中国信科 中信科移动 CICT Mobile

Auto-RAN
意图驱动的6G无线自智网络

AUTO RAN: INTENT-DRIVEN 6G WIRELESS AUTONOMOUS NETWORKS

基于北极星领航通信大模型的AUTO-RAN技术



中信科移动通信技术股份有限公司

无线移动通信全国重点实验室(中国信科)

Auto-RAN：意图驱动的 6G 无线自智网络

—基于北极星领航通信大模型的 Auto-RAN 技术白皮书—

本白皮书版权专属中信科移动通信技术股份有限公司(以下简称“中信科移动”)和无线移动通信全国重点实验室所有，并受法律保护。如需基于非商业目的引用、转载、传播或以其他方式合理使用本技术白皮书的全部或部分内容，应完整注明来源。违反前述声明者，中信科移动将追究其法律和商业道德之责任。

摘要

面向 6G，移动通信服务对象正由人与设备进一步拓展至智能体，业务需求由连接与数据传输延伸至通信、感知、计算和智能融合，网络运行也由人工配置与规则驱动逐步转向意图驱动和自主闭环。与此同时，空天地网络融合、业务需求差异化以及网络功能持续扩展，使无线网络面临性能提升边际收窄、建设运营成本承压和系统复杂度快速上升等挑战。传统依赖功能叠加、局部优化和人工运维的演进方式，已难以支撑 6G 网络对动态环境、复合任务和确定性体验的保障需求。

为应对上述变革，本白皮书提出 Auto-RAN（Autonomous Radio Access Network，自智无线接入网）技术体系，以用户业务意图为起点，将端到端服务目标逐级转化为无线网络可理解、可编排和可执行的保障任务，推动网络从“保障连接”向“意图即结果”演进。

Auto-RAN 遵循 Agentic、Unified、Trusted、Open 四项设计理念，构建意图任务解析与智能编排中枢，以及区域网络智能、多站协同智能和空口内生智能三级智能体系。编排中枢负责任务状态管理、能力选择、资源编排和冲突仲裁；不同层级的网络智能体根据作用范围和实时性要求协同执行，实现从业务意图解析、无线专业意图生成到策略执行、效果评估和持续优化的闭环。

在关键技术层面，Auto-RAN 以北极星领航通信大模型提供面向通信网络的语义理解、知识推理和候选方案生成能力，以多智能体协作执行跨专业、跨层级任务，以数字孪生支撑策略验证、风险评估和演进训练，并通过分级部署、能力开放的算智基础设施满足不同任务的时延、算力和可靠性需求。大模型、智能体、专业模型、传统算法和网络工具在统一编排机制下各司其职，共同形成可演进、可验证、可管控的无线网络智能体系。

通过上述架构，Auto-RAN 旨在实现体验可承诺、能效可持续、性能自优化、故障自恢复、能力可开放和价值新增长，提升复杂场景下网络资源的协同效率与服务确定性，并为网络能力产品化和新型商业模式提供基础。本白皮书进一步阐述 Auto-RAN 的设计理念、总体架构、关键技术、典型场景及从 5G-A 试点验证到 6G 规模应用的演进路径，推动形成开放协同、安全可信的 6G 智能内生产业生态。

目录

Auto-RAN：意图驱动的 6G 无线自智网络 0

摘要 1

1 从人工运维时代走向意图驱动算网智能时代的产业变革 5

1.1 变革一：服务对象与业务需求重构 5

1.2 变革二：网络角色与基础设施形态重构 6

1.3 变革三：网络运行与生产方式重构 7

2 6G 智能内生网络的价值和愿景 9

2.1 6G 智能内生网络的产业价值 9

2.2 6G 智能内生网络的愿景 10

2.3 6G 网络自智面临的技术挑战 12

3 Auto-RAN：意图驱动自智无线接入网 14

3.1 Auto-RAN 设计理念 14

3.2 Auto-RAN 的技术架构 16

(一) 总体架构 17

(二) 意图驱动流程 18

(三) 意图任务解析与智能编排中枢 19

(四) 三级智能层：面向不同网络任务分层注智 20

(五) 跨层协同与信息交互 21

(六) 可信执行与目标达成评估.....21

3.3 Auto-RAN 的核心价值.....22

(一) 体验可承诺.....22

(二) 能效可持续.....23

(三) 性能自优化.....24

(四) 故障自恢复.....24

(五) 能力可开放.....25

(六) 价值新增长.....25

4 Auto-RAN 的关键使能技术.....26

4.1 Agentic-多智能体协作.....26

4.1.1 智能体矩阵与分层.....26

4.1.2 多智能体协作机制.....28

4.2 Unified-北极星领航通信大模型.....30

4.2.1 通信领域数据集构建.....30

4.2.2 意图理解与编排增强.....31

4.2.3 大-小模型分层协同.....32

4.3 Trusted-数字孪生.....33

4.4 Open-算力分级部署与开放.....35

5 Auto-RAN 路线图：从 5G-A 试点验证迈向 6G 规模商用.....37

6 参考文献 39

1 从人工运维时代走向意图驱动算网智能时代的产业变革

移动通信产业正由连接规模扩张转向智能能力升级。服务对象从人与终端扩展到机器、物和智能体，网络资源从频谱与连接拓展到感知、算力、数据和模型，运行方式由人工配置和规则自动化走向意图驱动的分层自治。这既增加了网络复杂度，也为通信设备企业带来产品重构、能力延伸和价值增长的新机遇。

1.1 变革一：服务对象与业务需求重构

6G 智能内生网络打破了传统通信网络“通信为主、AI 外挂”架构模式，将环境感知、智能计算、算力调度、数据治理等多维能力嵌入网络底层基座，构建自感知、自决策、自优化、自闭环的原生智能体系，既是 6G 通感算一体网络自主全域优化的基础支撑，也是垂直行业差异化场景落地的核心载体。其对移动通信产业的核心价值主要体现在网络自主赋能和场景价值延伸两个维度。

长期以来，移动通信网络主要围绕人与智能终端构建，重点满足语音、网页、视频和社交等信息消费需求，网络建设与优化主要关注覆盖、容量和数据流量。进入智能时代，网络连接对象正由个人用户拓展至人、机、物和具身智能等多元主体，网络功能也由信息传输向通信、感知、计算、智能和安全等能力融合演进[1]。

截至 2026 年 7 月，我国移动物联网终端用户达到 29.93 亿户，移动电话用户为 18.47 亿户，前者相当于后者的 1.62 倍[2]。2025 年，移动物联网终端占移动网终端连接数的 61.3%，其中公共服务、车联网、智慧零售和智慧家居终端分别达到 11.64 亿、5.29 亿、3.65 亿和 3.23 亿户；移动物联网终端接入流量同比增长 42.7%，明显高于移动互联网接入流量 17.3% 的增幅[3]，机器、设备和感知终端已经成为移动网络的重要使用主体。

到 2026 年 6 月，我国已建成超过 8 万个 5G 行业虚拟专网，5G 和千兆光网应用融入 94 个国民经济大类，“5G+工业互联网”在建项目超过 2.6 万个，建成 1260 家 5G 工厂[1]。网络开始承载工业传感器、机器视觉设备、AGV、机械臂、无人机和生产控制系统，连接质量日益直接影响任务执行和生产过程。

人工智能将进一步推动服务对象由“联网机器”走向“自主智能体”。到 2026 年底，集成任务型 AI 智能体的企业应用比例预计将由 2025 年的不足 5%提升至 40%；到 2028 年，专业智能体将能够跨多个应用和业务系统动态协作，用户可直接提出目标，而不必逐一操作具体应用[4]。智能体任务将通过各类终端、边缘节点和云平台转化为网络连接、数据传输与资源调用。

未来网络面对的不只是流量增长，更是需求性质的改变。传统人联网业务多为下行主导和尽力而为，机器与智能体业务则强调上行感知、双向交互、任务时限、可靠执行和跨节点协同。网络服务目标需要由“传送数据”扩展为“保障任务完成”，根据业务目标协同无线、计算和智能资源。

1.2 变革二：网络角色与基础设施形态重构

传统移动通信网络以连接为中心，核心生产资源主要包括频谱、站址、无线设备、传输网络和核心网平台。进入智能时代，网络还需处理数据采集、环境感知、算力供给、模型推理和任务控制，向通信、感知、计算、智能和安全融合演进。

截至 2026 年 6 月，我国智能算力规模达到 2185 EFLOPS，同比增长 177%[1]。截至 2025 年底，三家基础电信企业对外提供服务的数据中心机架达到 93.8 万架，较上年增加 10.8 万架，可调度智能算力规模超过 94.4 EFLOPS，同比增长 87.6%[3]。两组数据统计范围不同，不宜直接比较，但都表明算力正在成为与连接能力并列发展的基础设施资源。

运营商的资源体系和业务结构也在改变。2025 年，中国电信自有及接入智算规模达到 91 EFLOPS，天翼云、AIDC 和智能业务收入分别达到 1207 亿元、345 亿元和 123 亿元，并逐步形成涵盖算力、平台、数据、模型和应用的智能云体系[5]。中国联通同期智算规模达到 45 EFLOPS，算力业务收入占服务收入的比例超过 15%，2026 年计划将超过 35%的资本开支投入算力领域，经营布局也由连接业务扩展到连接、算力、服务和安全等领域[6]。算力、数据、模型和智能服务正在进入运营商的核心生产体系。

面向实时推理、工业控制、低空应用和具身智能等业务，将计算任务全部传送至中心云，可能增加传输开销、时延和数据安全压力。未来基础设施将更多采用中心云、区域云、网络边缘、基站和终端协同的分布式形态，根据业务要求选择处理位置。

无线接入网位于连接用户、汇聚数据和感知环境的前沿。基站及其边缘节点可在条件允许时承载环境感知、位置服务、无线数据处理、轻量化模型推理和本地智能控制，由通信网元逐步演变为连接、数据和边缘智能交汇的综合节点。

这种变化并不要求每个基站都部署大规模 GPU，也不意味着无线接入网将取代数据中心。到 2030 年，全球数据中心需求可能增长至目前的三倍以上，达到至少 170 GW；与此同时，生成式人工智能和智能体业务正在推动推理算力向用户侧分布，运营商已有的光纤网络、边缘站址、机房空间、电力资源和高可靠运营能力，可能成为建设分布式智能基础设施的重要基础[7]。分布式算力和 AI-RAN 仍面临需求规模、架构适配和总体拥有成本等不确定性，其部署范围最终取决于业务需求和投入产出[7]。

网络角色变革的核心，是由主要提供连接资源，走向通信、感知、计算、数据和智能资源协同供给，根据业务目标确定数据获取、模型运行、算力分配和网络保障方式。这推动 Auto-RAN 由无线连接管理走向无线侧综合资源智能编排。

1.3 变革三：网络运行与生产方式重构

传统移动通信网络主要采用以专业分工、参数配置和固定流程为核心的运行方式。网络人员根据告警、性能指标和用户投诉分析问题，依靠经验调整网络参数。现有自动化系统虽可执行批量配置和预设流程，但目标、触发条件和步骤仍需人工定义，本质上是将专家经验固化为规则。

当前网络规模和管理对象已经显著变化。除传统公网外，5G 行业专网、边缘计算、卫星通信、低空网络和行业应用不断进入运营体系；同一网络并存多种制式、频段、站型、终端和业务等级，网络状态又随用户移动、业务潮汐、无线环境和算力负荷持续变化[2]。人工分析和静态规则越来越难以及时处理海量状态及其复杂关联。

人工智能已经应用于网络节能、故障诊断、流量预测、体验分析和资源优化，但多数仍局限于单一场景。当前几乎所有受调研运营商都在开展 AI 试点，但只有 57%将 AI 扩展到多个业务域，仅 16%认为 AI 已经成为日常运行中的常态；超过四分之三的运营商将运行模式不成熟、数据基础不足和变革管理滞后视为规模应用的主要障碍[9]。局部算法和分析工具尚难改变网络整体运行方式。

产业实践正在由“AI 辅助人”向“人机协同闭环”演进。2025 年上半年，中国移动已有 8 个网络运营高价值场景达到 L4 级自智能力[8]。2025 年，中国电信在云网运营中部署 900 余个数字员工，使网络故障处置效率提高 30%，云网运营自智水平达到 L4 级[5]。这些成果主要集中在边界明确、目标可度量、风险可控制的场景，已经能够形成从状态感知、问题分析到策略执行和结果评估的自动闭环。

AI 驱动的网络运营可使总体运营支出降低 15%～30%；在实现规模应用的场景中，故障处理工单可减少 30%～70%，网络运营中心成本可降低 55%～80%，平均修复时间可缩短 30%～40%[10]。这些收益不仅来自模型能力提升，也来自故障发现、根因定位、方案

生成、任务派发和执行验证等流程的重新组织。网络智能化的重点正由单项算法优化转向端到端流程重构。

新的运行方式首先体现为控制入口的变化。意图驱动模式允许用户、业务系统或运营人员直接描述目标、作用范围、资源边界和安全约束，由网络完成意图解析、任务分解、资源评估和方案生成，人的职责转向目标、规则和权限管理。

其次，网络任务将由固定流程执行转向智能体动态协作。覆盖、容量、能效、移动性、故障和体验等功能根据业务目标与实时状态参与任务，协调资源竞争和策略冲突，并根据执行效果调整、回退或重新规划，形成“目标—感知—决策—执行—评估”的持续闭环。

这种变革不意味着由一个大模型直接控制所有网元，也不意味着取消人工干预。更现实的形态是由大模型承担意图理解、知识推理和任务编排，专业模型和网络智能体负责领域决策，传统算法及协议功能完成确定性实时执行，并通过权限控制、策略验证、异常熔断和人工接管保障安全。

网络运行方式的根本变化，是将人、模型、智能体、传统算法和网络系统重新组织为新的生产体系。人的角色由操作执行者转变为业务目标制定者、自治规则管理者和运行风险监督者；网络则在授权和安全边界内自主完成更多分析、决策、执行和反馈任务。由专家参数和固定流程驱动走向意图驱动的分层自治闭环，正是 Auto-RAN 需要承接的核心产业变革。

2 6G 智能内生网络的价值和愿景

2.1 6G 智能内生网络的产业价值

6G 智能内生网络打破了传统通信网络“通信为主、AI 外挂”架构模式，将环境感知、智能计算、算力调度、数据治理等多维能力嵌入网络底层基座，构建自感知、自决策、自优化、自闭环的原生智能体系，既是 6G 通感算一体网络自主全域优化的基础支撑，也是垂直

行业差异化场景落地的核心载体。其对移动通信产业的核心价值主要体现在网络自主赋能和场景价值延伸两个维度。

6G 智能内生网络将解决传统网络人工运维成本高、优化滞后、适配性弱的行业痛点。通过内置原生智能能力，可实现网络全生命周期的自主管控与动态优化，保障网络核心性能和运维水平持续演进。在网络运行层面，依托内生 AI 引擎可完成无线环境自适应识别、干扰智能抑制、小区负载均衡、移动性动态优化等基础能力，替代传统人工参数配置与静态调度模式，大幅提升全网运行稳定性、资源利用率与复杂环境适配效率。在运维管控层面，基于标准化数据采集、智能故障研判与根因分析能力，实现网络故障主动预警、快速定位、极简自愈，有效降低全网运维人力成本与网络故障时长。

6G 智能内生网络将突破传统通信网络仅提供连接服务的单一价值，实现从“通用连接”向“场景化智能服务”的产业价值跃迁，是推动移动通信网络从公共基础服务向垂直行业深度赋能升级的核心支撑。传统移动通信网络以全域通用覆盖、标准化组网服务为核心，能力范式统一、功能迭代固化，难以适配各垂直行业碎片化、定制化、高专属的业务需求。6G 智能内生网络依托多级算力架构与分层优化能力，可面向工业专网、低空安防等垂直细分场景的差异化需求，构建高适配、轻量化、可落地的场景智能服务体系，实现网络服务能力与行业需求的高效匹配，完成场景化业务赋能。

2.2 6G 智能内生网络的愿景

5G-A 已经将人工智能应用于网络规划、节能、负载均衡、移动性管理、故障诊断和用户体验优化等场景，初步验证了 AI 提升网络性能和运行效率的价值。但现有智能能力多数仍以单一场景、单一网元或单一指标为中心，模型之间缺乏统一目标和协同机制，AI 主要作为传统网络功能的增强工具存在。

面向 6G，人工智能将进一步融入网络架构、协议功能和运行流程，成为具有网络感知、决策、控制和演进的基础能力。智能内生不是由某一种模型替代现有网络功能，也不是将所有任务集中交由大模型处理，而是在网络设计之初就统筹考虑数据、模型、知识、算力和闭环控制，使智能能力能够部署在适合的网络层级和位置，并贯穿网络规划、建设、运行、优化和维护全过程。

6G 智能内生网络首先需要实现由业务意图驱动网络运行。传统网络管理主要通过配置参数、操作指令和预设策略控制网络行为，高层业务需求需要经过多级人工分析才能转化为网络动作。意图驱动模式下，用户、业务系统或网络运营人员主要描述希望网络达到的业务目标、服务指标和约束条件，网络自主完成意图理解、任务分解、资源评估、策略生成和执行验证。

例如，面对特定区域内的应急通信、工业控制或低空业务需求，业务方可以直接提出覆盖范围、上行速率、端到端时延、可靠性和保障时长等要求。网络根据实时资源状态评估目标的可实现性，协调无线、传输、核心网和计算资源，形成相应的保障方案，并在业务运行期间持续监测目标完成情况。当网络环境或业务负荷发生变化时，系统能够动态调整策略，维持业务所需的服务水平。

其次，6G 智能内生网络需要形成分层、分布式的智能体系。无线网络同时存在中长期规划、区域级资源协同和毫秒级空口控制等不同时间尺度的任务，不同任务对数据范围、推理能力、执行时延和可靠性具有不同要求。网络应根据任务特征，将智能能力分布在中心云、网络管理域、边缘节点和基站等不同位置，由全局智能负责目标理解、知识推理和长期规划，由区域智能负责资源预测和跨网元协同，由网元侧智能负责实时感知和快速执行。

不同层级的智能闭环既需要保持相对独立的决策和执行能力，又需要共享必要的网络状态、业务目标和知识信息。上层智能向下层传递目标、约束和策略边界，下层智能根据实时

环境自主选择执行方式，并将运行结果和状态变化反馈至上层。通过不同时间尺度闭环之间的协同，网络能够兼顾全局资源效率与局部实时性能。

再次，6G 智能内生网络需要具备持续学习和自主演进能力。网络运行环境具有明显的动态性，不同区域、时段、用户和业务所形成的数据分布持续变化，固定规则和静态模型难以长期保持最佳效果。网络需要持续获取运行反馈，评估模型和策略的实际效果，识别环境变化和模型性能退化，并在满足安全条件的前提下完成模型更新、策略调整和知识积累，使智能能力能够随网络和业务共同演进。

智能内生还必须建立在安全可信的基础上。网络需要明确不同智能单元的数据访问范围、决策权限和操作边界，对模型输入、推理结果和执行策略进行必要的校验。对于影响范围较大或风险较高的网络操作，应支持执行前验证、运行中监测、异常熔断和策略回退，并保留完整的决策及执行记录。网络智能化程度越高，对可解释、可审计和可接管能力的要求也越高。

在这一愿景下，人的作用不会被简单替代，而是由具体操作的执行者转变为业务目标、运行规则和安全边界的制定者。网络在授权范围内自主完成分析、决策和执行，当任务超出能力或权限范围时主动请求人工介入。人与网络由传统的指令执行关系，逐步转变为目标设定、自治运行和过程监督相结合的以业务意图驱动 6G 网络自智架构。

2.3 6G 网络自智面临的技术挑战

从 5G-A 的场景化 AI 走向 6G 智能内生网络，需要实现从单点功能增强到系统性智能协同的跨越。其难点不仅在于提高模型精度和扩大应用范围，更在于建立能够贯穿业务目标、网络决策和空口执行的智能运行体系。

首先，高层业务意图与底层网络操作之间存在明显的语义和能力差异。用户及业务系统通常以自然语言或服务指标描述业务目标，而无线网络实际执行的是功率、载波、波束、调度权重、切换门限和资源预留等具体操作。业务意图可能存在表达不完整、指标冲突或超出网络能力等情况。网络需要准确识别意图中的目标、约束、优先级和作用范围，将其转换为可度量、可分解和可验证的网络任务，并在资源不足时提供合理的目标协商机制。

其次，不同智能功能之间存在协同与冲突问题。网络中的性能优化、能效控制、体验保障、故障恢复和资源调度具有不同目标，单独运行时均可能获得局部收益，但多个功能同时作用于同一网络对象时，可能造成资源竞争、策略冲突或反复调整。例如，节能策略可能减少可用网络资源，性能优化策略则可能提高功率或开启更多硬件资源。如何建立统一的目标体系和冲突协调机制，使各类智能功能在业务目标和全网价值约束下协同工作，是实现网络自智的重要前提。

第三，不同时间尺度的智能闭环需要保持协调一致。网络规划和模型训练可能以小时或天为周期，跨小区协同和业务保障可能以秒或分钟为周期，而调度、波束和链路自适应需要在毫秒级完成。集中式智能具有全局信息和较强推理能力，但难以满足严格实时性要求；网元侧智能能够快速响应，却受到算力、数据范围和模型规模限制。需要合理划分不同层级智能的职责、信息范围和执行权限，并解决目标下发、状态反馈和策略一致性问题。

第四，大模型的通用推理能力与通信网络的专业性、实时性和确定性要求之间仍存在差距。通信网络包含大量专业概念、配置关系、时序数据和因果联系，通用大模型难以仅依靠自然语言知识准确理解复杂网络状态。同时，大模型输出具有一定不确定性，无法直接承担要求严格确定性和极低时延的空口控制任务。因此，需要构建通信领域知识增强、专业模型协同和工具调用机制，形成大模型、任务专属模型与传统通信算法分工协作的模型体系。

第五，数据和模型的全生命周期管理仍需完善。无线网络数据来源广泛、类型多样，不同厂家设备、网络区域和运行阶段之间存在数据口径及分布差异。模型训练和运行需要解决数据质量、样本标注、隐私保护、权限控制和跨域流动等问题。模型部署后还需要持续监控性能，识别模型漂移和环境变化，并完成模型版本管理、更新验证及异常回退。数据和模型管理能力将直接决定网络智能能否由实验室验证走向现网长期运行。

第六，智能决策必须满足通信网络对安全性和可靠性的严格要求。网络操作可能影响大范围用户和关键业务，模型误判、工具调用错误或智能体之间的异常交互都可能放大运行风险。智能系统需要具有明确的权限控制、动作约束和安全保护机制，并能够对关键决策进行验证。对于不可接受的策略或异常状态，系统应及时停止执行、恢复至安全配置，并支持人工接管。

第七，智能能力需要与现有网络平滑演进。5G-A 网络已经形成相对稳定的协议体系、网管流程和设备架构，6G 智能内生能力不可能一次性替代现有系统。新的智能能力需要兼容现有网络接口和运行机制，支持按场景、按区域、按网络层级逐步部署，并能够与传统规则、现有自动化功能和人工操作长期共存。只有建立清晰的能力边界和渐进式演进路径，才能降低现网引入智能技术的成本与风险。

上述挑战表明，6G AI 的核心并非简单增加模型数量，而是建立一套能够理解业务目标、协调多类智能、适配不同时间尺度并保证安全可信的网络智能体系。该体系需要统一组织数据、模型、知识、算力和网络资源，形成从感知、认知、决策到执行和反馈的完整闭环。

3 Auto-RAN：意图驱动自智无线接入网

3.1 Auto-RAN 设计理念

面向算网融合、业务差异化和 6G 智能内生发展趋势，本白皮书提出 Auto-RAN（Autonomous Radio Access Network，自智无线接入网）技术体系。Auto-RAN 以业务意图为牵引，以北极星领航通信大模型为认知与规划基础，以专业网络智能体和任务模型为执行实体，通过分层分布式智能闭环，推动无线网络由人工配置和局部自动化，向目标驱动、协同决策和持续演进的自智运行模式转变。

Auto-RAN 的核心目标是打通业务需求与无线网络能力之间的转换链路。用户、行业应用或运营系统提出业务目标后，Auto-RAN 识别其中的业务对象、服务指标、时间空间范围、优先级和资源约束，将高层意图逐步转换为可执行的网络任务。系统根据网络实时状态评估目标可实现性，选择适当的智能能力和网络工具完成任务，并持续监测执行结果，形成“意图输入、任务规划、资源协调、策略执行、结果验证和反馈调整”的闭环。

Auto-RAN 不是以单一大模型取代现有无线网络功能。不同网络任务对实时性、专业性、算力和可靠性的要求不同，需要采用分层分工的智能实现方式。北极星领航通信大模型主要承担意图理解、知识组织、任务规划和多任务协调；专业网络智能体面向规划、优化、能效、体验和故障等具体领域完成任务执行；任务专属模型负责状态预测、特征识别和策略生成；对于毫秒级空口控制任务，则由基站侧轻量化模型和确定性通信算法完成实时处理。

Auto-RAN 遵循 Agentic、Unified、Trusted 和 Open 四项基本设计理念。

Agentic 强调以智能体为载体组织网络智能能力。Auto-RAN 将相对独立的网络任务封装为具有明确目标、知识、工具和权限的专业智能体。智能体能够感知网络状态、理解任务要求、调用网络工具并评估执行结果，同时可以在统一目标约束下开展协作。智能体不脱离现有网络功能独立运行，其行为范围由所属网络域、业务职责和授权边界共同确定。

Unified 强调数据、模型、知识、算力和网络资源的统一协同。Auto-RAN 建立统一的意图描述、网络语义和运行上下文，使不同层级、不同专业的智能单元能够准确理解业务目

标和网络状态。系统根据任务特征选择适当的数据、模型和算力资源，并协调不同网络资源，减少信息割裂、重复建设和局部优化之间的冲突。

Trusted 强调网络智能全过程可信可控。Auto-RAN 在意图输入、任务规划、模型调用、策略生成和网络执行等环节设置安全与可靠性约束。通过身份认证、权限控制、策略校验、仿真验证、运行监测、异常熔断和策略回退等机制，保证智能决策不突破网络安全边界。系统记录关键任务的决策依据、工具调用和执行结果，为运行审计、问题追溯和模型改进提供支撑。

Open 强调智能能力和网络能力的开放演进。Auto-RAN 通过标准接口、API、Skill 和智能体能力封装，对内连接不同网络域和专业系统，对外支持行业应用及合作伙伴按权限调用网络能力。开放并不意味着无边界访问，而是在统一身份、权限和安全策略控制下，实现数据、工具和能力的规范化调用，使新的模型、智能体和业务应用能够持续接入和迭代。

基于上述设计，Auto-RAN 形成高层认知与底层执行分工、全局目标与局部自治协同、智能决策与安全约束并重的技术体系。其目标不是追求完全脱离人工控制的网络，而是在明确业务目标、运行规则和安全边界的基础上，使网络能够自主完成更多复杂任务。网络运行人员由具体参数配置和重复操作的执行者，逐步转变为业务目标制定者、自治规则管理者和运行过程监督者。

Auto-RAN 将以现有 5G-A 网络为基础逐步演进，通过单场景智能应用、跨功能协同闭环和端到端意图驱动等阶段不断扩大自治范围，积累数据、模型和现网运行经验，为 6G 智能内生无线接入网的技术验证和规模部署奠定基础。

3.2 Auto-RAN 的技术架构

面向 6G 业务形态持续丰富和网络能力多样化的发展趋势，传统以网元参数

配置为主的网络管理方式难以直接承接用户和行业应用对服务效果的要求。Auto-RAN 构建面向业务目标的意图驱动网络架构，在用户、核心网、操作维护中心（OMC）与无线接入网之间建立统一的目标传递和结果反馈关系，使网络能够从业务需求出发组织无线服务，并以目标达成情况衡量网络运行效果。

(一) 总体架构

意图任务解析与智能编排中枢是 Auto-RAN 的组织核心，承接 OMC 下发的 RAN 保障子意图，将其解析为无线专业意图，进一步生成可执行的自智任务与执行策略。中枢统一配置任务目标、控制对象、资源需求、风险等级和评价指标，选择适配的模型、智能体、算法和工具，并确定执行位置及协同关系。

三级智能层按照网络任务的作用范围和处理要求形成分工。区域网络智能层面向区域网络和跨站群任务，多站协同智能层面向基站群和局部区域协同，空口内生智能层面向用户、链路和协议处理过程。基站、无线协议功能及相关网络节点依据决策结果完成网络执行，并将网络状态、无线测量、业务体验和异常信息反馈至相关智能层和意图任务解析与智能编排中枢。

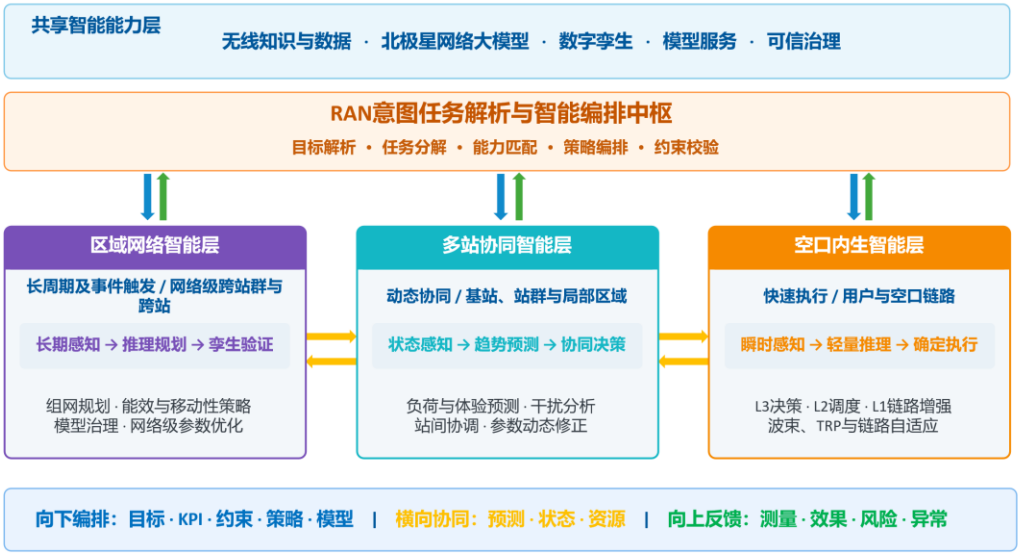


图 3-1 Auto-RAN 无线专业意图解析编排与三级智能协同

(二) 意图驱动流程

意图从业务语义到底层执行依次经过用户业务意图、端到端服务意图、RAN 保障子意图和无线专业意图四个层级。用户业务意图由用户、行业应用或终端智能体提出，可采用自然语言、结构化指标或业务模板描述服务对象、业务类型、作用区域、时间范围、体验指标、优先级和约束条件，重点表达期望获得的业务体验和服务结果，无需直接指定功率、波束、调度权重和切换门限等无线参数。

核心网结合用户签约、业务特征、会话状态和网络能力，对用户业务意图进行准入、协商和服务映射，形成可度量的端到端服务意图。端到端服务意图明确业务连续性、时延、可靠性、体验和资源等总体要求，并划分核心网、无线接入网、传送网和计算资源需要承担的目标，为跨域分解提供统一语义。

OMC 承接端到端服务意图，结合网络运营策略、无线资源状况和跨域协同关系，完成面向无线接入网的目标分解与编排，形成 RAN 保障子意图。RAN 保障子意图明确无线接

入网需要承担的保障目标、服务对象、作用范围、时间窗口、资源边界、优先级和成功条件，并承载网络能力信息与执行结果的双向传递。

Auto-RAN 承接 RAN 保障子意图，由意图任务解析与智能编排中枢结合网络状态、专业知识和能力边界进行语义解析，形成无线专业意图。无线专业意图采用覆盖、容量、时延、可靠性、移动性、能效和故障恢复等无线专业语义表达目标，并明确目标 KPI、控制对象、观测要求、动作范围、资源约束、生效条件和评价规则；在此基础上，中枢生成自智任务与执行策略，组织三级智能层完成决策和执行。由此形成“用户业务意图—端到端服务意图—RAN 保障子意图—无线专业意图—自智任务与执行策略”的下行传递。面向节能、故障治理等网络内部自治场景，OMC 可直接生成 RAN 保障子意图，后续仍统一进入无线专业意图和自智任务处理链路。

(三) 意图任务解析与智能编排中枢

无线网络中的区域规划优化、多站协同和空口控制具有不同的作用范围、处理周期和资源需求。意图任务解析与智能编排中枢以无线专业意图为输入、自智任务为基本执行单元，面向区域网络智能层、多站协同智能层和空口内生智能层组织并行闭环，使上层目标能够转化为职责清晰、边界明确、可观测和可评价的网络任务。

在意图接入与语义解析方面，中枢对 RAN 保障子意图进行身份确认、格式规范、语义消歧和完整性检查，识别保障目标、对象范围、时间要求、KPI、优先级及约束条件，并结合无线知识、网络能力和实时状态，将跨域服务语义转换为无线专业意图。对于目标冲突、约束缺失或能力不匹配的意图，中枢通过协商、补充或拒绝机制给出处理结果。

在任务生成与智能编排方面，中枢将无线专业意图分解为一个或多个相互关联的自智任务，统一配置目标 KPI、控制对象、观测需求、动作范围、资源需求、生效条件、风险等级、回退方式和评价规则；根据任务作用范围、处理时限、数据位置和能力需求，选择模型、智能体、算法、工具及数字孪生能力，确定执行位置、先后关系和跨层协同方式，并对并行任务的目标、资源和动作冲突进行协调。

在运行管控与闭环优化方面，中枢维护意图、任务、能力和资源的关联关系及生命周期状态，持续汇聚执行进度、网络观测、模型置信度、资源占用、风险状态和目标达成情况。当业务目标、网络状态、资源条件或模型质量发生变化时，中枢可触发任务调整、能力替换、资源迁移、重新编排或回退，并将执行结果按照无线专业意图、RAN 保障子意图、端到端服务意图和用户业务意图逐级反馈。

(四) 三级智能层：面向不同网络任务分层注智

区域网络智能层面向区域网络和跨站群任务，围绕覆盖、容量、移动性、能效、体验、故障等具体功能，主要承担网络状态认知、长期趋势判断、候选方案生成、覆盖容量规划、跨站资源协调、能效与移动性策略制定、模型治理和高风险方案推演。该层依托边缘云的大规模算力，输出区域网络方案、资源预算、策略边界、模型适用范围和风险控制要求，为多站协同和空口控制提供整体指导。

多站协同智能层面向基站、站群和局部区域，主要承担负荷、体验、干扰、移动性、资源占用和空口统计等多源数据分析，形成局部网络态势与环境预测，并开展站间资源协调、参数协同和策略校正。预测结果包含作用对象、空间范围、时间窗口、有效期和置信度，为跨站协同和空口行动提供依据。

空口内生智能层面向用户、链路和协议处理过程，主要承担 L3 决策、L2 调度、L1 链路增强、波束管理、TRP 选择、链路自适应、功率控制和无线资源分配等任务。该层采用经过训练、验证和授权的轻量模型或确定性算法，在规定的动作范围内完成快速决策与执行。当出现模型超时、输入异常、分布漂移或置信度不足时，系统切换至经过验证的传统算法或保守策略，保障网络稳定运行。

(五) 跨层协同与信息交互

为避免多个智能任务同时调整同一网络对象，Auto-RAN 对无线专业意图及其任务目标、控制对象、KPI 要求、动作范围、资源上限、生效条件、控制权限、优先级、风险等级、回退条件和评价规则进行统一配置。意图任务解析与智能编排中枢按照业务优先级、网络安全要求和策略边界，对不同任务的执行策略进行协调和冲突处理。同一网络对象在同一时刻由一个有效策略实施控制。局部偏差可在既定范围内自主修正，涉及控制目标、资源预算或动作权限变化时，系统重新进行任务编排。

三级智能层围绕同一无线专业意图并行运行，通过意图任务流、能力模型流、状态预测流和执行反馈流实现协同。意图任务流传递任务目标、KPI、约束条件和评价要求；能力模型流完成模型、Agent、Skill、算法和工具的选择、部署与回退；状态预测流交换网络观测、环境预测、数字孪生结果、适用范围和置信度；执行反馈流汇聚控制动作、KPI 变化、风险、异常和回退结果。各智能层根据事件和运行反馈触发相关任务调整，实现闭环之间的动态协同。

(六) 可信执行与目标达成评估

任务执行前，系统依据控制对象、动作类型、影响范围和业务重要性开展风险分级，并依次完成身份与权限校验、规则与策略校验、资源条件检查以及动作边界检查。低风险任务

在授权范围内自动执行；中高风险任务通过数字孪生或沙箱开展效果推演、冲突检测和安全验证，必要时由人工审批确认后分级下发，确保执行主体、执行对象、动作范围和生效条件均处于可信边界内。

任务执行过程中，系统持续监测输入数据质量、网络状态、模型置信度、执行时延、资源占用、KPI 变化和动作边界，识别策略冲突、指标恶化、模型漂移、执行超时和越权操作等异常。当风险超过设定阈值时，按照影响程度采取告警、限流、异常熔断、能力降级、策略回退或任务终止等措施；对于无法自动处置、影响范围扩大或涉及关键业务的异常，及时转入人工接管，防止局部智能决策引发网络级风险。

任务执行后，意图任务解析与智能编排中枢按照用户业务意图标识、端到端服务意图标识、RAN 保障子意图标识、无线专业意图标识和任务标识关联运行记录，对 KPI 改善、业务体验、资源消耗、动作影响和风险处置结果进行综合评估。目标达到后结束任务或转入持续监测；目标未达到时开展偏差归因，调整局部任务、重新编排能力或生成新的执行策略，必要时恢复至已验证的稳定状态。通过执行前可验证、执行中可控制、执行后可评价，形成完整的可信执行闭环。

3.3 Auto-RAN 的核心价值

Auto-RAN 为运营商与行业客户带来六大核心价值：Experience 体验可承诺、Efficiency 能效可持续、Performance 性能自优化、Resilience 故障自恢复、Openness 能力可开放、Value 价值新增长。六大价值覆盖网络体验、运营效率、技术性能、网络韧性、生态开放与商业增长六个维度，构成 Auto-RAN 的完整价值体系。

(一) 体验可承诺

面向工业、应急、无人机等行业场景，Auto-RAN 接收自然语言业务诉求，通过意图解析将业务需求转化为无线侧 KPI 预测，网络自动完成资源调度与实时保障，为业务提供可承诺的差异化确定性 SLA，保障速率、时延、可靠性等无线侧指标。



图 3-1 体验可承诺技术实现路径

关键价值指标：端到端业务开通时间从传统人工配置的小时级缩短至分钟级；无线侧 SLA 达成率显著提升；差异化业务保障能力覆盖从尽力而为到确定性保障的多等级需求。

(二) 能效可持续

Auto-RAN 实现通信-算力协同智能节能，根据业务负荷动态调整基站、边缘算力硬件状态，在保障用户体验的前提下降低全网运维能耗，支撑运营商绿色低碳目标。

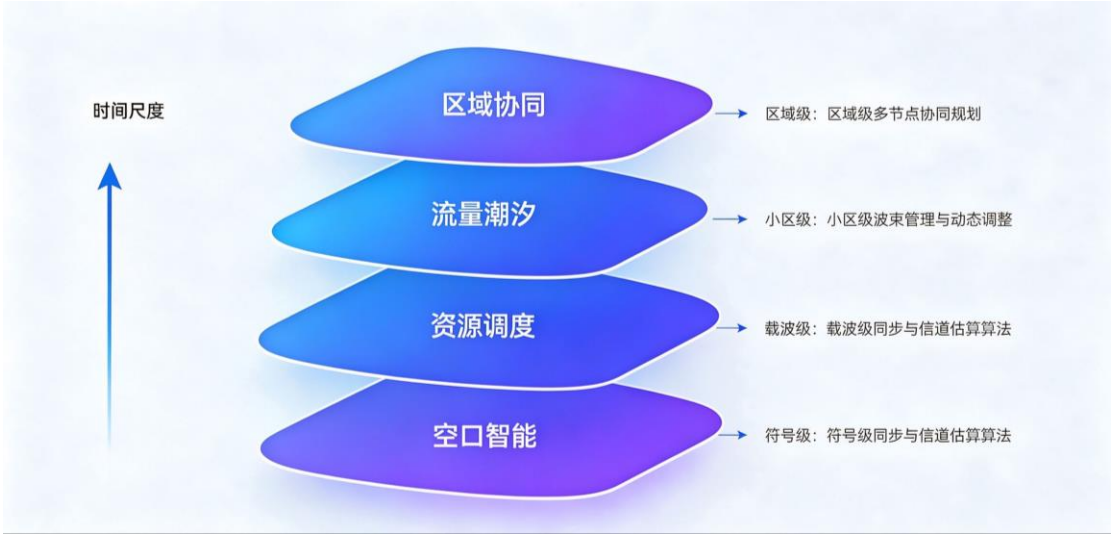


图 3-2 能效可持续技术实现路径

关键价值指标：全网能耗降低；单比特能效持续下降；在业务体验不劣化的前提下实现节能与体验的平衡。

(三) 性能自优化

网络持续感知覆盖、干扰、移动性、信道状态，自主完成参数与策略调优，提升星地融合一体化网络中稀缺频谱资源的利用效率。

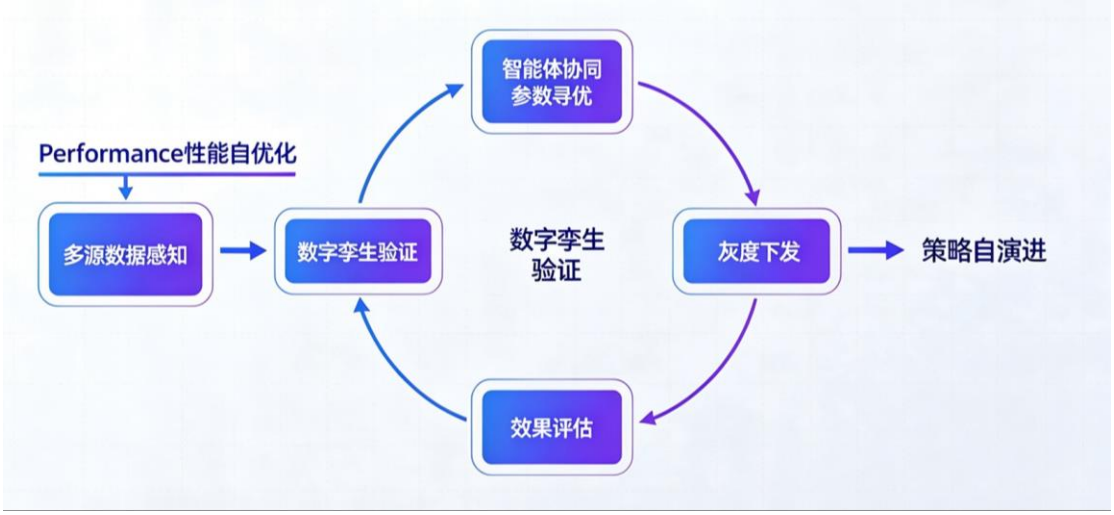


图 3-3 性能自优化技术实现路径

关键价值指标：频谱利用效率提升；用户体验速率与网络容量提升；人工参数优化工作量大幅减少；网络性能随运行持续自演进。

(四) 故障自恢复

Auto-RAN 实现故障提前预测、自动根因定位，触发自愈闭环，减少人工运维介入，提升网络韧性，降低故障对业务的影响。

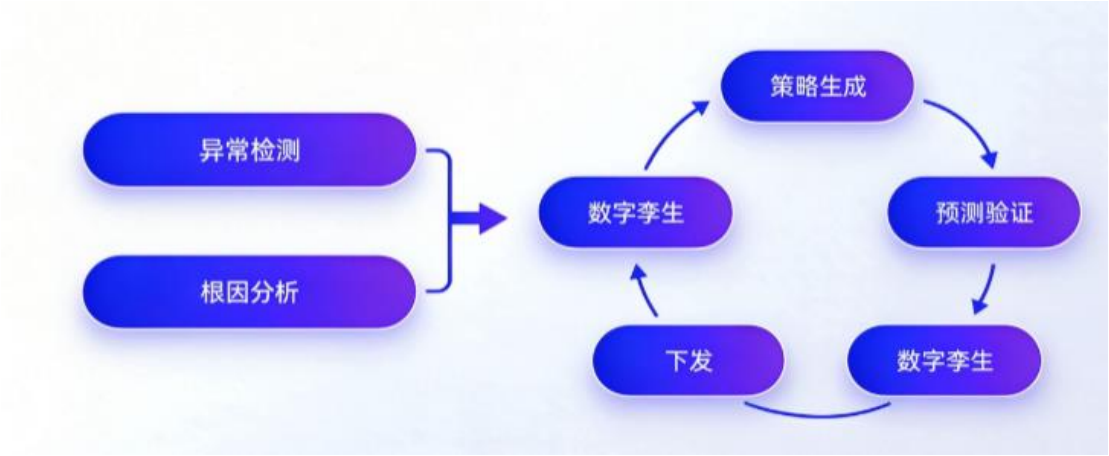


图 3-4 故障自恢复技术实现路径

关键价值指标：故障发现时间从分钟级缩短至秒级；故障恢复时间大幅缩短；自愈覆盖率持续提升；人工运维介入频次显著降低。

(五) 能力可开放

Auto-RAN 将网络能力通过 Skill、API、Agent 对外开放，赋能行业第三方开发者，基于无线网络智能能力开发行业应用，构建开放共赢的产业生态。通过开发者生态运营，持续丰富行业应用，构建开放共赢的产业生态。

关键价值指标：开放网络能力数量持续增长；第三方行业应用接入数增加；网络能力从“封闭运维工具”转变为“可被行业调用的服务组件”。

(六) 价值新增长

Auto-RAN 推动无线侧网络能力产品化与服务化开放，将无线覆盖、空口速率、时延保障、可靠性等级等 RAN 能力封装为可订购、可调度、可度量的服务组件，为运营商端到端行业解决方案提供无线侧能力底座。同时通过网络能力开放 API/Skill/Agent，赋能第三方行业应用调用无线智能能力，拓展 RAN 能力的商业价值。

关键价值指标：RAN 能力产品化种类持续增长；行业客户服务订购数增加；基于无线能力的新收入占比提升；RAN 从“封闭运维成本中心”转变为“可开放的价值服务中心”。

4 Auto-RAN 的关键使能技术

4.1 Agentic-多智能体协作

Auto-RAN 面向网络自智从“单点智能”向“群体智能”演进，构建以北极星领航网络大模型提供智能认知与推理能力、以意图任务解析与智能编排中枢负责协同组织、以多层级网络智能体为自治执行单元的多智能体协作体系。通过具备感知、推理、决策、执行和反馈能力的智能体动态协作，实现从单智能体自治、多智能体协同到网络级自智的逐级演进。

4.1.1 智能体矩阵与分层

Auto-RAN 采用“一级领域智能体+二级功能智能体”的两级组织方式，并与北极星领航通信大模型 共同形成“全局认知—领域协调—功能自治”的分层体系。北极星领航通信大模型 负责跨领域意图理解与高层协调；一级领域智能体负责任务承接、能力组织和冲突协调；二级功能智能体围绕覆盖、容量、移动性、能效、体验、故障等具体功能执行自治闭环。

按照自治任务的触发方式和运行节奏，一级智能体分为优化域、运维域和保障域，对应三种典型网络智能运行范式。三类领域并非传统业务部门划分，而是对不同自治任务运行机制的抽象。

（一）优化域：事件触发的运行态。面向覆盖、容量、干扰、移动性、能效和体验等性能变化，当指标偏离目标时触发相关智能体协同优化，目标达成或策略收敛后退出，实现按需触发、动态组合的事件驱动优化。

（二）运维域：任务驱动的离线态。面向故障工单、巡检、规划、诊断和知识分析等任务，以工单或任务为基本处理单元，围绕任务上下文组织异常分析、根因定位、方案生成和处置，直至任务完成或转交其他领域。“离线态”强调任务驱动的处理方式，并不意味着相关能力只能部署于离线计算环境。

（三）保障域：持续监控的实时保障态。面向 VIP 用户、大型活动、节假日和应急通信等场景，在保障窗口内持续监测 SLA、用户体验和关键业务目标，并按需联动优化与资源调度能力实施干预。

三类范式可随任务演化跨域接力。例如，保障域发现体验劣化后可触发优化域，若进一步识别出设备或网络异常，则转入运维域处置，恢复后再回到保障域持续观察。

Auto-RAN 在一级领域智能体之下构建覆盖网络规、建、维、优全生命周期的二级功能智能体矩阵。各智能体具有明确的专业领域、能力边界、适用场景和输入输出规范，可独立形成“感知—推理—执行—反馈”闭环，并参与多智能体协同。

表 4-1 Agenti-多智能体协作

智能体名称	所属领域	核心职责	典型应用场景
故障监控智能体	运维域	多源数据异常检测、故障根因定位、自愈策略生成	故障预测、根因分析、自动恢复
性能优化智能体	优化域	覆盖/容量/干扰/移动性参数寻优、策略自演进	区域寻优、负载均衡、移动性优化
网络运行智能体	运维域	网络状态监控、运行	运行态势感知、规建

		分析、知识问答辅助	维优问答
端网协同智能体	保障域	终端与网络协同优化、端到端体验保障	VIP 保障、大型活动保障、应急通信
能效优化智能体	优化域	通信-算力协同节能、分层能效策略生成	全网节能、小区级节能、符号关断
体验保障智能体	保障域	用户体验监测、SLA 保障、劣化主动干预	差异化 SLA 保障、体验劣化修复
规建维优智能问答	运维域	网络规划/建设/维护/优化知识问答、专家经验辅助	运维知识检索、方案辅助生成

4.1.2 多智能体协作机制

Auto-RAN 多智能体协作不是固定流程串联，而是以业务意图为起点，根据网络状态、约束条件和执行结果动态选择智能体并组织协作，核心包括意图解析、约束注入、协同调度和冲突仲裁。



图 4-1 多智能体协作架构

意图解析将高层业务诉求转化为网络任务，并识别相关领域、任务边界及跨领域任务关系；约束注入结合配置规则、安全策略、资源边界、业务 SLA 和网络状态，提前限定可执行决策空间，使智能体在可信边界内自主决策；协同调度根据任务目标、能力需求和依赖关系，动态形成串行、并行、条件触发或事件驱动等协作方式；冲突仲裁则面向性能、能效、体验和故障恢复等目标之间可能产生的资源竞争和策略冲突，通过领域内协调与跨领域综合仲裁保障网络级一致性。

在此基础上，Auto-RAN 将模型、智能体、传统算法、网络工具和外部系统组织为稳定可控的执行体系，形成从智能推理到真实网络执行的工程化支撑，并建立“目标—行动—效果—再决策”的持续闭环。网络状态和执行效果重新影响目标理解、能力选择和后续编排，使系统能够根据环境和业务变化自主进入下一轮协同。

专业智能体以“感知—推理—执行—反馈”为基本自治范式，持续感知网络和业务状态，结合规则、模型和历史经验生成策略，通过网络能力或数字孪生环境执行，并依据 KPI

和业务效果调整后续决策。短期状态记忆与长期经验记忆支持连续任务上下文保持和经验积累，为多智能体协同提供基础。

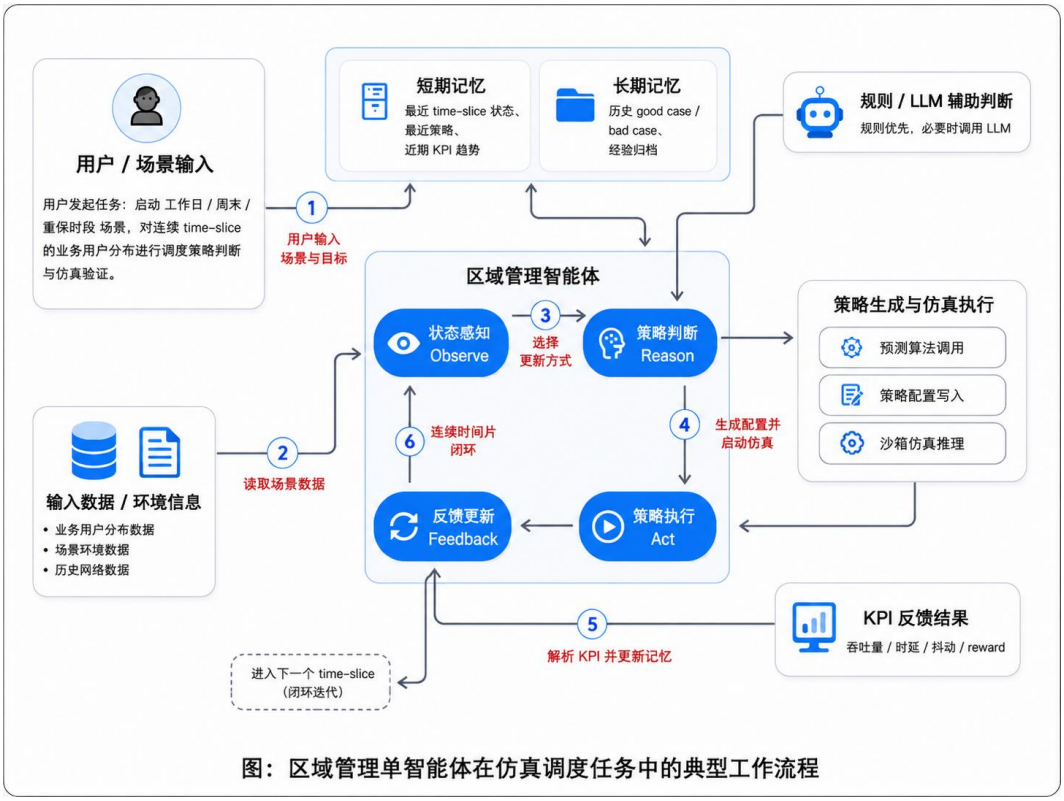


图 4-2 单智能体自治闭环

4.2 Unified-北极星领航通信大模型

北极星领航通信大模型是面向移动通信场景的通信领域专用大模型，作为智能体的全局网络大脑，承担意图理解、知识增强、任务规划与跨域数据推理等核心能力。北极星领航大模型是智能体 Unified 设计原则的核心承载——通过统一的通信领域数据集、统一的意图理解与编排范式、统一的知识增强架构和统一的大-小模型分层协同体系，实现数据、模型、知识和算力的统一协同。

4.2.1 通信领域数据集构建

高质量的通信领域数据集是模型微调和能力评估的基础。北极星领航通信大模型构建了覆盖多类型、多任务的通信领域数据集体系，为 6G 领域模型微调、任务适配和能力评估提供数据基础支撑。

领域数据类型已覆盖标准手册类（3GPP/ITU 标准协议、运维手册、专家经验文档）、运行监测类（KPI 指标、MR 测量数据、流量数据、告警日志）、故障案例类（故障工单、根因分析报告、处理记录）、配置拓扑类（网元配置参数、网络拓扑、参数依赖关系）等多类通信领域数据集，具备较好的基础支撑能力。任务导向聚焦于知识问答、区域寻优、故障诊断、性能优化等具体通信任务构建，确保训练数据与实际业务场景紧密对齐。

4.2.2 意图理解与编排增强

北极星领航通信大模型以意图理解为核心入口，通过知识增强、任务理解与策略推理、后训练对齐等能力，为意图任务解析与智能编排中枢提供语义理解与智能推理支撑，辅助完成业务意图向网络任务及候选策略的转化，并持续提升大模型的网络认知与推理能力。

(一) 意图理解增强

意图理解是北极星领航通信大模型承接业务诉求的首要能力。北极星领航通信大模型通过结构化输出采用通信意图 DSL（Domain Specific Language）标准化输出范式；通过指令微调、场景化适配与后训练对齐等大模型训练技术，持续提升对通信领域复杂业务问题的理解能力，实现从自然语言业务意图到标准化网络任务描述的精准转换。

基于通信领域意图-任务配对数据进行指令微调，使模型学习通信领域的意图表达习惯、业务术语与任务映射关系；通过后训练阶段的人类反馈与专家对齐，由无线优化、运维排障等领域专家对模型输出进行标注与排序，重点强化“参数范围合规、策略逻辑自洽、故障因果可追溯”等对齐目标。

(二) 任务编排能力提升

任务编排是北极星领航通信大模型从意图理解到策略生成的核心推理环节。北极星领航通信大模型通过大模型训练技术提升任务编排的自闭环推理能力，使模型能够基于业务意图自主完成任务分解、依赖识别、冲突检测与执行规划，输出高质量的任务编排方案。

基于通信领域编排案例进行指令微调，使模型学习典型业务场景的编排模式与最佳实践；通过领域专家对编排方案的标注与反馈进行后训练对齐，重点强化“分解逻辑合理、依赖识别准确、冲突检测全面、方案可执行性强”等对齐目标

4.2.3 大-小模型分层协同

北极星领航通信大模型聚焦全局高层推理规划实现全域网络智能；中模型实现秒级、分钟级区域网络协同；毫秒级实时任务下沉部署在基站侧专用任务专属空口小模型执行从而实现空口智能。整个系统形成“中心规划-边缘决策-空口执行”的分层协同关系，各层级模型的定位与分工如下图所示：

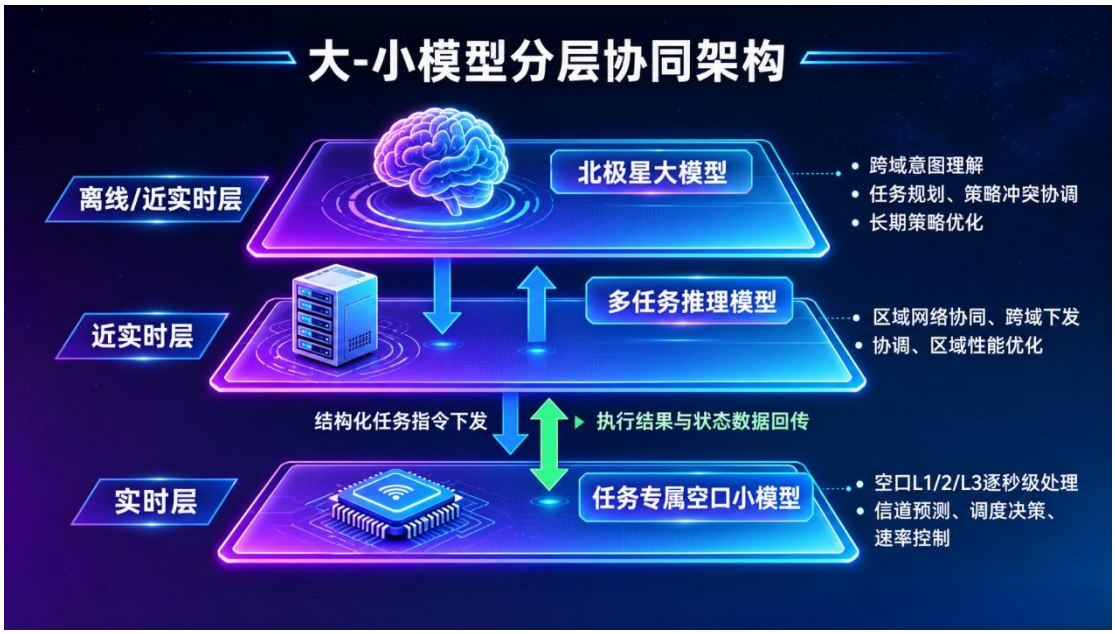


图 4-3 大-小模型分层协同架构

4.3 Trusted-数字孪生

数字孪生是 Auto-RAN Trusted 设计原则关键技术，实现网络策略可验证、智能体决策可信、物理网络可闭环演进。数字孪生确保了智能体决策的全生命周期可信，为其提供模型与策略的可信校验，为 Auto-RAN 从人工驱动运维向 AI 驱动智治范式跃迁提供使能路径。

数字孪生依托多维数据采集与感知、网络和无线侧高保真动态建模等关键技术，构建起覆盖 Auto-RAN 全生命周期的数字化底座。

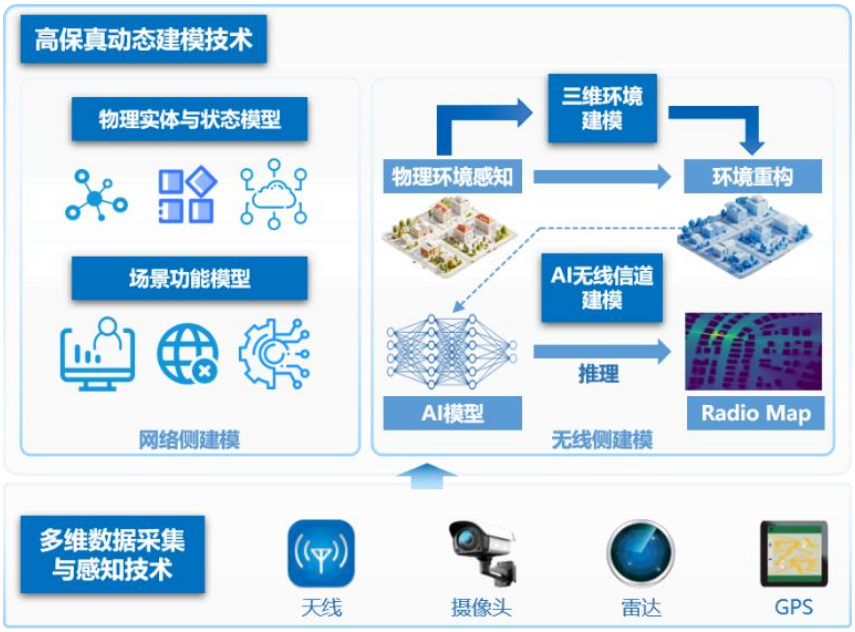


图 4-4 数字孪生关键技术体系

(一) 多维数据采集与感知技术

作为数字孪生构建的首要环节，其核心任务在于实时、精准、合规地获取物理网络多维度数据，以数据压缩或语义提取技术降低带宽压力，以 AI 增强保障数据质量，并以标准化接口规范破解厂商数据孤岛。

(二) 网络侧高保真动态建模

旨在构建与物理网络运行状态实时一致、可自更新的镜像。基础模型表征网络拓扑、网元特征与资源分布等核心要素，功能模型提供运行态势评估与决策优化能力，进一步与 AI 技术融合，模拟用户分布、业务流量及资源状态并预测趋势。

(三) 无线侧高保真动态建模

聚焦无线传播环境与电磁信道的高保真复现。三维环境动态建模以多模态传感设备与通信终端协同，在线采集并更新环境几何、动态特征与电磁参数。AI 信道建模克服经验模型、射线追踪与实测方法弊端，依托扩散模型等先进模型，融合射线追踪及历史实测数据等多源先验知识，快速生成无线信道预测，并以少量实测数据在线真值校准，持续提升精度可靠性，大幅降低建模时间与成本。

(四) 智能体决策与虚实映射闭环

在智能体决策、孪生验证、物理网络反馈协同运行范式下，数字孪生不仅是物理网络的镜像，更成为智能体的数字沙盒和决策验证引擎，提前预测问题并主动优化，实现真正自智网络。

数字孪生依托“孪生空间内闭环验证-物理网络外闭环反馈”双环协同，搭建起“虚拟仿真-决策验证-现网部署-数据回流-模型迭代”运行链路。智能体基于全域状态感知完成意图任务解析，智能编排中枢将其分解自智任务；自智任务先进入内闭环，依托孪生空间复刻网络拓扑、无线环境与用户行为，评估方案性能与风险，并通过注入扰动检验鲁棒性、自动追溯优化；自智任务排除高风险操作并达成关键性能指标后，随后经策略编排外闭环下发至物理网络执行，运行结果与实时数据回流孪生空间，驱动模型与决策持续迭代。

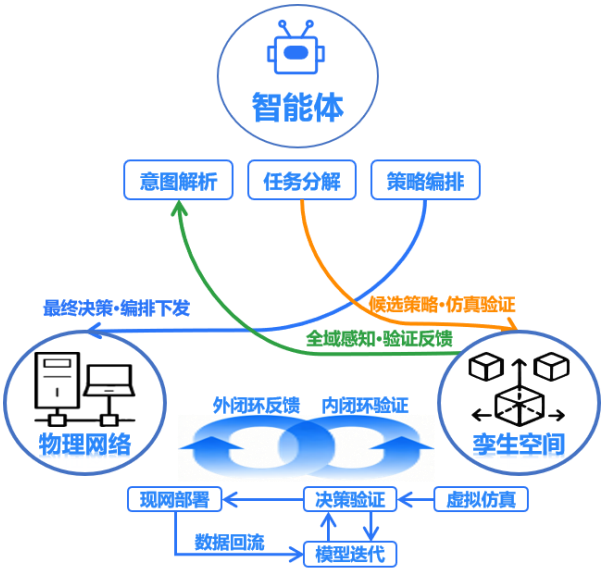


图 4-5 智能体决策与虚实映射闭环

这一融合范式核心价值是破解网络自智的两大结构性矛盾：其一，AI 高频探索与物理网络高可靠性不容试错之间的矛盾，以实时同步的虚拟镜像提供零现网影响的试验环境；其二，网络高维动态性与网络控制静态规则化带来服务质量提升瓶颈，数字孪生与智能体协同为 Auto-RAN 的动态策略提供的迭代优化环境。

4.4 Open-算力分级部署与开放

Auto-RAN 以开放（Open）为基本设计理念，面向算力形态多样、智能能力快速演进和应用场景持续扩展的发展需求，构建分级部署、统一纳管、按需调度、能力开放的算智基础设施。开放不仅体现为接口开放，更体现为算力资源、模型能力、智能体能力、数据服务和网络工具的标准化封装与灵活组合，使不同来源、不同形态的智能能力能够在统一规则下接入、部署和协同运行，为 Auto-RAN 持续演进提供可扩展的承载基础。

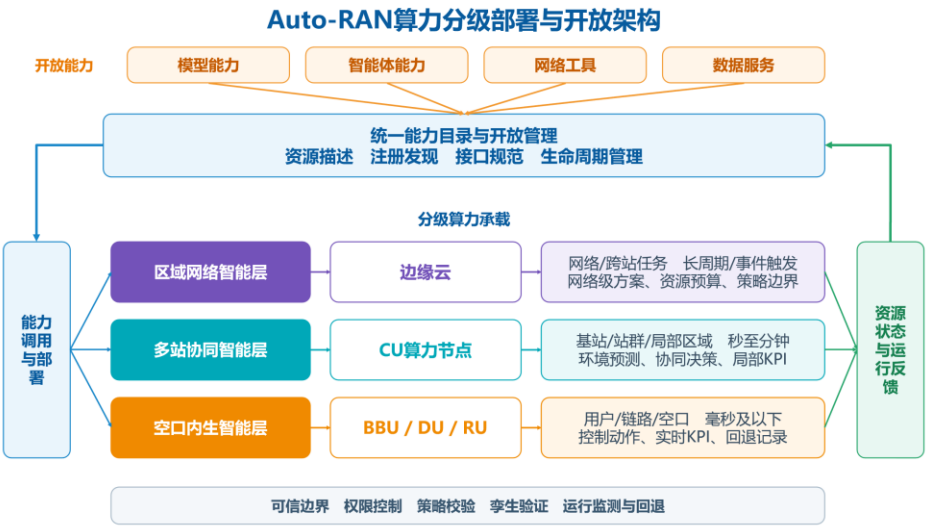


图 4-6 Auto-RAN 开放算智底座与算力分级部署架构

表 4-2 Auto-RAN 三级智能层与物理算力位置映射

智能层	可选位置	部署判据	说明
区域网络智能层	跨站群/跨站；以长周期和事件触发任务为主	边缘云	模型规模、跨站群数据、计算复杂度、风险验证
多站协同智能层	基站/站群/局部区域；以秒至分钟任务为主	CU 算力节点	预测窗口、局部数据、多任务协同、资源负载
空口内生智能层	用户/链路/空口；以毫秒及以下任务为主	BBU/DU/RU	最坏时延、动作权限、协议约束、确定性要求

表 4-2 给出了三级智能层的任务职责、主要算力位置、可选位置和部署判决。分级部署以该职责划分为基础，根据任务处理范围和时效要求配置算力，使智能处理靠近数据产生位置和网络执行位置，同时保障关键任务的确定性。

(一) 分级部署，匹配差异化网络智能任务

区域网络智能层面向网络及跨站的长周期或事件触发任务，主要依托边缘云开展全局认知、规划推演、资源与能效策略生成和模型治理；多站协同智能层面向基站、站群及局部区域的秒至分钟级任务，主要部署于 CU 算力节点，开展局部态势预测、干扰分析、站间协调和参数动态校正；空口内生智能层面向用户、链路及空口的毫秒及以下任务，依托 BBU、DU 和 RU 完成 L3 决策、L2 调度、L1 增强以及波束和链路自适应。部署过程中综合考虑最坏时延、数据本地性、资源负载、能耗、安全等级和动作权限，并通过资源预留、本地执行及回退机制保障通信关键任务稳定运行。

(二) 能力开放，构建可扩展的算力基础设施

Auto-RAN 对 CPU、GPU、NPU 和通信专用处理器等异构资源进行统一抽象，形成涵盖计算能力、存储容量、网络连接、资源负载、能耗、可信等级和模型适配性的资源描述。模型、智能体、网络工具和数据服务按照统一规范注册到能力目录，通过标准接口完成发现、调用和组合，使不同厂商、不同专业及不同部署位置的能力能够按需接入。

开放能力实行全生命周期管理，统一记录版本、依赖关系、适用范围、资源占用和运行状态，支持上线、升级、停用和退出。外部能力接入通过身份认证、权限控制、接口校验和安全隔离限定数据、网络对象及资源范围；涉及网络配置或关键参数调整的能力，经策略校验和数字孪生验证后分级执行，并持续反馈执行状态和效果指标，形成安全可靠、可持续扩展的开放能力体系。

5 Auto-RAN 路线图：从 5G-A 试点验证迈向 6G 规模商用

Auto-RAN 采用分阶段演进策略，基于现有 5G-A 产品底座逐步验证，从部件级试点到端到端原型联调，最终在 6G 时代实现 AI-Native 原生网络规模商用。整体演进路线分为

三个阶段：5G-A 试点验证阶段（近期）、原型联调与行业 POC 阶段（中期）、6G 商用部署阶段（远期）。

阶段一：5G-A 试点验证依托现有 5G-A 网络和产品体系，优先选择技术基础较成熟、业务价值较明确的场景开展试点验证。在硬件侧，推进智能融合板、边缘智算设备等产品部署，初步构建网络侧 AI 推理与任务承载能力；在软件侧，围绕智能运维、节能优化、性能保障和故障诊断等场景，验证区域网络智能层与多站协同智能层闭环。

北极星领航通信大模型重点面向规建维优知识问答、故障分析和任务辅助等场景开展 POC，通过与专业网络智能体协同，验证通信知识理解、任务规划和工具调用能力。

本阶段以单点能力和局部闭环验证为主，重点评估技术可行性、运行安全性和实际业务价值，逐步积累现网数据、模型和工程实践经验，为后续 Auto-RAN 端到端原型建设奠定基础。

阶段二：端到端行业 POC

阶段二以“以智筑网、智算融合、价值拓进”为方针，是从部件级验证走向端到端系统验证的关键阶段。核心目标是完成 Auto-RAN 端到端原型系统搭建，打通三级智能闭环与意图驱动全链路，实现 AI 能力从“外挂式增强”向“内生式嵌入”的跨越。

该阶段打通从意图输入到空口资源兑现的完整链路，包含意图输入接口、北极星领航通信大模型推理服务、多智能体运行框架、三级智能闭环引擎、数字孪生验证平台等关键组件，持续迭代北极星领航通信大模型、网络智能体与数字孪生验证能力，开展智算融合深度优化，验证创新商业模式可行性。

阶段三：6G 商用部署

阶段三以“智能内生、多维融合、万物智联”为方针，是 Auto-RAN 作为 6G AI-Native 原生架构实现规模商用的阶段。核心目标是实现意图驱动网络的规模化运行，推动多智能体

协同的去中心化治理，全面迈向“通感算智融合”的网络生态。多网络智能体作为执行载体实现分布式自主决策，三级智能闭环实现从网络级到符号级的全维度自智，数字孪生作为验证底座实现策略安全发布。

网络运营从“人工配置驱动”全面转向“用户意图驱动”，运营商与行业客户通过自然语言意图直接订购无线网络服务，Auto-RAN 自动完成意图解析、资源编排、SLA 保障与效果评估，支持海量并发业务意图的实时解析、调度与保障，实现“所想即所得”的网络运营范式。

6 参考文献

[1] 国务院新闻办公室:《2026 年上半年工业和信息化发展情况新闻发布会文字实录》，2026 年 7 月 20 日。

[2] 工业和信息化部：《2026 年前 7 个月通信业经济运行情况》，2026 年 8 月。

[3] 工业和信息化部：《2025 年通信业统计公报解读：通信业运行平稳，新动能加快培育》，2026 年。

[4] Gartner:《Gartner Predicts 40% of Enterprise Apps Will Feature Task-Specific AI Agents by 2026》，2025 年 8 月 26 日。

[5] 中国电信股份有限公司：《2025 年年度报告》，2026 年 3 月。

[6] 中国联合网络通信（香港）股份有限公司：《2025 年年度报告》，2026 年 3 月。

[7] McKinsey & Company：《AI Infrastructure: A New Growth Avenue for Telco Operators》，2025 年 2 月 28 日

[8] 中国移动有限公司：《2025 年中期报告》，2025 年 8 月。

[9] McKinsey & Company：《Telcos’ AI Inflection Point: What Leaders Do to Capture Value》，2026 年 2 月 26 日。

[10] McKinsey & Company：《AI-Driven Telecom Networks》，2026 年 2 月 27 日。

中国信科

中信科移动

中信科移动通信技术股份有限公司
CICT Mobile Communication Technology Co.,Ltd.

网址：www.cictmobile.com