



中国通信学会

CHINA INSTITUTE
OF COMMUNICATIONS

算力网络前沿报告

(2020年)

中国通信学会

2020年12月

版权声明

本前沿报告/白皮书版权属于中国通信学会，并受法律保护。转载、摘编或利用其它方式使用本报告文字或者观点的，应注明“来源：中国通信学会”。违反上述声明者，本学会将追究其相关法律责任。

CIC中国通信学会

专家组和撰写组名单

专家组：

组长：

赵慧玲 工信部科技委专职常委

副组长：

唐雄燕 中国联通网络技术研究院首席科学家

撰写组：

单位	姓名
中国联合网络通信有限公司	曹畅
中国联合网络通信有限公司	李建飞
中国联合网络通信有限公司	张帅
中国联合网络通信有限公司	刘莹
中国联合网络通信有限公司	耿庆鹏
中国联合网络通信有限公司	何涛

前 言

新一代信息网络正在从以信息传递为核心的网络基础设施，向融合计算、存储、传送资源的智能化云网基础设施发生转变。算力网络正是为应对这种转变而提出的新型网络架构。算力网络基于无处不在的网络连接，将动态分布的计算与存储资源互联，通过网络、存储、算力等多维度资源的统一协同调度，使海量的应用能够按需、实时调用泛在分布的计算资源，实现连接和算力在网络的全局优化，提供一致的用户体验。支撑算力网络实现的关键技术，可以归类为云、网、芯三个方面。

从云化技术来看，需要在目前已实现对虚拟资源编排的基础上，向容器编排和算力编排演进，并针对网络中异构算力资源并存的情况，探索计算能力的统一纳管与提供服务的方式。

从网络技术来看，需要在目前已实现云网拉通和统一配置的基础上，向基于 SRv6 技术的网络切片能力增强演进，并增强资源感知与应用感知能力，提升算力触达用户的广度和精度。

从芯片技术来看，需要在传统固定流程 ASIC 芯片架构的基础上，向具备可编程能力的新一代交换芯片架构演进，并增强网络转发面编程能力以便与计算芯片相配合，赋予设备更灵活的功能。

本报告主要研究算力网络的发展趋势、技术特征、以及所面临的挑战，希望抛砖引玉，共同推进我国计算产业和网络产业的协同发展与融合创新。

本报告的主要撰写人为唐雄燕、曹畅、李建飞、张帅、刘莹、耿

庆鹏、何涛、王友祥等，在此对这些专家的贡献表示感谢！

中国通信学会信息通信网络技术委员会

主任委员：



2020年12月

CIC中国通信学会

目 录

一、	研究概述	1
二、	计算与网络产业全球发展态势	1
三、	算力网络在我国的发展现状	2
四、	算力网络关键技术	4
	(一) 算力网络技术体系	4
	(二) 算力网络承载技术	5
	(三) 算力网络编排技术	7
	(四) 算力网络转发技术	9
五、	算力网络工程难题	10
	(一) 算力资源的感知与度量	10
	(二) 集中式控制与分布式控制的协同	11
六、	政策建议	12
	(一) 技术政策建议	12
	1. 大力推进算力全网感知与端管云协同	12
	2. 积极推动算力网络的标准化工作	13
	(二) 产业政策建议	14
	1. 促进计算产业和网络产业融合与能力互补	14
	2. 加快推进算力向边缘下沉	14
	3. 注重算力网络产业的自主可控	14

一、研究概述

随着 5G、大数据、人工智能、区块链等信息通信技术的推广应用，经济社会向数字化转型升级的趋势愈发明显。2020 年以来，国家发布了以“新基建”为导向的一系列政策，旨在通过加快建设数字化基础设施，提升各行业的“联接+计算”能力，引领重大科技创新、重塑产业升级模式，为社会发展注入更强动力。“联接+计算”能力的提升，需要计算和网络两大产业的有机协同，相互配合，在计算服务与网络服务均面向泛在化发展的今天，探索云-边-端多级计算资源和服务能力如何通过网络实现智能调度和高效分配十分必要，也是必须攻克的技术难题。

二、计算与网络产业全球发展态势

数字化、智能化正在加快推动计算产业的创新。数字化浪潮正在改变着世界经济格局，IDC 预测，2023 年全球 GDP 的 62% 将是数字经济产值。世界各主要经济体国家纷纷通过国家战略来抢占数字经济产业链的制高点，如美国的《美国国家网络战略》、德国的《高技术战略 2025》、日本的《日本制造业白皮书》等。此外，当前以云计算、大数据、移动超宽带为核心的智能化创新正在变革生产、管理和营销模式，升级传统生产力，加速数字经济的发展。智能化、数字化催生了海量的场景和应用，促使满足多样化计算需求的融合计算架构出现，以及云-边-端结合的泛在计算模式兴起。这些变化都将结合容器、微服务、云原生和 DevOps 等开发模式直接推动计算产业的快速变革。

行业数字化转型和产业智能化升级促使网络产业向着超宽管道、

泛在连接、场景化需求的方向发展。从业务看，5G 特性将带动 VR、AR、云游戏等大带宽、低时延业务走向商业化，同时奠定 Cloud+X 发展基础。从流量看，未来移动数据流量将会迎来再次剧增。爱立信预测，2025 年全球移动数据流量较 2019 年将增加 5 倍，其中 45% 将会由 5G 网络承载。网络带宽的扩大与业务类型的不断拓展将增加网络中的视频流量，预计 2025 年占比增至 76%，以国内电信运营商中国联通为例，2020 年 5 月视频流量占比已达 72%。

新应用、新技术、新计算架构，百亿级联接、爆炸式数据增长将重塑 ICT 产业新格局，催生新的生态体系。一方面要求新计算架构必须具备海量数据处理分析能力，各种应用场景下人工智能训练和推理能力，另一方面要求网络对于大规模联接场景下的安全和实时数据传输处理等能力。预计 2021 年，将有超过 500 亿的终端设备，产生 847ZB/年的数据量，2023 年 IP 设备将达 29.3 亿，M2M 的连接数将超过 14 亿。这将产生海量的数据，带动海量数据分析处理，并围绕数据创造价值，这就同时需要泛在计算的支持和优质的网络保障。由此判断，新的 ICT 格局将向着泛在联接与泛在计算紧密结合的方向演进。

三、算力网络在我国的发展现状

近年来，中国政府相继出台《国家信息化发展战略纲要》、《“十三五”国家信息化规划》、《中国制造 2025》等重大战略纲领，明确“数字中国”建设发展的路线图和时间表，全面推进国民经济各行各业的数字化和智能化。IDC 预测，到 2023 年，数字经济产值将占到

中国 GDP 的 67%，超过全球平均水平。

伴随着我国数字经济的蓬勃发展，新一代信息技术间的融合效应渐显，“5G+云+AI”将成为推动我国数字经济持续发展的重要引擎。结合未来计算形态云-边-端泛在分布的趋势，计算与网络的融合将会更加紧密，因此，我国率先提出了“算力网络”的概念。算力网络需要网络和计算高度协同，将计算单元和计算能力嵌入网络，实现云、网、边、端、业的高效协同，提高计算资源利用率。在算力网络中，用户无需关心网络中的计算资源的位置和部署状态，而只需关注自身获得的服务即可，并通过网络和计算协同调度保证用户的一致体验。

目前来看，我国信息通信行业对“5G+云+AI”的探索处于世界领先地位，这些都带动了全网的算力密集分布，快速下沉并且逐步实现联网服务。至今为止，算力网络的愿景已在业界得到广泛的认可，算力网络在标准制定、生态建设、试验验证等领域均取得了一定进展，并且作为我国的一项原创成果，开始走向国际舞台。标准制定方面，国内三大运营商中国移动、中国电信与中国联通分别在 ITU-T SG11 与 13 组立项了 Y.CPN, Y.CAN 和 Q.CPN 等系列标准，在 IETF 开展了 Computing First Network Framework 等系列研究，华为联合国内运营商在 ETSI 和宽带论坛（BBF）也启动了包括 NWI、城域算网在内的多个项目。在国内的 CCSA，“算力网络需求与架构”以及“算力感知网络关键技术研究”两项研究也在有序开展。面向未来 6G 时代，算力网络已经成为国内 IMT-2030 6G 网络组的研究课题之一，正在开展算力网络与 6G 通信技术的融合研究。生态建设方面，国内未来数

据通信研究的主要组织——网络 5.0 产业联盟专门成立了“算力网络特设工作组”，MEC 领域的多个开源组织也发起了 KubeEdge、Edge-Gallery 等开源项目，2019 年底，在中国联通、中国移动和边缘计算网络产业联盟（ECNI）均发布了算力网络领域相关白皮书，进一步阐述了算网融合的重要观点。试验验证方面，2019 年中国电信与中国移动均已完成算力网络领域的实验室原型验证，并在 GSMA 巴塞罗那展、ITU-T 和 GNTC 相关展会上发布成果。中国联通也在推进算力网络平台的自主研发，并积极策划现网试点工作。

2018 年以来，国家提出一系列加快 5G 商用步伐，加强人工智能、工业互联网、物联网等新型基础设施建设的政策建议，“新基建”的概念由此产生。2020 年 4 月，国家发改委首次对新基建的具体含义进行了阐述，在信息基础设施部分，提出构建以数据中心、智能计算中心为代表的算力基础设施。从新基建政策的导向来看，给予了算力提供者、网络运营者、服务提供者和服务使用者等不同角色引入多方参与的空间，同时也给以算力网络技术为基础的转-算-存主体分离、联合服务的新商业模式提供了宝贵的尝试空间。

四、算力网络关键技术

（一）算力网络技术体系

从算力网络所倡导的技术理念中可以看出，算网一体是结合 5G、泛在计算与 AI 的发展，在云网拉通和协同基础上的下一个阶段，即云网融合 2.0 阶段。云网融合 2.0 是在继承云网融合 1.0 工作的基础上，强调结合未来业务形态的变化，在云、网、芯三个层面持续推进

研发，实现应用部署匹配计算，网络转发感知计算，芯片能力增强计算，服务算力网络时代云游戏、千人千面直播、自动驾驶、智能安防与工业机器人视觉等新业态，其技术内涵如图 1 所示。

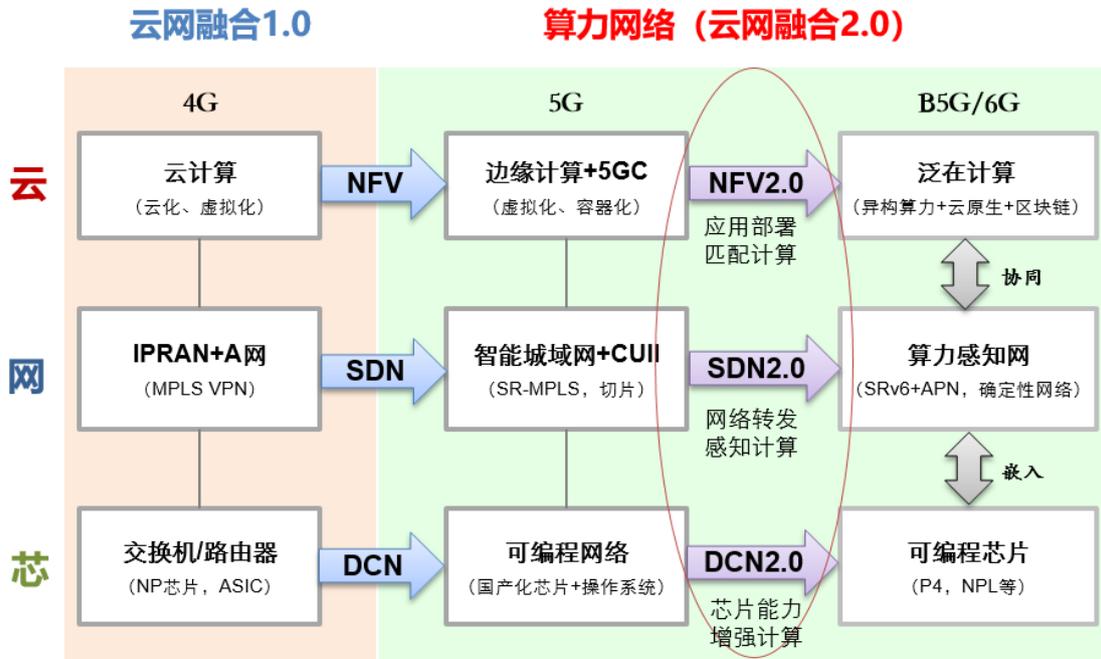


图 1 算力网络与云网融合 2.0 示意

从图中可以看出，算力网络的技术内涵是在现有 SDN/NFV 技术基础上的发展和升华，相关技术可概括为承载、编排和转发三个方面，分别对应 SDN2.0，NFV2.0 和 DCN2.0，本章节的后续部分将会逐一进行阐述。

(二) 算力网络承载技术

随着 5G、云服务和物联网等新兴业务的发展，更多网络设备的接入对于地址扩展的需求和网络可编程的需求都在增加，“IPv6+”是面向 5G 和云时代的智能 IP 网络，可以满足算力网络灵活组网、优化用户体验按需服务等需求。

算力承载网以 SRv6 技术为底座，在网络切片能力的基础上，引

入网络感知技术，解决当前网络难以感知业务需求，算力和服务难以良好匹配的问题。在算力网络中，业务网关进一步下沉，并通过算力网关将南北向流量提前转化为东西向流量，同时利用 IPv6 可扩展头丰富的可编程空间，开展 IPv6+网络新技术包括但不限于 VPN+（网络切片）、IFIT（随路网络检测）、SFC 等和新应用开发，实现城域算力基础设施互联，通过业务的部署和资源调整来保证应用的 SLA 要求，以此提供业务链服务，如图 2 所示。

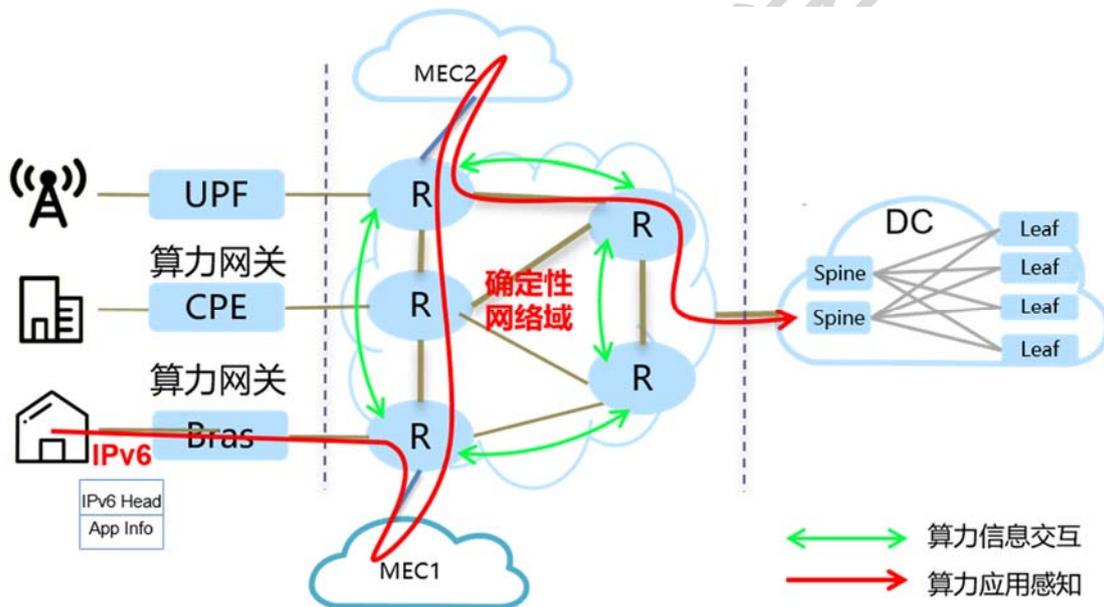


图 2 算力网络资源感知和信息交互示例

当前，CCSA（中国通信标准化协会）TC3 正在开展算力网络需求与架构的研究工作，其思路是通过网络、存储、算力等多维度资源的统一管理和协同调度，实现连接和算力在网络的全局优化，技术实现上可分为集中式方案、分布式方案、及混合式方案等多种技术路线。除基于数据中心 SDN 集中调度的方案，还可基于电信运营商承载网分布式控制能力，结合承载网网元自身控制协议扩展，复用现有 IP 网

络控制平面分布式协议的方式实现算力信息的分发与基于算力寻址的路由，如图 3 所示，同时综合考虑实时的网络和计算资源状况，将不同的应用调度到合适的计算节点处理。

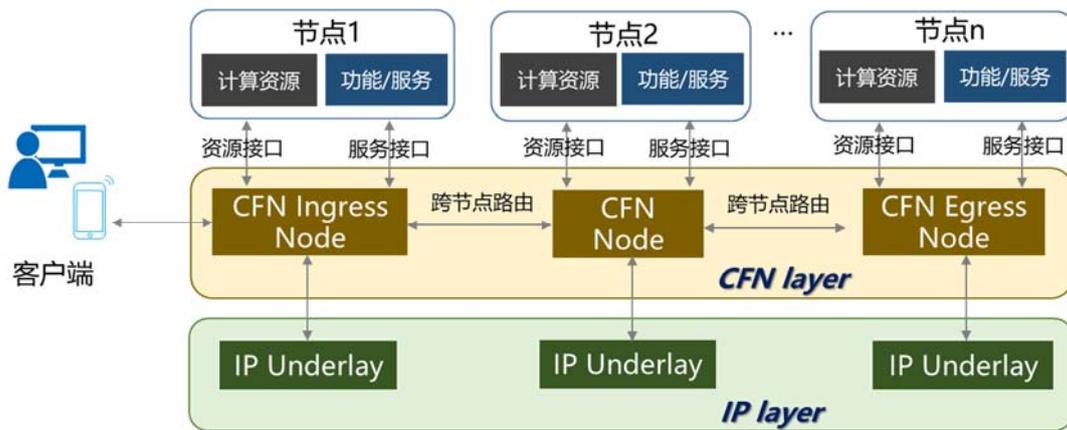


图 3 算力网络分布式控制方案

同时，面向高质量算力服务要求，算力网络还需要引入确定性网络技术，通过资源预留和队列管理算法来避免高优先级报文之间的冲突，提供显式路径，对报文的路由进行约束，最后通过冗余保护能够在某一条路径发生断路丢包时无损切换到另一条路径，保证业务的高可靠传输，实现对特定业务端到端时延的控制，打造确定性且有差异化的算力连接。

(三) 算力网络编排技术

算力网络是融合计算、存储、传送资源的智能化新型网络，通过全面引入云原生技术，实现业务逻辑和底层资源的完全解耦。需通过打造如 Kubernetes 的面向服务的容器编排调度能力，实现服务编排面向算网资源的能力开放。同时，可结合 OpenStack 的底层基础设施的资源调度管理能力，对于数据中心内的异构计算资源、存储资源和网络资源进行有效管理，实现对泛在计算能力的统一纳管和去中心化的

算力交易，构建一个统一的服务平台，如图 4 所示。



图 4：基于云原生的算力建模与服务编排示例

图 4 中，算力是设备/平台处理、运行业务的关键核心能力。在算力网络中，算力的提供方不再是专有的某个数据中心或集群，而是将云边端这种泛在的算力通过网络化的方式连接在一起，实现算力的高效共享。因此，算力网络中的算力资源将是泛在化的、异构化的。

算力网络中的算力资源包括传统的 X86 通用服务器架构下的 CPU 计算单元，专门适用处理类似图形图像等数据类型统一的 GPU 并行计算芯片，专业加速处理神经网络的 NPU 或 TPU，广泛应用于边缘侧嵌入式设备的 ARM，半定制化处理器 FPGA 等等。由于所运行算法的不同，所涉及的数据计算类型不同，从业务的角度出发，将算力分为可提供逻辑运算的算力、可提供并行计算的算力和可提供神经网络加速计算的算力。

对于异构算力资源，算力网络架构采用基于“K8S+轻量化 K8S”的两级联动的架构来实现统一的算力资源调度纳管。K8S 采用中心的资源调度统一平台对于整体的基础资源进行统一管理和集群管理，而轻量化 K8S 集群主要是作为边缘侧得资源调度平台对于边缘计算集群进行调度和管理。

泛在算力资源的统一建模度量是算力调度的基础。针对泛在的算力资源，通过模型函数将不同类型的算力资源映射到统一的量纲维度，形成业务层可理解、可阅读的零散算力资源池。为算力网络的资源匹配调度提供基础保障。将业务运行所需的算力需求按照一定分级标准划分为多个等级，这样可为算力提供者设计业务套餐时进行参考，也可作为算力平台设计者在设计算力平台时根据所需运行的业务对平台算力的选型设计提供依据。

(四) 算力网络转发技术

传统的网络设备采用转控一体的工作模式，其转发芯片的功能相对固化，“紧耦合”的网络设备难以支撑算力网络对设备灵活性及可编程性的需求。新一代高性能可编程数据包处理芯片加上 P4 等编程语言的出现，让网络拥有者、工程师、架构师及管理员可以自上而下地定义数据包的完整处理流程。除了帮助算力网络实现最适合其自身需求的具体网络行为外，可编程芯片还能使芯片供应商专注于设计并改进那些可重用的数据包处理架构和基本模块，而不必纠结于特定协议里错综复杂的细节和异常行为。因此，可编程芯片技术产生，为算力网络转发面提供了相应的技术支撑。

目前两种主流的数据平面可编程语言为 P4 和 NPL，其中基于 P4 语言描述的处理逻辑如图 5 所示。该语言主要用于定义新协议和网络相关功能（例如带内网络遥测，VNF 卸载规则，传输层负载均衡方案等），并以高级编程语言的形式呈现。

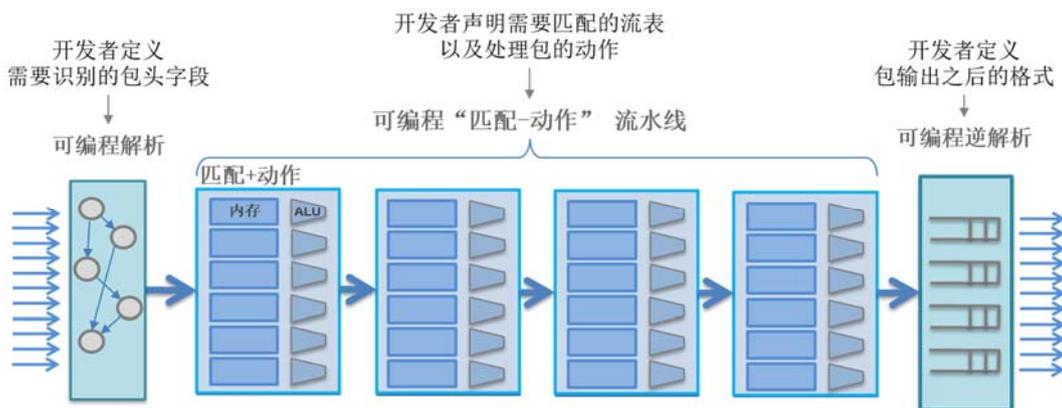


图 5 基于 P4 的协议无关可编程架构

在基于 P4 的协议无关可编程架构下，用户通过 P4 语言实现全流水线可编程，当算力网络需要引入诸如 SRv6 等新协议时，用户只需向 P4 程序添加新的逻辑即可，而基于 NPL 语言实现的编译器可编程芯片，可支持系统工作时的各种功能更新。两种方式都在很大程度上缩短了设备的开发周期，同时也提供了更快的数据包处理速度用以支撑算力网络的后续部署。最新一代 P4 可编程芯片的总带宽已达到 12.8Tbps，而最新一代 NPL 可编程芯片总带宽范围是 2Tbps 到 12.8Tbps。

五、算力网络工程难题

(一) 算力资源的感知与度量

算力网络在工程实际应用中首先面临的是算力的感知与度量，进而才能实现对算力的编排并合理快速匹配业务需求。目前，如何感知算力，并通过有效建模形成统一度量的算力资源，并能够合理编排来满足业务需求，是算力网络研究的重点和难点之一。

随着 5G 人工智能等技术的发展，算力网络中的算力提供方不再是专有的某个数据中心或计算集群，而是云边端这种泛在化的算力通

过网络连接在一起，实现算力的高效共享。因此，算力网络中的算力资源将是泛在化的、异构化的。目前市面上不同厂家的计算芯片类型形式各异，如英伟达 GPU，寒武纪的 ASIC，以及近年出现的 NPU、TPU 等，这些芯片功能和适用场景各有侧重，如何准确感知这些异构的泛在芯片的算力大小、不同芯片所适合的业务类型以及在网络中的位置，并且有效纳管、监督。针对泛在算力的纳管方法也不再适合采用统一的调度方式，业界提出的两级联动调度模式正在研究中。

再者，所感知的算力需要映射到统一的度量单位，屏蔽不同算力类型的差异性，以形成上层资源调度器或编排器可理解、可量化的资源层。但是对于业务运行，不光要有足够的算力，同时也需要配套的存储能力、网络能力，甚至还可能需要编解码能力、吞吐能力等来联合保障用户的业务体验。然后基础算力难以度量，无法直观合理的给出基础算力的统一度量方法，建议从微服务的角度来衡量算力，并建立相应的资源调度分配的标准和机制，降低算力网络中业务和应用部署复杂度，简化业务管理流程和机制。

(二) 集中式控制与分布式控制的协同

算力网络控制方案的实现有集中式和分布式两种。集中式控制方案是基于数据中心 SDN 集中调度的方案，即在云数据中心内部，由多个分布式应用服务器节点构成集群，分担业务计算与存储请求，同时云数据中心向城域网扩展，与边缘云相连接，通过集中式的 SDN 控制器和 NFV MANO 实现中心云及边缘云间的算力网络的统一管理和协同调度。分布式控制方案即基于电信运营商承载网分布式控制

能力，结合承载网网元自身控制协议扩展，复用现有 IP 网络控制平面的方式实现算力信息的分发与基于算力寻址的路由，同时综合考虑实时的网络和计算资源状况，将不同的应用调度到合适的计算节点处理，实现连接和算力在网络的全局优化。

对比集中控制与分布控制两种方案，前者能够做到算力节点的路由可达，配置通过集中式的 SDN 控制器可快速实现。但该方案的问题是计算节点无法快速与网络属性联动，也较难与 underlay 网络联动，因此建议作为算力网络初期的实现方案。另一种分布控制方案，需要根据具体的业务需求选择 BGP 扩展的种类和形式，实现比较复杂，也尚未标准化。但是该方案充分调动了承载网中 IP 路由器节点的控制能力，应用可以感知路径中沿途的所有节点的服务质量，是真正意义上的计算需求向网络开放，建议用于算力网络的远期实现，同时分布式方案也更适合具有基础网络资源的电信运营商采用。

六、政策建议

(一) 技术政策建议

1. 大力推进算力全网感知与端管云协同

“IPv6+”拉通端管云以实现统一的网络配置，可以满足云网融合的灵活组网、业务快速开通、确定性传输、优化用户体验按需服务等需求。国内运营商和设备商为了满足 IPv6+ 应用场景如 5G 与边缘计算等，选择计算、存储等基础 IT 资源离散部署或多级部署的方式，虽然短期看有利于快速满足应用场景需求，但长期看，随着资源节点出现泛在化的特点，还需解决以下问题，一是如何使得多方、异构的

资源整合在一起，解决云和网的灵活对接、云网资源的统一管控；二是如何在满足用户业务需求的情况下，实现资源利用的整体最优化的问题。在国内运营商和设备商的积极推动下，目前算力网络协同的标准正在制定中，需要研究云服务应用感知、算力资源及时调用与网络能力开放之间的协调机制，以便更好地推进云网融合，促进算力的端管云拉通。

2. 积极推动算力网络的标准化工作

算力网络的标准化工作处于前期开展阶段，我国的研究成果目前处于领先状态。后续建议国内运营商和设备商结合自身的标准研究与应用实践，将标准推向国际化，进一步推动算力网络技术的标准化程度。以解决对泛在计算和服务感知、互联和资源控制和调度的问题及满足未来新应用场景需求研究为导向，重点推进以下标准化工作：

算力网络架构及接口：从网络算力服务调度的需求出发，开展网络、算力、应用等多维度资源感知的融合架构设计，制定相关的总体技术要求及接口标准。

应用及算力感知研究：研究算力、网络和存储等多维资源感知，实现多维资源感知、调度的协同机制。

算力需求量化与建模研究：针对泛在的算力资源，通过模型函数将不同类型的算力资源映射到统一的量纲维度，形成业务层可理解、可阅读的零散算力资源池。

算网资源可信与协同：解决资源可信与协同问题，为需求方提供更多选择，促使算力流动起来，促进应用发展。

(二) 产业政策建议

1. 促进计算产业和网络产业融合与能力互补

当前，我国以数据中心为代表的算力基础设施建设尚无法完全满足智能化业务对于服务高实时性、多场景适配和资源高效利用的要求，须通过研究网络架构创新，来解决“局部算力过剩”和“全局算力不足”等问题，并实现以算联网，以网促算。以算网一体为核心，构建弹性开放、高效协同的计算基础设施，促进计算产业和网络产业融合发展与能力互补。

2. 加快推进算力向边缘下沉

随着智能化业务对计算实时性要求的不断提升，算力向边缘下沉已成为计算产业的发展趋势。目前，以边缘计算为代表的相关技术研究已经取得了系列成果，应用推广已逐步开展，建议产业界以典型应用为切入点，探索边缘计算典型商业模式，挖掘新机会和新应用点，以示范应用带动整个技术的成熟和产业化发展。在国内，以相关产业联盟为典型代表，通过成员间互通协作，将行业使用场景和相关应用相结合，围绕电信运营商边缘计算、企业与物联网边缘计算、工业边缘计算三种场景，进一步构建创新的边缘计算业务形态与解决方案和应用，同时推动边缘计算的开发者生态，加速产业发展，突破创新。

3. 注重算力网络产业的自主可控

在当前国际竞争背景下，网络领域的自主可控是一项突出问题，算力网络也不例外，要把算力网络技术的自主可控作为研究的重要内容贯穿始终。这个过程不仅需要国家相关部门牵头组织、政策性扶持，

同时更需要产学研用各个参与方的积极推动。但是，大力发展自主可控并不意味着故步自封，闭门造车，自主可控的策略应该是在我国企业掌握核心竞争力的基础上，以积极开放的态度拥抱开源，在全球范围内共建共享算力网络生态。

CIC中国通信学会

中国通信学会

地址：北京市海淀区万寿路 27 号院 8 号楼

邮政编码：100840

联系电话：010-68203021、68203019

传真：010-68203004

网址：<https://www.china-cic.cn/>

