ZTE中兴

IP 网络未来演进技术白皮书 2.0

—— 开放服务互联网络

IP 网络未来演进技术白皮书 2.0 —— 开放服务互联网络

版本	日期	作者	备注
V1.0	2021/05/22	ZTE	新建
V2.0	2022/08/28	ZTE	更新—提出开放服务互联网络解决方案和三大关键技术：服务感知网络（SAN）、增强确定性网络（EDN）、网络内生安全

重要贡献单位：

中国信息通信研究院
中国移动研究院
中国电信研究院
中国联通研究院

目 录

图

1 前言

IP 技术作为互联网的技术基础，支撑着互联网的蓬勃发展。目前，IP 网络已经逐渐从 IPv4 向 IPv6 过渡。在未来 10 年乃至更长时间，IPv6 技术将如何演进，是业界的热门课题。2021 年 6 月，中兴通讯联合中国信息通信研究院及移动、电信、联通三大运营商（以下简称三大运营商）联合发布《IP 网络未来演进技术白皮书》[01]（以下简称白皮书 2021），提出了 IP 网络技术未来仍将平滑演进的预判。

白皮书 2021 基于未来网络面临的需求和痛点，提出三大愿景：万维互联、算网融合、精准网络。这三大愿景与国际电信联盟网络 2030 焦点组提出的未来网络 12 个应用场景相对应[02][03]。基于这些愿景和应用场景，白皮书 2021 阐述了未来网络五大关键技术需求：网络确定性需求、内生安全需求、移动性管理需求、算力感知与调度需求、多语义多标识需求。

白皮书 2021 提出，传统 IP 技术所遵循的设计原则，如端到端原则、分层解耦原则，总体上是非常成功的。由于端到端和分层解耦的架构设计，极大降低了互联网业务的创新门槛，增加了业务部署的便利。同时，这种架构提升了网络的扩展性、健壮性和适应性，有利于跨越不同类型的底层介质，实现全球互联的目标。但按照这两个原则设计出来的 IP 网络架构，存在的最大问题就是业务和网络处于相对割裂的状态，导致绝大多数业务只能按照“Best Effort”的模式运行。

随着互联网业务向纵深推进，尤其是产业互联网的发展，业务和网络的割裂状态越来越不能满足需求。白皮书 2021 提出未来 IP 网络的目标是实现业务和网络的协同：持续增强 IP 网络能力，为共性化需求提供内生的、网络主导的解决方案，为业务和应用提供支撑。为此，白皮书 2021 提出了改进后的未来 IP 网络设计原则：服务化网络赋能的端到端原则，智能控制面支撑的瘦腰模型；以及基于 IPv6 演进的未来网络参考架构。

在白皮书 2021 发表之后（2021 年 6 月），中兴通讯与信通院、三大运营商等业内伙伴于 2021 年 9 月共同启动了“未来网络战略合作项目”，持续深入推进未来 IP 网络的研究。

近两年来，算和网的融合成为行业发展的重要方向。随着数字经济成为国家战略，算力、5G 网络、光纤网络等作为数字经济的基础设施，重要性日益凸显。三大运营商也紧抓算力时代大机遇，加快推进数字基础设施融合升级，把云网融合/算网融合发展提升到了公司战略的高度[04][05][06]，积极由通信服务运营商向数字服务提供商转型升级。

在互联网整个技术架构中，通常来说算对应着上层的应用，网对应着底层的连接，IP 技术作为中间层，起到承上启下的枢纽作用。算和网的融合必然对应着应用和网络的融合，在这个过程中，IP 网络将起到什么作用，或者说 IP 网络技术将如何演进以适应算网融合的需求，是未来 IP 网络演进的最重要课题。因此，中兴通讯在未来 IP 网络的研究中，将白皮书 2021 所提的未来网络三大愿景中的算网融合作为最核心的场景，另外两大愿景（万维互联和精准网络）作为辅助和支撑。

本白皮书在白皮书 2021 提出的未来 IP 网络目标、设计理念、参考架构的基础上，以“算网融合”为核心场景，从“主机互联”到“服务互联”，提出了未来 IP 演进的具体技术方案，即：开放服务互联网络。

本白皮书第二章首先分析了数字经济和产业互联网的发展，带来了算网一体基础设施服务的需求，由此给网络运营商带来了为整个互联网提供一体化算网共性服务的机会。针对一体化算网服务的技术架构如何实现，第二章分析了一体化算网服务的技术需求，并且提出了在云原生架构基础上进行扩展，以网络为中心的“开放服务互联网络”架构。通过分析现有 IP 技术支撑开放服务互联的不足之处，基于 IP 网络平滑演进的思想，提出了开放服务互联网络的整体方案。

第三章描述了开放服务互联网络的关键使能技术，包括服务感知网络（SAN）、大规模确定性网络和网络内生安全技术。

第四章描述了 SAN 的应用示例，包括云游戏场景和人脸识别场景下的 SAN 实例化方案。

第五章对整个白皮书做了总结。

2 面向一体化算网服务的开放服务互联网络

2.1 一体化算网服务的愿景

数字经济的发展，产业互联网的应用，对于算力和网络都提出了很高的要求，由此催生出了“算力网络”这种新的技术形态和运营模式。算力网络的未来发展路径，将是算和网逐步融合的过程，初期是运营层面的融合，中期是控制面的融合，后期目标是实现技术架构和设备形态的融合。

算网深度融合有两大驱动力，一是需求侧，实现算力和网络的协同调度，满足业务对算力资源和网络连接的一体化需求。比如，高分辨率的 VR 云游戏，既需要专用图形处理器（GPU）计算资源完成渲染，又需要确定性的网络连接来满足 10 ms 以内的端到端时延要求。二是供给侧，依据 IT 和 CT 技术融合演进的趋势，传统网络设施逐渐向融合计算、存储、资源提供的新型智能化设施转变，而云计算也由中心向边缘、超边缘以及端计算等分布式泛在演进，形成全方位多维立体的算力布局。借助于网络设施天生的无处不在的分布式特点，算网深度融合可以助力算力资源也实现分布化部署，满足各类应用对于时延、能耗、安全的多样化需求。

算网深度融合的发展趋势，使得网络运营商有成为“算网一体基础设施提供者”的重要机会。从国家战略看，“十四五”新基建战略提出构建以新一代通信网络为基础、以数据和算力

设施为核心、以融合基础设施为突破的新型数字基础设施体系。基于算网一体设施向全社会提供互联网共性服务，有利于加快业务创新、降低业务部署成本、提升资源利用率，促进整个社会的数字经济发展。消费互联网时代，云服务商是云计算基础设施的提供者；到了产业互联网的时代，网络运营商最有可能成为算网一体基础设施的提供者。这不仅是因为运营商在网络基础设施和数据中心等算力基础设施的综合能力上具有优势，而且运营商也是更值得信任、更有经验的国家级关键设施运营者。

业内曾有多方尝试互联网共性服务的运营模式。比如，运营商曾提出“智能管道”的理念，试图把网络功能开放出来供业务调用，但只有短信等少数功能的开放取得成功；云服务商推行 PaaS、SaaS 平台已有多年，但云服务商的平台属于封闭式、烟囱式的单方算力系统，不对外开放，且以中心算力为主，不能满足泛在计算、网络 SLA 等需求。

本文提出的一体化算网服务，定位于社会级算网能力开放公共平台，不仅向各类互联网应用以及社会和产业应用提供 AI、音视频、大数据等算力共性服务，也统一接纳全社会多样性算力(公有、私有)，同时也提供包括确定性、内生安全等在内的网络基础共性服务。这一公共服务定位意味着需要将计算由当前封闭私有的资源向开放公共资源转变。

与已有的各种平台型服务相比，本文提出的一体化算网服务具有以下三个特征：

1．全面的共性服务

既包括算力类服务，如 IaaS、PaaS、SaaS、FaaS 等多种算力服务模式，也包括网络连接类服务，如确定性连接服务，同时能支持一体化算网服务，为应用同时满足算力和连接的需求。还可以进一步扩展其他互联网共性能力的服务（比如安全能力、AI 能力、大数据等）。

2．开放的泛在服务

如果说云仍带有某种程度的私有和有限域的属性，而算则必须完全脱离私有和封闭域的限

制成为向整个社会和产业开放的公有资源。一体化算网服务是跨运营商的，统一定义的，面向整个社会的开放的服务，在这一点上，电信行业有自身的优势，不仅有成熟的标准组织和体系，还有互联互通的文化和传统。为了满足各类应用、各行各业对于低时延、可靠性、安全性的需求，一体化算网服务是泛在提供的，即各类算力和网络服务无处不在、按需提供。服务使用方存在于各用户终端、企业网关以及各类产业社会端的应用，服务提供方包括但不限于中心云、区域云、边缘云、终端等多种混合场景以及 CPU/GPU/NPU 等多样性异构算力的提供形式。

3．有质量保障的服务

网络不仅能感知业务的 SLA（Service Level Agreement）需求，提供高质量、高可靠的性能保障，而且还能感知应用对算力的需求，以及算力资源的部署情况和性能状态，形成算力和网络的高度协同和融合调度与编排，确保算力和网络资源的最优化和最高效利用。

一体化算网服务将使运营商从目前的“管道型”服务向计算、网络、存储一体化的新型基础设施服务演进，构建运营商自身发展第二曲线，为收入增长注入新的动能。

2.2 一体化算网服务的技术需求

一体化算网服务在技术架构上如何实现？互联网从计算机网络发展而来，在设计之初就让算和网处于相对分离的状态。近年来随着 SDN、NFV 的发展，算和网的技术有了部分结合，但离算网融合的目标还远远不够。在白皮书 2021 中，我们提出未来网络的目标是实现应用和网络的协同，这与算网融合的目标是相一致的。

一体化算网服务的技术架构，要满足算和网两方面的技术需求。其中，算力的需求是核心，算力技术的发展特点决定了一体化算网服务的整体技术性需求。当前在算力领域主要有两方面的发展趋势，一是泛在化的算力供给，二是基于云原生的算力服务框架。

泛在算力有两个含义，一是算力分布由集中式云向分布式云、边缘计算的方向发展。分布式云不仅能以集群的方式打破中心单点式算力堆叠的性能限制从而提升整体算力的规模和极限，而且通过立体全覆盖模式更靠近用户，满足产业和用户对于低时延、高安全、敏捷响应的服务需求。二是算力类型由 CPU 为主，向 CPU/xPU/DSA 软硬结合的异构多样性算力演进，随着云向大规模计算密集型演进，原有以 CPU 作为通用适配器的软件定义+缓存/IO 的算力构建模式遇到了严重的性能瓶颈，单纯依赖芯片堆积的集中式算力模式已无法满足不断增长的算力需求。因此，根据算法任务的不同，可利用多样化异构硬件加速处理芯片卸载和承担不同的计算任务。目前普遍采用以计算为中心的 CPU+DSA 通用异构计算卸载架构，和以数据为中心的 DPU+DSA 业务加速架构的 2 条技术路线。CPU+DSA 方式适合于灵活性和计算性强但对于时效性要求不严格的中心云计算场景，而 DPU+DSA 方式则比较适用于大流量、高交互性业务的加速和卸载，以边缘计算为主要场景。

算力服务框架关系到如何能够为用户提供高效的算力封装和算力调度。当前算力服务以容器化和微服务化的云原生架构为发展方向。云化算力资源的管理和编排单元从原来的裸机、虚机，逐渐向以容器、微服务、FaaS 无服务器等更细颗粒度任务级别的算力服务模式发展。容器化使得算力资源不仅能够实现灵活弹缩和扩容，而且可实现秒级甚至毫秒级的响应，这对于高交互的算力服务至关重要。而微服务则使应用内部功能的组织架构也能够像容器一样实现灵活组合和弹缩。云原生是在容器和微服务基础之上，进一步规范使用接口，屏蔽底层异构算力资源，提供弹性、多样化、分布式统一的算力服务，使得应用通过云原生架构可以接入任何合适的一致性资源服务而无需考虑底层资源的差异性以及分布性，从而实现应用的轻量化。

云原生作为下一代云计算的架构“内核”首次拉齐了应用和云之间的交互和重组，成为算力服务提供、使用和编排的事实标准，打通了应用到算力的端到端交付。应用基于云原生改造

和部署使业务能以更敏捷、更低成本的方式聚焦价值开发，大幅提升云的易用度和效能。因此，未来一体化算网服务的技术架构，必须以云原生架构为基础，构建统一的分布式算力供给技术体系，屏蔽底层异构算力差异，在公共平台上实现对算力的发布、交易和流动。

但是云原生服务架构毕竟是从集中式云计算模式发展而来的，尽管已经向分布式云的方向发展，但离一体化算网服务所需要的“全面的、开放泛在的、有质量保障”服务的目标还有差距。主要体现在如下方面：

1．无法提供全面的服务。目前的云原生服务架构只是纳管了算力类服务，而无法纳管网络类服务，比如确定性连接、移动性管理、内生安全等。同时也无法满足未来产业应用所需要的算网一体需求。

2．无法提供开放的服务。当前云原生架构本质仍然属于云服务商自身技术体系内的封闭平台解决方案，缺乏一个脱离封闭孤岛体系的公共服务视角，这显然无法满足未来数字经济所需要的开放性基础设施平台的需求。

3．无法实现全网泛在的调度。作为云原生核心的微服务架构，目前主要适用于云内的服务，即云内的应用和共性服务之间的调用架构。而一体化算网服务是跨端、网、云的泛在服务，服务使用方和服务提供方都可能是云、网、边、端的任何一方。虽然目前也有分布式云的方案，但从外部访问云内服务以及集群内外的互通只能通过 API 网关和 7 层域名解析的方式，存在效率低、响应慢的问题。在服务提供方是全网泛在的情况下，最佳的服务调度点不应该在云的入口，而是应该在网的入口，也就是最靠近服务使用方的网络边缘位置。

4．无法提供有质量保障的服务。目前的云服务是以 Overlay 的方式叠加在底层的承载网络上，无法感知和调度 Underlay 网络的资源。因此无法提供类似确定性时延、带宽的有质量保障的网络连接服务。

5．云网分治的现状无法实现算网协同资源调度。当前网络架构云、边、端、网生态体系相对独立，业务、算力、网络等资源各自独立规划和治理，无法整合各类 ICT 基础设施为用户提供一致性的服务，无法满足全局资源优化利用的目标。

解决以上问题关键是在技术架构上实现算和网的融合。目前有两种方案路线可以选择：一种是云为中心，网络向云管理系统开放自身状态信息。另一种是网络为中心，网络整合和集成算、存、网各类资源，向应用提供特定的数据处理任务(比如图形的渲染、视频的编解码等)。这两种方式各有优劣。从构建泛在、开放、共享的算网一体基础设施的最终目标看，以网络为中心的技术路线更加合适。网络天然就是广覆盖、分布式模式，为算网资源的泛在部署提供了很好的基础；网络设备更靠近用户，可以更快地响应需求；网络技术领域在开放性、标准化、互联互通方面一直都有很好的传统。本白皮书重点描述以网络为中心的技术路线。

因此，未来一体化算网服务的技术架构将在借鉴云原生架构的基础上，结合泛在调度、质量保障等目标进行扩展。我们提出一体化算网服务架构的三大技术需求：

2.2.1 构建端、网、云一体化泛在算力供给技术体系

以容器和微服务作为异构泛在算力互联的基础底座，构建多云、算网一致和云网端一体化的分布式算力供给体系。统一纳管端网云所有底层计算资源、存储资源、网络资源，为用户提供统一的公共算力资源。

- 泛在算力使能，支持云边、多云、混合多种算力部署形式，算力资源可灵活部署、平滑迁移，构建多样、泛在的立体多维分布式算力供给布局。
- 异构计算使能，除构建和纳管 CPU/DPU+DSA 等多样性 IaaS 异构算力资源之外，还提供面向应用的算法 FaaS/Serverless 等更敏捷、弹性的精细化算力体系使能。

- 调度编排使能，将分布的算力联合网络、存储等多要素资源按需、实时、弹性动态调度和编排，实现连接和算力的全局优化和一致的用户体验。

2.2.2 以云原生构建端、网、云统一的算力服务框架

以云原生算力服务为核心，通过定义共性的算力服务满足差异化的业务场景需求，与特定应用和场景松耦合。通过规范统一的跨端网云算力服务接口，使用者能够在任何地方、任何时间和任何场景下使用统一、无差别、有质量保障的算力服务，而无需关心算力服务的位置和类型。

- 多层次算力服务模式，支持 IaaS、PaaS、SaaS、FaaS 等多种共性算力服务模式。
- 屏蔽异构算力差异，以云原生作为异构算力的标准化接口，实现货架式无差别算力使用，用户只需通过统一服务接口调用算力资源，而无需关心底层异构算力类型。
- 分布式服务治理，支持泛在和异构算力服务的统一度量、发现、注册、鉴权、计费和交易，屏蔽底层的多样性和差异化，以及算力服务的统一版本、交付、观察、测量和可视化管理等。

2.2.3 构建以网为中心的扁平化算网服务和资源供给

网络作为算网服务的入口和底座整合算力供给和算力服务，提供可靠、安全、精准的算力连接，保证算力有序、合理和高效的调度，确保算力服务能够精准对接算力资源。

- 网络感知算力服务，应用通过云原生算力服务接口向网络传递精细化算力关键特征，网络不仅能够感知算力服务的资源需求，还能感知算力服务的网络需求，为应用提供统一的一体化算网服务。

- 网络感知算力供给，网络向下承接泛在异构的一体化算力供给体系，网络不仅能够感知泛在的算力供给，也能够感知算力的异构分布，为应用选择最佳算力路径和最优的资源利用。
- 精准承载使能，网络能够对每种算力服务都能够提供精准匹配的网络承载策略，支持容器和微服务级别的算力路由，并以确定性、SRv6 可编程、切片等支撑算力服务精准送达。
- 算网深度融合，网络协议融合控制面的算网协同和数据面的算力感知。

2.3 开放服务互联网络

网络整合算力供给和算力服务架构，以提供公共、泛在、高质量的算力服务为目标，以算力与网络深度融合作为顶层设计逻辑。基于上节提出的一体化算网服务的技术需求，本节提出未来网络将构建以网为中心，以云原生为基础的一体化算网服务实现架构，称之为“开放服务互联网络”。如图 1 所示。

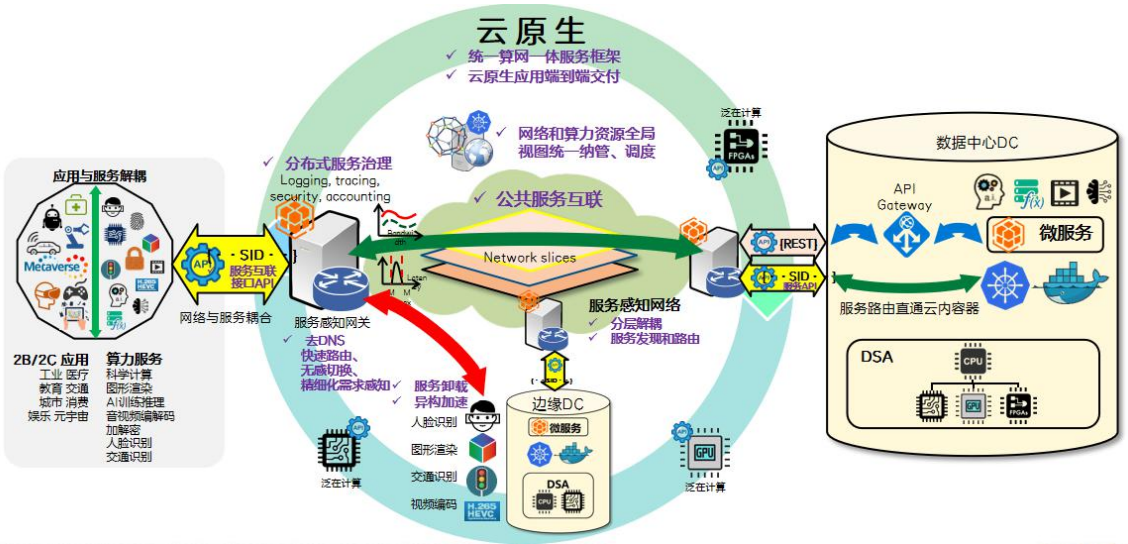

图 1 开放服务互联网络示意图

开放服务互联定义为一种新型云端网一体互联网顶层设计架构和互联范式。随着算力、算法、微服务、比特币、区块链、数字人等整个信息社会脱“实”向“虚”的大背景趋势下，互联网由原来的主机(物理)互联向服务(虚拟)互联演进。信息的传递趋向行为化，不是单纯的获取信息和信息传递，而是包含了对信息和数据的任务处理，比如计算、图像渲染等。IP 协议能够统一主机位置的互联，但很难统一面向千行百业的复杂行为的互联，而且行为的泛在性与基于位置的 IP 地址并不相容。但如果把复杂行为分解为单一行为的服务，而且这些服务可以与特定应用完全解耦作为公共的服务，网络只需要提供服务之间的互联就可以构建一种新的统一开放的行为互联范式。

开放服务互联网络作为一体化算网服务的公共平台，向上向各类应用提供公共服务接口，包括算力、内容、元数据、微服务、网络连接等各类共性服务；向下支持以容器为基础的各类泛在、异构算力和网络资源。开放服务互联将各类云计算、边缘计算以及各种泛在的算力和网络资源深度融合在一起，并根据应用需求进行动态调度和编排，将原本分散的云内、云间资源合理优化、深度整合，以开放公共服务的形式开放给各种应用，使得云网边界模糊，算网架构进一步扁平化。利用容器、微服务等技术快速部署和升级应用，实现云边协同支持多种服务发布、更新、推进，形成完整的模型闭环。应用无需考虑算资源以及网资源，开放服务互联架构根据用户需求自动匹配和调度一体化算和网资源。

在开放服务互联网络中，服务是一个具体的可视、可寻址、可调度、可路由、可衡量、可管理的基本单元。服务跨端、网、云全局定义和管控。从端侧，服务指应用需要调用的非本地的资源和数据处理任务，应用以服务为单位请求远程任务调用和网络连接；从网络侧，服务作为网络可视、可路由、可管理的数据连接；从云侧，服务即所有可以虚拟化、可寻址、可编排、可调度、可视的资源和处理对象。同一个服务标识 ID 可用于应用的任务调用、端到端连接、网

络服务路由、算力的寻址和调度，形成闭环端网云数据流。

2.3.1 开放服务互联网络的两大设计要素

从整体架构设计层面，开放服务互联网络必须具备两个核心要素：

1．以云原生架构为基础，构建全网的一体化算网服务

未来需要以云原生为基础，从全网的视角构建一体化的算力供给和算力服务，提供面向共性算力服务的全局资源编排和优化方案。围绕这个目标，未来云原生架构需要在要素、空间、异构、治理和拓扑等方面进一步扩展以建立以开放服务为核心的新模式。

- 要素扩展，从以计算为主到算、转、存并重。当前在云计算的发展过程中，网络常常成为被忽略或者次要的要素。但当算力调度从云内向全网扩展的时候，网络资源将成为整体服务质量的关键。比如在分布式云中，云间协同计算需要网络提供精准连接和快速响应；在异构算力模式下，需要解决异构计算任务的广域快速疏导和精准引流。因此未来需要从计算为主向算、转、存并重的思路演进，通过对包括网络、存储在内的多要素融合适应算力的发展需求，最大程度的发挥算网一体在共性服务、泛在和高质量方面的核心优势。
- 空间扩展，从云边有限域扩展到云网边端全域。一体化算网服务提供的是面向未来整个社会数字基础设施的公共算力服务，因此将包含各行各业使用者、广泛的算力提供者、多家网络运营商在内的庞大的生态体系。对于共性算网服务的使用方广泛存在于云、网、边、端的各类应用，同时共性算网服务的提供方也包含云、网、边、端全域的计算资源。
- 异构扩展，云原生架构需要从服务端异构扩展到泛在、社会化的异构资源统一纳管，

支持多云、混合云、边缘云等多种形式泛在的算力部署，统一纳管屏蔽底层异构网络和算力，构建弹性、高可用、自动化、自愈性和可观测的统一算力供给体系。全面支持云原生应用接口，应用可一键接入 CPU/GPU/NPU 等多种计算资源，以及 IaaS/PaaS/FaaS 等多种资源颗粒度。

- 治理扩展，从云边统一到云网边端统一，开放服务互联支持基于全局服务标识的分布式服务治理，实现对云网边端资源的一体化管理、度量、命名/标识、发现、注册、解析、可信等，从而强化端到端一体化服务与资源（算力、存储、算法、存算一体等）的提供和部署的能力，拉通跨云、跨网络、跨应用场景的一体化服务链条，满足云原生应用交付。
- 拓扑扩展，从云内 Service Mesh 到全域泛在 Service mesh。当前应用基于云原生的端到端交付的链条中，现有 K8S 的容器网络通过 overlay 网络的方式把底层网络完全屏蔽掉，这使得 service mesh 主要受限在云内互访方式，无法适应未来泛在、动态高频的算力灵活部署，因此未来云原生的拓扑将从云内扩展到全域范围内的服务间互联。

2．以网络作为一体化算网服务的入口和底座，实现服务的泛在感知和调度。

在开放服务互联架构中，网络作为一体化算网服务的入口和底座扮演了很重要的角色，需要承担全网泛在、异构的算力服务的感知和调度，需要为应用提供高 SLA 的网络连接级服务，还要在泛在的、无边界的服务环境中保障各类应用的安全。

- 服务标识深度融合算力服务和计算标准，开放服务互联网络引入服务标识作为跨应用、跨网络和跨算力的唯一服务命名，实现服务的统一抽象、度量和感知，统一网和算的资源分配和调度。作为应用、云、网络之间的轻量级服务传递接口，通过开放标准化服务标识以及接口函数，可以纳管多方、异构资源和服务，实现跨应用、跨网、跨云

一体化资源调度和编排。这样网络不仅能够感知算的需求而且能够感知算的能力，提供最优算力路径和服务策略。应用可以基于服务标识快速发起业务连接，而无需事先通过 DNS 获取目的 IP，从而可以极大消除 DNS 延迟。服务标识基于全局注册和认证机制保证安全可控。服务标识可以进行单独封装，也可以基于 IPv6 进行扩展，在网络层内嵌服务标识及必要的参数，使得网络设备具备服务感知能力。

- 在网络边缘引入“服务感知网关”，作为服务资源感知、服务需求感知和调度的枢纽节点。服务感知网关既部署在服务提供方（中心云、边缘云等），也就近部署在服务使用方（终端、园区等）。对于服务提供方，服务感知网关搜集算力服务部署信息、状态信息，建立服务路由表；对于服务使用方，服务感知网关提供服务调用接口，应用通过服务标识可以直接发起服务调用。
- 算力服务通告和路由。算力服务在经过注册和鉴权之后，可以基于控制面的 ETCD 数据库或者数据面的 BGP 协议扩展等方式，将分布式的异构算力通过统一的服务标识进行全局通告。服务感知网关基于算力通告信息建立算力服务路由表和转发表。
- 算力服务的快速响应、无缝切换以及精细化需求传递，对于连接类服务，网络可以快速感知识别服务类型建立业务所需要的连接，并精准提供确定性保障。对于算力服务，网络可以为应用快速准确的选择最佳服务端和路由，降低呼叫建立时间，同时满足算力和网络的双重服务需求。由于服务与位置的解耦，因此可支持终端或者服务发生迁移情况下的无缝切换。
- 全网异构算力的泛在调度。服务感知网关通过 3 层算力路由直达的方式将应用需求引流到分布式的算力服务实例(如边缘容器或虚机)，避免在 API 网关进行 7 层连接的终结，提高响应速度和卸载性能。这样整个网络可以看作是一个超级 I/O，通过高效的 3

层路由机制支持广域异构算力的卸载和加速，实现全网位置无关的算力调度。

2.3.2 现有 IP 技术支持开放服务互联网络的不足

正如白皮书 2021 所述，目前 IP 互联网的两大设计原则是端到端原则和分层解耦原则。尽管互联网取得了很大的成功，但这两个设计原则把业务和网络过分去耦合，使得两者处于互相割裂的状态。随着互联网业务的纵深演进，尤其是产业互联网的发展，业务和网络的这种完全解耦模式越来越不能满足业务需求。

上一节提出的开放服务互联网络，目标是实现算力和网络的协同调度，同时满足应用对于算力资源和网络资源的需求。开放服务互联网络的一个关键要素是引入了服务感知网关，作为一体化算网服务的入口和底座，实现服务的泛在感知和调度。服务感知网关主要工作于 IP 层，既需要感知应用的算力需求、网络 SLA 需求，也需要感知算力服务的分布情况，算力资源的使用状况等信息。现有 IP 技术如要满足开放服务互联的需求，主要存在以下困难：

1．无法基于服务感知网关实现泛在调度的需求

在算力资源全网泛在分布的情况下，高效的调度方式是在靠近服务发起方的网络边缘位置进行服务感知和分发调度，这也是开放服务互联架构引入服务感知网关的主要原因。但是服务感知网关工作在网络层，无法感知应用层的服务标识，也无法感知算力资源的分布情况。

现行互联网采用了固定的主机互联模式，IP 地址与主机、位置以及业务端口强绑定。应用调用某个服务时，需要首先通过 DNS 查询服务名对应的 IP 地址，在泛在分布式的算力服务模式下，DNS 往往需要重定向很多次才能找到最合适的服务节点，无法满足业务对快速响应的需求。而且，基于 DNS 和 GSLB 的集中调度模式，数据中心按照全算力/全服务模式部署，通过 DNS 进行 DC 入口调度，DC 内再进行服务调度，资源利用率不优，算力布局粗放低效。

2．网业分离机制难以满足未来流级服务 SLA 保障要求

在网业分离机制下，网络协议不感知应用信息。IP 地址与端口号不携带足够的业务特征信息，因此网络只能被动的根据地址+端口号的组合(5-tuple)推断业务类型，但基于 5 元组的应用级别的颗粒度也难以区分不同微服务算力的精细化 SLA 要求，无法为算力服务选择最佳的服务资源和网络路径。现网采用 DPI 技术，甚至采用 AI 大数据来推断业务类型，但面对应用层加密识别效果有限。

因为网络感知服务接口存在不足，现有 QOS 运营机制采用手工模式、运维复杂，响应慢。网络自智想解决这个问题，但是依然需要有一个前置的意图学习阶段，无法实时区分新应用。

3．IP 网络资源管理和路由技术无法满足算网深度融合资源调度的需求

当前网络架构云、边、端、网生态体系相对独立，计算、存储等云资源的调度和网络资源的调度处于互相隔离的状态。云资源调度侧重算力最优体验，网络资源调度侧重网络 SLA 的保障，无法形成用户体验的整体最优。

目前的 IP 路由协议只感知网络拓扑、链路状态、带宽资源等网络层的信息，无法感知和路由应用层面的算力服务、算力资源信息。无法实现算力、网络资源的协同调度。

4．IP 技术体系的基础能力不足

IP 技术按照“尽力而为”的原则设计，存在 QoS 保障、安全性、移动性等基础性能力的不足，无法满足全网泛在服务的要求。比如无法提供工业级应用所需要的时延、抖动有界，零丢包的确定性连接。无法在无边界的环境中提供安全有保障的服务能力。

因此，现网 IP 技术的业务和网络完全去耦合模式不能满足未来的一体化算网服务需求。在未来网络的架构中，业务和网络必须以某种方式“再耦合”。这既能保持业务的独立性，又可使得网络能够感知业务的关键需求，实现服务的感知和调度。如何建立业务和网络之间的桥梁，

是算网融合给未来 IP 网络架构和协议提出的重要需求。本白皮书提出的开放服务互联网络是面向算网融合需求的 IP 网络演进方案。

2.3.3 面向 IP 平滑演进的开放服务互联网络架构设计

开放服务互联网络的核心思想是基于泛在的算网共性服务构建一个新的能力平台，在这个平台上，终端、网络和云端实现以服务为中心的请求、路由、调度和交付。其中增强的 L3 层网络是关键的使能模块。在 L3 层增强基于服务的算网感知和路由能力，在实现算网深度融合并精细化使能算网资源调度的同时，在基础网络转发流程中高效实现了面向服务的算网路由，使能算网资源敏感、高效交互类新型业务场景和应用。

开放服务互联网络架构如图 2 所示，该架构的核心功能是以网络为中心实现广域服务互联，无缝拉通服务使用方和服务提供方之间的高效连接，从而形成面向服务互联的端到端方案。

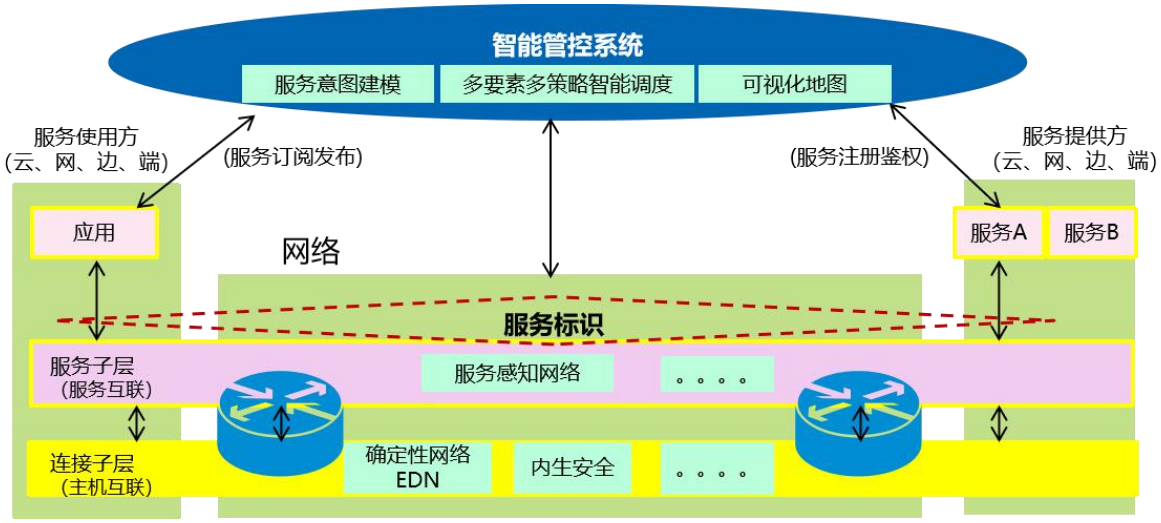

图 2 开放服务互联网络的整体设计

该架构有三个核心设计要素：横向贯穿端网云、纵向打通应用和网络的服务标识；实现服务互联的服务子层；能力增强的连接子层。

1）横向贯穿端网云、纵向打通应用和网络的服务标识

- 引入全新的基础共性服务及其标识，作为端、网、云统一的服务治理对象。包括连接类服务和算力服务。对于连接类服务，网络快速建立业务级颗粒度的连接，提供确定性保障。对于算力服务，网络为应用快速选择最优服务节点，同时满足算力和网络双重需求；
- 基于服务标识实现标准化的共性服务，结合订阅发布模式实现服务的安全可控。
- 应用层直接用服务标识发起位置无关传输层连接，无需 DNS 过程，大大缩短了服务响应时间，并且内含了对移动性的支持。
- 服务标识关联了网络连接的 QoS 需求，使能网络对服务需求的感知和提供有质量保障的连接。

2）在 IP 网络层引入服务子层（3.5 层），以网络为中心实现服务互联

- 在保留传统 IP 主机路由的基础上，开放服务互联网络引入了以服务标识为中心的服务子层，使能网络对服务使用方的算力需求的感知，和服务提供方的算力资源状态的感知，从而通过服务路由，实现服务需求到服务资源的高效连接。从连接和路由模式视角看，开放服务互联网络实现了网络从传统模式下的主机互联到服务互联的演进。
- 开放服务互联网络是基于 IPv6 的兼容性增量创新方案，保留并继承 IPv6 的全部架构和优势特征，基于 IPv6 扩展架构实现服务子层。位于服务子层的基于 IPv6 扩展的服务感知和路由技术，称之为服务感知网络（Service Aware Network，SAN）

3）能力增强的连接子层

- 确定性连接能力增强。服务的网络连接级 SLA 需求由连接子层满足。连接子层保留传统的 IP 主机路由，重用 IPv6 基础路由协议以及基于 IPv6 的各种 UnderLay 技术，比如切片、组播、OAM、SRv6 等。开放服务互联网络面向下一代业务场景，如 AR/VR、

元宇宙等，这些新兴业务场景往往对网络有不同类型的确定性网络需求，比如有界时延、抖动。因此，连接子层使能增强的确定性网络技术，称之为增强确定性网络（Enhanced Deterministic Network，EDN）。

- 内生安全能力增强。开放服务互联网络在传统分组数据网络的基础上同时引入了网路域之外的节点和资源，如算力及加载其上的服务。同时，终端用户跟网络的接入控制接口除了传统的身份鉴权之外，新增了服务请求的接入控制。因此，连接子层同步引入对应的安全机制，在 L3 层基础接入和路由流程中提供基于可信标识的内生安全解决方案。

综上所述，位于服务子层的服务感知网络（SAN），和位于连接子层的增强确定性网络（EDN）、内生安全是开放服务互联网络的三大关键使能技术，在统一架构下满足业务的端到端多样性需求。

3 开放服务互联网络关键使能技术

3.1 SAN、EDN、内生安全等与 IPv6+

如第 2 章所述，开放服务互联网络全新引入了服务功能子层，增强和扩展了连接子层的确定性网络能力和内生安全能力。同时，开放服务互联网络是一种基于 IPv6 的功能增强和扩展架构。近年来，IPv6+创新技术体系和演进路线已经成为行业共识，在既有 IPv6 功能基础上，从网络可编程、SLA 灵活保障、服务驱动等能力上分阶段扩展和演进。如下图所示，在 IPv6+阶段 2，IPv6 分别扩展和增强在大规模网络的确定性路由转发能力（EDN）和基于身份可信的内生安全能力，在 IPv6+阶段 3，SAN 服务路由和服务感知增强和扩展应用感知能力，内生安全

则向服务可控进一步扩展。

图 3 开放服务互联网络关键技术与 IPv6+

3.2 服务感知网络（SAN）

服务感知网络（SAN）架构如下图所示。SAN 的关键设计要素包括如下几个方面：协议功能设计（在 L3 网络层之上引入 SAN 服务子层并向端、云两端延伸）、服务标识的设计、网元功能增强设计（SAN 网关和 SAN 中继）。

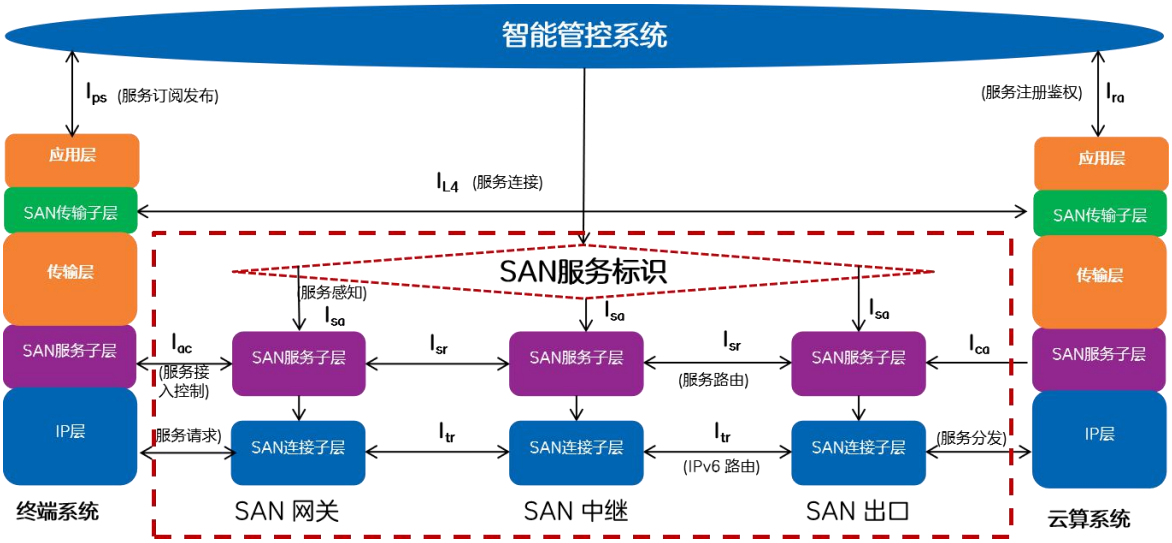

图 4 服务感知网络（SAN）架构示意图

3.2.1 SAN 架构

3.2.1.1 SAN 架构概述

SAN 服务功能子层使能基于 L3 层网络的服务互联，终端得以通过服务标识发起位置和归属无关的业务连接，保障移动场景下的业务连续性。以服务标识为接口，网络得以高效感知服务算网 SLA 需求。同样，以服务标识为索引，网络得以动态感知算力资源状态。因此，SAN 内生支持泛在算力场景下的广域服务互联。

(1) 业务功能高效调度和路由机制设计

SAN 服务功能子层在逻辑上构成了一个完整的路由协议体系，所不同的是这个子层只执行服务“选点”功能，并不直接转发和路由业务报文，跟 IP 路由协议层是解耦和叠加的关系。由于 SAN 服务功能子层拥有以服务标识为寻址对象的控制面和转发面，因此“端”对服务的请求和“云”对服务的提供得以通过服务功能子层无缝高效的连接协同起来，SAN 本质上成为一个 L3 层网络维度的服务路由架构，即异构算力、多方算力、泛在算力之上的基础共性服务，通过 SAN 进行高效、动态、智能的调度和路由。除此之外，为了确保与现网 IP 架构的平滑兼容，避免算侧资源导致的路由表项膨胀和路由震荡，SAN 引入了层次化算力路由的聚合机制，将慢变算力状态通告维护在 SAN 远端节点，将快变算力状态通告维护在 SAN 本地节点，端到端转发和路由流程上实现多级路由缝合机制（服务路由的详细描述参见 3.2.3 节）。

(2) L3 层路由延伸至集群内的微服务

泛在算力和泛在服务的全新业务场景和模式之下，实现服务以及支撑服务的算网资源的高效发现、调度和寻址路由，是实现算网深度融合的核心要素。传统微服务架构主要解决云内应

用的服务高效治理和路由调度，对外网路由而言，仅仅应用或应用网关可见，即网络的横向资源和服务调度的颗粒度被限定在应用甚至 DC，这显然无法满足算网融合模式下服务平台化、资源虚拟化的全新范式，即网络不仅需要发现比应用颗粒度更细的微服务，还需要对其执行灵活高效的调度、寻址和路由。

SAN 将微服务模式向云外和云间进行延伸，网络层面以服务标识为索引构建新型协议功能子层，支持以端、网、云全局语义的服务标识为寻址和路由对象的新型路由架构，即 SAN 以服务标识为全局路由对象，实现对服务及支撑服务的算力资源的全局调度。

3.2.1.2 服务标识的设计

SAN 架构下，服务标识由智能管控系统统一纳管（含注册/鉴权/校验/发布/订阅/策略配置），其生命周期涉及注册、发布、策略下发、订阅、服务请求（封装）、服务路由、服务分发、服务更新、服务撤销。服务标识拥有终端、网络、云算全局语义，所有接口均以服务标识为中心索引。

(1) 服务标识的设计原则如下：

- 服务标识在终端、网络、云端具有全局语义，以服务标识为中心无缝拉通业务、网络和云算系统，实现算网深度融合；
- 服务标识仅适用于基础通用的服务类型。服务标识的语义空间不是全覆盖，仅涵盖复用率较高的共性服务类型；
- 服务标识主要适用于对算网资源有高于“尽力而为”和“一般性计算处理”的服务类型，即 SAN 在 L3 层提供一体化算网统一服务；
- 服务标识支持类型聚合。聚合式标识结构有利于索引和查表效率提升；

- 服务标识可选支持同一应用下多种子业务的关联和同步，以及同一子业务会话的上下行数据流；
- 服务标识可选支持业务功能链的语义设计；
- 服务标识可选支持用户语义设计，以实现具体业务流量的精细化识别。

(2) 服务标识的封装：

- 主机侧方案下的服务标识封装。服务标识在 IPv6 扩展头或 SAN 专用转发头中封装，为兼容主流硬件的转发架构，推荐标识长度为 32，64，128 比特；
- 网络侧方案下的服务标识封装。服务标识重用目的地址字段，即使用 IPv6 地址结构标识服务语义，通过特定前缀唯一表征服务标识并区别于普通 IPv6 地址。

(3) 服务标识在端、网、云的语义

- 在终端，服务标识是用户对服务的请求接口。对用户而言，无需关心服务提供方及其位置，甚至无需关心服务参数，订阅服务并获取服务标识，向 SAN 发起业务请求。
- 在 SAN 网络域，服务标识是算网 SLA 策略的唯一索引，也是网络 UnderLay 服务策略的映射标识。
- 在云端，服务标识是云侧服务调度和路由的对象。

3.2.1.3 SAN 网元功能设计（SAN 网关和 SAN 中继）

服务感知网关（SAN 网关）入节点接受智能管控系统下发的服务算网 SLA 参数，并根据控制面的算网资源状态生成以服务标识为索引的服务路由表项。接收并解析用户业务报文的服务标识，并基于服务标识索引控制面的一级路由表项，首包模式下生成流亲和表，完成业务流的路由和转发。SAN 架构采用层次化算力路由表机制，在入节点维护粗颗粒度算力路由表，在

出节点维护细颗粒度算力路由表。因此，SAN 网关出节点解析用户业务报文的服务标识，并基于服务标识索引控制面的二级路由表项，首包模式下生成流亲和表，完成业务流的最后一跳转发。

SAN 中继节点在需要服务标识递归路由或不同域 UnderLay 拼接场景下执行入节点功能，其他场景无需解析和识别服务标识，无需维护服务标识状态表项。

3.2.2 SAN 服务调度与应用层服务调度

(1) 算网资源维度的服务类型：算网敏感型服务和算网非敏感型服务

服务对算网 SLA 需求敏感与否，决定是否需要 SAN 在 L3 层提供路由和服务保障，即从算网资源以及 SLA 的维度，可将服务类型分为算网敏感型和算网非敏感型。

(2) SAN 在 L3 层聚焦算网敏感型服务

SAN 在 L3 层引入服务标识，为业务提供精细化的算网双维服务，因此 SAN 聚焦算网敏感型服务，即服务对选算和选网均有精细化的需求。SAN 通过服务标识精细化感知服务对网络和算力的需求。因此 SAN 为业务提供高效率高质量服务，解决了传统网络与业务分离导致的网络连接策略颗粒度不够精细化的问题，同时 SAN 在 L3 层通过直通路由取代了传统应用层的 DNS 和 GSLB 等绕行流程，提供了可观的时延收益。

(3) 应用层聚焦算网非敏感性服务

如上所述，对算网 SLA 非敏感类业务，则无需经过 SAN 执行算网融合路由，通过灵活且扩展性良好的应用层协议执行主机侧端云交互，IP 网络提供普通的“尽力而为”连接服务。

3.2.3 SAN 层次化服务路由

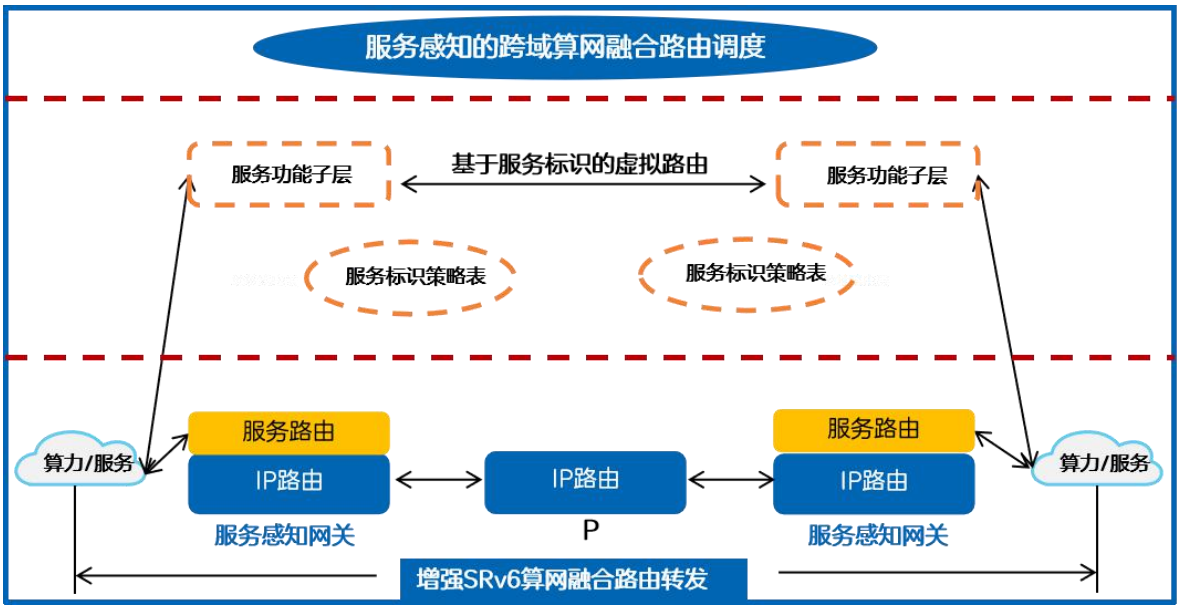

图 5 SAN 层次化服务路由架构示意图

1）以服务标识为索引的服务功能子层

如前所述，SAN 引入了以服务标识为索引的控制面和转发面（轻量级服务标识封装），从而构成了一个虚拟的服务功能子层，在功能上完成以服务标识为寻址对象的服务路由。但是服务路由子层并不实际转发和路由业务报文，而是在传统 IP 拓扑路由层基础上叠加服务节点优选即服务路由功能。

2）服务与 IP 拓扑路由解耦架构

从云算系统的视角看，服务路由的本质是根据算力和网络状态为用户在分布式的服务实例和泛在算力资源中优选服务节点，服务节点优选聚焦的云算侧的资源调度，网络路径的优选聚焦的是 SAN 网络域的资源调度。因此，服务和网络拓扑路由一定程度上是一种解耦关系，SAN 在一体化算网服务路由机制上执行算网解耦方案。

3）层次化分段服务路由架构

SAN 感知并维护算网双维资源和状态，然而算力和网络资源状态在类别、稳定性等方面存在极大的差异，比如 IP 网络拓扑相对比较稳定，因此其对应的路由表也比较稳定，但是算力资源尤其是服务实例资源状态却高度动态，极端情况下可能出现毫秒级的状态变更频率。这种算网资源特性的不一致，势必会给网络带来路由表项膨胀以及路由震荡等严重的衍生问题。SAN 采用层次化路由架构，仅将慢变的服务资源状态通告和同步到 SAN 远端节点，实现一定的资源聚合，将快变的服务资源状态则维护在 SAN 本地节点，从而保护了传统 IP 网络的路由稳定性，并限制了 SAN 节点表项的尺寸。

4）可编程算网融合路由转发方案

服务标识作为业务的算网 SLA 需求感知接口，在 SAN 域内快捷高效的映射到服务的算网策略，SAN 网关节点基于算网路由表执行算网双 SLA 路由。基于源路由技术的算网功能可编程特征，算网端到端路由策略可在一个转发平面统一完成。

SAN 架构下的服务标识是网络和业务之间的高效交互接口，服务标识以及服务标识所在的服务功能子层无缝拉通了业务和网络，算网系统发布服务标识，同时提供对应服务所需的算网资源，以满足这种服务对应的精细化算网需求。如下图所示，算网系统以服务标识为索引，向 SAN 网络下发以服务标识为索引的算网 SLA 需求参数或端到端算网路由策略，后者结合算网状态生成算网路由表。在用户业务流量到达 SAN 入口节点时，该节点解析报文并根据报文中携带的服务标识查上述算网路由表，并对业务流量执行对应的转发和路由策略。

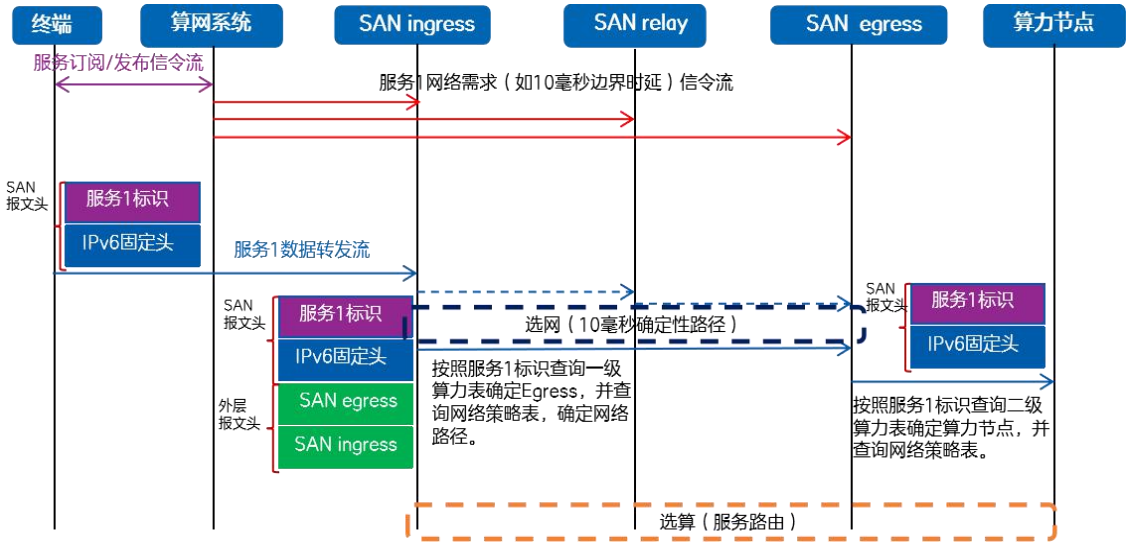

图 6 SAN 算网协同路由流程示意图

3.2.4 SAN 基本业务流程

SAN 架构融合了终端、网络和云三方业务和资源系统，网络域聚焦以服务标识为新型寻址对象的路由协议，终端和云侧除了服务标识的交互和处理机制之外，还涉及到客户端和服务端协议栈。为此，SAN 将业务流程和方案分为主机侧方案和网络侧方案，分别对应主机侧协议栈需要更新扩展和仅需要 SAN 网络节点扩展增强的部署场景。

- **SAN 主机侧方案业务流程**

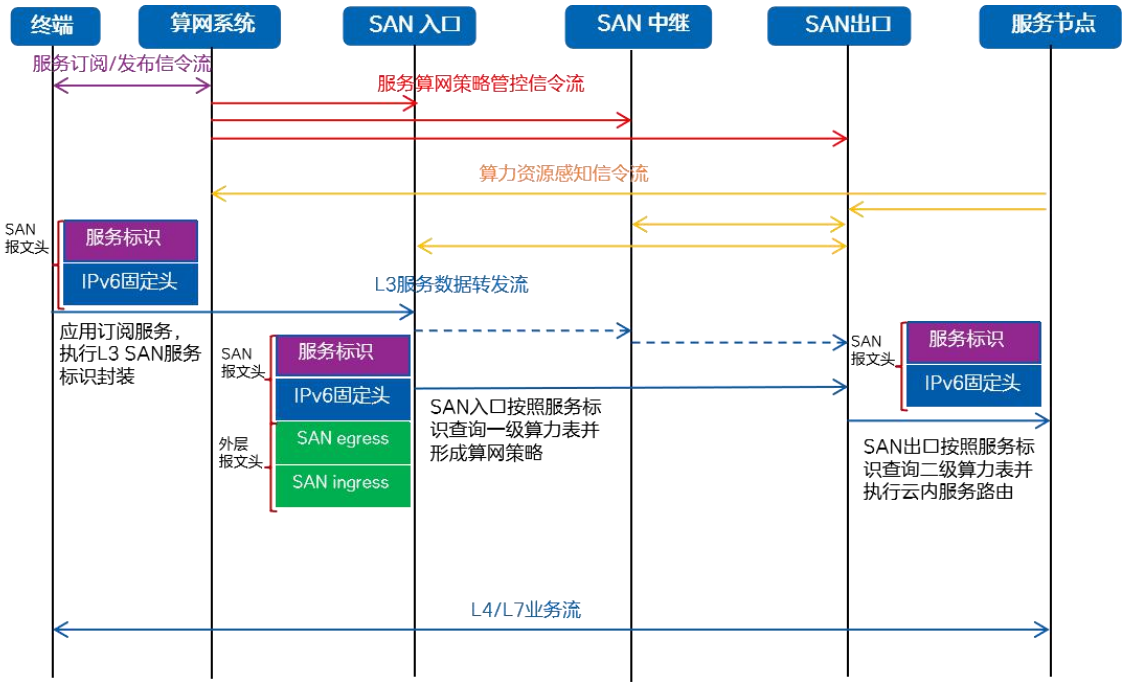

图 7 SAN 主机侧方案业务流程

1）服务标识的发布和订阅。终端（应用客户端）向算网运营系统订阅所需的服务标识，完成身份认证和鉴权。客户端将服务标识封装在 SAN L3 报文头中。

2）服务标识对应的算网需求及策略管控。算网运营系统向 SAN 网络节点下发服务标识的需求及服务策略；

3）服务关联的算力资源感知。云算系统以服务标识为索引，向 SAN 网络域通告服务以及支撑服务的算力状态。

4）服务数据转发和路由。终端向 SAN 网络发送携带服务标识的数据报文，SAN 入口节点根据服务标识关联查找一级算力路由表，确定符合服务需求的服务节点（算力优选）以及 SAN 域网络路径，执行报文转发。SAN 出口节点执行 SAN 域隧道协议操作（如剥掉外层报文头），并查找二级算力路由表，执行网云服务路由。

5）L4 层以源地址，服务标识，源端口号，目的端口号维持业务链接。由于客户端和服务

端引入了服务标识以及以服务标识创建的 L4 层业务链接，因此两端主机侧均需要进行协议栈升级。

- **SAN 网络侧方案业务流程**

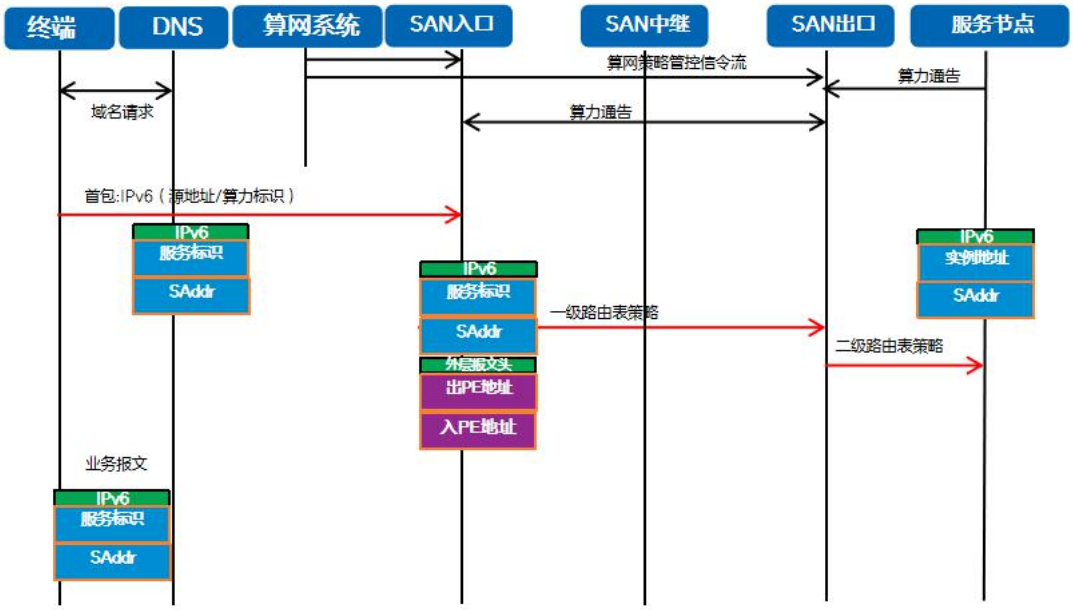

图 8　SAN 网络侧方案业务流程

1）域名解析。终端以服务域名向 DNS 发起解析请求，DNS 向终端返回与 IP 地址同长度的服务标识，终端将服务标识封装在 L3 层目的地址字段，并以服务标识为 L4 链接的目的地址参数。

2）服务标识对应的算网需求及策略管控。算网运营系统向 SAN 网络节点下发服务标识的需求及服务策略；

3）服务关联的算力资源感知。云算系统以服务标识为索引，向 SAN 网络域通告服务以及支撑服务的算力状态。

4）服务数据转发和路由。终端向 SAN 网络发送携带服务标识的数据报文，SAN 入口节点识别目的地址字段的服务标识，并根据服务标识查找以及算力路由表，确定符合服务需求的服

务节点（算力优选）以及 SAN 域网络路径，执行报文转发。SAN 出口节点执行 SAN 域隧道协议操作（如剥掉外层报文头），并查找二级算力路由表，执行网云服务路由。

5）L4 层以源地址，服务标识，源端口号，目的端口号维持业务链接。为了让终端 L4 层协议状态机始终以服务标识维护业务链路，SAN 网络节点需要针对 L4 层首包回复报文（如 TCP ACK）执行实例地址到服务标识的映射。

由于服务标识重用 L3 层目的地址字段，因此主机侧的终端和服务端无需升级协议栈，但是 SAN 网络节点，尤其是入口节点需要识别目的地址字段为服务标识，并执行相应的服务算网路由策略。

3.2.5 SAN 端到端服务互联流程

端到端服务互联流程包括服务治理和服务流量转发两部分。如下图所示，服务标识的注册、发布、订阅等治理流程在服务提供方、服务使用方和服务运营平台三方之间完成，涉及服务编排、以服务为索引的算网资源编排以及服务标识运营。服务流量的转发和路由则由 SAN 完成，其中 SAN 管理面通过服务标识从服务运营系统获取包括服务算网 SLA 在内的特征数据，并下发给 SAN 控制面，后者结合算网资源状态生成以服务标识为索引的服务路由表。算网编排系统向 SAN 控制面下发算网资源调度策略，如弹性均衡策略、就近策略等。SAN 控制面感知云算侧的动态算力资源状态。

服务使用方订阅服务，并在 L3 用户报文中封装对应的服务标识，向 SAN 网络发起服务请求，SAN 转发面和控制面执行基于服务标识的路由和转发，服务提供方根据服务标识完成对应的服务交付。

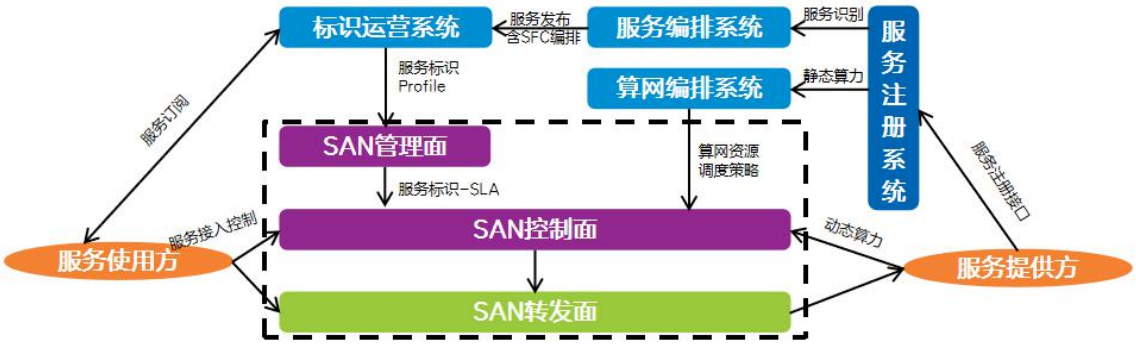

图 9 SAN 端到端服务互联流程

3.3 增强确定性网络（EDN）

开放服务互联需要满足业务多样化的连接需求，包括确定性的连接需求；开放服务互联架构需要支持从局域、城域到广域大规模增强确定性网络（EDN：Enhanced Deterministic Network）技术。

3.3.1 大规模确定性网络

确定性网络技术当前已经在众多的特定网络域中率先发展和应用，不同的标准组织正致力于开发对应的技术标准，比如在工业以太网领域，IEEE 已经完成系列 TSN 标准制定；在工业无线网络领域，IEEE、3GPP 制定了无线 TSN 相关的标准；在 IP/MPLS 分组网络领域，IETF 也发布了 Detnet 的系列 RFC；在数据中心网络中，IEEE 也有无损网络相关的标准制定；这些特定网络的确定性技术正在逐步得到部署应用。

开放服务互联网络提供的服务将会使用不同的异构确定性技术，也会跨不同的特定确定性网络域，需要一种大规模确定性网络技术支持不同的异构确定性技术和不同的确定性网络域。而且，开放服务互联网络提供的确定性服务是日益差异化的，需要多样化的确定性技术相匹配；

而且开放服务互联网络架构中需要支持的确定性服务的规模也是逐步增长的，同样需要大规模确定性网络技术支持扩展性的要求。

3.3.2 面向开放服务互联的确定性网络架构

面向开放服务互联大规模确定性网络面临众多的的挑战：从需求角度看，大规模确定性网络需要满足差异化的确定性 SLA 连接需求，这些差异化的需求需要在相同的基础网络设施中提供服务；从技术角度看，已有不同的异构确定性技术提供不同的确定性连接质量；从网络角度看，大规模网络跳数多，链路时延大，跨越多个不同的网络域；从端到端角度看，端到端连接的路径计算和路由机制需要支持确定性属性和能力。

针对上述问题，面向开放服务互联的大规模确定性网络可以从资源层，路由层和业务层等三个层次分别提供三维确定性，即资源确定性、路由确定性和时间确定性，建立统一的大规模确定性网络架构。资源层维护了全网的确定性资源，对确定性资源进行统一建模形成确定性链路，屏蔽异构资源能力的差异；路由层基于资源层统一建模的确定性链路计算内生确定性路由，提供确定性承载能力；业务层通过规划业务流的流量特性，在入口节点做流量监管，并映射至确定性的路由以满足不同类型不同级别业务的时间要求。

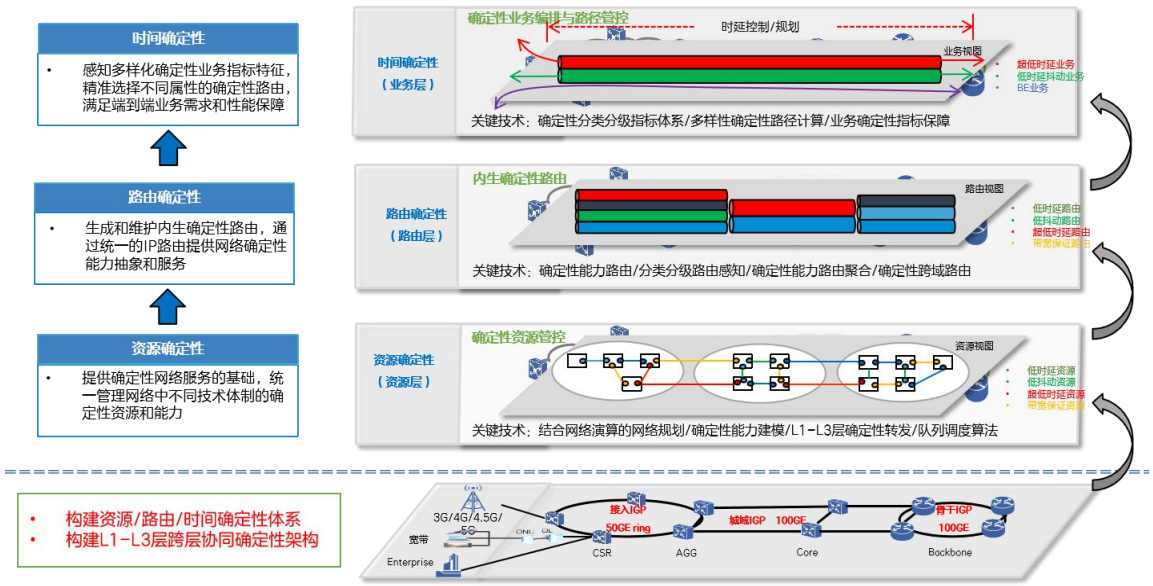

图 10　面向开放服务互联的大规模确定性网络架构

(1) 资源确定性

确定性业务需求的多样化导致网络需要提供的确定性能力不同，与确定性能力相关的资源也是多样化的，网络需要屏蔽网络能力的差异性。资源确定性是提供确定性网络服务的基础，是指满足节点内、链路处理的确定性指标达成的资源以及对应资源的处理机制（比如链路带宽、队列及其调度算法）。需要对网络进行整体资源规划，统一建模异构确定性资源形成统一的确定性链路，为不同级别的确定性转发能力提供保障。确定性链路可以是提供确定性传输的子网（Sub-network），也可以是 P2P（Point-to-point）链路。当网络中的现有资源不足以满足业务的 SLA 需求时，需要重构虚拟网络。

(2) 路由确定性

传统的路由只具备可达性，确定性需求如时延抖动等只作为算路约束条件，路径随网络拓扑的实时变化而发生改变，不具备 SLA（Service Level Agreement）能力，无法满足多种确

定性级别的需求。为了满足分类分级确定性业务的需求，基于统一建模后的确定性链路资源，进一步创建具有不同 SLA 等级的确定性路由。

确定性路由可以基于严格显式路由或松散路由，前者适用于有控制器的集中式场景，后者适用于没有控制器的分布式场景。集中式场景下，确定性业务的源宿 PE 处于有限物理范围的广域网的两端时，由单个控制器（单域时）或多个控制器（跨域时）根据收集的确定性资源，预先按照典型的业务流量特征计算一条或多条具有确定性 SLA 的路径，或者在有业务需求时按需根据业务流量特征动态计算。建议生成两条时延非常接近的不相交路径形成 1+1 保护，双发双收，并在 egress PE 上做复制消除。分布式场景下，在设备侧通过路由协议计算内生确定性松散路由。域内使用 IGP（Interior Gateway Protocol）基于确定性时延度量（Deterministic-delay）计算确定性路由，跨域使用 BGP（Border Gateway Protocol）基于精确的时延/抖动指标计算确定性路由。

(3) 时间确定性

时间确定性包括时延确定性与抖动确定性。时间确定性是在资源确定性与路由确定性的基础上，确定性业务流量在网络中端到端的时延/抖动将被严格限制在有界的区间内。不同的业务等级由于对时延/抖动的界限需求不同，则映射至的确定性路由所使用的资源与路由机制也不同。

3.3.3 增强确定性网络 EDN 方案

基于大规模确定性网络架构，增强确定性网络 EDN 从两个方面支持面向开放服务互联的大规模确定性网络承载方案，一是对确定性业务的差异化特征进行分类分级，二是为分类分级的

差异化业务提供关键技术保障。

3.3.3.1 分类分级差异化业务特征

根据确定性业务分类分级需求，需要定义确定性分级服务模型，根据时延和抖动指标，对网络中确定性业务的服务质量（Quality of Service，QoS）进行分级，明确确定性服务等级中每个等级对应的服务级别协议 SLA 指标。将确定性业务映射到确定性服务等级，在报文头中封装确定性分类分级信息，包括流标识或分类分级标记和 SLA 指标参数等，确定性网络按照业务的确定性服务等级保障不同业务的确定性转发，满足确定性业务多样化需求。从确定性业务需求的角度，可将网络中确定性业务特征分成 5 种类型或级别：

- Level-1：带宽保证类，指标要求包括基本带宽保障及一定的丢包容忍度，时延上限无要求，时延抖动无要求，典型业务包括下载类，如 FTP 等；
- Level-2：抖动保证类，指标要求包括，时延<300ms，抖动<50ms，典型业务包括同步语音类，如语音电话等；
- Level-3：时延保证类，指标要求包括，时延<50ms，抖动<50ms，典型业务包括实时通信类，如视频、生产监控、通信等；
- Level-4：低时延/低抖动保证类，指标要求包括，时延<20ms，抖动<5ms，典型业务包括视频交互类，如 AR/VR，全息通信，云视频，云游戏等；
- Level-5：超低时延/超低抖动保证类，指标要求包括，时延<10ms，抖动<100us，典型业务包括生产控制类，如电力保护、远程控制等。

确定性网络分级 QoS 服务模型可以满足分类分级后的最差情况下的端到端时延，抖动，及

丢包等，可通过支持资源确定性、路由确定性及时间确定性的增强确定性网络架构实现。

3.3.3.2 增强确定性网络 EDN 关键技术

基于分类分级确定性业务的多样化承载需求，采用三个维度的大规模确定性网络架构的增强确定性 EDN 方案对于资源确定性、路由确定性和时间确定性提供下面的关键技术增强。

1）增强确定性“链路”技术

EDN 将确定性资源统一抽象为确定性“链路”，不同的资源类型或资源能力都是为确定性路由层提供不同的“链路”，这些链路具有确定性路由可感知的确定性属性。通过确定性链路的抽象，对路由层屏蔽不同资源的细节，由确定性链路集合构成具有确定性服务能力的逻辑拓扑。

2）内生确定性路由技术

在确定性链路的基础上，生成和发布确定性路由，区别于传统可达路由，内生确定性路由本身都是由确定性链路构成，通过路由选择匹配不同的确定性链路资源，从而提供不同的路由转发服务。

3）基于分类分级的差异化业务特征支持大规模时间确定性

分类分级的差异化业务特征通过 SDN 控制器进行端到端确定性链路资源预留，然后通过 SAN 服务信息选择相匹配的内生确定性路由，从而支持分类分级的差异化时间确定性。

3.4 网络内生安全

开放服务互联网络从原有的集中式数据中心内服务模式转为泛在、开放的分布式服务模式，

在满足多样化业务泛在化连接需求的同时，却使得服务对外的暴露面持续增加，易于引发非法者窃取信息、攻击资源等安全问题。为快速响应用户需求，服务愈发下沉至客户近端，边缘网络的系统防护能力不足或缺失，导致服务安全更加难以得到保障，造成攻击者非法访问服务、以及各类攻击造成业务不可用等安全风险。此外，工业互联网、远程医疗等产业网络未来 IP 化和智能化的趋势，打破了原有的安全防护边界，不仅给产业服务带来了攻击风险，也使得攻击入口急剧增加。再者，开放网络场景中多样的业务、海量的终端和泛在的连接均为非法入侵者提供了更多的攻击条件，严重加剧攻击程度。而现有的攻击防御方案通常采用额外部署多类硬件设备实现联动清洗，先检测流量再过滤的被动防护机制，效果滞后且占用大量网络资源。另外，借用防火墙这类补丁式方案，需要新增设备部署的同时，也在业务面引入了新的攻击点。因此，现有防御技术难以满足网络发展过程中实时、高效、系统的网络可信安全防护需求。

当前网络安全不应仅仅局限于对数据中心、应用边界防护的关注，而需要通过全局视角系统性审视安全防御，从以下多个方面提供全面性、体系化防范能力。

- 实时性：借助网络层为业务面提供接入域、传输域、目的域等多点防范，尽可能达成自动化接入和靠近攻击源的防护。
- 高效性：打造高效防御阻断系统，减少业务冲击，为未来网络赋予高性价比、精准化攻击检测能力。
- 系统性：整体上面向全系统构建攻击防护方案，设计系统性的信任链传递机制，多维度、多层次、多位置提供可信可控防护机制。

3.4.1 网络内生安全可信防御整体架构

基于攻击防护安全需求识别，设计网络内生安全可信防御框架体系。整体架构主要分为身

份可信和服务可控两大类方向，如下图所示。其中身份可信涵盖基于可信平台模块的终端身份可信存储以及基于业务源的 IP 真实性控制关键技术，从发起端设备自身以及网络接入角度实现用户身份的安全可信，保障接入过程的可信认证与转发。服务可控涉及面向业务的网络层防御机制，可针对通信业务数据在网络层面实现近源近目的多点业务防护，提升系统防御全面性和攻击免疫性。

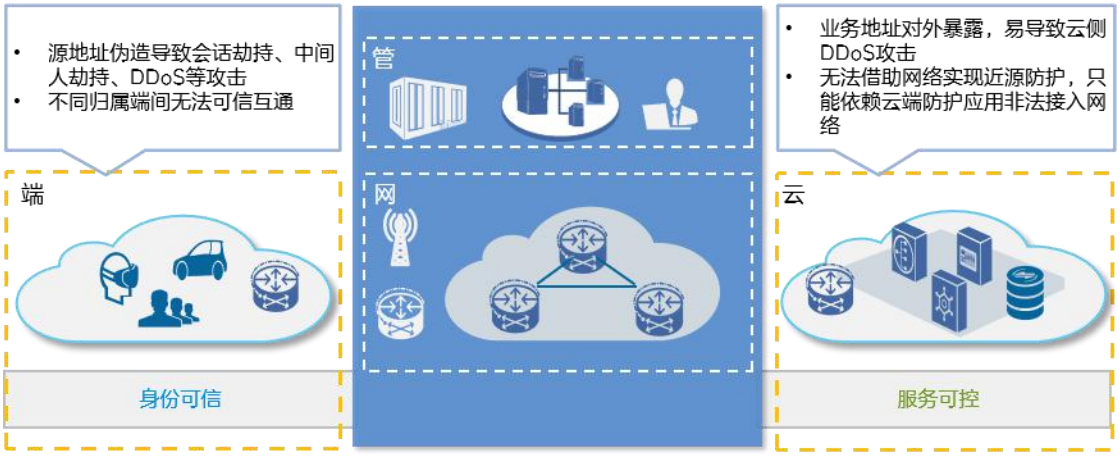

图 11 网络内生安全可信防御框架

基于终端、服务、网元身份和位置的数字化统一表达，充分考虑通信业务特点及其安全防护需求，在设计顶层进行内生安全攻击防御体系化建模。整个内生安全攻击防护体系涵盖多种关键技术，具体涉及基于可信平台模块的终端身份可信存储、基于业务源的 IP 真实性控制、以及面向业务的网络层防御机制，从而构建网络可信安全通行证防控体系。

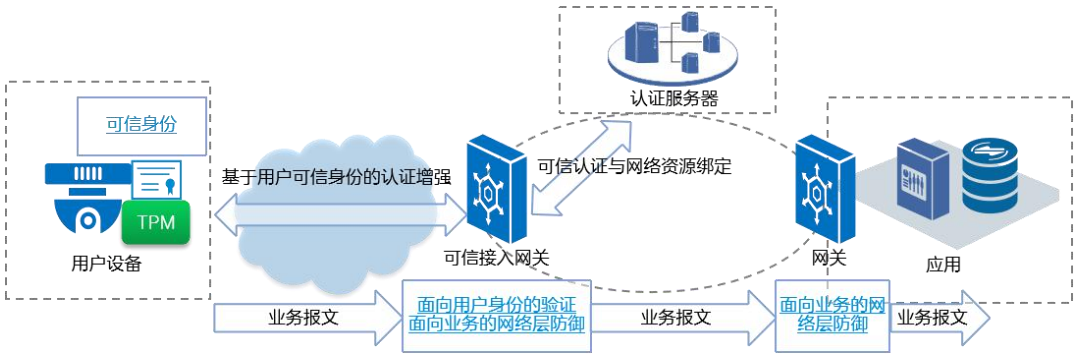

图 12 内生安全攻击防护体系

3.4.2 身份可信

由于目前攻击防御方案主要考虑在目的边界进行防护，无法从业务发起源实现根本性的控制，未能提供真正意义的端到端防护能力。借助可信平台模块，利用可信执行环境对业务发起源端进行可信增强，个人隐私信息的加密保护，提供信息的安全存储和执行环境，从源头上尽可能消除终端信息假冒的风险，系统性实现信任链传递、设备真实身份安全保障。

传统 IP 网络缺乏基本的安全性设计，仿冒源地址引发的攻击层出不穷，而现有的可信通信技术验证开销大、保护机制考虑不够周全，难以满足拥有多样化应用、海量终端的泛在网络安全需求，因而需要考虑如何系统性构建高效的业务真实源验证安全机制。基于业务源的 IP 真实性控制机制提供基于用户可信身份的认证增强、用户资源的可信分配授权、面向用户身份的接入网验证技术。

- 基于用户可信身份的认证增强通过自主式便捷身份可信认证方式，以可信身份为基础增强终端接入认证流程，避免繁杂的人机交互过程，实现终端自动化可信接入认证、提升入网效率，为终端安全接入网络提供保障前提。
- 用户资源的可信分配授权保证了终端资源的可信分配与获取，通过可信认证与网络资源紧密绑定，确保终端成功认证后再授权地址分配，避免未经认证而申请 IP 地址的非法行为。与此同时，通过在 IPv6 后 64 位中放入可信身份以构造完整的终端 IPv6 地址，实现终端位置和身份强绑定，为后续的业务验证提供报文发起端身份和位置的可信性保障。
- 在业务传输中，面向用户身份的接入网验证技术基于访问控制的高效验证机制、基于密码学的验证码快速生成和验证机制，由接入网关验证报文中发起端身份和位置的协同验证与控制，确保发起端的真实性。通过多层级信任传递和多类型验证方式确保业

务源地址转发真实性，避免地址假冒引发的网络攻击、信息非法获取等异常操作，实现高性能轻量化的实时验证、精细化管控和网络安全可信传输。

3.4.3 服务可控

充分保证业务地址源真实的同时，目标可访问性也应得以关注。端到端通信业务中如何确保业务获取目的端访问授权、如何高效识别数据报文的合法与否均值得深入研究。

面向业务的网络层防御通过全方位构建目标域鉴权授权机制，及时阻止无合法访问权限的真实地址用户非法访问业务行为。利用验证信息的统一化生成、一致性表达、动态更新管理以及轻量化抗重放机制，为业务访问提供高效的合法性验证依据，增强网络主动和被动抵御攻击的能力。在网络层面实现近源、中间传输节点以及近目的的多点攻击防范，完善整套面向未来网络的攻击主动防范和溯源灵活阻断技术方案，达成流量实时高效检测控制的安全防护系统。

4 开放服务互联网络的应用示例

4.1 云游戏应用场景下的实例化方案

云游戏由三个主要环节组成，第一是客户端指令的上传，第二是服务端根据指令进行游戏操作，执行抓屏、渲染、编码等计算任务，第三是将处理完成的视频流下发给客户端。如下图所示，云游戏的上下行业务是两个独立的业务链接，但二者之间存在业务上关联。因此，从 SAN 架构的视角看，上下行业务分别对应两个不同的服务标识：SID1，SID2.

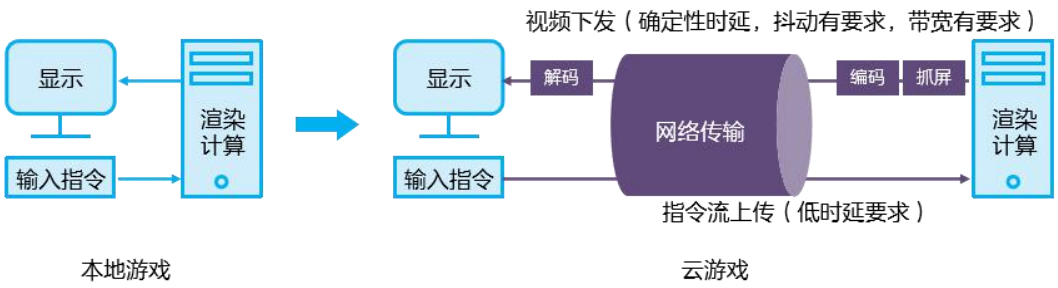

图 13 云游戏服务化示例

1．SAN 方案一（仅提供网络连接类服务）

云游戏上下行都需要网络提供精细化连接级服务。这种模式下对应的 SAN 业务流程如下：

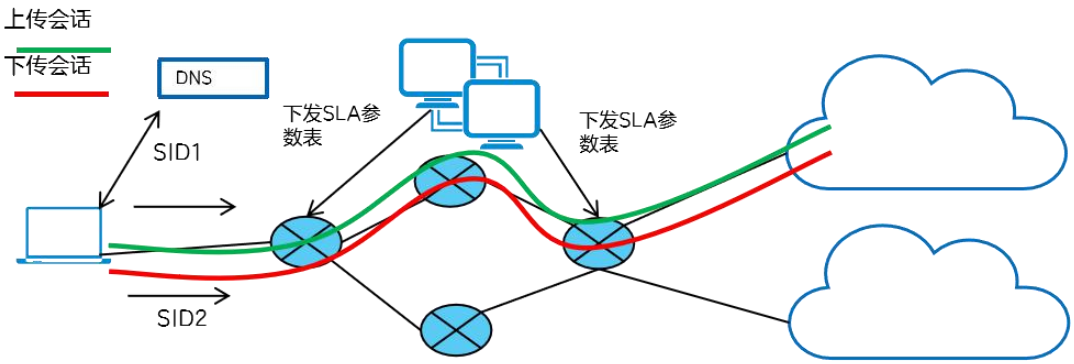

图 14 云游戏 SAN 流程之一

1）客户端从 DNS 获取到服务端的 IP 地址；

2）同时建立两个到服务端的会话，上行会话（指令流上传会话）使用 SID1 封装，下行会话（视频流下行发送）使用 SID2 封装；

3）网络根据 SID1 和 SID2 查找控制面下发的 SLA 参数表，进行路径选择和引流，并按照匹配的网络策略转发业务流量。

2．SAN 方案二（网络提供算网服务）

SAN 为上下行业务提供一体化算网服务，终端通过 SID1，SID2 分别向云游戏服务端发起两个会话连接，SAN 基于这两种服务标识执行选算（服务节点优选）和选网（网络连接策略匹配）。

对应的基本业务流程如下：

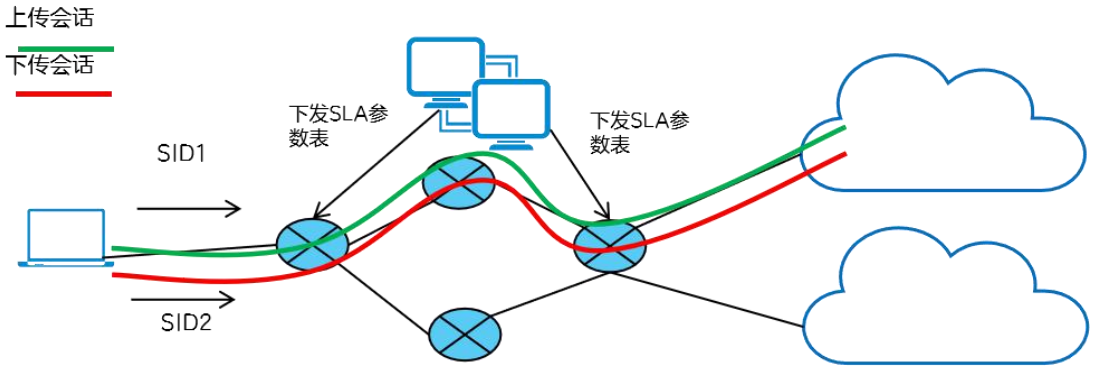

图 15 云游戏 SAN 流程之二

1、客户端无需通过 DNS 获取具体服务节点的 IP 地址；

2、客户端创建上行会话（指令流上传会话）使用 SID1 封装，下行会话（视频流下行发送）使用 SID2 封装。这里 SID1，SID2 同属一个云游戏应用，因此 SAN 需要确保二者“选算”结果为统一服务节点；

3）SAN 根据 SID1 和 SID2 查找控制面下发的 SLA 参数表，进行路径选择和引流。由于同一个业务会话的上下行流量对网络的需求不同，因此 SAN 需要识别并区别对待不同方向的业务流量。

4.2 人脸识别场景下的实例化方案

人脸识别通常包括人脸检测，特征提取，人脸对比三种类型的子业务。SAN 架构下，这三种子业务分别对应的服务标识为 SID1，SID2，SID3. 人脸识别业务流程如下图所示。

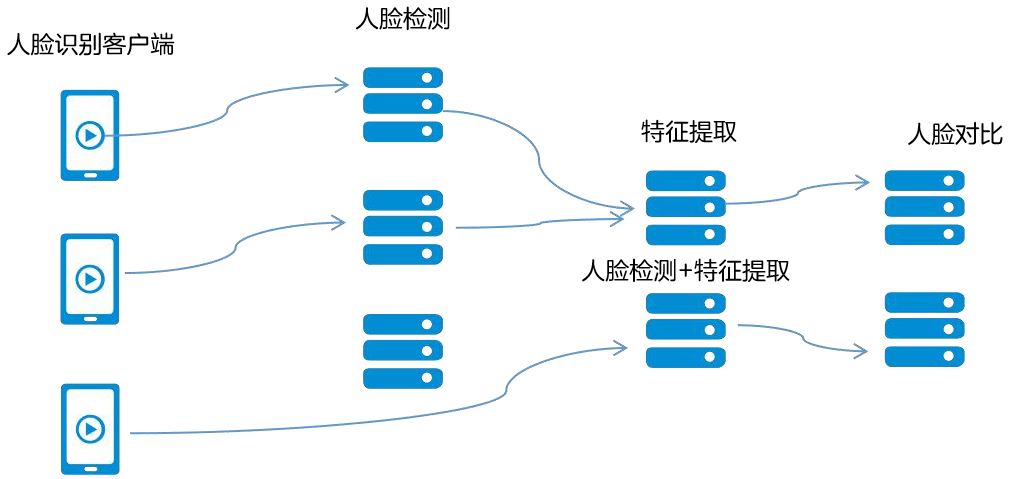

图 16 人脸识别业务示例

实际部署中，三种子业务可能按一定的时序依次执行，或者合并统一执行。

SAN 架构下的人脸识别流程示意图如下图所示。

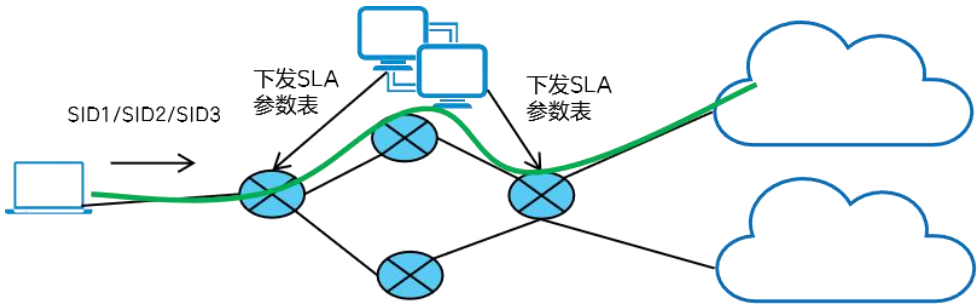

图 17 人脸识别 SAN 业务流程

1、客户端无需通过 DNS 获取 IP 地址

2、客户端创建一个到服务端的会话，使用三个 SID 分别进行服务发起，或者三个 SID 形成业务链进行分发。

3）SAN 根据 SID1/SID2/SID3 查找控制面下发的 SLA 参数表，然后进行路径选择和引流。由于同一个子业务会话的上下行流量对网络的需求不同，因此 SAN 需要识别并区别对待不同方向的业务流量。

5 总结

本白皮书针对 IP 网络未来如何演进，在白皮书 2021 所提出的“横向”、“纵向”两个设计原则的基础上，以算网融合为核心，从主机互联到服务互联，提出未来 IP 演进方案——开放服务互联网络。

开放服务互联网络的目标是提供开放的、泛在的、有质量保障的一体化算网基础设施服务。针对 IP 网络技术存在的四个不足之处，开放服务互联网络在连接子层保留 IP 主机互联并增强连接能力，新增服务子层，引入服务标识和服务互联，实现以网络为中心的服务感知和路由。

开放服务互联网络的三大核心技术是服务感知网络（SAN）、增强确定性网络（EDN）和网络内生安全，未来还将同业内各方共同合作，在服务子层和连接子层研究更多的技术，更好地满足未来业务发展的需求。

6 缩略语

- BGP：Border Gateway Protocol 边界网关协议
- CPU：Central Processing Unit 中央处理器
- Detnet：Deterministic Networking 确定性网络
- DPU：Data Processing Unit 数据处理器
- DNS：Domain Name System 域名系统
- DPI：Deep Packet Inspection 数据包深度检测技术
- DSA：Domain Specific Architecture 特定领域架构
- EDN：Enhanced Deterministic Network 增强确定性网络
- ETCD：etc（存储配置信息的目录）+d（distribution 分布式） 分布式存储数据库
- FaaS：Function-as-a-Service 功能（或函数）即服务
- GPU：Graphics Processing Units 图形处理单元
- GSLB：Global Server Load Balance 全局负载均衡
- IaaS：Infrastructure-as-a-Service 基础设施即服务
- IGP：Interior Gateway Protocol 内部网关协议
- K8S：Kubernetes 开源的容器编排引擎
- NFV：Network Function Virtualization 网络功能虚拟化
- NPU：Neural-networks Process Unit 神经网络处理单元
- PaaS：Platform-as-a-Service 平台即服务
- QoS：Quality of Service 服务质量
- SaaS：Software-as-a-Service 软件即服务

- SAN：Service Aware Network　服务感知网络
- SDN：Software Defined Networking　软件定义网络
- SLA：Service Level Agreement　服务保障协议
- SRv6：Segment Routing IPv6　基于 IPv6 转发平面的段路由
- TSN：Time-Sensitive Network　时间敏感网络

7 参考文献

[01]中兴通讯股份有限公司：IP网络未来演进技术白皮书，2021年6月

[02]ITU-T FG-NET2030：Representative Use Cases and Key Network Requirements for Network 2030，2020年1月

[03]ITU-T FG-NET2030：Additional Representative Use Cases and Key Network Requirements for Network 2030，2020年6月

[04]中国移动：算力网络白皮书，2021年11月

[05]中国电信：云网融合2030技术白皮书，2020年11月

[06]中国联通：中国联通CUBE-Net3.0网络创新体系白皮书，2021年3月

[07]中国信息通信研究院等：IPv6+技术创新愿景与展望白皮书，2021年11月

[08]中国信息通信研究院：云计算发展白皮书(2022)，2022年7月。

[09]云原生产业联盟：云原生发展白皮书(2020),2020年7月。

[10]中国算力大会，中国算力白皮书(2022)，2022年7月

[11]张亮等：未来架构，从服务化到云原生，电子工业出版社，2019

[12]IETF：draft-huang-service-aware-network-framework, Daniel Huang, “Service Aware Network Framework”, May 2022,
<https://datatracker.ietf.org/doc/draft-huang-service-aware-network-framework/>

[13] IETF：draft-huang-two-segment-routing-solution-of-can, Daniel Huang, “Hierarchical segment routing solution of CAN”, July 2022,
<https://datatracker.ietf.org/doc/draft-huang-two-segment-routing-solution-of-can/>

[14] IETF：draft-liu-can-computing-resource-modeling, Peng Liu, “Computing Resource Modeling for CAN”, July 2022,
<https://datatracker.ietf.org/doc/draft-liu-can-computing-resource-modeling/>