

数据中心光互联中低时延路由频谱分配算法

赵继军*, 郭宏**

河北工程大学信息与电气工程学院, 河北 邯郸 056038

摘要 数据中心光互联通过 IP 网和光网络联合优化, 实现了业务跨层疏导, 但由于数据中心内部业务的突发性, 突发业务被疏导到光网络, 导致光网络业务的排队时延和建路时延增加, 为此提出一种低时延路由频谱分配 (RSA) 算法。该算法将光纤中的频谱资源分为多个频谱层, 在各频谱层并行建立光路, 在频谱层中再用多个线程并行建路。各频谱层中当前建路时用的点和边组成了着色虚拟拓扑, 将着色虚拟拓扑作为频谱层中资源使用的约束条件, 并根据当前正在建路、未建路、新到达业务请求的变化在各频谱层重构着色和可用虚拟拓扑, 来适应突发业务, 快速疏导聚集在光网络中的业务请求。仿真结果表明, 该算法减少了业务请求的光路平均建立时延, 减小了光路建立时延方差和阻塞率, 提高了频谱利用率, 提升了用户体验质量 (QoE)。

关键词 光通信; 弹性光网络; 路由频谱分配; 虚拟拓扑重构; 低时延; 二维资源分配模型

中图分类号 TN929.11

文献标识码 A

doi: 10.3788/LOP55.080602

A Routing and Spectrum Assignment Algorithm with Low Delay in Optical Interconnection of Data Center

Zhao Jijun*, Guo Hong**

School of Information and Electrical Engineering, Hebei University of Engineering, Handan, Hebei 056038, China

Abstract Requests grooming in multi-layer architecture are realized by the joint optimization of IP and optical networks in optical interconnection of data center. Due to the burstiness of requests in intra-data center, burst requests are groomed to optical network, which leads to the longer delay of queue and establishment of light paths of requests. In order to reduce the establishment delay of optical paths, we propose a routing and spectrum assignment (RSA) algorithm with low delay, which separates the spectrum resource of fibers into spectrum layers, and multi threads can set up optical paths in different spectrum layers in parallel. The algorithm sets up colored virtual topology with the points and edges of establishing optical paths in every spectrum layer, and the topology is used as constraints of resource utilization in current spectrum layer. The colored and usable virtual topologies in spectrum layers are reconfigured according to transformations of establishing, unestablished and new arrived requests, which can adapt bursts of requests and promptly groom the requests swarmed in optical network. The simulation results demonstrate that the proposed algorithm can efficiently reduce the average delay, delay variance of establishment of optical paths and blocking probability of requests, and improve the rate of spectrum utilization and quality of experience (QoE).

Key words optical communications; elastic optical network; routing and spectrum assignment; reconfiguration of virtual topology; low delay; two-dimensional model of resources assignment

OCIS codes 060.4510; 060.4264; 060.4258; 060.4256

1 引言

在当前以高清视频、远程医疗和金融数据等为代表的互联网业务驱动下, 以网络为基础的高性能

数据中心应用受到了网络运营商的强烈关注, 光互联凭借其低时延、高带宽、高能效和灵活性等优势, 成为适应大规模数据中心组网的关键技术, 其中时延性能是衡量数据中心服务质量的重要指标, 成为

收稿日期: 2018-01-25; 修回日期: 2018-03-12; 录用日期: 2018-03-21

基金项目: 河北省自然科学基金(F2017402068)、河北省高等学校科学技术研究青年基金(QN2016090)

* E-mail: zhaojijun@ieee.org; ** E-mail: hb_peric@163.com

研究关注的重点^[1]。

在传统的数据中心网络互联中,IP 分组网络和光网络相互独立运营,IP 网的集群路由技术提升了骨干路由器的缓存能力和交换容量,弹性光网络实现了大容量、超长距、低时延的传送^[2],但是 IP 网和光网络的分开管理和维护,造成网络资源浪费和运维成本升高。文献[3]提出了数据中心的跨层联合优化控制架构,利用 OpenFlow 协议将软件定义网络(SDN)控制器扩展到光层^[4],中心 SDN 控制器维护全网络的拓扑资源信息,通过分析业务粒度和当前网络资源,决定 IP 层和光层中哪些资源更适合为业务请求提供服务^[5],以此来提高资源使用效率和业务的快速响应能力^[6]。但由于数据中心应用的突发性和 IP 数据业务的不确定性^[7-9],突发业务被疏导到光网络^[10-11]。光网络中尽管频谱资源充足,但是业务请求大量聚集,导致光网络控制器无法及时为聚集业务建立光路,业务请求建路时延逐渐增加,严重影响了网络的时延性能,降低了用户体验质量(QoE)。

叶飞等^[12]预先在光网络单元上设置波长分配表,简化波长分配的流程,降低波长分配时延,提高了用户体验质量。李非^[13]在对业务进行波长路由分配时,删除达到预定阈值的节点或光纤,动态调整网络拓扑,减少波长路由分配的时间,提升用户体验质量。Guo 等^[14]通过预先并行计算出路由频谱分配(RSA)方案,来减少业务请求光路建立时延,该算法在预先计算期间能充分结合静态和动态 RSA 算法的优点。Lu 等^[15]在业务请求所需频谱资源不足时,通过正交频分复用(OFDM)技术,在多径传输的时延差限制下,将业务在多条路径中进行传输,提高了网络对业务的疏导能力。Yang 等^[16]通过压缩或扩展业务请求的带宽,来减少数据中心业务请求的光路建立时延或数据传输时延。文献[17]通过将光交换开关集成到芯片内部,实现芯片内光互联,并行进行光输入输出端口的连接,减少光路建立时延。文献[18]在进行频谱重构时,判断频谱重构的次序和相互独立性,将频谱重构分解为多个相互独立的子问题,来减少光路重新建立时延。上述算法虽取得了很好的效果,但是由于业务请求光路逐个建立,遇数据中心跨层疏导突发业务时,算法无法减少排队时延的影响。

为减少数据中心光互联中业务的光路建立时延,本文在数据中心跨层联合优化架构的基础上,提出数据中心光互联中基于虚拟拓扑重构的低时延路

由频谱分配(LDRVT-RSA)算法,并对该算法性能进行仿真验证。

2 问题描述

2.1 跨层联合优化架构中时延描述

光网络的时延主要包括信号的传输时延、网络设备的处理时延和控制器的策略处理时延等。文献[19]对光传送网 L_0, L_1, L_2, L_3 层硬件设备的时延进行对比,发现设备所在层次越低,处理时延越小, $L_0, L_1, L_2/L_3$ 层网元分别引入纳秒、微秒、毫秒级的时延。硬件设备的处理时延和前向纠错(FEC)编解码、封装和映射算法有关,在设备设计阶段已经形成,其中光器件和算法时延较低且较难进一步提升。信号的传输时延与所选路由的距离成正比,控制器的策略处理时延和所选的 RSA 算法有关。由于光网络时延中信号的传输时延和控制器的策略处理时延是决定性的因素,减少光网络时延的主要方法是减少所选光路的距离和控制器的策略处理时延。在跨层联合优化架构中,光网络需要承担 IP 网的业务疏导任务,导致大量的业务请求在光网络聚集,后面到来的业务请求,排队时延逐渐增加,严重影响了网络的时延性能。因此,在跨层联合优化的场景下,信号的传输时延、控制器的策略处理时延和业务请求的排队时延成为光网络时延的决定性因素。

2.2 传统 RSA 算法和多线程 SDN 控制器描述

传统 RSA 算法^[20-21],例如最短路径首次频谱适应算法(SPFF-RSA),具有计算开销小、复杂度低的特点,但是算法在进行光路建立时,受算法控制模型的限制,业务请求的光路逐个建立,遇突发业务是突发业务将严重影响网络的时延性能。为减少业务请求光路建立时延,文献[22]将静态业务的路由频谱分配映射成多处理器系统的多个子问题,各处理器可以同时处理各自的任务,提升了 RSA 效率。文献[23]介绍了多个厂家的 SDN 控制器,其中 NOX、Floodlight、MUL、Beacon、Maestro 的 SDN 控制器支持多线程技术,文献[24-25]通过对比单线程和多线程 SDN 控制器吞吐量、可靠性、安全性等方面的指标,证明多线程控制器在吞吐量等方面的优势明显。

由于 SDN 控制器可较好地支持多线程技术,本研究设计了单边约束多线程路由频谱分配(SECM-RSA)算法,该算法虽然能同时建立不同业务请求的光路,但是由于缺少合理的约束条件,多个线程对网络资源的使用变得复杂和无序,无法高效使用全局资源。多个线程在建路过程中使用相同的频谱资源时,

部分线程需撤销已分配的资源,这不仅导致程序执行不稳定、算法时延增加,还可能导致频谱资源的重复利用。其原因为 SECM-RSA 算法用单个资源作为全局资源使用的约束条件,约束条件具有单维度性和相互独立性,约束条件预先生成,不在资源的使用过程中形成和连续变化。因此,LDRVT-RSA 算法还需要更合理的约束条件、资源互斥使用模型和控制模型来梳理算法,从而打破业务请求光路逐个建立的限制,减少业务的排队时延,实现不同业务请求光路的同时建立。

2.3 符号定义

为了描述算法的思路,算法中所使用符号及意义如表 1 所示。

表 1 网络元素符号及其含义

Table 1 Symbols of network elements and their implications

Symbol	Implication
L_n	n^{th} spectrum layer
V_i	Node i , the set of all nodes is called as V
$E_{i,j}$	Edge from node i to node j , the set of all edges is called as E
$G(V,E)$	Physical topology of network (G for short)
$G_n^*(V_n^*, E_n^*)$	Coloring virtual topology in L_n (G_n^* for short), V_n^* and E_n^* are the sets of coloring nodes and edges in L_n , respectively
$G_n'(V_n', E_n')$	Usable virtual topology in L_n (G_n' for short), V_n' and E_n' are the sets of usable nodes and edges in L_n , respectively
V_i^n	Virtual node i in L_n
$E_{i,j}^n$	Virtual edge from virtual node i to j in L_n
$R(i,j,n_{\text{slot}})$	Request from node i to node j , need n_{slot} slots (spectrum element); the set of all requests is called as R
$R^n(i,j,n_{\text{slot}})$	Request with affordable resource in L_n from node i to node j , need n_{slot} slots
S_{variable}	Variable, the unit of the symbol is ms
L_{times}	Times that the function of sleep (S_{variable}) need to be executed in a spectrum layer
T_{gen}	Interval to generate burst requests
R_n	Hop threshold of long distance requests
R_l	Count of long distance requests
R_u	Threshold to serve long distance requests
R_d	Threshold to stop serving long distance requests
$\lambda_{i,j}^n$	Used times of the n^{th} slot in $E_{i,j}$

3 算法设计

3.1 用虚拟拓扑重构实现多线程资源互斥使用模型

网络虚拟化技术是建立新一代互联网体系架构的新兴技术^[26],它允许创建多个具有不同拓扑且相互隔离的虚拟网络,同时共存于公共的基础设施上,以充分利用底层物理网络资源^[27-28],有效地节约网络资源的消耗^[29-30],为用户提供差异化的端到端定制服务。创建虚拟拓扑是实现网络虚拟化技术的关键问题,RSA 算法引入的虚拟拓扑能够充分利用网络可用带宽^[31-32],将网络资源的使用情况和节点间业务的分布作为约束条件,选择光通道构成方案,提高网络对业务需求变化的适应性^[33]。

在 LDRVT-RSA 算法中,需考虑光网业务请求的动态性和网络资源的动态变化。虚拟拓扑作为资源使用的约束条件需通过重构来适应未建路业务、新到业务和正在建路业务的变化。虚拟拓扑重构要考虑 3 个问题:1)确定在什么情况下需要重构;2)重构虚拟拓扑,既要考虑满足新的业务请求,又要使虚拟拓扑改变尽量少;3)如何切换到新的虚拟拓扑而又不对正在建立的光路有任何损伤,即实现无损伤重构。

数据中心光互联由于时间、频谱、空间和偏振等多维度资源的引入,令逻辑拓扑的灵活性能脱离物理线路的硬约束,为解决多线程并行建路问题提供了可能。为实现多线程对网络资源互斥式使用,LDRVT-RSA 算法中概念定义如下:

1) 光纤链路中的频谱平均分为多个宽度的频谱资源,单个宽度的频谱资源称为频谱层。

2) 频谱层中需要建立的多条光路,如果没有公共的点(端)或边,本研究称之为路径不重合,如图 1 所示,在 NSFNET 拓扑中, $V_2 \rightarrow V_3$ 和 $V_4 \rightarrow V_5$ 是同一个频谱层中两个路径不重合的业务请求,可以在不同的线程中同时建立光路。

3) 各个频谱层中着色虚拟拓扑由各个频谱层中正在建路用的点和边集合组成,反映各频谱层中频谱资源的实时使用。全网拓扑删除着色虚拟拓扑剩余可用的拓扑称为可用虚拟拓扑。

4) 业务的线程着色是指线程通过路径查找算法,为业务请求计算路径,查找能为业务建路提供资源的频谱层,在可用频谱层中把路径上的点集合用线程号着色。线程着色实质是通过各频谱层临界区和信号量集的组合逻辑精确地从全局资源中占用路径和频谱资源。

5) LDRVT-RSA 算法多个线程从 R 中选择 $R^n(i, j, n_{\text{slot}})$ 时,被着色虚拟拓扑互斥且跳数大于 R_n 的业务请求称为长距离业务请求,为长距离业务请求建立光路的线程,简称长距离 RSA 线程。

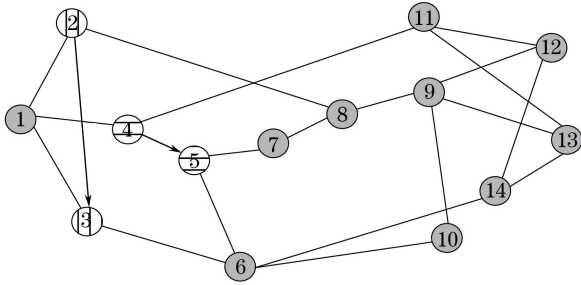


图 1 NSFNET 拓扑

Fig. 1 Topology of NSFNET

LDRVT-RSA 算法用各个频谱层中着色虚拟拓扑作为多个路径不重合业务资源互斥使用的约束条件,同时为多个路径不重合业务请求建立光路来减少突发业务的排队时延,算法资源使用模型设计如下:

1) 算法在光路建立过程中,用点信号量集合实现路由的分配,用边信号量集合实现频谱资源的分配。先占用频谱层中点集合资源,判断频谱资源后,根据判断结果占用或不占用频谱层中边集合的资源。算法对资源使用的约束条件是在光路的建立过程中形成的,这是二维的资源使用方式,构成基于点和边的二维资源互斥式分配模型。

2) 模型中资源使用的实质是对频隙资源的使用,但如果用频隙资源作为资源使用的单位,大量的判断和读写操作将降低频谱资源的使用效率,因此算法根据 RSA 算法的频谱一致性约束,将光纤链路中频谱资源平均分为多个频谱层,不同的线程在不同频谱层中建立光路,增加光路并行建立效率。算法不考虑节点可能存在的带波长转换功能的器件,增加了算法的通用性。

3) 业务请求新建光路时不能使用着色虚拟拓扑中的资源,只能在可用虚拟拓扑中查找合适的资源,对业务请求进行线程着色。光路建立后,使用的点和边集合将从频谱层着色虚拟拓扑中删除,然后和频谱层中可用虚拟拓扑连接。

4) 只有进入频谱层着色虚拟拓扑中且颜色相同的节点间的边资源才能被对应的线程判断频谱资源是否充足,模型通过点互斥来规避 SECM-RSA 算法的多个线程对频谱资源重复判断和无效占用。算法在没有成功进行路由频谱分配时,能及时刷新着色虚拟拓扑,释放路径上着色的节点,减少对边信

号量的操作和频谱资源的占用,减少算法的时延。

5) 因为判断点集合相对于判断边集合来说减少了算法执行的时间复杂度,模型首先判断路径上的点集合,增加业务请求查找可用资源的速度。另外,模型如果只用点集合作为资源使用的约束条件,不能确定唯一路径,例如已知节点集用到节点 V_1 、 V_2 和 V_3 ,路径可能为 $V_1 \rightarrow V_2 \rightarrow V_3$ 或 $V_1 \rightarrow V_3 \rightarrow V_2$ 。模型如果只用频谱层中边信号量集合的组合逻辑也可实现资源互斥使用,但是查找路径不重合业务请求的效率不高。

3.2 算法控制模型

为了描述不同控制模型的性能,对控制模型中所使用符号的定义如表 2 所示。

表 2 控制模型中符号定义及其意义

Table 2 Symbols in control model and their meanings

Symbol	Meaning
N_r	Number of requests
N_{th}	Number of threads in all spectrum layers
T_i	Delay of scheme shaping for the i^{th} request
T	Delay constant
$Q(n)$	Queue delay of the n^{th} request
$f(n)$	Establishment delay of optical path of the n^{th} request
$A(n)$	Average establishment delay of optical path of n requests

传统 RSA 算法,以 SPFF-RSA 算法为例,在突发业务大量聚集的时候,只能逐条为业务请求建立光路,在为单个业务请求建立光路期间,所有资源全部被主函数线程占用,通常有如下步骤:1)业务请求到达;2)主函数线程占用全部资源为业务请求建立光路;3)业务请求光路建立成功或阻塞。算法在图 1 拓扑中建立光路时节点资源锁定时序如图 2 所示。

图 2 由节点、时间、主函数线程构成,单个业务请求大约需要两个时间间隔,在 15 个时间间隔内,共有 7 个业务请求逐条得到响应。虚线内时间是为一个业务请求建立光路的时延,方块为计算 RSA 方案的时延,虚线内其他时间开销包括设置定时器、刷新系统资源和计算方案提供时延。在这个时间区间内,网络中所有节点全部由主函数线程锁定,但不是所有节点在被锁定期间都会用到,这导致网络资源浪费,增加业务请求的平均光路建立时延。本研究用业务请求光路建立时延的数学平均值来计算业务请求平均光路建立时延。根据 $M|M|1$ 排队模型,第 n 个业务请求的光路建立时延为排队(等待)时延 $Q(n)$ 和策略处理(服务)时延 T_n 之和:

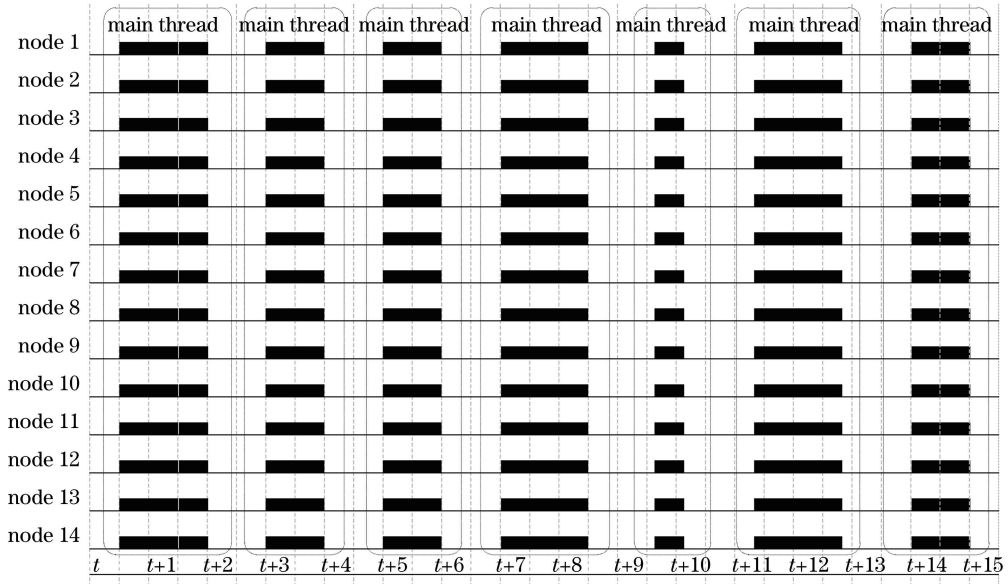


图 2 SPFF-RSA 算法的节点资源使用时序

Fig. 2 Time sequence of nodes resource utilization to SPFF-RSA algorithm

$$f(n) = Q(n) + T_n. \quad (1)$$

根据实验数据测得,各业务请求的策略处理时延相差不大,为理论上粗略推导业务请求平均光路建立时延,将单个业务请求的策略处理时延近似为 T ,即 T_i 近似为 T :

$$Q(n) = \sum_{i=1}^{n-1} T_i, \quad (2)$$

将(2)式代入(1)式,得到第 n 个业务请求的光路建立时延为

$$f(n) = nT. \quad (3)$$

N_r 个业务请求的平均光路建立时延为

$$A(N_r) = \frac{1}{N_r} \sum_{i=1}^{N_r} f(i), \quad (4)$$

由(4)式得到 N_r 个业务请求的平均光路建立时延为

$$A(N_r) = T(N_r + 1)/2. \quad (5)$$

由(5)式可知,SPFF-RSA 算法的平均光路建立时延与业务请求个数 N_r 呈线性关系,LDRVT-RSA 算法的主要目标就是降低 $A(N_r)$ 随 N_r 增加的速度,以适应跨层联合优化架构中突发业务疏导的场景.LDRVT-RSA 算法在上述资源互斥使用模型的基础上并行进行业务请求的光路建立,打破业务请求由单个线程逐个计算的限制,算法的控制模型步骤为:1)业务请求到达;2)多个路由频谱计算线程或长距离线程为业务请求建立光路;3)业务请求光路建立成功或阻塞。

由于算法在各频谱层中建立光路时使用相同的

资源分配模型,本研究用单个频谱层的节点资源锁定时序来说明 LDRVT-RSA 算法的控制模型,并和 SPFF-RSA 算法进行效率比较.LDRVT-RSA 算法在图 1 拓扑中建立光路时单个频谱层节点资源锁定时序如图 3 所示。

图 3 由节点、时间、多个线程构成,在 15 个时间间隔内,共有 19 个业务请求得到响应。在任一时间点,不同路径上的节点被不同的线程通过资源互斥使用的方式锁定,这增加了网络中资源的使用效率,减少了业务请求的平均光路建立时延。图 3 中有长距离线程在为长距离业务请求建立光路,这是因为着色虚拟拓扑将物理拓扑分割成几个部分,长距离业务请求不易被响应。为考虑公平,统计长距离业务的个数,当 R_l 大于 R_u 的时候,为长距离业务请求建立光路。为粗略计算 LDRVT-RSA 算法在所有频谱层中的业务平均光路建立时延,忽略长距离业务请求建立光路对算法效率的影响,忽略算法中各个线程计算量的差别,设线程个数为 N_{th} ,则单个线程的业务请求个数近似为 N_r/N_{th} ,代入(5)式,得到 LDRVT-RSA 算法的业务请求平均光路建立时延为

$$A(N_r) = T(N_r/N_{th} + 1)/2. \quad (6)$$

由(6)式可知 LDRVT-RSA 算法的平均光路建立时延与 N_r 成正比,与 N_{th} 成反比,有效降低 $A(N_r)$ 随 N_r 增加的速度。虽然因为存在长距离线程,且多个线程并行使用资源存在互斥现象,图 3 中的业务请求没有理论计算得多,但是对比图 2 和图 3 可以看出,在同等的时间区间内,LDRVT-RSA 算

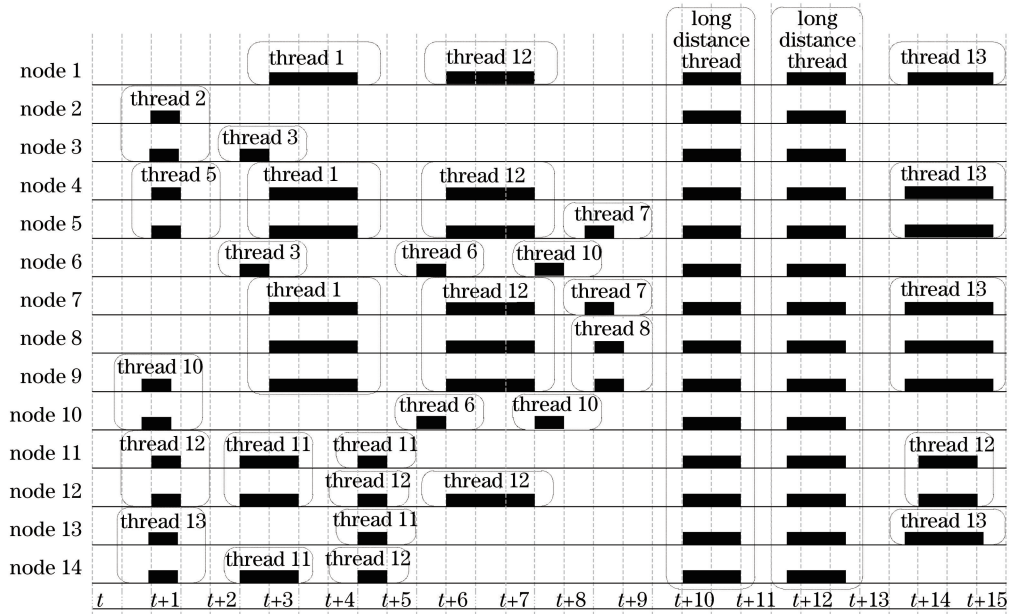


图3 LDRVT-RSA算法的节点资源使用时序

Fig. 3 Time sequence of nodes resource utilization to LDRVT-RSA algorithm

法为更多的业务请求建立了光路。算法对频谱层的设置和着色虚拟拓扑的运用,提高了并行建立光路的效率,充分发挥了多线程SDN控制器的优势。

3.3 算法线程及流程图

LDRVT-RSA算法流程图如图4所示,该算法包括多个线程,按线程优先级由高到低分别为:

1) 算法主函数线程,负责创建其他线程,设置线程优先级,开启线程,设置定时器等。

2) 业务产生线程,负责产生业务,业务的属性包括源点、目标节点、业务级别、业务产生时间、持续时间和光路建立时限等。为测试算法应对突发业务的实用性和可靠性,线程中设计了突发业务产生函数和泊松间隔业务产生函数。突发业务产生函数一次产生1000个业务,泊松间隔业务产生函数按泊松间隔产生业务。

3) 长距离RSA线程,负责长距离业务请求的RSA方案,建立光路后累计业务请求的光路建立时延。为充分运用全局资源建立长距离路径,线程设计为一个非重入线程,使用全网资源。最短路径法(FLOYD)和首次适应频谱分配(FF)算法分别用来分配路由和频谱资源。 R_1 大于 R_u 线程被启动,小于 R_d 线程被休眠。

4) RSA线程,负责源点为线程序号(或线程序号的线性函数值)的业务请求的光路建立,建立光路后累计业务请求的光路建立时延。为减少代码的规模,线程被设计为一个可重入线程,各个重入的线程

优先级相等,重入的各个线程有自己的线程资源,同时互斥地共享全局资源。

3.4 算法实例

如前所述,LDRVT-RSA算法在运行的过程中,能够规避当前频谱层中正在建路业务使用的资源,重构频谱层中可用虚拟拓扑,在频谱层可用虚拟拓扑中实现多线程同时建立光路。本研究用图3中节点、时间、线程、线程号来说明算法二维资源互斥分配模型和控制模型,图3中几个时间点的着色虚拟拓扑和可用虚拟拓扑如表3所示。

从图3中可以看出,在时间点 $t+1$,线程2,5,10,12,13同时建立5条光路,线程号2,5分别代表不同的颜色,被用到 $R(2,3,2)$, $R(4,5,4)$ 的光路上的点 V_2 、 V_3 和 V_4 、 V_5 以及边 $E_{2,3}$ 和 $E_{5,4}$ 的着色,着色虚拟拓扑如图5所示,基于二维资源互斥式分配模型重构的可用虚拟拓扑如图6所示。

假设 $R(1,9,1)$ 在时间点 $t+1$ 前到达,在时间点 $t+1$,线程1为 $R(1,9,1)$ 建立光路时被着色虚拟拓扑互斥,在时间点 $t+2.6$,线程1在可用虚拟拓扑中判断到可用资源,开始为 $R(1,9,1)$ 建立光路。

在时间点 $t+5$,基于二维资源互斥式分配模型重构的可用虚拟拓扑如图7所示。在时间点 $t+10$,主函数线程判断到 R_1 大于 R_u ,休眠所有RSA线程,启动长距离RSA线程,为长距离业务请求建立光路。在时间点 $t+13$, R_1 小于 R_d ,主函数线程休眠长距离RSA线程,RSA线程被唤醒。

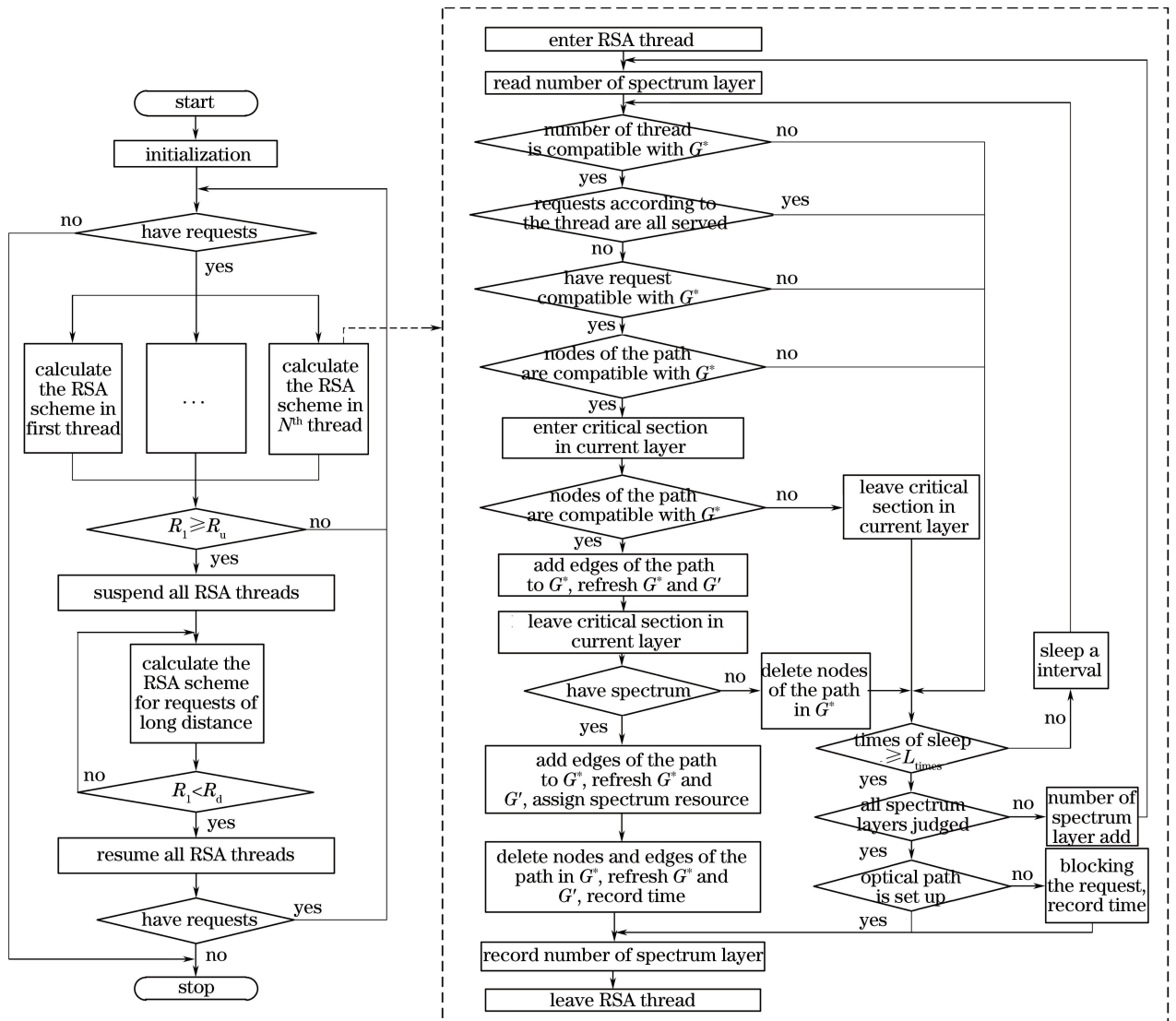


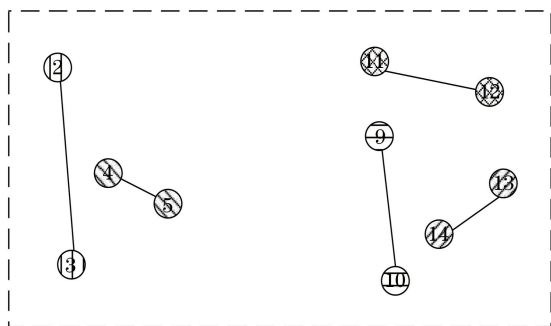
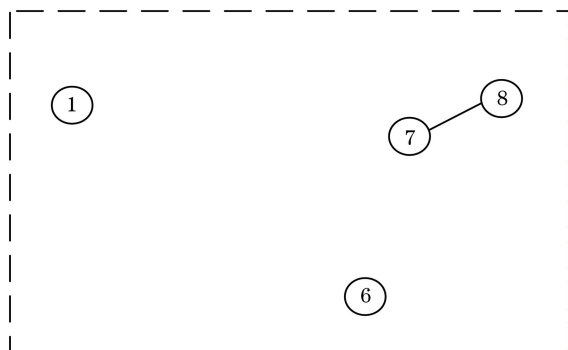
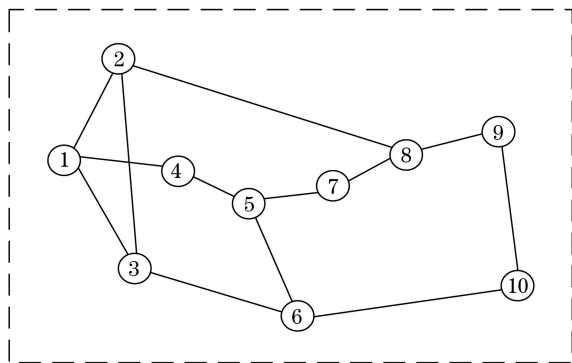
图 4 LDRVT-RSA 算法流程图

Fig. 4 Flow chart of LDRVT-RSA algorithm

表 3 着色虚拟拓扑和可用虚拟拓扑重构

Table 3 Reconfiguration of the coloring and usable virtual topologies

Time	{Nodes} thread in colored virtual topology	{Path} thread in colored virtual topology	Usable nodes in usable virtual topology	Usable edges in usable virtual topology
	$\{V_2, V_3\}2, \{V_4, V_5\}5$	$\{2 \rightarrow 3\}2, \{5 \rightarrow 4\}5$		
$t+1$	$\{V_9, V_{10}\}10$ $\{V_{11}, V_{12}\}12$ $\{V_{13}, V_{14}\}13$ $\{V_1, V_4, V_5, V_7, V_8, V_9\}1$	$\{10 \rightarrow 9\}10$ $\{12 \rightarrow 11\}12$ $\{13 \rightarrow 14\}13$ $\{1 \rightarrow 4 \rightarrow 5 \rightarrow 7 \rightarrow 8 \rightarrow 9\}1$	V_1, V_6, V_7, V_8	$E_{7,8}$
$t+3$	$\{V_3, V_6\}3$ $\{V_{11}, V_{12}, V_{14}\}11$	$\{3 \rightarrow 6\}3$ $\{11 \rightarrow 12 \rightarrow 14\}11$	V_2, V_{10}, V_{13}	$E_{1,2}, E_{1,3}, E_{1,4}$ $E_{2,3}, E_{2,8}, E_{3,6}$ $E_{4,5}, E_{5,6}, E_{5,7}$ $E_{6,10}, E_{7,8}, E_{8,9}, E_{9,10}$
$t+5$	$\{V_{11}, V_{13}\}11$ $\{V_{12}, V_{14}\}12$	$\{11 \rightarrow 13\}/11$ $\{12 \rightarrow 14\}/12$	V_1, V_2, V_3, V_4, V_5 $V_6, V_7, V_8, V_9, V_{10}$	
$t+10$	All nodes are used	All edges are used		

图 5 时间点 $t+1$ 的着色虚拟拓扑Fig. 5 Coloring virtual topology at $t+1$ 图 6 时间点 $t+1$ 的可用虚拟拓扑Fig. 6 Usable virtual topology at $t+1$ 图 7 时间点 $t+5$ 的可用虚拟拓扑Fig. 7 Usable virtual topology at $t+5$

4 仿真结果与分析

4.1 仿真参数设置

本研究中的仿真程序用 VC++ 编写,在 VS2010 开发平台上运行。算法采用 NSFNET 网络拓扑,该拓扑包含 14 个节点和 21 个双向光连接链路,如图 1 所示。算法运行的参数配置具体如下:光纤链路中有 8.95 THz 频谱资源,12.5 GHz 是划分的最小频隙,则可用的频谱区域有 716 个频隙,频谱被分为 5 个频谱层,各频谱层有 140 个频隙,其余为频谱层间隔,线程总数为 17 个。到达网络的业务请求由突发业务和到

达网络的时间间隔服从泊松分布的业务组成,持续时间服从负指数分布。业务请求光路建立分别需要 1, 2, 4 个频隙,比例分别为 25%, 50%, 25%。在上述的仿真环境下,分别多次运行 SPFF-RSA 算法、SECM-RSA 算法、LDRVT-RSA 算法, FLOYD 和 FF 算法分别被用来分配路由和频谱资源。

4.2 评价指标与时间标尺

本研究提出基于虚拟拓扑重构的低时延 RSA 算法,旨在降低数据中心跨层联合优化架构中光路平均建立时延,快速疏导聚集在光网的突发业务。为衡量算法性能,选取业务请求平均光路建立时延、光路建立时延方差、业务请求阻塞率、频谱利用率为评价指标对算法进行仿真。

上文对 SPFF-RSA 算法和 LDRVT-RSA 算法的运行时间已粗略进行了理论计算,但是在实验仿真中需要更精确的时间标尺,来对比不同算法光路建立的平均时延。VC++ 内部自带的时延函数能精确到毫秒级,运用 VS2010 开发平台中 VC++ 的两个函数通过系统调用的方式访问计算机的精确定时器,来精确测量业务请求的光路建立时延。在进行定时之前,先调用 Query Performance Frequency () 函数读取机器内部定时器的时钟频率 (OSCCLK, f_{OSCCLK}),然后在第 n 个业务请求到达时和应答或阻塞业务请求后分别调用 Query Performance Counter () 函数得到 T_{n1} 和 T_{n2} ,利用两次获得的计数之差及时钟频率,计算出业务请求光路建立的精确时延 T_{act} :

$$T_{\text{act}}(n) = (T_{n1} - T_{n2}) / f_{\text{OSCCLK}}, \quad (7)$$

成功实现十微秒级时延的精确测量是证明算法性能达到预期目标的关键。

LDRVT-RSA 算法为确保虚拟拓扑实现的资源互斥使用模型有效,需要对算法进行正确性测试。在算法的测试程序中,在拆除业务请求光路时,将 $\lambda_{i,j}^n$ 减 1,为每个业务请求分配频谱资源时,将业务请求的频谱资源变量 $\lambda_{i,j}^n$ 加 1,并判断 $\lambda_{i,j}^n$ 的数值,如果 $\lambda_{i,j}^n$ 大于 1,说明频谱资源存在重复使用。程序在 $\lambda_{i,j}^n$ 大于 1 的位置设置断点,算法反复多次运行,未进入断点,说明算法通过运用二维资源互斥分配模型,有效规避了并行建立光路对频谱资源的重复使用。

4.3 结果分析

各算法单独进行时延统计,以减少统计阻塞率和频谱利用率的代码对时延统计的干扰,算法的时延性能比较如图 8 所示。与 SPFF-RSA 算法相比, LDRVT-RSA 算法和 SECM-RSA 算法的光路平均

建立时延更小,当网络负载为 500 erl(其中 erl 是通用的网络负载衡量单位,),光路平均建立时延分别减少 28% 和 15%,这是因为 LDRVT-RSA 和 SECM-RSA 算法运用多线程技术同时为多个业务请求建立光路,减少了业务请求的排队时延;与 SECM-RSA 算法相比,LDRVT-RSA 算法的光路建立平均时延更小,这是因为 LDRVT-RSA 算法运用二维资源互斥分配模型,提高了查找路径不重合业务请求的效率,规避了多个线程对同一条边的频谱资源重复判断和无效占用,减少了多线程算法建路策略处理时延。

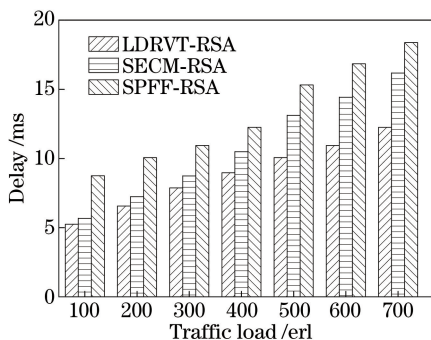


图 8 光路平均建立时延

Fig. 8 Average establishment delay of optical paths

根据各算法的光路建立平均时延统计光路建立时延方差,如图 9 所示,与 SPFF-RSA 算法相比,LDRVT-RSA 算法和 SECM-RSA 算法的光路建立时延方差更小。这是因为 SPFF-RSA 算法中先后到达业务请求的排队时延差值较大,同时和光路建立时延均值相差较大,而 LDRVT-RSA 和 SECM-RSA 算法运用多线程技术同时为多个业务请求建立光路,减少了业务请求的排队时延,从而减少了光路建立时延方差;与 SECM-RSA 算法相比,LDRVT-RSA 算法的光路建立时延的方差较小,这是因为点和边集合分别在路由和频谱分配时,多次对资源使用进行有效的约束,从而有效梳理了算法,增加了算法的稳定性,减少了光路建立时延方差。

频谱利用率比较结果如图 10 所示,与 SPFF-RSA 算法和 SECM-RSA 算法相比,LDRVT-RSA 算法的频谱利用率更高,当网络负载为 600 erl,LDRVT-RSA 算法的频谱利用率分别增加了 9% 和 4%,这是因为 LDRVT-RSA 算法光路建立时延较低,二维资源互斥分配模型提高了频谱分配效率,频谱分层技术提高了频谱使用效率,从而提高了频谱资源利用率。阻塞率结果比较如图 11 所示,LDRVT-RSA 算法的阻塞率低于 SPFF-RSA 和

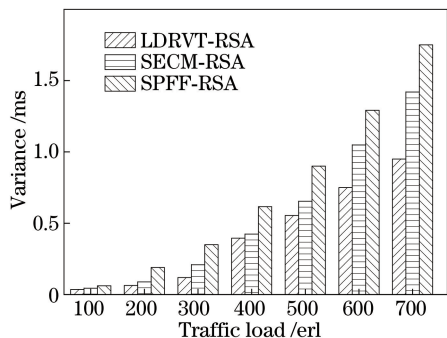


图 9 光路建立时延方差

Fig. 9 Variance of establishment delay of optical paths

SECM-RSA 算法的阻塞率,当网络负载为 650 erl,LDRVT-RSA 算法的阻塞率分别减少 6% 和 5%,这是因为 LDRVT-RSA 算法具有更高的频谱资源利用率和更低的光路建立时延,间接减少了业务请求阻塞率。与 SPFF-RSA 算法和 SECM-RSA 算法相比,LDRVT-RSA 算法用多个线程同时建立光路,光路的平均建立时延较低,频谱使用效率更高,具有更高的为业务请求建立光路的能力,能快速疏导聚集在光网的业务请求,提升了用户体验质量。

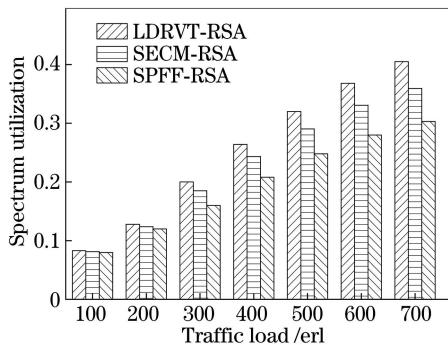


图 10 频谱利用率

Fig. 10 Rate of spectrum utilization

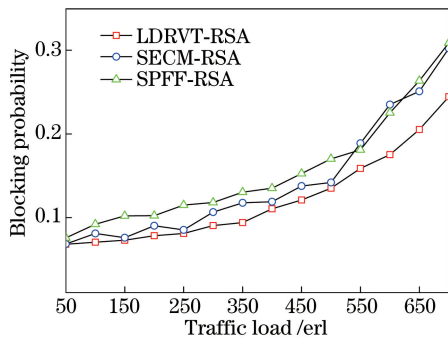


图 11 阻塞率

Fig. 11 Rate of blocking probability

5 结 论

提出了数据中心跨层联合优化架构中低时延

RSA 算法,通过对仿真结果分析可知 LDRVT-RSA 算法通过二维资源互斥分配模型、控制模型和频谱分层技术,精确地实现了全局资源共享,各线程同时为业务请求建立光路。与 SPFF-RSA 算法和 SECM-RSA 算法相比,LDRVT-RSA 算法具有更低的光路平均建立时延、光路建立时延方差、阻塞率,更高的频谱利用率,提升了用户体验质量。由于多线程 SDN 控制器是光网络的发展方向,后续的工作将研究在频谱分区中运用多线程 RSA 技术实现低物理损伤低时延路由频谱分配。

参 考 文 献

- [1] Yang H. Research of software defined networking and integrated control for data center optical interconnection[D]. Beijing: Beijing University of Posts and Telecommunications, 2014: 56-60.
杨辉. 面向数据中心光互联的软件定义组网与集成控制研究[D]. 北京: 北京邮电大学, 2014: 56-60.
- [2] Jiang X K, Zhao F, Fan Y Q, *et al.* Frequency assignment algorithm for elastic optical network with multi-cores considering crosstalk [J]. *Laser & Optoelectronics Progress*, 2017, 54(6): 060601.
江祥奎, 赵峰, 范永青, 等. 考虑串扰的多纤芯弹性光网络中的频谱分配算法[J]. *激光与光电子学进展*, 2017, 54(6): 060601.
- [3] Yang H, Zhang J, Zhao Y L, *et al.* Performance evaluation of multi-stratum resources integrated resilience for software defined inter-data center interconnect[J]. *Optics Express*, 2015, 23(10): 13384-13398.
- [4] Yamashita S, Yamada A, Nakatsugawa K, *et al.* Extension of OpenFlow protocol to support optical transport network, and its implementation[C]//2015 IEEE Conference on Standards for Communications and Networking (CSCN), October 28-30, 2015, Tokyo, Japan. New York: IEEE, 2015: 263-268.
- [5] Yang H, Zhang J, Zhao Y L, *et al.* Software defined clustered-optical access networking for ubiquitous data center optical interconnection [C] // 2016 International Conference on Computing, Networking and Communications (ICNC), February 15-18, 2016, Kauai, HI, USA. New York: IEEE, 2016: 1-5.
- [6] Yang H, Cheng L, Deng J, *et al.* Cross-layer restoration with software defined networking based on IP over optical transport networks [J]. *Optical Fiber Technology*, 2015, 25: 80-87.
- [7] Zhang L, Tizghadam A, Bannazadeh H, *et al.* Iterative traffic engineering in the data plane of multimedia IP communications [C] // 2016 IEEE NetSoft Conference and Workshops (Net Soft), June 6-10, 2016, Seoul, South Korea. New York: IEEE, 2016: 107-111.
- [8] Lord A, Soppera A, Jacquet A. The impact of capacity growth in national telecommunications networks[J]. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 2016, 374(2062): 20140431.
- [9] Xuan H J, Wang Y P, Xu Z Q, *et al.* Core selection algorithm for multi-core elastic optical networks[J]. *Acta Optica Sinica*, 2016, 36(12): 1206005.
宣贺君, 王宇平, 徐展琦, 等. 多纤芯弹性光网络中纤芯选择算法[J]. *光学学报*, 2016, 36(12): 1206005.
- [10] Rozic C, Klonidis D, Tomkos I. Latency-aware multi-layer network optimization in IP-over-WDM core networks[C]//Proceedings of ECOC 2016; 42nd European Conference on Optical Communication, September 18-22, 2016, Dusseldorf, Germany. Frankfurt: VDE, 2016: 1-3.
- [11] Yi P, Ramamurthy B. Provisioning virtualized cloud services in IP/MPLS-over-EON networks [J]. *Photonic Network Communications*, 2016, 31(3): 418-431.
- [12] Ye F, Gao J H, Zhou X P. Wavelength distribution method in optical communication system, optical network unit and optical network system: CN102388557A[P]. 2012-03-21.
叶飞, 高建河, 周小平. 光通信系统中的波长分配方法、光网络单元及光网络系统: CN102388557A[P]. 2012-03-21.
- [13] Li F. Wavelength routing method and system for optical network: CN101754058A[P]. 2010-06-23.
李非. 一种用于光网络的波长路由方法及系统: CN101754058A[P]. 2010-06-23.
- [14] Guo H, Yang R L, Wang W Y, *et al.* Physical layer impairment-aware and less delay RSA algorithm in flexible optical networks[J]. *Optical Communication Technology*, 2017, 41(5): 15-18.
郭宏, 杨润利, 王文彦, 等. 基于物理损伤感知的低时延 RSA 算法[J]. *光通信技术*, 2017, 41(5): 15-18.
- [15] Lu W, Zhou X, Gong L, *et al.* Dynamic multi-path service provisioning under differential delay constraint in elastic optical networks[J]. *IEEE Communications*

- Letters, 2013, 17(1): 158-161.
- [16] Yang H, Zhang J, Zhao Y L, *et al.* Time-aware software defined networking for OpenFlow-based datacenter optical networks [J]. *Network Protocols and Algorithms*, 2014, 6(4): 77-91.
- [17] Magalhães F G D, Priti R, Nikdast M, *et al.* Design and modelling of a low-latency centralized controller for optical integrated networks [J]. *IEEE Communications Letters*, 2016, 20(3): 462-465.
- [18] Zhang M Y, You C S, Zhu Z Q. On the parallelization of spectrum defragmentation reconfigurations in elastic optical networks [J]. *IEEE/ACM Transactions on Networking*, 2016, 24(5): 2819-2833.
- [19] Meng F. Implementation scheme of low latency for optical transport networks [J]. *Telecommunications Science*, 2016, 32(S1): 238-244.
孟凡. 低时延光传送网实现方案 [J]. *电信科学*, 2016, 32(S1): 238-244.
- [20] Wan X, Hua N, Zhang H, *et al.* Study on dynamic routing and spectrum assignment in bitrate flexible optical networks [J]. *Photonic Network Communications*, 2012, 24(3): 219-227.
- [21] Talebi S, Alam F, Katib I, *et al.* Spectrum management techniques for elastic optical networks: a survey [J]. *Optical Switching and Networking*, 2014, 13: 34-48.
- [22] Alaskar R W, Ahmad I, Alyatama A. Offline routing and spectrum allocation algorithms for elastic optical networks [J]. *Optical Switching and Networking*, 2016, 21: 79-92.
- [23] Zhao Y, Iannone L, Riguiedel M. On the performance of SDN controllers: a reality check [C] // 2015 IEEE Conference on Network Function Virtualization and Software Defined Network (NFV-SDN), November 18-21, 2015, San Francisco, CA, USA. New York: IEEE, 2015: 79-85.
- [24] Shalimov A, Zuikov D, Zimarina D, *et al.* Advanced study of SDN/OpenFlow controllers [C] // Proceedings of the 9th Central & Eastern European Software Engineering Conference in Russia (CEE-SECR), October 24-25, 2013, New York, NY, USA. New York: ACM, 2013: 1.
- [25] Scott-Hayward S. Design and deployment of secure, robust, and resilient SDN controllers [C] // 2015 1st IEEE Conference on Network Softwarization (NetSoft), April 13-17, 2015, London, UK. New York: IEEE, 2015: 1-5.
- [26] Mijumbi R, Serrat J, Gorricho J L, *et al.* Network function virtualization: state-of-the-art and research challenges [J]. *IEEE Communications Surveys & Tutorials*, 2016, 18(1): 236-262.
- [27] Peng L, Chen M, Park K, *et al.* Virtual-pod-assisted routing and resource assignment in elastic all-optical intra-datacenter networks [J]. *IEEE Access*, 2016, 5: 406-420.
- [28] Zhao J J, Zheng D, Hu J H, *et al.* Node-first virtual network mapping algorithm in elastic optical network [J]. *Chinese Journal of Lasers*, 2017, 44(9): 0906003.
赵继军, 郑斗, 胡劲华, 等. 弹性光网络中节点优先的虚拟网络映射算法 [J]. *中国激光*, 2017, 44(9): 0906003.
- [29] Ghazisaeedi E, Huang C. Energy-aware node and link reconfiguration for virtualized network environments [J]. *Computer Networks: The International Journal of Computer and Telecommunications Networking*, 2015, 93(P3): 460-479.
- [30] Eramo V, Ammar M, Lavacca F G. Migration energy aware reconfigurations of virtual network function instances in NFV architectures [J]. *IEEE Access*, 2017, 5: 4927-4938.
- [31] Wang W, Zhao Y L, He R Y, *et al.* Continuity aware spectrum allocation schemes for virtual optical network embedding in elastic optical networks [J]. *Optical Fiber Technology*, 2016, 29: 28-33.
- [32] Huang H B, Huang S G, Yin S, *et al.* Virtual network provisioning over space division multiplexed optical networks using few-mode fibers [J]. *IEEE/OSA Journal of Optical Communications and Networking*, 2016, 8(10): 726-733.
- [33] Hsu C F, Chang Y C, Sie S C. Graph-model-based dynamic routing and spectrum assignment in elastic optical networks [J]. *IEEE/OSA Journal of Optical Communications and Networking*, 2016, 8(7): 507-520.