

文章编号:2095-6134(2016)04-0554-08

# SDN 网络中基于业务资源偏好的批路由策略<sup>\*</sup>

房秋生, 洪佩琳<sup>†</sup>

(中国科学技术大学信息科学技术学院中国科学院无线光电通信重点实验室, 合肥 230027)  
(2015 年 12 月 9 日收稿; 2016 年 2 月 29 日收修改稿)

Fang Q S, Hong P L. Batch routing in SDN networks with resource preference consideration[J]. Journal of University of Chinese Academy of Sciences, 2016, 33(4): 554-561.

**摘 要** SDN 网络中有限的交换机流表资源限制流经节点的业务流数目, 带宽资源限制流经链路的业务流的数据流量. 本研究基于业务流特点, 提出业务资源偏好的概念; 基于 SDN 网络集中控制的特点, 提出使用批路由策略处理多个同时到达 SDN 控制器的业务流请求. 设计了 SDN 网络中基于业务资源偏好的批路由策略 BRP-SA. 仿真结果表明, BRP-SA 算法有效地均衡流表资源和带宽资源的使用, 使网络接纳更多的业务流请求.

**关键词** 软件定义网络; 批路由; 业务资源偏好; 流表资源

中图分类号: TN929.5      文献标志码: A      doi:10.7523/j.issn.2095-6134.2016.04.018

## Batch routing in SDN networks with resource preference consideration

FANG Qiusheng, HONG Peilin

(Key Laboratory of Wireless-Optical Communications of Chinese Academy of Sciences, School of Information Science and Technology, University of Science and Technology of China, Hefei 230027, China)

**Abstract** In SDN networks, the limited flow tables at switches confine the number of flows which can pass through the OpenFlow switches. The limited bandwidth resources confine the data traffic which can pass through the network links. A concept of resource preference is suggested based on traffic characteristics, and batch routing strategy is proposed to process multiple flow requests simultaneously. We first model and formulate the batch routing optimization problem, and then present a heuristic algorithm called BRP-SA to execute batch routing algorithm with resource preference consideration. Simulation results show that BRP-SA effectively balances the utilization of both flow table and bandwidth resources, and then the network accepts more flow requests.

**Key words** SDN; batch routing; flow resource preference; flow table resource

SDN 网络定义了一个新的网络架构, 其转发平面和数据平面分离, 并通过控制器实现对整个网络逻辑上的集中控制. SDN 控制器不仅可以获取全局网络拓扑信息, 而且可以通过向交换机发送询问消息获取网络状态统计信息<sup>[1]</sup>. 与传统网络中的路由器不同, SDN 网络中的交换机缺乏智

<sup>\*</sup> 国家自然科学基金面上基金(61390513, 61170231)资助  
<sup>†</sup> 通信作者, E-mail: plhong@ustc.edu.cn

能性.对新到达的业务流,如果交换机中没有对应的流表项匹配,则此交换机会通过安全通道向控制器发送一个 packet-in 消息,以指示一个新的业务流请求.某一时刻,SDN 控制器可能收到网络中各个交换机发来的多个业务流请求,因此本文考虑提出 SDN 网络中的批路由策略用于同时解决多个业务流的路由请求问题.批路由在传统网络中已有研究,但受到传统网络分布式架构的限制,应用起来十分困难.而 SDN 网络采用集中式的网络架构,可以通过控制器对网络业务流提供集中式的管理,这使得批路由在 SDN 网络中的实际应用成为可能.批路由在现实中的应用场景主要体现在批路由队列网络,网络中的各种业务流顺序到达,形成一个业务流请求队列.如果选择批处理方式,则 SDN 控制器可以实现对队列中的多个业务流请求同时处理;否则 SDN 控制器根据队列中的业务流请求的优先级逐个计算业务路由路径.

在 SDN 架构中,交换机根据控制器下发的流表项对分组实现匹配转发(如 OpenFlow<sup>[1]</sup>).按照现有 SDN 控制器与交换机的南向标准 OpenFlow 协议,一张流表可以使用任意的字段组合,这种基于匹配转发的模式要求交换机掩盖不需匹配的字段.在当前的商业芯片设计中,只有使用三态内容寻址存储器(ternary content addressable memory, TCAM)才能实现. TCAM 是一种昂贵的资源,具体表现在芯片面积大、功耗大.现有的交换机只支持大约  $1 \sim 2 \text{ k}^{[2-3]}$  的 TCAM 流表项,而且短期内不会增加太多<sup>[4]</sup>.因此,与传统网络路由器的路由表项不同,交换机的流表资源也是极其昂贵的稀缺资源.

SDN 网络中链路带宽资源的不合理使用会导致瓶颈链路的存在,从而引起链路拥塞;流表资源的不合理使用同样会导致瓶颈节点的存在.例如在极端情况下,如果某个交换机上的流表资源已经耗尽,则控制器无法向此交换机下发流表,交换机亦无法再转发新的业务流.若此交换机的邻接链路的带宽资源依然富余,则这些富余的链路带宽受交换机流表资源的限制亦无法再利用.即瓶颈节点的出现亦会导致带宽资源的浪费.因此,在设计 SDN 网络的路由策略时,需合理利用带宽和流表资源,同时防止瓶颈链路和瓶颈节点的存在.

本文基于业务流的特点,对业务流进行分类,并引入业务资源偏好的概念,旨在设计一种 SDN 网络中基于业务资源偏好的批路由策略.通过业务资源偏好,根据业务流特点合理使用网络中的带宽和流表资源;通过批路由策略实现对多个业务流路由的全局调度.从而更加均衡地利用网络资源,以获得更好的网络性能.

## 1 相关工作

近来已有一些工作开始考虑流表资源的限制对 SDN 网络的影响,并提出了一些机制用于合理利用流表资源和带宽资源. Cohen 等<sup>[5]</sup>研究流表资源对 SDN 网络使用率的影响,并研究受流表资源有限性的影响,路径度数受限的网络最大流问题.研究结果显示,SDN 网络交换机中的流表资源越少,网络可以接收的总流量也越少.

Luo 等<sup>[6]</sup>提出一种快速流表聚合算法(FFTA)用于解决 SDN 网络中 TCAM 有限的问题. FFTA 是一种离线的流表聚合解决方案,它通过将多个流表项聚合成一个而不改变转发规则,但是流聚合的方式在细粒度处理业务流请求时并不适用. Feng 等<sup>[7]</sup>通过对不同的应用程序计费,并基于计费提出一个适用于多个应用程序的联合流表和带宽资源的比例公平资源调度算法. Zhang 等<sup>[8]</sup>通过混合路由同时为多个流量矩阵实现负载均衡,其与显示路由相比可以极大地减少流表项的使用.

Banerjee 和 Kannan<sup>[9]</sup>指出可采用混合 TCAM 和 SRAM(static random access memory)方式匹配流表项. TCAM 存储方式可提供快速查找( $O(1)$ ),但是 TCAM 芯片面积大、功耗大,因此其可用大小受限. TCAM 资源少意味着可用流表数目少,这是流表资源受限的根本原因. SRAM 与 TCAM 相比,会降低 50% 的功耗,但会提高 20 多倍的匹配延迟.考虑到实际网络应用的需求,SDN 网络环境更希望交换机有较少的匹配延迟,因此在匹配性能方面 TCAM 优于 SRAM.联合 TCAM/SRAM 的流表存储方式,可以在一定程度上实现功耗和延迟的折中,缓解流表资源的受限问题.

现有网络是基于分布式的路由决策模型,而批路由更适合 SDN 集中控制的网路架构,同时也更好地利用了集中控制的优点. Trivisonno<sup>[10]</sup>提出一种虚拟链路映射算法用于同时网络中的多个业

务流请求. 结果显示虚拟链路映射算法在带宽分配上性能优于最短路径优先算法.

与现有工作不同, 本文考虑了业务流特点和网络资源使用之间的关系, 并据此引入业务资源偏好的概念. 然后结合业务资源偏好和 SDN 集中控制的优点, 设计了一种基于业务资源偏好的批路由策略.

## 2 问题描述

### 2.1 SDN 网络模型

本文将 SDN 网络建模成一个有权无向图  $G = (V, E, \rho, \mu)$ ,  $\rho = V \rightarrow N^+$ ,  $\mu = E \rightarrow R^+$ , 其中  $V$  和  $E$  分别代表网络中交换机和链路的集合. 在设计路由策略时考虑网络中交换机流表资源和链路带宽资源的限制. 在无向加权图中, 流表资源抽象成交换机上的可用流表数目  $\rho$ ; 带宽资源抽象成链路上的可用带宽  $\mu$ . 对任意节点  $i \in V$ ,  $\rho_i$  表示节点  $i$  上的可用流表数; 任意链路  $ij \in E$ ,  $\mu_{ij}$  表示链路  $ij$  的可用带宽.

### 2.2 业务资源偏好

SDN 网络中有限的流表资源限制流经 OpenFlow 交换机的业务流的数目; 与此同时, 网络中的带宽限制流经网络链路的数据流量. 换言之, 业务流对流表资源的消耗与其业务流量大小无关, 只与业务流数目多少有关; 而业务流对带宽资源的消耗与业务数据流量呈正相关. 根据调研, 网络流量服从 Zipf 分布<sup>[11]</sup>, 少数的业务流产生了大量的业务数据流量, 多数业务流的业务数据流量都很少. Cai 等<sup>[12]</sup>通过挖掘网络中的业务流特性证明了上述规律. 这在现代社会中尤其如此, 越来越多的手持终端设备产生了大量的即时消息、心跳消息等, 这些消息的流量很小但是数目巨大. 本文以业务数据流量(带宽需求)为标准将业务流分类为长流<sup>[13]</sup>(视频流等)和短流<sup>[14]</sup>(即时消息等). 根据以上分析可以得出如下结论:

- 1) 长流: 流量大, 流数少, 耗带宽多, 耗流表少;
- 2) 短流: 流量小, 流数多, 耗带宽少, 耗流表多.

由此可以看出长流更偏好带宽资源, 而短流更偏好流表资源. 基于流分类的业务资源偏好的提出, 使得业务流路由策略更符合业务流特点, 以均衡带宽资源和流表资源的使用.

### 2.3 业务流批到达

在 SDN 网络中, 当一个新的业务流到达且在交换机中没有对应的流表项匹配, 则交换机会向控制器发送一个 packet-in 消息. 每个 packet-in 消息表示一个新的业务流请求. 如图 1 所示, 网络中所有的交换机在某一时刻都可能新的业务流到达, 则控制器会收到批到达的业务流请求<sup>[15]</sup>.

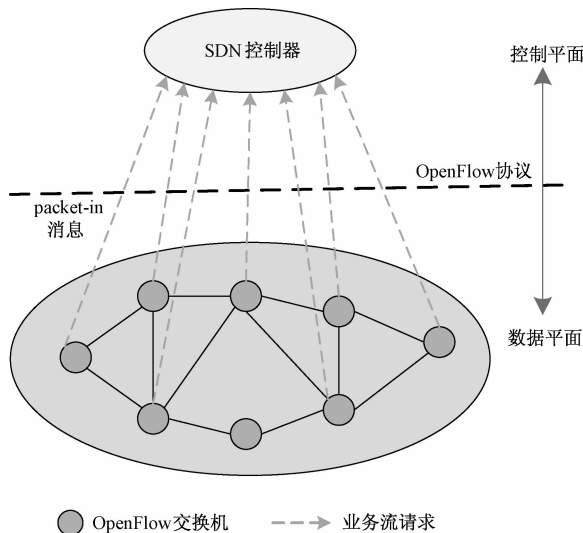


图 1 业务流请求批到达示意图

Fig. 1 Batch arriving flow requests

某一时刻控制器中批到达的新业务流请求总数记为  $K$ ,  $f_k = (s_k, t_k, d_k)$  表示第  $k$  个业务流请求<sup>[16]</sup>,  $k = 1, 2, \dots, K$ . 其中  $(s_k, t_k)$  表示业务流请求  $f_k$  的源-目的地址对,  $d_k$  表示带宽需求.  $F = \{f_k, k = 1, 2, \dots, K\}$  表示批到达的业务流请求的集合.

## 3 系统模型

本节首先基于业务流分类和流量特点定义业务资源偏好和单位路由开销; 然后参考最小开销多商品流模型<sup>[17]</sup>, 为 SDN 网络中基于业务资源偏好的批路由问题建立优化模型.

### 3.1 业务资源偏好度定义

本文基于网络业务流区分提出业务资源偏好度的概念. 业务资源偏好度本质上是基于网络流量特点和业务流特点定义的经验值, 其定义基于业务流分类. 不失一般性, 本文根据业务流带宽需求将其分为长流和短流. SDN 网络中的路由策略需同时考虑链路带宽资源和交换机流表资源, 因此对长流和短流需分别定义带宽资源偏好度和流

表资源偏好度。

根据网络流量特点可以分别获取网络中长流和短流的流数比例和平均带宽需求。其中长流和短流的平均带宽分别用  $d_e$  和  $d_m$  表示,流数比例分别用  $\gamma_e$  和  $\gamma_m$  表示,且  $\gamma_e + \gamma_m = 1$ 。由于网络流量服从 Zipf 定律<sup>[11]</sup>,则有如下关系式成立:

$$\begin{aligned} d_e &> d_m, \\ \gamma_e &< \gamma_m, \\ \gamma_e \times d_e &> \gamma_m \times d_m. \end{aligned} \quad (1)$$

SDN网络中流表资源的使用只与流数有关,而带宽资源使用与业务流的带宽需求和流数均有关。因此,带宽资源偏好度  $\alpha$  和流表资源偏好度  $\beta$  定义成如下形式:

$$\text{长流: } \alpha_e = g\left(\frac{\gamma_e \times d_e}{\gamma_m \times d_m}\right), \beta_e = 1, \quad (2)$$

$$\text{短流: } \alpha_m = 1, \beta_e = g\left(\frac{\gamma_m}{\gamma_e}\right). \quad (3)$$

其中,函数  $g(\cdot)$  表示资源偏好度定义函数,  $\alpha_e$  和  $\beta_e$  对应长流的带宽资源偏好和流表资源偏好,  $\alpha_e > \beta_e$  (长流偏好带宽资源);  $\alpha_m$  和  $\beta_m$  对应短流的带宽资源偏好度和流表资源偏好度,  $\alpha_m < \beta_m$  (短流偏好流表资源)。

### 3.2 单位路由开销定义

本文引入单位路由开销的概念用于衡量链路和节点负载。当业务流在某条路径上传输时,业务流会消耗路径上的链路带宽和交换机流表资源,因此会产生链路开销和节点开销。单位路由开销包含链路开销和节点开销,其定义基于资源使用率、业务流需求和资源偏好度。

用  $\rho_i^0$  表示交换机  $i \in V$  的最大流表资源容量,  $\mu_{ij}^0$  表示链路  $ij \in E$  的最大带宽资源容量。则对业务流请求  $f_k = (s_k, t_k, d_k)$ , 节点流表资源使用率  $\eta_i^k$  和链路带宽资源使用率  $\eta_{ij}^k$  可以定义成如下形式:

$$\eta_i^k = 1 - \frac{\rho_i^k}{\rho_i^0}, \eta_{ij}^k = 1 - \frac{\mu_{ij}^k}{\mu_{ij}^0}. \quad (4)$$

其中,  $\rho_i^k$  表示业务流请求  $f_k$  流经交换机  $i$  时当前可用流表资源,  $\mu_{ij}^k$  表示链路  $ij$  的当前可用带宽资源。  $\eta_i^k$  和  $\eta_{ij}^k$  分别反映当前节点和链路的负载大小。

为实现节点流表资源和链路带宽资源的均衡使用,节点开销应该与交换机流表资源使用率  $\eta_i^k$

成正比;链路开销应该与链路带宽资源使用率  $\eta_{ij}^k$  成正比。节点开销  $c_i^k$  和链路开销  $c_{ij}^k$  定义如下:

$$c_i^k = w(\eta_i^k, \beta^k), c_{ij}^k = w(\eta_{ij}^k, \alpha^k). \quad (5)$$

其中,  $w(\cdot)$  表示开销函数,资源使用率体现当前网络负载情况。业务资源偏好度  $\alpha^k, \beta^k$  由业务流请求  $f_k$  的业务流类型决定,参考式(2)、式(3)。根据式(5),若一条链路的资源使用率已知且固定,受业务资源偏好的影响,不同类型的业务流  $f_k$  流过这条链路,其单位路由开销亦不同。

### 3.3 优化模型

本部分参考最小开销多商品流模型<sup>[17]</sup>对基于业务资源偏好的批路由策略建立数学优化模型,且优化目标为最小化  $K$  个批到达的业务流请求的总路由开销。该优化问题是一个混合整型线性规划问题,其优化模型(边-节点公式)如下所示:

变量说明:

$y_{ij}^k$ : 二进制变量,如果业务流请求  $f_k$  流经链路  $ij$ , 则  $y_{ij}^k = 1$ ; 否则  $y_{ij}^k = 0$ 。

$x_i^k$ : 二进制变量,如果业务流请求  $f_k$  流经交换机  $i$ , 则  $x_i^k = 1$ ; 否则  $x_i^k = 0$ 。

$b_i^k$ : 指示业务流请求  $f_k$  的流供给变量,若  $i$  是路由路径的源节点,则  $b_i^k = 1$ ; 若  $i$  是目的节点,则  $b_i^k = -1$ ; 若  $i$  是中间节点,则  $b_i^k = 0$ 。

$$\begin{aligned} (P1) \quad & \min \sum_{y_{ij}^k, x_i^k} \sum_{k=1}^K \sum_{ij \in E} \{d_k \times c_{ij}^k \times y_{ij}^k + c_i^k \times x_i^k\} \\ \text{s. t.} \quad & (a) \sum_{k=1}^K d_k \times y_{ij}^k \leq \mu_{ij}, \quad \forall ij \in E, \\ & (b) \sum_{k=1}^K x_i^k \leq \rho_i, \quad \forall i \in V, \\ & (c) \sum_{ij \in E} y_{ij}^k - \sum_{ji \in E} y_{ji}^k = b_i^k, \quad k = 1, 2, \dots, K, \\ & (d) x_i^k = y_{ij}^k, \quad k = 1, 2, \dots, K, i \neq t_k, \\ & (e) x_i^k \in \{0, 1\}, y_{ij}^k \in \{0, 1\}, \quad k = 1, 2, \dots, K. \end{aligned} \quad (6)$$

其中,单位链路开销  $c_{ij}^k$  和节点开销  $c_i^k$  的定义参考式(5),其定义基于业务资源偏好和当前网络资源负载。假定批处理的业务流数目为  $K$ ,其优化目标  $P1$  是最小化  $K$  个业务流请求的总的路由开销,这是一个全局优化问题。约束条件(a)表示带宽资源约束,即流经链路的业务流的总带宽需求不能超过当前链路可用带宽资源  $\mu_{ij}$  的限制。约束条件(b)表示流表资源约束,即流经交换机的总的

业务流数目不能超过当前交换机可用流表资源  $\rho_i$  限制. 约束条件(c) 是流守恒约束, 以保证路由路径的连通性和有效性. 约束条件(d) 表示网络拓扑约束, 即流经链路  $ij$  的业务流必然流经交换机  $i$ . 约束条件(e) 表示不可分割流约束, 即每条业务流只在一条路由路径上传输, 不可分割.

优化问题 P1 变量和约束较多, 求解复杂. 为简化问题求解, 引入路径开销  $c_p^k$ , 其定义如下

$$c_p^k = \sum_{ij \in p} c_{ij}^k \times d_k + \sum_{i \in p} c_i^k. \quad (7)$$

此时可将边 - 节点公式 P1 转化成基于路径的优化公式 P2. 如下所示:

$$\begin{aligned} \text{(P2)} \quad & \min \sum_{y_p^k} \sum_{k \in K} \sum_{p \in P(k)} c_p^k \times y_p^k \\ \text{s. t.} \quad & \text{(a)} \sum_{k \in K} \sum_{p \in P(k)} d_k y_p^k \delta_{ij}^p \leq \mu_{ij}, \quad \forall ij \in E, \\ & \text{(b)} \sum_{k \in K} \sum_{p \in P(k)} \delta_{ij}^p y_p^k \leq \rho_i, \quad \forall i \in V, \\ & \text{(c)} \sum_{p \in P(k)} y_p^k = 1, \quad k = 1, 2, \dots, K, \\ & \text{(d)} y_p^k \in \{0, 1\}, \quad k = 1, 2, \dots, K, \\ & \text{(e)} \delta_{ij}^p \in \{0, 1\}. \end{aligned} \quad (8)$$

其中,  $P(k)$  表示业务流请求  $f_k$  的所有源节点  $s_k$  到目的节点  $d_k$  可用路径集合.  $y_p^k \in \{0, 1\}$  是一个二进制变量, 若业务流请求  $f_k$  流经路径  $p \in P(k)$ ,  $y_p^k = 1$ ; 否则  $y_p^k = 0$ . P2 中(a)(b) 分别对应链路带宽约束和交换机流表约束. (c)(d) 对应不可分割流约束. 约束条件(e) 是网络拓扑约束,  $\delta_{ij}^p$  用于指示链路  $ij$  是否在路径  $p \in P(k)$  上.

4 BRP-SA 算法

本节基于模拟退火法<sup>[19]</sup> 提出一个启发式算法 BRP-SA, 用于求解基于业务资源偏好的批路由优化问题 P2 (NP - 难问题<sup>[18]</sup>).

4.1 算法描述

BRP-SA 基于业务流分类和业务资源偏好, 因此首先要实现算法预处理过程: 将 SDN 控制器中批到达的业务流分为长流和短流, 并根据式(2)、式(3) 设定资源偏好度, 根据式(4) 计算单位链路和节点开销.

优化问题 P2 的解是所有  $K$  个业务流请求的路由路径的集合. BRP-SA 算法详细流程如表 1 所示:

步骤 1: BRP-SA 算法设定初始化参数, 包括初始温度  $T_{\max}$ , 最低温度  $T_{\min}$  和算法最大迭代次数  $I$  等.

步骤 2: BRP-SA 算法为优化问题 P2 设定初始解, 即顺序处理批到达的  $K$  个业务流请求, 并根据式(7) 采用 Dijkstra 以此分别计算最小开销路径. 这  $K$  个业务流路径的集合即为 P2 的初始解  $S_0$ . 根据 P2 的优化目标即可计算初始路由总开销  $C(S_0)$ . 当前最优解  $S = S_0$ .

步骤 3: 打乱  $K$  个批到达业务流的处理顺序, 为 P2 计算新解  $S'$  及对应路由总开销  $C(S')$ . 若新解路由总开销比当前最优解小, 则接受新解; 否则以 Metropolis 准则<sup>[19]</sup> 接受新解. 若接受新解, 则更新当前最优解  $S = S'$ . 每次新解处理结束后要降低当前温度.

表 1 BRP-SA 算法  
Table 1 BRP-SA algorithm

1) 初始化算法参数: 设定初始温度 $T_{\max}$ , 最低温度 $T_{\min}$ , 迭代次数 $I$ ;
2) 初始解状态: 设置初始解 $S = S_0$ , 并计算初始路由开销 $C = C(S_0)$ ;
3) 迭代求解: 对 $i = 1, 2, \dots, I$ 迭代第 4) - 7) 步;
4) 产生新解: 获取新解 $S'$ , 计算新解的总路由开销 $C(S')$ ;
5) 决定是否接受新解: $\Delta C' = C(S') - C(S)$ , 若 $\Delta C < 0$ , 则接受新解; 否则以概率 $\exp(-\Delta C'/T)$ 接受新解, $S = S'$ ;
6) 若达到终止条件, 则输出当前解作为最优解;
7) 降温: $t_f = -\ln(T_{\max}/T_{\min})$ , $t = \exp(t_f \times \frac{i}{I})$ .

步骤 4: 迭代执行步骤 2, 并不断更新当前最优解. 若达到终止条件, 则算法终止, 当前最优解  $S$  即为 BRP-SA 算法输出的最优解. 终止条件设定为连续若干个新解都没有接受或者迭代终止.

业务流请求处理顺序会影响网络资源使用率, 因为先处理的业务流会优先占用资源, 增加链路带宽负载和交换机流表负载, 必然会导致单位链路开销和节点开销的增加. 基于引理 4.1<sup>[22]</sup>, 在步骤 3 中引入一种特殊的流处理顺序, 即对批到达的业务流请求根据其带宽需求排序, 优先处理带宽需求比较大的业务流.

引理 4.1 假设  $0 < a_1 < a_2 < \dots < a_n$ , 并且  $0 < b_1 < b_2 < \dots < b_n$ , 则对  $[1, 2, \dots, n]$  的任意排列  $\pi$ , 有以下不等式成立:

$$\sum_{i=1}^n a_i b_{n-i+1} \leq \sum_{i=1}^n a_i b_{\pi(i)} \leq \sum_{i=1}^n a_i b_i. \quad (9)$$

4.2 算法性能分析

BRP-SA 算法设计基于模拟退火法, 通过设

置合理的参数可以在较短时间内收敛到近似最优解. 批处理的业务流个数  $K$  影响算法的迭代次数  $I$ ; 网络拓扑大小用节点数  $n$  表示, 其影响路径计算复杂度  $T_p$ . BRP-SA 算法批处理  $K$  个业务流的总时间复杂度  $T_K$  计算如下:

$$I = O(K), T_p = O(n^2),$$
$$T_K = I \times K \times O(n^2) = O(K^2 \times n^2).$$

(10)

由式(10)可以看出, BRP-SA 可以在多项式时间内求得较优解. 批处理的业务流请求数目  $K$  越大, 算法复杂度也越高, 因此要设定合适的批路由数目, 以实现算法性能和算法复杂度的折中.

## 5 BRP-SA 算法性能分析

### 5.1 仿真设定

仿真基于 Erdos\_Renyi 拓扑模型<sup>[20]</sup>产生一个包含 15 个节点和 34 个边的 SDN 网络拓扑, 设置交换机流表资源容量  $p_i^0$  为 1 500<sup>[2-3]</sup>, 链路带宽资源容量为  $\mu_{ij}^0$  为 1 Gbps. 业务流请求基于网络流特点<sup>[11,21]</sup>设定, 具体参数如表 2 所示.

表 2 业务流请求设定  
Table 2 Flow requests

参数	值
批处理业务流数目 $K$	1, 4, 8
长流平均带宽 $d_c$ /Mbps	6
短流平均带宽 $d_m$ /Mbps	0.2
长短流比例 $\gamma_c : \gamma_m$	0.2 : 0.8
带宽资源偏好度 $\alpha_c, \alpha_m$	$\alpha_c = 5, \alpha_m = 1$
流表资源偏好度 $\beta_c, \beta_m$	$\beta_c = 1, \beta_m = 5$

### 5.2 仿真结果

本文优化目标为最小化总体路由开销, 其本质上是基于业务资源偏好、考虑当前链路负载, 通过批路由方式实现链路带宽资源和节点流表资源的均衡利用. 仿真结果中的“瓶颈节点使用率”和“瓶颈链路使用率”从一定程度上反映当前网络的资源均衡情况, 是最小化总体路由开销后流表资源和链路带宽资源均衡程度的反映. 而本文中的另一个仿真结果“业务流请求接受率”则是在流表和带宽资源均衡使用后的直接结果, 即网络可以接纳更多的业务流请求.

图 2 比较了随着网络业务负载增加, 贪心路由算法 GA 和 BRP-SA 算法在不同的  $K$  取值时的瓶颈交换机流表资源使用率情况, 其中 GA 为传

统考虑链路带宽的最小开销路径算法. 仿真结果显示 BRP-SA 算法在均衡交换机流表资源使用的性能上优于 GA 算法. 但是, BRP-SA 在不同  $K$  值下的流表资源均衡性能并无明显差别. 这表明 BRP-SA 路由算法由于考虑到流表资源的有限性和业务资源偏好, 可以更加均衡地使用交换机流表资源. 但是, 批路由策略并不能明显提高流表资源的均衡使用率, 这是因为流表资源的使用只与业务流数目和业务流路径长度(跳数)有关, 而 BRP-SA 选择的是最小总开销的路由路径集合, 并不能保证路径长度最短.

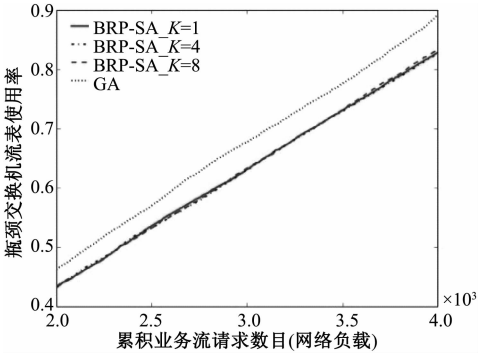


图 2 瓶颈交换机流表使用率

Fig. 2 Bottleneck switch flow table utilization

图 3 比较了随着网络负载增加, 贪心路由算法 GA 和 BRP-SA 算法在不同  $K$  取值时的瓶颈链路带宽资源使用率. 仿真结果表明, BRP-SA 算法可以在考虑流表资源限制的同时, 较好地均衡链路带宽资源的使用. BRP-SA 算法均衡带宽资源的性能随着  $K$  值的增大逐渐提高. 另外, 结果同时表明, 若不考虑批路由策略, 受交换机流表资源的影响, 算法在均衡流表资源的同时, 会降低链路带宽资源的性能.

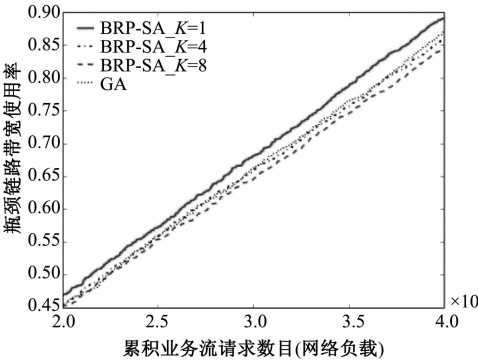


图 3 瓶颈链路带宽使用率

Fig. 3 Bottleneck link bandwidth utilization

图 4 比较了随着网络业务负载增加,贪心路由算法 GA 和 BRP-SA 算法在不同  $K$  取值时的网络业务流请求接受率. 仿真结果表明采用 BRP-SA 算法可以使 SDN 网络接收更多的业务流请求,而且业务流请求的接受率随着  $K$  值的增加而提高.

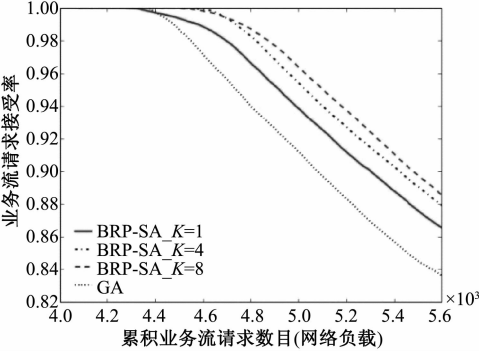


图 4 业务流请求接受率  
Fig.4 Acceptance ratio of flow requests

结合图 2、图 3、图 4 的仿真结果,可以得出如下结论:业务资源偏好和批路由具有互补性,基于业务流特点和业务资源偏好可均衡使用流表资源,基于批路由策略可均衡使用带宽资源. BRP-SA 算法通过基于业务资源偏好的批路由策略,可同时实现这两种网络资源的均衡使用,从而极大提高网络资源使用率. 因此, BRP-SA 可使网络接受更多的业务流请求.

6 结论

本文提出一种 SDN 网络中基于业务资源偏好的批路由策略 BRP-SA. 首先, BRP-SA 考虑 SDN 网络中的一种特殊的网络资源,即流表资源(TCAM)的有限性. 其次,基于网络流特点实现业务流分类,提出业务资源偏好的概念,以根据业务流偏好分配带宽和流表资源. 最后, BRP-SA 考虑 SDN 网络架构集中控制的特点,提出批路由策略用同时处理多个业务流请求. 仿真结果表明 BRP-SA 可同时均衡带宽资源和流表资源的使用,防止由于网络资源不合理使用导致的现瓶颈节点和瓶颈链路问题. BRP-SA 的设计使得 SDN 网络可接受更多的业务流请求.

参考文献

[ 1 ] McKeown N, Anderson T, Balakrishnan H, et al. OpenFlow: enabling innovation in campus networks[J]. ACM SIGCOMM

Computer Communication Review, 2008, 38(2): 69-74.  
[ 2 ] Curtis A R, Mogul J C, Tourrilhes J, et al. DevoFlow: scaling flow management for high-performance networks[J]. ACM SIGCOMM Computer Communication Review, 2011, 41(4): 254-265.  
[ 3 ] Huang D Y, Yocum K, Snoeren A C. High-fidelity switch models for software-defined network emulation [ C ] // Proceedings of the second ACM SIGCOMM workshop on Hot topics in software defined networking. ACM, 2013: 43-48.  
[ 4 ] Moshref M, Yu M, Govindan R, et al. DREAM: dynamic resource allocation for software-defined measurement [ J ]. ACM SIGCOMM Computer Communication Review, 2015, 44(4): 419-430.  
[ 5 ] Cohen R, Lewin-Eytan L, Naor J S, et al. On the effect of forwarding table size on SDN network utilization [ C ] // INFOCOM, 2014 Proceedings IEEE. IEEE, 2014: 1 734-1 742.  
[ 6 ] Luo S, Yu H, Li L. Practical flow table aggregation in SDN [ J ]. Computer Networks, 2015, 92: 72-88.  
[ 7 ] Feng T, Bi J, Wang K. Joint allocation and scheduling of network resource for multiple control applications in SDN[ C ] // Network Operations and Management Symposium (NOMS), 2014 IEEE. IEEE, 2014: 1-7.  
[ 8 ] Zhang J, Xi K, Luo M, et al. Load balancing for multiple traffic matrices using sdn hybrid routing [ C ] // High Performance Switching and Routing (HPSR), 2014 IEEE 15th International Conference on. IEEE, 2014: 44-49.  
[ 9 ] Banerjee S, Kannan K. Tag-in-tag: Efficient flow table management in sdn switches [ C ] // Network and Service Management (CNSM), 2014 10th International Conference on. IEEE, 2014: 109-117.  
[ 10 ] Trivisonno R, Vaishnavi I, Guerzoni R, et al. Virtual links mapping in future sdn-enabled networks [ C ] // Future Networks and Services (SDN4FNS), 2013 IEEE SDN for. IEEE, 2013: 1-5.  
[ 11 ] Sarrar N, Uhlig S, Feldmann A, et al. Leveraging Zipf's law for traffic offloading [ J ]. ACM SIGCOMM Computer Communication Review, 2012, 42(1): 16-22.  
[ 12 ] Cai Y, Wu B, Zhang X, et al. Flow identification and characteristics mining from internet traffic with hadoop[ C ] // Computer, Information and Telecommunication Systems (CITS), 2014 International Conference on. IEEE, 2014: 1-5.  
[ 13 ] Xiao P, Qu W, Qi H, et al. An efficient elephant flow detection with cost-sensitive in SDN[ C ] // Industrial Networks and Intelligent Systems (INISCom), 2015 1st International Conference on. IEEE, 2015: 24-28.  
[ 14 ] Hegde S, Koolagudi S G, Bhattacharya S. Scalable and fair forwarding of elephant and mice traffic in software defined networks[J]. Computer Networks, 2015, 92: 330-340.  
[ 15 ] Min G, Wu Y, Wang L, et al. Performance modelling of

- adaptive routing in hypercubic networks under non-uniform and batch arrival traffic [ C ] // Local Computer Networks, 2007. 32nd IEEE Conference on. IEEE, 2007: 583-590.
- [ 16 ] A Allalouf M, Shavitt Y. Centralized and distributed algorithms for routing and weighted max-min fair bandwidth allocation [ J ]. Networking, IEEE/ACM Transactions on, 2008, 16(5): 1 015-1 024.
- [ 17 ] Barnhart C, Hane C A, Vance P H. Using branch-and-price-and-cut to solve origin-destination integer multicommodity flow problems [ J ]. Operations Research, 2000, 48 ( 2 ): 318-326.
- [ 18 ] Chowdhury N M, Rahman M R, Boutaba R. Virtual network embedding with coordinated node and link mapping [ C ] // INFOCOM 2009, IEEE. IEEE, 2009: 783-791.
- [ 19 ] Yang X Q, Mees A I, Campbell K. Simulated annealing and penalty methods for binary multicommodity flow problems, Progress in Optimization [ M ]. New York: Springer US, 2000: 93-105.
- [ 20 ] Ganesh A, Massoulié L, Towsley D. The effect of network topology on the spread of epidemics [ C ] // INFOCOM 2005. 24th Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings IEEE. IEEE, 2005, 2: 1 455-1 466.
- [ 21 ] Benson T, Akella A, Maltz D A. Network traffic characteristics of data centers in the wild [ C ] // Proceedings of the 10th ACM SIGCOMM conference on Internet measurement. ACM, 2010: 267-280.
- [ 22 ] Hardy G H, Littlewood J E, Pólya G. Inequalities [ M ]. Cambridge: Cambridge university press, 1952.