

基于软件定义的时间敏感网络跨域调度机制

汪硕^{1,2}, 黄玉栋¹, 黄韬^{1,2}, 霍如³, 刘韵洁^{1,2}

(1. 北京邮电大学网络与交换国家重点实验室, 北京 100876; 2. 网络通信与安全紫金山实验室, 江苏 南京 211111;
3. 北京工业大学信息学部, 北京 100124)

摘要: 针对时间敏感网络的跨域调度问题, 提出了基于软件定义的时间敏感网络跨域调度机制。该机制基于 3 种广域时钟同步技术来保障纳秒级域间时间同步。提出了时隙无冲突跨域调度模型, 解决了域间链路时延带来的时隙对齐问题, 利用时隙无冲突约束严格限制跨域流量在边缘节点的到达时间, 最终在线计算得到融合各域的全局调度门控列表。实验结果表明, 在典型的跨域远程工控场景下, 所提模型能调度上百条时间敏感流量, 并具有可行的秒级逐流执行时间。

关键词: 时间敏感网络; 软件定义; 跨域调度; 门控列表; 服务质量

中图分类号: TP393

文献标识码: A

DOI: 10.11959/j.issn.1000-436x.2021182

Software-defined cross-domain scheduling mechanism for time-sensitive networking

WANG Shuo^{1,2}, HUANG Yudong¹, HUANG Tao^{1,2}, HUO Ru³, LIU Yunjie^{1,2}

1. State Key Laboratory of Networking and Switching Technology, Beijing University of Posts and Telecommunications, Beijing 100876, China

2. Purple Mountain Laboratories, Nanjing 211111, China

3. Information Department, Beijing University of Technology, Beijing 100124, China

Abstract: For the cross domain scheduling problem of time-sensitive network, a software-defined cross-domain scheduling mechanism for time-sensitive networking was proposed, which was based on three kinds of wide area clock synchronization technology to ensure nanosecond inter-domain time synchronization. Then, a time slot conflict-free cross domain scheduling model was proposed to solve the problem of time slot alignment caused by inter-domain link delay. The time slot conflict-free constraints were used to strictly limit the arrival time of cross-domain traffic at the edge node. Finally, the gate control list of fusion domains for global scheduling was calculated online. The simulation results show that the model can schedule hundreds of time-sensitive traffic in a typical cross-domain remote control scenario, and has a feasible algorithm execution time in seconds.

Keywords: time-sensitive networking, software-defined, cross-domain scheduling, gate-controlled list, QoS

1 引言

在许多物理网络环境中, 如工业自动化、车载

网络、航空航天网络, 关键实时应用对网络的带宽、时延、抖动、丢包率有严格的要求。一方面, 当前部署在物理网络环境中的传统总线存在带宽速率

收稿日期: 2021-05-06; 修回日期: 2021-07-19

基金项目: 广东省重点领域研发计划基金资助项目 (No.2018B010113001); 国家重点研发计划基金资助项目 (No.2018YFB1800500); 国家自然科学基金资助项目 (No.61902033); 北京市自然科学基金资助项目 (No.4204105); 中央高校基本科研业务费专项资金资助项目 (No.2021RC02)

Foundation Items: The Research and Development Program in Key Areas of Guangdong Province (No.2018B010113001), The National Key Research and Development Program of China (No.2018YFB1800500), The National Natural Science Foundation of China (No.61902033), Beijing Natural Science Foundation (No.4204105), The Fundamental Research Funds of the Central Universities (No.2021RC02)

低、标准互不兼容的缺点。另一方面，以太网相比于传统总线具有支持更高带宽速率、更低开销，以及更易于与现有 IT 系统整合的特点，但因其尽力而为的转发模式存在长尾时延效应，无法满足关键实时应用的传输要求。因此，时间敏感网络 (TSN, time-sensitive networking) 致力于在以太网的基础上增强确定性转发能力，从而实现时间敏感流量和尽力而为流量的共网传输。

时间敏感网络工作组提出的全网时钟同步、时间感知门控等机制，已经能够满足局域网范围内的实时应用传输需求^[1-2]。但随着 5G 超可靠低时延、云化 PLC (programmable logic controller)、远程控制、工厂互联等大规模时间敏感应用的兴起，如何实现时间敏感网络的跨域调度成为当前的一个重要挑战。

时间敏感网络的工业自动化标准^[3]提出了时间敏感网络域的概念，一个时间敏感网络域是一个终端与交换设备的管理组。在机器与机器通信时，如控制器与执行器、主动机械臂和从动机械臂、控制器与控制器间通信，存在机器属于不同域的场景，跨域的应用需要运行在整个融合的时间敏感网络域的顶层。在数据平面，Dorr 等^[4]提出用域 ID 标识 TSN 域，并将 TSN 域抽象为虚拟网桥，然后根据目的 MAC 地址、VLAN 值、优先级值这 3 个参数在发往邻域的出端口进行时间敏感流的识别与转换。在控制平面，Chen^[5]提出了一个新的配置实体间协议 (CCP, config-entity to config-entity protocol) 来进行不同域的网络控制器间的协商与信息传输，使用链路层发现协议来寻找相邻的 TSN 域。Böhm 等^[6]设计了一个时间敏感网络控制器间的东西向协议，并基于 Ryu 控制器进行了跨域的门控列表配置下发的概念原型验证。

跨域调度的一个关键难题是如何生成多域融合的门控列表配置信息。以前的相关工作大多讨论局域网内的调度生成机制，其将链路时延视为零，且认为单域内设备时隙严格对齐。在最新的时间敏感网络跨域研究中，文献[7]通过部署确定性 IP 路由器来保证每跳转发时延有界且小于 30 μs ，成功验证了可以在工厂边缘利用虚拟 PLC 跨域远程控制执行器，但其只讨论了一条时间敏感流与尽力而为流共传的场景，未讨论多条时间敏感流跨域调度的问题。文献[8]将工厂车间网络划分为多个子网，并提出一种利用工厂网络层次结构的

等时流量调度方法，其构建了域内域间分开调度的两步调度思想，但并未给出严格意义的跨域调度框架和数学模型。

软件定义网络将网络控制逻辑与其底层硬件分离，能够提供网络的全局视图，实现网络可编程。本文提出了基于软件定义的时间敏感网络跨域调度机制，联合调度路径、队列和时隙资源，保证跨域场景下的流量有界时延指标。该机制首先融合全球导航卫星系统 (GNSS, global navigation satellite system)^[9]、广域精确授时 (WPT, wide-area precise timing)^[10]、地面光纤高精度授时 WR-PTP (white rabbit precision time protocol)^[11] 技术，实现跨域场景下的 1~100 ns 域间时间同步，达到时间敏感网络 802.1AS 标准所要求的精度。然后，为解决域间设备长距链路时延带来的时隙对齐问题，提出了时隙无冲突跨域调度模型，利用时隙无冲突约束严格限制跨域流量在边缘节点出端口的到达时间，从而在叠加链路时延后得到精确的邻域到达时间，最终在线计算生成融合各域的全局调度门控列表，通过精确的全局门控列表控制来保证流量时延有界。实验结果表明，在典型的跨域远程工控场景下，该模型能调度上百条时间敏感流量，并具有可行的秒级的算法执行时间。

2 背景与动机

2.1 全网时钟同步

精确的全网时钟同步是基于时隙的调度整形的基础。时间敏感网络的 802.1AS-Rev 标准定义了通用精准时钟协议 (gPTP, generalized precision time protocol)，gPTP 利用最佳主时钟选择算法选出主时钟，并以主时钟为根节点形成树状拓扑，在频率对齐的基础上，对设备和链路时延进行实时测量补偿校准，其他节点以主时钟为参考进行匹配调节和主从同步，最终实现纳秒级的全网时钟同步。gPTP 与精确时钟同步 PTP 的区别在于它是一个完全基于二层网络的非 IP 路由的协议，其精简了无关的二层机制，增强了二层的时间同步机制。

在工业领域，时钟同步主要有普遍时钟和工作时钟这两种同步要求^[12]。普遍时钟存在于工厂内部和车间范围，采用外接时钟源来实现精度不低于 100 μs 的工厂内时间同步；工作时钟主要应用于生

产线内的设备和机器间的精度不低于 $1 \mu\text{s}$ 的时间同步，例如具有精准事件序列的机器人运动控制、实时数据传输等。因为时间敏感网络是一个局域网标准，目前还没有涉及工厂间、跨域及广域场景下的时钟同步方案。

2.2 时间感知整形

为严格保障关键应用的时延和抖动要求，在全网时钟同步的基础上，时间感知整形机制被提出。时间感知整形 (TAS, time-aware shaper) 是增强的出端口流量调度整形机制，采用类似时分复用的方式保证时间敏感的业务流在传输中具有逐跳的确定性时延。

TAS 在每一个优先级队列后增加了一个由门控列表里的门控条目驱动控制的门。当门是打开状态时，对应队列中业务流进行出队传输；当门是关闭状态时，对应队列中业务流不可被传输。在时间敏感流的队列门打开前，TAS 通过设定保护带宽，保证其他流量的传输不阻碍时间敏感流的传输。时间感知整形的配置对象为门控列表条目，每个条目需要配置的参数为时间点以及对应时刻的门状态（打开或者关闭），并可计算选配最大为 1 522 B 的保护带宽大小。在局域网范围内，基于时间感知整形的相关硬件设备已趋于成熟，相关局域调度和优化算法已有大量的研究。Steiner^[13]自 2010 年开始利用 SMT/OMT 求解器研究了时间触发以太网中的时间触发流的调度合成问题。Dürr 等^[14]设计了一种基于整数线性规划公式的车间工件调度模型来计算单局域网内的门控列表，并提出了一种软件定义的多周期流传输调度方法^[15]。

2.3 动机

时间敏感网络工作组提出的全网时钟同步、时间感知门控等机制当前被限制在 7 跳网络半径的局域网范围内。但新兴的 5G 超可靠低时延、云化 PLC、远程控制、工厂互联、边缘实时计算等大规模时间敏感应用要求实现跨域的实时流量调度。例如，云化 PLC 场景需要将 PLC 从生产线迁移到工厂边缘甚至另一个工厂，从云端经过城域网、接入网接入工厂内部对机器设备进行实时操作和运动控制。基于 5G 的车路协同、虚拟现实交互场景，要求网络达到 1~10 ms 的端到端时延和微秒级抖动。

网络中的一跳时延包括链路的传播时延、交换机内部的处理时延、出端口队列的排队时延，

以及出端口发送的传输时延。因为局域网内链路长度通常小于 100 m，其调度算法将设备间的链路时延设为零，并认为全网是严格时间同步的。然而，在跨域调度中，首先各个子域的设备时钟只与该域的主时钟进行同步，域间缺少时钟同步机制；其次，跨域传输的链路时延不可忽视，从一个域的出端口传输到另一个域的出端口，存在跨域端口的门控列表的时隙对齐问题，即如果包的到达存在抖动，则无法根据包在一个域的发送时间确定另一个域的门控列表的开关时间，时隙错位严重时会导致跨域关键流量包丢失。

为解决上述问题，本文提出了基于软件定义的时间敏感网络跨域调度机制，通过软件定义网络技术将网络控制逻辑与其底层硬件分离，利用其提供的网络全局视图，对跨域时间敏感流量实现确保有界时延抖动的全局调度。软件定义产生的控制器开销以及软件定义网络 (SDN, software defined network) 对 TSN 的影响可参考控制架构相关研究^[6,16]。

3 软件定义的时间敏感网络跨域调度机制描述

3.1 软件定义的时间敏感网络

软件定义的时间敏感网络跨域调度机制如图 1 所示，本文采用 802.1Qcc 标准定义的全集中式控制架构。控制平面组件包括中心化用户配置 (CUC, centralized user configuration) 和中心化网络控制器 (CNC, central network controller)。CUC 相当于编排器，负责采集终端业务的带宽时延抖动等网络服务质量需求，并将其转换后通过北向接口发给中心化网络控制器；CNC 相当于控制器，包含计算拓扑路径等网络功能，并通过南向接口下发更新路由路径、门控列表等配置信息给时间敏感网络交换机。

数据平面设备包括时间敏感交换机和终端。时间敏感交换机支持时钟同步和时间感知整形机制，每个交换机的出端口有 0~7 共 8 个优先级队列，其中时间敏感流量被标记为最高优先级 7，尽力而为流量被标记为剩余的低优先级。终端软件间的语义互操作采用开放平台通信统一架构 (OPC-UA, open platform communications united architecture) 的发布订阅模式，以保证发布者和订阅者能在跨域调度代理机中基于统一数据结构

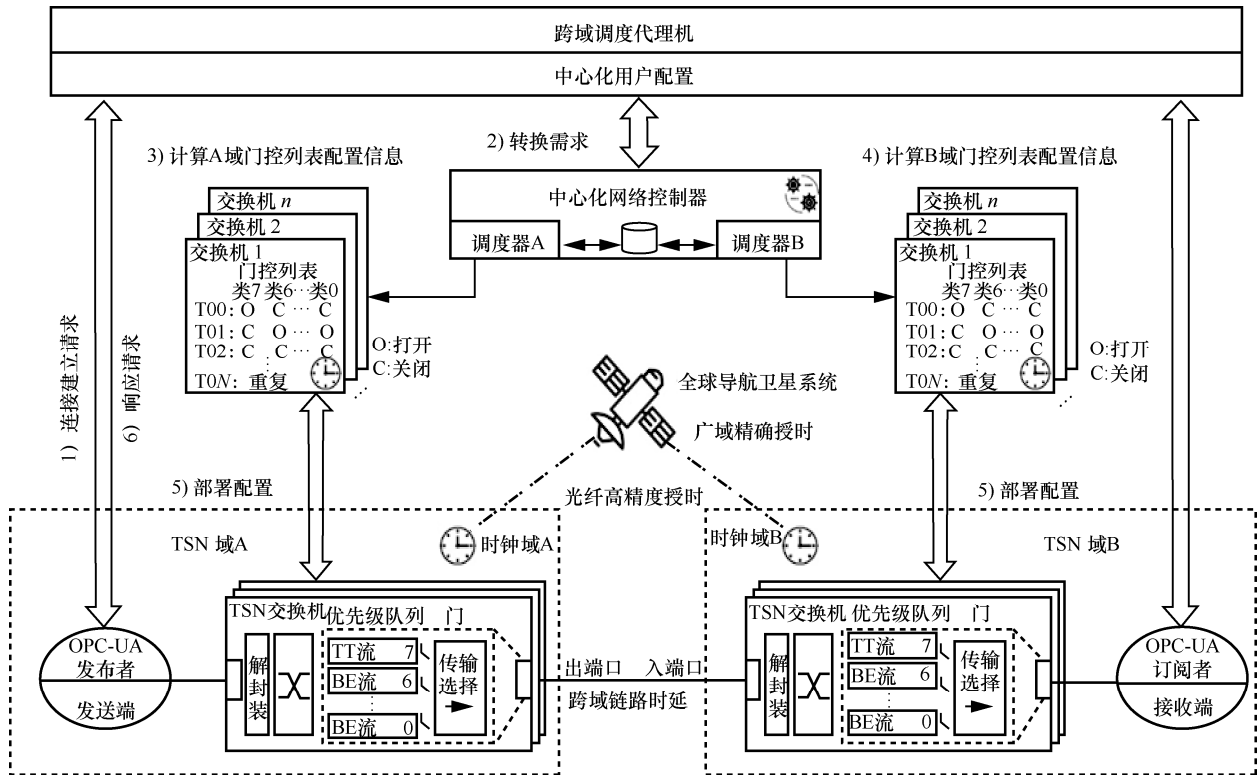


图 1 软件定义的时间敏感网络跨域调度机制

的主题进行交互。

时间敏感网络的跨域调度工作流程如下。1) TSN 域 A 中的发送端向中心化用户配置发起连接建立请求; 2) 中心化用户配置将发送端的业务服务质量请求转换为具体的网络需求参数, 并下发给中心化网络控制器; 3) 中心化网络控制器根据 A 域的拓扑和时隙资源计算跨域流量在 A 域的门控列表配置信息; 4) 中心化网络控制器根据 A 域的出端口门控时间以及跨链路时延, 计算得到流量到达 B 域的时间, 再根据 B 域的拓扑和时隙资源计算跨域流量在 B 域的门控列表配置信息; 5) 若跨域调度成功, 则分别对两域的交换机下发更新门控列表配置; 6) 中心化用户配置响应请求, 返回具体的发包开始时间给发送端。

3.2 域间时钟同步

本文机制首先基于 GNSS^[9]、WPT^[10]、WR-PTP^[11]技术来协同保障跨域场景下的 1~100 ns 域间时间同步, 从而达到时间敏感网络 802.1AS 标准所要求的精度。

在 TSN 域内, 主节点配置高级晶振主时钟(铯原子钟或铷原子钟), 其授时精度近似于世界协调时间这一基准时间, 可作为域内的主时钟; 其他

节点为从时钟, 主从时钟间通过通用精准时钟协议实现精确时间同步。

在 TSN 域间, 可采用包括美国的全球定位系统 GPS 和中国的北斗卫星系统在内的全球导航卫星系统进行卫星授时。卫星时钟和晶振时钟的误差特性互补。卫星时钟的秒时钟误差服从正态分布, 随机误差大, 累积误差小; 而晶振时钟与国际标准时间相比, 每秒的时间偏移很小, 且偏移量较稳定, 其随机误差小, 累积误差大。卫星时钟与晶振时钟通过具有信号跟踪能力的数字锁相环进行结合, 产生高精度和高稳定性的时钟。此外, 基于 GNSS 载波相位观测量及状态空间域改正的广域精确授时技术, 在中国科学院国家授时中心经过连续数月的测试后验证了授时精度能够达到 0.48 ns。最后, 基于光纤网络的地面高精度授时系统能实现 1 000 km 覆盖范围内的纳秒级时钟同步, 也可作为域间时钟同步的补充。

3.3 调度时隙设计

为了实现时间敏感流的有界时延抖动传输, 时间控制的门根据门控列表条目在队列中每个时隙的开始处执行打开或关闭操作。以图 1 为例, 当时间敏感流在 T00 时刻到达时, 队列 7 的门打

开，使时间敏感流无排队地开始传输，而其余门关闭。在时间敏感流传输完成之后，其余的门在 T01 时刻打开以传输尽力而为流，而队列 7 的门关闭。在时间敏感流再次到达之前，在 T02 时刻，所有门都关闭以形成最大传输单元大小的保护带。最后，重复执行门控列表的条目以进行周期传输调度。

因此，流的调度由门控列表的计算决定，而计算的优化取决于时隙划分的设计和时隙敏感流发送开始时间的确定。为满足中心化调度的实时性，本文设计了在线增量式调度方法和变长调度时隙。一方面，以前的离线调度方法^[14,17-18]要求提前已知所有流的信息，需要花费数小时甚至数天的时间进行计算，以合成接近最优的调度方案；本文采用增量式调度方法，按流的固有顺序逐流计算流的调度方案，从而减少计算时间开销，保证中心化控制器调度的实时性。另一方面，离线调度为简化计算参数，常将时隙大小设为固定值，如传输最大传输单元大小的包所需的时间^[19-20]，甚至传输 7 跳长度所需的时间^[15]，造成极大的带宽和时隙资源浪费；本文采用变长调度时隙，根据包的实际大小计算出包的传输时延，并将其作为模型的调度时隙，从而提高带宽利用率。

4 时隙无冲突的跨域调度模型

在域间保证严格的时钟同步后，针对时间敏感流量跨域调度的时隙对齐问题，本节设计了时隙无冲突的跨域调度数学模型。跨域调度需要解决 3 个问题：1) 如何设计跨域流量在域 A 内的传输调度，以保证流能准时准确地到达域边节点；2) 如何获得跨域到达时刻，以计算域 B 的门控列表，并保证两域的门控时隙对齐；3) 如何压缩域间流量所使用的时隙资源，以保留更多的时隙资源给域内流量。

4.1 输入和决策变量

本文使用花体字母表示集合，使用大写粗体字母表示矩阵，使用小写粗体字母表示向量。网络拓扑建模为有向图 $\mathcal{G}=(\mathcal{V},\mathcal{E})$ ，其中， \mathcal{V} 是节点集合， \mathcal{E} 是边的集合。此外， $\mathcal{V}=(\mathcal{S}\cup\mathcal{H})$ ，其中 \mathcal{S} 和 \mathcal{H} 分别是交换机和主机的集合。 \mathcal{E} 是一组表示网络中链路的元组， $\mathcal{E}=\{(i,j)|i,j\in\mathcal{V},i\neq j\}$ ，并且节点 i 和节点 j 之间存在一条边。跨域的时间

敏感流被表示为元组 $f=(s_f,d_f,p_f,t_f,dl_f)$ ，其中， s_f 是源节点， d_f 是目的节点， p_f 是发送周期， t_f 是流的传输时延， dl_f 是流的截止时间。

表 1 给出了用于调度问题的拓扑相关和流相关的参数^[21]。

表 1 拓扑及流相关参数	
参数表达式	参数说明
$\mathcal{E}=\{0,1,2,\dots\}$	链路集合
$\mathcal{V}=\{0,1,2,\dots\}$	节点集合
$\mathbf{A}_{EE}\in\{0,1\}^{ \mathcal{E} \times \mathcal{E} }$	稀疏邻边矩阵
$\mathcal{F}_{od}\subset\mathbb{N}$	跨域时间敏感流集合
$s_F\in\mathcal{V}^{ \mathcal{F} }$	源节点向量，如果 $f\in\mathcal{F}$ 开始于 $v\in\mathcal{V}$ ，则 $s_F[f]=v$
$d_F\in\mathcal{V}^{ \mathcal{F} }$	目的节点向量，如果 $f\in\mathcal{F}$ 到达于 $v\in\mathcal{V}$ ，则 $d_F[f]=v$
$p_F\in\mathbb{N}^{ \mathcal{F} }$	流发送周期向量，如果 $f\in\mathcal{F}$ 有周期 $p_f\in\mathbb{N}$ ，则 $p_F[f]=p_f$
$t_F\in\mathbb{N}^{ \mathcal{F} }$	传输时延向量，如果 $f\in\mathcal{F}$ 有 $t_f\in\mathbb{N}$ ，则 $t_F[f]=t_f$
$dl_F\in\mathbb{N}^{ \mathcal{F} }$	截止时间向量，如果 $f\in\mathcal{F}$ 有 $dl_f\in\mathbb{N}$ ，则 $dl_F[f]=dl_f$
$h_F\in\mathbb{N}^{ \mathcal{F} }$	一跳时延向量，如果 $f\in\mathcal{F}$ ， $h_f=t_f+t_{prop}+t_{proc}$
$h=\text{LCM}(p_F)$	超周期，等于所有流周期的最小公倍数

超周期等于所有流量周期的最小公倍数。在每个超周期中，传输模式从全局角度重复，因此只需要确保时间敏感流在一个超周期长度上满足调度约束。下面介绍跨域时隙无冲突调度的决策变量、约束条件和求解目标。

为了描述边与边之间的关系，本文使用边-边邻接矩阵 \mathbf{A}_{EE} 的元素的值，并在行 $e_p\in\mathcal{E}$ 和列 $e_n\in\mathcal{E}$ 有如下定义。

$$\mathbf{A}_{EE}[e_p][e_n]=\begin{cases} 1, & e_p=(*,v)\text{且}e_n=(v,*) \\ 0, & \text{其他} \end{cases} \quad (1)$$

式(1)表明，如果节点 $v\in\mathcal{V}$ 既是边 e_p 的目的节点又是边 e_n 的源节点，则 $\mathbf{A}_{EE}[e_p][e_n]$ 值为 1；否则为 0。 $|\mathcal{E}|$ 表示边集合中的元素数量。

为了描述流与边之间的关系，本文使用一组二进制决策变量 $u\in\{0,1\}^{|\mathcal{F}|\times|\mathcal{E}|}$ 并有如下定义。

$$u[f_{cd}][e] = \begin{cases} 1, & \text{流 } f_{cd} \text{ 占用边 } e \\ 0, & \text{流 } f_{cd} \text{ 未占用边 } e \end{cases} \quad (2)$$

本文采用最短路径算法来确定路由，且对于每一个 $u[f_{cd}][e]=1$ ，有一个传输开始时间矩阵的值 $t[f_{cd}][e]$ 表示流 f_{cd} 在边 e 上的传输开始时间。

整数线性规划是一个被广泛用于建模网络调度问题的数学工具，且有成熟商用的高性能求解器。该整数线性规划模型的输入变量和决策变量如下。

- 1) 输入变量
 - 全局网络拓扑图为 $\mathcal{G}=(\mathcal{V},\mathcal{E})$ 。
 - 跨域时间敏感流集合为 $\mathcal{F}_{cd} \subset \mathbb{N}$ 。
- 2) 决策变量
 - 流所流经的边的向量为 $u[f_{cd}][e]$ 。
 - 流在边上的传输开始时间为 $t[f_{cd}][e]$ 。

4.2 约束条件

为保证跨域流量在域 A 内能准时准确地到达域边节点，本文采用时隙无冲突调度模型，即每个包在交换机的出端口进行传输的开始互不冲突，一个包完全传输完成后再传输下一个包。这样实现了出端口最大队列长度为 1，因此排队时延为 0，包在到达域边节点时不会产生抖动。其约束条件如下。

1) 传输开始时间约束

如果流在边上的传输开始时间 $t[f_{cd}][e]$ 大于流的发送周期，则它应该对流的周期取模以符合周期性调度，如式(3)所示。式(4)确保每个输出端口上的流传输在下一个传输周期开始之前结束，即流在超周期中各个发送周期内的传输互不影响。时间表 $t[f_{cd}][e]$ 中的每一列都代表边的门控状态改变的时间点，用该时间点可合成每个输出端口上的门控列表。

$$t[f_{cd}][e]=t[f_{cd}][e] \bmod p_F[f_{cd}], t[f_{cd}][e] \geq p_F[f_{cd}] \quad (3)$$

$$0 \leq t[f_{cd}][e] \leq p_F[f_{cd}] - t_F[f_{cd}] \quad (4)$$

2) 时隙无冲突约束

在任何周期内，对于流 f_{cd1} 和任何流 f_{cd2} ，每个交换机的出端口的传输时间都不重叠。该约束定义为

$$\forall f_{cd1} \in \mathcal{F}_{cd}, f_{cd2} \in \mathcal{F}_{cd}, e \in \mathcal{E}:$$

$$\forall x \in \mathcal{X}, y \in \mathcal{Y}:$$

$$\text{若 } u[f_{cd1}][e] + u[f_{cd2}][e] \geq 2, \text{ 则}$$

$$\begin{aligned} t[f_{cd1}][e] + xp_F[f_{cd1}] &\geq t[f_{cd2}][e] + yp_F[f_{cd2}] + \\ t_F[f_{cd2}] &\text{ 或 } t[f_{cd2}][e] + yp_F[f_{cd2}] &\geq t[f_{cd1}][e] + \\ xp_F[f_{cd1}] + t_F[f_{cd1}] & \end{aligned}$$

其中

$$\begin{aligned} \mathcal{X} &= \left\{ x \in \mathbb{N} : 0 \leq x \leq \frac{h}{p_F[f_{cd1}]} \right\} \\ \mathcal{Y} &= \left\{ y \in \mathbb{N} : 0 \leq y \leq \frac{h}{p_F[f_{cd2}]} \right\} \end{aligned} \quad (5)$$

3) 域内邻边约束

在域内，链路传播时延可视为 0，又因为采用时隙无冲突调度，排队时延也为 0，因此一跳时延等于交换机内处理时延加上端口传输时延，是一个恒定值。前一条边的传输开始时间 $t[f_{cd}][e_p]$ 加上一跳时延 $h_F[f_{cd}]$ （如果超出流周期，则错开一个周期），等于下一跳的传输开始时间 $t[f_{cd}][e_n]$ 。该约束定义为

$$\begin{aligned} \forall e_p \in \mathcal{E}, \forall e_n \in \mathcal{E} : e_p \neq e_n, \text{ 且 } A_{EE}[e_p][e_n] &= 1, \\ \forall f_{cd} \in \mathcal{F}_{cd} : \\ \text{若 } u[f_{cd}][e_p] + u[f_{cd}][e_n] &\geq 2, \text{ 则} \\ t[f_{cd}][e_n] &= t[f_{cd}][e_p] + h_F[f_{cd}] \\ \text{或 } t[f_{cd}][e_n] &= t[f_{cd}][e_p] + h_F[f_{cd}] - p_F[f_{cd}] \end{aligned} \quad (6)$$

4) 域间邻边约束

在域间，链路传播时延不可视为 0，假设域间链路时延为 λ ，则到达域 B 的传输开始时间 $t[f_{cd}][e_n]$ 等于域 A 边缘节点的传输开始时间 $t[f_{cd}][e_p]$ 加上一跳时延 $h_F[f_{cd}]$ 以及域间链路时延 λ 。根据式(7)可获得跨域到达时刻以计算域 B 的门控列表，并保证两域的门控时隙对齐。该约束定义为

$$\begin{aligned} \forall e_p \in \mathcal{E}, \forall e_n \in \mathcal{E} : e_p \neq e_n, \text{ 且 } A_{EE}[e_p][e_n] &= 1 \\ e_p \in \text{域A}, e_n \in \text{域B} \\ \forall f_{cd} \in \mathcal{F}_{cd} : \\ \text{若 } u[f_{cd}][e_p] + u[f_{cd}][e_n] &\geq 2, \text{ 则} \\ t[f_{cd}][e_n] &= t[f_{cd}][e_p] + h_F[f_{cd}] + \lambda \\ \text{或 } t[f_{cd}][e_n] &= t[f_{cd}][e_p] + h_F[f_{cd}] + \lambda - p_F[f_{cd}] \end{aligned} \quad (7)$$

5) 截止时间约束

对于每一条跨域流量，其端到端时延由域 A 内逐跳时延之和、跨域链路时延以及域 B 内逐跳时延之和三部分组成。该约束要求流在截止时间之内到

达目的节点，即

$$\sum_{\text{域A}} h_F[f_{cd}] + \sum_{\text{域B}} h_F[f_{cd}] + \lambda \leq \mathbf{dl}_F[f_{cd}] \quad (8)$$

4.3 目标函数

为保留更多的时隙资源给域内流量，本文压缩域间流量所使用的时隙资源，将模型的目标函数设置为最小化跨域时间敏感流在发送第一跳的边 e_s 的发送开始时间，该发送时间定义为 SST，目标函数可以表示为

$$\text{SST} = \text{mini}(t[f_{cd}][e_s]) \quad (9)$$

5 仿真实验

本节评估在第 4 节所述时隙无冲突跨域调度模型的性能。首先给出仿真设置；然后评估模型求解性能，包括流调度能力、可扩展性和执行时间。

5.1 仿真设置

本文仿真使用 Python 实现时隙无冲突跨域调度实例，使用 Networkx 图形库创建复杂的网络拓扑，生成不同的流量模式，使用 IBM ILOG CPLEX 优化求解器的 Python API (Linux 版本 12.10.0) 求解整数线性规划公式。计算服务器配有 4 个英特尔 i5-8259U 中央处理单元和 16 GB 的随机存取存储器。

1) 拓扑选择

线形、环形和雪花形是工业控制网络的主要拓扑类型^[22]。本文将 A/B 域各有 8 个节点的这 3 种拓扑作为测试拓扑结构，在域的边缘随机选择一个节点作为跨域连接节点。链路带宽被设置为 1 Gbit/s。为了验证算法的可扩展性，将跨域链路时延集合设置为 {0.2, 1, 5} ms，分别代表接入网、城域网、广域网 3 种不同距离的网络中^[23]的跨域调度场景。

2) 流量生成

对于集合中的每个流 f ，通过从顶点集中均等概率地选择 2 个值来创建源节点和目的节点，以降低路由算法对调度结果的影响。所有流量都在广域监视和控制系统(由 IEC 61850 定义)以及 DetNet 用例的工业机器与机器通信的流量特性的指导下生成^[24]。数据包长度为 100~1 500 B。通常流的发送周期和截止时间为毫秒量级，从集合 {1, 2, 4, 8} ms 中随机抽取一个值来模拟发送周期 p_F ，从 20~50 ms 随机选择一个值来模拟流的截止时间 \mathbf{dl}_F ，以保证实验的普遍适用性。

5.2 调度能力仿真结果

调度能力表示可调度跨域时间敏感流的数量。设置跨域链路时延为 0.2 ms，将流数量划分为 200~800 条的 4 个级别，每个级别进行 5 次实验，取平均值作为结果。当前工厂互联、远程控制等跨域时间敏感应用需求远小于域内应用，因此上百条流已经可以满足当前的调度需求。探索调度上千条跨域流量将作为未来工作。

环形拓扑下，不同时隙大小对调度能力的影响如图 2 所示。在调度 200 条流时，变长时隙和最大传输单元 (MTU, maximum transmission unit) 大小的时隙均能调度接近 100% 的流量，而以 7 跳 MTU 大小为时隙的方案只能调度约 40% 的流量。随着流数量增加，各方案调度能力均有所下降，但变长时隙一直保持最高的可调度流数量；相比于 MTU 固定时隙，变长时隙的调度能力最大提升了 18%，相比于 7 跳 MTU 固定时隙，变长时隙的调度能力最大提升了 69%。因此，后面的调度仿真均采用变长时隙方案。

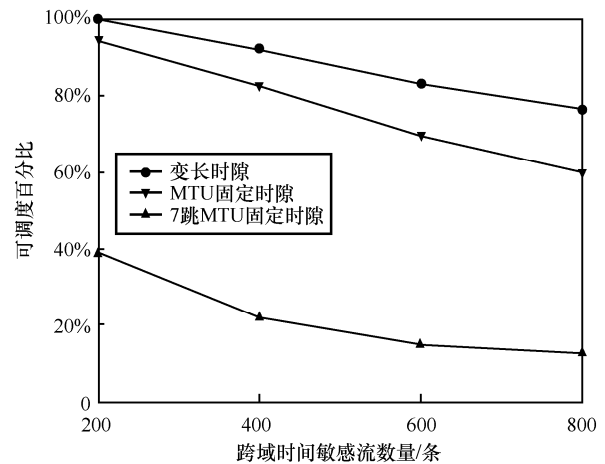


图 2 不同时隙大小对调度能力的影响

不同拓扑对调度能力的影响如图 3 所示，雪花形拓扑中可调度的流数量最多，线形拓扑中可调度的流数量最少，环形拓扑与雪花形拓扑接近。其原因是线形拓扑流调度时流经的跳数更多，单条流会占用更多的出端口时隙资源，导致调度能力下降。在调度 800 条流时，雪花形拓扑中的调度能力比线形拓扑提升了 14.5%。

5.3 可扩展性仿真结果

环形拓扑下，将流数量划分为 200~800 条的 4 个级别，不同跨域距离对模型可扩展性的影响如图 4 所示。接入网、城域网、广域网这 3 种网络中，模

型的可调度流数量相差不大，最低也保持了 77.5% 的调度能力。因此，随着跨域链路时延增大，模型的调度能力没有下降，该模型具有良好的可扩展性。

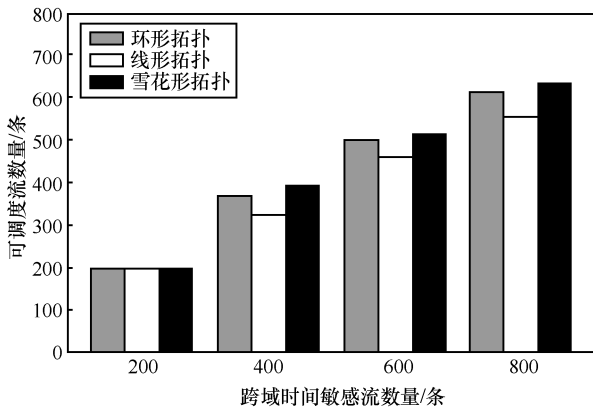


图 3 不同拓扑对调度能力的影响

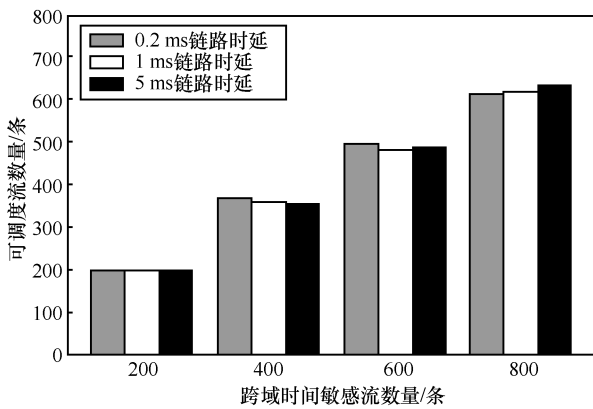


图 4 不同跨域距离对模型可扩展性的影响

此外，本节设置了两组发送周期，一组高频周期为 {1,2,4,8} ms，一组低频周期为 {4,8} ms。0.2 ms 跨域链路时延下，不同发送周期对模型可扩展性的影响如图 5 所示，低频周期组在 4 个流级别均能调度 100% 的流量。这是因为低频流在一个超周期内占用的时隙资源相比高频流更少，产生调度冲突的概率大大降低。随着流的发送周期增大，模型的调度能力将显著提高。

5.4 执行时间仿真结果

本节在环形拓扑、0.2 ms 跨域链路时延下，将流数量划分为 200~800 条的 4 个级别，对比了在线调度和离线调度的执行时间。因为执行时间的值变化很大，将 log₁₀ 的对数函数作为执行时间，其累积分布函数如图 6 所示，随着调度流数量增多，在线调度执行时间不断增大，但最少有 60% 的流量逐流调度时间小于 1 s，所有流的逐流调度时间均小于 10 s，即具有秒级的在线调度执行时间；而离线

调度需要数小时，平均每条流需要几十秒的执行时间，因此离线调度的累积分布函数是一条直线，远大于在线调度方法。

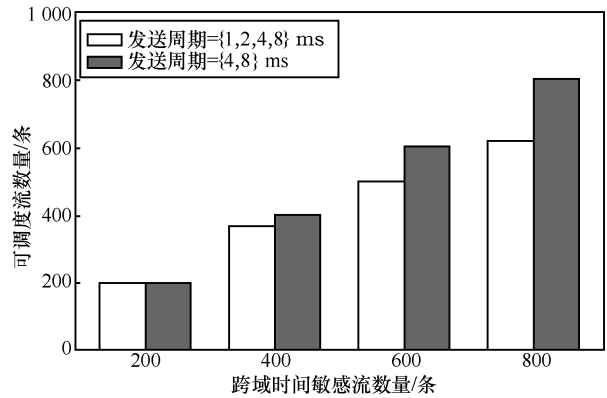


图 5 不同发送周期对模型可扩展性的影响

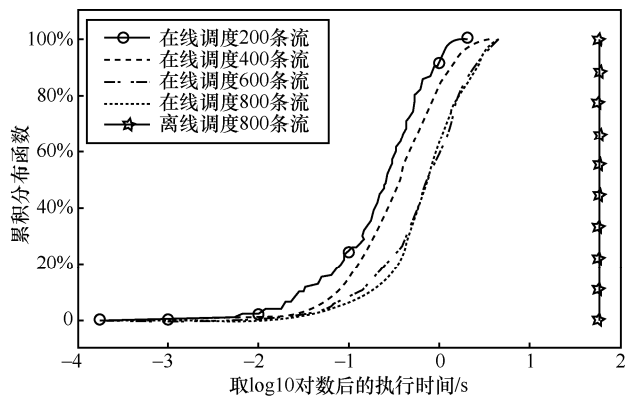


图 6 在线和离线调度的执行时间对比

6 结束语

本文提出了基于软件定义的时间敏感网络跨域调度机制，融合了 3 种广域时钟同步技术来保障纳秒级域间时间同步，设计了时隙无冲突跨域调度模型，解决了域间链路时延带来的时隙对齐问题。实验结果表明，在典型的跨域工控场景下，该模型能调度上百条时间敏感流量，并具有秒级的逐流执行时间。

参考文献:

[1] 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.

[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] IEC/IEEE. IEC/IEEE 60802 Time-sensitive networking profile for industrial automation[S]. IEEE, 2020.
- [4] DORR J, HÖME S, KERSCHBAUM S. TSN domain boundary considerations[S]. IEEE, 2020.
- [5] CHEN L. TSN configuration interaction[S]. IEEE, 2019.
- [6] BÖHM M, OHMS J, WERMSEER D. Multi-domain time-sensitive networks - an east-westbound protocol for dynamic TSN-stream configuration across domains[C]//Proceedings of 2019 24th IEEE International Conference on Emerging Technologies and Factory Automation (ETFA). Piscataway: IEEE Press, 2019: 1363-1366.
- [7] BADAR A, LOU D Z, GRAF U, et al. Intelligent edge control with deterministic-IP based industrial communication in process automation[C]//Proceedings of 2019 15th International Conference on Network and Service Management (CNSM). Piscataway: IEEE Press, 2019: 1-7.
- [8] HELLMANN D, GLAVACKIJ A, FALK J, et al. Scaling TSN scheduling for factory automation networks[C]//Proceedings of 2020 16th IEEE International Conference on Factory Communication Systems (WFCS). Piscataway: IEEE Press, 2020: 1-8.
- [9] 赵廷, 李泽文, 邹彬, 等. 卫星时钟与网络时钟互备的广域时间同步方法[J]. 电力系统自动化, 2017, 41(14): 202-207.
ZHAO T, LI Z W, ZOU B, et al. Wide-area time synchronization method for mutual preparation of satellite clock and network clock[J]. Automation of Electric Power Systems, 2017, 41(14): 202-207.
- [10] 于佳亮, 施闯, 张东, 等. 广域精确授时技术及其应用[J]. 电信工程技术与标准化, 2020, 33(8): 82-87.
YU J L, SHI C, ZHANG D, et al. Wide-area precise timing technology and its application[J]. Telecom Engineering Technics and Standardization, 2020, 33(8): 82-87.
- [11] 李恩, 宋文忠, 罗青松, 等. 光网络高精度授时体系建设策略研究[J]. 光通信技术, 2019, 43(9): 58-62.
LI E, SONG W Z, LUO Q S, et al. Research on construction strategy of optical network high precision time service system[J]. Optical Communication Technology, 2019, 43(9): 58-62.
- [12] 朱瑾瑜, 程娜. 时间敏感网络时间同步特性综述[J]. 信息技术与政策, 2019(12): 70-73.
ZHU J Y, CHENG N. Overview of time-sensitive network time synchronization characteristics[J]. Information and Communications Technology and Policy, 2019(12): 70-73.
- [13] STEINER W. An evaluation of SMT-based schedule synthesis for time-triggered multi-hop networks[C]//Proceedings of 2010 31st IEEE Real-Time Systems Symposium. Piscataway: IEEE Press, 2010: 375-384.
- [14] DÜRR F, NAYAK N G. No-wait packet scheduling for IEEE Time-sensitive networks (TSN)[C]//Proceedings of Proceedings of the 24th International Conference on Real-Time Networks and Systems. New York: ACM Press, 2016: 203-212.
- [15] NAYAK N G, DÜRR F, ROTHERMEL K. Incremental flow scheduling and routing in time-sensitive software-defined networks[J]. IEEE Transactions on Industrial Informatics, 2018, 14(5):2066-2075.
- [16] 王敬超, 高先明, 黄玉栋, 等. 时间敏感网络的控制架构[J]. 北京邮电大学学报, 2021, 44(2): 95-101.
WANG J C, GAO X M, HUANG Y D, et al. Research of control framework in time-sensitive network[J]. Journal of Beijing University of Posts and Telecommunications, 2021, 44(2): 95-101.
- [17] CRACIUNAS S S, OLIVER R S, CHMELÍK M, et al. Scheduling real-time communication in IEEE 802.1Qbv time sensitive networks[C]//Proceedings of the 24th International Conference on Real-Time Networks and Systems. New York: ACM Press, 2016: 183-192.
- [18] POZO F, STEINER W, RODRIGUEZ-NAVAS G, et al. A decomposition approach for SMT-based schedule synthesis for time-triggered networks[C]//Proceedings of 2015 IEEE 20th Conference on Emerging Technologies & Factory Automation (ETFA). Piscataway: IEEE Press, 2015: 1-8.
- [19] SCHWEISSGUTH E, TIMMERMANN D, PARZYJEGLA H, et al. ILP-based routing and scheduling of multicast real-time traffic in time-sensitive networks[C]//Proceedings of 2020 IEEE 26th International Conference on Embedded and Real-Time Computing Systems and Applications (RTCSA). Piscataway: IEEE Press, 2020: 1-11.
- [20] ATALLAH A A, HAMAD G B, MOHAMED O A. Routing and scheduling of time-triggered traffic in time-sensitive networks[J]. IEEE Transactions on Industrial Informatics, 2020, 16(7): 4525-4534.
- [21] FALK J, DURR F, ROTHERMEL K. Exploring practical limitations of joint routing and scheduling for TSN with ILP[C]//Proceedings of 2018 IEEE 24th International Conference on Embedded and Real-Time Computing Systems and Applications (RTCSA). Piscataway: IEEE Press, 2018: 136-146.
- [22] YAN J L, QUAN W, JIANG X Y, et al. Injection time planning: making QCF practical in time-sensitive networking[C]//Proceedings of IEEE INFOCOM 2020 - IEEE Conference on Computer Communications. Piscataway: IEEE Press, 2020: 616-625.
- [23] KROLIKOWSKI J, MARTIN S, MEDAGLIANI P, et al. Joint routing and scheduling for large-scale deterministic IP networks[J]. Computer Communications, 2021, 165: 33-42.
- [24] IETF DetNet. Deterministic networking use cases[S]. IEEE, 2019.

[作者简介]



汪硕 (1991–), 男, 河南灵宝人, 博士, 北京邮电大学讲师, 主要研究方向为确定性网络、数据中心网络、软件定义网络、网络流量调度等。

黄玉栋 (1998–), 男, 重庆人, 北京邮电大学博士生, 主要研究方向为未来网络体系架构、确定性网络、软件定义网络等。

黄韬 (1980–), 男, 重庆人, 博士, 北京邮电大学教授, 主要研究方向为路由与交换、软件定义网络、内容分发网络等。

霍如 (1988–), 女, 黑龙江哈尔滨人, 博士, 北京工业大学讲师, 主要研究方向为计算机网络、信息中心网络、网络缓存策略与算法、工业互联网、标识解析技术等。

刘韵洁 (1943–), 男, 山东烟台人, 中国工程院院士, 主要研究方向为未来网络体系架构、网络融合与演进等。