



中国移动
China Mobile



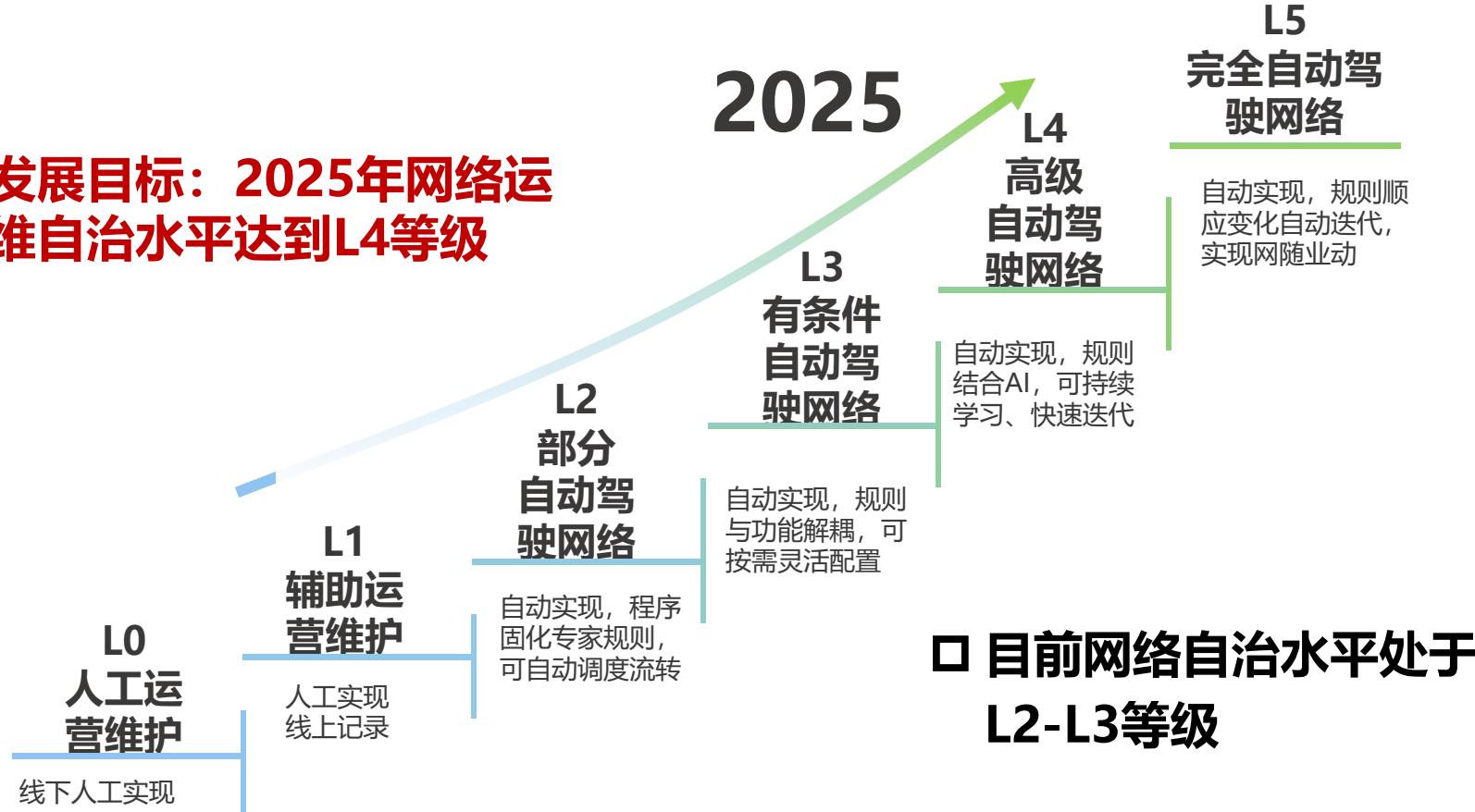
基于数字孪生和内生AI的网络自治

邓娟

中国移动研究院 未来研究院



□ **发展目标：2025年网络运维自治水平达到L4等级**





架构：集中式智能→多形态智能

- 网管集中式数据采集和模型训练→网管/网元设备间数据和模型协同机制
- 重塑网络数据和AI计算资源的分布架构，云边端多形态智能，提高模型训练和响应效率



效果：高度自治网络

- 降本增效的运营智能→端到端高度服务智能
- 响应速度：1小时/15分钟→秒/毫秒/微秒级
- 优化提升维度：网元→用户、信道
- QoS级服务智能保障



场景：多样性发展需求

- 以ToC业务为主的智能化应用相对成熟
- ToB场景业务智能化水平仍需提升，包括场景、数据、算法、效果，进一步增强网络切片差异化服务能力



环境：强大的模拟复现能力

- 网络资源/参数的智能调度决策依赖精准的环境反馈
- 缺乏可供智能体训练的试错环境，可依托海量历史数据建模，还原网络不同性能表现

数字孪生和内生AI实现6G网络高水平自治



6G网络以实现L5等级的自动驾驶网络为目标，面向消费者和垂直行业客户提供全自动、零等待、零接触、零故障的创新网络服务与ICT业务，打造自服务、自治愈、自优化、自演进的通信网络

智能应用

“0”接触运维



自服务



自优化



自治愈



自演进

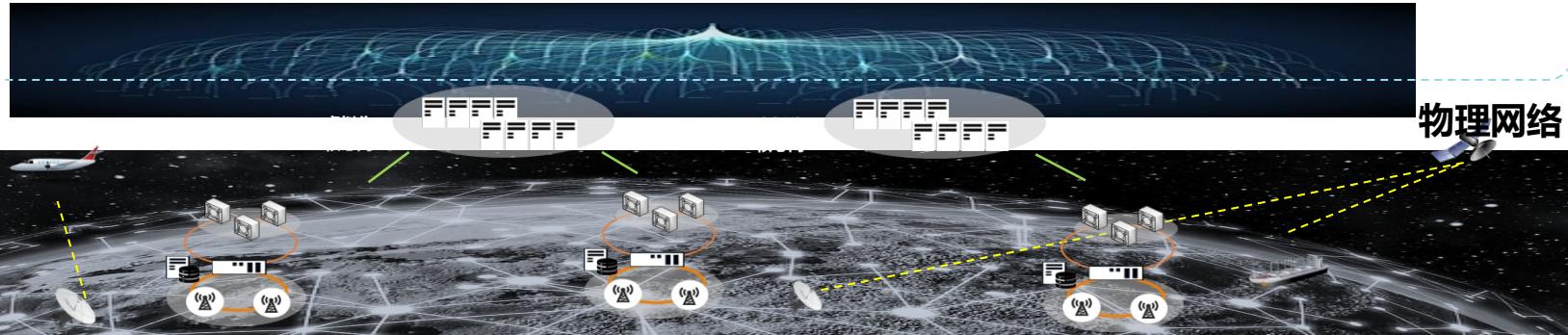


内生AI

网络数字孪生

数据采集

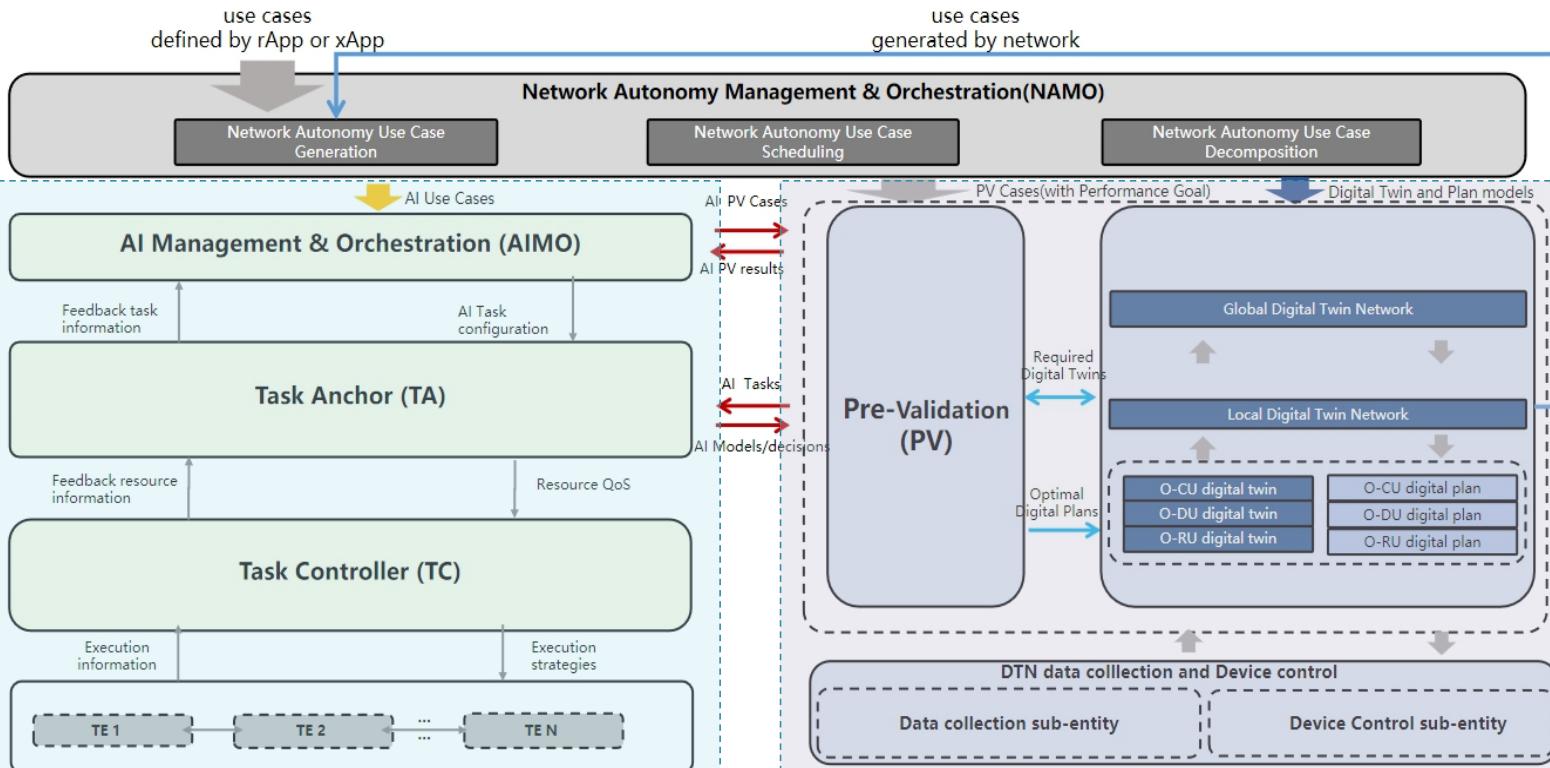
物理网络



数字孪生和内生AI实现6G网络高水平自治



相比5G，6G新增内生的智能面和虚实交互的数字孪生网络，并通过智能面与数字孪生网络的交互与融合，实现6G网络全生命周期（规、建、维、优）的高水平自治

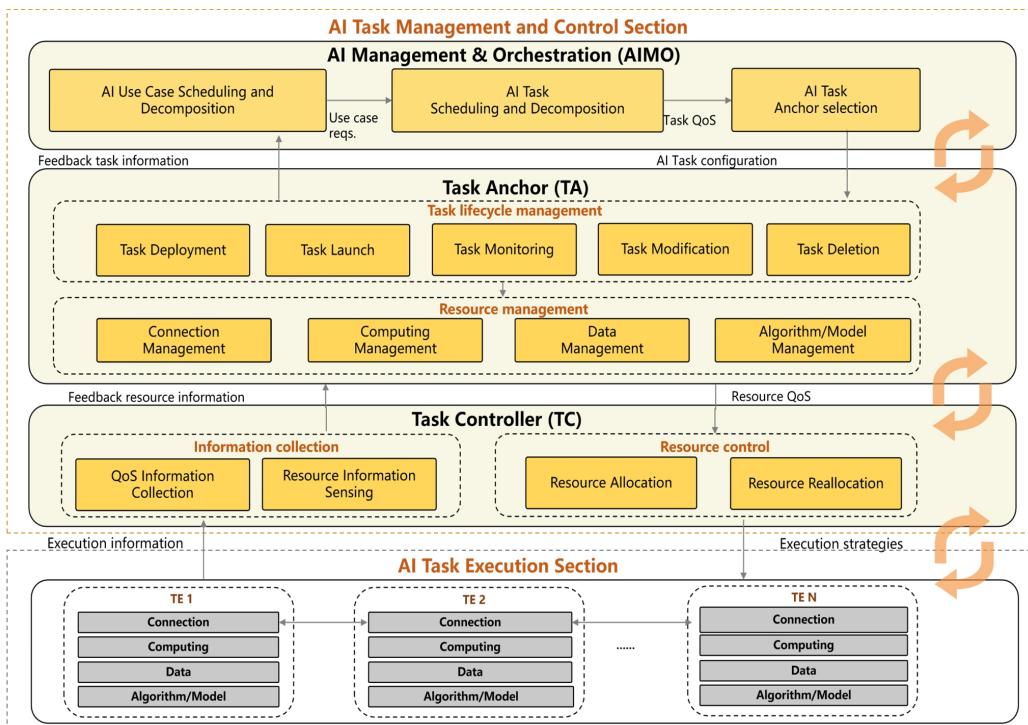


数字孪生网络生成数据样本，为AI工作流和模型提供预验证和优化环境；内生AI为数字孪生网络提供AI能力支撑，优化孪生网络性能

6G内生AI网络逻辑功能架构设计



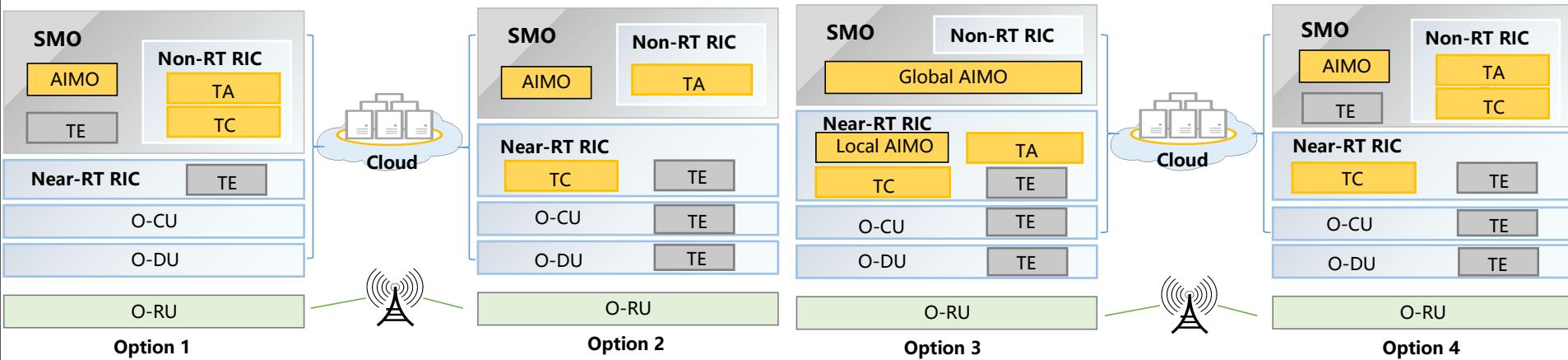
6G内生智能面的逻辑功能主要由AI管理编排（AIMO）、任务锚点（TA）、任务控制（TC）和任务执行（TE）组成



序号	技术特征
1	以任务为中心 <ul style="list-style-type: none">从以会话为中心到以任务为中心包括AI任务的拆解、管理、资源配置、调度
2	三层闭环QoAIS保障 <ul style="list-style-type: none">QoAIS指标体系：突破传统通信QoS指标维度三层模型：服务QoS、任务QoS、资源QoS管控协同的QoAIS保障
3	多维资源调度 <ul style="list-style-type: none">连接、计算、数据和模型连接与计算融合的控制

内生AI架构将AI三要素与网络连接一样下沉为网络内部的基本资源，使网络通过多维资源的协同，直接、便捷地为用户提供有QoS保障的AI服务，完成网络自治场景所需的AI用例

内生AI网络逻辑功能部署选项——以ORAN为例



Option 1

- 管理面部署方案
- 实时性要求较低，资源和精准度要求较高场景
- 任务执行在网管，可能涉及NRT-RIC协作

Option 2

- 管理面部署方案
- 实时性要求较低、所需资源仅存在于网元
- 任务执行在网元

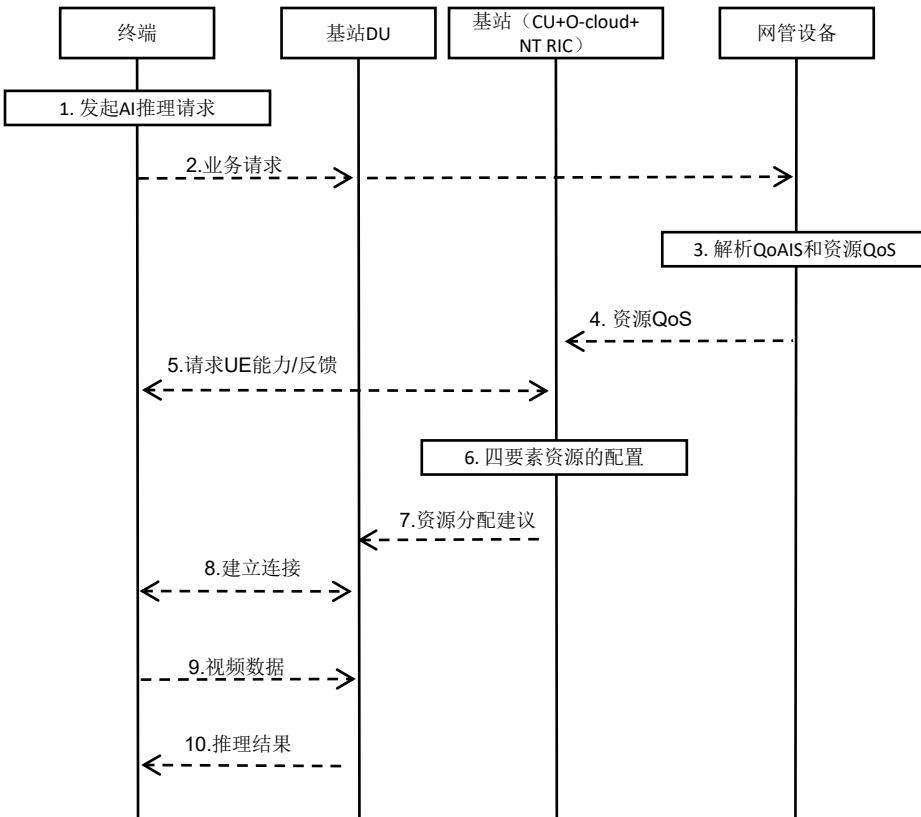
Option 3

- 管控两级用例方案
- 自治区域规模较大，AI用例拆解为本地和全局
- 本地实时性要求较高，任务执行在网元
- 本地执行结果上报给网管

Option 4

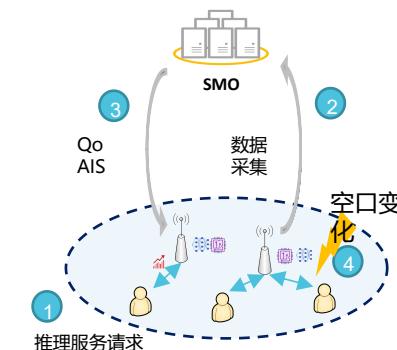
- 管控两级执行方案
- 实时性要求较低，所需部分资源仅存在于网元
- 资源调度上多级协同

AI训练服务示例



■ AI推理服务用例具体流程：

1. 用户发起AI推理请求；
2. 导入AI用例到网管设备；
3. 解析AI服务对应的QoAIS，由网管设备下发至基站；
4. 将QoAIS分解为资源QoS需求，明确所需四要素资源的需求，包括连接、计算、数据和算法/模型；
5. 基站实时决定并调整计算的分配、优化通信连接质量、决定并采集处理所需数据，以及决定并更换或优化算法模型，以保证任务QoS的达成，从而保证QoAIS的达成；
6. 基站将结果反馈给用户。

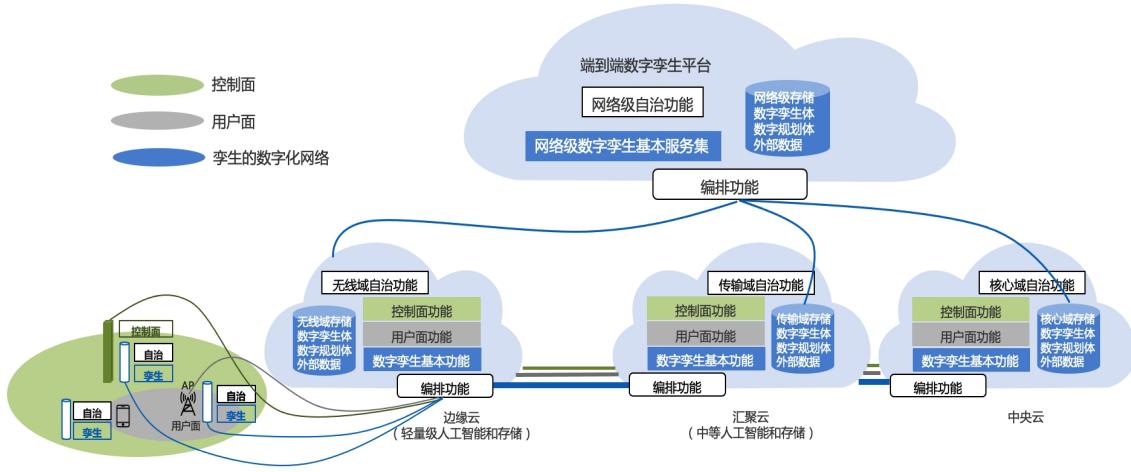


■ 功能模块：

1. 网管：解析QoAIS和资源QoS，根据用户业务请求分解出四要素资源使用量。
2. 基站：四要素资源的配置，根据资源QoS针对一个请求配置资源，并执行AI推理。

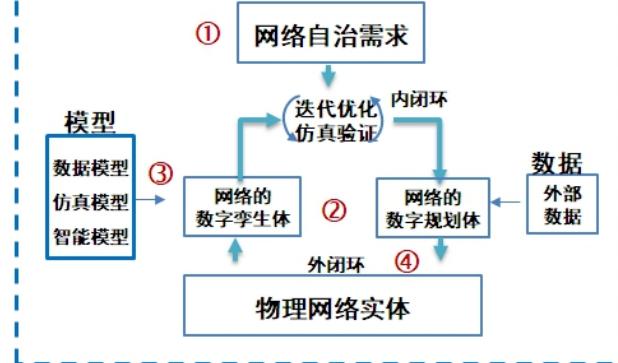
数字孪生网络基本功能

- 生成和解析网络自治需求
- 构建、编排和调整数字孪生体和数字规划体模型
- 生成和更新网络的数字孪生体
- 生成和实施网络的数字规划体



数字孪生网络端到端架构

数字孪生网络

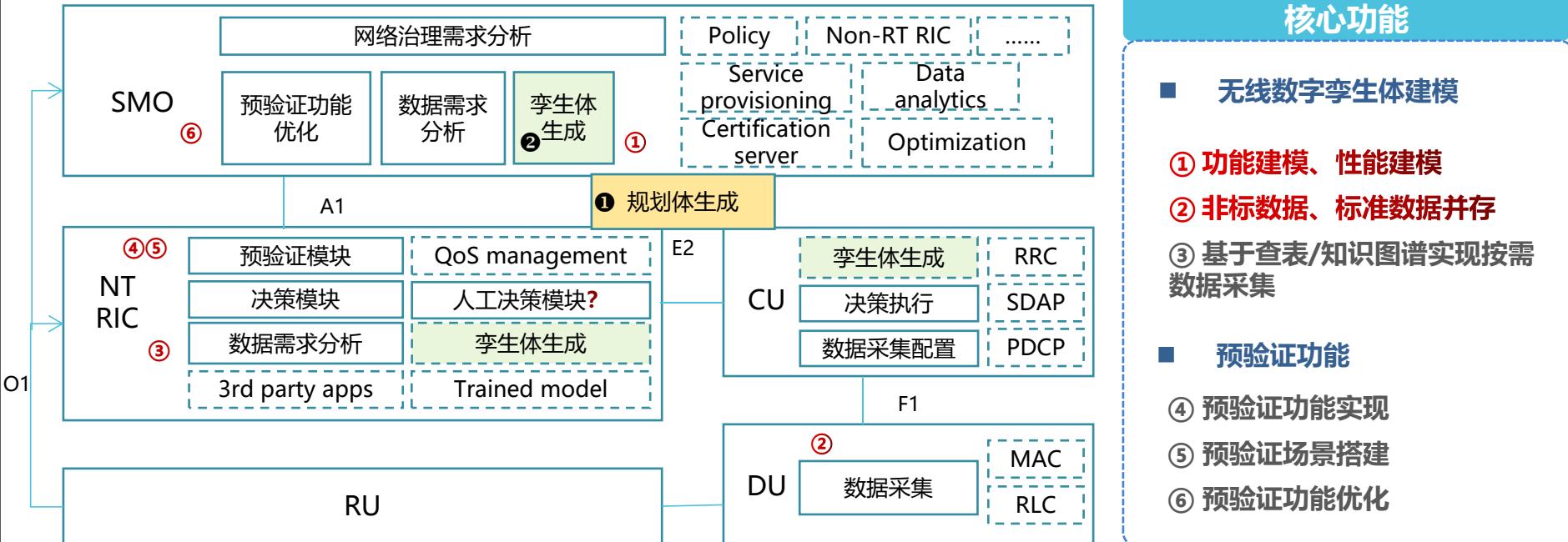


数字孪生网络基本概念间的关系

关键技术特征

- 集中与分布相结合
- 按需建模、高效交互的数字孪生体
- 低成本预验证试错环境
- 内闭环数字域寻优，外闭环校正

数字孪生网络逻辑功能部署选项——以ORAN为例

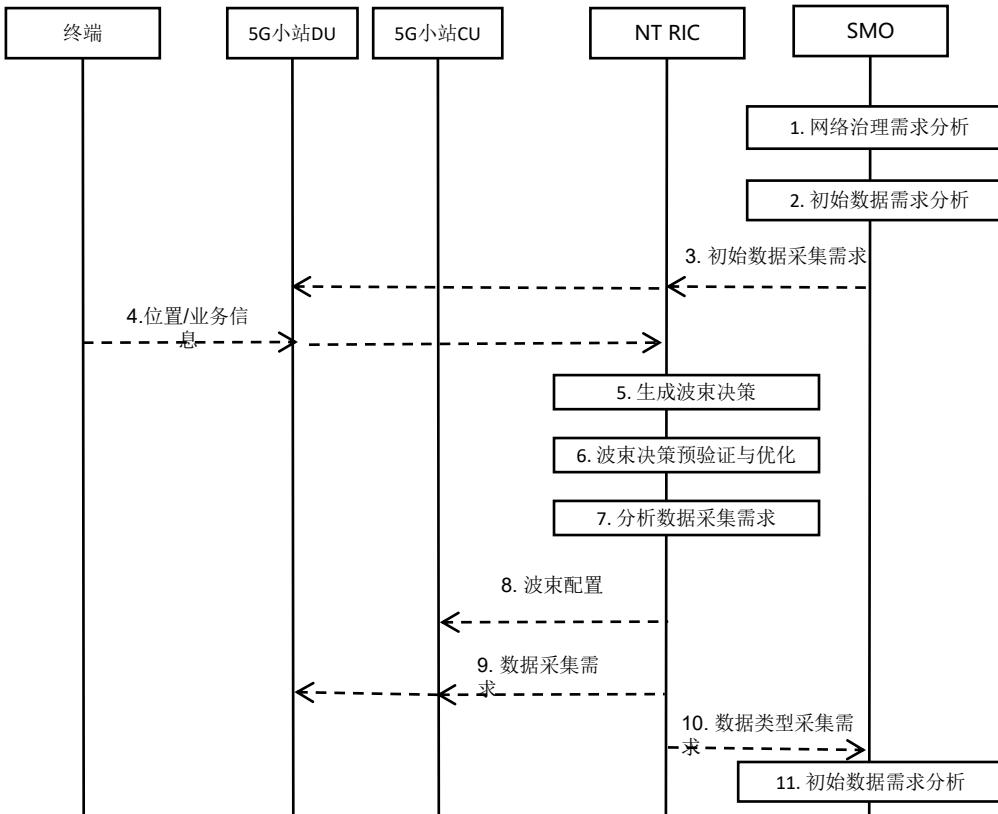


孪生体/规划体生成

① 对于实时性要求高的用例，“规划体生成”位于NT RIC或CU；实时性要求低的用例，“规划体生成”位于SMO；规划体生成涉及决策算法的迭代优化

② “孪生体生成”模块的位置依据用例的实时性需求确定

大规模天线权值优化用例示例



■ MIMO用例具体流程：

1. SMO生成波束权值优化用例；
2. SMO分析初始数据采集需求；
3. NT RIC根据初始数据采集需求设置小站数据采集配置；
4. 终端上报位置和业务信息；
5. 波束决策模块生成波束决策；
6. 波束决策在预验证模块中进行预验证，波束决策模块根据预验证的性能优化波束决策，直至满足需求；
7. 数据需求分析模块根据预验证性能分析数据采集需求，如数据采集类型与频次；
8. 波束配置决策下发至5G小站；
9. 新的数据采集需求下发至5G小站；
10. 如新的数据采集需求涉及采集数据类型变化，新的数据类型采集需求上传至SMO，更新初始数据采集需求；

■ 功能模块：

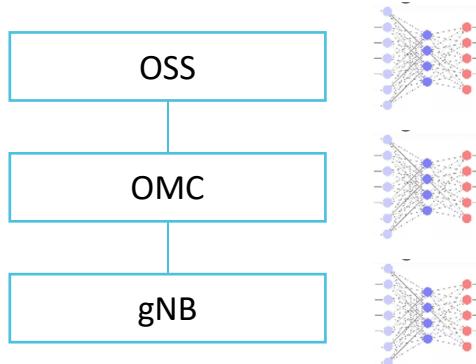
预验证：可预测未来网络状态；接收采集、增广与生成数据更新预验证环境；与波束权值生成模块交互，返回预验证的性能，辅助波束配置优化；

数据需求分析：基于查表/知识图谱分析需要采集的数据类型；

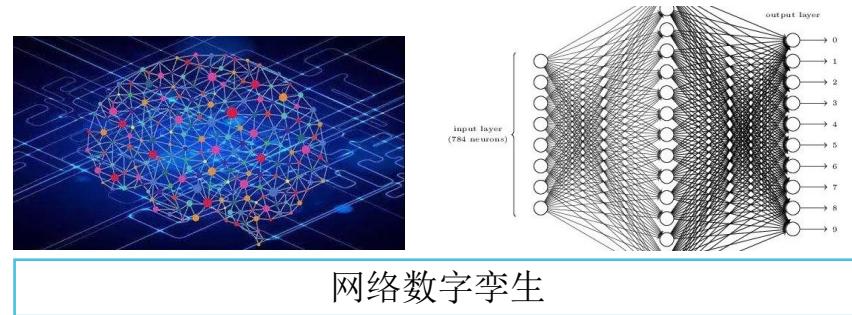
■ 新增接口信息：

1. AI接口：模型参数；
2. E2接口：数据采集需求，含类型、频次等；

5G：独立的、局部应用所驱动的小模型



6G：通用网络智能大模型



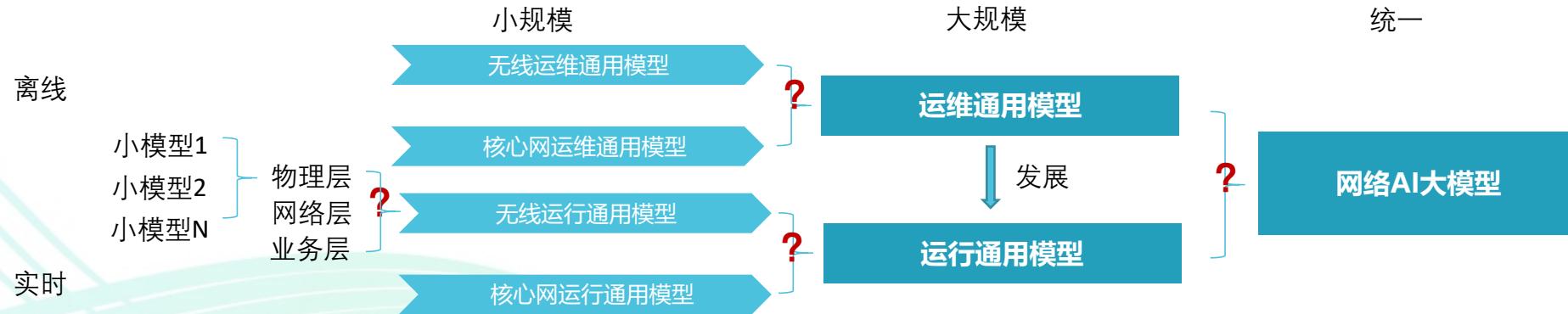
网络数字孪生

- 技术特点：模型参数较少，结构简单，复杂度低，训练推理速度快，可解释性较高
- 优势：轻量化，无需网络架构适配AI部署
- 劣势：模型迁移和泛化性不佳，建模场景有局限性，分域而治导致重复性研发

- 技术特点：模型参数大，数据规模大，算力大
- 优势：通用性好，容易裁剪成小模型，针对部分特定场景只需要少量调参
- 劣势：成本极高（包含算力、数据集构建、人工调优等），需要极为庞大的预训练数据

- 网络智能化应用工作分散到不同厂家和领域，场景数据规范性不相通，存在数据重复采集、资源浪费等问题。
 - 网络智能化应用的场景种类繁多导致场景碎片化。
 - 网络智能化应用的应用分散，AI 模型的通用性较低，各个应用之间缺少协同优化，模型无法复用和积累。

大模型



网络运营

网络直接调用**封装的大模型服务**，利用大模型多模态机器学习、语言理解与文本生成能力，**提升网络的信息通信服务能力**。

- 增强智能客服的语义理解、情感识别
- 增强分析，实现智能化经营分析
- 智能营销

网络运维

通信网络可以通过网络内部接口进行交互，大模型作为一种**网络内部功能或者网元**，协助智能运维AIOps 助力运维人员**快速处理系统故障**。

- 故障识别
- 故障定位
- 故障自愈
- 网络优化

网络运行

无线信道建模：根据无线信道上输入的数据实现信道生成，利用多模态输入输出技术补充或完善已有的数据/图片/视频集，并可根据场景描述生成场景。

语义通信：利用ChatGPT作为系统中的编码器（encoder）和解码器（decoder）模块来实现语义编码和解码。

通信网络大模型：通信挑战

场景	对数据的要求	对模型的要求	对网络架构的影响	对网络的影响程度	大模型研发工作量	可行性
网络运营	多模态数据:通信行业属性、客户服务对话、行业标准和法规文件、在线音频、巡检图像、渠道信息	多模态机器学习、语言理解,逻辑推理、文本生成	大模型作为网络外部的一种 远程服务 ,可以直接调用	小	小	高
网络运维	标准化数据: KPI、XDR、MR、拓扑等资源 非标准化数据: 日志、图、文档/案例等	非标数据治理、数据对齐, 多模态、自然语言理解,逻辑推理、代码生成	大模型作为一种 网络内部功能 ,通信网络可以通过网络内部接口进行交互,需要在已有大模型自然语言理解能力上增强丰富的运维能力,故障的解决方案或修复策略知识库	中	中	待评估
网络运行	非标准化物理层数据:包括上下行信道、语义信息、语义特征、小区内/间干扰 标准化数据网络层,业务层,服务级包括: MR测量数据、MDT数据,业务数据流信息数据 (IP五元组、URL、PFD等),控制面信令数据、业务数据流信息数据	逐步推理,模型压缩(减少资源),自适应推理(跳过不重要层,速度快)	将大模型功能 嵌入 已有的应用和服务中,或者将大模型作为网元形式与网络架构进行融合,以保证实时性。	大	大	待评估

有待研究的技术问题

序号	内生AI
1	如何基于三层QoAIS指标体系，设计面向网络自治用例的QoAIS指标？
2	面向QoAIS保障，如何以AI任务为中心，实现灵活的多维资源编管控？
3	网络运维和网络运行是否存在通用智能大模型？对数据有何要求？代价如何？若存在，网络如何满足运行态大模型对存储、算力和数据的需求？是否影响网络架构？

序号	数字孪生网络
1	如何对无线网元和网络进行功能和性能建模？如何依据不同网络自治场景的需求，按需生成网元和网络的数字孪生体？
2	如何构建低成本、满足精度和实时性需求的数字孪生预验证环境和迭代优化闭环？
3	通用网络AI大模型与网络数字孪生的关系？与网络数字孪生模型的关系？

谢 谢！

