DOI：10.19473/j.cnki.1008-4940.2019.01.015　网络出版时间：2019-03-27 09:53:55
网络出版地址：http://kns.cnki.net/kcms/detail/35.1333.g4.20190326.1610.030.html

数据中心网络风险规避算法的设计

郑爱媛

（福建商学院　信息工程系，福建　福州，350012）

［摘　要］数据中心网络（DCN）在承载突发载荷时所面临的风险问题已成为当前的研究热点。为确保良好服务质量（QoS），针对 DCN 提出一种基于软件定义网络（SDN）架构的风险规避算法。该算法经由控制器来实施全局网络和突发数据流实时特征的监视，以优化控制层和转发层有限的资源。同时计算出闲置超时最优值来转发突发数据流，进而实现风险规避的目标。通过多个指标的考察与比较，风险规避算法均表现出了传统算法无可比拟的相对优势。

［关键词］数据中心网络；软件定义网络；风险；计算；评估

［中图分类号］TP393　　［文献标识码］A　　［文章编号］2096-3300（2019）01-0095-06

0　引言

互联网+时代，DCN 作为云计算和存储的核心网络基础设施连接着大量服务器集群[1]。由于数据业务不可预见，构建 DCN 时将链路中数据流量的峰值作为网络带宽建设的目标。随着云技术、虚拟技术、物联网技术不断融合发展，海量宽带数据流量将以指数级呈现在 DCN 上，因此业界纷纷着手研究如何控制 DCN 风险以确保数据业务的 QoS[2]。

就网络结构而言，采用多根树型胖树拓扑的传统网络中，由于存在核心层交换机无法应对突发网络重载的风险而无法推广。因此，采用多路由层次型胖树拓扑无疑是首选方案。该方案下的边缘层和汇聚层归属于不同的集群，两个层的交换机彼此相连，汇聚层仅和核心层部分交换节点互连，此举可使 DCN 平台上的服务器集群高效率地转发数据流量载荷。从控制层面来说，传统网络在实施控制功能前需事先在全网中部署路由算法，并将数据流转发模块和路由模块统一在一起；网络设备定义的接口不具开放性；数据流转发只对请求节点负责而缺乏对最优路由的计算选择。因此，实现高效控制无疑要将控制模块独立于转发模块，以便使控制器统一管理设备资源和网络资源，进而规避网络风险。基于上述分析，本文提出一种软件定义网络（SDN）架构下 DCN 风险规避的控制算法。

1　风险规避思想

SDN 有别于传统网络架构，将控制模块和转发模块分离，如图 1 所示。其中 Ryu 控制单元根据网络实时状态为每个条数据流估算出最佳转发路由并下发至转发层中的 OpenFlow 交换机[3]，交换机更新流表后再转发该数据流。整个转发过程的实施为：首先，突发数据包载荷到访 OpenFlow 交换机，由于交换机既有的流表中并未包含与该突发数据包匹配

收稿日期：2018-11-01

作者简介：郑爱媛（1977-），女，福建福州人，讲师，硕士，研究方向：计算机技术。

的流表项，于是交换机向控制层中的 Ryu 控制器提交请求为该突发数据包估算出转发的路由；其次，Ryu 控制器计算出该突发数据包的转发路由并响应给转发层中交换机；然后，交换机在其内存流表中添加该流表项；最后为该突发数据包载荷执行转发动作。显然，SDN 架构可用于应对突发网络中突发数据流的匹配请求，同时采用层次型胖树拓扑结构的 DCN 所具备的多径路由特征也很适合 SDN 优势的发挥。

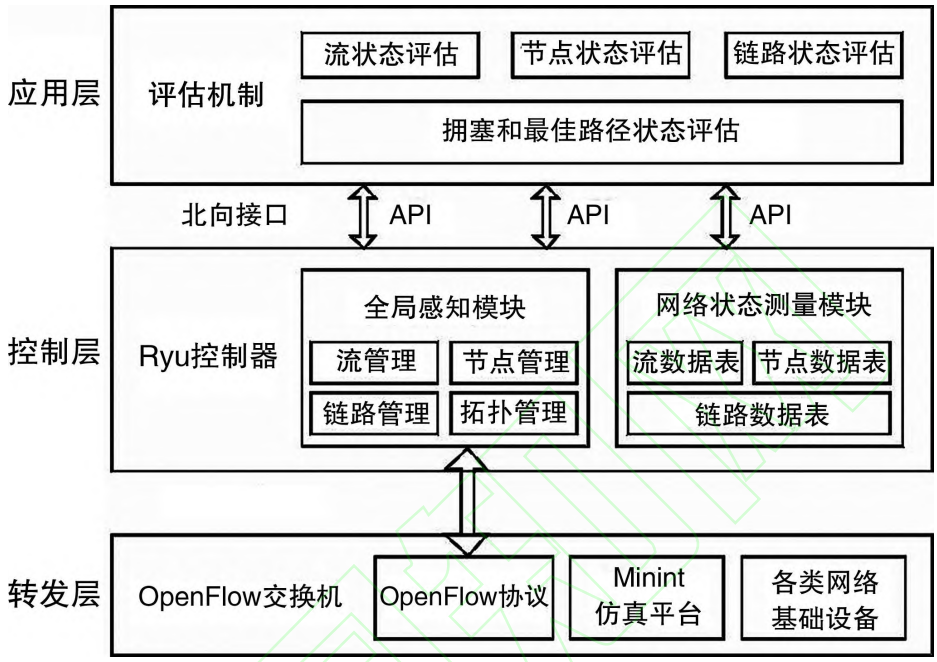

图 1　SDN 体系架构

Fig. 1　SDN architecture

但是，受限于转发层中交换机寻址存储器有限的流表项储存空间[4]，以及控制层中 Ryu 控制器响应数据包转发路由计算请求的规模，往往存在新生突发数据流无法被顺利地分配到流表项的风险，使得新生突发数据流无法被顺利转发，最终导致阻塞率、丢包率、传输时延等与服务质量（QoS）相关的指标劣化。尤其在忙碌时段 DCN 重载情形下甚至可能引发全网瘫痪。基于此，本文从交换机存储空间流表优化的角度提出一种控制机制来规避突发数据包无法被分配到流表项的风险。该机制主要通过统计转发层中交换机存储空间内的流表闲置资源以及控制层中 Ryu 控制单元闲置的计算资源，然后根据突发数据流特性，计算出分配流表项过程中的闲置超时 T_{idle} 的最优值来删除无用的流表项，进而优化整张流表，提高突发数据流和流表项的匹配成功率，实现快速转发。同时通过计算并设置闲置超时 T_{idle} 的最优值，可充分地利用 Ryu 控制器的计算资源改善数据包转发路由计算请求的响应率。这样的控制机制不仅能优化整个 DCN，也可充分地应对突发网络潜在的流表资源匮乏的风险。

2　风险规避原理

根据交换技术原理，突发载荷流量被划分成多个子集抵达交换机。当网络为该突发载荷流量分配足够带宽资源进行传输时，将该突发载荷流量子集记作 $\mathrm{DP}\{\mathrm{DP}_1, \mathrm{DP}_2, \cdots \mathrm{DP}_{n-1}, \mathrm{DP}_n\}$，且服从帕累托分布，并定义帕累托分布参数 τ、M、α[5]；当网络未能为该突发载荷流量授予足够带宽而等候交换机响应时，在这等候期间流表项暂时闲置，将闲置的时间间距子集记作 $\mathrm{It}_n\{\mathrm{It}_1, \mathrm{It}_2, \cdots, \mathrm{It}_{n-1}, \mathrm{It}_n\}$，遵循负指数 F 分布。于是，在突发网络重载环境下上述两个子集遵循函数：

$$\begin{cases} \mathrm{DP}(\mathrm{DP}_n > \tau) = (M/\tau)^{\alpha} \\ \overline{\mathrm{DP}_n} = M\alpha/(\alpha - 1) \end{cases} \tag{2.1}$$

$$\begin{cases} \mathrm{It}_n \sim \mathrm{EXP}(F) \\ \overline{\mathrm{It}_n} = 1/F \end{cases} \tag{2.2}$$

根据上述分析可知，在网络未能为该突发载荷流量授予足够带宽而等候交换机响应的情形下，如果流表项闲置的时间间隔长度 It_n 值超过了超时值 T_{idle}，则交换机中的流表项将被剔除；相反，如果流表项闲置的时间间隔长度 It_n 值低于超时值 T_{idle}，流表项则被保存下来。

当突发数据流量载荷传输的空闲时间间隔值太大，随着流表项的剔除将陆续出现转发路由计算的请求向 Ryu 控制单元提交。Ryu 控制单元的计算资源能够在多大程度上响应该突发数据流量载荷的后续数据包子集的计算请求，取决于该突发数据流中数据包子集的规模 $\sum_{n=1}^{n}\varepsilon(\mathrm{It}_n - T_{idle})$[6]。令 Ryu 控制单元为突发数据流量载荷的每一个数据包子集分配的计算资源值为 σ，那么响应整个突发数据流量载荷所需的归一化 Ryu 控制器计算资源 $\mathrm{ar}(\mathrm{It}, T_{idle})$ 表征为：

$$\mathrm{ar}(\mathrm{It}, T_{idle}) = \lim\{[\sum_{n=1}^{n}\varepsilon(\mathrm{It}_n - T_{idle})\cdot\sigma]/n\} \tag{2.3}$$

显然，在闲置的时间间距子集 $\mathrm{It}_n\{\mathrm{It}_1, \mathrm{It}_2, \ldots, \mathrm{It}_{n-1}, \mathrm{It}_n\}$ 内，当无效的等待时间 $U(\mathrm{It}, T_{idle})$ 越长，闲置流表项规模越大，交换机中流表的匹配率也就越低，网络出现 QoS 异常的风险也就越高。为了描述无效等待时间对交换机流表生存性的影响力，对无效的等待时间做归一化表征[7]：

$$u(\mathrm{It}, T_{idle}) = \lim\{[\sum_{n=1}^{n}\mathrm{It}_n\varepsilon(T_{idle} - \mathrm{It}_n) + T_{idle}\cdot N(\mathrm{It}, T_{idle})]/\sum_{n=1}^{n}\mathrm{It}_n\} \tag{2.4}$$

根据交换技术原理可知，突发网络在实施轻载的情形下数据流无需划分为多个子集即可成功转发。当轻载规模较大时对转发层中交换机的流表资源要求较高；反之，突发网络在实施重载时，数据流将被划分多个数据包子集方可转发。然而由于重载规模不一，交换机存储器中流表内的流表项也将不可避免地出现大量的无效等待时间，使得流表项因超时而被剔除，进而向控制层中的 Ryu 控制单元提交更多转发路由计算的请求。故，突发重载流量对 Ryu 控制单元的计算资源要求较高。假设在满足所有突发流量载荷路由需求的前提下，突发网络可为 N 个平均带宽为 $\bar{B}$ 的数据流载荷提供带宽资源。重载下的突发数据流被划分为 n 个数据包子集，且网络需该条突发数据流分配带宽为 B 的网络资源。据此参数计算出突发数据流对 Ryu 控制单元计算资源依赖度 C 和交换机流表资源依赖度 S：

$$\begin{cases} C = H[(B\cdot n)/(\bar{B}\cdot N)] \\ S = H(n/N) \end{cases} \tag{2.5}$$

3　风险规避策略

风险规避目标旨在通过统计转发层中交换机的流表资源响应度、控制层中 Ryu 控制单元的计算资源响应度和突发数据流量载荷对转发层和控制层资源的依赖度[8]，分析出闲置超时 T_{idle} 的最优值，使突发数据流量载荷在短期内被顺利转发，排除因阻塞率、丢包率、传输时延等指标劣化而导致全网瘫痪的风险。

据此分析，将突发数据流量载荷对全网资源的依赖度 $Y(\mathrm{It}, T_{idle})$ 描述为 $C\cdot\mathrm{ar}(\mathrm{It}, T_{idle}) + S\cdot u(\mathrm{It}, T_{idle})$，可得闲置超时 T_{idle} 的最优值为：

$$T_{idle} = \mathrm{argmin}[C\cdot\mathrm{ar}(\mathrm{It}, T_{idle}) + S\cdot u(\mathrm{It}, T_{idle})] \tag{2.6}$$

根据风险规避思想和原理，在实施风险规避策略时应首先初始化网络参数 F、每一个数据包子集的计算资源 σ 等，并设置闲置超时 T_{idle} 的初始值；其次，统计承载了突发数据流量载荷的网络在当前传输状态下的实时网络参数，如 $\bar{B}$、N 等，并根据抵达交换机的突发数据流特征收集实时数据流载荷参数 B、n，再根据网络实时状态调整突发数据流对 Ryu 控制单元计算资源依赖度 C 的参数和交换机流表资源依赖度 S 的参数；然后评估整个突发数据流量载荷所需的归一化 Ryu 控制器计算资源 $\mathrm{ar}(\mathrm{It}, T_{idle})$ 和归一化无效等待时间对交换机流表生存性的影响力 $U(\mathrm{It}, T_{idle})$[9]。最后求出优化的闲置超时 T_{idle} 值，同时更新交换机中的流表项信息使突发数据流被顺利转发，进而规避了因流表项不匹配或控

制器响应度不高所造成的 QoS 劣化风险。

4 风险规避评估

4.1 评估模型

风险规避策略效能的评估在 Mininet 2.2.1 环境中开展。评估采用 4 元胖树的拓扑结构，含 16 个终端和 20 个 OpenSwitch 交换机。通过调用命令：$ sudo mn --custom ~ /mininet/custom/fattree.py -topo = mytopo -switch = ovsk --mac --controller = remote，ip = Ryu 控制器远程主机 IP 地址，port = 6633，使 Ryu 控制单元和 Mininet 运行在两个主机上。借助于灌包软件 Iperf 协作生成突发数据流量载荷。载荷在区间［10Mb/s，10^6 Mb/s］内呈现线性递增的趋势。每次突发数据流量生成时间间隔为 5sec.，每个突发数据流量载荷长度为 60 sec.。每个 OpenSwitch 交换机上的突发数据流量载荷规模不超过 10^6 条，并遵循 0.05 的负指数参数。

同时，OpenSwitch 交换机提供 2Gbit/s 的带宽资源，其流表存储空间为 2 000 条，其中转发数据流的阈值缺省为 70%。

目前，在基于 SDN 架构的 DCN [10] 上实施数据流转发请求的策略有动态链路均衡算法和多路由均衡算法。因此评估方案从突发数据流请求响应度、交换机流表使用率、全网吞吐量三个方面将风险规避策略与动态链路均衡算法、多路由均衡算法做出对比。

4.2 评估方案

（1）突发数据流请求响应度

由前文描述可知，当突发网络处于轻载时，控制层中 Ryu 控制单元计算资源较为充裕，完全可以响应 OpenSwitch 交换机上的突发数据流子集向上层提交的转发路由计算请求；反之，当突发网络处于重载时，随着 Ryu 控制单元计算资源耗尽，响应度势必减弱。这样的特征在如图 2 所示的三种算法策略仿真曲线图中均得到了体现。同时，由于三种算法策略对突发数据流采用不同的处理方式，使得各自产生突发数据流子集规模也不一致，向 Ryu 控制单元提交转发路由计算的请求机会也有所差异。相对于多路由均衡算法和动态链路均衡算法，本文提出的风险规避机制能够结合突发数据流特征优化 Ryu 控制单元计算资源，为相同规模的突发数据流量载荷子集提供更多的响应度。

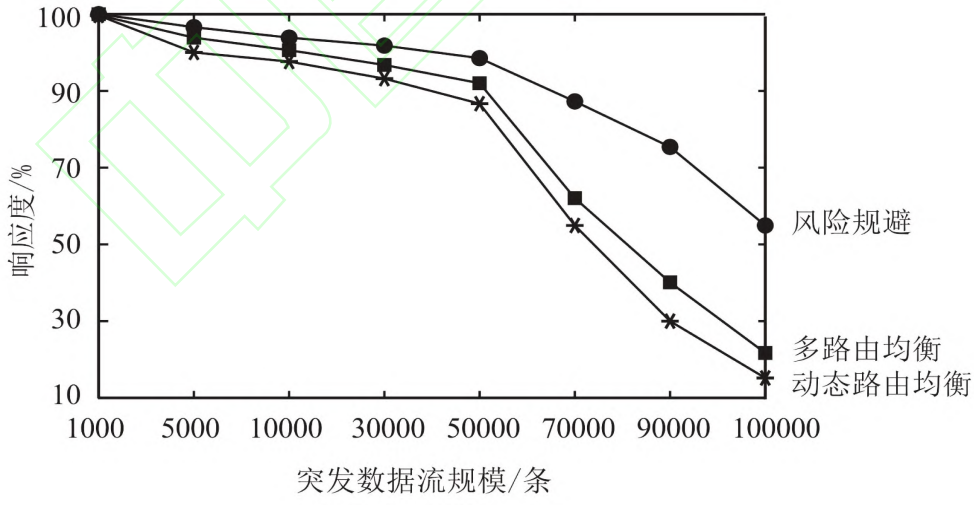

图 2 三种算法策略下突发数据流请求响应度

Fig. 2 Response of burst data stream requests under three algorithms

（2）交换机流表使用率

如图 3 所示，当突发网络中生成的突发数据流规模较小时，转发层中 OpenSwitch 交换机内的流表空间存放的流表项能够良好地满足数据流子集的适配请求。随着突发数据流规模线性递增，后续的突发数据流无法从有限的流表资源中得到匹配的机会，

只能向上层控制单元提交请求等候控制器为其分配流表项，这样的情形显著降低了流表资源的使用率。较之传统算法，本文提出的风险规避机制通过一定的计算机制设定最优化的超时值，从而显著改善了 OpenSwitch 交换机内的流表资源使用率。

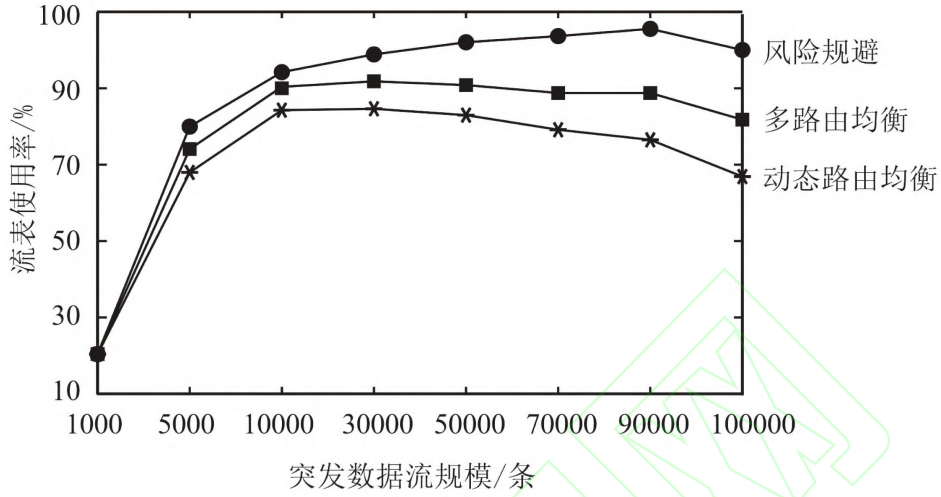

图 3　三种策略下交换机流表使用率

Fig. 3　Utilization of switch flow table under three policies

（3）全网吞吐量

理论上，网络吞吐量与突发数据流载荷流量呈正比。然而受限于设备和网络性能，实际吞吐量并不总是呈现递增趋势，而是在达到一定峰值后有所减缓。图 4 所示的本组实验结果总体上也符合这样的规律。图中不难看出，随着负载增至 10 000，动态链路均衡算法首先达到了峰值。数据流量进一步增加，阻塞也变得愈加严重，使得整体吞吐量下降。相对于多路由均衡算法，本文提出的风险规避机制能够结合突发数据流量载荷实时特征，统计其对全网资源的依赖度来高效管理 OpenSwitch 交换机内的流表项和 Ryu 控制单元计算资源，从而使得在相同网络环境参数下，表现出了吞吐量优越性。

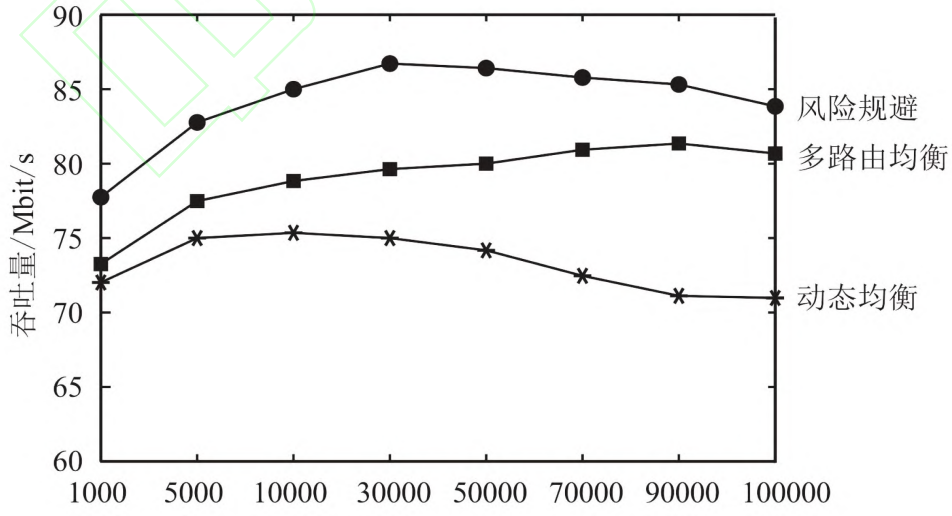

图 4　三种策略下全网吞吐量

Fig. 4　Network throughput under three policies

5 结论

DCN 在传统的算法机制下存在丢包、拥塞等 QoS 异常现象。本文借助 SDN 架构优势，将其部署在 DCN 中并提出一种风险规避算法，该算法结合突发数据流特征，通过计算 OpenFlow 交换机资源和 Ryu 控制器资源来求解超时优化值，以便顺利转发突发数据流从而达到化解网络阻塞和丢包的风险。该风险规避算法借助 Floodlight 和 Mininet 环境展开模拟评估，实验表明，相对于动态均衡算法和多路由均衡算法，该机制在突发数据流请求响应度、交换机流表使用率、网络吞吐量等方面具有良好的优势。

参考文献：

[1]朱丹红．基于 SDN 数据中心网络的时限感知的拥塞控制算法[J]．计算机工程与应用，2018(3)：68-73.

[2]周飞．基于 SDN 的 QoS 研究[J]．计算机技术与发展，2018(3)：6-10.

[3]邱翔．基于 OpenFlow 的 SDN 架构研究与仿真分析[J]．电子科技，2016(12)：85-88.

[4]席孝强．一种保持 OpenFlow 功能完整性的 TCAM 流表压缩模型[J]．计算机应用研究，2018(5)：1464-1469.

[5]闻长江．SDN 原理解析—专控分离的 SDN 架构[M]．北京：人民邮电出版社，2016.

[6]孟飞．基于博弈的中心骨干网带宽分配策略[J]．计算机研究与发展，2016(6)：306-313.

[7]余庚．基于约束的 ASON 生存性探讨[J]．光通信技术，2018(1)：13-15.

[8]马海燕．基于 SDN 的 QoS 时延控制研究综述[J]．中国传媒大学学报(自然科学版)，2018(2)：24-29.

[9]于笑．软件定义网络技术的发展与应用研究[J]．光通信技术，2017(3)：5-8.

[10]许志聪．DCN 中多播数据传输方案[J]．计算机工程与设计，2016(6)：157-163.

Design of Risk Avoidance Algorithm of Data Center Network

ZHENG Aiyuan

(Department of Information Engineering, Fujian Business University, Fuzhou 350012, China)

Abstract: The risk of data center network (DCN) carrying burst load has become a hot research topic. In order to ensure good quality of service (QOS), a risk aversion algorithm based on software defined network (SDN) architecture is proposed for DCN. The algorithm implements real-time monitoring of global network and burst data flow via a controller to optimize the limited resources of the control layer and the forwarding layer. At the same time, the optimal value of idle timeout is calculated to forward the burst data stream, and then the goal of risk aversion is realized. Through the investigation and comparison of several indexes, the risk aversion algorithm shows the unparalleled comparative advantage of the traditional algorithm.

Key words: data center network (DCN); software defined network (SDN); risk; evaluation; estimation

(责任编辑：杨成平)