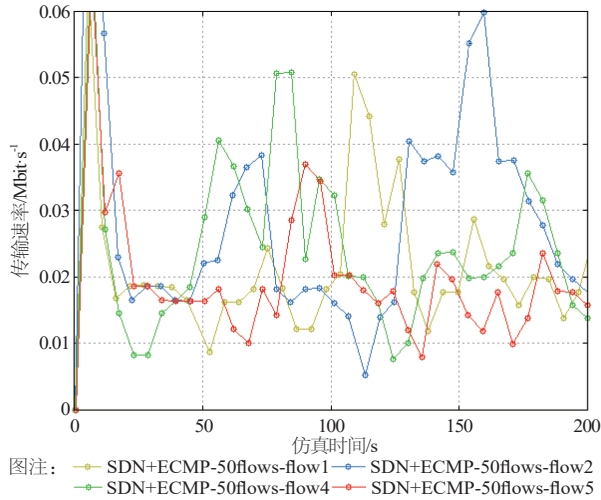


图7 ECMP算法50个数据流仿真图

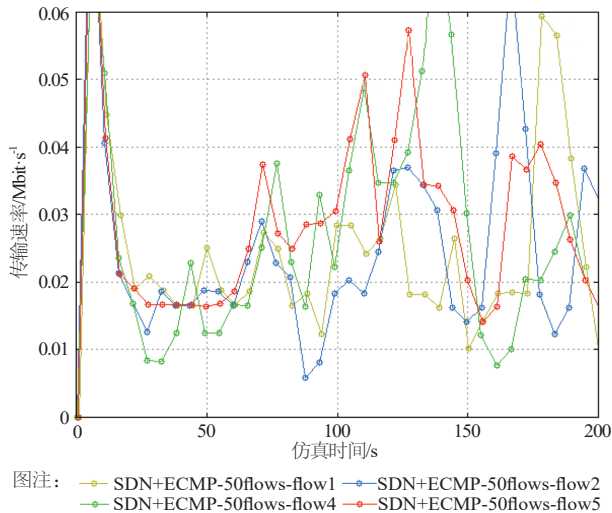


图8 SDN-MPTCP算法50个数据流仿真图

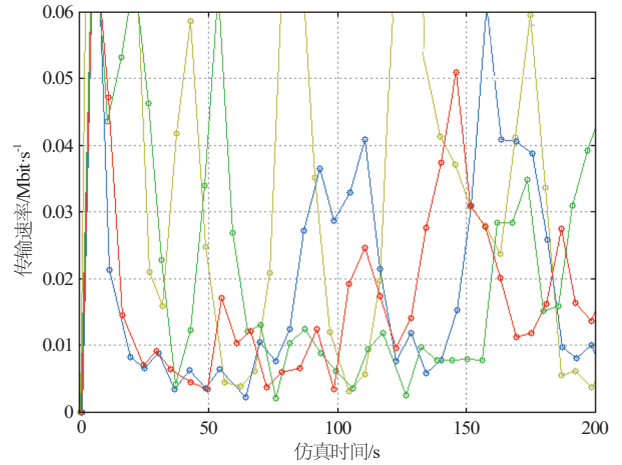


图9 TCP协议50个数据流仿真图

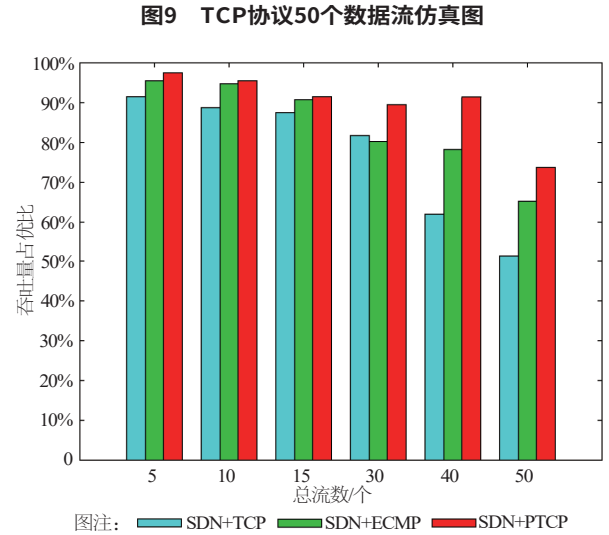


图10 不同算法吞吐量占优比仿真图

很差,无法满足系统性能要求。ECMP算法选择所有最短路径进行传输,使用随机散列矩阵,导致部分链路出现拥塞。而MPTCP协作SDN算法选用不相交路径进行数据传输,使用多个子流,数据流数量较小时增益不明显,当数据流增加时,可看出MPTCP协作SDN算法相比于ECMP算法数据流传输速率占优比明显增加。因此MPTCP协作SDN算法应用在大规模LEO卫星组成的网络拓扑中,当数据量激增时,针对系统的吞吐量有明显的增益,保证大部分数据流得到较好的传输。

5 结束语

本文主要将MPTCP与SDN结合,解决由于用户激增导致的空间网络中的链路负载增大,使得系统性能下降的问题。设计并实现了MPTCP协同SDN算法,此算法

包括了 K_SP 与 K_DP 算法,即 K 条最短路径算法和 K 条不相交路径算法,在SDN控制器中实现MPTCP协议,有效地利用网络中的冗余链路,减少链路出现拥塞情况,减少当多个子流传输时传输速率极低等情况,提升网络性能,更好地实现数据公平的传输。

参考文献：

- [1] GARDIKIS G, COSTICOGLOU S, KOUMARAS H, et al. NFV applicability and use cases in satellite networks[C]// European Conference on Networks & Communications. IEEE, 2016.
- [2] SHENG M, WANG, LI J, et al. Toward a flexible and reconfigurable broadband satellite network: resource management architecture and strategies[J]. IEEE Wireless Communications, 2017:2-8.
- [3] FERRIS R, KOUMARAS H, SALLENTA O, et al. SDN/NFV-enabled satellite communications networks: Opportunities, scenprios and challenges[J]. Physical Communication, 2016,18(2):95-112.
- [4] AGAPIOU G, FERRUS R, KOUMARAS H, et al. SDN and NFV for satellite infrastructures[C]// in Proc. IEICE ICTF Inst. Electron, Inf. Commun. Eng.(IEICE),2016:1-4.
- [5] FERRUS R, SALLENT O, AHMED T, et al. Towards SDN/NFV-enabled satellite ground segment systems: End-to-End Traffic Engineering use case[C]//2017 IEEE International Conference on Communications Workshops (ICC Workshops). IEEE, 2017.
- [6] HA S, RHEE I, XU L. Cubic: a new tcp-friendly high-speed tcpvariant[J]. ACM SIGOPS Operating Systems Review, 2008,42(5):64-74.
- [7] REDA W, BOGDANOV K, MILOLIDAKIS A. Analysis of an equal-cost multi-path algorithm[J]. ACM SIGCOMM Computer Communication Review, 2000, 109(1):S265.
- [8] DU P, LI X, LU Y, GERLA M. Multipath TCP over LEO satellite networks[C]//In Wireless Communications and Mobile Computing Conference(IWCMC), 2015 International. IEEE, 2015:1-6.
- [9] BONFIM M S, DIAS K L, FERNANDES S F L. Integrated NFV/SDN architectures[J]. ACM Computing Surveys, 2019, 51(6):1-39.
- [10] QU, ZHUANG K, YE W H, et al. Dynamic flow migration for embedded services in SDN/NFV-enabled 5G core networks[J]. IEEE Transactions on Communications. 2020: 1-1.
- [11] YONG K, JOON G, KIM D. A location-aware network virtualization and reconfiguration for 5G core network based on SDN and NFV[J]. International Journal of Communication

Systems, 2020.

- [12] ZHUANG W, YE Q, LYU F, et al. SDN/NFV-empowered future IoV with enhanced communication, computing, and caching[J]. Proceedings of the IEEE, 2019 (99):1-18.
- [13] JIA M, ZHU S, WANG L, et al. Routing algorithm with virtual topology toward to huge numbers of LEO mobile satellite network based on SDN[J]. Mobile Networks and Applications, 2017.
- [14] 张九龙, 鄂广增. 我国移动卫星星座设计、覆盖分析及动态仿真[J]. 南京邮电大学学报(自然科学版), 2002(2):11-15.
ZHANG J L, FENG G Z, My country mobile satellite constellation design, coverage analysis and dynamic simulation[J].Journal of Nanjing University of Posts and Telecommunication (Natural Science Edition), 2002(2):11-15.
- [15] LI T, ZHOU H, LUO H, et al. Using SDN and NFV to implement satellite communication networks[C]// 2016 International Conference on Networking and Network Applications (NaNA). IEEE, 2016.

[作者简介]



贾敏 (1982-), 女, 黑龙江哈尔滨人, 哈尔滨工业大学教授、博士生导师, 主要研究方向为天地一体化认知通信、组网与接入技术。



郭家 (1997-), 男, 山东潍坊人, 哈尔滨工业大学硕士研究生, 主要研究方向为软件定义网络、机器学习。



徐媚琳 (1995-), 女, 黑龙江哈尔滨人, 哈尔滨工业大学硕士生, 主要研究方向为SDN、网络虚拟化技术。



郭庆 (1964-), 男, 黑龙江哈尔滨人, 哈尔滨工业大学电子与信息工程学院院长、教授、博士生导师, 主要研究方向为卫星通信和宽带多媒体通信。



顾学迈 (1957-), 男, 黑龙江哈尔滨人, 哈尔滨工业大学教授、博士生导师, 主要研究方向为卫星通信和天地一体化测控。