

# 算力网络四面三级算力度量技术体系



## Three-Level and Four-Aspect Computing Measurement System in Computing Force Network\*

杜宗鹏/DU Zongpeng, 李志强/LI Zhiqiang, 陆璐/LU Lu

(中国移动通信有限公司研究院, 中国 北京 100053)  
(China Mobile Research Institute, Beijing 100053, China)

DOI: 10.12142/ZTETJ.202304003

网络出版地址: <http://kns.cnki.net/kcms/detail/34.1228.TN.20230717.1807.006.html>

网络出版日期: 2023-07-18

收稿日期: 2023-05-26

**摘要:** 算力网络要实现广泛的算力互联, 就离不开统一的算力度量和算力建模机制。算力度量技术主要解决算力描述的标准化, 提供方便的跨厂商互通能力、算力资源协同管控能力。对算力网络中算力度量需求进行分析, 并提出四面三级算力度量技术体系, 同时还探索了在算力路由决策中, 网络按需进行算力信息通告等关键技术, 为后续的算力度量研究提供参考。

**关键词:** 算力网络; 算力度量; 算力建模; 算网融合; 算力感知流量调度

**Abstract:** Unified measurement and modeling mechanisms are the guarantee for the computing force network to realize extensive computing interconnection. The computing measurement technology mainly solves the standardization of computing information description and provides convenient cross-vendor communication ability and collaborative control ability of computing resources. The three-level and four-aspect measurement system is proposed, and the key technologies such as on-demand computing information transmission for the computing aware routing decision are also explored, to form a reference for future computing measurement research.

**Keywords:** computing force network; computing measurement; computing modeling; computing and network convergence; computing-aware traffic steering

**引用格式:** 杜宗鹏, 李志强, 陆璐. 算力网络四面三级算力度量技术体系 [J]. 中兴通讯技术, 2023, 29(4): 8-13. DOI: 10.12142/ZTETJ.202304003

**Citation:** DU Z P, LI Z Q, LU L. Three-level and four-aspect computing measurement system in computing force network [J]. ZTE technology journal, 2023, 29(4): 8-13. DOI: 10.12142/ZTETJ.202304003

随着互联网技术的不断进步, 当今社会的发展逐渐呈现数字化、智能化的趋势。智能运算需要大量的算力来完成数据处理。为了缓解云数据中心的计算压力, 获得更快的业务响应速度, 算力逐渐从中心走向边缘, 形成网络中泛在的算力资源。为了支持分散算力资源的统一感知、统一决策和统一管理, 算力网络的概念被提出<sup>[1-5]</sup>。算力网络将提供算力和网络的一体化服务, 支持算力资源和网络资源的联合优化, 从而充分利用有限的算网资源, 为用户提供高品质的算力服务。

不同算力业务的算力需求有所不同, 不同算力节点的服

务能力也各不相同。要实现灵活高效的业务调配和服务映射, 算力网络首先需要对算力进行感知, 而算力度量就是算力感知的基础之一。正如热力学温标的提出者开尔文勋爵曾提到的: “如果你无法测量它, 那么你就无法改进它”。

传统运营商主要提供网络资源, 在用户需求、网络能力度量方面也有相对成熟的方案。网络度量常见的度量方式包括带宽和流量等。运营商围绕带宽和流量提供网络服务, 进行网络运营, 同时也可以基于时长、网络质量, 与用户进行签约, 并提供网络服务。

算力网络提供的是算力和网络的综合服务, 这时需要一种方便对算力进行度量的机制。该机制一方面可以支持用户对算力需求的描述, 另一方面也可以支持运营商、服务提供商或第三方对所提供的算力能力的描述。

**基金项目:** 北京邮电大学-中国移动研究院联合创新中心基金资助 (CMYJY-202000332)、东南大学-中国移动研究院联合创新中心基金资助 (CMYJY-202100163)

\* 作者确认算力网络译为 computing force network

## 1 算力度量的需求架构

在算力网络中，算力度量的目标是关联和整合网络中的异构计算资源，使能多维资源的统一协同管理，从而通过统一的算力度量体系和异构计算资源的服务映射机制，实现算网资源的合理分配和高效调用。

目前，中国通信标准化协会（CCSA）TC3 已经布局了算力网络的相关标准，包括算力网络总体技术要求、算力路由、算网编排、算力交易等。根据相关的工作<sup>[2, 6]</sup>，算力网络包括算力服务层、算力路由层、算网基础设施层和算网编排管理层。图1展示了算力网络总体架构中与算力度量相关的功能模块。在算力网络中，算力用户需求和算网资源的映射决策可能发生在算力路由层的策略决策模块或是算网编排管理模块。

目前，针对算力网络系统中网络资源的感知和通告，我们可以使用已有的网络操作维护管理（OAM）机制进行测量和收集。本文中，我们主要关注算力资源的感知和通告，以及算网信息的联合决策，尤其是算力度量信息在这些功能或流程中的需求和作用。

算力资源的有效管理，首先需要有一个统一的算力资源模型，即对异构的算力资源进行建模。算力建模的目标是构建异构算力资源的统一描述方式。算力建模的具体信息包括数值型和非数值型。算力度量是算力建模的一部分，而算力资源的统一建模是算力度量的前提。算力度量中的算力评估值，可以更为方便地作为算力决策模块的输入参数，从而影响算力服务的部署和调度。算力度量和建模的相关需求，主要包括以下3个方面：

1) 支持算力资源的度量和建模，对运营商、服务提供商及第三方算力资源节点所提供的算力进行可量化的能力描述；

2) 支持算力用户需求的度量和建模，对用户算力业务进行可分类分级的需求描述；

3) 算力路由策略决策模块或算网编排管理模块可基于用户需求，合理分配算力节点来完成计算任务。

其中，前2个方面主要涉及算力度量和建模信息的抽象和通告，即把算力信息或算力需求信息按照约定的格式发送至算力路由策略决策模块或算网编排管理模块；第3个方面主要涉及算力度量信息的使用，即算力路由策略决策模块或算网编排管理模块如何利用算力度量和建模信息进行决策和管理，执行算网感知调度。

## 2 算力度量信息的使用

算力路由策略决策模块和算网编排管理模块都需要收集算力度量和建模信息，并据此进行策略决策。算力路由策略决策模块主要生成路由相关的策略，可根据感知到的算力服务的总体状态信息（尤其是算力服务节点当前的算力业务会话接入能力等动态参数），进行快速决策。算网编排管理模块支持更复杂的处理逻辑，例如与人工智能（AI）分析平台互通，根据更全面的算网信息进行网络和算力资源的全局优化。

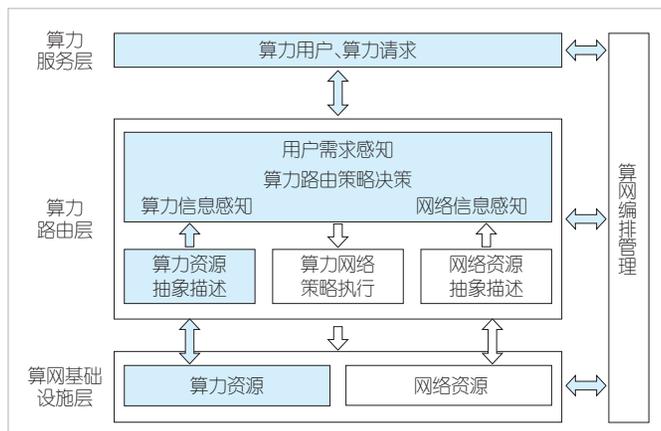
算力路由策略决策模块可根据收集到的算力信息和网络信息，生成策略并执行，引导算力业务流量的高效转发。算力业务的流量指的是用户请求算力业务时所使用的连接建立报文，以及所需传给服务节点的数据报文，例如需要被分析的图片数据信息等。同时，算力路由层也支持算力业务的用户需求感知，以按需提供算网资源。

算力路由层的功能可以在网络控制器/网络转发设备的控制面上实现。算力路由层的策略决策模块可将收集到算力信息保存在一个或者多个专用的表中，并且支持按需更新。网络转发设备中默认的路由转发架构可以保持不变。这时，算力网络的路由策略优先级更高，即优先按照算力网络策略对算力业务流量进行转发。

算网编排管理模块支持算力调度策略的生成，甚至可与算力路由决策模块进行策略交互。算网编排管理模块感知到的信息会更加全面，因此能够支持更加复杂的决策逻辑，例如利用AI机制进行训练和推理，优化全网的算力业务。同时，算网编排管理模块需要根据感知到的算力资源情况，选择合适的算力节点，进行算力服务实例的部署。

### 2.1 算力服务部署

算力网络服务在部署时，需要对云、边、端各级架构中泛在异构设备进行纳管，还需要对中央处理器（CPU）、图



▲图1 算力网络中算力度量需求

形处理器（GPU）、内存、磁盘、网络等多维算力资源进行感知和维护。当有服务部署需求时，算力网络系统需要基于维护的资源视图，通过适配算法为用户找到合适的算力资源以完成服务部署。

算力网络中各种异构的算力资源散落各地，用户不同的部署服务对各维度算力资源也有着不同的需求。算力服务部署要解决的问题是在掌握资源状态信息的情况下，如何合理调度资源，部署服务。算力服务的部署是算力网络实现低时延、高可靠优质服务调度的前提。

在云计算或边缘计算的领域，算力数据模型相关的多项工作都可以作为算力网络中算力度量和建模的参考。例如，用于云服务器带外管理的 Redfish 项目从不同维度定义了大量算力资源相关的模型；创建于 2021 年的 Anuket 项目支持云原生和虚拟网络功能、基础设施和服务的快速部署，致力于整合 OpenStack 和 Kubernetes（K8s）等知名工具的不同架构。

云原生中流行的编排管理工具 K8s 支持集群内部的服务部署和资源调度。一个 K8s 集群包含至少一个主节点和若干从节点，主节点为集群管理节点，能根据各节点的资源状态完成容器到物理节点的分配<sup>[7]</sup>。K8s 中的最小调度单位为 Pod，一个 Pod 可以包含多个容器。每个 Pod 都可以在部署时指定其 CPU 和内存的数量。K8s 的资源调度是指将 Pod 任务映射到物理资源的过程。在该过程中，主节点中的调度器 Kube-scheduler 负责决策一个 Pod 应该被调度到集群中的哪一个节点上。在服务节点的具体选择时，可以指定不同的节点优选策略，例如优选算力资源占用比较小的节点，或者是优选 CPU 和内存使用率接近的节点。

资源的感知和监控是资源调配的前提。有效的资源感知能够精确地传递资源状态信息，从而为后续的资源调度、服务迁移等业务奠定基础。例如，作为一款基于时序数据库的监控告警系统<sup>[8]</sup>，开源软件 Prometheus 支持多维数据模型，收集由度量名和键值对组成的时间序列数据，被广泛应用于云原生中的资源监控中。

算力网络的服务部署和资源调度模式可以参考上述云原生中的编排管理工具，在面向异构立体泛在的算力场景方面进行增强。通过了解各个算力节点的算力资源和剩余可用资源的情况，并根据算力业务需求，按需制定合适的部署策略，灵活地调度资源，从而完成算力服务部署，提升算力服务能力。

## 2.2 算力服务选择

在完成算力服务部署后，算力网络需要为算力用户找到

并接入一个合适的算力服务节点。该节点需要离用户较近，满足用户的时延需求，还需要有充足的算力资源，支持算力业务的快速完成。这个过程被称为算力服务选择，或算力服务节点的负载均衡。算力服务节点的状态是实时变化的，因此在网络中实时同步这些信息存在一定的挑战。

在算力网络服务场景中，不同网络位置能够提供相同的算力服务，然而，此时不同服务节点的计算资源通常是不同的，计算资源的负载情况也在变化。尤其是在多接入边缘计算（MEC）的场景中，计算资源相对有限，因此 MEC 之间的协同就显得更加重要。

算力感知网络（CAN）<sup>[3,9]</sup>为 MEC 的协同提供了一种基于网络的负载均衡机制。在相关场景中，整个网络就像一个虚拟的负载均衡设备一样运作。在 CAN 的机制中，策略决策点是一个网络设备，位于网络转发节点或是网络管控节点上。然而，对于一个网络设备而言，了解算力的状态还是一个新的技术领域。

## 3 四面三级算力度量技术体系

网络所提供的服务主要聚焦于转发，例如 IP 网络的核心就是按照目的地址，把报文转发到目标网络。计算所对接的服务会更加复杂，影响计算性能的因素也更多，因此较难给出单维度的度量单位。网络中的算力资源分散，异构算力资源种类繁多，是算力统一度量所面临的主要困难。

算力芯片的种类包括 CPU、GPU、嵌入式神经网络处理器（NPU）、张量处理器（TPU）等。这些芯片都有各自的性能指标和应用场景，可以针对特定业务进行计算优化。例如，通用算力以 CPU 承载为主，主要面向通用软件应用，执行逻辑运算；人工智能算力以 NPU/TPU/GPU 承载为主，主要面向 AI 应用，逻辑简单，但计算密集、并发程度高；超算算力以 CPU/GPU 承载为主，主要面向科学计算、工业仿真等场景，计算复杂且对计算精度要求高。

在计算领域中，常见的一个度量单位是 FLOPS，即每秒执行的浮点运算次数。但计算业务的执行效率和能力，并不仅仅取决于计算节点的浮点运算能力，其影响因素还包括输入/输出（I/O）效能、内存架构、缓存速率等。因此，在算力通告和服务调度中，需要综合考虑如算力节点的负载等的多种因素。

算力服务的性能受到芯片、存储、网络、平台软件等多维因素的影响，因此我们需要对算力服务节点进行综合评估。算力的度量和建模不仅可以参考制造商提供的规格算力值，例如 FLOPS，还可以参考一些可以全面反映服务支持能力的综合指标值。为了描述算力网络中服务节点的算力，我

们提出了“四面三级”算力度量技术体系，从节点的计算、通信、内存和存储能力4个方面，按照三级指标的方式来建模和描述算力网络中异构立体泛在的算力，如图2所示。“四面”指标具体如下：

1) 计算。节点资源的计算能力评估主要是指对CPU、GPU等计算资源运算能力的评估。根据计算程序的不同特点和需求，计算性能可分为整数计算性能、浮点计算性能以及哈希计算性能。

2) 通信。节点资源的通信能力评估主要是指对节点连接到外部网络的通信能力的评估。从单个算力节点的角度考虑，节点的通信能力指标主要参考节点连接到外部网络的带宽大小。

3) 内存。节点的内存（缓存）能力评估主要是对节点的内存（缓存）单元的性能评估。节点的内存（缓存）指标主要涉及节点内存容量和内存带宽。

4) 存储。节点的存储能力评估主要是指对节点的外存储器（如硬盘）的性能评估。节点的存储指标主要涉及节点的存储容量、存储带宽和每秒读写操作数（IOPS）。

“三级”指标具体如下：

1) 一级指标代表异构硬件算力度量，包括整数计算速率、浮点计算速率等；

2) 二级指标代表节点服务能力度量，包括节点可提供的计算、通信、内存、存储4方面的能力等；

3) 三级指标代表节点对业务的支撑能力度量，包括节点业务处理能力度量等，可以按照不同的服务类型进行描述。

上述度量体系中的各种指标可以根据业务需求的不同，按需提供给其他的算力网络节点，如决策节点。下文中，我们将简要介绍上述相关指标的获取方式，以及一些示例性的指标参数。

1) 一级指标可以通过不同的评估方法来获得，如通过

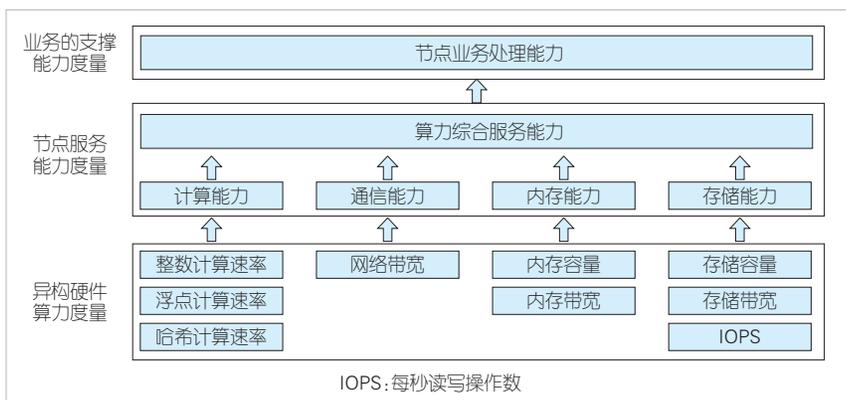
基准测试程序测试、直接从厂商获取、公式计算等方式。该级别的指标可以是厂家提供的性能参数，例如CPU的型号、主频、内核数，GPU的型号、单精度浮点性能等，也可以是常用的基准测试程序的测试值。厂商提供的性能参数也可以包含动态的资源状态，例如目前可用的CPU/虚拟CPU（vCPU）核数等信息；基准测试程序的测试可分为整数计算性能测试、浮点计算性能测试等。

2) 二级指标由一级指标抽象得到，主要用于对算力的综合评价，可提供节点某方面的或是综合的服务能力。节点某方面的能力可以通过执行对应方面的多次测试，并且基于输出结果进行运算后得到；综合服务能力可以基于多个方面的评估值，再次进行运算处理得到。文献[10]提供了一种基于PageRank的节点性能评价算法。在相关流程中，首先对每个节点使用主流基准测试进行评价，然后采用PageRank算法处理每个基准测试的执行结果，从而得到节点某方面的性能指标。

3) 三级指标与节点上部署的业务相关，围绕业务的维度进行描述。三级指标提供的主要是与业务相关的评价指标，例如针对特定的算力业务，节点能够提供的实际处理能力等。文献[11]提供了一种“有效算力”的计算方法，用真实业务软件在一定规模的信息通信技术（ICT）基础设施系统上（含计算、存储、网络、软件中间件等）进行性能测量，之后通过几何加权平均，对多个真实业务性能测试结果进行整合。例如，在人工智能领域，对特定训练作业的有效计算能力是单位时间内训练过程消耗的样本数量，或称为业务吞吐量。在视觉类测试中，业务吞吐量单位是图片数每秒；在自然语言处理类测试中，单位为句数每秒。

上述度量体系中各个维度的信息，可以用于网络算力服务的调度和算力业务的转发决策，从而提高算力资源和网络资源的利用率。同时，上述评估机制反应了算力节点对计算任务的支撑能力，牵引最终用户从使用效果的角度，进行算力节点的规划和建设。

算力度量的指标包括静态指标和动态指标。前者通常与分配给服务的算力资源有关，后者通常与服务器在运行时能够提供的可用资源有关，例如，服务器当前的可用会话数或CPU利用率。在算力服务的部署和节点选择中，上述的动态指标更具参考性。例如，在部署算力服务时，需要了解当前算力节点可用的CPU/vCPU核数，并将其视为一种较为动态的一级指标；在选择算力服务节点时，业务相关的动态指



▲图2 算力资源抽象描述的四面三级度量体系

标包括当前服务节点状态等，可将其视为动态的三级指标。

这些动态指标可以与上文所述的规格算力信息、综合节点算力信息、业务处理能力信息综合考量，从而影响算力服务的决策。这些动态指标可以指定不同的更新频率与算力网络的决策点进行通信，更新相关的动态算力信息。

## 4 算力度量的关键技术

算力网络的定位是面向未来网络中泛在的算力资源，支持网络和算力的融合统一，提供算网一体化服务的新型信息基础设施。算力网络的愿景是推动算力像水电一样，逐步成为“一点接入、即取即用”的社会级服务<sup>[1]</sup>。算力度量在算力网络中的作用类似于水网、电网中水表和电表的作用，是算力网络的重要基础技术之一。

为了实现异构算力资源的统一度量和统一建模，支持跨厂商算力网络节点互通，方便算力网络的管理和运营，算力度量和建模机制标准化工作已经被提上日程。目前，中国通信标准化协会（CCSA）已经有了一些正在进行的算力度量相关的标准化项目。同时，国际互联网工程任务组（IETF）的CATS工作组也在推进该领域的工作。

算力信息的格式设计与使用场景密切相关，不同的使用场景有着不同的算力度量指标的组合和选择。本文中，我们围绕“四面三级”算力度量技术体系，对算力信息如何在网络中传递也进行了探索。其核心的两个问题为：1) 算力信息的格式，即应该传递什么样的算力信息到网络中；2) 算力信息的通告方式，即在网络中以什么样的方式通告算力信息。

### 4.1 算力度量信息的格式设计

算网编排管理模块可以通过定义算力相关的YANG模型来传递算力信息；算力路由模块可以通过扩展目前的路由协议来传递算力信息。本节将主要探索后者的格式，并给出一种示例性的描述方式。路由协议携带的算力信息会更加简单，一方面这样可以减少网络中的信息传递，避免不必要的信息暴露；另一方面，可以使路由决策流程快速完成。

不同的业务可能有不同的计算需求，需要使用不同的算力路由决策依据。因此，算力度量的指标体系应该可扩展、可编程，支持灵活自定义，以便能够反映不同业务甚至未来业务需求。对于特定的业务，决策点应能订阅想了解的算力度量指标，以及指定指标的相关更新频率。也就是说，决策点可以按需选择感兴趣的各级细节指标进行了解，并作为决策的依据。

作为参考，在K8s的Pod资源调度中，调度器会执行初

选和优选两个步骤<sup>[7]</sup>。初选时，调度器会过滤掉不满足Pod资源需求的节点；优选时，调度器会根据设定的策略对初选后的节点进行评分，并据此决定调度节点。在算力网络中，对于一个选择服务节点的业务而言，首先要确定的是目标服务节点是否还能够接入新的会话，其次需要按照业务特定的策略，优选综合性能更高的算力节点来提供服务，例如有些业务有较快的业务完成时间的需求。在文献[12]中，建议使用动态的参数来反应服务状态，使用相对静态的参数来反应节点执行该服务的能力信息。

本节中，我们提出了一种示例性的算力度量信息的格式设计思路：首先要携带算力节点的服务状态信息；其次，需要按照业务需求或优化目标，携带一个或者几个算力信息相关的类型-长度-值（TLV）。这里不同的业务类型可以根据业务需求，自定义不同的TLV组合。

在服务节点选择时，算力路由的决策点首先要感知备选的算力节点是否还能接入新的算力用户。例如，文献[9]提到了一种算力节点忙闲状态的通告方式。该方式设定了3种状态：红色代表繁忙状态，绿色代表空闲状态，黄色代表即将繁忙的状态。我们可以在服务节点上设定阈值，触发这些状态的改变，例如接入的用户数达到阈值，或某种资源（如CPU、内存等）的利用率达到阈值。

如果多个算力节点都标记了绿色的空闲状态，同时算力业务又有较快的业务完成时间需求，那么也可以进一步地按照算力节点的综合性能来决策接入哪个节点。这些性能的综合评估也是算力度量体系的一部分，一般体现为三级指标或二级指标，或多级指标的组合等。

如果面向的业务较为明确，并且存在相关业务的有效算力测试结果，那么可以优先考虑使用三级指标中的有效算力来评估算力节点的计算能力。对于相同的业务，如果算力用户之间的计算需求差异较大，仅使用三级指标并不能很好地反映节点的计算能力；或是由于面向新的业务，没有可参考的有效算力结果评估，那么可以考虑二级指标中的综合服务能力指标。在该指标中，针对不同的业务，在计算、通信、内存、存储4个方面有不同的系数，即影响因子。例如，目标业务如果是计算密集型的，就可以适当调高计算维度的影响因子。

如果业务对于能耗比较敏感，除了算力相关的信息外，也可以要求算力节点提供能耗相关的信息。这时需要对算力节点的能耗进行统一建模，从而将其作为算力服务调度的参考因素之一，即在一定程度上优先使用较为节能的算力节点来完成计算任务。

### 4.2 算力度量信息的通告

CAN是算力网络中的关键技术之一。图3为CAN算力通告流程：首先，对于泛在的算力节点，需要统一度量值的描述方式；其次，通过某种方法，将这些节点的算力能力通告到网络中；最后，算力网络的决策点可以参考收到的算力信息来进行算力服务调度。上一节主要关注的是第一步，即描述算力能力和服务状态的方法。一个标准化的算力能力的描述方法能够在很大程度上提升算力网络不同节点之间的协同能力。本节中，我们主要关注第二步，这其中也会涉及相关的标准化工作。

通常，算力节点和CAN并不在同一个网络域中，因此算力信息的通告经常会考虑使用边界网关协议（BGP）。BGP可以用于网络中边缘路由器（PE）节点之间的路由信息交流，是计算机网络实现广泛互联的基础协议之一。针对特定的算力服务，可以考虑扩展BGP协议，如在通告该服务的路由信息时，BGP消息中携带节点的忙闲信息，以及节点的业务执行能力TLV。

在路由协议的相关扩展中，网络中的多个算力节点会共享同一个Anycast（任播）地址，并同时发布自身相关算力信息。一方面，多点部署的Anycast地址本身就不支持按照前缀汇聚；另一方面，不同的算力节点的算力情况不同，在BGP路由优选时，需要对BGP协议进行扩展以支持对同一个Anycast地址存在不同的Nexthop（下一跳）以及算力信息的情况。

### 5 结束语

算力度量是算力网络技术的重要基础之一。统一的度量标准、描述方式，有助于算力网络的各个网元之间进行高效协作。目前，算力网络业务以及针对相关业务的度量还处于发展的初期阶段。未来，随着算力网络业务的普及，统一的算力度量机制将促进算网融合共生，避免“算力孤岛”，实现全网算力的高效协同，从而使整个算力网络能够像一台虚拟的计算机一样统一调度、统一管理，支撑千行百业的数智化转型。



▲图3 算力感知网络中算力通告的流程

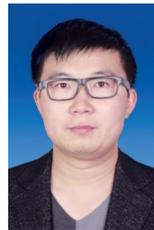
### 参考文献

- [1] 中国移动. 算力网络白皮书 [R]. 2021
- [2] CCSA. 算力网络 总体技术要求:CCSA TC3 WG3 [S]. 2021
- [3] 姚惠娟, 陆璐, 段晓东. 算力感知网络架构与关键技术 [J]. 中兴通讯技术, 2021, 27(3): 7-11. DOI: 10.12142/ZTETJ.202103003
- [4] 雷波, 赵倩颖, 赵慧玲. 边缘计算与算力网络综述 [J]. 中兴通讯技术, 2021,27(3):
- [5] 唐雄燕, 张帅, 曹畅. 夯实云网融合 迈向算网一体 [J]. 中兴通讯技术, 2021,27(3):
- [6] CCSA. 面向算网融合的算力度量与算力建模研究 [R]. 2021
- [7] ZHANG W G, MA X L, ZHANG J Z. Research on kubernetes' resource scheduling scheme [C]//Proceedings of the 8th International Conference on Communication and Network Security. ACM, 2018: 144-148. DOI: 10.1145/3290480.3290507
- [8] LEE S, SON S, HAN J, et al. Refining micro services placement over multiple kubernetes-orchestrated clusters employing resource monitoring [C]//Proceedings of 2020 IEEE 40th International Conference on Distributed Computing Systems (ICDCS). IEEE, 2021: 1328-1332. DOI: 10.1109/ICDCS47774.2020.00173
- [9] DU Z P, LI Z Q, DUAN X D, et al. Service information informing in computing aware networking [C]//Proceedings of 2022 International Conference on Service Science (ICSS). IEEE, 2022: 125-130. DOI: 10.1109/ICSS55994.2022.00027
- [10] SONG D, CHEN X S, RUI L L, et al. Resource allocation algorithm based on modeling of ubiquitous network node capability [J]. Journal of physics: conference series, 2022, 2224(1): 012089. DOI: 10.1088/1742-6596/2224/1/012089
- [11] 中国电子技术标准化研究院. 计算中心有效算力评测体系白皮书 [R]. 2022
- [12] DU Z P, FU Y X, LI C, et al. Computing information description in computing-aware traffic steering: draft-du-cats-computing-modeling-description-00 [S]. 2023

### 作者简介



**杜宗鹏**，中国移动通信有限公司研究院基础网络技术研究所研究员；研究方向为算力网络、未来IP网络、确定性网络等；发表论文10余篇。



**李志强**，中国移动通信有限公司研究院基础网络技术研究所研究员；长期从事未来IP网络演进、标准和技术研究工作，主要涉及下一代IP网络、算力网络、云网融合、SDN/NFV、5G核心网等。



**陆璐**，中国移动通信有限公司研究院基础网络技术研究所副所长、中国通信标准化协会TC5核心网组组长；长期从事移动核心网策略、演进、标准和技术研究工作，主要涉及未来网络架构、智能管道、边缘计算、算力网络等领域。