

确定性网络研究综述

黄韬, 汪硕, 黄玉栋, 郑尧, 刘江, 刘韵洁

(北京邮电大学网络与交换国家重点实验室, 北京 100876)

摘 要: 现有的互联网面对激增的视频流量和工业机器应用, 存在着大量的拥塞崩溃、数据分组时延等问题, 而许多网络应用, 例如工业互联网、远程医疗、无人驾驶、VR 游戏等, 需要达到 1~10 ms 时延, 微秒级抖动, 但传统的网络只能将端到端的时延减少到几十毫秒。在这样的背景下, 对网络端到端时延的控制如何从“尽力而为”到“准时、准确”, 成为当前全球关注的热点领域。通过介绍确定性网络的应用场景与需求, 描述当前该领域的主要研究成果, 总结分析了该领域的研究发展趋势和核心问题, 期望对该领域的研究起到参考和帮助作用。

关键词: 确定性网络; 时间敏感网络; 5G 移动通信; 工业互联网; 服务质量; 超低时延

中图分类号: TP393

文献标识码: A

doi: 10.11959/j.issn.1000-436x.2019119

Survey of the deterministic network

HUANG Tao, WANG Shuo, HUANG Yudong, ZHENG Yao, LIU Jiang, LIU Yunjie

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

Abstract: The current Internet faces a huge increase in video traffic and industrial machine applications, causing a lot of congestion and packet delays. Besides, many network applications, such as industrial Internet, telemedicine, driverless, VR games, etc, require a latency from 1 to 10 milliseconds and jitters within microsecond. But traditional network can only reduce the end-to-end latency to tens of milliseconds. Under this background, how to transit from “best effort” to “punctuality, accuracy”, controlling the end-to-end delay has become a hot field of global network research. The application scenarios and requirements of the deterministic network were introduced, and the main research results of the current parties were described as well as the trend of development and core issues were summarized. It hopes to provide a reference and help for the research in this field.

Key words: deterministic network, time sensitive network, 5G mobile communication, industrial internet, quality of service, ultra-low latency

1 引言

以太网自 20 世纪 70 年代诞生以来, 由于其简单的网络连接机制、不断提高的带宽以及可扩展性和兼容性而被广泛使用, 根据全球移动数据流量预测报告^[1]显示, 到 2020 年全球 IP 网络接入设备将达 263 亿台, 其中工业和机器连接设备将达 122 亿

台, 相当于总连接设备的一半, 同时高清和超高清互联网视频流量将占全球互联网流量的 64%。

激增的视频流量和工业机器应用, 带来了大量的拥塞崩溃和数据分组时延。同时, 许多网络应用, 例如工业互联网中的数据上传和控制指令下发、远程机器人手术、无人驾驶、VR 游戏等, 需要将端到端时延控制在 1~10 ms, 将时延抖动控制在微秒

收稿日期: 2018-12-26; 修回日期: 2019-02-10

基金项目: 中国博士后科学基金资助项目 (No. 2018M641281); 国家自然科学基金资助项目 (No.61872401)

Foundation Items: China Postdoctoral Science Foundation Project (No.2018M641281), The National Natural Science Foundation of China (No.61872401)

级,但传统的网络只能将端到端的时延减少到几十毫秒。面对时延敏感性业务的迫切需求,如何从“尽力而为”到“准时、准确”地控制端到端的时延对IP网络提出了新的挑战。

虽然在工业领域,有几个扩展以太网提供了初步的确定性解决方案,例如PROFINET(process field net^[2]、EthernetCAT(ethernet control automation technology)^[3-4]、TTEthernet(time-triggered ethernet)^[5]、HaRTES(the hard real time switch architecture)^[6],但它们要么不能相互兼容,要么不能与标准以太网设备集成,很难满足工业控制系统的实时确定性要求。深入分析发现,因为以太网缺乏时钟同步机制、带宽预留等管理机制,数据分组优先级等过滤机制,从而无法为应用提供时延和抖动的服务质量(QoS, quality of service)保障。

目前,电子电气工程师学会(IEEE, Institute of Electrical and Electronics Engineers)和互联网工程任务组(IETF, Internet Engineering Task Force)已提出了新的确定性网络技术。IEEE 802.1工作组(WG, Work Group)致力于时间敏感网络(TSN, time sensitive network)的标准化,时间敏感网络是当前最为成熟的实现局域确定性网络的技术,通过IEEE 802.1AS时钟同步、IEEE 802.1Qcc流预留、IEEE 802.1Qch循环排队等技术保证物理层和链路层的确定性时延;IETF的DetNet(Deterministic Network)工作组专注于网络层(L3)及更高层次的广域确定性网络技术。此外,5G标准化工作组已将目标定为总时延1 ms或更低,应用层的开放通信平台OPC UA(object

linking and embedding for process control unified architecture)也在积极寻求与TSN的结合,确定性网络的发展充满了机遇与挑战。

2 确定性网络技术应用场景与需求

确定性网络技术已成为当今学术界和产业界研究和关注的热点之一,不仅在学术领域有广阔的研究空间,而且在产业化方面也有巨大的市场前景,因此,研究其在特定场景下的具体需求对确定性技术的发展落地具有重要意义。

2.1 工业互联网

工业互联网是互联网与工业系统全方位深度融合所形成的产业和应用生态,是工业智能化发展的关键综合信息基础设施。如图1所示,工业互联网网络连接架构分为下层网络互联和上层数据互通两部分,其中,网络互联又包括工厂内网、工厂外网。以下将从工厂内网、工厂外网和数据互通三方面阐述工业互联网对确定性时延和抖动的要求。

1) 工厂内网

工厂内网用于连接工厂内的各种要素,包括人员(如生产人员、设计人员、外部人员)、机器(如装备、办公设备)、材料(如原材料、在制品、制成品)、环境(如仪表、监测设备)等。工厂内网与企业数据中心及应用服务器互联,支撑工厂内的业务应用。

如图2所示^[7],当前,工厂内网呈现“两层三级”的结构,“两层”是指存在“工厂IT(information technology)网络”和“工厂OT(operational technology)网络”两层技术异构的网络;“三级”

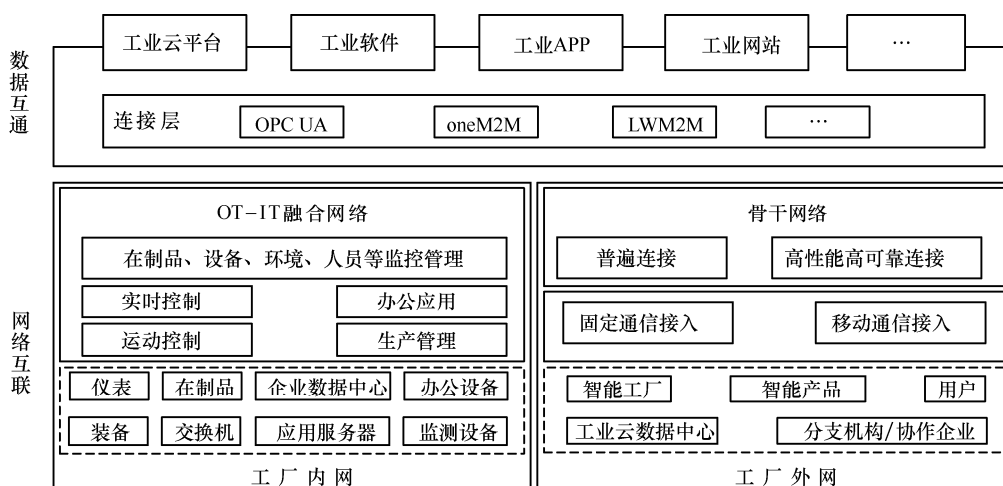


图1 工业互联网网络连接架构

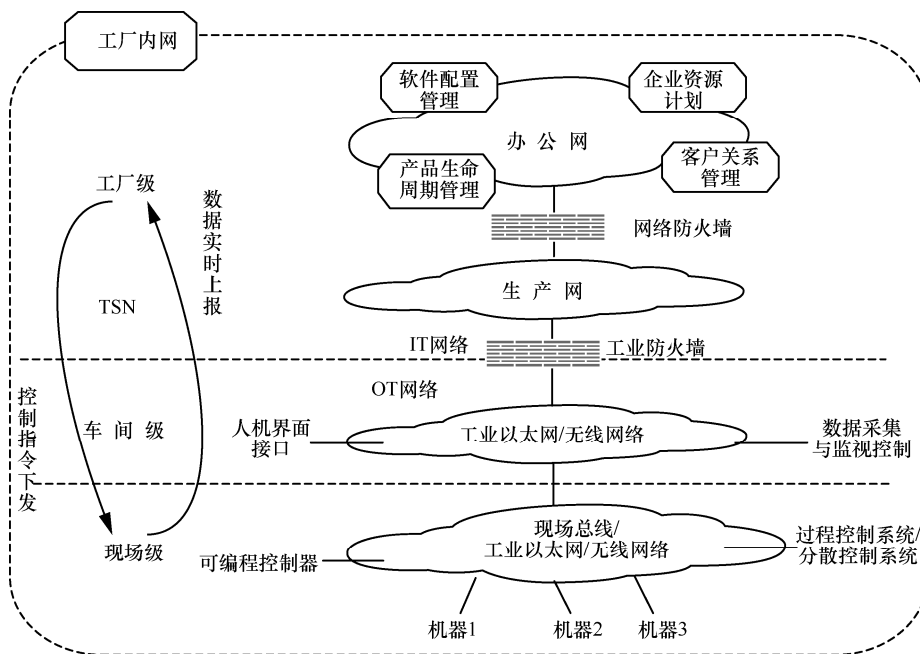


图2 工厂内网确定性时延要求

是指根据目前工厂管理层级的划分，网络也被分为“现场级”“车间级”“工厂级”这3个层次，每层之间的网络配置和管理策略相互独立。

其中，三级中的工厂级的IT管理运营系统对现场实时工艺过程数据和设备状态数据有着强烈需求，比如数据实时上报、控制指令下发等。如何实现现场级与工厂级之间高实时性、高可靠性数据通信，是目前工业网络领域普遍关注的焦点问题。比如钢铁厂、炼油厂和海上钻井平台等实施复杂的工业流程，数千个现场传感器向工厂控制中心报告温度、压力和油箱填充水平，中心以自动或人为干预的方式使用该信息来控制执行器、启动新的生产阶段、安排维护或触发警报，其中传感器、执行器和控制中心之间的通信需要1~10 ms确定性时延。

当前，满足该要求的通常做法是修改工业以太网协议或者在关键生产流程部署独立的专用以太网。然而，这类方式存在的互通性、扩展性和兼容性不够的问题，在从传统工厂控制网络升级到工业互联网的过程中日益明显，目前主流的解决方案是TSN时间敏感网络和6TiSCH(IPv6 over time slotted channel hopping) [8]。

TSCH(time slotted channel hopping)是2012年IEEE发布的802.15.4e MAC层的一种时隙信道跳频模式，该模式使用及时同步通信和信道跳变来消除信道衰落和干扰的影响。6TiSCH是IETF成立的工作组，

旨在研究IEEE 802.15.4e协议的TSCH模式下的IPv6连接，使其具有低时延、低抖动、低功耗和高可靠性的工业级确定性特性。

如图3所示，6TiSCH上层堆栈(IETF 6LoWPAN、RPL和CoAP) [9]具有最先进的低功耗无线网络通信技术。6top为操作子层，用以绑定IETF上层和IEEE 802.15.4e TSCH，它通过创建标准方法来构建和维护调度，执行TSCH配置和控制过程，并定义了最小的6TiSCH配置，以便实现所有设备间的互操作性。

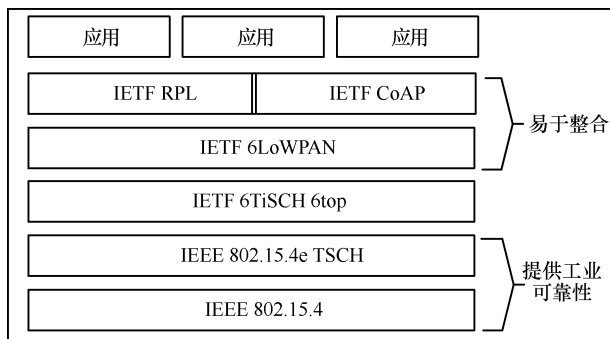


图3 6TiSCH网络架构

目前，6TiSCH工作组在6top操作子层、集中式调度机制、6TiSCH远程管理接口设计以及实现骨干网确定性方面的标准尚未完成。

发电厂目前主要采用6TiSCH协议以确保实时通信。智能电网应用涵盖了发电、输电、配电和客

户驻地领域。在智能电网中,实时信息和可靠的电力输送使电力系统更加智能,并克服了诸如组件故障、容量限制和影响电力输送的灾难等挑战。例如,实时监控和数据收集要求通过 IEC 61850 协议传输的 6 种指定消息,具有以下 3 组可接受的传输时间:快速消息、原始数据和特殊任务需要小于 10 ms,中等消息需要小于 100 ms,而慢速消息和非关键命令需要小于 500 ms^[10]。具有灵活调度的 TSCH MAC 层可以满足 IEC 61850 消息传输时间要求。

图 4 展示了在发电、配电和家用电表网络中利用 6TiSCH 概念的智能电网架构,其中实时数据采集和智能电子设备的控制(如自动化本地站的断路器控制器、自动化变电站的电压调节器和其他自动化设备)需要满足极低的时延要求。

2) 工厂外网

工厂外网用于连接智能工厂、分支机构、上下游协作企业、工业云数据中心、智能产品与用户等主体。智能工厂内的数据中心/应用服务器通过工厂外网与工厂外的工业云数据中心互联。

工厂外网的需求主要包括 4 个方面:工业实体的互联网接入需求、跨区域之间的互联与隔离需求、工业网络与混合云互联的需求、工业互联网对广域承载网络的差异化需求(QoS、安全/保护等),当前,满足以上需求并广泛使用的运营商专线业务主要包括 MPLS(multi-protocol label switching)VPN 专线和基于 OTN(optical transport network)的光网专线。

MPLS VPN 虚拟专网为用户在公共 MPLS 网

络上构建企业的虚拟专网,满足其不同城市(国际、国内)分支机构间安全、快速、可靠的工业化通信需求,并能够支持办公、数据、语音、图像等高质量、高可靠性要求的多媒体业务。基于 OTN 的智能光网络是大颗粒宽带业务传送的理想解决方案,如果企业外部专网的主要调度颗粒达到 Gbit/s 量级,可以考虑优先采用 OTN 技术进行网络构建。

随着智能工厂的发展,各机器之间通过交换彼此的信息,或者使用超级控制器来描绘网络拓扑结构以及各种状态信息。工厂内网需要建立在 IP 网络之上,并且满足确定性时延和抖动以及远程控制的要求。而目前的工业控制网络主要局限在局域网的范围,不能满足跨局域网、多实时边缘网络互连的确定性业务传输需求,IETF 的 DetNet 工作组目前正在解决这个问题。

3) 数据互通

数据互通是指建立数据的结构和规范,使传递的数据能被有效地理解和利用,对应协议栈的 TCP/UDP 层到应用层。

如图 5 所示,OPC UA^[11-14]是一个开放通信平台,可以实现从 PLC(programmable logic controller)到云、PLC 之间以及 PLC 到上层应用的数据互通,已被选为工业 4.0 的参考标准。其中,上层应用包括 ERP(enterprise resource planning)、MES(manufacturing execution system)、SCADA(supervisory control and data acquisition),接口为 HMI。对于信息交换,OPC UA 提供 2 种通信机制:第一种是客户端-服务器模型,客户端通过定义的服务访问服务器信息;第二种

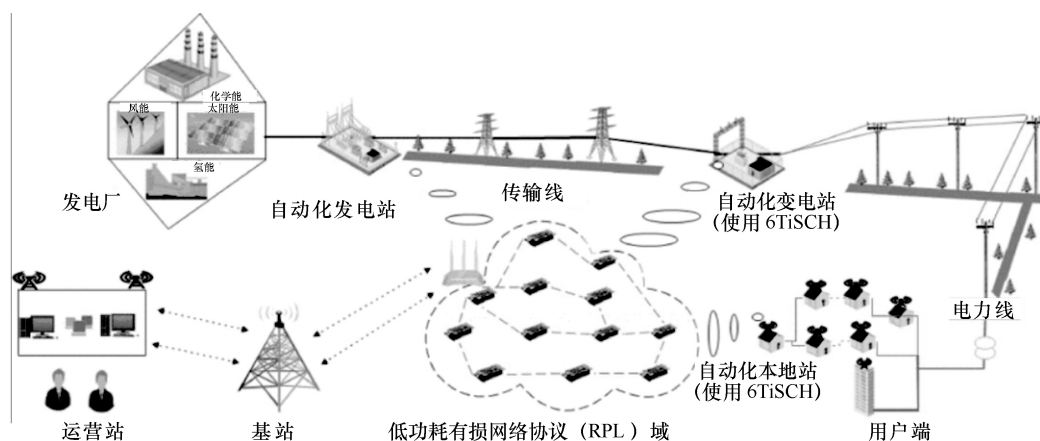


图 4 6TiSCH 在智能电网架构中的使用

通信方法是 OPC UA PubSub, 即发布者-订阅者模型, 这种方法允许数千个传感器和云之间的多播通信, 以及机器之间的协调通信。

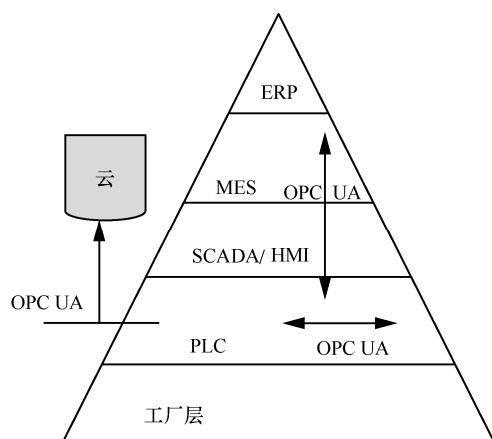


图5 工厂自动化中的对象连接与嵌入处理控制统一架构

为提高数据互通的实时性和可靠性, 向现场设备端延伸, OPC UA 正积极与 IEEE TSN 相结合, 支持用于时间同步的 IEEE 802.1 AS-Rev 和用于调度的 IEEE 802.1 Qbv 协议, 实现不同供应商的工业控制器之间的开放式数据交换, 以及实时的机器到机器通信。OPC UA PubSub 标准目前还处于试运行中, 但已被使用 Profinet、EtherCat、PowerLink 等架构的大公司采用。

2.2 5G 移动通信

2015 年 6 月 24 日, 国际电信联盟 (ITU) 公布 5G 技术的正式名称为 IMT-2020。IMT-2020 是第五代移动电话行动通信标准, 传输速度是 4G 网络的 40 倍, 且对时延有非常高的要求。如表 1 所示^[15], 4.9G 对应于 LTE 向 5G 的优化工作, 实现了减少时延 90% 以上, 并且 5G 标准化工作组已将目标定为总时延为 1 ms 或更低^[15]。各种标准化组织为 5G 的发展做出了贡献, 包括 IEEE、IETF 及第三代合作伙伴计划 (3GPP, 3rd Generation Partnership Project) 和欧洲电信标准协会 (ETSI, European Telecommunications Standards Institute), 相关标准将在 2020

年制定完成。

1) 网络架构

如图 6 所示, 整个 5G 系统可以分为无线接入段、前传段, 以及到核心网络的回程段。无线接入段负责将无线设备连接到基站, 前传段将基站连接到云无线接入网, 而回程段将云无线接入网连接到核心网络。核心网络连接整个互联网, 包括数据中心, 为设备提供端到端的服务。

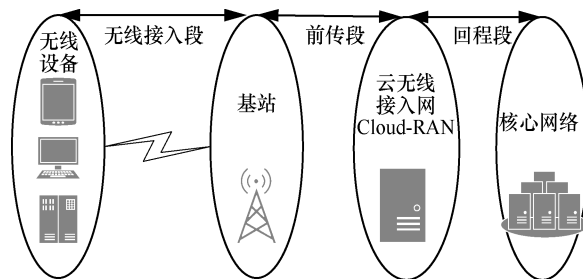


图6 5G网络架构

2) 网络应用

低时延 5G 的应用包括远程医疗、自动驾驶、环境监测、娱乐和工业自动化等。远程机器人手术需要保证 1~10 ms 的时延保证, 自动驾驶需要高速率以及 10 ms 以内的低时延以确保快速响应不断变化的道路状况, 还有用于游戏和娱乐的增强现实 (AR, augmented reality) 和虚拟现实 (VR, virtual reality) 需要高速率的视频传输和极低时延, 以避免视频和音频中的抖动。预计未来 5G 将不断发展以支持超低时延的端到端连接, 如图 7 所示, 现有的 3 种 5G 切片服务方案包括 eMBB (增强移动宽带)、URLLC (超可靠低时延通信) 和 mMTC (大规模机器类型通信)^[16]。

eMBB 侧重于高数据速率服务, 如高清视频、虚拟现实、增强现实和固定移动融合。

URLLC 侧重于对时延敏感的服务, 如自动驾驶车辆、远程手术或无人机控制。

mMTC 专注于对连接密度有很高要求的服务, 如智慧城市和智能农业。

表 1

3G、4G、4.9G 和 5G 的时延比较

标准	时延/ms							
	下行	上行	帧对齐	调度	手机处理	基站处理	核心网处理	目标总时延
3G	2	2	2	1.3	8	3	2	20
4G	1	1	1	0~18	4	2	1	10~28
4.9G	0.14	0.14	0.14	预调度	0.5	0.5	0.5	1.5
5G	0.125	0.125	0.125	预调度	0.25	0.25	0.1	1

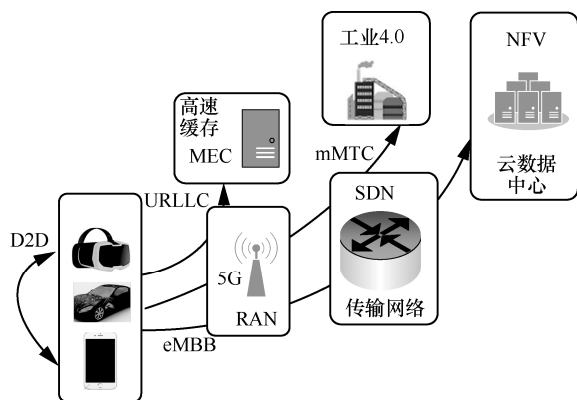


图7 5G中网络切片的应用图

5G 承载网络可以使用确定性网络来提供跨切片和切片内的传输。例如有 2 个切片 A 和 B，确定性网络通过 URLLC-A 和 URLLC-B 进行传输服务。URLLC-A 和 URLLC-B 具有单独的带宽保留，它们在不同的逻辑网络中，没有资源冲突，可以保证带宽和时延。

2.3 智慧建筑

智慧建筑的自动化系统（BAS, building automation system）^[17]可以管理建筑的设备和传感器，以改善居民的舒适度，减少能源消耗，并探测紧急情况，比如定期测量房间的温度、湿度，远程控制门和灯的开关，对设备的异常状态报警等。

1) 网络架构

如图 8 所示，BAS 中通常有两层网络，上层为管理网络，下层为现场网络。楼宇管理服务器（BMS, building management server）和人机接口（HMI, human machine interface）通过管理网络连接到本地控制器（LC, local controller），本地控制器通过现场网络连接到设备。管理网络中使用基于 IP 的通信协议，现场网络使用非 IP 的现场协议。在现场网络中有各种物理接口，如 RS232C 和 RS485，它们具有特定的时序要求。因此如果现场网络被以太网或无线网络替换，这种替换网络必须支持确定性流。

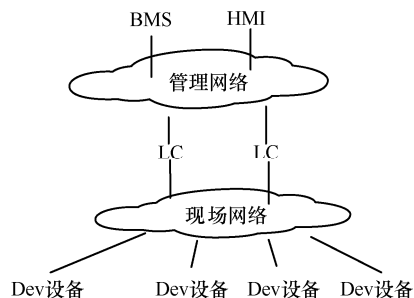


图8 楼宇自动化系统（BAS）架构

2) 网络应用

在环境监测中，BAS 系统中的 BMS 以 100 ms 的最大测量间隔轮询每个 LC，然后执行指定的操作。每个 LC 需要在一次间隔中测量几百个传感器，所以需要极低的测量时延，其可用性预计为 99.999%。

在火灾探测中，当发现火灾时，BAS 中的 BMS 须关闭空气调节系统（HVAC, heating ventilation and air conditioning）、关闭火灾百叶窗、打开消防喷淋头、发出警报等。在这一过程中，BMS 需要管理每个 LC 的大约 10 个传感器，需要实现 10~50 ms 的测量间隔、10 ms 以内的通信时延，以及 99.999 9% 可用性。

在反馈控制中，BAS 以各种方式利用反馈控制，其中最需要确定的是控制直流电机，这需要极短的反馈间隔（1~5 ms）、极低的通信时延（10 ms）和抖动（小于 1 ms），反馈间隔取决于设备特征和目标控制质量值。通常每个 LC 约有 10 个这样的设备，且可用性需达到 99.999 9%。

2.4 4K/8K/AR/VR 音视频

4K/8K/AR/VR 音视频的应用场景包括：音视频制作、直播、广播、电影院、现场音乐会、大型场所（机场、体育场馆、教堂、主题公园）的公共广播媒体和应急系统。这个行业正在从点对点的硬件互连转向无线互联，从而降低成本、提高灵活性。

1) 网络架构

为提供多媒体流所需的 QoS 保障，2011 年 IEEE 发布了 IEEE 802.1 AVB^[18]（audio/ video bridging）标准，其架构如图 9 所示。它在传统以太网的基础上，通过保障带宽（IEEE 802.1Qat）、限制时延（IEEE 802.1Qav）和精确时钟同步（IEEE 802.1AS）^[19-20]这 3 个方面的具体协议，实现在二层局域网创建确定性流。



图9 音视频桥接网络架构

2) 网络应用

① 实时音视频。文件传输一般采用差错重传保证服务质量。实时音视频不同于常见的文件传输,因为发现丢失或损坏的数据分组时再执行重传为时已晚,虽然现有的缓冲机制可用于提供足够的时延以允许一个或多个时间重试,但这不是真正有效的解决方案,需要保证带宽来消除拥塞,使用冗余路径来提供更高的可靠性。

② 同步流播放。Pro AV (professional audio video) 是在拍摄时对音频和视频进行时间同步,然后通过播放系统的不同路径进行传输。为保证接收端音频与视频一致,一个音频/视频同步的常见容差是一个 NTSC (national television standards committee) 视频帧 (约 33 ms),同时为保持观众对唇部同步的感知,时延需要在一定的合理容差范围内 (比如 10%) 保持一致。

以美国 ESPN 为例,2014 年 6 月,其耗资 1.25 亿美元建造了一个 1.8 万平方米,拥有最先进广播和后期制作技术的广播工作室名为 DC2。它具有 2 个 Evertz EXE 光纤路由器,吞吐量为 46 Tbit/s,可以在工厂的 1 770 km 光纤上同时处理超过 60 000 个信号。在 DC2 控制台核心内部,音频可嵌入视频或独立信号,可立即提供同步流给任何配有 AVB 接口的 Lawo mc2-56 音频控制台。AVB 基础设施取代了 ESPN 之前的全 MADI 音频传输系统。

③ 消除回声。如图 10 所示,实线表示从有人对着麦克风说话到声音从扬声器出现的网络时延,包括模数转换时延、传输时延、处理时延和数模转换时延。虚线表示声音播放总时延,如果总时延超过 10~15ms,将会产生声学回声,则扩声系统将无法使用,且总时延界限包括信号的所有路径,而不仅仅是网络,所以网络时延必须显著小于 15 ms。

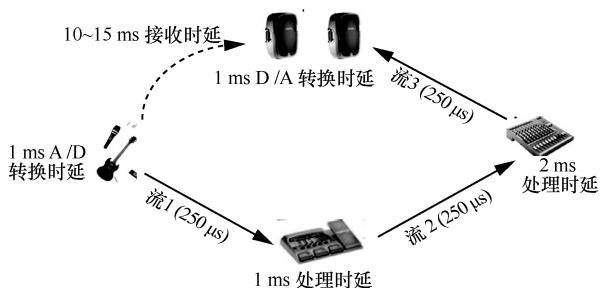


图 10 现场 Live 端到端时延模型

目前, Avid^[20]的 S6L 系列产品已支持以太网音

视频桥接端口,提供低时延的统一现场声音平台。

2.5 需求小结

本节所述确定性网络技术的应用场景有如下几点共性的需求。

- 1) 需要有界的且极低的时延、抖动和分组丢失率,保证端到端交付。
- 2) 需要标准化的数据流信息模型。
- 3) 需要集中式网络配置和控制系统。确定性网络的大小没有限制,但局限于集中管理和明确的网络,需排除互联网这样的无限制的分散网络。
- 4) 基于 IPv4 对以太网进行扩展,以开放标准统一多个专有确定性网络。
- 5) 确定性网络和以太网共存,未使用的确定性网络带宽可用于以太网。
- 6) 需要高可靠性和可用性。

3 确定性网络主要研究成果

随着实时流量的增加和对多网融合的需求,传输网提出了对确定性路径传输的需求。ATM (asynchronous transfer mode) 为整合电信网,提出了映射底三层网络的整体传输方案,但是技术复杂,设备昂贵。传统以太网凭借“尽力而为”的简洁思想逐渐在传输网络中成为主流,但因此会导致不可控的路由路径、丢失分组率和传输时延。基于以太网成为主流的趋势,2005 年,IEEE 的 802.1 任务组成立 AVB 任务组,用于局域网时延敏感的音视频业务的传输。2012 年,AVB 任务组改名为 TSN 任务组,主要应用于各种支持低延时及基于时间同步数据传输的以太网协议。TSN 与 OP-CUA 等相关联盟积极协作,促进多家标准的统一。IETF 在 2015 年 10 月成立了 DetNet 工作组,其专注于在第二层桥接和第三层路由段上操作的确定性数据路径,目标在于将确定性网络通过 IP/MPLS 等技术扩展到广域网上。2016 年 5 月,光联网论坛 (OIF, optical internetworking forum) 定义了 FlexE 技术的接口帧结构,实现带宽的捆绑、通道化,为 5G 切片网络中的确定性路径提供了保证。本节按照 IOS 五层模型的顺序,由下向上介绍确定性网络的关键技术。表 2 对比了 3 种关键技术现有成果的特点。

3.1 FlexE

3.1.1 概述

随着云计算、视频以及 5G 移动通信等业务的兴起,人们对 IP 网络的诉求从以带宽为主逐渐转移

表 2 关键技术现有成果对比

技术	应用范围	应用场景	是否支持 SDN	主要功能
FlexE	物理层与数据链路层之间	5G 网络的云服务、网络切片, AR/ V R/ 超高清视频等	是	网络切片, 子端口隔离
AVB/TSN	数据链路层	实时音频/视频, 工业自动化与控制系统 (IACS) 应用等	是	时钟同步, 局域网范围有界的时延抖动和丢失分组
Det/Net	网络层	实时音频/视频, 工业自动化与控制系统 (IACS) 应用等	是	网络层的操作维护管理, 端到端的确定传输

到业务体验、服务质量和组网效率上。为满足上述需求, 2011 年 1 月, OIF 成立灵活以太网研究小组, 2015 年 7 月发布草案, 2016 年 3 月发布了 FlexE 的 1.0 标准内容 (OIF-FLEXE-01.0), 2017 年第一季度 FlexE For IP/MPLS 标准正式在 BBF (broad band forum) 立项, 该标准对二层、三层业务基于 FlexE 接口的应用模型加以定义, 将 FlexE 接口扩展至 IP/MPLS 网络。在 OIF 2018 年第三季度会议上, FlexE 2.0 标准正式发布, 包括帧格式、时间同步方案等。FlexE 也称为灵活以太网, 是由 OIF 发布的通信协议, 在以太网 L2 (MAC, media access control) /L1 (PHY, physical layer) 之间的中间层增加了 FlexE Shim 层, 它通过时分复用分发机制, 将多个 client 接口的数据按照时隙方式调度并分发至多个不同的子通道, 使网络即具备类似于时分复用 (TDM, time division multiplex) 的独占时隙、隔离性好的特性, 又具备以太网统计复用、网络效率高的特性。FlexE 在以太网技术的基础上实现了业务速率和物理通道速率的解耦, 客户业务不一定在一个物理通道上传递, 还可能由多个物理通道捆绑形成的一个虚拟的逻辑通道传递。图 11 展示了 FlexE 数据传输的逻辑结构。在网元节点中, 配置了 3 个 FlexE 隧道, 其中隧道 1 和隧道 2 从左

侧的 FlexE A 组交叉传输到了右侧的 FlexE B 组, 隧道 3 交叉穿通到了右侧的 FlexE C 组。FlexE 端到端隧道通过 FlexE client 交叉实现, FlexE shim 通过解映射恢复出各 FlexE client 的 66 bit 码块流, 根据 FlexE 交叉单元配置的连接关系, 输出到对应出向的 FlexE client 单元, 通过其 FlexE shim 映射到 FlexE group 发送出去, 从而完成整个隧道的连通。

3.1.2 关键技术/标准

FlexE 技术可以实现 3 种应用模式: 链路捆绑模式、子速率模式和通道化模式^[21]。链路捆绑模式是将多个物理通道捆绑起来, 形成一个大的逻辑通道, 实现大流量的业务传输。子速率模式是指单条客户业务速率小于一条物理通道速率时, 将多条客户速率汇聚起来共享一条物理通道, 提高物理通道的带宽利用率。通道化模式是客户业务在多条物理通道上的多个时隙传递, 客户业务分布在多条不同物理通道的多条时隙上, 多个客户共享多条物理通道。实现这些应用的 FlexE 的关键技术包括实现网络切片的 FlexE shim 层结构、实现 FlexE 端到端传输的交叉传送、监控端到端传输的 OAM (operation administration and maintenance) 机制和提供可靠性的隧道保护技术。

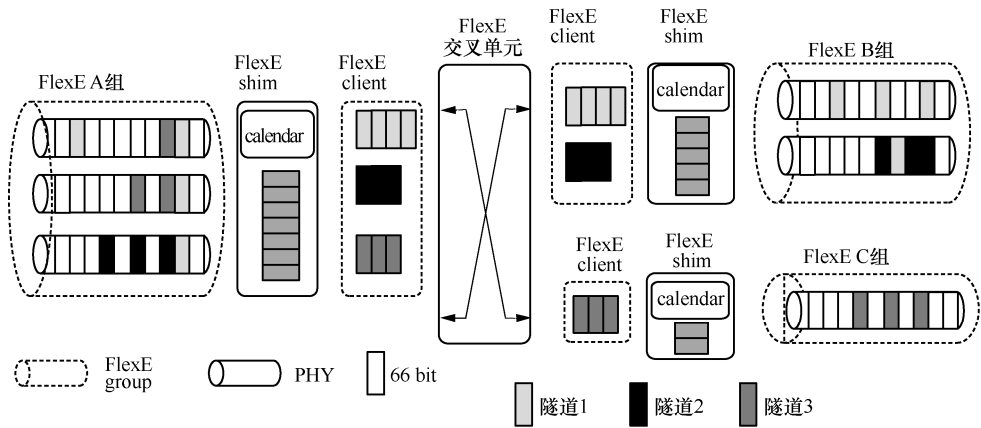


图 11 灵活以太网技术的数据传输逻辑结构

1) FlexE shim 层

FlexE 协议定义了一个时分复用的 FlexE shim 层。FlexE shim 层通过多个绑定的 PHY 来承载各种 IEEE 定义的以太网业务 (FlexE client)。FlexE shim 层可以支持多种以太网 MAC 分组, 包括大于或小于单个物理 PHY 速率的以太网分组。在承载客户业务时, 先对以太网分组进行 64/66 编码, 然后通过插入和删除空闲块实现速率适配, 将业务的时隙配置插入 master calendar 中, master calendar 将所有时隙分配成多个成员 (即 sub calendar), 添加 FlexE 开销, 扰码后经过 PMA (physical medium attachment)、PMD (physical medium dependent) 发送出去。FlexE 协议定义每个物理成员 PHY (注: 标准为 100 GE) 上传递一个 sub calendar, sub calendar 循环分配时隙编号来划分 66 bit 码块顺序, 同一编号的码块在逻辑上组成一个独立的时隙, 在 FlexE shim 中作为一个独立物理带宽资源单元分配使用。

2) FlexE 业务的交叉传送

端到端 FlexE tunnel 传送技术是构成 L1 层承载网络的基础^[21]。端到端 FlexE 隧道实现的核心思路是 FlexE client 能直接在 L1 层从一个 FlexE 组交叉到另一个 FlexE 组承载, 而不是上传 MAC 定帧后再进行分组交换。比如, 业务码流从一个 FlexE 组的 PHY 送到对应的 FlexE 组 master calendar, 按配置规则提取对应业务码块流后根据系统交叉配置送到交叉单元, 直接输出映射到另一个 FlexE group 的 master calendar 对应的时隙通道, 通过 PHY 转发往下一个节点。

3) OAM 机制

采用 FlexE tunnel 隧道技术的承载网络实现客户业务的端到端传输时, 需实现增加 OAM 开销, 以实现传输管道的端到端监控。在承载网络分层结构中, FlexE 技术涉及两层网络: FlexE 通道层和 FlexE 段层。

FlexE 通道层位于 FlexE 客户数据层和 FlexE 段层之间, 实现客户数据的接入/恢复、增加/删除 OAM 信息、数据流的交叉连接, 以及通道保护的功能。FlexE 通道层形成的端到端连接管道称之为 FlexE 隧道, 它是 FlexE 网络中传递的一条逻辑承载管道, 客户业务从源节点映射到 FlexE shim, 经过逻辑承载管道后, 再从 shim 中解映射到目的节点。FlexE 通道层 OAM 信息需要进行扩展实现,

在客户业务复用进入 FlexE shim 层前, 在客户业务流 (由 66 bit 码块组成的 TDM 码流) 中按某种固定周期插入 OAM 信息块。

FlexE 段层位于 FlexE 通道层和物理层之间。在 FlexE 段层中, 实现接入数据流的速度适配、数据流在 FlexE shim 上映射与解映射、FlexE 帧开销的插入与提取的功能。FlexE 段层的 OAM 信息来自标准定义的 FlexE 开销帧的内容, 通过开销帧头、复帧帧头信息可以提供等效 CC (continuity check) /CV (connectivity verification) 检测, 开销帧中 RPF (remote PHY fault) 信息可以提供远端成员缺陷指示 RDI (remote defect indication), 通过 PHY map、client calendar A/B 等字段来交互链路带宽以及相关时隙配置业务类型等。

4) 隧道保护

在 FlexE 通道层提供保护功能, 提高客户业务在 FlexE 隧道中传输的可靠性。保护方式分为 “1+1” 保护和 “1:1” 保护^[21]。当客户业务在一条隧道中出现故障时, 快速将客户业务切换到另外一条隧道中进行传输。

当客户业务正常工作时, 在 “1+1” 保护中, 客户业务可以同时 2 条隧道中传输, 在目的点同时检测 2 条隧道的业务服务质量状况, 从服务质量高的隧道中接收客户业务, 网络承载客户业务的带宽利用率只有 50%。在 “1:1” 保护中, 有 2 条承载通道隧道: 主通道隧道和备通道隧道。在正常工作时, 客户业务在主通道隧道传输, 备通道隧道可以传输低优先级客户业务。当主通道隧道出现故障时, 发送端和接收端协商并决策, 将客户从主通道隧道切换到备通道隧道中传输。由于 1:1 保护模式在正常工作状态下备用通道隧道可以传递其他低优先级客户业务, 网络承载客户业务的带宽利用率可以达到 100%。

3.1.3 趋势

网络切片技术可以让运营商在一个硬件基础设施中切分出多个虚拟的端到端网络, 每个网络切片在设备、接入网、承载网及核心网方面实现逻辑隔离, 适配各种类型服务并满足用户的不同需求。对每一个网络切片而言, 网络带宽、服务质量、安全性等专属资源都可以得到充分保证。采用 FlexE 技术的网络具有弹性带宽、灵活分配的硬管道, 可以实现业务的物理隔离和可靠的服务质量, 天然地实现了网络切片功能。FlexE 技术的物理管道捆绑、

子速率、通道化的应用模式可以承载各类速率需求的客户业务,提高了网络承载带宽的利用率,降低了网络设备的成本,逐步完善的 OAM 功能满足网络维护管理需要,这些优势很好地满足了 5G 承载网络的技术需求。

3.2 AVB/TSN

3.2.1 概述

AVB (audio video bridging), 即音视频桥接技术,是 IEEE 802.1 任务组 (TG, task group) 在 2005 年基于以太网架构制定的一套用于实时音视频的二层传输协议集^[22-23]。由于传统以太网是基于“尽力而为”的思想设计,在数据传输时会出现分组丢失和不确定性的时延,无法满足确定性网络对确定性端到端传输路径与时延的需要。AVB 标准通过时钟同步、资源预留和流量整形等技术有效降低了音视频数据在以太网传输中的最差时延,同时保持 100% 向后兼容传统以太网。

随着对确定性网络需求的增加,AVB TG2012 年更名为 TSN (time sensitive network) TG,即时间敏感网络任务组。与 AVB 相比,TSN 对 AVB 已有的协议进行了补充,应用范围更加广泛,TSN 主要有 4 个应用方向:专业音视频 (Pro AV)、汽车控制领域、商用电子领域和需要实时反馈的工业领域。TSN 还可以用于支持大数据的服务器之间的数据传输。

TSN 可以通过在以太网中的部署支持实时的 IACS (industrial automation and control system) 应用。IEEE 802.1 TSN TG 基于桥接局域网扩展了传

统以太网数据链路层的标准和协议,保证数据具有确定的低时延和抖动,满足数据传输的 QoS 要求,同时提供了足够的措施来实现网络中的端到端通信的安全性,这些标准还在修订并不断更新,新的标准会覆盖取代旧有标准,图 12 展示了 AVB/TSN 标准的发展历程。

3.2.2 关键技术

为了满足部分流量对确定性时延的要求,首先,要提供精确的时间同步机制;其次,要提供确定的传输路径。AVB/TSN 的思路是:首先将网络中需求不同的流量^[24]分成不同的优先级流,将有确定性需求的流量与其余流量区分开;然后以类似时分复用的思想,通过不同的流量整形机制为高优先级流量提供确定的传输时隙,以保证时间敏感流量有一条确定的传输路径。图 13 展示了 TSN 技术的架构。为了在两层网络中实现确定性网络,TSN 标准^[25]提供了精确的网络时间同步机制,调度不同优先级流量的网络管理机制,保证确定性时延的 Qos 机制和配置以上标准的执行机制。接下来,分别就这 4 个方面进行介绍。

1) 时间同步机制

确定性网络的一个基本要求是提供精准的网络时间同步机制。在 TSN 协议中,IEEE 802.1AS^[26-27]和更新的修订版本 IEEE 802.1AS-REV 可以实现亚微秒级的时间同步。IEEE 802.1AS 采用 IEEE 1588-2008 (1588v2)^[28]中的通用精确时间配置协议 (gPTP, generic precision time protocol)。gPTP 协议通过 BMCA (best master clock

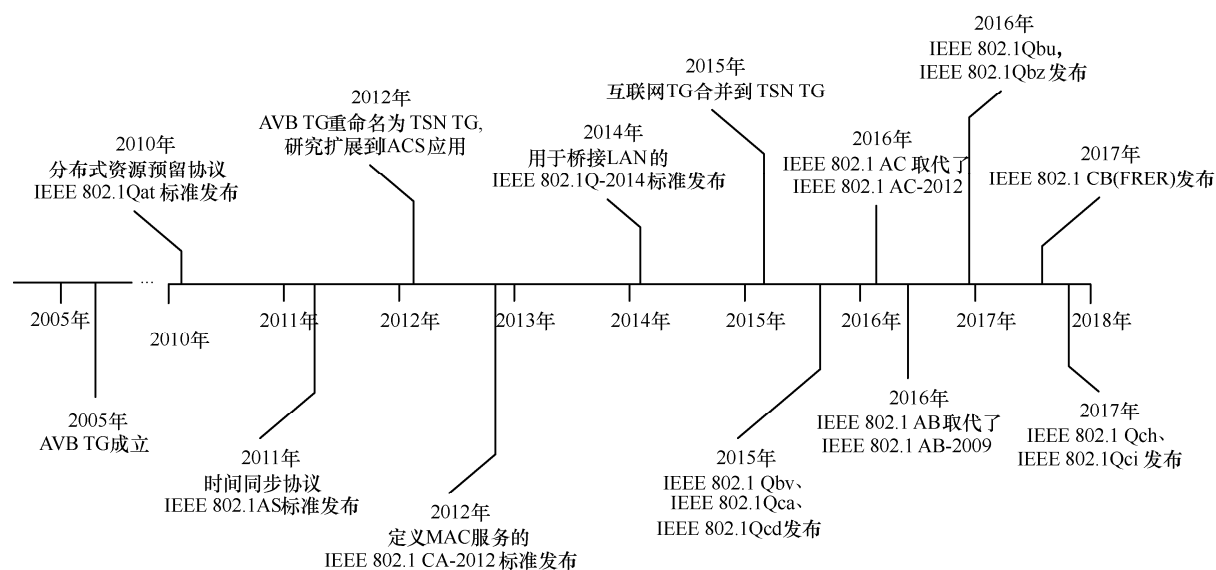


图 12 AVB/TSN 标准的发展历程

algorithm)^[28]建立主从结构形成 gPTP 域, 然后选出最精确的时钟源 GM (grand master) 时钟。在 gPTP 域内, 主时钟和从时钟之间不断传递时间信息, 并将时间与 GM 时钟进行同步^[29]。IEEE802.1 AS-REV 增加了在多个时域进行时间同步的功能, 既能在某域内 GM 时钟发生故障时实现快速切换到其他域的功能, 又能提高时间测量精度。

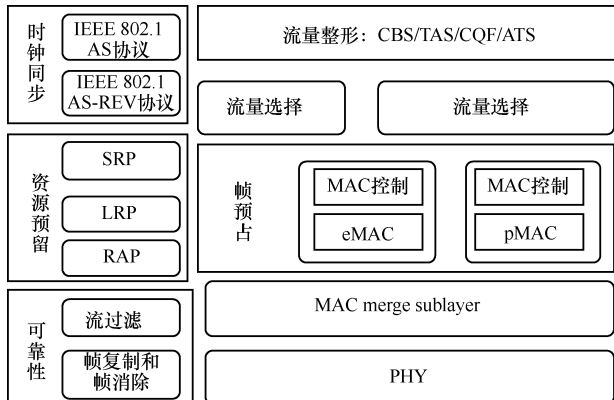


图 13 TSN 技术架构

2) 资源预留机制和准入框架

为了增强 TSN 网络的可扩展性, TSN 可以使用 UML 语言配置, IEEE 802.1 Qcp 协议标准化了 TSN 的 YANG 模型^[30-31], 为周期性监控和报告以及配置 802.1 网桥和组件提供框架, 比如配置媒体访问控制 (MAC) 网桥, 双端口 MAC 中继, 虚拟局域网 (VLAN, virtual local area network) 网桥^[32]等。

为了实现对时间敏感流量的优先调度, 需要有针对性 TSN 流的资源预留和准入控制机制。TSN 协议中的 IEEE 802.1 Qat 流预留协议 (SRP, stream reservation protocol)^[33]、IEEE 802.1 Qcc 增强型 SRP 和集中管理协议^[34]、IEEE 802.1 CS 本地链路保留协议 (LRP, link-local reservation protocol)^[35]和 RAP 协议 (resource allocation protocol), 负责路径预留和带宽限制, 接下来详细介绍这些协议。

AVB TG 提出了目前已经合并到 IEEE 802.1Q 的 IEEE 802.1 Qat 流预留协议 (SRP, stream reservation protocol)^[33]。该协议基于 TSN 流的资源要求和当前网络可用资源规定了准入控制架构, 通过多址注册协议^[24] (MRP, multiple registration protocol), 使用 48 位扩展唯一标识符, 也称作 StreamID, 来识别和注册业务流, 为 AVB 流提供足够的资源预留。此外, IEEE 802.1Qat 规定了在全双工以太网

中保留网络资源和广告流的框架。

由于现有的 IEEE 802.1Qat 采用分布式的注册和预留方式, 注册请求的变更有可能使网络过载从而导致关键流量类的时延。因此, TSN TG 引入了 IEEE 802.1Qcc^[36]标准, 通过减小预留消息的大小和频率来改善现有 SRP, 使更新仅由链路状态或预留改变触发。此外, IEEE 802.1Qcc 提供了一套集中式的全局管理和控制网络的工具, 可通过远程管理协议 (如 NETCONF^[37]或 RESTCONF^[38]) 执行资源预留、调度和其他配置。IEEE 802.1Qcc^[36]仍支持完全分布式配置, 允许集中管理系统和分散的 Ad Hoc 系统共存。当与 IEEE 802.1 Qca 路径控制和保留机制和流量整形结合时, 该方案可以提供确定性的端到端时延和零拥塞丢失。

虽然 MRP^[39]提供了有效的注册流方法, 但它保存流状态信息的数据库限制在大约 1 500 B。随着更多业务流共存以及网络规模的增加, 数据库成比例地增加, SRP 和 MRP 由于注册流状态信息的数据库有限而无法扩展到具有实时性 IACS 应用的大型网络。TSN TG 引入了本地链路预留协议^[35], 在点对点链路的两端之间有效地复制 MRP 数据库, 并在网桥报告新的网络状态时逐步复制更改。LRP 还提供清除过程, 当此类数据库的源无响应或数据过期时, 删除复制的数据库。经过优化, LRP 可有效处理大约 1 MB 的数据库。LRP 与作为信令协议的 RAP 一起以分布式的方式支持可扩展的 TSN 网络的资源预留。

3) 确定性时延的 QoS 机制

为了实现确定时延, TSN 利用帧抢占和流量整形机制在以太网链路中实现确定的传输路径, 目前已经发布了几种流控制标准。比如 IEEE 802.1 Qav^[40-41]采用 CBS (credit-based shaper) 机制, IEEE 802.1 Qbv 采用 Tas (time-aware shaper) 机制^[40], IEEE 802.1 Qch 采用 CQF (cyclic queuing and forwarding) 机制, IEEE 802.1 Qcr 采用 ATS (asynchronous traffic shaping) 机制。流量整形机制通过为高优先级流量提供确定的传输时隙来提供确定的传输时延, 避免突发流量造成的重传和分组丢失的影响。

CBS 主要应用于 AVB 应用, 它利用信用这个指标将传输时间分为允许高优先级流量和普通优先级流量传输这 2 个时隙。CBS 结合 SRP 可以将每个网桥出现的时延限制在 250 ms, 但是 CBS 使网络平均时延增加。因此, TSN TG 提出了 TAS 机制,

配合帧抢占机制一起提供更好的 QoS, 这也是目前最常使用的流量整形机制。对于符合 IEEE 802.3 的以太网帧, 算入单个 IEEE 802.1Q VLAN 标记和帧间间隔的帧的总长度为 1 542 B, 其中包括前导码和帧起始定界符共 8 B, 以太网地址、以太网类型或长度和 CRC 共 18 B, VLAN 标记 4 B, 帧有效载荷 1 500 B, 帧间隔 12 B。与传统以太网帧相比, 增加了 4 B 的 IEEE 802.1Q VLAN 标记。

TAS 机制要求所有时间触发的窗口时间同步, 然后利用门控列表技术控制不同优先级的队列的传输或等待。为了减少在时隙转换时, 低优先级流量对高优先级流量的干扰, TSN 采用 IEEE 802.1Qbu 帧抢占^[42]机制, 利用保护频带为 TAS 确保传输信道对于下一个高优先级流量的传输是可用的, 同时显著减小保护频带的字节数。使用帧抢占技术前, 保护频带的持续时间必须与安全传输的最大帧尺寸一样长, 即 1 542 B 的传输时间; 加入帧抢占后, 保护频带可以减少到帧的最小尺寸, 即最小帧 64 B 和不能抢占的剩余长度 63 B 的和, 总共 127 B。帧预抢占造成的低优先级帧中断只发生在链路层, 在下一个网桥的接口处, 被中断的帧会被重新整合成为完整的帧。

TAS 虽然有效减小了传输时延, 但是配置复杂, 对网络节点的时间同步要求很高, CQF (cyclic queuing and forwarding), 也称蠕动整形器 (PS, peristaltic shaper), 可以通过同步入口和出口的队列操作来降低 TAS 配置的复杂性, 实现与网络拓扑无关的零拥塞丢失和有界时延, 但是会导致更高的时延, 对时间同步的要求也很高。

为了有效利用网络带宽, ATS (asynchronous traffic shaper) 基于 UBS (urgency-based scheduler)^[43-44], 通过在每一跳重塑 TSN 流, 提供不需要严格时间同步的确定性时延。总体来说, 这些提供确定性时延的流量控制机制的高效动态配置需要更多的研究和实践。

4) 配置 TSN 流量

TSN 流会根据应用需求在以太网报头中的 802.1Q VLAN 标记中的 PCP (priority code point) 和 VID (VLAN ID) 中定义流的不同优先级。

TSN 中有多个流管理标准, 包括 IEEE 802.1CB、IEEE 802.1Qca、IEEE 802.1Qci, 负责提供路径冗余、多路径选择及队列过滤。尽管标准以太网可以通过生成树协议提供冗余能力, 但

是对于实时 IACS 应用而言, 在发生故障时的收敛时间太长。因此, FRER (frame replication and elimination for reliability) 在不相交的路径上发送关键流量的重复副本, 用于主动实现无缝数据冗余, 代价是额外的带宽消耗。PCR (path control and reservation) 提供显式转发路径控制所需要的协议, 如预定义的保护路径、带宽预留、数据流冗余、流同步和流控制信息的控制参数的分配^[45], PCR 与 FRER、IEEE 802.1Qcc 结合使用时达到快速恢复、高效路径冗余和动态流量管理。PSFP (per-stream filtering and policing) 通过 StreamID 识别流, 执行相关策略, 负责管理控制并防止恶意流程恶化网络性能。

3.2.3 趋势

TSN 在二层网络通过时分复用的思想为高优先级流量提供了确定性网络需要的传输路径和传输时延, 但是会导致低优先级流量的时延增加, 一种方法是通过适当的接入控制, 利用统计复用可以为时延界限提供统计保证^[46]。如何协调传输时延的最大时延和平均时延是一个加速 TSN 应用部署的关键问题。除此之外, 如何部署 TSN 网络, 是采用分布式部署还是结合 SDN 等技术进行集中式部署, 以及如何互联多个封闭的 TSN 网络也是未来要考虑的关键问题。

3.3 DetNet

3.3.1 概述

确定性网络工作组 (DetNet WG, Deterministic Networking Working Group), 是由 IETF 在 2015 年 10 月成立的小组, DetNet 目标是在第二层桥接和第三层路由段上实现确定传输路径, 这些路径可以提供时延、丢失分组和抖动的最坏情况界限, 以此提供确定的时延。相比于 TSN, DetNet 的工作范围更加广泛, 通过 MPLS/IP 技术, 以期实现三层的确定性传输。

DetNet 在二层网络的确定性路径的实现主要依靠 TSN 标准实现。可以看到, 活跃在 TSN 标准制定组织中的成员也广泛活跃在 DetNet 的标准制定组织中。

DetNet 工作组尚未建立 IETF RFC, 但目前有多个 IETF 草案可供参考。本文接下来会介绍 DetNet 工作组提出的确定性网络的操作、管理和维护的标准化的整体架构, 其支持多跳路由的时间同步、管理、控制和安全操作, 以及各种形式

的动态网络配置和多路径转发。图 14 展示了 DetNet 的主要架构, 图中不同线型表示承载不同功能的抽象链路。

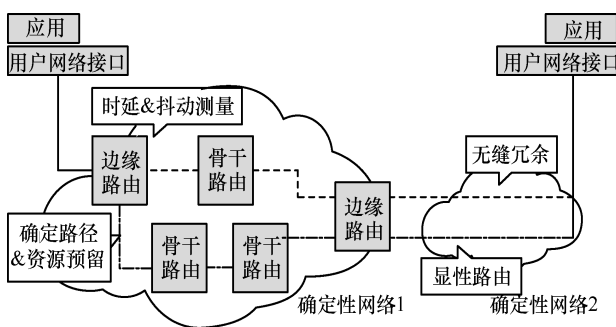


图 14 DetNet 的架构

3.3.2 关键技术

DetNet 在二层网络的确定性时延主要通过 TSN 机制实现, 本节不再赘述。如何在三层网络部署 DetNet 还在讨论阶段, 目前的几种备选方案是结合 UDP/TCP 服务层协议或结合基于 MPLS 的服务层协议实现三层的确定性网络。接下来从 DetNet 流定义、资源规划、流量工程和配置模型几个方面介绍 DetNet。

1) DetNet 流定义

DetNet 的时钟同步主要通过 TSN 机制实现, 它指定通过实现网络实体之间的亚微秒级时间同步和在程序包中嵌入执行时间字段^[47]来减少抖动。

DetNet 流按其 QoS 类别分类, 通过最大和最小端到端时延, 以及丢失分组概率要求来定义每个流的 QoS, 目前确定了 4 种主要的 DetNet 流类型^[47]。

DetNet 堆栈模型架构分为 DetNet 服务层以及 DetNet 传输层。DetNet 服务层是负责特定 DetNet 服务的层, 例如分组排序, 流复制/重复消除和分组编码, 而 DetNet 传输层负责可选地通过底层网络提供的路径为 DetNet 流提供拥塞保护^[42]。DetNet 可以有多个分层拓扑, 其中每个下层拓扑为更高层拓扑服务。DetNet 节点之间相互连接形成子网络, 这些子网络, 例如二层 TSN 网络或者点对点光传输网络 (OTN, optical transport network)^[47], 可以通过兼容服务支持 DetNet 流量。

为了低层流量能够更精确地针对高层流量类型来实施不同的排队、整形和转发策略, DetNet 模糊了网络层和链路层的界限, 二层网络可以通过流 ID 和 DetNet 控制字 (CW, control word) 识别

DetNet 流类型和对应属性相关的上层信息。为此, DetNet 流需要标准化跨层或异构网络的流属性映射, DetNet 考虑了 3 种主要的转发方法: IP 路由、MPLS 标签交换、以太网桥接。对于在异构网络的转发, 每个 DetNet 数据分组都附加或封装有多个流 ID (IP、MPLS 或以太网)。这使 DetNet 能够在 IP 和非 IP 网络之间进行路由和转发, 从而实现网络互操作性。

2) 资源规划

在资源管理方面, DetNet 有集中式和分布式这 2 种路径设置方式。集中路径设置类似于 IEEE TSN 的集中管理模型, 利用 PCE (path computation element) 和基于分组的 IP 或非 IP 网络的信息传播来实现全局网络的优化。分布式路径设置利用内部网关协议流量工程 (IGP-TE) 信令协议开发了类似于 IEEE 802.1Qat、IEEE 802.1Qca 和 MRP 信令协议的初始设计规范。

为了补充 DetNet 流量控制机制 (包括整形、调度和抢占), 每个节点 (或集中设置的中央控制器) 需要有和附近网络共享网络状态的能力^[47]。例如共享当前节点的资源使用状态、邻居节点及其关系的属性等。目前还没有规定如何实现这种能力, 但是这对于全局规划流量、实现确定性网络至关重要。

3) 流量工程

IETF 流量工程架构和信令工作组考虑将流量工程 (TE, traffic engineering) 架构用于分组和非分组网络^[48], 定义控制和管理 DetNet 流的关键概念、功能以及不同层面之间的关系, 使用户和操作人员可以动态地轻松控制、测量和管理流, 还引入了 QoS 参数的快速恢复和确定性边界。DetNet WG 采用类似于软件定义网络范例的方法, 为 DetNet 起草了一套 TE 架构, 与 IEEE TSN 的 IEEE 802.1Qcc 管理方案和集中式 SDN 方法具有相似之处。整个架构分为应用层面、控制层面和网络平面。在控制层面对 DetNet 流进行全局规划。

4) DetNet 配置模型

DetNet 能够在各种支持 DetNet 的网络实体之间实现无缝配置和重配置, 草案^[49]定义了 DetNet 分布式、集中式和混合式的配置模型及其相关属性, 还介绍了在集中配置模型中传递网络配置参数的 YANG 模型。

在分布式配置模型中, 控制信息通过 IGP 和

RSVP-TE 等协议来执行, 草案中没有详细介绍。集中式配置模型的控制信息通过 CUC (centralized user configuration) 和 CNC (central network controller) 配置, 集中式配置模型^[49]被定义了以下主要属性。

① DetNet 拓扑属性, 指定拓扑相关属性, 例如节点类型, 是否具有数据分组复制和消除功能 (PREF, packet replication and elimination function) 以及排队管理算法。

② DetNet 路径配置属性, 指定网络路径相关属性, 例如约束条件 (所需的最小/最大时延) 和使用 PCE (具有 PREF) 的显式路由。

③ DetNet 流配置属性, 指定 DetNet 流属性, 例如流 ID、优先级、流量规范和封装方法。

④ DetNet 状态属性, 指定流状态反馈属性, 例如流性能 (时延、丢失分组、监管/过滤) 和 PREF 状态。

DetNet 流量的大多数控制功能通过和 IEEE TSN TG 相同的机制实现, 目前, DetNet 服务和传输层协议正在考虑各种协议和技术选项。根据文献[50], 数据平面协议的 2 个最突出的部署候选者是基于本地 IP 传输层的 UDP/TCP 服务层和在分组交换网络传输层上基于 MPLS 的基于伪线^[51]的服务层。

3.3.3 趋势

与给定 L2 网段中包含的 TSN 流控制操作和服务相比, 预计 DetNet 流控制操作将具有更大的规模和更高的复杂性。DetNet 流量控制将在互操作性、控制数据开销以及保证各种 L2 网段的 QoS 指标方面带来若干挑战。此外, 不同网段的所有者之间可能会出现 QoS 服务水平协议的不同要求。

4 确定性网络发展趋势分析

4.1 实现跨广域网的确定性业务

当前 TSN 标准已经相当成熟, 已有厂商推出支持 TSN 特性的交换芯片和支持部分特性的 TSN 交换机, 而 DetNet 目前进展还处于场景、需求、架构的前期阶段, 离提供 L3/L2 融合的确定性服务的标准和方案还有一定距离。目前已发布的工作组草案包括: 基本的 DetNet 架构、用例、安全, DetNet 数据平面的 IP 和 MPLS 方案, DeNet 配置模型中的数据流信息模型、YANG 模型等。DetNet 工作组下一步计划包括: 数据平面流的控制和转发; 数据配

置信息模型, 包括 YANG 模型、DetNet OAM 模型; DetNet QoS 保障技术。

确定性业务在跨域场景并且多条确定性业务流场景下, 每条流的特性配置包括带宽、时延、分组长度、发送频率、在端口的入时间窗口和出时间窗口, 以及每个节点间的出、入时间窗口的匹配; 在节点内部针对确定性业务流的资源分配包括时隙分配、循环间隔、带宽预留、门状态控制、流队列映射、抢占状态、多条流的资源冲突判断等。这些增加的技术复杂度是否导致应用的局限性, 比如局限在一定范围内的局域网内, 还有待进一步的研究分析。

4.2 处理好革新式架构和演进式部署的关系

目前, 确定性网络的部署采取平滑演进的方式, 基于传统以太网部署, 在新加入的流标识和流机制中都兼容了传统机制, 虽然浪费了部分网络资源, 但是降低了部署以太网的成本, 大大加快了确定性网络的部署和研究。随着确定性网络部署程度的增加, 如何减少为兼容传统网络造成的开销也是一个值得考虑的方向。

部署确定性网络的另一个关键问题是采用集中式部署还是分布式部署, 具体在建立同步时钟、流量控制、资源预留等方面都有体现。集中式部署和分布式部署分别具有特定的部署优势和缺点。TSN 基础设施和协议必须支持确定的端到端时延和可靠性, 为了支持各种相关协议, 网络基础设施变得更加复杂。因此, 简化的 TSN 管理机制对于降低复杂性同时满足确定性网络应用的关键需求至关重要。集中式部署可以从 SDN 的实施和管理中受益, 例如在建立同步时钟时, 为了实现精确的时间同步而在网络实体之间周期性进行的定时信息的交换会在控制层面引入额外的开销, 而采用集中式的时间同步系统, 如基于 SDN^[52-53]的设计, 仅需在中央控制器之间进行控制信息的交换, 可以帮助控制平面减少开销。但是, 集中式部署可能导致运营商的新基础设施成本, 且容易发生单点故障, 而分布式方案虽然可以避免这类问题, 但是在控制平面需要更多的开销。或许可以在集中式和分布式的性能, 以及现有基础架构的使用和新基础架构的部署之间权衡, 确立一种混合应用的模型。

4.3 确定性网络应设计故障和容错等安全机制

到目前为止, IEEE 802.1 安全工作组已经解

决了 IEEE 802.1 网络中通用的安全和隐私问题,即支持网络实体(即终端站和网桥)之间的安全通信的功能。这个 TG 详细介绍了许多标准和修订,专注于提供身份验证、授权、数据完整性和机密性。安全协议和标准与支持 TSN 的网络的集成需要在未来的研究和标准化中得到解决。例如,需要调查安全堆栈开销对 TSN 流的影响以及安全开销对在以太网 LAN 上运行的 OT 相关应用程序的影响^[8]。

由于 DetNet 集成了 IT(物理数据中心)和 OT(物理操作点),因此安全性是 DetNet 架构和协议的一个重要方面。先前的 OT 网络拓扑结构和设计具有“空中缺口”,即与外部世界完全隔离的 OT 网络,因此 IT 和 OT 的融合也将重点放在安全协议上,需要可扩展的、灵活的以及可以移植到 OT 网络组件的高效安全堆栈。此外,随着新兴的“雾”计算平台的发展,即基本上将 IT 移动到 OT 附近,必须密切检查交通和监控条件,因为任何入侵都可能导致灾难性的情况。尽管已经发布了许多针对确定性网络的标准和推荐实践,仍需要进行更多的基准测试,以便为行业 and 消费者市场提供保证。

4.4 确定性网络应确保各层技术间的融合

在 IP 网中划分出时隙提供硬管道是确保低时延的关键,不同网络层次上有不同的解决方案: FlexE 是 MAC 层以下技术,实现业务的管道隔离,不解决同一管道内的流量抢占问题; TSN 是链路层技术,基于流方式按照时间片进行流的刚性调度,可以解决同一管道内不同流传输的瞬时冲突,比如通过 IEEE802.1 Qat 流预留协议、IEEE 802.1Qcc 协议、IEEE 802.1CS 协议和 RAP 协议提供资源预留机制,通过 IEEE 802.1Qav、IEEE 802.1Qbv、IEEE 802.1Qch 和 IEEE 802.1Qcr 等协议提供帧抢占和流量整形机制,来保证确定性业务流的实时传送; DetNet 借鉴 TSN 的二层流的确定性传输技术,扩展到 IP 层,重点解决 IP 层的确定性业务传送,比如类似 TSN IEEE 802.1Qcc 的集中路径设置和 IEEE 802.1Qat 的分布路径设置,但还没有流量控制机制的设计。将各层间的技术有效融合,以及结合机器学习研究检查网桥中资源预留要求的预测模型,从而有效地管理队列和调度,有效地利用网络资源,降低技术复杂度和实现成本,需要产业界和学术界的更多探讨。

5 结束语

确定性网络已经得到全球学术界和产业界的充分重视,针对现有以太网“尽力而为”、无法保证提供 QoS 保障的状况,关于确定性网络架构的设计方法也各有特点,呈现出百家争鸣的现状,然而确定性网络的最终目标是一致的,即要建设一个提供确定性时延、抖动和低分组丢失率,为用户提供更好服务质量的可靠网络。本文介绍了目前确定性网络领域的主要研究成果,由于篇幅所限,并没有将所有的项目一一列举,只选择了其中比较具有代表性的技术和架构。期望通过综述该领域的已有研究成果,探讨分析研究目标和方法,总结研究思路,从而为相关领域的研究人员提供参考和帮助。

参考文献:

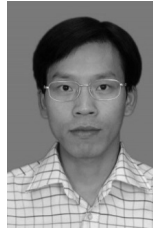
- [1] Cisco. Cisco visual networking index global mobile data traffic forecast update, 2015-2020[R]. Cisco Systems, 2015.
- [2] ANDRE L, GUILHERME S, AFONSO C. Panorama, challenges and opportunities in PROFINET protocol research[C]//The 13th International Conference on Industry Applications. IEEE, 2018: 186-193.
- [3] MARTIN R, JOSEPH E, DMITRY D. EtherCAT enabled advanced control architecture[C]//2010 Advanced Semiconductor Manufacturing Conference. IEEE/SEMI, 2010: 39-44.
- [4] VINH Q, JAE W. EtherCAT network latency analysis[C]//2016 International Conference on Computing, Communication and Automation. 2016: 432-436.
- [5] ZHAO L, HE F, LI E. Comparison of time sensitive network(TSN) and TTEthernet[C]//The 37th Digital Avionics Systems Conference. IEEE/AIAA, 2018: 1-7.
- [6] SOUSA R, PEDREIRAS P, GONÇALVES P. Enabling IIoT IP backbones with real-time guarantees[C]//The 20th Conference on Emerging Technologies & Factory Automation. IEEE, 2015: 1-6.
- [7] All. White paper on industrial Internet network connection[R]. Alliance of Industrial Internet, 2018.
- [8] KARAAGAC A, HAXHIBEQIRI J, MOERMAN I. Time-critical communication in 6TiSCH networks[C]// IEEE Wireless Communications and Networking Conference Workshops. IEEE, 2018: 161-166.
- [9] DUJOVNE D, WATTEYNE T, VILAJOSANA X. 6TiSCH: deterministic IP-enabled industrial Internet (of things)[J]. IEEE Communications Magazine, 2014, 52(12): 36-41.
- [10] PAVENTHAN A, DARSHINI D, KRISHNA B H, et al. Experimental evaluation of IETF 6TiSCH in the context of smart grid[C]//The 2nd World Forum on Internet of Things. IEEE, 2016: 530-535.
- [11] DRAHOŠ P, KUŠERA E, HAFFNER O. Trends in industrial communication and OPC UA[C]//2018 Cybernetics & Informatics. IEEE, 2018: 1-5.

- [12] YANG C, LI H, LIU Z. Implementation of migrations from class OPC to OPC UA for data acquisition system[C]//2012 International Conference on System Science and Engineering, 2012:588-592.
- [13] SCHWARZ M, BORCSOK J. A survey on OPC and OPC-UA: about the standard, developments and investigations[C]//2013 XXIV International Conference on Information, Communication and Automation Technologies. IEEE, 2013:1-6.
- [14] ECKHARDT A, MULLER S, LEURS L. An evaluation of the applicability of OPC UA publish subscribe on factory automation use cases[C]//The 23rd International Conference on Emerging Technologies and Factory Automation. IEEE, 2018:1071-1074.
- [15] NASRALLAH A, AKHILESH S, ZIYAD T. 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.
- [16] ZANZI L, SCIANCALEPORE V. On guaranteeing end-to-end network slice latency constraints in 5G networks[C]//15th International Symposium on Wireless Communication Systems. IEEE, 2018:1-6.
- [17] GROSSMAN E. Deterministic networking use cases draft-IETF-detnet-use-cases-18[S]. IETF, 2018.
- [18] LIM H, HERRSCHER D, WALTL M J, et al. Performance analysis of the IEEE 802.1 ethernet audio/video bridging standard[C]// The 5th International ICST Conference on Simulation Tools and Techniques. ICST, 2012.
- [19] ZHAO L, POP P, ZHENG Z. Time analysis of AVB traffic in TSN networks using network calculus[C]//2018 Real-Time and Embedded Technology and Applications Symposium. IEEE, 2018:25-36.
- [20] VERDUZCO H, CUIJPERS P, CAO J. Work-in-progress: best-case response time analysis for ethernet AVB[C]//Real-Time Systems Symposium. IEEE, 2017:378-380.
- [21] Huawei. Guangdong research institute of China telecom: white paper on flexible ethernet technology(2018)[R]. Shenzhen:Huawei, 2018.
- [22] EVELEENS J. Ethernet AVB overview and status[C]// SMPTE Technical Conference Exhibition. IEEE, 2014: 1-11.
- [23] TEENER M, FREDETTE A, BOIGER C. Heterogeneous networks for audio and video: using IEEE 802.1 audio video bridging[J]. Proceedings of the IEEE, 2013, 101(11): 2339-2354.
- [24] WG802.1. IEEE standard for local and metropolitan area networks—bridges and bridged networks: IEEE Std 802.1Q-2014[S]. IEEE, 2014.
- [25] FARKAS J, BELLO L, GUNTHER C. Time-sensitive networking standards[J]. IEEE Communications Standards Magazine, 2018, 2(2): 20-21.
- [26] WG802.1. IEEE standard for local and metropolitan area networks - timing and synchronization for time-sensitive applications in bridged local area networks: IEEE Std 802.1AS-2011[S]. IEEE, 2011.
- [27] STANTON K. Distributing deterministic, accurate time for tightly coordinated network and software applications: IEEE 802.1AS, the TSN profile of PTP[J]. IEEE Communications Standards Magazine, 2018, 2(2): 34-40.
- [28] WG802.1. IEEE standard for a precision clock synchronization protocol for networked measurement and control systems: IEEE Std 1588-2008[S]. IEEE, 2008.
- [29] AIJAZ A, SIMSEK M, DOHLER M. Shaping 5G for the tactile internet[M]. Switzerland: Springer International Publishing, 2017: 677-691.
- [30] BJORKLUND M. YANG - a data modeling language for the network configuration protocol (NETCONF) [S]. The Internet Engineering Task Force, 2010.
- [31] BIERMAN A. Guidelines for authors and reviewers of YANG data model documents[S]. The Internet Engineering Task Force, 2011.
- [32] HOLNESS M. IEEE draft standard for local and metropolitan area networks - media access control (MAC) bridges and virtual bridged local area networks amendment: YANG data model: IEEE P802.1Qcp/D0.7[S]. IEEE, 2016.
- [33] WG802.1. IEEE standard for local and metropolitan area networks—virtual bridged local area networks amendment 14: stream reservation protocol (SRP) : IEEE Std 802.1Qat-2010[S]. IEEE, 2010.
- [34] WG802.1. IEEE draft standard for local and metropolitan area networks - media access control (MAC) bridges and virtual bridged local area networks amendment: stream reservation protocol (SRP) Enhancements and Performance Improvements: IEEE P802.1Qcc/D2.0[S]. IEEE, 2017.
- [35] FINN N. IEEE draft standard for local and metropolitan area networks - media access control (MAC) bridges and virtual bridged local area networks amendment: link-local registration protocol: IEEE P802.1CS/D1.2[S]. IEEE, 2017.
- [36] WG802.1. IEEE draft standard for local and metropolitan area networks - media access control (MAC) bridges and virtual bridged local area networks amendment: stream reservation protocol(SRP) enhancements and performance improvements: IEEE P802.1Qcc/D2.0[S]. IEEE, 2017.
- [37] ENNS R, BJORKLUND M, BIERMAN A. Network configuration protocol (NETCONF)[S]. The Internet Engineering Task Force, 2011.
- [38] BIERMAN A, BJORKLUND M, WATSEN K. RESTCONF protocol[S]. The Internet Engineering Task Force, 2017.
- [39] WG802.1. IEEE standard for local and metropolitan area networks—bridges and bridged networks: IEEE Std 802.1Q-2014[S]. IEEE, 2014.
- [40] WG802.1. IEEE standard for local and metropolitan area networks - bridges and bridged networks - amendment 25: enhancements for scheduled traffic: IEEE Std 802.1Qbv-2015[S]. IEEE, 2016.
- [41] WG802.1. IEEE standard for local and metropolitan area networks - virtual bridged local area networks amendment 12 forwarding and queuing enhancements for time-sensitive streams: IEEE Std 802.1Qav-2009[S]. IEEE, 2009.
- [42] WG802.1. IEEE standard for local and metropolitan area networks - bridges and bridged networks - amendment 26: frame preemption:

- IEEE Std 802.1Qbu-2016[S]. IEEE, 2016.
- [43] SPECHT J, SAMII S. Urgency-based scheduler for time-sensitive switched ethernet networks[C]//Euromicro Conference on Real-Time Systems. IEEE, 2016: 75-85.
- [44] SPECHT J, SAMII S. Synthesis of queue and priority assignment for asynchronous traffic shaping in switched ethernet[C]//IEEE Real-Time Systems Symposium. 2017:178-187.
- [45] WG802.1. IEEE standard for local and metropolitan area networks—bridges and bridged networks - amendment 24: path control and reservation: IEEE Std 802.1Qca-2015[S]. IEEE, 2016.
- [46] CHEN Y, ZHANG H, FISHER N. Probabilistic per-packet real-time guarantees for wireless networked sensing and control[J]. IEEE Transactions on Industrial Informatics, 2018, 14(5): 2133-2145.
- [47] FINN N, THUBERT P, VARGA B. Deterministic networking architecture[R].[2019-04-25].
- [48] LEE Y, CECCARELLI D, MIYASAKA T. Requirements for abstraction and control of TE networks[R].The Internet Engineering Task Force, 2018.
- [49] GENG X, CHEN M. DetNet configuration YANG model[R]. The Internet Engineering Task Force, 2017.
- [50] KORHONEN J, ANDERSSON L, JIANG Y. DetNet data plane encapsulation[R]. The Internet Engineering Task Force, 2017.
- [51] PATE P, BRYANT S. Pseudo wire emulation edge-to-edge (PWE3) architecture[R]. The Internet Engineering Task Force, 2005.
- [52] ALVIZU R, MAIER G, KUKREJA N. Comprehensive survey on T-SDN: software defined networking for transport networks[J]. IEEE Communications Surveys & Tutorials, 2017, 19(4): 2232-2283.
- [53] NUNES B, MENDONCA M, NGUYEN N. A survey of

software-defined networking: past, present, and future of programmable networks[J]. IEEE Communications Surveys & Tutorials, 2014, 16(3): 1617-1634.

[作者简介]



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。

汪硕 (1991—)，男，河南灵宝人，博士，北京邮电大学在站博士后，主要研究方向为数据中心网络、软件定义网络、网络流量调度等。

黄玉栋 (1998—)，男，重庆人，北京邮电大学硕士生，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。

郑尧 (1998—)，女，吉林四平人，北京邮电大学硕士生，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。

刘江 (1983—)，男，河南郑州人，博士，北京邮电大学副教授，主要研究方向为网络体系架构、网络虚拟化、软件定义网络、信息中心网络等。

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。