

边缘计算环境下低时延网络架构研究进展

张坚实

长春职业技术大学信息学院, 吉林长春, 130033;

摘要: 随着物联网、工业互联网、自动驾驶和增强/虚拟现实 (AR/VR) 等新兴应用的蓬勃发展, 对网络服务的实时性要求日益严苛, 低时延成为衡量网络性能的关键指标。边缘计算通过将计算、存储和网络资源下沉至网络边缘, 靠近数据源和终端用户, 为解决传统云计算架构下的高时延问题提供了有效途径。然而, 在边缘计算环境下构建高效、可靠的低时延网络架构仍面临诸多挑战。本文系统梳理了边缘计算环境下低时延网络架构的研究进展。首先, 阐述了边缘计算与低时延需求的内在关联, 分析了传统网络架构的局限性。其次, 从网络架构设计 (如多级边缘、分布式架构)、资源管理与任务卸载 (如计算、通信、缓存资源协同优化, 智能卸载策略)、网络功能与协议优化 (如低时延路由、拥塞控制、确定性网络技术) 以及新兴技术融合 (如网络切片、软件定义网络 SDN、网络功能虚拟化 NFV、人工智能 AI 赋能) 等维度, 综述了当前主流的技术方案与研究动态。最后, 探讨了当前研究面临的挑战, 包括动态环境适应性、跨域协同、安全性与隐私保护、能效优化等, 并展望了未来的研究方向, 旨在为构建更智能、高效、可靠的低时延边缘网络提供参考。

关键词: 边缘计算; 低时延; 网络架构; 任务卸载; 资源管理; 网络功能虚拟化; 软件定义网络; 人工智能

DOI: 10. 64216/3080-1508. 25. 08. 036

引言

以物联网 (IoT)、工业 4.0、车联网 (V2X)、远程医疗、沉浸式交互 (AR/VR) 为代表的新兴应用正以前所未有的速度改变着社会生产和生活方式。这些应用普遍具有海量数据、高带宽、低时延、高可靠性的核心需求, 尤其是对端到端时延的要求极为苛刻 (通常需达到毫秒级甚至亚毫秒级)。例如, 自动驾驶要求车辆与路侧单元或云端的决策响应时延必须极低, 以确保行车安全; 工业自动化中的实时控制环路对时延抖动极为敏感; AR/VR 应用需要极低的交互时延以避免用户眩晕感。传统的集中式云计算架构, 虽然具备强大的计算和存储能力, 但由于数据需要长距离传输至远端数据中心进行处理, 再将结果返回, 不可避免地引入了显著的网络传输时延和处理排队时延, 难以满足上述新兴应用的严格时延要求。边缘计算 (Edge Computing) 作为一种新兴的分布式计算范式, 应运而生。其核心思想是将原本集中在云数据中心的计算、存储和网络功能下沉到更靠近数据产生源头 (如终端设备、传感器) 和用户 (如基站、接入点、本地服务器) 的网络边缘。

1 边缘计算环境下的低时延网络架构研究进展

1.1 异构网络架构设计

异构网络架构 (Heterogeneous Network Architec

ture) 是指在边缘计算环境中, 整合多种类型、不同性能、位于不同层级的计算、存储和网络资源, 形成一个协同工作的复杂系统。它打破了单一、同质化网络的局限, 融合了多级边缘 (如微边缘、接入边缘、区域边缘)、分布式对等网络、蜂窝网络 (4G/5G/6G)、Wi-Fi、有线网络、卫星网络等多种接入和传输技术, 以及专用硬件 (如 GPU、FPGA、AI 加速器)、通用服务器、嵌入式设备等多样化计算节点。其核心在于“异构性”和“协同性”, 旨在通过资源的多样化和智能调度, 灵活应对不同应用场景的差异化需求。

1.2 异构网络架构的关键技术

软件定义网络 (SDN) 与网络功能虚拟化 (NFV): SDN 实现了控制平面与数据平面的分离, 提供了集中的、可编程的网络控制能力, 是实现异构网络资源统一调度和灵活编排的核心。NFV 将传统的网络功能 (如防火墙、负载均衡器) 从专用硬件中解耦, 以软件形式在通用服务器上运行, 使其能够按需部署在任何边缘节点, 极大地提升了网络功能的灵活性和部署效率。SDN+NFV 是构建可编程、可重构异构边缘网络的基石。

网络切片 (Network Slicing): 基于 SDN/NFV 技术, 网络切片允许在同一个物理网络基础设施上, 创建多个逻辑上隔离的、端到端的虚拟网络。每个切片可以

拥有独立的网络拓扑、资源分配、协议栈和 QoS 策略，专门服务于特定类型的应用（如为自动驾驶创建一个超低时延高可靠的切片，为视频流创建一个高带宽的切片）。这是实现“一网多用”和差异化服务的关键技术。

1.3 边缘计算节点部署策略

边缘计算节点的选择和部署位置节点选择：选择何种类型的节点作为边缘计算载体，取决于应用场景、性能需求和成本约束。常见选择包括：专用服务器：部署在基站机房、汇聚机房或小型数据中心，性能强，扩展性好，适合处理复杂任务。通用网关/路由器：具备一定计算能力的网络设备，成本较低，易于部署，适合轻量级任务处理。基站集成（MEC）：直接将计算模块集成到 4G/5G 基站中，位置最优，时延最低，是移动边缘计算的主要形式。云服务提供商边缘节点（如 AWS Wavelength, Azure Edge Zones）：大型云厂商在运营商网络边缘部署的节点，便于与公有云无缝集成。

2 边缘计算网络协议优化

2.1 现有的网络协议在边缘计算环境下的挑战

高时延与低吞吐量：TCP 的拥塞控制机制（如慢启动、拥塞避免）在无线网络中容易将丢包误判为拥塞，导致窗口急剧减小，造成高时延和低吞吐量。连接建立开销大：三次握手过程在短连接、高时延场景下开销显著。队头阻塞（Head-of-Line Blocking）：单个数据包的丢失或延迟会阻塞后续所有数据包的交付。路由效率低：传统路由协议（如 OSPF, BGP）主要优化路径成本或跳数，对时延优化不足，且收敛速度慢。无线链路不稳定：无线信道易受干扰、衰减和多普勒效应影响，导致链路质量波动大，传统协议难以快速适应。缺乏确定性保障：传统 IP 网络提供“尽力而为”（Best-Effort）服务，无法为关键业务提供严格的时延上限和抖动保证。

2.2 低时延网络协议的设计和优化

2.2.1 传输层优化

采用 QUIC：基于 UDP，集成了 TLS 加密，支持 0-RTT/1-RTT 快速连接建立，多路复用避免队头阻塞，可自定义拥塞控制算法（如 BBR），是替代 TCP 的理想选择。

开发专用低时延协议：针对特定场景（如工业控制）设计极简、高效的专用协议，最大限度减少协议开销。

改进拥塞控制：设计适用于无线边缘环境的拥塞控制算法，如利用显式拥塞通知（ECN）、基于延迟的算

法（如 BBR）或 AI 驱动的智能算法。

2.2.2 网络层与数据链路层优化

低时延路由协议：设计以最小化端到端时延或跳数为首要目标的路由算法，如最短路径优先（SPF）的变种，或结合链路质量（如 ETX）的路由。

确定性网络（DetNet）：通过资源预留（带宽、缓冲区）、流量整形（如门控调度、信用整形）、路径复制/消除、减少抖动等技术，在 IP/MPLS 网络中提供有界低时延、低抖动和高可靠性的服务。

时间敏感网络（TSN）：在以太网层面，通过时间同步（802.1AS）、流量调度（802.1Qbv）、帧抢占（802.1Qbu/802.3br）等标准，实现微秒级的确定性低时延通信，是工业边缘网络的核心。

5G uRLLC：5G 新空口（NR）通过缩短传输时间间隔（TTI）、使用更短的 HARQ-ACK 反馈周期、免调度资源分配、重复传输等技术，实现 1ms 级的空口时延和 99.999% 的可靠性。

2.2.3 跨层优化与智能协议

SDN/NFV 赋能：SDN 的集中控制为跨层优化提供了可能。控制器可以获取全网视图，综合应用需求、网络状态和计算资源，进行端到端的最优决策。

AI/ML 驱动的协议：利用机器学习模型预测网络状态（如链路质量、流量模式），动态调整协议参数（如拥塞窗口、重传超时、路由选择），实现自适应、智能化的网络行为。

2.3 低时延网络协议的性能评估和比较

在评估和比较不同低时延协议时，需在时延、抖动、可靠性、吞吐量、复杂性、可扩展性、兼容性等多个维度进行权衡。例如，TSN 在局域网内能提供极佳的确定性，但跨域扩展性有限；DetNet 旨在解决跨域问题，但实现复杂；QUIC 能显著提升 Web 性能，但对现有网络设备有一定要求。选择何种协议需根据具体应用场景的需求来决定。

3 边缘计算环境下的低时延网络架构应用案例

3.1 工业自动化领域

实时过程监控与控制：在生产线、化工厂、电力系统等场景，传感器实时采集海量数据（如温度、压力、振动、电流）。边缘计算节点部署在车间或控制室，对这些数据进行毫秒级的本地处理和分析，实现对设备运

行状态的实时监控,并直接向 PLC(可编程逻辑控制器)或执行器发送控制指令,确保生产过程的稳定、安全和高效。

预测性维护(Predictive Maintenance):通过边缘节点对设备(如电机、泵、风机)的运行数据(振动、温度、噪声)进行实时分析,利用机器学习模型识别潜在的故障模式。一旦发现异常,立即发出预警,安排维护,避免非计划停机,减少维护成本。

3.2 智能交通领域

车联网(V2X)通信与协同感知:路侧单元(RSU)作为边缘节点,接收来自车辆(V2V)、行人(V2P)和基础设施(V2I)的信息(如位置、速度、意图)。边缘节点进行融合计算,生成更全面的交通态势图,并将预警信息(如前方碰撞、交叉路口冲突、道路危险)低时延地广播给相关车辆,辅助驾驶决策。自动驾驶决策支持:对于 L3 及以上级别的自动驾驶,车辆自身的算力可能不足以处理所有复杂场景。边缘计算节点(如 RSU、MEC)可提供额外的算力,处理高清地图更新、复杂交通流预测、路径规划等计算密集型任务,并将结果快速反馈给车辆,弥补单车智能的不足。

3.3 虚拟现实领域

将 VR/AR 应用中复杂的 3D 图形渲染任务从终端设备(头戴显示器、手机)卸载到边缘计算节点。边缘节点完成渲染后,将生成的视频流低时延地传输回终端。这解决了终端设备算力不足、功耗高、发热严重的问题,实现了高质量、沉浸式的体验。

实时交互与内容分发:对于多人在线 VR 游戏或社交应用,边缘节点作为本地服务器,处理用户间的交互逻辑、状态同步和音视频通信,减少与远端数据中心的通信延迟,保证交互的实时性和流畅性。

4 边缘计算环境下的低时延网络架构未来发展趋势

4.1 新一代网络技术的融合

以 5G、物联网(IoT)、人工智能(AI)为代表的新一代信息技术正加速融合,形成强大的技术合力,共同驱动社会数字化转型。这种融合不再是简单的叠加,而是深度的协同与赋能,催生了全新的应用场景和商业模式。5G 的大连接(mMTC)和低时延高可靠(uRLLC)特性,为海量物联网设备的接入和实时控制提供了理想

的通信基础。物联网设备产生的海量数据通过 5G 网络高效回传,为上层应用提供数据支撑。5G 网络产生的海量、高维、实时的网络数据(如信道状态、用户行为、流量模式)为 AI 模型的训练和优化提供了丰富的“燃料”。同时,AI 可以赋能 5G 网络,实现智能的无线资源管理、网络切片编排、故障预测与自愈等,提升网络效率和性能。三者的融合催生了更多对极致低时延要求的应用,如基于 AI 的实时视频分析(安防、质检)、AI 赋能的自动驾驶决策、海量 IoT 设备的协同控制等。这要求网络架构必须持续优化以满足更严苛的时延指标。

4.2 边缘计算环境下的低时延网络架构创新

从“连接”到“算力+连接”融合:未来的网络架构不仅是数据传输的管道,更是计算能力的延伸。创新方向是将网络功能与计算功能深度融合,如在路由器、交换机中集成轻量级计算单元,实现“在网计算”(In-Network Computing)或“近数据计算”,减少数据移动,直接在传输路径上进行简单处理(如过滤、聚合、AI 推理),进一步降低端到端时延。

确定性网络(DetNet/TSN)的广域化与融合:将局域网内的 TSN 技术与广域网的 DetNet 技术深度融合,构建端到端的确定性网络,为跨域的工业控制、车联网等关键业务提供从终端到云端的、有严格保障的低时延、低抖动服务。

4.3 边缘计算环境下的低时延网络架构挑战 and 解决方案

用户移动性导致网络拓扑和链路质量快速变化;IoT 设备产生的流量具有突发性和不可预测性;边缘节点的计算负载波动剧烈。这种动态性使得网络状态难以准确预测,给低时延保障带来巨大挑战。边缘节点的计算、存储、带宽和能源资源有限且分布不均。同时,网络和计算资源高度异构(不同厂商、不同性能)。如何在资源受限且异构的环境中高效协同,实现低时延目标,是核心难题。涉及运营商、云服务商、企业、设备商等多个利益相关方。实现跨管理域的资源发现、服务编排、安全策略统一和计费结算,技术上和商业上都存在巨大复杂性。

5 结论

研究系统梳理并整合了异构网络架构设计、边缘节

点部署策略、网络协议优化等核心维度,构建了一个相对完整、逻辑清晰的边缘计算低时延网络架构研究框架。该框架不仅涵盖了从物理部署到协议设计的多层次技术,还深入分析了其在工业自动化、智能交通、虚拟现实等关键领域的应用,为后续研究提供了系统的理论参考。当前的 AI 应用多集中于单点优化。未来研究应致力于开发更强大的端到端 AI 模型,能够跨网络、计算、应用多层进行联合感知、预测与决策。研究可解释性 AI (XAI) 在网络管理中的应用,使 AI 决策过程更透明、可信。关注量子计算、类脑计算等前沿计算范式的发展,探索其未来可能为边缘计算和低时延网络带来的革命性变革,如超高速计算能力或全新的信息处理模式。

参考文献

- [1] 汪明,黄杰,仲从乐,等. 通信网络与信息安全标准化体系架构探究[J]. 中国标准化,2025, (15): 75-79.
- [2] 李婷. 基于 5G 网络的车联网系统架构与性能优化研究[J]. 汽车电器,2025, (07): 36-38. DOI: 10. 13273/j.

cnki. qcdq. 2025. 07. 020.

- [3] 陈德进,王丽波,卜小琳,等. 基于 5G uRLLC 的工业网关端到端低时延通信研究[J]. 邮电设计技术,2025, (06): 30-33.
- [4] 程国明. 基于生物实验楼无源光网络系统的应用与分析[J]. 安装,2025, (08): 104-107.
- [5] 王睿秋雨,张洪硕,张孙烜,等. 基于通感协同的配电网数据高频采集与低时延互动方法[J/OL]. 电工技术学报, 1-13[2025-08-25]. <https://doi.org/10.19595/j.cnki.1000-6753.tces.250075>.

作者简介: 张坚实 (1975.06.25—), 性别: 男, 民族: 汉, 籍贯: 吉林省长春市, 单位名称: 长春职业技术大学信息学院, 学历: 研究生, 职称 副教授, 主要研究方向: 计算机科学技术, 计算机网络技术。

项目基金: 课题号: 2024-AFCEC-687, 课题名称: 以数字能力和素养为导向的网络基础及应用课程教材建设。