网络切片在 5G 无线接入侧的动态实现和发展趋势

Realization and Trend of Network Slicing in 5G NG-RAN

杨立 /YANG Li，李大鹏 /LI Dapeng
（中兴通讯股份有限公司，广东 深圳 518057）
(ZTE Corporation, Shenzhen 518057, China)

摘要：网络切片功能是未来 5G 蜂窝移动网络的必选功能，如何在 5G 蜂窝移动网络无线接入侧实现端到端（E2E）的网络切片功能是业界的研究重点。针对该问题，介绍了第三代合作伙伴计划（3GPP）中的相关标准进展，并系统阐述了其背后的技术发展脉络、缘由，以及各自优劣点。认为 5G 网络已初步具备 E2E 的网络切片能力，但仍有一些技术需要进一步发展和提升。

关键词：网络切片；5G 蜂窝移动网；无线接入网；功能编排；资源共享

Abstract: The network slicing function is the required function of the future 5G cellular mobile network. How to realize and implement the end to end (E2E) network slicing function in 5G celullar next generation-radio access network (NG-RAN) is one of the hot topics. Based on various facts of 3rd Generation Partnership Project (3GPP) standardization process, the comprehensive views about its development are given and advantages and disadvantages of its related essential technical issues are analyzed. It is considered that the 5G network has the capability of E2E network slicing, but there are still some technologies that need to be further developed and promoted.

Keywords: network slicing; 5G cellular mobile network; RAN; function orchestration; resource sharing

DOI：10.12142/ZTETJ.201906002
网络出版地址：http://kns.cnki.net/kcms/detail/34.1228.TN.20191211.0845.002.html

网络出版日期：2019-12-11
收稿日期：2019-06-14

1 网络切片的需求

随着蜂窝移动通信进入到 5G 时代，一系列新的行业特征和移动通信技术也随之而来。从产业生态的角度看，5G 实现了真正意义上的全球蜂窝移动通信统一技术标准，因此全球信息通信技术（ICT）产业力量将得到空前聚合，同时产业能力也将得到集中释放；业界将专攻于 5G 唯一制式——蜂窝网络，这更有利于把它做得更好、更精。从业务应用的角度看，5G 旨在更高效地统一支持增强移动宽带（eMBB）、海量机器类通信（mMTC）、高可靠低时延通信（uRLLC）3 大类业务，这涉及到不同的行业领域、千差万别的动态服务质量（QoS），以及更宽范围的业务服务需求。这有利于推进传统运营商的业务升级、跨行业间的资源协作整合，并带来各种新应用的快速商业部署。从网络节点形态的角度看，5G 网络节点的建立基于异构融合化、模块解耦化、功能虚拟化、软件定义化等先进的技术理念；因此在产品业务和服务的开发、构建、部署、应用等方面，将都变得更加灵活且适配，成本和性能将更加按需可控。总之，5G 蜂窝移动网络将成为未来人类社会强大的通信基础设施，为各行各业的生产力和生产关系再发展注入更多的向上使能。未来各种更新潮、更高级的移动业务应用都离不开 5G 网络的支持保障。过去，出于节省网络的资本性支出

（CAPEX）/ 运营成本（OPEX），达到快速部署业务应用的目的，蜂窝移动运营商们可利用无线接入网（RAN）共享的合作模式，即一套物理无线网络设备（包括基站、空口和传输资源等）由多个协议运营商共享地使用和运维，来自不同运营商的签约用户可同时接入和被服务于同一个物理服务小区。通过这种共享方式，能尽可能地实现网络资源复用，避免无线网络资源的闲置浪费。由于网络共享并不能有效地实现多运营商间的不同用户、不同业务的应用部署，不同资源配置、运维计费等方面的充分隔离和差分定制化操作；因此在用户管理、业务功能编排、资源分配、调度协同等方面难以做得很好。这同时也会造成 RAN 资源无法被最优化地复用和综合利用，不同用户的体验无法得到充分保障。运营商们逐渐意识到，基于专有核心网（DCN）技术而进行的核心网功能剪裁去冗和不同用户业务群差分归类处理，可以带来精细化管理。于是，业界很快达成共识，网络端到端的切片化功能，更有利于网络资源的优化利用，可以对不同用户业务管理和服务质量提供保障。作为 5G 蜂窝移动网络端到端的必要一部分，5G RAN 切片是整个网络切片在无线接入侧的功能实现，它能保证 RAN 对外呈现出网络切片的相关重要特征 [1-2]。

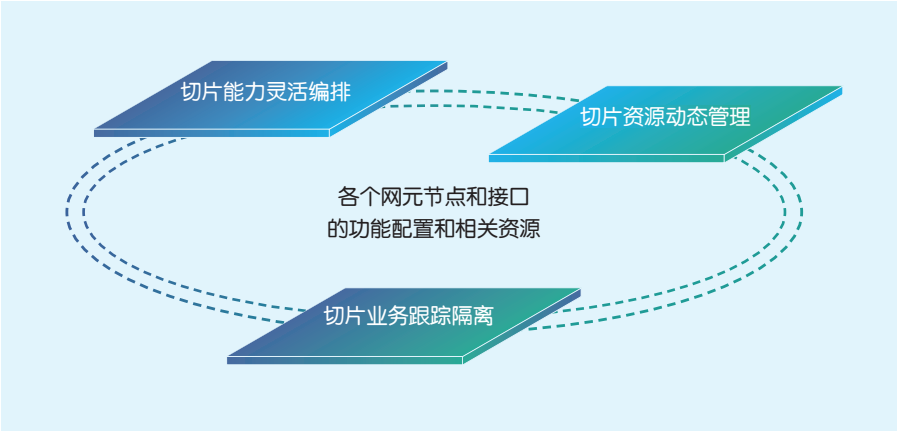
5G 面向万物互联，需要承载指数级增多的无线连接数和更多元的移动业务应用。不同运营商通常面向不同的市场用户业务群，有着可归类描述的服务等级协议（SLA），例如，不同的市场用户服务质量的差异、在特定区域内的容量覆盖要求、传输带宽和时延要求、链路安全等级和可靠性等属性。上述可归类的用户业务属性集合，为网络切片提供了天然的细分市场需求。随着 5G 时代新商业合作模式的发展，如虚拟运营、大数据运营，以及基于 RAN 共享的概念，运营商们期待一套 5G 物理网络基础设施，能够同时提供多个虚拟的端到端子网络，动态共享着各类网络实体资源，但又能相对独立地存在和被管理且彼此隔离。这些虚拟子网络被称为网络切片。网络切片功能的实现可用图 1 中的 3 个特征来概括：切片能力灵活编排、切片资源动态管理、切片业务跟踪隔离。换言之，非网络切片化就无法高效地支持上述 3 大特征，并会导致运营商无法灵活地部署业务并充分利用网络资源 [3]。

（1）切片能力灵活编排：建立在对网络功能模块和虚拟化的基础上，每个网络切片可以按需配置特定的网络功能或能力集合，去除不需要的网络功能或能力，减少一切配置的冗余。

（2）切片资源动态管理：每个网络切片运行中所消耗的资源，应尽量和切片实时业务量相关，不同切片间要尽量共享地使用资源，以提升资源利用率。

（3）切片业务跟踪隔离：对某切片业务的服务在网络移动中能保持一定的物理连续性和一致性，不同的切片业务间必须保持彼此不干扰，服务质量不互相影响。

如图 2 所示，类似于云计算服务的分层框架，5G 蜂窝移动网络提供的各类服务可抽象地分为 3 层：网络基础设施层、通信平台使能层、网络切片编排层。网络基础设施层尽量采取软件定义网络（SDN）的方式去实现网络基本功能，以延续网络各类硬件资源的生命力；通信平台使能层基于下层的网络基础设施，进一步构建出虚拟化的网络功能或能力，从而形成虚拟化网络功能（VNF）；网络切片编排层基于下层的通信平台使能，进一步灵活编排出特定的网络切片，从而对外提供有针对性且定制



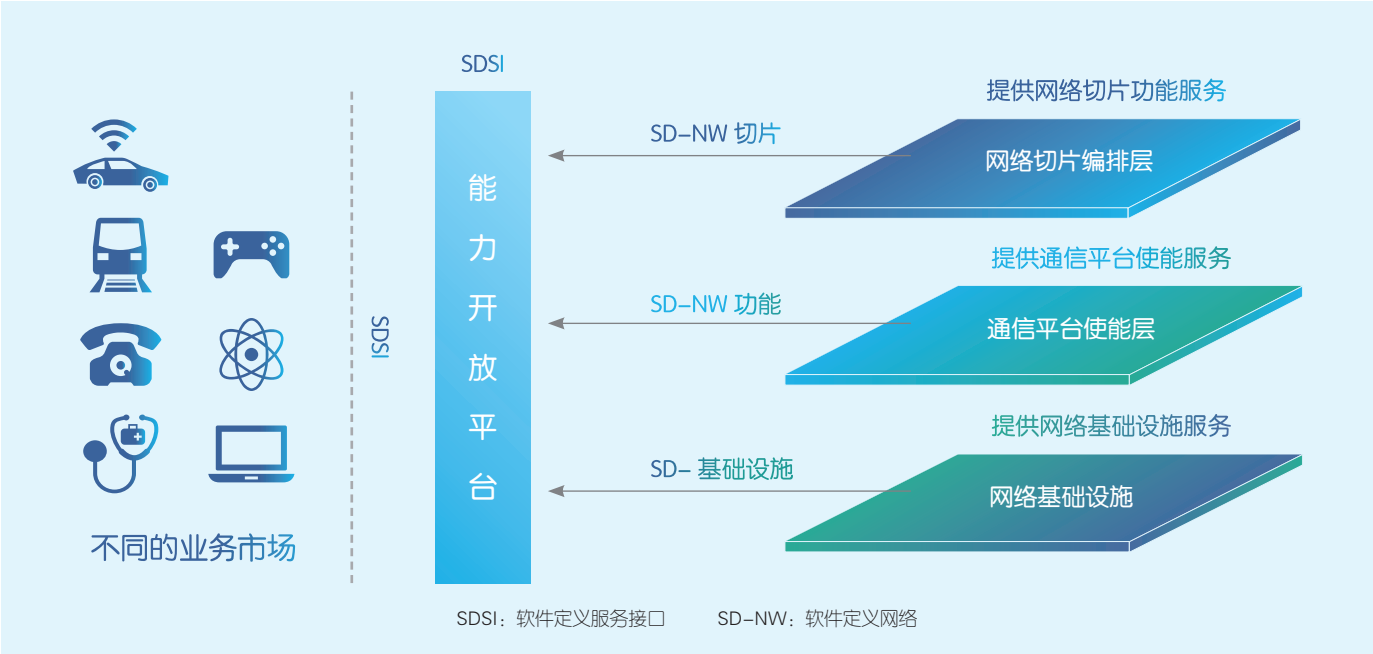
▲图 1 端到端网络切片功能的 3 大实现特征

化的切片业务服务。上述 3 大层，都可各自通过能力开放平台对外提供服务；但网络切片编排层提供的服务，无疑是最有针对性和定制感的，因此对于各类行业用户和应用开发者而言是最简洁方便的。基于 SDN/NFV 技术、统一运营管理（OAM）和切片编排器系统，运营商可在同一物理基础设施上，根据各种主客观因素和具体业务需求，按需动态地去构建编排网络切片。强大的切片编排器，可通过数据驱动和软件模板等方式，快速创建生成所需的业务切片，实现切片参数集的适配定制化。此外，切片还必须具备资源量弹性伸缩和还原能力释放，以对切片服务质量进行实时监测控制分析和诊断。综上所述，5G 网络切片能按需独立灵活且非常高效地去定制化配置网络功能，动态分配资源，承载大动态范围下不同的用户业务应用，以实现对 5G 网络基础设施各类资源的最优化利用[4]。

2 网络切片的标准化发展

第三代合作伙伴计划（3GPP）从 2015 年开始，启动了对网络切片的标准化研究和制定工作：SA1 最先研究了场景业务，SA2 随后制定了系统框架，SA5 也对网络切片有专门的课题研究。2016 年 3 月，随着 5G 下一代无线接入网（NG-RAN）/ 新空口（NR）制式研究立项的启动，网络切片功能对 NG-RAN 的协议影响和落地实现，进一步地被标准化推进。2018 年 6 月，随着 Rel-15 5G 的工作立项（WID）顺利完成，网络切片端到端功能的第 1 个版本也正式完成。

SA2 确定了 5G 核心网（5GC）的控制面接入移动管理功能（AMF）不需要切片化，而用户面功能（UPF）及控制面会话管理功能（SMF）要被切片化。结合 5G 新的协议数据单元（PDU）会话承载管理架构，SA2 规定单个 PDU 会话只能从属服务于单个网络切片，即切片业务在网络侧承载管理的最小粒度是 PDU 会话，如此的粒度较好地平衡了网络承载灵活度、复杂度、差异化管理这三者之间的综合要求。SA2 并没具体规定 NG-RAN 该如何去实现切片功能，RAN2/3 研究规范了 NG-RAN 如何实现切片功能。网络“业务类型”和“租户”是决定网络切片划分的 2 个主要因素，因此 SA2 规定单个网络切片选择辅助信息（S-NSSAI）包含切片服务类型（SST）和切片区分租户（SD）2 个属性。S-NSSAI 可用来唯一标识特定的网络切片，由运营商内部协同进行赋值。目前业界已完成试验的 5G 典型切片，包括监测控制终端群物联网切片、车联网 / 无人机切片、公共安全管理切片、远程医疗服务



▲图 2 5G 蜂窝移动网络服务提供的分层框架

切片、工业自动化切片等[5-6]。

SA2并没对RAN侧切片功能的具体实现方式做精确定义，也未对RAN侧的功能配置、资源分配和隔离保护的具体做法进行规定。RAN侧的功能包含NG-RAN各个协议层所支持的全部特征功能或其组合，例如，支持不同的波段频段组合、多输入多输出（MIMO）多天线技术、半静态调度（SPS）、非连续接收（DRX）、载波聚合（CA）、双连接（DC）、非激活新状态（INACTIVE）等；RAN侧的资源主要包含NG-RAN基站内部的基带计算存储资源、射频天线和功率资源、空口载波无线时频资源、基站外部网络接口相关的传输资源等（下文简称为RAN资源）。在实际的RAN实现和部署应用中，网络切片根据对RAN资源使用的动态性和弹性程度，可划分为以下3大类：

（1）像特定行业的专网或专用类基站那样，某NG-RAN的RAN资源被某切片专门地占据使用。如果没有相关业务数据发生，这些RAN资源就是被闲置浪费的。当该基站需要去服务其他新切片业务时，必须通过宕机和重定义编排的方式，重新做配置和资源划分预留。这种方式称为切片的静态方式。

（2）某NG-RAN的RAN资源在某段时间内，被某切片暂时预留而占据使用。如果该切片的业务量变化较大，相应的资源就会通过网络上层接口信令做半静态地调整，或被回收或增配。当该基站需要服务其他新切片业务，不必通过宕机和重定义、重编排，可直接进行新切片的创建和相关资源的暂时预留。这种方式被称为切片半静态方式。

（3）某NG-RAN的RAN资源在任何时间内，理论上可供任何切片被分配和使用。无论该切片的业务量如何变化，相应的资源就会通过网络下层信令的方式进行动态调整，实现资源快速回收或增配。当该基站需要服务其他新切片业务时，不必通过宕机或重定义、重编排，也无须相关资源预留，就可直接进行新切片创建。这种方式称为切片动态方式。由于运营商们很看重RAN资源的动态共享性，不希望因为构建网络切片而导致空口容量减少，资源利用率降低；因此上述的第3类动态方式成为RAN切片功能最重要的实现目标，也是众多运营商和3GPP追求的标准化方向。理论上，任何网络切片上的业务承载，都能充分利用基站内外硬软件资源和空口无线资源。因此，NG-RAN基站的内部和接口处理必须时刻跟踪和区分当前处理的切片业务数据块和哪个切片相关联，且为了保证切片彼此间的隔离，需要执行一些必要的约束操作，如无线资源管理（RRM）算法配置、资源复用方式、无线时频资源调度处理等，同时这些约束操作又不能降低切片业务QoS或移动性等方面的服务质量。

NG-RAN在空口仍基于数据无线承载（DRB）粒度的承载方式和基站侧主控资源分配调度的原则，因此NG-RAN天然具备对RAN资源的细切分和强管控能力。为了以必选的方式支持端到端网络切片功能，NG-RAN必须要能支持图1中的NG-RAN切片能力灵活编排（主要通过网管和编排器预配置、在线重配的方式）、NG-RAN切片资源动态管理（基于承载切片业务的DRB及QoS参数信息，依赖于基站RRM和调度增强）、NG-RAN切片业务跟踪隔离（依赖于基站对业务所属的切片标识的及时跟踪和管理）等功能。

从用户终端（UE）的角度看，单个UE可同时和8个不同的网络切片连接并参与相关的切片服务，每个切片业务可同时被多个PDU会话所承载，如UE同时连接切片S-NSSAI1（被PDU会话1 and 2承载相关的切片业务数据）和切片S-NSSAI2（被PDU会话3 and 4承载相关的切片业务数据）。虽然UE在用户面同时连接多个网络切片（即消耗着不同的用户面资源），但在控制面仍然只有一个无线资源控制（RRC）连接和网元间的信令连接（如NG-C、Xn-C、F1-C）。不同切片在空口共享着相同的信令无线承载（SRB），而在网元间的接口控制面仍然共享着相同的流控制传输协议（SCTP）连接承载（即消耗着共同的控制面资源）。网络切片在5G无线接入侧的实现架构如图3所示。针对单个UE，NG-RAN可通过内部实现增强，保证不同的切片业务在用户面上的充分隔离，但无法做到在控制面上的隔离。例如，当某UE发生无线链路失败（RLF），该UE所有的切片业务都要中断。但针对不同的UE，

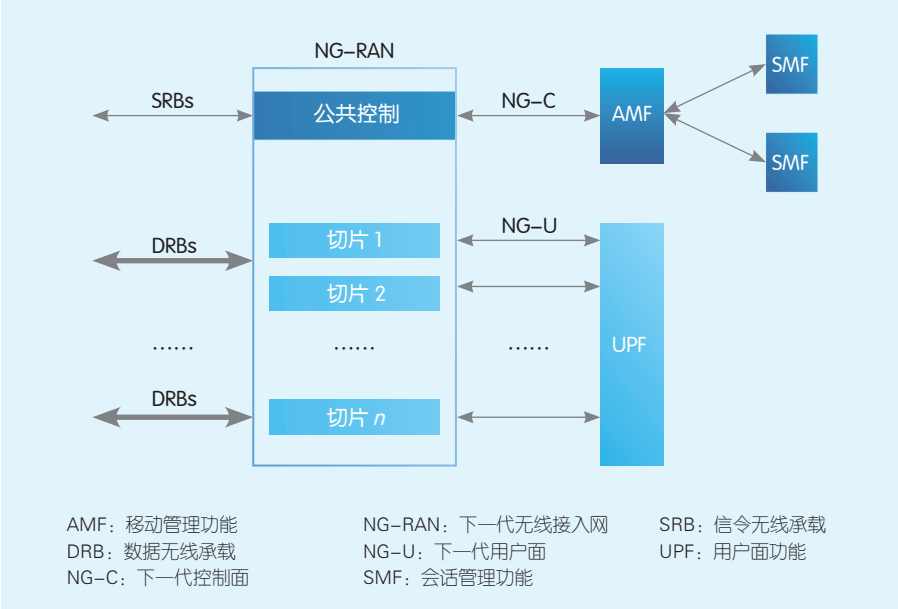
NG-RAN 仍然可充分保证它们在用户面和控制面上的彼此隔离，例如，几个不同 UE 即使在同一位置经历了相同的无线环境，当一个 UE 发生了 RLF，另一个 UE 也还可能保持正常的通信连接。

3 5G 无线接入侧的具体实现

3.1 空口物理层

5G NR 工作立项 WID 并没有针对网络切片功能做物理层的专门设计，因为 5G NR 支持非常灵活的空口无线时频资源的使用，例如，引入了时分双工（TDD）动态帧结构、不同的 Numerology/ 传输时间间隔（TTI）、部分带宽（BWP）、增强的异步调度机制。在 5G 网络中，为了更好地支持不同业务动态范围的 QoS 需求，不同的 NG-RAN 基站可部署在不同的频带和频点上，本地服务小区可配置不同的帧结构和 Numerology/TTI 等，从而可灵活地适配于不同的网络切片功能。例如，某些 eMBB 切片可大量地利用高频无线资源，而 uRLLC 切片尽量地利用优质的低频无线资源；某些下行流量为主的切片尽量多配置 TDD 下行子帧资源，而上行流量为主的切片可多配置 TDD 上行子帧资源。

和过去 4G 演进的通用陆地无线接入（E-UTRA）制式不同，5G NR 从第 1 个版本 Rel-15 开始就已针对各种移动全业务类型，面向各种部署场景，旨在充分利用更广的频带资源（sub-100 GHz）；因此，5G NR 比 4G E-UTRA 具备更灵活、更精细的空口无线资源管理和利用能力。从空口物理层的角度看，5G NR 更能有效地支持端到端网络切片功能在无线接入侧的灵活实现；而 E-UTRA 即使在 5G 系统框架内不断地演进，在某些切片服务能力方面可能亦有所不及 NR。



▲图 3 网络切片在 5G 无线接入侧的实现架构

3.2 空口高层

3.2.1 空口高层控制面

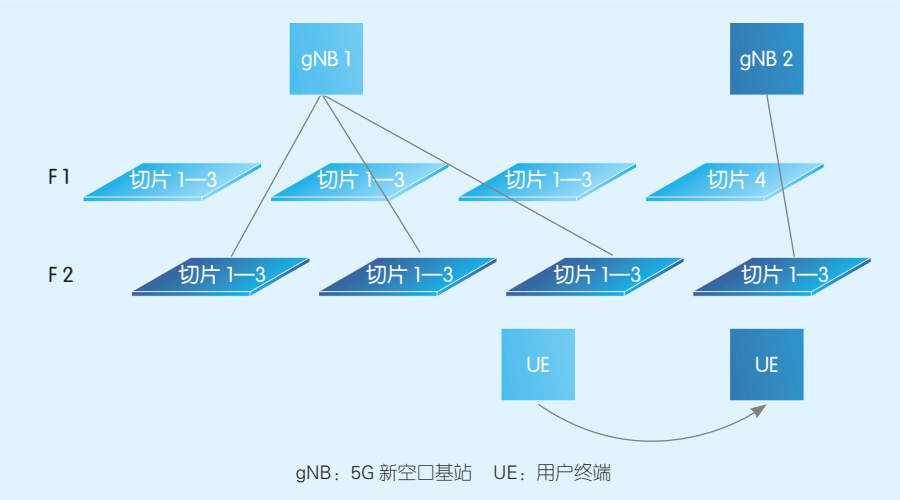
当前，NG-RAN 基站的服务小区，不要求显式地广播它所支持的切片能力信息，因此 UE 无法通过读取系统消息来提前确定是否能从当前驻留小区中获得想要的切片业务。当服务小区选择或重选时，UE 无须提前获知目标小区支持的切片信息。NG-RAN 基站中小区支持的网络切片能力，是由基站网管系统提前预配置的。为了实现 UE 在物理移动时对切片能力的一致性支持，在网络部署阶段，UE 入网注册归属的注册区（RA）内必须支持相同的切片能力集合。在注册区域内移动时，UE 无须通过跟踪区域更新（TAU）流程更新切片能力信息。切片部署可能无法一直物理连续（某些区域支持的切片在另一区域中不支持），因此 UE 在移动到新的注册区域时，需要通过 TAU 流程更新新切片的能力信息。切片部署的不连续性典型的场景如图 4 所示，在不同区域可能不同的频点支持不同的切片能力集合。gNB1 的 F1 频点上所有的小区都支持切片 1、切片 2 和切片 3（图 4 中切片 1—3），gNB2 的 F1 频点上的小区仅支持切片 4。当 UE 支持切片 1—3 而不支持切片 4 时，如图 4 所示，UE 移动后继续驻留和使用 F1，会导致无法获得想要的切片业务，造成业务中断；因此基站需要为 UE 重选目标频点，这个功能被称为切片相关的频率优先级。具体过程为，基站通过 RRC

连接释放消息中携带 UE 应优先接入的频率列表，引导 UE 选择到切片支持的目标频点上。如果 UE 接入时初始使用的是 F1，但 UE 需要接入切片 1—3，基站 2 可为 UE 设置优先的频率 F2。基站 2 可通过核心网配置的制式率选择优先级（RFSP）配置信息，结合本地小区的切片部署信息，为 UE 选择更合适的 F2。

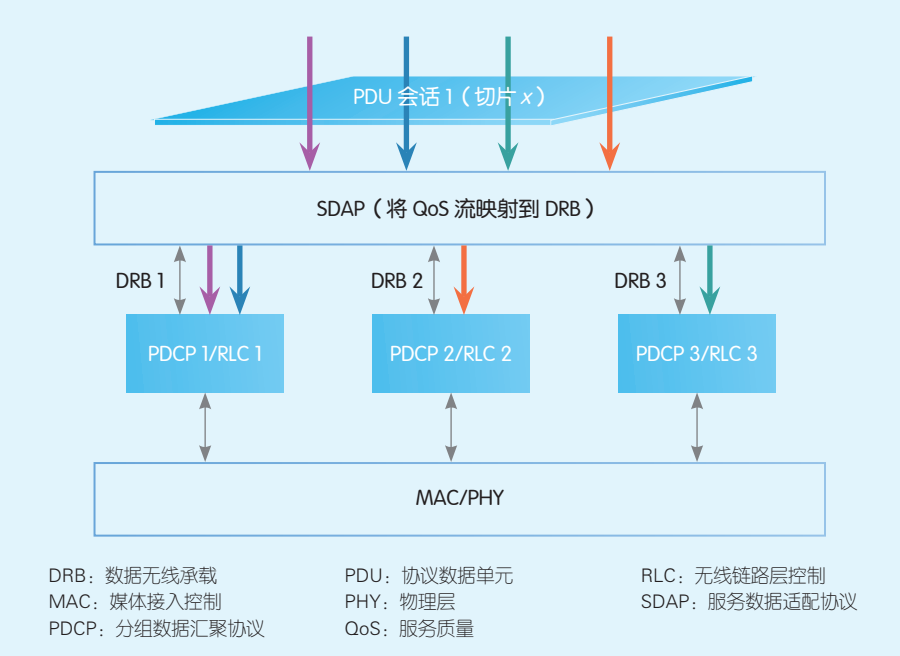
随着业务量增加，越来越多的 UE 业务会消耗掉有限的网络切片 RAN 资源，切片有可能处于拥塞状态。后续接入到切片的业务请求，将会进一步加剧设备的拥塞状态。为了解决上述问题，NG-RAN 基站可用多种手段来降低 UE 切片业务请求对基站的冲击。当 UE 支持的某切片发生拥塞时，NG-RAN 基站可以调整和切片相关的接纳控制类型所归属的接纳控制参数。UE 在发送切片业务申请之前，在 UE 侧先经过 UE 内接纳控制功能的判决。接纳控制判决决定了请求的切片业务是否可以被发起。如果判决阻止发起请求，UE 只能等一段时间后，再尝试发送切片业务申请。另一个解决拥塞的手段是，NG-RAN 基站拒绝或释放终端请求的切片业务。根据消息 5 中携带的 S-NSSAI 信息，基站可先知道 UE 请求的业务归属的切片信息。如果此时请求的切片属于资源拥塞状态，NG-RAN 基站可直接释放该 UE 刚刚建立的 RRC 无线连接。

3.2.2 空口高层用户面

NG-RAN 网络侧处理切片业务的最小粒度是 PDU 会话，单个 PDU 会话可包含最多 64 条 QoS 业务流，而这些 QoS 业务流必须从属于同一个切片；因此 PDU 会话总和唯一的 S-NSSAI 关联标识。NG-RAN 空口侧通过服务数据适配协议（SDAP）将 QoS 业务流进一步映射到一条或多条 DRB 上，进而对应到一个或多个分组数据汇聚协议（PDCP）/ 无线链路层控制（RLC）协议实体上。SDAP 可根据不同切片需求进行不同方式的映射，当某条 QoS 业务流没被显式地映射到某条 DRB 上时，它将被自动映射到默认 DRB 承载上。如图 5 所示，某 PDU 会话 1 包含 4 条 QoS 业务流，其中红色和蓝色的 QoS 业务流由于和 QoS 的参数接近，被映射到同一 DRB1 上，而橙色和绿色 QoS 业务流被分别独立地映射到 DRB2



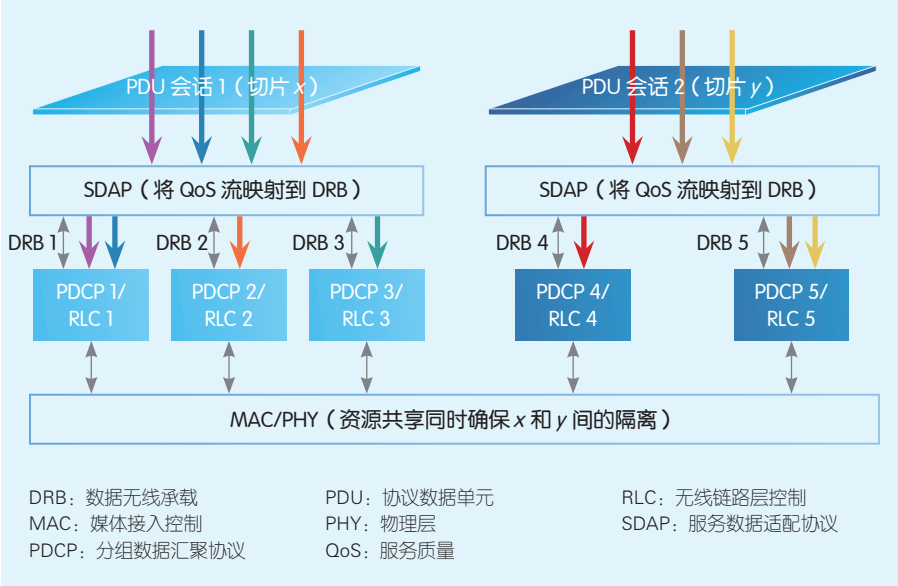
▲图 4 网络切片在不同频率上的部署示意



▲图 5 PDU 会话为粒度的单切片处理

和 DRB3 上。进一步地，SDAP 下层的 PDCP、RLC、MACPHY 协议实体将根据不同切片的需求，产生对应的功能和参数配置。例如，如果该 PDU 会话属于 uRLLC 类业务切片，PDCP 可被配置复制传输功能，RLC 可配置较高的逻辑信道优先级，自动重传请求（ARQ）的重传次数可设置为 1，MAC 层可配置基于载波聚合 CA 的复制操作，PHY 可选择低阶调制编码率编码策略（MCS）和短 TTI 等配置。再如，如果该 PDU 会话属于 4K/8K 高清视频业务切片，PDCP 可配置互联网协议（IP）包头压缩功能，RLC 可选择中等的逻辑信道优先级，MAC 可将相关的逻辑信道映射到较大的工作带宽和 Numerology 之上，PHY 可配置高阶 MIMO 和调制与 MCS 配置等。总之，NG-RAN 空口用户面的各个协议实体，须充分考虑相关切片的不同需求，进而产生合理的配置，以能达到切片服务定制化的目的。

如果单个 UE 同时连接到 2 个或更多不同的业务切片（如图 6 所示），除了上述属于切片 x 的 PDU 会话 1，另还有从属于切片 y 的 PDU 会话 2。PDU 会话 2 包含有 3 条 QoS 业务流，其中粉色的 QoS 业务流被独立地映射到 DRB4 之上，而紫色和灰色的 QoS 业务流由于 QoS 参数接近，被联合映射到 DRB5 承载之上。需要注意的是，不同切片所属的 QoS 业务流被映射到不同的 DRB 承载上，以保证充分隔离。进一步地，NG-RAN 基站内单个媒体接入控制（MAC）和 PHY 实体，须联合处理来自 2 个或多个不同切片的所有逻辑信道数据块。通过内部实现的约束机制，一方面要保证基站本地的基带计算和存储资源、射频天线和功率资源，以及空口无线资源被共享高效地使用，同时还要尽量保证不同切片业务间的隔离。例如，对应于不同切片的逻辑信道不要被映射到同一个传输块（TB）中，但可被映射到不同的 PRB 资源块中。



▲图 6 PDU 会话为粒度的多切片处理

如果单个 UE 同时连接到 2 个或更多不同的业务切片（如图 6 所示），除了上述属于切片 x 的 PDU 会话 1，另还有从属于切片 y 的 PDU 会话 2。PDU 会话 2 包含有 3 条 QoS 业务流，其中粉色的 QoS 业务流被独立地映射到 DRB4 之上，而紫色和灰色的 QoS 业务流由于 QoS 参数接近，被联合映射到 DRB5 承载之上。需要注意的是，不同切片所属的 QoS 业务流被映射到不同的 DRB 承载上，以保证充分隔离。进一步地，NG-RAN 基站内单个 MAC 和 PHY 实体，需要联合处理来自 2 个或多个不同切片的所有逻辑信道数据块。通过内部实现的约束机制，一方面要保证基站本地的基带计算和存储资源、射频天线和功率资源，以及空口无线资源被共享高效地使用，同时还要尽量保证不同切片业务间的隔离。例如，对应于不同切片的逻辑信道不要被映射到同一个 TB 中，但可被映射到不同的 PRB 资源块中。

5G NR MAC 实体，可支持多种不同的 Numerology/TTI 持续时长的混合配置，即可根据不同切片的需要，把不同的逻辑信道映射到不同的 Numerology/TTI 上，以实现对不同时频资源块的适配性使用，充分体现出不同切片服务的资源配置差异化。同一条逻辑信道，其实也可分时地映射到多个不同的 Numerology/TTI 上。从内部实现的角度看，单个 MAC 实体逻辑上可划分为 2 个 MAC 子层，上子层是不同切片间的公共功能或相同配

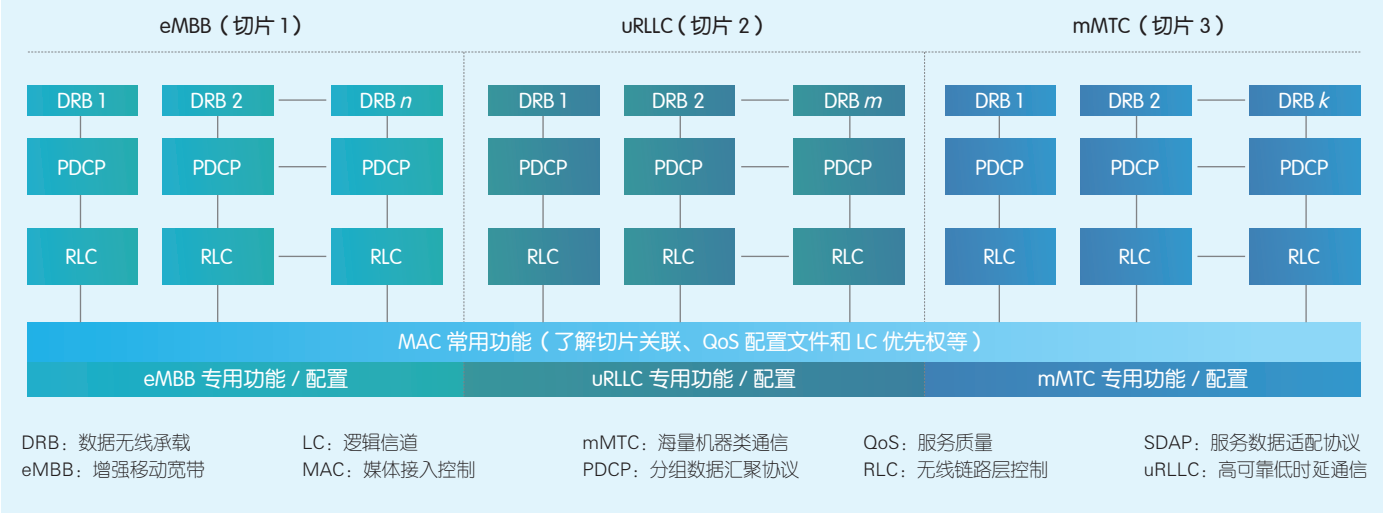
置，它们对所有切片相同；而下子层可以是不同切片所特有的功能，相关配置和处理可针对不同的切片进行定制化操作。该原理如图 7 所示，即使某 UE 同时连接 eMBB 类型的切片 1 和 uRLLC 类型的切片 2，以及 mMTC 类型的切片 3，内部也可只配置有单个 MAC 实体。

3.3 网络接口高层

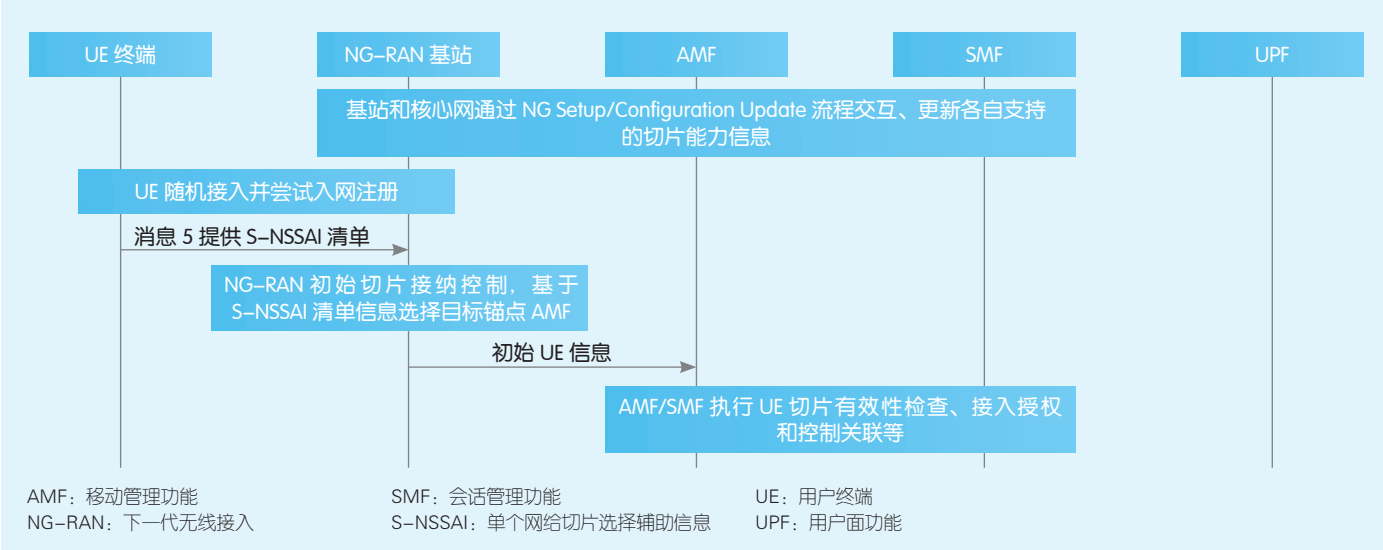
在 5G 部署组网阶段，运营商首先通过 OAM 和切片编排器系统对 NG-RAN 进行切片化的功能编排和资源预配置。为了适配不同切片的各自需求，NG-RAN 需要能为各个切片配置独立的 RRM 策略、相关功能参数集；NG-RAN 基站内的每个协议实体，须有效及时地跟踪并锁定住待服务对象的切片关联标识，进行有针对性的资源动态分配和调度，同时能够保证不同切片间的隔离处理，从而 NG-RAN 对外呈现出网络切片化功能。

编排预配置完成后，网络节点间通过接口公共流程交互和更新各自支持的切片能力信息。物理覆盖上以 TA 为最小粒度，TA 内的所有本地服务小区必须支持相同的切片能力集合。该原则同样适用于 NG、Xn 和 F1、E1 等网络侧接口，从而保证通信端到端所经历的所有网络节点都能为同一目标切片服务。

当 UE 初始驻留在某 NG-RAN 服务小区，并完成了随机接入流程的前消息 1—4 后，如图 8 所示，UE 通过消息 5 携带 S-NSSAI 清单



▲图 7 单 MAC 实体支持多切片业务并发的架构



▲图 8 网元之间切片能力协同和 UE 初始接入

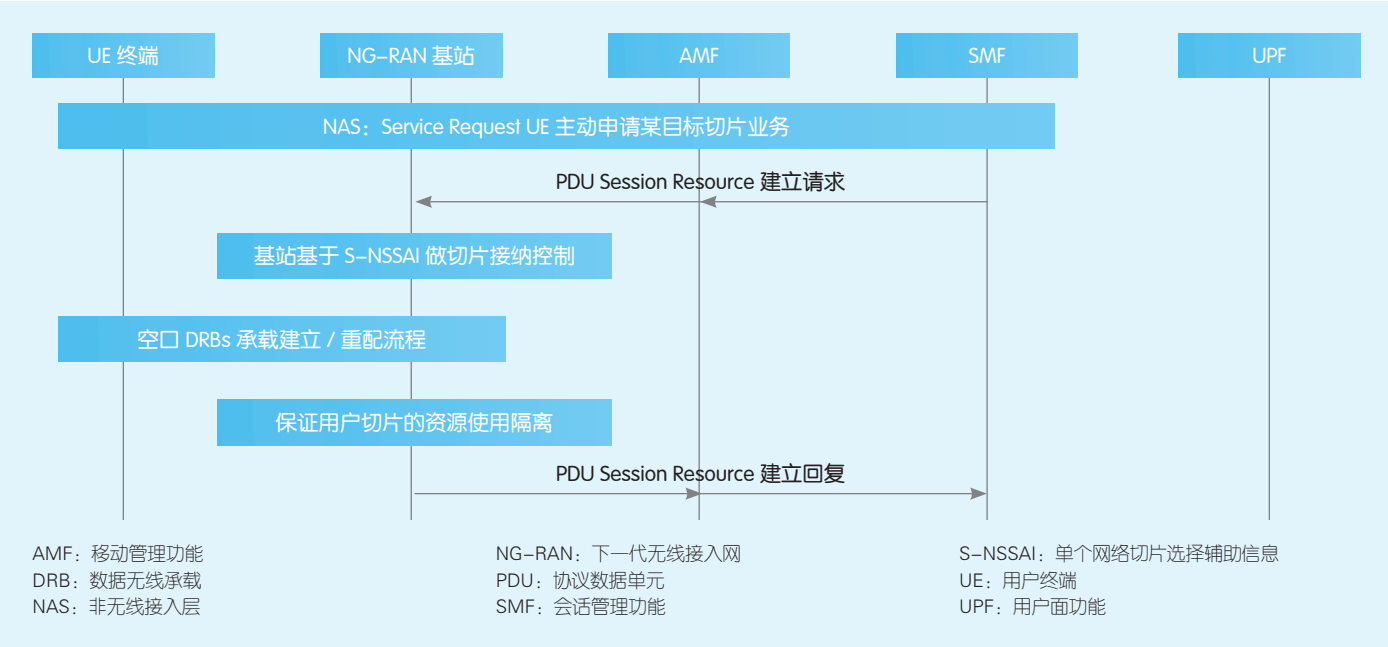
信息。NG-RAN 基站完成 UE 的初始切片接纳控制，再基于 S-NSSAI 清单信息，路由选择到目标锚点 AMF 实体（如果选择失败，则选择默认 AMF 实体），然后再向上发送初始 UE 消息。AMF/SMF 进一步完成 UE 切片的有效性检查和接入鉴权和授权等操作。对于特定终端 UE，当它成功地入网注册即选定好了锚点 AMF 和服务基站后，在当前 TA 区域内所能进行的切片业务也就被限定了。

当 UE 成功完成 5G 系统入网注册之后，会被锚点 AMF 分配唯一的 5G 辅临时终端标识（5G-S-TMSI），进行 UE 非接入层（NAS 上下文）管理。如图 9 所示，如果 UE 请求发起某切片业务，将通过 NAS 业务请求消息向当前锚点 AMF 申请。如果锚点 AMF/SMF 能支持 UE 当前申请的切片业务，则向 NG-RAN 服务基站发起 PDU 会话建立流程；如果当前 NG-RAN 服务基站也能支持 UE 当前申请的切片业务，则进一步进行本地 RRM 配置、分配空口资源等。NG-RAN 服务基站需要保证该切片业务资源和其他切片的隔离。如果锚点 AMF/SMF 不能支持 UE 当前申请的切片业务，则可以执行 AMF 重定向（相当于 UE 重新注册）；如果当前 NG-RAN 服务基站不能支持 UE 当前申请的切片业务，则可拒绝失败 PDU 会话建立请求，或触发 UE 重定向或切换到能支持目标切片业务的 TA 中 [7-9]。

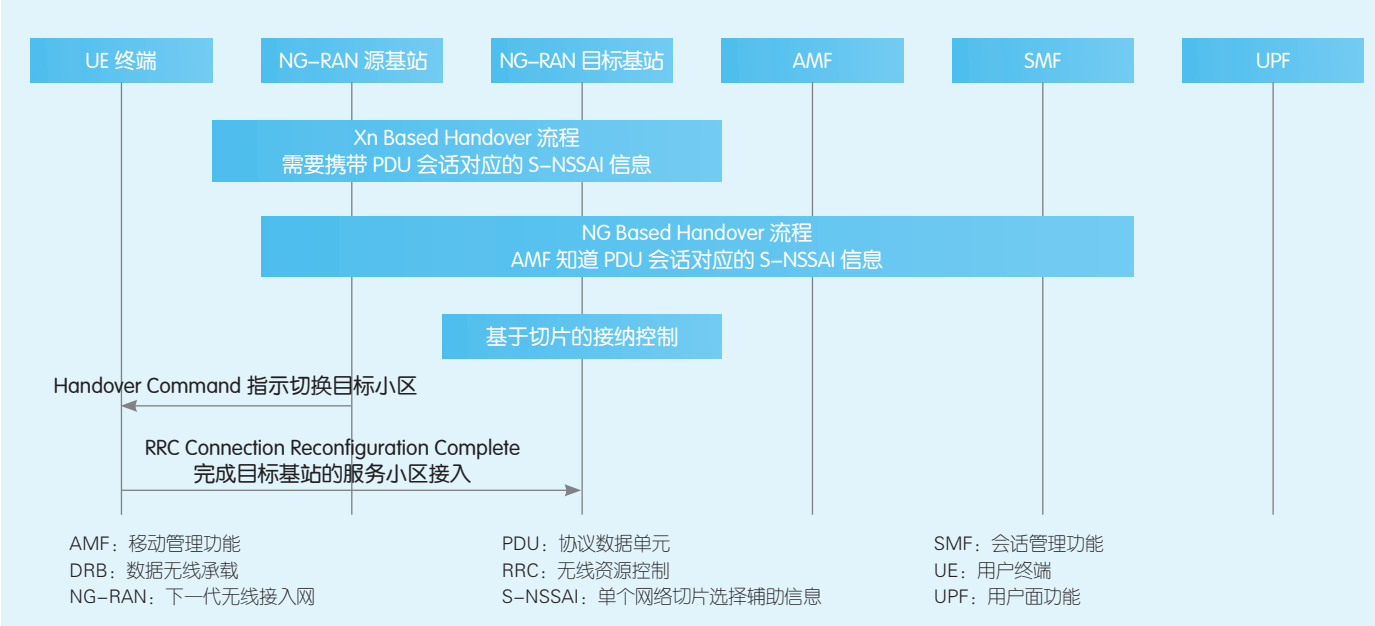
3.4 切片业务移动性支持

如图 10 所示，5G Rel-15 支持 NG 接口切换（类似于 LTE S1 切换）和 Xn 接口切换（类似于 LTE X2 切换）。随着 UE 移动或其他 RRM 策略而触发的切换，目标 NG-RAN 基站需要执行基于切片的接纳控制和资源分配预留。在跨 TA 移动的情况下，当目标 NG-RAN 基站不支持源基站某 PDU 会话所对应的网络切片 S-NSSAI 之时，目标 NG-RAN 基站将直接丢弃该 PDU 会话内所有 QoS 业务流。Rel-15 规定在切换过程中，目标基站不能重匹配 Remapping 和变更源侧的 S-NSSAI 信息，只能继承前后保持一致 [10]。

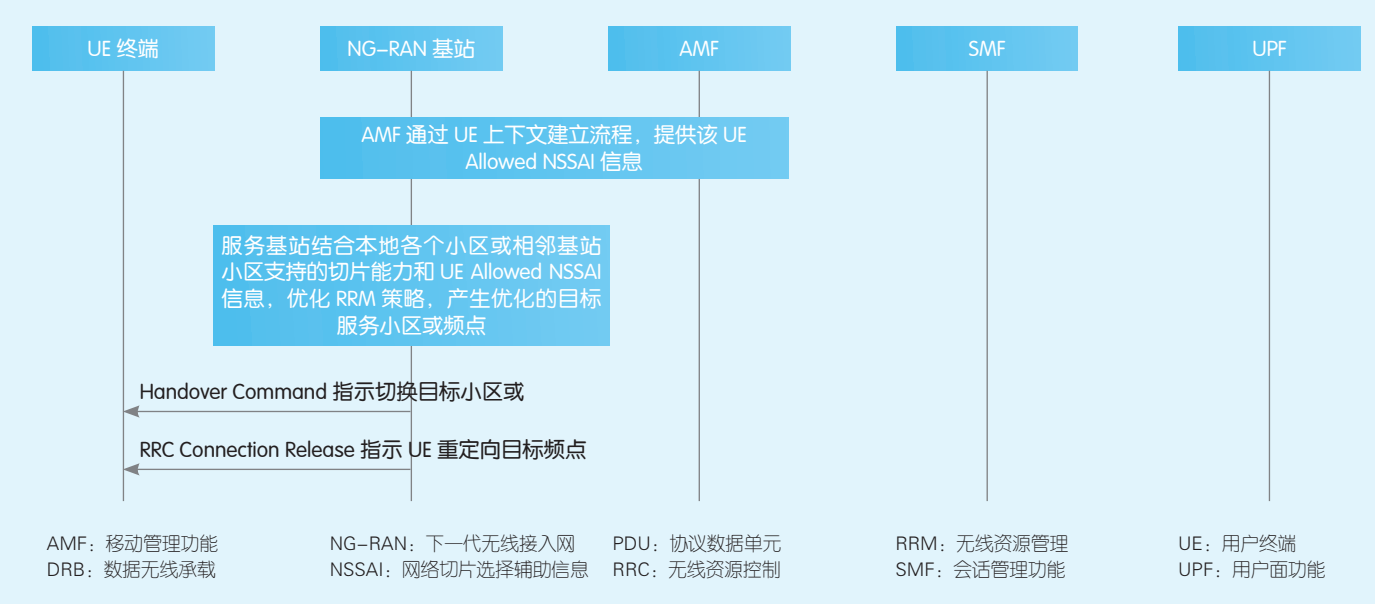
如图 11 所示，UE 在当前驻留 TA 所能发生进行的切片业务，由于受到当前服务基站和锚点 AMF 能力的限制，不等同于 UE 签约能允许申请的全部切片业务。因此，锚点 AMF 可在 UE 上下文建立流程中，发送给服务基站该 UE 允许的网络切片选择辅助信息。它表示除了当前 TA 所能提供的切片业务之外，UE 在其他 TA 环境下，还能被允许发起的切片业务。基于 UE 允许的网络切片选择辅助信息，



▲图 9 UE 切片业务的请求成功建立



▲图 10 UE 切片业务的连续性管理



▲图 11 UE 切片业务的可达性管理

服务基站可以更好地执行 RRM 移动选择策略，例如把 UE 迁移到更高优先级或更全的切片业务所属 TA 和频点之上。

4 网络切片的增强发展趋势

随着 Rel-15 网络切片功能在 RAN 侧标准化的完成，5G 网络已初步具备端到端的网络切片能力，为新商业模式的部署应用而开辟先河；但仍有一些技术方面需进一步发展和提升，这也是 Rel-16/17 网络切片功能潜在增强能做的方面。在当前 5G 不同网元节点之间，如相邻 NG-RAN 基站间，集中单元（CU）和分布单元（DU）网元节点间并不能有效地交互各个 NG-RAN 侧切片的工作状态和负荷等信息；因此尚不能在标准层面，做到跨节点针对切片业务的 RRM 策略处理优化，从而导致切片资源的非优化利用。特别在多网元节点“异频垂直”重叠部署的场景下，需要进一步规范如何整体优化利用各个

网络切片层的资源。这种优化可以通过类似自组织网络（SON）的公共流程等来实现。

在 5G 超密集网络中，网络切片能力的部署可通过多异构节点的方式来实现，即并不需要每个基站节点支持所需的全部网络切片能力；但它们的合集可包含所需支持的全部网络切片能力。终端和多个不同切片的服务，也可通过双或多连接的方式来实现。未来异构部署需求，一些运营商希望能够实现 UE 跨不同 TA 的不同基站间的双多连接操作，因此核心网需能提前感知 UE 在双多连接操作下的基站侧切片能力环境。例如，主基站支持 eMBB 类型的切片，而运营商在某些楼宇内的辅基站，仅仅部署了时延鲁棒性要求更强的 uRLLC 业务类型的切片。如果核心网无法感知到辅基站上的切片能力，在为 UE 建立 PDU 会话时是无法使用那些辅基站上的 uRLLC 切片资源，从而造成辅基站相关切片资源的闲置；或者为了能使用辅基站上的切片资源，必须把终端 UE 强行切换到辅基站之上，进而触发切换和 TAU 流程。UE 在移动过程中，只有进入到辅基站的覆盖范围内，才可能使用 uRLLC 切片业务资源，在辅基站的覆盖范围之外仍然是无法使用该切片业务的。因此在下一阶段工作中，可考虑如何让核心网能实时地提前感知到辅基站和主基站上不同的切片能力组合，进而更好地让 UE 利用网络不同 TA 下的切片资源。

如前所述，NG-RAN 基站可利用 PDU 会话中 S-NSSAI 信息对用户业务数据做到隔离区分，但 NG-RAN 控制面资源的隔离还需进一步增强。按照目前的设计，RAN 控制面并不要严格地区分切片。例如，当大量用户业务请求达到基站时，需要考虑如何解决高优先级的用户切片请求优先被服务的需求。因此这在空口 RACH 资源配置和使用优化、寻呼 Paging 资源配置和使用优化、逻辑信道优先级的使用中，都可进一步增强基站控制面对 RAN 切片的差分处理和感知隔离能力。对于一些特殊的电信垂直行业，如交通控制、安防救灾、油田矿山等，某些应用切片对业务数据安全私密性和资源保障性有更强的要求。这需在共享网络基础设施和充分利用网络资源的前提下，考虑对切片资源的专用隔离，进一步提升 UE 对这些特殊切片能力的及早感知获取和接入尝试。例如，专有切片 1，2 均是对切片资源量和保护有更强要求的，因此可给切片 1，2 各自静态预留一部分专用资源，同时再共享其他的公共资源。这样既能保证这些切片业务的特殊要求，保障业务能力，又能提升网络资源的利用率。其次，在空口通过系统广播机制，进一步提升 UE 及早感知特殊切片的业务能力，可进一步降低 UE 接入和业务延时，从而提升用户使用切片的效率和感受。

参考文献

[1] 3GPP. Feasibility Study on New Services and Markets Technology Enablers – Network Operation, Stage 1 (Release 14)[S]: 3GPP TR 22.864. 2018
[2] 3GPP. Study on Architecture for Next Generation System (Release 14)[S]: 3GPP TR 23.799. 2018
[3] 3GPP. Study on New Radio Access Technology Radio Interface Protocol Aspects (Release 14)[S]: 3GPP TR 38.804. 2017
[4] NGMN Alliance. Description of Network Slicing Concept, Version 1.0[R]. 2016
[5] 3GPP. System Architecture for the 5G System (5GS) (Release 15)[S]: 3GPP TS 23.501. 2018
[6] 3GPP. NR; Overall Description; Stage-2 (Release 15)[S]: 3GPP TS 38.300. 2018
[7] 3GPP. Clarification on Network Slicing in RAN[S]: 3GPP RAN3, R3-161106[S]. 2016
[8] 3GPP. Key Principles for Support of Network Slicing in RAN: RAN3, R3-161751[S]. 2016
[9] 3GPP. Principle of Network Slicing for 5G: RAN3, R3-161888[S]. 2016
[10] 3GPP. Resource Management Between Slices: RAN3, R3-162463[S]. 2016

作　者　简　介

杨立，高级工程师，中兴通讯股份有限公司算法部技术预研资深专家，担任 3GPP 5G 规范协议 TS38.414 主编；长期从事 3GPP 标准 3G/4G/5G 网络高层技术的研究，涉及 ICT 生态战略、移动网络架构演进、节点功能和接口流程标准化、网络大数大维智能化、卫星通信和垂直行业应用等；出版个人专著 1 部，撰写 3GPP 国际标准会议文稿 1 000 余篇（超过 300 篇被采纳），拥有授权专利 40 余项。

李大鹏，中兴通讯股份有限公司算法部技术预研高级系统工程师；长期从事 3GPP 蜂窝产品开发及无线关键技术研究工作，涉及移动网络架构演进和流程增强、SON 智能化和大数据利用、网络切片等；撰写 3GPP 会议文稿 500 余篇，拥有 20 余篇授权专利，发表论文 2 篇。