

算网协同：构建新型信息基础设施和新型服务模式

王晓云¹，孙滔²，崔勇³，Rajkumar BUYYA⁴，郭得科⁵，

黄群⁶，Hassnaa MOUSTAFA⁷，田臣⁸，王尚广⁹

¹中国移动通信集团有限公司，中国北京市，100032

²中国移动通信有限公司研究院，中国北京市，100053

³清华大学计算机科学与技术系，中国北京市，100084

⁴墨尔本大学计算与信息学院 CLOUDS 实验室，澳大利亚维多利亚州墨尔本市，3010

⁵国防科技大学信息系统工程国家重点实验室，中国长沙市，410073

⁶北京大学计算机学院，中国北京市，100871

⁷英特尔公司，美国加利福尼亚州圣克拉拉市，95054

⁸南京大学计算机软件新技术国家重点实验室，中国南京市，210023

⁹北京邮电大学网络与交换技术国家重点实验室，中国北京市，100876

计算，作为信息处理的基石，在数字服务时代起着至关重要的作用。“网络”和“计算”分别用于信息传输和处理，传统上属于不同的利益主体，并独立演进。近期，网络和计算协同、融合的演进趋势得到工业界和学术界广泛关注。相应地，“算力网”和“算力网络”等概念被陆续提出，电信标准化组织国际电信联盟（ITU-T）也已启动算网协同（coordination of networking and computing, CNC）的标准化工作，尤其关注架构和框架设计。网络与计算协同将带来几个直接好处：

1. 能够在广域内有效调度资源，实现节能。中国在 2022 年启动的“东数西算”国家工程就是例证，该工程旨在引导产业界在有着丰富能源的中国西部地区进行数据处理和存储，高效利用能源，特别是可再生能源。

2. 促进算力成为一种公共基础设施服务。这是人工智能先驱约翰·麦卡锡（John McCarthy）在 20 世纪 60 年代提出的愿景。其目的是使算力像家用水电一样容易获得，以弥合算力资源可获取性的差距。孙凝晖院士及其团队将上述能力命名为“computility”，亦即基于网络提供泛在的信息处理能力。

3. 显著提升用户体验。例如，沉浸式通信是以大流量、低延迟、频繁交互和高计算复杂性为特征的通信模式，其服务质量的评估需同时考虑数据的传输和处理。连接和计算的协同可以实现优化的平衡，例如选择合适的计算节点和网络传输路径。

本专题包括 6 篇文章——1 篇前瞻观点、1 篇综述和 4 篇研究论文。该专题可以看作对长久以来的经典问题——将系统设计为一个网络连接的超级计算机——的一次再探讨。通过对终端、边缘云和中心云进行计算资源的协同管理，实现端到端服务的优化至关重要。其核心理念是“以计算换传输”，通过增强的算力减少数据处理时延，从而降低网络传输过程中对超低延迟传输的苛刻要求。通过系统性协同和集成，让网络 and 计算能力产生协同放大作用。为达到上述优势，需解决 3 个方面的重要问题：

1. 对网络与算力的需求及资源进行标识，识别应用所需的算力和网络传输能力。了解如何识别和建模这些需求至关重要，特别对于异构和分布式计算资源。挑战在于确保有效分配原始数据或处理逻辑，以使用可用的计算资源。

2. 及时与联合的感知算力和网络资源，收集与网络或算力相关的负载信息。由于采用不同监控方法，算力指标通常获取自算力基础设施组件，如云服务、虚拟机或容器。相比之下，网络连接指标通常依赖于端到端数据路径传输的测量。

3. 计算和网络资源的协同调度需要系统性优化。促进终端设备、边缘计算和云服务之间的协作，涉及到为网络建立一个统一的“大脑”，使得该网络化的“分布式计算系统”中的资源提供者可协作地应对多维资源调度的挑战。

除以上核心方面外，这些系统的实现还涉及任务，比如应用任务分解。此外，评估网络

和计算资源能否在信息处理全生命周期——包括生成、传输、处理和消费——中协作至关重要。在数字化转型背景下，网络与计算协同预示着一个有着广泛前景的崭新研究领域，还有许多挑战有待发现和解决。

系统的架构设计是实现这些新功能的关键。王晓云等提出算力感知网络（CAN）概念，这是一个新的框架，集成了促进广域计算和网络协同的感知平面。这个框架确定了 3 项关键技术：算力路由（CATS）、弹性广播和广域高吞吐传输。尽管将算力引入路由策略还需进一步调研，国际互联网工程任务组（IETF）已启动关于算力路由的讨论，以探索潜在场景和架构设计。弹性广播是为广域网络中“一对多”通信定制设计的，这对于跨数据中心的人工智能模型训练和推理至关重要。尽管高吞吐传输并非新概念，但长距离造成的物理延迟、数据丢包和服务限制，为其在广域网中的应用带来不小挑战。

由于高功率计算对环境和成本的影响，能源使用效率至关重要。近年来业界关注的联邦学习就需要采用节能的方法，特别是在边缘计算环境中。颜康等对边缘联邦学习中的节能策略作了全面调研，包括系统和能源消耗模型。探讨了 3 类策略，包括基于学习的策略、资源分配策略和用户选择策略，并指出几种潜在的节能联邦学习研究方向。

有效提供计算服务或解决复杂问题，需要采用协作的方法进行任务分解。这包括在用户设备、边缘网络和云数据中心之间的任务卸载，促进在云、边缘、终端设备等基础设施之间的协作。白小军等将不同任务分为延迟敏感类和延迟容忍类，利用连续时间马尔可夫链对系统建模。通过对边缘网络缓冲区访问阈值的优化，实现了在平均延迟和阻塞率之间的平衡。

当有多个计算和网络资源服务提供者时，选择适当的提供者对于完成任务至关重要。付月霞等利用基于 β 分布函数的信誉模型来评估资源提供者的信誉，并引入基于性能的信誉更新模型。该方法被建模为一个受约束的多目标优化问题，通过改进的快速精英非主导的排序遗传算法确定可行解。大量仿真实验证实了模型在提高用户满意度和资源使用率方面的有效性和效力。

网络和计算协同的本质在于对计算资源与网络资源分配的联合考虑。韩雪莹等提出一种将深度强化学习与图神经网络相结合的智能资源分配方法。该方法通过优化跨网络拓扑中的网络和计算资源，在结构性变化情况下也能克服算力网络中路由的挑战。

在各种场景和领域中，算力与网络之间的协同作用至关重要。蔡逸卓等致力于提高 6G 网络中联邦学习场景的通信效率，调研了传统和对等（peer-to-peer）联邦学习架构。文章证明基于实时资源状态优化的算力调度可以显著提高性能。

本专题得到学术和产业界大力支持，涵盖了网络与计算协同的多个方面，如互联网协议（IP）、蜂窝网络等。需要指出的是，相关研究仍处于起步阶段，许多问题和技术值得进一步深入研究。

最后，衷心感谢所有作者、评审人及期刊编辑的支持和贡献。

<https://doi.org/10.1631/FITEE.2430000>