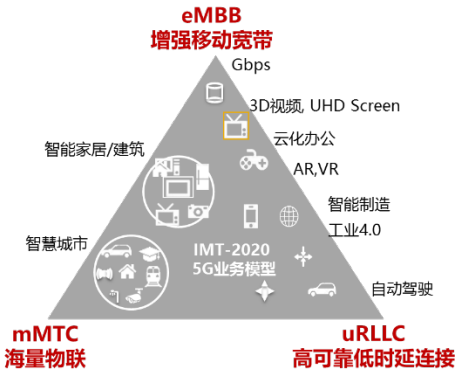
5G承载灵活切片技术保障垂直行业差异化SLA

摘要：5G新技术使能垂直行业应用，网络切片是保障业务SLA的基础。5G承载的信道化子接口+FlexE技术可以提供垂直行业业务SLA在5G承载网中端到端的保障。

5G连接新时代，一个数字的变化，一个时代的开启。移动通讯技术从3G到4G，再演进到5G，聚焦点从人与人的连接，到人与物的连接，再到物与物的连接。移动业务也从3G、4G时代的2C（个人用户）场景自然延伸到2B（垂直行业）场景，也就是说5G开启了移动通信技术在垂直行业广泛应用的新时代。从标准上看，3GPP定义了5G三大业务场景：

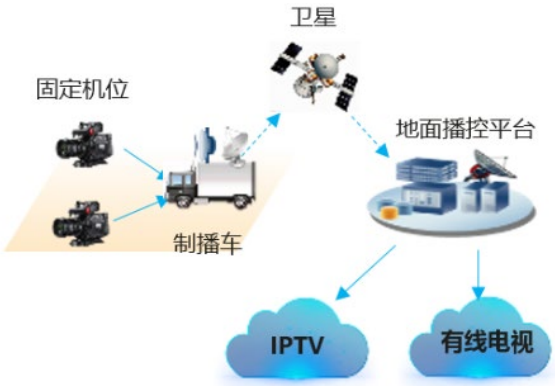


其中eMBB要求把单用户的接入峰值速率提到1Gbps，uRLLC要求业务端到端超低时延1ms，mMTC要求在每平方公里的连接数可以达到一百万。这三大业务场景的定义是由5G诸多新技术做为实现支撑，比如Massive MIMO，新的空口编码，控制面和转发面分离，核心网分层次部署等，其中网络切片技术是实现5G时代业务隔离，端到端超低时延的基础。

网络切片是3GPP标准中的一个概念，其历史可以追溯到R13/R14，并将在5G时代逐步完善和商用。网络切片利用网络功能虚拟化（NFV）和软件定义网络（SDN）将单个物理网络划分为多个虚拟网络，其中一个切片代表一个独立的虚拟化的从无线接入、到承载网，再到核心网的端到端网络，以满足多种业务场景对网络的不同需求。本文从承载网的角度，阐述信道化技术和FlexE技术如何实现承载网切片。

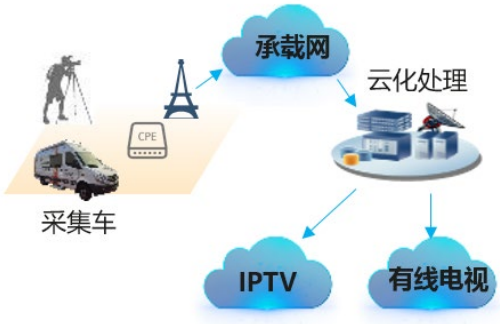
首先我们从三个垂直行业应用对网络的需求来看为什么需要网络切片。

第一个垂直行业应用是高清视频直播。如下图所示：



高清视频直播一般都是由事件驱动，比如两会、春晚外场、体育赛事、现场演唱会、游园会，或其他突发事件。视频清晰度要求高、突发临时性和场地不确定性是这类业务的主要特点。传统做法是在直播现场部署一台装备着制作播放设备和卫星通讯设备的制播车，现场编辑由多个固定机位摄像机拍摄的视频，制作成直播节目并通过卫星回传到地面播控平台，然后再分发到不同的直播网络中。

那么在5G时代，高清直播业务是怎么做呢？以今年央视春晚外场直播为例，如下图所示：

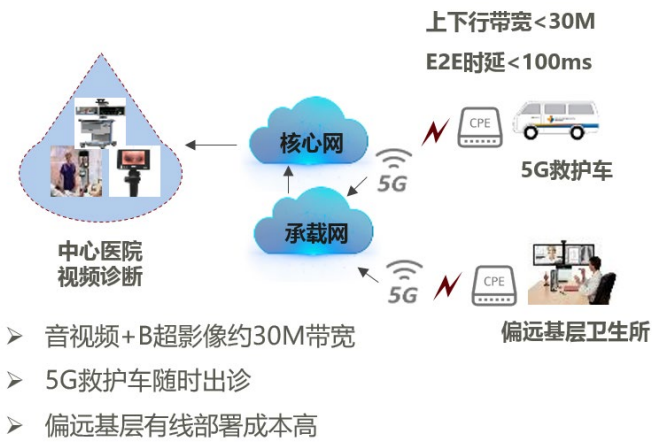


直播现场不再部署专用的制播车，活动机位摄像机（摄像师或视频采集车）与CPE连接，将拍摄到的视频通过5G基站经承载网/核心网回传到制播中心进行云化处理，然后再分发到不同的直播网络中。活动机位可以灵活移动，从不同角度捕捉到更多的视频信息，制播云化处理可以节省数百万的车载制播设备和卫星通讯设备。

高清视频直播业务对5G网络的带宽需求主要是多路4K视频上传（每路约30Mbps），时延要求是网络端到端小于30ms。同时由于直播节目的实时性要求，高清视频业务在网络中一般由UDP协议承载。而网络中的丢包现象会导致UDP承载的视频业务出现花屏现象，如下图：

目前的直播视频质量标准要求每4小时只能出现一次花屏，也就是说要求网络的丢包率小于$1.1*10^{-6}$ 。

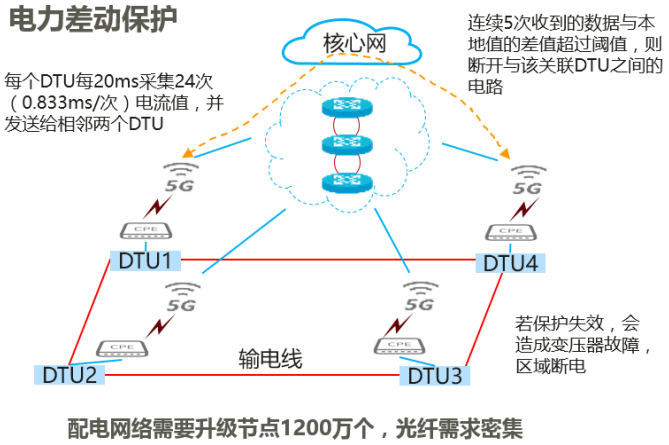
第二个垂直行业应用是远程医疗，如下图所示：



5G救护车在现有救护车上装备CPE和音视频系统，随时现场出诊，救护人员通过音视频系统将现场病人的情况实时回传到中心医院，主任专家远程指导医疗救护；偏远基层卫生所，部署光纤成本高，用5G网络将诊室的音视频和B超等影像回传到中心医院，各科专家集中会诊，病人不用来回奔波。

远程医疗业务对5G网络的带宽需求是音视频+B超影像约30Mbps，业务端到端时延小于100ms。同时，远程医疗要求视频图像质量必须清晰，不能出现花屏，因此一般采用TCP协议承载。如果网络稳定性不好，也就是网络时延抖动大，TCP协议承载的视频容易出现卡顿现象。

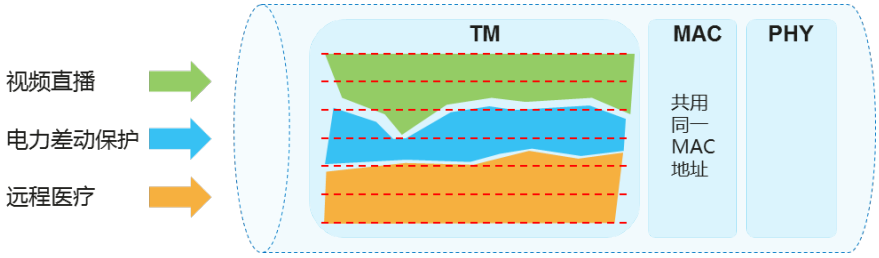
第三个垂直行业应用是电力差动保护，如下图所示：



电力的配电网络由配电站和输电线组成，一般配电站里面会有输电设备、变压器、配电保护设备(DTU)等。如果全网按照有线连接的方式部署差动保护，光纤需求量大成本高。按照目前差动保护业务的要求，每个DTU每隔20ms采集24次（0.833ms/次）本地电流值，通过5G网络发送给关联的DTU。如果一个DTU连续5次收到的电流值与本地的差值超过阈值，则认为发送端有故障，断开相连的电路以保护设备和供电安全。从试验站点测试结果上看，5G网络端到端时延15ms，带宽5M，即可满足差动保护业务要求。

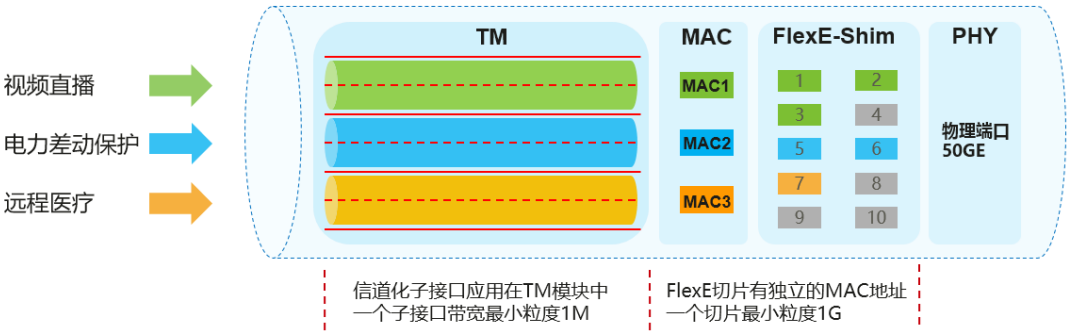
由上述三个行业应用案例可见，视频直播要求大带宽，低丢包率；远程医疗要求网络稳定，提供确定性时延；电力差动保护要求超低时延。这些业务流量同时通过一个5G承载网传输，如何保证它们要求的不同SLA（服务等级约定）？

现在的移动承载网是基于以太/IP技术，以统计复用和存储转发为主要特征，用VPN+QoS技术为不同业务提供不同的SLA。VPN技术实现软隔离，业务流量在虚拟网络中传输；QoS技术通过流量监管/整形，拥塞管理/避免等基于共享缓存队列调度的机制实现不同业务的差分服务。如下图所示：



视频直播、远程医疗和电力差动保护三个业务同时在一个承载网管道里传输数据，绿色、蓝色、橙色分别代表这三种业务的VPN流量。承载设备内的TM模块（流量管理）负责不同业务的QoS处理，红色虚线代表不同的QoS缓存队列。可以这样简单的理解：QoS技术在一个管道中划分出不同的车道，业务流量根据优先级，在相应的车道中传输。每个车道代表一定大小的缓存队列，车道分割线是虚线，在QoS优先级调度机制的作用下，业务流量可以变更车道。由于缓存资源和MAC地址对各个业务来说都是共享的，当流量突发时，各业务间互相抢占资源，不可避免的会产生拥塞现象，业务的SLA也就无从保证。

为了实现5G标准的网络切片规范， 5G承载网采用信道化子接口+FlexE的灵活切片技术，满足5G时代垂直行业严格的SLA要求。如下图所示：

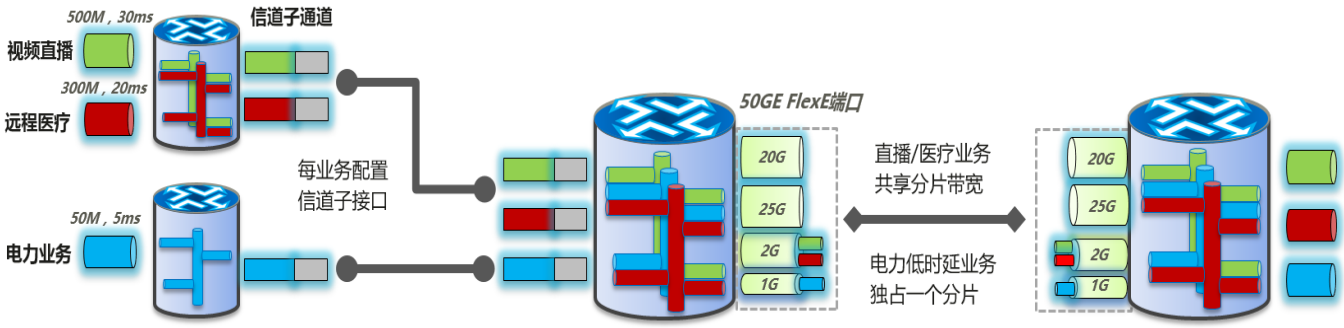


信道化子接口技术作用于承载设备的TM模块中。根据业务的SLA要求，为每个业务分配相应的硬件缓存资源，相当于在TM模块中为每个业务划分车道，业务车道之间是实线，业务流量传输过程中不能并线换车道，业务为车道之内是虚线，采用QoS技术进行缓存调度。业务车道的最小粒度是1M。缓存资源的独占，使得各个业务在TM模块中得到硬隔离，有效避免流量突发时，出现各业务争抢缓存资源，导致业务SLA劣化。

FlexE（灵活以太端口子通道）技术是在承载设备的MAC和PHY层之间定义一个FlexE shim层，对物理端口带

宽进行基于时间片的切分。即在大管道物理端口上通过FlexE的时隙复用划分出若干个子通道端口，把这些子通道端口切片划分网络切片的不同切片中，通过硬件的时隙复用实现各个切片之间的业务在转发层面上完全隔离。比如50GE端口，分为10个时隙，每个时隙带宽为5G，按照5G颗粒倍数进行带宽分配。目前华为设备的FlexE技术可以做到1G的时隙粒度。每个FlexE切片都有独立MAC地址，相当于每个切片都是一个物理端口，切片之间是硬隔离互不影响。

将信道化子接口和FlexE技术结合，灵活配置，可以保障承载网端到端的业务SLA。假设视频直播要求带宽500M，时延30ms，丢包率10^{-6}；远程医疗要带宽300M，时延20ms，抖动1ms；电力业务要求带宽50M，时延5ms。如下图所示：



在承载网接入层的低成本设备上，为每个业务按照SLA的要求分别配置一个信道化子接口，例如视频直播信道CIR（承诺保障速率）500M，远程医疗信道CIR 300M，电力业务信道CIR 50M。业务流量经各自的信道子接口传输到汇聚层设备。汇聚层设备根据各业务的SLA，将业务信道子接口跟相应的FlexE切片绑定，比如将要求大带宽的视频直播和远程医疗绑定到一个2G的切片，将要求低时延的电力业务绑定到一个1G的切片。业务信道子接口嵌套在FlexE切片中，业务流量在硬管道中传输，直到承载网络出口。

综上所述，信道化子接口+FlexE是5G承载网在网络切片领域的实现技术，分割独立缓存资源，物理端口基于时隙切片，两者灵活组合，业务流量硬隔离，实现业务从承载网络入口到出口的端到端SLA保障。