



专题：新型网络

MEC 业务连续性技术研究与应用

毛磊, 王卫斌, 罗鉴

(中兴通讯股份有限公司, 江苏 南京 210012)

摘要: 分析了当前移动性面临的难题以及可行的解决思路, 同时研究了 3GPP 网络侧现有的业务连续性能力和关键特性, 包括会话和服务连续模式、超高可靠与低时延通信、流量引导等。提出了一种基于网络侧和应用侧协同机制, 满足低时延业务在边缘快速移动涉及跨 MEC 切换场景下的业务连续性要求。并探讨了移动性管理后续研究的总体思路 and 分阶段实施路径。

关键词: 会话和服务连续模式; 多接入边缘计算; 无状态设计; uRLLC

中图分类号: TP929

文献标识码: A

doi: 10.11959/j.issn.1000-0801.2020193

Research and application of MEC service continuity technology

MAO Lei, WANG Weibin, LUO Jian

ZTE Corporation, Nanjing 210012, China

Abstract: The current mobility problems and feasible solutions were analyzed, and the existing service continuity capabilities and key features on the network-side which were defined by 3GPP were studied, including SSC mode, uRLLC, traffic routing and so on. A network-side and application-side collaboration mechanism was proposed to satisfy the low latency services involving cross-MEC handover for fast mobility at the edge scenarios. The general idea of the follow-up study on mobility management and the path of phased implementation were also proposed.

Key words: SSC mode, multi-access edge computing, stateless design, uRLLC

1 引言

5G 万物互联, 催生了众多类似云 XR、云游戏、V2X、无人机巡检等新兴应用。各类低时延和高带宽的业务要求大量的数据能够在本地消化, 而传统集中式的算力处理模式难以满足 5G 新应用差异化的需求。

MEC (multi-access edge computing, 多接入边缘计算) 作为 5G 的核心技术之一, 是云计算由中心向边缘的一种延伸, 可以在接近数据侧提供低时延的连接和海量的计算能力。通过中心云的算力下沉到用户边缘, 结合云边协同显著提升网络效率、用户体验, 降低响应时延。随着 5G 和 IoT 的成熟商用, MEC 具备低时延、低成本、优

算力的特性, 将成为新兴的热点技术, 使能 5G 新商业^[1]。

低时延且高速移动性的业务, 在满足端到端超低时延要求的基础上, 还要求网络通信具备高可靠性和可用性。传统 TCP/IP 网络标识(包括身份标识、位置标识等)和 IP 地址绑定, IP 地址具备网络层路由地址和传输层通信链接标识的双重身份, 在 IP 地址变化时网络路由地址发生变化, 同时传输层通信链接标识也随之变化, 传输层改变使业务无法连续, 如果把网络层路由地址和传输层通信链接标识拆分开, 各自独立标识, 那么在 IP 地址发生变化时, 传输层标识就可以保持不变, 从而达到业务连续的目标。这种方案对现有的网络协议改动较大, 短期内不具备可实施性。

当前在移动通信网络中最直接的体现就是位置标识和 IP 地址绑定, 即终端所在的位置确定了对应的 IP 地址。当用户从一个小区移动到另一个小区, 涉及重定位和切换, 移动锚点的变化对应 IP 地址的改变, 会对应用时延和可靠性造成影响, 应用也需要重新放置到贴近用户的位置, 连接点的切换和 MEC 重定位会导致较高的时延。

边缘移动性作为 MEC 的一个重要特性, 为边缘高速移动业务的连续性提供支持, 当应用在网络中移动超出了当前 MEC 服务覆盖范围时, 涉及了跨 MEC 之间的切换, 此时 IP 地址变化会影响终端业务的 QoE (quality of experience)^[2], 对于要求超低时延和高可靠性的应用是不可接受的。因此为了支持低时延应用的移动连续性, MEC 移动性需要能够解决以下难题, 从而将切换引起的时延降到最低, 甚至达到无缝水平。

- 网络侧: 实时保证网络连接的高可靠和 IP 会话的连续性、应用的移动性感知、实时和超精准的重定位, 并能将网络状态信息通过能力开放平台给第三方应用调用。
- 应用侧: 上层应用平台能够做到在业务忍

受的中断时间范围内完成应用层的业务迁移(例如, XR 业务通常需要应用平台切换在 20 ms 以内完成), 采用云原生理念, 通过无状态设计做到应用和数据分离^[3], 数据独立迁移保证源和目的一致性, 应用本地加载可降低有状态的数据在跨 MEC 迁移时的同步, 降低对源和目的 CPU 一致性的要求。

- 应用和网络协同: 网络感知应用的位置变化时通知应用平台, 通过应用平台和网络协同快速在目标 MEC 上分配新应用实例, 实现新用户面功能 (UPF) 锚点和目标应用的位置一致性。

2 移动会话连续性关键技术

2.1 关键技术 1: 会话和服务连续模式

3GPP 从网络侧角度出发, 5G 核心网提供了 3 种 SSC 模式(session and service continuity mode, 会话和服务连续模式)^[4], 如图 1 所示, 保障业务连续或保持 IP 地址不变。策略控制功能 (PCF) 提供 SSC MSP (SSC mode selection policy) 给 UE, 会话管理功能 (SMF) 根据 UE 请求 SSC 模式, 签约本地策略, 确定 SSC 模式。

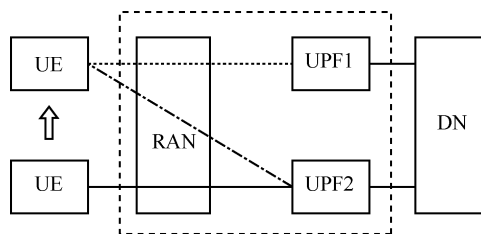


图 1 3 种业务连续性模式

会话和业务连续性模式特点和应用场景见表 1。

SSC mode1 移动锚点是固定的方式, 终端 IP 地址不变化, 在移动过程中, 流量存在迂回, 传输时延大, 满足不了低时延业务需求。SSC mode2/3 移动锚点是变化的方式, 终端 IP 地址随锚点变化, 业务链接存在拆链和重建的过程, 业



表 1 会话和业务连续性模式特点和应用场景

SSC 模式	模式特点	典型应用场景
SSC mode1	始终保持协议数据单元（PDU）会话建立时的锚点 UPF 不变	时延不敏感和会话连续性要求高的业务：运营商语音、视频等
SSC mode2	PDU 会话的锚点 UPF 可以改变，网络触发 PDU 会话的释放，并指示 UE 立即建立到同一 PDN 的 PDU 会话连接	时延不敏感和会话连续性要求不高的业务：互联网浏览数据类等
SSC mode3	PDU 会话的锚点 UPF 可以改变，且保持业务持续	业务感知（业务需要参与流程）且有严苛业务连续性要求的业务

务连续性、可靠性难保证。因此，通常语音类业务采用 SSC mode1，需要低时延的业务则上层应用结合 SSC mode3 解决移动业务连续性问题。

SSC mode2 用户移动导致应用要求重新选择会话锚点时，PDU 会话先拆后建，并且重新为 UE 分配 IP 地址，适用于对连续性要求不高的场景。PDU 会话 SSC mode2 基本流程如图 2 所示，具体介绍如下。

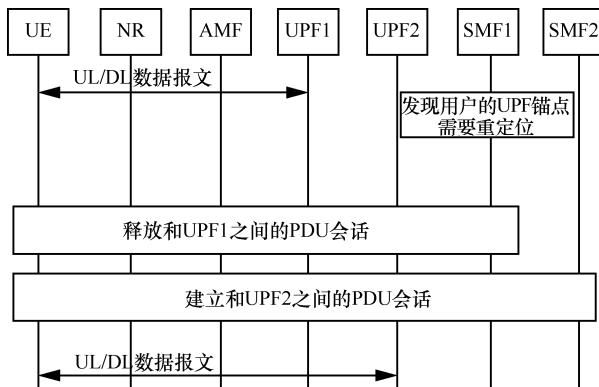


图 2 PDU 会话 SSC mode2 基本流程

- (1) 由 SMF 确定 UPF 的重定位；
- (2) SMF 向 UE 发送非接入层（NAS）消息，携带 PDU 会话 ID，要求会话重新进行建立；
- (3) SMF 执行 PDU 会话释放流程；
- (4) UE 收到释放命令，根据原因值确定 PDU 会话重建，UE 发起 PDU 会话重建过程；
- (5) AMF 选择 SMF，由 SMF 为用户重新选择 UPF。

SSC mode3 用户移动导致应用要求重新选择会话锚点时，PDU 会话先建后拆，并且重新为 UE

分配 IP 地址，终端上层保障业务连续性，适用于对连续性要求高的场景。PDU 会话 SSC mode3 基本流程如图 3 所示，具体介绍如下。

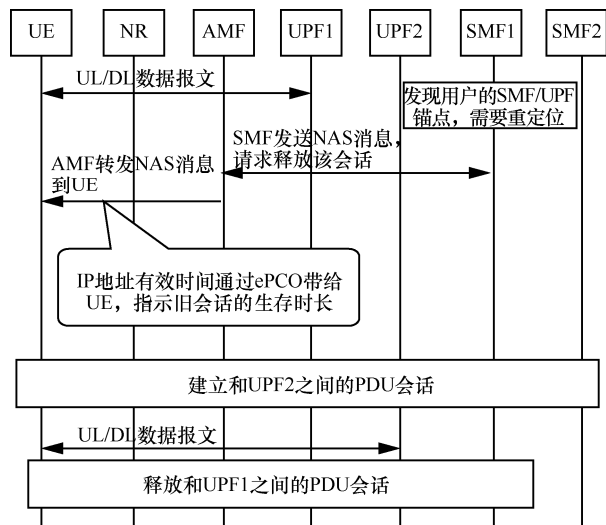


图 3 PDU 会话 SSC mode3 基本流程

- (1) 由 SMF 确定 UPF 需要重定位，或者 SMF 需要改变；
- (2) SMF 向 UE 发送 NAS 消息，携带 PDU 会话 ID，SMF 重分配请求指示，要求释放该会话，并提供定时器指示该会话保留时间；
- (3) SMF 启动 PDU 会话释放定时器；
- (4) AMF 转发 NAS 消息到 UE，UE 将释放定时器值提供给上层；
- (5) UE 接受 PDU 会话修改命令，执行 PDU 会话建立流程，AMF 根据 SMF 重分配指示，选择新 SMF，由新 SMF 为用户重新选择 UPF；
- (6) SMF 根据旧会话 ID，为用户重新选择

PDU 锚点 UPF ;

(7) 一旦 PDU 会话重新建立, UE 使用新的 IP 地址和前缀转发数据流;

(8) SMF 或者 UE 根据释放定时器, 释放原有旧会话。

综合 SSC mode1/2/3 对比分析, 针对低时延又有移动性需求的场景, 当应用变动的时候, 即用户位置改变, 需要 MEC、应用和 5G 网络结合 SSC mode3 共同解决。

2.2 关键技术 2: 本地路由及流量引导

MEC 与 5G 网络的集成^[5], 如图 4 所示, 5G 核心网 CU 分离后, 可以将 UPF 下沉至 MEC 部署, SMF 通过选择策略, 基于上行分类器 (UL-CL)、多归属 (multi-homing) 和本地数据网络 (LADN) 方案实现边缘 UPF 的选择及特定数据业务本地分流, 实现最小化的访问时延。非本地流量, 从本地 UPF 通过 N9 发送给中心的 UPF, 满足云边业务灵活的路由访问。

- UL-CL 基于目的 IP 地址进行分流, 当 PDU 会话类型为 IPv4 或 IPv6 或 IPv4v6 或 Ethernet 时, SMF 可决定在 PDU 会话的数据路径中插入或删除一个 UL-CL。UL-CL 采用流过滤规则 (例如检查 UE 发送的上行 IP 数据分组的目的 IP 地址/前缀) 决定数据分组如何路由, UE 对于 UL-CL 无感

知, 并且不参与 UL-CL 规则插入删除。

- multi-homing 基于目的 IP 地址前缀进行分流, 一个 PDU 会话可关联多个 IPv6 前缀, 提供通过多个 PDU 会话锚点到 DN 的接入。到不同 PDU 会话锚点的不同用户面路径在公共的 UPF 形成分支, 公共的 UPF 被称为支持“分支点”功能的 UPF。分支点转发上行流量到不同 PDU 会话锚点, 并聚合发送到 UE 的下行流量, 即聚合从不同 PDU 会话锚点发送到 UE 的流。multi-homing 仅适用于 IPv6 类型的 PDU 会话, 并且 UE 须支持 IPV6 和参与 PDU 会话锚点改变过程。
- LADN 基于 UE 只有本地数据网络的服务区才能通过连接本地数据网络的 PDU 会话访问本地数据网络。本地数据网络服务区是一个 TA 集, LADN 的服务是由服务 PLMN 提供的。SMF 向 AMF 订阅针对 LADN DNN 的“UE 位置变化通知”, 并根据从 AMF 收到的 UE 位置和 LADN 服务区关系 (在服务区、不在服务区、不确定在不在服务区) 的信息, 对 PDU 会话进行处理。

IPv6 多归属 PDU 会话可用于支持先建后拆业务连续性, 以支持 SSC mode3, 如图 5 所示。此时, 当应用变动导致会话锚点重定位时, 5G 核

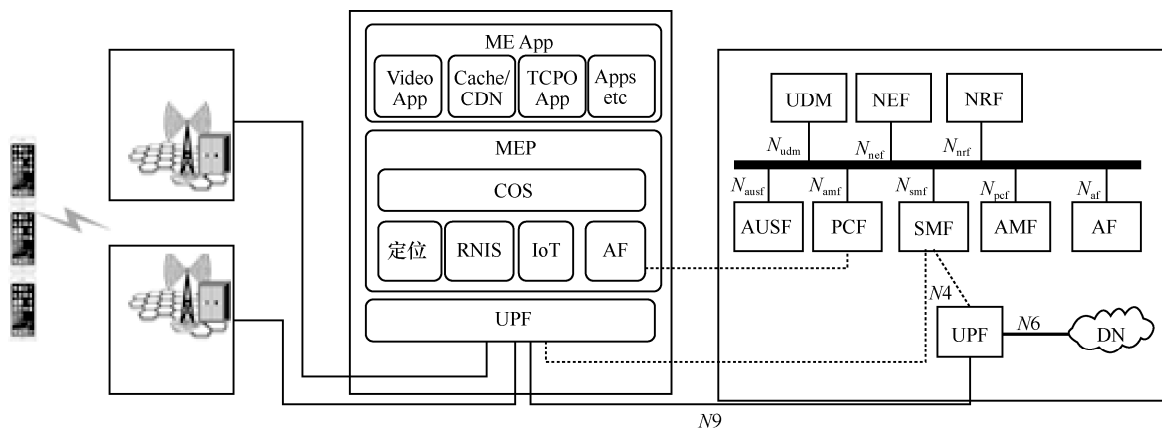


图 4 5G 网络集成 MEC 部署



心网除了先建一条新的 PDU 会话，同时还要通过能力开放功能（NEF）通知上层的应用平台，使得应用平台实时将边缘应用迁移到新的 UPF 锚点附近，满足低时延的需要。

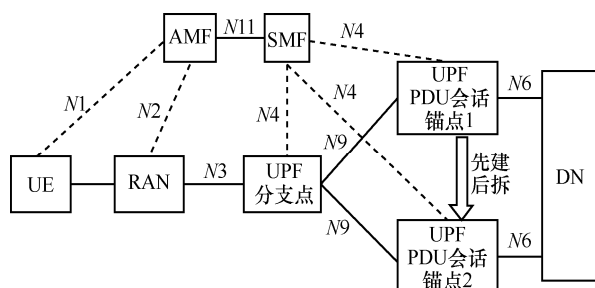


图5 多归属 PDU 会话下的业务连续性示例

2.3 关键技术 3：网络和应用的通知协同机制

运营商和第三方应用之间可通过 PCF 与 AF/NEF 间的消息实现相互通知的协同机制，如图 6 所示，从而实现第三方应用基于业务的特点向 5G 网络发送对网络的需求以及 UE 终端或网络的状态改变时及时通知到应用侧^[6]。网络和应用之间的通知协同机制有利于加快应用对终端行为、网络状态的感知，以便为跨 MEC 切换提前进行预准备操作，是实现低时延业务在移动切换过程中快速应用迁移和业务连续性保障的重要技术手段。

有了网络和应用间的通知协同机制之后，还需要应用侧支持跨 MEC 的服务迁移和应用上下文的快速迁移才可以解决跨 MEC 移动的业务连续性，两者缺一不可。

应用实例在 MEC 上部署完成之后，可以通过 AF request 消息，代表应用将影响路由策略的控制信息发送给 PCF（可信域）或者经由 NEF/UDR 发送给 PCF（非可信域），运营商也可以选择所有 AF 交互消息全部通过 NEF。

- 应用侧可以通过 AF request 消息传递应用对网络资源的需求，从而实现网络侧对该业务流保障业务所需的时延、带宽需求的保障。

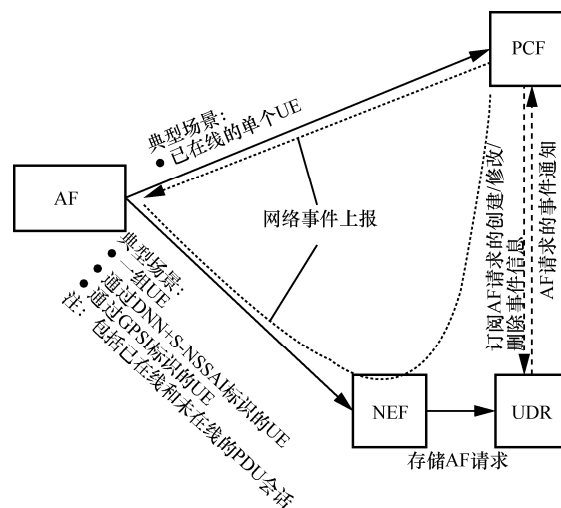


图6 网络和应用间的通知和协同机制

- 应用侧也可以通过 AF/NEF 向网络订阅事件通知消息，以便当网络检测到订阅的事件触发时，立刻通知 AF。
- 应用侧还可以通过 AF 发起终端数据路由的修改流程，调整 UL-CL/multi-homing 策略实现增加、修改或删除业务路由所经的 UPF。

当网络侧在感知到 AF 订阅的事件发生时，例如：SMF 感知到 PDU 会话的 UPF 锚点发生改变、DN 接入标识符（DNAI）改变等事件发生，网络侧将发送通知消息给 AF，并且携带改变后的目标数据网络标识 DNAI，以便应用侧可以得知将应用服务迁移至哪个 MEC 中。

然后，AF 可以发起应用侧的跨 MEC 迁移，将该终端的应用上下文迁移至新的 UPF 锚点所在 MEC 的应用服务中。应用侧迁移期间，网络侧将维持之前的用户面数据路径，保持业务不中断。

当应用完成迁移后，AF 将通知 SMF 切换已完成，然后 SMF 发起切换用户面数据路径，将终端的数据业务从新的 UPF 锚点转发到新的 MEC 应用服务中，从而实现跨 MEC 平滑切换的移动性管理。

2.4 关键技术 4：uRLLC 低时延高可靠性保障

ITU-R 为 5G 定义了增强型移动宽带（eMBB）、

超高可靠与低时延通信 (uRLLC) 和海量机器类通信 (mMTC) 三大场景, 其中 uRLLC^[7] 支持双连接来构建高可靠低时延网络, 冗余的传输路径机制, 实现移动性流程中的 PDU 会话锚点重定位的用户会话连续性增强, 满足低时延场景下数据通信的高可靠性要求。主要定义了以下 3 种高可靠业务保障方案。超高可靠性业务的 3 种方案如图 7 所示。

- UE 双连接方案主要是从端到端解决可靠性, 包括空口和核心网、承载网端到端, 这个方案利用两条冗余的 PDU 会话传输数据, 由应用层进行分组复制和重复分组检测。对终端、应用服务器基本没有影响、对网络影响较小。
- 回传网双链接方案有两个变种, 一个是双 N3/N9 隧道, 另外一个单 N3/N9 隧道但

是在传输层上进行冗余传输。这个方案假设空口已经满足高可靠的要求 (这个前提很难达到), 前一种在 GTP-U 层实现分组复制和重复分组检测, 对协议影响较大, 后一种在传输层实现分组复制和重复分组检测, 对 3GPP 协议没有影响但是对传输有要求。

- 单设备双终端方案和 UE 双连接方案类似, 通过两个终端 PDU 会话来实现冗余传输, 这时候从 UE、基站、AMF 到 SMF、UPF 都是双路径, 可靠性最高。

3 种方案的可靠性对比见表 2, 分析可知, 回传网双链接方案存在空口和 UPF 单点故障可能, 对于低时延、高可靠性的业务是不可接受的。单设备双终端方案可靠性最高, 并且对网络影响较小, 短期内可实施性最强。

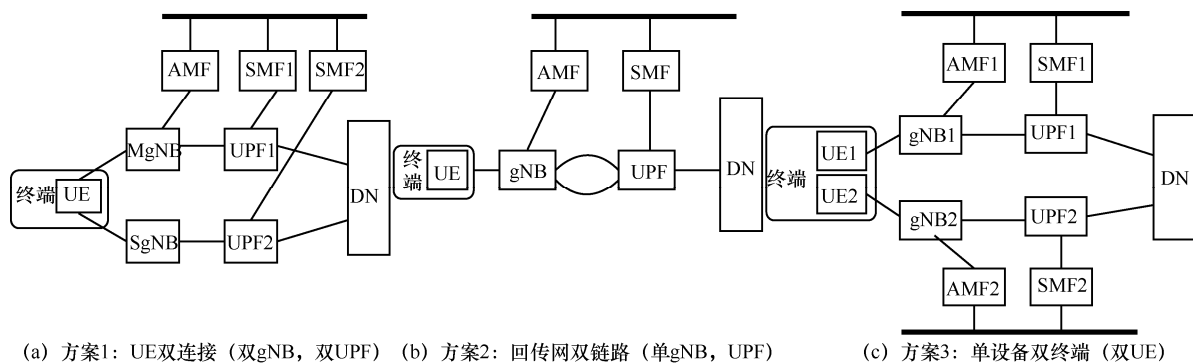


图 7 超高可靠性业务的 3 种方案

表 2 uRLLC 3 种方案可靠性对比

	方案 1: UE 双连接	方案 2: 回传网双链路		方案 3: 单设备双终端
		双 N3/N9 隧道	单 N3/N9 隧道	
对终端要求	UE 接入层基本无感知	UE 无要求		UE 接入层基本无感知
对无线要求	支持双连接和 RSN (redundancy sequence number) 处理	假设空口已经满足高可靠的要求		支持双连接和 RSN 处理
对网络要求	应用层进行分组复制和重复分组检测, 对核心网影响小	对 GTP-U 实现分组复制和重复分组检测, 对协议影响大	在传输层实现分组复制和重复分组检测, 对传输有要求; 对核心网影响小	对核心网影响小
可靠性分析	高	中, 存在空口和 UPF 单点故障		最高
切片要求	一个或两个切片	一个切片		一个或两个切片



3 MEC 业务连续性优化方案

车联网是一种典型的边缘 MEC 高速移动连续性业务，场景包括车路协同安全辅助、自动驾驶协同感知控制、交通管理等，移动性强、时延低、云边协同和网络高可靠是车联网的核心^[8]。V2X 场景和关键需求见表 3。

表 3 V2X 场景和关键需求

场景	关键需求
车路协同安全辅助	时延 2~20 ms，多数据融合，边云协同
自动驾驶协同感知控制	数据安全高可靠、弹性扩容，边云协同控制
交通管理	本地化服务、弹性扩容、应用平滑升级
信息服务	带宽>100 Mbit/s，远程升级、弹性扩容

对于这类在高速移动的网络中低时延应用，除了下沉到 MEC 部署，还涉及多个 MEC 之间业务连接的连续性，应用实例跨 MEC 的快速实例化和状态信息同步也至关重要^[9]。在高速移动过程中，不论是太迟/太早定位，还是重定位到错误的 MEC 节点，都是不可接受的。

MEC 业务连续性优化方案通过网络侧和应

用侧之间的协同机制加快和优化应用跨 MEC 的切换处理，以满足低时延、空间无限移动的业务场景需求。上层应用侧结合网络侧 SSC mode3 能力解决移动性和低时延业务的连续性，并采用 uRLLC 双连接特性保障高可靠，如图 8 所示。当高速、低时延的业务需要跨 MEC 切换时，网络侧实时感知并上报终端的位置、状态等信息，应用平台通过网络侧的信息提前在目标 MEC 上创建新的应用实例，并同步应用上下文信息，完成到目标 MEC 新的应用实例的切换，从而解决低时延、高速移动过程中因 PDU 会话锚点导致 UE IP 地址变化后上层业务连续性的问题。

- 高可靠保障：结合 uRLLC 双链路特性，采用双 PDU 会话不同切片，UE 和 DN 间使用两条独立的网络路径，一条路径上任意网络节点故障，包括 RAN、UPF、交换机、路由器等设备，UE 和 DN 间数据帧零丢失。
- 跨 MEC 快速实例化：支持承载容器化边缘网元和 IT 业务，应用跨 MEC 切换时，能够在目标 MEC 上快速分配新的应用，实现秒级实例化。
- 减少状态信息的同步：提供 PaaS 公共组件，

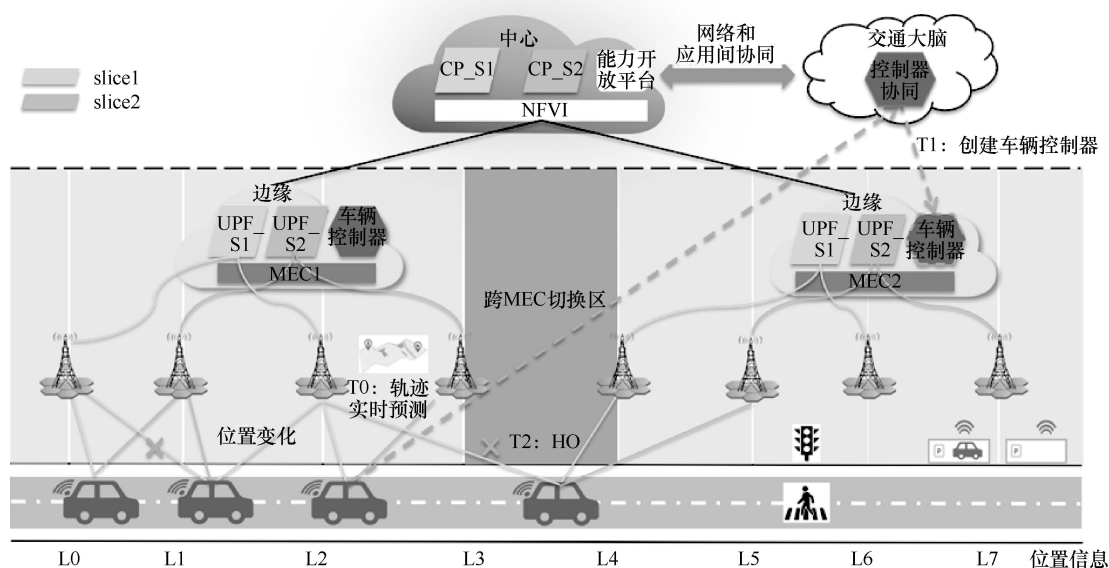


图 8 MEC 高速移动业务连续性优化方案

助力第三方应用无状态设计,使上层应用更轻薄。

- 高精度定位:精准地感知应用位置信息,基于 AI 的移动轨迹预测,并通过标准化接口开放网络的实时状态信息,包括无线网络信息、位置信息、用户信息等。
- MEC 和移动应用的融合:MEC 为应用提供辅助计算、数据存储等支持,降低端到端数据传输时延,满足实时交互的应用,对于应用平台中心控制系统移至中心集中部署,进行统一调度和决策。

4 结束语

5G 的一大强劲优势就是可以支持超低时延,作为整个社会经济数字化、自动化转型的利器,只有更低的时延、更高的可靠性才能够满足垂直行业的需要,低时延场景下关键的业务连续性问题需要网络设备商、第三方应用和科研机构一同携手探索来解决。通过网络侧和应用侧之间的协同处理机制,共同解决在移动过程中业务连续性问题,助力工业互联、XR 等场景对未来网络要求高带宽、低时延、高可靠、空间无限制移动的需求。

移动性管理涉及界面广泛,需要不断地结合实际的应用场景进行实践完善。后续的研究需要分阶段实施,总体思路分成 3 部分进行,实现由特定解决方案向通用方案的演进。当前移动性管理主要聚焦特定场景业务连续性解决方案,典型的如车联网场景业务连续性解决方案。下一个阶段将聚焦业务迁移关键技术解决,如 MEC 上业务迁移关键方案。最后考虑移动业务迁移网络解决方案以及网络层/传输层协议改进提案。

参考文献:

- [1] 中兴通讯股份有限公司. Common Edge 边缘计算白皮书[R]. 2020.
ZTE Corporation. Common edge edge computing white paper[R]. 2020.
- [2] ETSI. Application mobility service API: ETSI GS MEC 021 V2.1.1[S]. 2020.
- [3] ETSI. MEC in 5G networks[S]. 2018.
- [4] 中兴通讯股份有限公司. 5G core network 技术趋势白皮书[R]. 2019.
ZTE Corporation. 5G core network technology trends white paper[R]. 2019.
- [5] 3GPP. System architecture for the 5G system: 3GPP TS23.501[S]. 2019.
- [6] 3GPP. Procedures for the 5G system: 3GPP TS23.502[S]. 2019.
- [7] 3GPP. Study on physical layer enhancements for NR ultra-reliable and low latency case (uRLLC): 3GPP TS33.824[S]. 2019.
- [8] IMT-2020. MEC 与 C-V2X 融合应用场景白皮书[R]. 2019.
IMT-2020. MEC and C-V2X convergence application scenarios white paper[R]. 2019.
- [9] 邱佳慧,周志超,林晓伯,等. 基于 MEC 的车联网技术研究及应用[J]. 电信科学. 2020, 36(6): 45-55.
QIU J H, ZHOU Z C, LIN X B, et al. Research and application of V2X based on MEC[J]. Telecommunications Science, 2020, 36(6): 45-55.

[作者简介]



毛磊(1989-),男,中兴通讯股份有限公司电信云及核心网产品规划经理,主要研究方向为 5G 核心网、边缘计算、网络切片等。

王卫斌(1973-),男,中兴通讯股份有限公司电信云及核心网产品总工程师,主要研究方向为 5G 核心网、边缘计算、云计算等。

罗鉴(1972-),男,中兴通讯股份有限公司有线研究院总工程师、高级工程师,CCSA TC3 副主席、CCF 互联网专业委员会常务委员,主要研究方向为分组数据网络、网络架构演进等。