



专题：数据网与算力网络

面向大规模确定性网络的全局循环排队与转发机制

莫益军, 杨子涵, 刘辉宇, 何天流

(华中科技大学计算机科学与技术学院, 湖北 武汉 430074)

摘要: 时延抖动敏感的工业控制和远程驾驶应用驱动网络由尽力而为服务向确定性转发服务转变。为满足确定性应用时延抖动的有界需求, 工业界和学术界在逐流过滤整形和队列转发控制方面进行了大量研究, 实现了小规模以太网环境下轻载稳定流的微秒级时延和亚微秒级抖动保障。多数研究因较少考虑长距离网络拓扑变化、传输时延波动和短时突发流量过载等影响, 难以满足大规模骨干网场景下的确定性要求。在循环排队转发(cyclic queuing and forwarding, CQF)基础上, 考虑了源目的路径特性和确定性流量强度因素, 提出了全局循环排队转发三队列(global cyclic queuing and forwarding 3-queue, GCQF-3), 并在基于 OMNeT++ 搭建的确定性网络仿真系统中, 比较了 CQF、CQF-3 和 GCQF-3 的转发控制机制, 实验表明 GCQF-3 能根据确定性流量状态及时调整门控时机和排队优先级, 在时延、抖动和网络利用率等方面都达到最佳。

关键词: 确定性网络; 大规模骨干网; 循环队列转发; 流量整形

中图分类号: TP393

文献标识码: A

doi: 10.11959/j.issn.1000-0801.2021240

Global cyclic queuing and forwarding mechanism for large-scale deterministic networks

MO Yijun, YANG Zihan, LIU Huiyu, HE Tianliu

School of Computer Science and Technology Huazhong University of Science and Technology, Wuhan 430074, China

Abstract: Delay and jitter sensitive industrial control and remote driving application-driven network by best-effort to deterministic service. To meet the deterministic application of delay and jitter demands, industry and academia have conducted detailed research on stream filtering shaping and queue forwarding control, implementing microsecond delay and microsecond jitter of light load steady in small-scale ethernet environment. However, the study seldom considers the effects of long-distance topology change, transmission delay fluctuations and short-time traffic overloads, challenging to meet deterministic requirements under large-scale backbone networks. For this challenge, based on cyclic queuing and forwarding, considers source-destination path properties and deterministic traffic intensity, the global cyclic queuing and forwarding three-queue was proposed. The deterministic network simulation system was based on OMNeT++. Experiments compare queue forwarding control mechanisms such as CQF, CQF-three and GCQF-three. Results show that GCQF-three can timely adjust the gatecontrol and queueing priority according to the deterministic traffic state, achieves the best in delay, jitter and network utilization.

Key words: deterministic network, large-scale backbone network, CQF, traffic shaper

收稿日期: 2021-09-20; 修回日期: 2021-10-14

基金项目: 国家重点研发计划项目 (No.2020YFB1805203)

Foundation Item: The National Key Research and Development Program of China (No.2020YFB1805203)



1 引言

随着工业 4.0、远程驾驶和远程手术等应用数量的不断增长，对网络超低时延和微小抖动要求越来越高^[1]。其中工业物联网要求端到端时延为微秒级到毫秒级，抖动为微秒级；触觉互联网（远程手术）要求端到端时延为 3~10 ms，抖动不超过 2 ms；辅助驾驶要求端到端时延为 100~250 μ s，抖动为几微秒。远程驾驶不仅需要低时延抖动，还需要更高传输速率^[2]。为满足上述应用对网络的要求，时间敏感网络（time sensitive network, TSN）和确定性网络（deterministic network, DetNet）分别对以太网的链路层和网络层进行优化，提升其对时间敏感流传输的支撑能力。

为保障确定性业务需求，TSN 主要从时间同步、资源预留、流量控制和排队转发等方面对以太网 L2 层进行优化，DetNet 则从动态网络配置、资源编排、路径规划、路由转发和多径转发等方面对以太网 L3 层进行优化。考虑到业务流排队转发是 L2 层 TSN 和 L3 层 DetNet 共同关注的焦点，且 DetNet 中，即便做了路径规划和资源预留，也会因突发流量和过载流量竞争而产生不可预知的时延和抖动，本文后续的工作将围绕流量整形和队列排队转发展开。IEEE 802.1tsn 标准中提出的流量整形调度包括 IEEE 802.1Q 中提出的基于信用的流量整形^[3]（credit based shaper, CBS）、IEEE 802.1Qbv 中提出的时间感知整形^[4]（time aware shaper, TAS）、IEEE 802.1Qcr 中提出的异步流整形^[5]（asynchronous traffic shaper, ATS）和 IEEE 802.1Qbu 中提取的帧内抢占整形（frame preemption shaper, FPS）等，IEEE 802.1Qch 中设计的队列和转发机制包括 CQF^[6]、CQF-3、增强型循环排队转发（paternoster）、循环特定排队转发（cyclic special queuing and forwarding, CSQF）和可扩展确定性排队转发（scalable deterministic forwarding, SDF）。

分析流量整形和排队转发方案可知，CBS 根据两类队列中的排队帧数和传输帧数调整端口速率，导致平均队列时延较高；TAS 采用门控机制对输出端口八队列进行控制，每队列可分别设置不同的传输策略，其时延抖动性能得到提升，其管理复杂度高，且抗过载流量能力较弱；CQF 在 TAS 门控队列基础上选取最后两个队列用于时延敏感业务流量的循环排队，其时延抖动性能得到进一步提升，但同样存在抗过载流量能力弱，且对时间同步要求较高。上述排队方案通过周期性控制节点自身门控队列开关保障时延敏感业务服务质量，仅考虑了交换节点自身调度频率和进出流量类型，只适用于小规模轻载局部网络。而大规模 DetNet 传输距离远、拓扑易变化和路径的不确定，难以保证下游交换节点在两倍时隙周期内接受到上游节点发出的数据报。另一方面，确定性时延敏感业务并非恒定速率，且存在流量微突发和流量过载情况。一旦流量过载，时间抖动性能急剧恶化^[7]。鉴于此，TSN 和 DetNet 研究者在 CQF 基础上将出口队列扩展至 3~4 个，分别承担接收、暂存和门控等功能，并结合端到端的路径分别确定各交换节点的排队调度策略。

针对大规模 DetNet 拓扑和传输特点^[8]，确定性业务流存在的微突发现象和现有队列机制保障过载确定性流量能力不足问题，本文在标准 CQF 方案基础上提出了 GCQF-3 方案，对其进行了 3 项优化，首先，将标准 CQF 中服务于确定性流的双队列扩展为三队列，增加的缓存队列用于对微突发业务流进行整形，并暂存部分高负载数据报和未及时转发数据报；其次，在对队列进行门控时，综合考虑端到端路径途径交换节点的出入度和流量负载分布，并在帧抢占式队列基础上，将周期性逐报文门控优化为逐帧门控；最后，通过增加确定性流队列数量和全局状态感知的灵活门控策略，降低了高负载情况的丢包率，提升时延抖动性能。

2 相关工作

实现确定性网络的过程中,单一 VID 难以满足需求,边缘节点生成的唯一 DetNet 流标识被用于确定性流的识别;同时周期性标识的嵌入使得全网精确的时间同步得以实现,这是 DetNet 和 TSN 的基础。但是网络中无序到达的流量对于业务的处理造成了困扰,引入的流量整形器保证确定性业务的优先级资源分配,使得流量在有界时间内进入下一跳。循环队列排队转发作为一种流量整形机制,实现了特定流量需求的细粒度 QoS 服务。

2.1 流标识及同步

大规模 DetNet 对确定性业务流的时延抖动进行保障,逐流确定其业务类别及确定性时延抖动属性,并以此为基础在交换节点上进行排队转发^[9]。IEEE 802.1tsn 中建议将以太网报头的优先级码点(PCP)字段和 VLAN ID (VID) 共同作为流标识。但远距离大规模 DetNet 需跨多段网络,VID 难以穿越端到端所有网络,需逐段进行 VID 的映射,且全网最大 VLAN 数不能超过 4 096,意味着使用 VID 作为 DetNet 流标识将限制了确定性业务流并发数量。另一方面,诸如远程驾驶和远程手术之类的确定性业务涉及多流和多转发路径,单一 VID 难以满足多流、多路径需求。鉴于此,可在确定性业务进入网络时,由边缘节点生成唯一 DetNet 流标识,将其封装在 DetNet SRv6 头部,以用于确定性流识别和排队。

DetNet 和 TSN 需要全网具备精确的时间同步,以满足时延和抖动需求^[10]。毫秒级的网络时间协议(network time protocol, NTP)因其时钟偏移和时间配置无法满足 DetNet 的时间同步要求。现有时间同步仅限于亚微秒级精度,跨网络定时信息的全局共享需要网络节点时间彼此精确地同步,大型对撞机通信网络设置要求亚纳秒级的时间精度。鉴于此,IEEE 802.1AS 采用通用精

确时间协议^[11](gPTP)实现网络设备时间同步。gPTP 通过时间主机和时间从机之间的消息传递形成时间感知网络(gPTP 域),基于对等路径延迟机制计算驻留时间和链路时延,以此校准时间实现本地时间与 gPTP 域的主时间同步,最终保证所有设备同步到全局时间,以用于门控队列的开闭控制。该时间同步精度取决于主时间频率、驻留时间和链路时延。各网络实体保持高时钟精度成本较高,且网络设备间频繁的定时信息周期性交换给控制面带来较大负载压力。为解决同步成本和网络开销问题,ATS 和 SDF 都提出在数据报中嵌入周期标识,只需网络节点频率同步即可消除节点队列控制对时间同步和时隙同步的苛刻要求。

2.2 流量整形控制

DetNet 和 TSN 中节点引入流量整形器对无序到达的各类流量进行平滑整形,保证在有界时间内进入下一跳^[12]。其流量整形器需要保证确定性业务的优先级资源分配,确定性流与尽力而为流的隔离和排队帧等待时间限制等。标准的流量整形方法包括 CBS、脉冲限制整形器、TAS 和 ATS 等。CBS 引入信用机制进行特定流量的排队整形,其信用随空闲斜率增加,随发送斜率降低,当信用大于或等于零时传输转发流量,分散了 CBS 通信量,避免了下游节点的流量突发;TAS 引入时分复用机制周期循环控制队列开闭,根据门控列表(gate control list, GCL)循环转发定时门打开状态的队列流量,节点出口流传输时选择特定的传输策略;ATS 则融合了逐跳逐流分类整形机制,引入令牌桶整形器控制流输出速率,防止下游节点流量突发。

2.3 队列排队转发

DetNet 和 TSN 的标准流量调度和队列管理方案分为分类队列和逐流队列^[13]。分类队列基于服务类别分配优先级,与以太网层优先级码点(priority code point, PCP) 3 位相对应,分类队列每个节点端口使用 8 个队列来处理服务类。分类



队列 QoS 保障粒度较粗, 共享相同 PCP 码的业务流排队时容易导致业务流突发, 且传递到下游相应队列; 逐流队列则在节点端口为每条流提供一个队列, 可逐流管理突发性及速率, 实现特定流量需求的细粒度 QoS 服务。但随着网络规模扩大, 逐流排队成本高、扩展性差, 难以实用。分类逐流混合队列方式是平衡分类队列和逐流队列性能的首选。

针对 DetNet 和 TSN 的队列控制, 对 CQF、CQF-3、Paternoster、SDF 和 CSQF 的分析如下。

(1) CQF

循环排队和转发由蠕动整形器^[14]发展而来, 以循环方式调度节点业务流入出队列。CQF 将队列按时隙划分复用, 使用 8 个队列中的两个对确定性业务流进行循环调度。两个队列中一个队列在某时隙对到达帧排队时, 另一队列将上个时隙排队帧取出并转发, 即任何时隙内到达帧在下一个时隙传输, 且在两个时隙内到达下一跳节点。这意味着上下游节点的传播时延需小于时隙间隔时长, CQF 门控时隙受制于网络规模和链路距离。网络规模越大, 其时隙间隔越大, 时延抖动越难保障, 且当入口帧到达间隔小于时隙间隔时, 容易引起累积流量突发。

(2) CQF-3

针对 CQF 存在的问题, 尤其是上游节点到达帧传输到下游节点的时延因链路传播和节点处理时延超过两个时隙周期时, 引入第 3 个分组队列充当缓冲, 以避免帧到达时隙错误或丢包。

(3) Paternoster 机制

同样针对标准 CQF 存在的问题, Paternoster 机制在每个节点出口设置一个流保留计数器和先前、当、后续和最后 4 个输出队列。每个时间窗口中, 先处理完先前队列后, 再选择当前队列进行传输, 直至窗口期满。在处理过程中, 新接收帧在当前队列排队。追加帧在后续队列和最后队列排队, 直至超过队列门限而被丢弃。每个时间

周期内, 4 个队列交替转换状态。该方案通过 4 个状态队列增强了 CQF 的容错能力, 降低了丢包率, 但随着 DetNet 跳数节点增加, 维护成本快速增加, 时延抖动性能有所下降。

(4) SDF

SDF 在保证节点具有同步频率基础上, 在数据报中增加周期标识符, 以时隙方式异步转发数据流, 以消除节点间的苛刻时间同步需求。在入口节点和出口节点增加流量整形门控, 仅在网关处进行逐流排队, 并为每类流量维护 3 个队列, 一个用于输出排队, 两个用于不同周期标识符的输入排队, 一个队列出列, 两个队列入列 (来自不同的周期标识符)。本节点根据接收到来自上游节点报文中的显式周期标识符进行节点间的周期关系映射和转发, SDF 时延上限同 CQF 一样节点间不超过两倍时隙周期, 全网不超过两倍总跳数时隙周期。

(5) CSQF

CSQF 针对标准 CQF 为满足帧转发时隙规则预留保护带宽而损失带宽利用率问题, 借助段路由标识符 (SID) 协调循环传输时间, 并明确在端到端路径上指定每个 DetNet 节点的传输周期。此外, CSQF 沿用了 CQF 的分类队列思想, 按流量分类在节点出口维护发送、接受和容忍 3 个队列, 且循环交替更换队列角色和状态。

综上所述, 为满足大规模确定性时延抖动需求, 在 CQF 基础上不断进行队列演进尝试, 其研究趋势主要体现为 4 点, 并用于指导 GCQF-3 设计。

- 引入时隙周期性标签降低时间同步要求。
- 采用混合分类排队和逐流排队机制以满足不同规模拓扑的扩展性和管理需求。
- 采用状态循环交替变换的多队列机制解决队列转发时隙周期错误和突发流量问题。
- 采用端到端全路径多跳排队代替单跳排队控制以及时响应网络节点的状态变化。

3 全局循环排队转发机制

3.1 全局循环排队转发框架

参考大规模 DetNet 中节点队列设计的四大趋势, 为保证 DetNet 流时延抖动需求, 本文提出 GCQF-3 排队转发框架如图 1 所示, 包括 4 个关键点。

- 当新发起确定性流进入入口网关时, 节点向控制平面注册 DetNet 流速率和时延抖动等网络资源需求和 QoS 需求, 以进行路径选择和资源预留。当 DetNet 流进行排队转发时, 携带 DetNet 流标识。
- 端到端路径上的交换节点向控制平面报告自身节点在途流标识、确定性流总速率及确定性流的入度和出度, 以指导全局循环排队转发。
- 每个交换节点为确定性流分配三队列, 分

别用于接收、缓存和传输。

- 出口节点根据时隙周期标识计算测量确定性流的端到端时延和抖动, 将异常差错报告给控制平面, 以调整全局循环排队转发策略。

3.2 三队列转发处理机制

考虑到 PCP 位数限制, 全局循环排队沿用标准 CQF 的八队列设计, 其中两队列用于确定性业务流量排队转发, 伴随确定性双队列的, 额外引入一个共享缓存队列。该队列不直接从输入端口接收数据, 仅接收两个确定性队列溢出时拟丢弃数据报进行排队, 用以实现对突发业务流的缓存和平滑调度。并与对应队列门控进行联动, 调节出口报文转发速率。随着确定性流量的增加, 可按两确定性队列加伴随缓存队列的配置成组增加用于确定性业务流量排队的队列。全局三队列循环排队转发机制如图 2 所示。

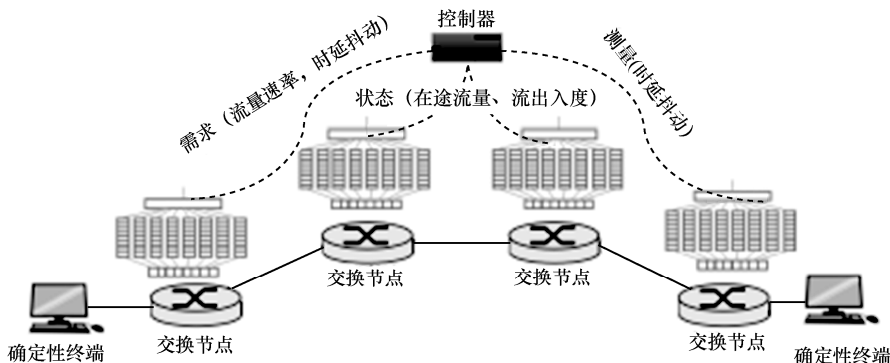


图 1 面向 DetNet 的全局循环排队转发框架

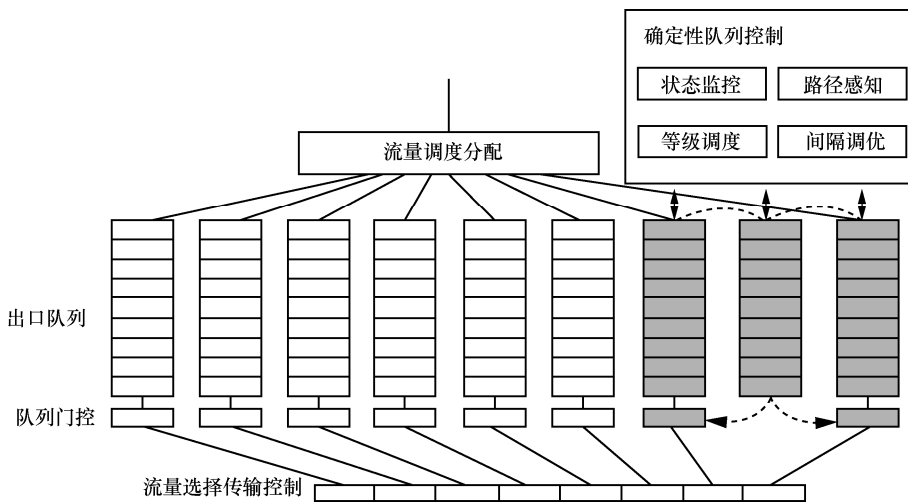


图 2 GCQF-3 循环排队转发机制



交换节点完成选路和交换后，基于以太网帧中的 VLAN 标签、PCP 分类码和流标识区分流量优先级，对流量数据进行分类，并将确定性流输出至相应的优先级队列进行排队。其中，确定性业务进入分发至两个确定性队列进行排队。各交换节点的确定性队列控制单元定期对两个确定性队列和缓存伴随队列的忙闲状态和流标识进行监控，将监控状态汇报至确定性控制器。确定性控制器对其计算确定性流的端到端路径，并对各交换节点承载的确定性业务流属性、数量和流量进行估计，并将与其反馈至与其相关联的交换节点。各交换节点根据各自队列状态、确定性流需求和路径各节点状态，控制缓存伴随队列的存取顺序和速度，调节确定性队列门控打开闭合时长。当排队的数据报未超出确定性队列容量时，GCQF-3 与 CQF 相同，两队列在奇偶周期轮流打开。当排队数据报超出确定性队列容量，GCQF-3 并不会将数据报丢弃，而是进入缓存伴随队列进行缓存，确定性队列中当前数据报转发结束后，按序逐流分等级转发缓存伴随队列中确定性等级高的数据报，其门控周期保持与 CQF 一致，但其数据报调度转发频次取决于端到端传播路径和路径节点状态。

3.3 队列门控策略及参数

(1) GCQF-3 时序

大规模确定性网络完成路径规划、资源切片和路由选择后，端到端路径和拓扑相对固定，单跳往返时延和端到端往返时延相对固定，但因两个确定性队列循环开闭周期和上游报文发送时间

与下游报文接收时间与与上下游往返时间相关，随着端到端路径传播距离和节点数量增加，节点排队时延和端到端时延都将增大。为降低确定性流时延，GCQF-3 针对数据报到达排队和门控转发需占用两个时隙周期的问题，直接将端到端传播距离较长的确定性流的数据报同时通过两个确定性队列进行排队，保证长距确定性流数据报无须耗费等待周期，一直处于可发送状态，但发送时机仍受排队长度和报文转发速度影响。

如图 3 所示，标准 CQF 以 T 为单位周期性轮流开闭队列门控。队列 6 奇数时隙打开，偶数时隙关闭；队列 7 奇数时隙关闭，偶数时隙打开。若平均每跳传输时延为 P ，则数据报每跳时延为 $[T+P, 2T+P]$ ，端到端 N 跳的时延为 $[NT+NP, 2NT+NP]$ ，端到端时延随跳数二倍增加。当队列负载较重和端到端路径跳数较长时，GCQF-3 除了在奇数时隙打开队列 6 门控，偶数时隙打开队列 7 时隙，将未传输的数据通过缓存伴随队列 HQ 转发，HQ 保持与打开状态的队列门控连接，则除首跳外，数据报每跳时延为 $[P, P+\delta]$ ，则端到端 N 跳的时延为 $[T+P+\delta, T+NP+\delta]$ ，即时延仅与传播时延相关。

(2) 门控策略

为简单起见，上述 GCQF-3 时序仅介绍了网络拓扑和流量恒定状态下的端到端队列转发时序。真实网络受节点加入退出、负载分布和服务能力影响，其端到端路径跳数和距离也随之变化。另一方面，确定性业务码流并非恒定速率，且受路径变化和确定性流非均匀到达的影响，高负载确定性网络中端到端路径上中间节点不可避免地

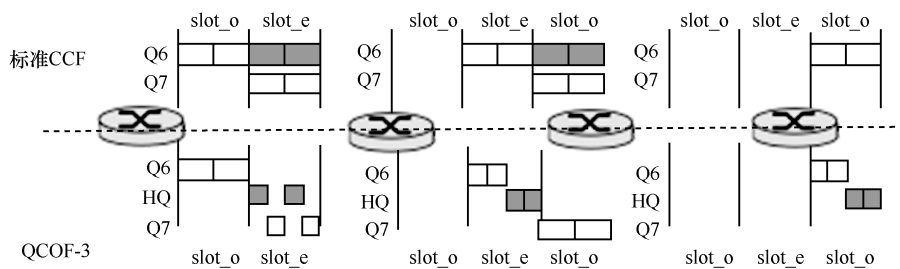


图 3 GCQF-3 循环排队转发机制

存在流量突发。为降低长距转发的时延,且兼顾平滑抑制流量突发问题,GCQF-3 综合平衡周期循环、快速转发和状态感知 3 类门控策略。

(1) 周期循环策略

周期循环门控策略适用于网络规模较小和端到端跳数较少的情况,GCQF-3 退化为 CQF,在奇偶周期交替分别打开确定性队列 6 和 7。

(2) 快速转发策略

快速转发策略是为适应 GCQF-3 时序描述中提到的大规模确定性网络转发而制定的策略,相对长距链路和多跳路径,GCQF-3 通过缓存伴随队列保证到达报文持续有机会被转发,无须等待,但其转发调度频率受制于队列长度及队列读取最小调度周期,该策略相当于以更细颗粒度执行周期循环策略。快速转发策略仅与三队列的排队顺序和时间戳有关,不区分确定性流的类型和网络状态的变化。

(3) 状态感知策略

状态感知策略是为适应拓扑变化、重载流量和流量突发制定的队列转发策略。重载流量和突发流量情况下,不考虑网络状态进行快速转发,将加剧排队拥塞和丢包,时延抖动难以保障。针对此问题,各交换节点的 GCQF-3 根据确定性控制器返回的端到端路径上逐跳往返传播时延,交换节点相应端口的入度、出度、流量和队列负载比,以及本节点相应端口的待转发流量,确定其转发调度频率,若感知的上述状态未超过门限值,其策略将简化为快速转发策略,甚至周期循环策略。

针对上述 3 种情况,GCQF-3 的排队转发策略算法的相关流程如算法 1 所示,其中涉及的参数符号见表 1。

算法 1 GCQF-3 队列门控策略

输入 Sid, Pid_i

输出 T, Repeat_k

获取确定性流的 Sid, Pid_i

通过 Pid_i 计算 Hop_i

计算 Length_i

$$\text{Length}_i = \sum_{j=0}^{\text{Hop}_i} \text{TTL}_{ij}$$

$$\text{if } \text{Length}_i > \alpha \text{Exp}D_i \text{ then } T_i = \frac{\text{Exp}D_i}{2 \times N_i}$$

$$\text{if } \frac{\text{count}(\text{Ratio}_{ij} \geq \beta)}{N_i} \geq \delta \text{ then}$$

$$\text{Repeat}_k = 1 - \frac{\sum_{j=0}^{\text{Hop}_i} \text{Ratio}_{ij}}{N_i}$$

else

$$\text{if } T_{r_{ij}} \leq \frac{\text{Cap}_{ij}}{T_i} \text{ then}$$

$$\text{Repeat}_k = \frac{T_{r_{ij}}}{\beta \text{Cap}_{ij}}$$

else

$$\text{Repeat}_k = 1$$

else

$$T_i = \frac{\sum_{j=0}^{\text{Hop}_i} \text{TTL}_{ij}}{N_i}, \text{Repeat}_k = 1$$

return result

表 1 GCQF-3 排队策略相关参数

类型	符号	描述
输入	Sid	流标识
	Pid _i	路径标识
输出	T	门控周期
	Repeat _k	单位周期调蓄次数
过程	ExpD _i	确定性流预期时延
	N _i	确定性总跳数
	Tr _{ij}	确定性流预期总流量
	ExpJ _i	确定性流预期抖动
	Hop _i	路径下游总跳数
	Length _i	路径下游总长
	TTL _{ij}	下游逐跳往返时间
	Cap _{ij}	下游逐跳队列容量
	Ratio _{ij}	下游跳队列负载率



4 实验分析

4.1 仿真框架

不同网络规模拓扑和确定性流量负载对循环转发机制的端到端时延与抖动性能影响较大,为快速评估循环转发队列的时延抖动性能,在事件驱动的 OMNeT++ (objective modular network tested in C++) 仿真平台^[15]基础之上,构建承载确定性流量排队转发的验证环境,其仿真验证框架如图 4 所示,包括流量拓扑构建、时隙分配映射、排队转发控制和仿真参数控制 4 个部分。

(1) 流量拓扑构建

工业互联网业务在确定性网络应用中占比较高,而本地工业网络存在线形、环形和星形 3 种拓扑。鉴于此,本仿真框架将生成线形拓扑、环形拓扑和星形拓扑 3 种以验证不同拓扑下的确定性流量排队转发性能。

(2) 时隙分配映射

时隙分配是确定性网络排队门控的基础,本仿真框架中不区分同步时钟和异步时隙映射,但将时隙划分为确定性周期、非确定性周期及保护周期,其周期占比分别为 70%、25%、5%。

(3) 排队转发控制

本仿真控制框架在 OMNeT++ 的 NeSting 中实

现了 CQF、CQF-3 和 GCQF-3 等多种框架,为保证 GCQF-3 的正常运行,还对路径和状态进行测量、监控和感知,对排队等级和周期调度参数等进行修改和控制。

(4) 仿真参数控制

为全面评估各种参数对确定性网络中排队转发策略的影响,该模块对表 1 中所示参数提供调整接口,并组合确定性流量、非确定性流量、背景流量和突发流量等各种流量进行打流仿真测试。

在上述确定性网络转发仿真框架下,对本文提出的 GCQF-3 排队机制与 CQF 和 CQF-3 等现有排队转发机制的时延抖动性能进行如下对比。此外,不同机制在网络通信开销、确定性流扩展性和调度复杂度等方面也相去甚远,本文还将从网络通信开销、灵活扩展性和调度复杂度等方面展开定性分析和总结。

4.2 性能比较

(1) 不同拓扑下排队策略的平均时延

确定性网络拓扑结构对排队转发性能影响较大,如图 5 所示,在相同策略下,环形拓扑下排队转发性能最差,星形拓扑下排队转发性能最优。在相同拓扑下,本文提出的 GCQF-3 排队转发性能最优,以线形拓扑为例,GCQF-3 相对 CQF 平均时延缩短 27%,而相对 CQF-3 平均时延缩短

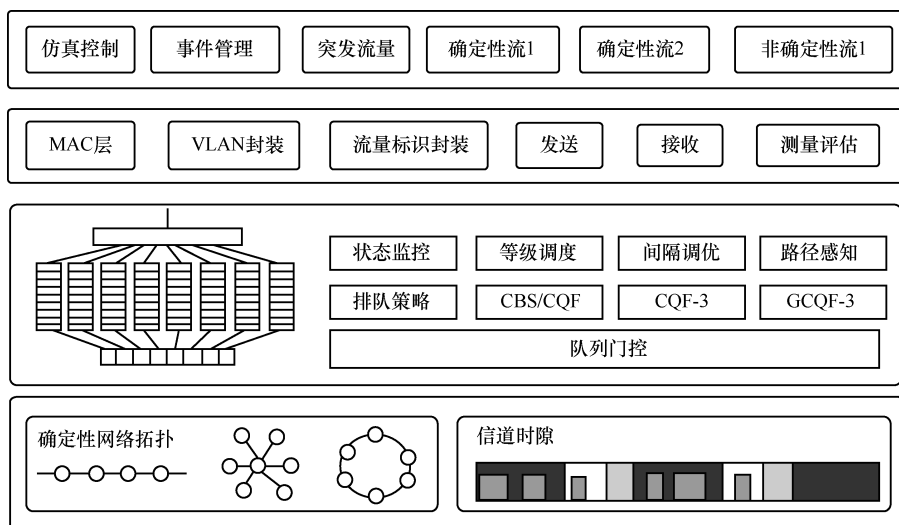


图 4 确定性网络排队转发仿真框架

33.5%。其主要原因是 CQF-3 通过增加的队列增加了抗突发能力,保证了包丢失,但未改变转发频率和周期,反而增加了排队长度,导致时延增加。GCQF-3 虽增加了缓存伴随队列,同时根据网络状态加快了转发频率,在不突破节点流量负载和队列容量的条件下,大幅降低排队转发时延。

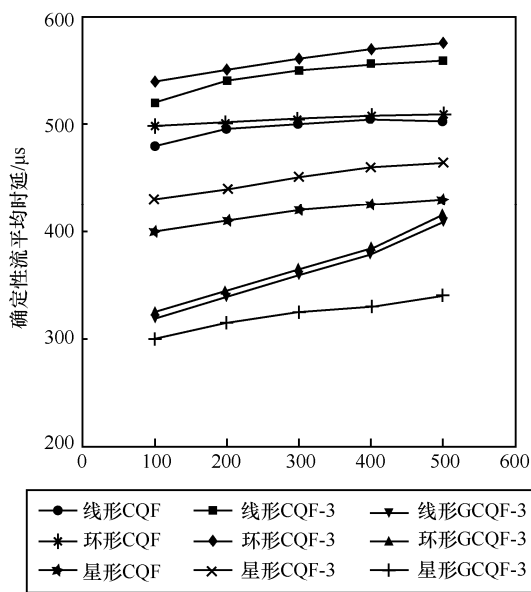


图5 不同拓扑下排队策略的平均时延

(2) 线形拓扑下各种排队策略的平均抖动

由确定性网络排队策略时延分析可知,线形拓扑因其路径长度和排队时序的特点,最能代表排队策略对确定性流性能的影响。主要以线形拓扑为主进行分析,CQF、CQF-3 和 GCQF-3 的平均抖动性能如图6所示。其中 GCQF-3 平均抖动介于 CQF 和 CQF-3 之间,鉴于 CQF 对流量突发和重载流量性能支撑不佳,而 CQF-3 通过三队列联动加大了队列缓存,进一步实现了对流量突发的平滑。而 GCQF-3 虽加大了队列缓存,但在门控打开周期增加了转发频率,导致下游节点流量汇聚突发的可能性加大。

(3) 线形拓扑下各种排队策略的平均丢包率

与确定性网络排队策略抖动分析类似,CQF、CQF-3 和 GCQF-3 的平均丢包率如图7所示。

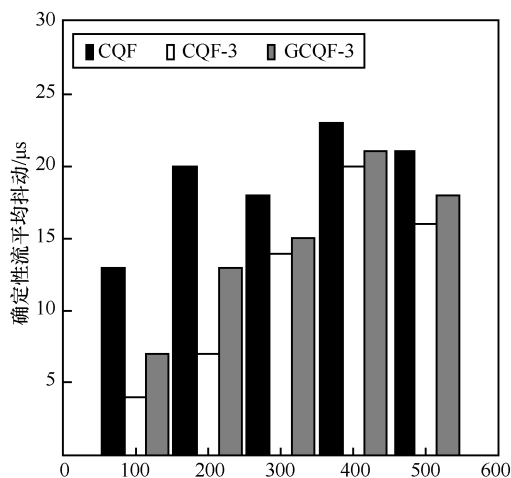


图6 线形拓扑下不同排队策略的平均抖动

GCQF-3 由于加快了队列转发频率,三队列出现拥塞情况较少,大大降低了丢包率,在重载情况下(非过载),GCQF-3 相对 CQF 丢包率降低了 78%,相对 CQF-3 其丢包率降低 61%。

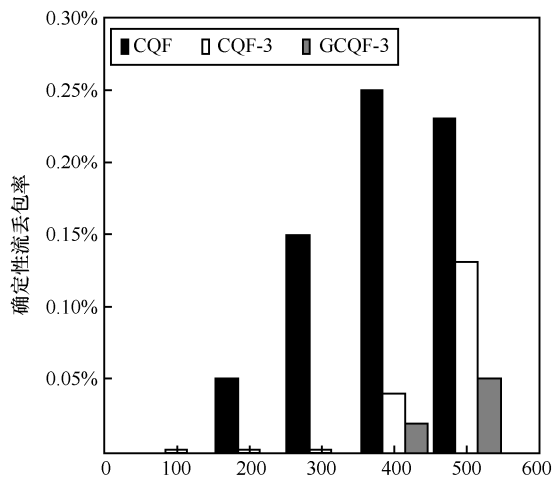


图7 线形拓扑下不同排队策略的平均丢包率

综合上述对比结果,GCQF-3 除了抖动略有增加之外,时延和丢包率都优于 CQF 和 CQF-3。

4.3 定性分析

(1) 网络通信开销

该开销主要来源队列转发决策时进行网络队列监控和状态传递的开销,不仅包括控制平面开销,也包括数据平台的开销。受测量控制分布式和集中式的影响,其网络通信开销各不相同,但总体趋势一致。因 GCQF-3 排队转发机制需获取



端到端全局队列状态，而另外两种队列转发机制属于独立节点自调节方式，GCQF-3 的网络通信开销则远大于 CQF 和 CQF-3，针对时间敏感长距离通信，后续将需进一步研究在低开销和时延及丢包率之间找到平衡的方法。

(2) 灵活拓展性

CQF 和 CQF-3 在解决时延和抖动问题时，并未充分考虑可伸缩性，尤其是随着确定性流增加时。GCQF-3 考虑流量增加的适应性和扩展性，但其状态监控和传递随着网络的增长呈指数增长，对于排队转发的时效性和扩展性带来新的挑战。

5 结束语

本文针对确定性业务需求分析了确定性网络的关键技术需求，并对工业界和学术界的过滤整形和队列转发控制方进行了简单综述。针对大规模确定性网络存在的时延抖动挑战，本文在 CQF 基础上，引入端到端网络状态对队列门控机制进行了优化，提出 GCQF-3 排队机制。并基于 OMNeT++ 搭建一套事件驱动的确定性网络性能评估仿真系统，分析了线形、环形和星形拓扑下 CQF、CQF-3 和 GCQF-3 的队列转发控制机制的时延抖动性能。虽 3 种方式各有优缺点，总体来说，GCQF-3 在满足抖动需求的情况，时延和丢包率都优于其他两种机制。但 GCQF-3 网络和队列状态监控的粒度及消息传递的时机和策略还需进一步优化，以平衡网络开销和时延抖动性能，并进一步提升扩展性。除此之外，后续还将进一步细化排队转发策略中流量负载率和出入度对 GCQF-3 排队转发策略的影响，以及通过增强学习来提高其排队转发策略的及时性和准确性。

参考文献：

[1] 戴彬, 曹园园, 莫益军. 未来网络场景及需求分析综述[J].

电信科学, 2019, 35(8): 39-48.

DAI B, CAO Y Y, MO Y J. A survey on future networks: application scenarios and requirements analysis[J]. Telecommunications Science, 2019, 35(8): 39-48.

[2] 黄韬, 汪硕, 黄玉栋, 等. 确定性网络研究综述[J]. 通信学报, 2019, 40(6): 160-176.

HUANG T, WANG S, HUANG Y D, et al. Survey of the deterministic network[J]. Journal on Communications, 2019, 40(6): 160-176.

[3] MOHAMMADPOUR E, STAI E, MOHIUDDIN M, et al. Latency and backlog bounds in time-sensitive networking with credit based shapers and asynchronous traffic shaping[C]//2018 30th International Teletraffic Congress (ITC 30). September 3-7, 2018, Vienna, Austria. IEEE, 2018: 1-6.

[4] NASRALLAH A, THYAGATURU A S, ALHARBI Z, et al. Performance comparison of IEEE 802.1 TSN time aware shaper (TAS) and asynchronous traffic shaper (ATS)[J]. IEEE Access, 2019(7): 44165-44181.

[5] ZHOU Z F, YAN Y, BERGER M, et al. Analysis and modeling of asynchronous traffic shaping in time sensitive networks[C]//Proceedings of 2018 14th IEEE International Workshop on Factory Communication Systems (WFCS). Piscataway: IEEE Press, 2018: 1-4.

[6] NASRALLAH A, BALASUBRAMANIAN V, THYAGATURU A, et al. Large scale deterministic networking: a simulation evaluation[EB]. 2019.

[7] 郑秀丽, 蒋胜, 王闯, 等. 对网络技术跨代发展的思考: 网络 5.0[J]. 信息通信技术, 2017, 11(6): 37-44.

ZHENG X L, JIANG S, WANG C, et al. A potential direction of next generation data communication network—network 5.0[J]. Information and Communications Technologies, 2017, 11(6): 37-44.

[8] 黄韬, 霍如, 刘江, 等. 未来网络发展趋势与展望[J]. 中国科学: 信息科学, 2019, 49(8): 941-948.

HUANG T, HUO R, LIU J, et al. Development trends and prospects of future networks[J]. Scientia Sinica (Informationis), 2019, 49(8): 941-948.

[9] LIN Y H, JIN X, ZHANG T, et al. Queue assignment for fixed-priority real-time flows in time-sensitive networks: Hardness and algorithm[J]. Journal of Systems Architecture, 2021, 116: 102141.

[10] NASRALLAH A, THYAGATURU A S, ALHARBI Z, et al. Ultra-low latency (ULL) networks: the IEEE TSN and IETF DetNet standards and related 5G ULL research[J]. IEEE Communications Surveys & Tutorials, 2019, 21(1): 88-145.

[11] PUTTNIES H, DANIELIS P, JANCHIVNYAMBUU E, et al. A

- simulation model of IEEE 802.1AS gPTP for clock synchronization in OMNeT++[C]//EasyChair. [S.l.:s.n.], 2018: 63-72.
- [12] BOROUJENY M K, MARK B L, EPHRAIM Y. Stochastic traffic regulator for end-to-end network delay guarantees[C]// Proceedings of ICC 2020 - 2020 IEEE International Conference on Communications (ICC). Piscataway: IEEE Press, 2020: 1-6.
- [13] Alexej Grigorjew, Florian Metzger, Tobias Hoßfeld, et al. Asynchronous traffic shaping with jitter control[EB]. 2020.
- [14] THIELE D, ERNST R. Formal worst-case timing analysis of Ethernet TSN's burst-limiting shaper[C]//Proceedings of the 2016 Design, Automation & Test in Europe Conference & Exhibition (DATE). Piscataway: IEEE Press, 2016: 187-192.
- [15] ADDANKI V, IANNONE L. Moving a step forward in the quest for Deterministic Networks (DetNet)[C]//2020 IFIP Networking Conference (Networking). Piscataway: IEEE Press, 2020: 458-466.

[作者简介]



莫益军 (1976-), 男, 博士, 华中科技大学计算机科学与技术学院副教授, 主要研究方向为智能网络、边缘计算。

杨子涵 (1998-), 男, 华中科技大学计算机科学与技术学院硕士生, 主要研究方向为机器学习、确定性网络。

刘辉宇 (1978-), 男, 博士, 华中科技大学计算机科学与技术学院讲师, 主要研究方向为网络、自然语言处理、推荐技术。

何天流 (1999-), 男, 华中科技大学计算机科学与技术学院硕士生, 主要研究方向为联邦学习、边缘计算。