



北京 2022 年冬奥会官方合作伙伴
Official Partner of the Olympic Winter Games Beijing 2022

异构算力统一标识与服务白皮书

中国联通算力网络产业技术联盟

2021 年 3 月

中國美術通史

版权声明

本白皮书版权由中国联通算力网络产业技术联盟拥有，并受法律保护。转载、摘编或利用其它方式使用本白皮书文字或者观点的，应注明来源。

编写组成员

中国联通研究院：唐雄燕、曹畅、李铭轩、李建飞

华为技术有限公司：丁肇辉、刘飞、何军

中兴通讯股份有限公司：熊先奎、朱堃

联想未来通信科技（重庆）有限公司：褚遵利、崔先锋

浪潮软件科技有限公司：高娴、谢奥

安谋：邹婷

寒武纪：程归鹏

前言

2021年3月，中国联通发布了《中国联通 CUBE-Net 3.0 网络创新体系白皮书》。CUBE-Net 3.0 是以打造算网一体为重要目标，融合云原生、边缘计算、人工智能、内生安全等新的技术元素，强化要素深度融合，构建支撑经济社会数字化转型的新一代数字基础设施。

早在2019年，中国联通就发布了业界首部《算力网络白皮书》，阐述了未来算力业务形态、平台经营方式、网络使能技术等方面的观点，对算力网络的技术研究和产业推动产生了重要影响。2020年11月，中国联通在业界率先发起成立了算力网络产业技术联盟，旨在联合产学研各方力量加强技术交流与合作，加快推进算力网络的生态建设、标准制定和应用推广。该联盟成立以来，各成员单位已就相关技术开展了大量研讨，本次发布的《异构算力统一标识与服务白皮书》就是联盟成员合作产出的成果之一。

算力作为国民经济数字化转型过程中重要的资源，是算力网络整个架构中不可缺少的重要元素。在整体技术架构和新型网络连接研究的基础上，需不断深入研究底层算力资源的管理和调度机制。在中国联通算力网络产业技术联盟成员单位的大力协作下，本白皮书结合目前算力资源异构化发展的产业趋势，多产权主体算力接入和共享的商业模式，对异构算力统一标识技术开展重点攻关，以期构建“可信、可溯源、可度量”的算力管理架构，为上层应用和最终用户提供安全、可靠的计算平台。同时，通过算力资源抽象和服务化，进一步屏蔽底层算力的差异化，为用户提供无差别的应用环境，进一步降低各种计算资源的使用门槛。

希望本白皮书能在算力资源管理、业务建模和服务能力构建等领域为行业发展起到抛砖引玉的作用，同时为我国的算力基础设施建设提供有益参考。

目 录

1 异构算力产业发展现状	1
1.1 算力网络构建“云、边、端”泛在计算协同.....	1
1.2 算力从通用架构走向专用定制.....	2
1.3 异构算力的主要应用场景.....	4
2 异构算力发展需求.....	7
2.1 异构算力需要统一标识和调度	7
2.2 异构算力统一标识设计原则	7
2.3 异构算力统一调度设计原则	8
3 异构算力统一标识和资源抽象.....	9
3.1 异构算力统一标识	9
3.1.1 异构算力统一标识技术架构	9
3.1.2 可信算力的认证鉴权	10
3.1.3 统一资源数据模型	11
3.1.4 异构算力资源管理	12
3.1.5 算力度量与算力推荐系统.....	12
3.2 异构算力资源抽象	14
3.3 异构算力统一调度与适配	15
3.3.1 异构算力统一调度技术架构.....	15
3.3.2 算力调度编排平台功能	17
3.3.3 统一调度核心流程	18
3.3.4 异构算力调度智能运营	20
4 总结与展望.....	22
4.1 异构算力的融合发展	22
4.2 算力能力进一步开放	22
4.3 算力网络促进算力合理化布局	23
4.4 异构算力的标准化建议	23
5 参考文献.....	25
6 缩略语.....	26

1 异构算力产业发展现状

1.1 算力网络构建“云、边、端”泛在计算协同

互联网的高速发展使得万物数据化，数据量和计算量呈指数爆发，赛迪顾问数据显示，到 2030 年数据原生产业规模量占整体经济总量的 15%，中国数据总量将超过 4YB，占全球数据量 30%。数据资源已成为关键生产要素，更多的产业通过物联网、工业互联网、电商等结构或非结构化数据资源来提取有价值信息。而海量数据的处理与分析需要强有力的算力提供支撑。

数据处理对于算力需求的日益剧增，需要云、边、端算力实现高效协同。高效的定义是“在相同或更短的时间里完成比其他人更多的任务，而且质量与其他人一样或者更好”。对于目前海量、分散的数据处理场景，仅仅数据中心内部或者终端的高效算力处理，都已无法满足业务需求。因此需要广泛分布在云、边、端的算力能够有效的协同，来满足当今数据处理需求。“云、边、端”泛在计算协同是目前的算力网络正在积极探索和研究的领域之一。随着算力网络的进一步发展，在促进新型网络技术发展的过程中，同样也促进了云计算技术在数据中心以及跨数据中心的协同部署发展，近年来由于边缘计算、分布式计算等新型计算业态的出现，新型网络连接形态更好的实现了不同数据中心之间的连接，因此数据中心的云资源部署逐渐从传统的以集约化为主的中心云方式，主要向云计算和边缘计算之间的协同调度方式演进，而随着未来移动互联网的发展，正逐渐促进了包括终端侧的泛在计算的协同和发展。因此，算力网络技术所带来的新型的网络连接实现了以“云、边、端”为主的新型的算力资源的协同和发展，围绕三者结合实现的泛在计算构建了不同层级的算力组成，能够实现低时延、低成本等效果。并且结合大数据和算法，能够实现在不同层级的算力满足不同的场景需求，即，云计算满足大体量数据分析和算法训练，边缘计算满足敏捷接入和本地计算，而端侧算力侧重现场体验和端侧智能等要求。

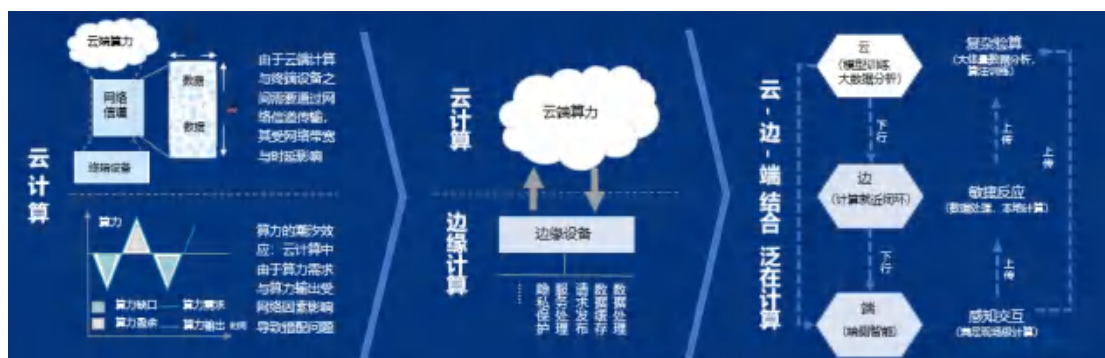


图 1-1 算力网络促进“云、边、端”协同

1.2 算力从通用架构走向专用定制

当前算力网络架构中的算力可以由不同的硬件架构组成，一般包括 CPU、GPU、FPGA 和 ASIC 等类型；CPU 主要有 X86 和 ARM，虽然 ARM 设计之初作为面向低功耗等场景推出的定制化的 ASIC 芯片，但是随着 ARM 在服务器和嵌入式终端的广泛应用，目前也是作为通用芯片来应用和部署；GPU 主要是快速实现矢量的图形化数据处理的专有架构；FPGA 作为可编程逻辑门电路在硬件加速等方面具有优势；而面向特定场景的处理需求需定制专用芯片来进行处理，比如当前针对深度学习设计的各种 TPU、NPU 属于领域的定制的 ASIC。

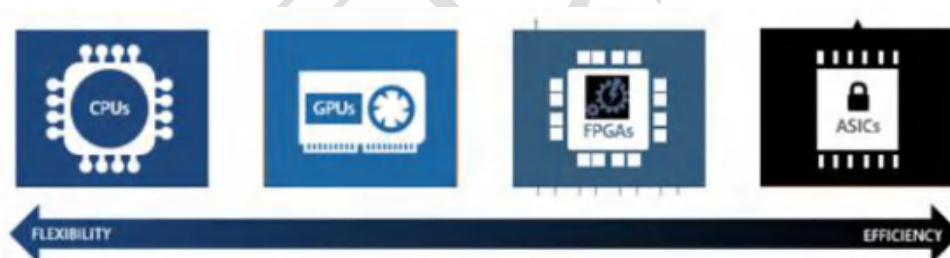


图 1-2 算力从通用走向定制

CPU 采用冯诺依曼架构，将计算分为取指、译码、发射、执行、写回等几个阶段，通过软件调度，可以完成任意特点计算。当前 CPU 的架构已经相当复杂，并且真正有效计算在 CPU 整体功耗比例中不到 10%，所以 CPU 适合控制复杂，而计算密度不高的应用场景。X86 CPU 在数据中心和云计算领域具有统治地位，而 ARM CPU 由于其低功耗、低成本的特点占据绝大部分终端市场。另一方面，随着国内在推动国产化服务器的布局，ARM 服务器也已经逐渐进入数据中心作为异构算力的组成部分。

GPU 采用的是 SIMD/SIMT 架构，虽然本质上还是冯诺依曼架构，但减少了取值、译码开销，GPU 同样具有很强的通用性，以 NVIDIA 为代表的 GPU 厂商，培

养了非常好的 GPU 生态系统，为用户提供了非常方便的开发环境，所以 GPU 在高性能计算、图像处理和 AI 领域都有非常广泛的应用。

FPGA 为现场可编程门电路，可以认为是细粒度可重构芯片，FPGA 非常适合数据流驱动的计算架构，具有高空间并发和低时延的特点，理论上 FPGA 可以实现任意功能，但 FPGA 开发周期比较长，同时对开发人员的专业技能要求也比较高，这些都影响了 FPGA 的应用场景。

ASIC 是一种专用芯片，与传统的通用芯片有一定的差异，是为了某种特定的需求而专门定制芯片。ASIC 芯片的计算能力和计算效率都可以根据算法需要进行定制，所以 ASIC 与通用芯片相比，具有以下几个方面的优越性：体积小、功耗低、计算性能高、计算效率高、芯片出货量越大成本越低。但是缺点也很明显，只能针对特定的某个或某几个应用场景，一旦算法和流程变更可能导致 ASIC 无法满足业务需求。

在摩尔定律和登纳德缩放定律的共同作用下，引领芯片行业 30 年的飞速发展，随着摩尔定律放缓，登纳德缩放定律失效，通用 CPU 的算力提升已经无法满足当前数据中心对算力增长的需求。

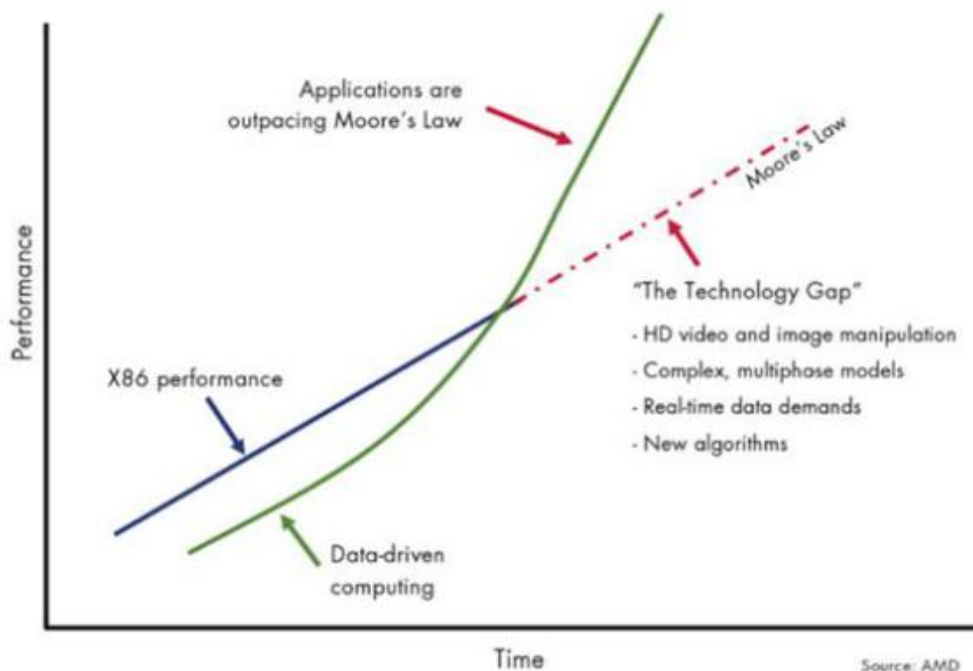


图 1-3 异构算力不同应用场景

基于领域定制（DSA）的异构计算，采用各种加速器来完成对特点业务的加速，从而达到提升系统算力和降低成本的目的，最典型的深度就是深度学习加速器，

不管是采用 GPU、FPGA 还是各类 NPU，相对于单纯 CPU 的解决方案，可以将系统的算力提升数倍。

1.3 异构算力的主要应用场景

算力的需求促进了算力架构的多样性和算力性能的不断提升，在不同的应用场景中，异构算力协同处理发挥最大化的计算效力。围绕以“云、边、端”为主体架构的三级算力调度需要满足不同场景下的算力需求：云端主要面向以高性能计算为主的传统集约化的性能计算，主要处理大流量、高并发的数据处理场景；边缘侧的数据中心主要考虑用户的快速接入和内容推送为用户提供快速处理和及时响应；而端侧的算力应用主要面向物联网、工业互联网为主要场景的嵌入式设备能够长期稳定运行，要求有低功耗和多连接等算力要求的场景。而围绕人工智能的算力需求场景是近年来新兴发展起来的应用场景，其算力需求是贯穿云端到边缘再到端侧的泛在算力需求，因此，在整个人工智能产业环境中其算力需求也是呈现异构性的特点，而且人工智能对于算力的需求也是目前推动异构算力发展的主要因素。

（1）算力与高性能计算

高性能计算 HPC，主要应用在气象、地震、勘探等科学计算领域，一般是建设一定规模的计算集群通过高速网络互联。高性能计算项目中的算力一般使用 X86 CPU 和中高端 GPU，近期 ARM 处理器的集群逐渐增多，在某些专用领域也可使用 FPGA 和 ASIC 来加速计算。

另外一种 HPC 的方案是分布式计算，利用分散的计算机和其它终端的闲置处理能力来解决大型计算问题，如生物病理研究、药物研究、寻找地外文明的信号等项目。主要利用的是 X86 CPU、消费级的显卡等算力资源。

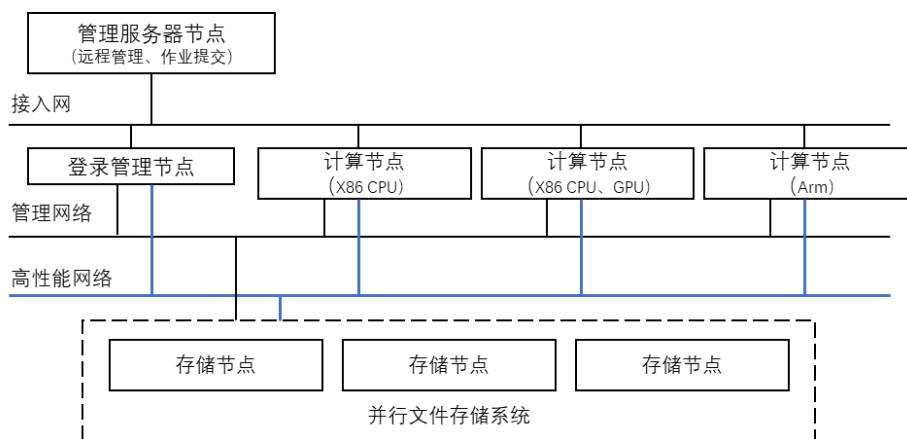


图 1-4 异构算力协同实现 HPC

(2) 算力与物联网

物联网（IOT）通过各种信息传感器设备，实时采集任何需要监控、连接、互动的物体或过程，采集其声、光、热等各种需要的信息，通过各类网络接入，实现物与物、物与人的泛在连接，实现对物品和过程的智能化感知、识别和管理。物联网的应用场景包括工业物联网、家庭物联网、车联网等各种应用场景。

物联网网关通过支持各种协议将数据收集、转换、传送到云端进行处理，不同应用场景对物联网网关的性能要求差别较大，传统的物联网网关采用配置低功耗 X86 或者 ARM 处理器的小型计算设备。物联网网关的趋势是承担更多的数据处理的任务，对算力的要求越来越高。

IOT 从传感器收集数据，网关做初步计算分析，在边缘云处做本地分析和处理，然后汇聚到数据中心，进行大数据处理和分析。

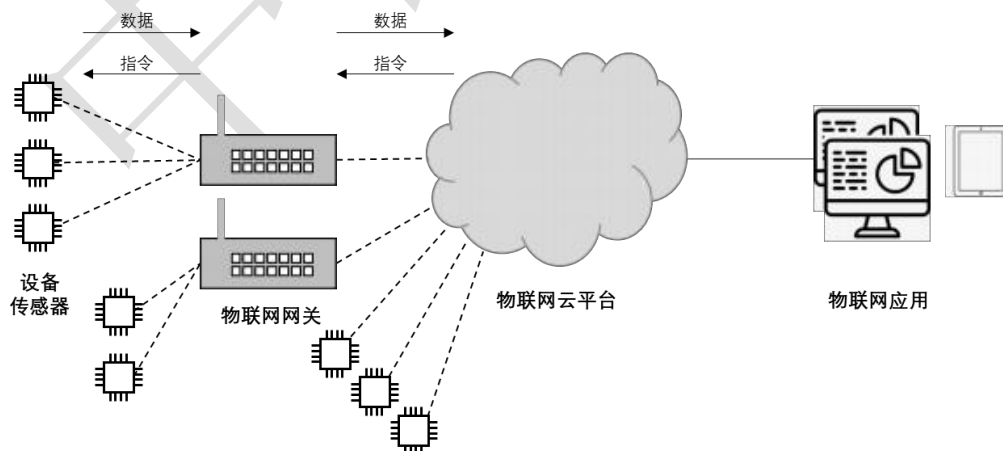


图 1-5 异构算力与物联网

(3) 算力与边缘计算

边缘计算和 5G 的大带宽、低延时相辅相成，边缘计算是在网络边缘为应用开发者和内容服务商提供所需的云端计算功能、互联网技术服务环境等。边缘计算在靠近数据源、终端设备的位置对数据进行处理。边缘计算对算力的要求较高，在视频和图像识别、语音识别等场景需要较大算力，需配置 SoC、GPU、ASIC 或 FPGA 等加速处理芯片。边缘计算所需算力分布广泛，可分布于智能装备、工业控制器、传感器、ICT 融合网关和边缘云等处。

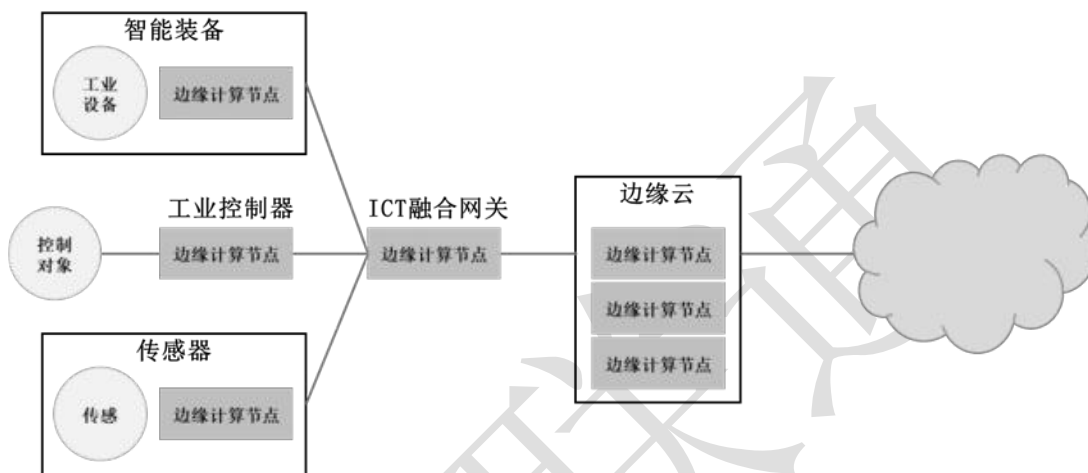


图 1-6 异构边缘计算节点协同

(4) 算力与 AI

人工智能研究的主要目标是使机器能够胜任一些通常需要人类智能才能完成的复杂工作。人工智能的研究领域包括机器人、语言识别、图像识别、自然语言处理和专家系统等。人工智能的三要素是算力、算法和数据。人工智能的算力主要是 GPU、ASIC、FPGA、TPU 等。人工智能在不同层面使用不同算力水平的 GPU，如手机终端上的 GPU、边缘计算中的推理 GPU、数据中心的训练 GPU。

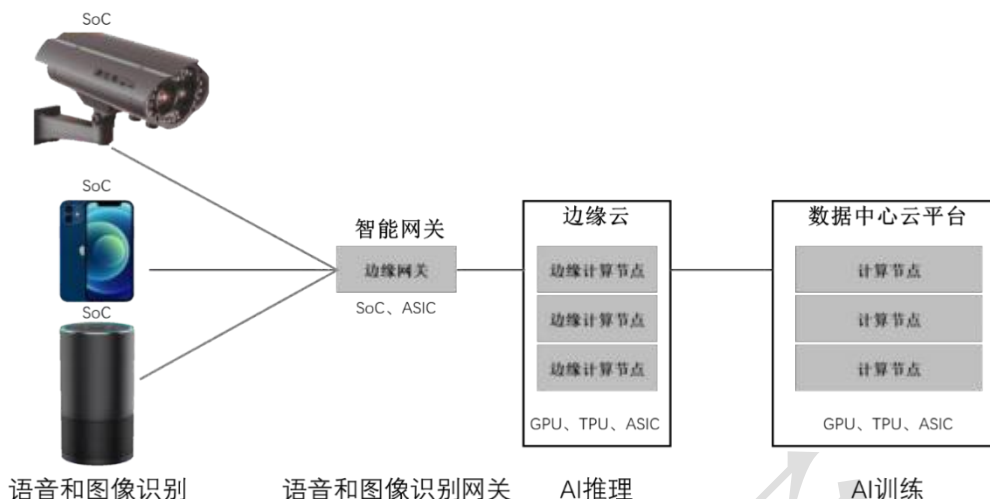


图 1-7 异构算力应用人工智能

2 异构算力发展需求

2.1 异构算力需要统一标识和调度

异构算力技术的发展能够满足不同应用场景的需求。通过异构算力的协同能够最大化的实现异构算力的效力。而若将底层算力作为资源进行有效的管理，首先需要建立异构算力的统一标识。异构算力的统一标识体系的建立是将底层异构算力作为资源服务进一步开放和共享的基础。算力网络的构建打破了原有的围绕数据中心内部算力资源实现共享的围墙，构建了基于新型网络连接为基础的异构算力接入的分布式计算形态，通过建立异构算力统一标识和网络标识的映射关系，可以进一步在网络中实现算力资源调度，并且也为算力交易提供底层的技术基础。

另外，在面向异构算力统一调度方面，基于“云原生+轻量化云原生”两级调度能力，一方面能够实现“云、边、端”底层资源的统一纳管和协同；另一方面，通过面向上层 PaaS 能力的下沉，建立算法能力库、计算能力库、存储能力库、网络能力库等异构算力能力库，从而实现异构算力服务的共享，用户也可以基于能力库快速实现业务代码的开发和部署。

2.2 异构算力统一标识设计原则

异构算力的统一标识是实现在算力网络连接范围内异构芯片的统一标识。实

现异厂家和多数数据中心算力资源接入算力网络内部以期能够共同打造算力流通，算力溯源以及确保数据在可信算力环境中可管可控提供切实有效的保障机制。

建立异构算力统一标识，其出发点主要基于以下三个方面：

- 可信算力：基于统一标识，实现算力可信接入、安全认证和鉴权，打造算力安全有效的计算环境，切实保障算力网络安全有效的生态环境；
- 算力溯源：基于统一标识，可以实现算力溯源，实现算力在云、边、端侧的溯源和安全等级；
- 算力度量：基于统一标识，构建算力度量，实现算力在分场景下的算力匹配和专家推荐，从而实现算力在联盟范围内的流通。

2.3 异构算力统一调度设计原则

为适应不同的应用场景，不同的负载特征，不同的负载组合，统一调度层应具备高度的抽象性、灵活性、可扩展性和实时性。统一调度层由一套统一调度系统框架和内置的调度策略组成。统一调度系统框架主要面向调度框架开发者、调度策略开发者及调度器使用者等群体，其设计目标具有以下特征：

- 抽象性

在异构性应用场景下，统一调度框架的北向接口对异构算力资源的申请、响应、控制和查询交互等具有高度的抽象性，以统一各应用场景对资源的需求描述，易于适配不同部门和组织的用户系统。

- 灵活性

异构算力的资源属性天然决定了作业资源调度策略的多样性，统一调度异构算力框架的调度流程及策略应具有可自主选择、可细粒度定制的特征，灵活满足不同部门组织、用户及应用场景的个性化资源调度需求，达到最优的资源分配效率。

- 可扩展性

统一调度框架对新应用、新算力应具有生态友好、开放等特征，开发者易于进行定制化二次开发以扩展作业排队、选择及资源调度策略，并且可支持企业第三方信息系统、云平台等系统与统一调度框架对接适配，达到资源最优配置及信息互通。

- 实时性

统一调度框架通过获取应用对资源的实时消耗、异构算力的硬件拓扑变化和软硬件故障等反映可获得的调度资源，资源更新的实时性直接决定了异构算力的资源利用效率。

3 异构算力统一标识和资源抽象

基于目前异构算力技术的快速发展以及在不同场景下的广泛应用，中国联通在打造新型网络连接实现数据中心内部以及跨数据中心的互联的同时，也在考虑围绕“云、边、端”的泛在计算场景下实现异构算力的统一标识和资源抽象，从而更好实现算力的统一管理，便于上层应用开发者能够调用相关的算力资源。

3.1 异构算力统一标识

3.1.1 异构算力统一标识技术架构

围绕上述异构算力统一标识的需求和场景，中国联通面向目前主流的算力架构提出了异构算力统一标识和认证体系架构，从而保障在一个算力联盟内部接入的算力资源提供切实可信的计算环境和算力溯源机制，从而让上层用户应用程序能够在保证数据安全、稳定、有效的计算环境中正常运行和处理；同时，结合不同应用场景提供算力度量机制，为算力流通提供可度量、可推荐的算力评价机制。

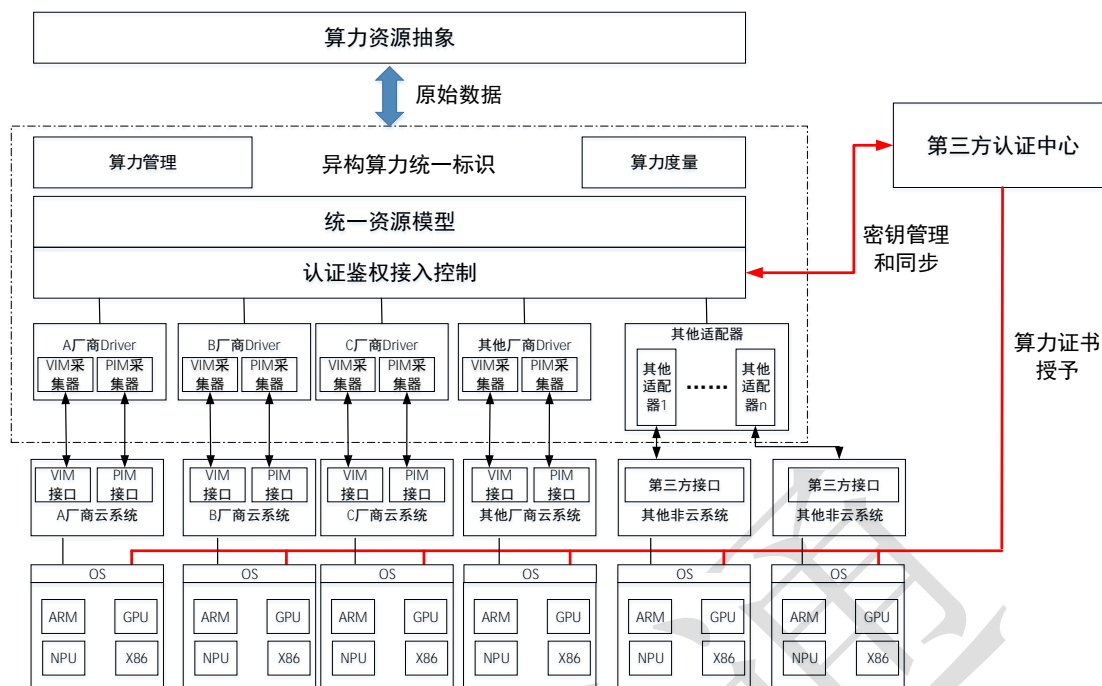


图 3-1 多厂家异构算力统一认证接入架构图

依据上述架构图，异构算力统一标识在原有多云接入方式下，通过各家厂商的云系统的 PIM 接口上报异构算力的具体信息，对于非云系统采用其他适配层的方式纳管算力资源。

在算力接入过程中需要通过第三方认证中心进行算力认证并且发放签名证书并通过操作系统保存在可信区中。当算力经过自有云系统或者第三方接口上报至统一资源模型时，需要凭借授权的证书进行认证鉴权，只有经过认证合法的算力才能够作为可信算力纳入到算力管理中，并且经过统一资源模型成为算力资源。

为了给上层用户提供安全可靠的异构算力使用生态环境，包括中国联通在内的整个行业实现异构算力公平可靠的流通，数据能够在可管可控的可信计算环境下安全有效的进行处理。需要具有公立的第三方安全可靠的鉴权认证中心实现可信算力的认证鉴权，并且为了能够降低算力使用者的使用门槛，需要构建统一的资源模型，进行有效的管理，并且能够建立标准的算力度量机制，从而保障算力公平、可靠的流通。

3.1.2 可信算力的认证鉴权

基于第三方的算力认证中心，在面向可信算力的认证方面具有独立的认证体系。从而能够实现算力本身的类型识别、算力来源合法性验证以及算力等级的度

量。该认证中心主要对于接入到算力联盟的异构算力进行独立认证，并且为异构算力签发证书。当异构算力系统在接收到认证中心签发的证书时，并将其通过操作系统保存在可信区域内。当底层算力在通过算力提供方的云资源或者非云系统的第三方接口接入到算力统一标识系统中时，统一标识系统首先需要进行认证鉴权。当认证其算力具有合法性时，则允许其算力资源有效接入到算力资源标识中以便于后续的资源抽象和统一调度。并且通过认证鉴权以后，在算力管理模块中为异构算力生成唯一的算力标识符。而该算力标识符作为统一资源模型的唯一索引关键值，能够匹配异构算力在整个算力网络异构资源池的唯一性和可溯源性。

3.1.3 统一资源数据模型

面向异构算力基于底层适配层的接入，中国联通算力网络基于自定义资源模型，提供统一的异构算力资源抽象模型。其优势在于：

- 模型通用性：支持异构算力资源的统一标识，基本涵盖了现有数据中心适用的异构算力资源，典型算力架构包括：CPU、NPU、GPU、FPGA 等；
- 定义灵活：支持资源扩展属于设置，如 NPU 等复杂资源，既可以将算力拓扑结构、利用率、告警等信息作为算力厂商资源的扩展属性，也可以将 SOC 内的内存等信息作为独立资源；
- 简化接口模型，无需编程：通过客户端、API 等即可实现自定义资源的扩展和使能。

基于现有的算力采集上报和多厂商异构算力接入适配层的统一数据模型，将上报的算力资源数据模型划分为九个数据域。其总体的数据模型架构如下表所示：

表 3-1 统一资源模型

属性	类型	必选	说明
资源名	String	是	资源名集群唯一
资源描述	String	是	对资源的描述
是否是动态资源	Boolean	是	定义资源类型，默认为 xx
是否可消费	Boolean	是	定义资源类型，默认为 xx
是否可共享	Boolean	是	定义资源类型，默认为 xx
资源值	Numeric	否	静态资源必须指定
资源获取命令	String	否	动态资源必须指定，输出格式为“ [资源值] [资源扩展属性] ”，[资源值]必须输出，[资源扩展属性]可选输出
资源采集时间	Integer	否	动态资源必须指定，默认为 xx

隔			
使用范围	String	否	1、如果是独享资源，指资源持有的对象，如 node1、node2 2、如果是共享资源，指资源的共享范围

基于上述自定义扩展统一资源模型，在多厂家异构算力接入过程中，实现统一的算力资源上报和数据采集，并且通过统一的数据模型建立算力资源统一标识，为后续的算力管理和算力度量提供数据。

同时统一资源模型的建立，也为后续的算力资源的抽象提供了技术支撑。随着目前在云计算和云原生的资源调度扩展接口的实现方式不同，需要对于算力资源节点的纳管和注册实现统一的管理，并且通过上层资源抽象为用户提供无差异的开发环境。

3.1.4 异构算力资源管理

算力资源管理面向异构算力建立统一的算力纳管。为 CPU、NPU、GPU 等异构性算力建立统一的算力管理体系。基于算力统一标识，建立算力和标识的索引和厂商归属关系，为算力接入厂商建立认证鉴权证书私钥，并且通过第三方认证机制为接入厂商提供合法的认证证书。

算力管理构建算力联盟范围内的统一算力类型和算力使用情况统计信息。通过统一算力标识来维护和管理算力和厂商的唯一性，接入时间、运行情况以及接入地点等算力的上架、算力运行以及算力增加等全生命周期的管理信息。

算力管理提供算力在整个使用过程中的数据统计分析以及为算力的应用场景的分析提供必要的的数据支撑和必要的运行维护数据，为后续的算力推荐和算力度量提供必要的的数据支持。

3.1.5 算力度量与算力推荐系统

不同的应用场景所需的算力需求的类型和量级也是不同的。如非实时非移动的 AI 训练类业务，这类业务训练数据庞大，神经网络算法层数复杂，若想快速达到训练效果，则需计算能力和存储能力都极高的运行平台或设备；对于实时性类的推理业务，一般要求网络低时延，而对计算能力的需求则可降低几个量级。因此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的依据。

算力是设备/平台处理、运行业务的关键核心能力，根据所运行算法的不同，涉及的数据计算类型不同，可以将算力分为逻辑运算能力，并行计算能力和神经网络计算能力。

- **逻辑运算能力：**这种计算能力是一种通用的基础运算能力。硬件芯片代表 CPU。CPU 的结构主要包括运算器（ALU, Arithmetic and Logic Unit）、控制单元（CU, Control Unit）、寄存器（Register）、高速缓存器（Cache）和它们之间通讯的数据、控制及状态的总线。这类芯片需要大量的空间去放置存储单元和控制单元，相比之下计算单元只占据了很小的一部分，所以它在大规模并行计算能力上极受限制，而更擅长于逻辑控制。度量单位一般的可以用 TOPS 来衡量其运算能力。

- **并行计算能力：**并行计算能力是指专门为了处理如图形图像等数据类型统一的一种高效计算能力。该类型的计算能力特别适合处理大量的类型统一的数据，不但在图形图像处理领域大显身手，同时也适合科学计算、密码破解、数值分析，海量数据处理（排序，Map-Reduce 等），金融分析等需要并行计算的领域。目前并行计算能力也是一种比较通用的计算能力。典型的硬件芯片代表如 GPU。从架构来看，GPU 的构成相对简单，有数量众多的计算单元和超长的流水线。度量单位：常用浮点运算能力来衡量。常用单位 TFLOPS/s 可以简单写为 T/s，意思是“1 万亿次浮点指令/每秒”。同时还有 MFLOPS（megaFLOPS），GFLOPS（gigaFLOPS），PFLOPS（petaFLOPS）。

- **神经网络计算能力：**神经网络计算能力主要针对近年来 AI 神经网络、机器学习类密集计算型业务，是一种用来对机器学习、神经网络等进行加速的计算能力。近年来各类发布的 AI 类芯片皆为加速神经网络计算设计。如华为提出的 NPU, Google 提出的 TPU。另外，机器学习、神经网络的本质是密集计算，如 Google 工程师举的例子 — 如果人们每天用三分之一的语音搜索，但运行没有 TPU 加持的语音识别人物的话，该公司将需要建造两倍多的数据中心。

算力网络将算力作为服务提供给用户，算力交易平台是一套基于区块链的去中心化、低成本、保护隐私的可信算力交易平台。该平台的计算节点由多种形态

的算力设备组成，包含大型 GPU 设备或者 FPGA 服务器集群或者中小型企业闲散的空余服务器以及个人闲置的计算节点等。该平台能承载海量计算产品底层算力需求，推动整个算力交易行业的发展。

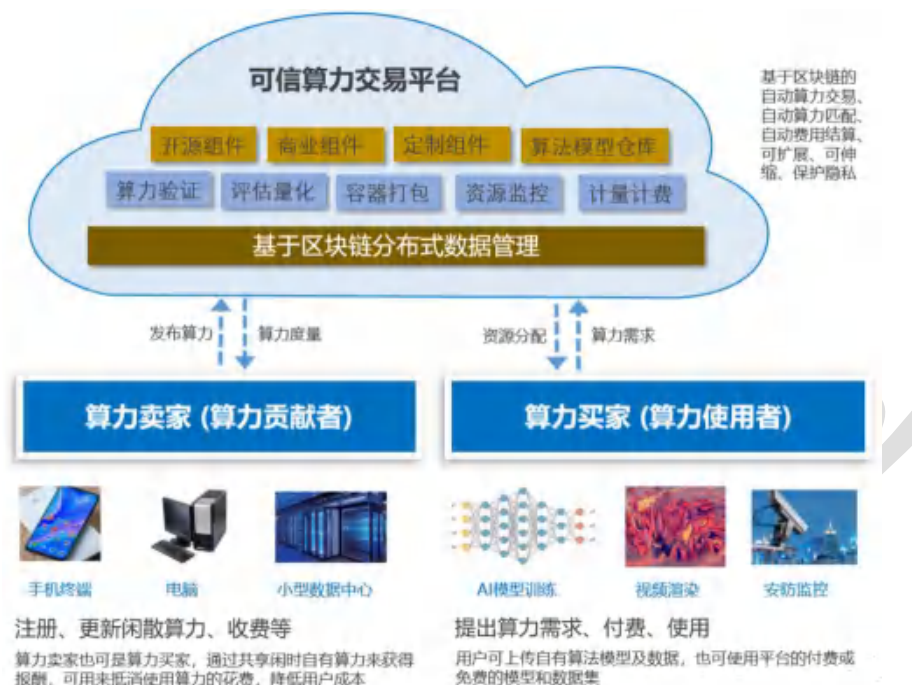


图 3-2：可信泛在算力交易平台示意图

算力交易过程中的有三种角色，如下图所示：1) 算力卖家：贡献闲散算力，如手机终端、电脑、游戏机、企业空闲时的小型数据中心等，卖家将零散算力通过注册或更新的方式告知算力交易平台；2) 算力买家：向算力平台提出算力需求，获得算力平台分配的容器节点，并进行付费，执行任务等操作；3) 算力交易平台：算力交易平台作为中间角色分别于算力的卖家和买家打交道，维护、纳管、调度算力资源，提供经济、高效、安全、可靠的算力服务。

3.2 异构算力资源抽象

异构算力资源的抽象（SDR）是为上层应用进一步屏蔽底层算力的差异化，使得用户更关注上层业务代码的开发，而无需关注底层差异性资源的申请和调度等方面。通过统一的接口和客户端来实现异构算力资源的抽象。在现有的云计算和云原生等开源社区中，纷纷基于扩展的异构资源接口来实现异构算力资源抽象。

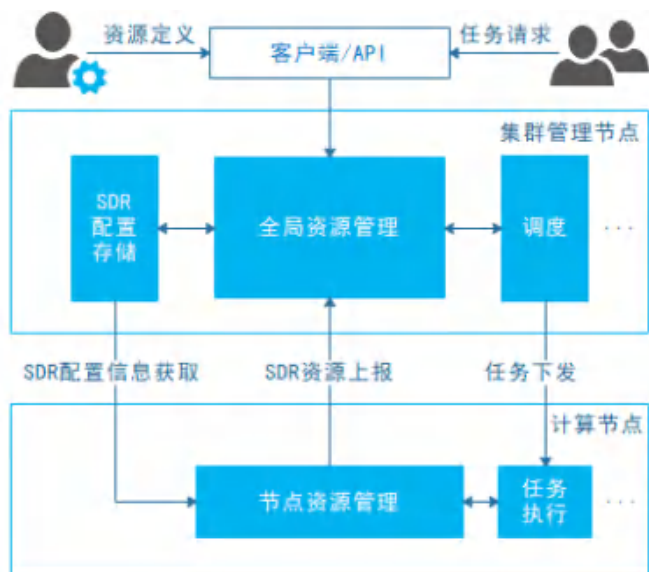


图 3-3 异构算力资源抽象方案示意图

基于全局资源管理定义算力资源抽象，并且通过 API 等方式提供给上层应用。上层应用通过任务请求等方式来实现调用，并且通过算力集群调度的方式将任务请求发下到指定的算力节点上进行任务执行。

采用异构算力资源抽象能够进一步屏蔽掉底层算力的差异性，并且上层开发者无需关心算力具体部署在哪个集群的节点上。而对于新增加的算力类型能够实时的同步更新到上层开发环境中，从而缩短了新算力上线到用户应用之间的使用周期，可以更好的为用户服务。

3.3 异构算力统一调度与适配

3.3.1 异构算力统一调度技术架构

异构算力的统一调度系统框架，按照分层结构，其总体技术架构如下图所示：

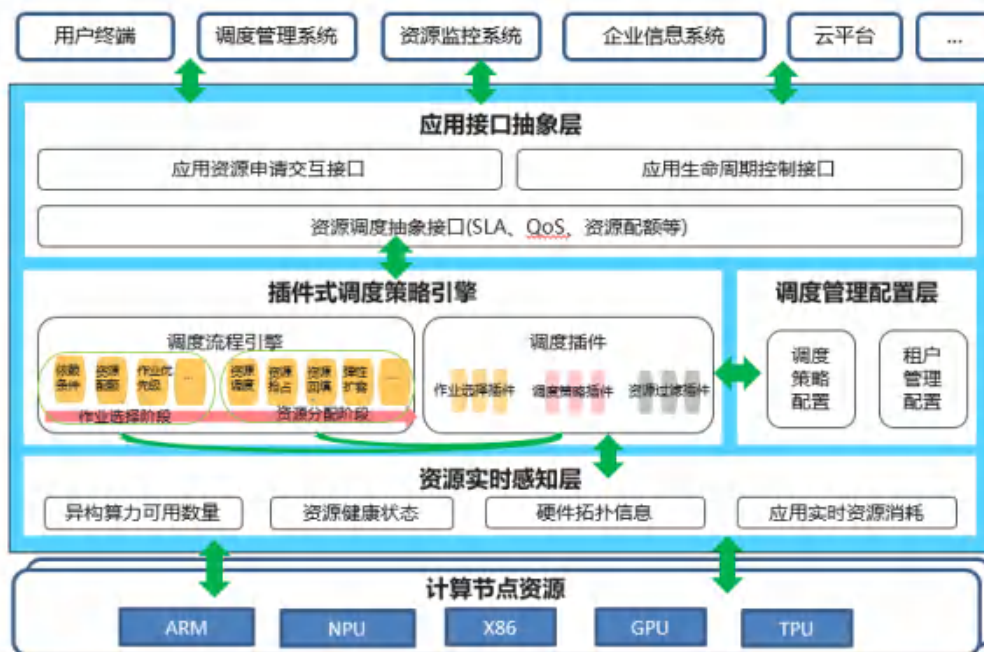


图 3-4 统一调度系统技术架构图

面向异构计算节点资源实现统一的资源实时感知，并且对上层提供资源抽象和应用调度抽象，以满足上层多场景需求的使用，具体如下：

- 应用接口抽象层

该层衔接应用业务对资源的需求与系统资源的统一调度，以用户友好的应用资源需求表示方式和交互界面屏蔽异构算力的资源调度复杂性，实现调度器使用者与统一调度系统之间对作业的联动控制和实时反馈。

- 插件式调度策略引擎

该引擎根据应用的资源需求匹配最佳的异构算力资源分配，作业调度流程和策略具有高度模块化、灵活组合、可插件式扩展等能力。一方面满足调度器管理者对系统资源分配目标的统一控制，如控制不同场景任务的资源分配额度、优化系统特定维度资源的使用效率等；另一方面满足调度器使用者的应用个性化资源和业务模型需求，如控制多个应用间的运行依赖行为、资源竞争关系等。

应用获取可使用资源的调度过程主要分为作业排队和资源配置两个阶段，其中作业排队用于决定当前优先获得调度资格的作业序列顺序，影响作业调度顺序的因素多种多样，如当前用户或所在组织的资源配额限制、作业之间的启动依赖关系、作业所在排队序列的资源分配策略及可调度份额等。

作业排队阶段选择待调度作业后，由资源配置阶段决定最终作业的最佳运行

节点及其异构算力资源分配。资源分配的复杂度受系统的可选择节点数量、异构算力的资源维度、作业的资源需求复杂度等多种因素影响。资源分配的决策等待时间直接影响到作业启动速度，进而影响系统的整体资源利用率。调度策略层应具有过滤策略的插件式定制能力，以降低作业匹配最佳资源的计算开销，提升统一调度框架的决策实时性。

- **调度策略配置层**

提供对插件式调度策略的控制面管理能力，根据调度器系统所服务的应用场景配置作业调度各阶段的策略启用行为，满足目标应用的异构算力资源调度需求；提供对租户面的资源配额、运行限制等多维度层次化管理配置能力。

- **资源实时感知层**

该层实时收集系统内各节点的异构算力资源数量，感知硬件拓扑及运行健康变化，反馈到调度引擎用于匹配作业的资源需求。资源信息感知的反馈及时性对调度引擎的决策实时性和准确性至关重要。

3.3.2 算力调度编排平台功能

为了构建可运营、可开放的算力网络资源调度和编排环境，构建面向上层的算力网络编排调度平台按照业务功能划分为：资源层、调度层、编排层、应用层、自服务运营门户、开发者门户和监控层几个模块。

(1) **资源层**：主要面向底层的计算、存储、网络等资源进行统一的纳管，这其中包括裸金属的管理，同时也包括虚拟机、容器、边缘集群等基础设施资源等；

(2) **调度层**：通过底层资源的抽象，在调度层主要专注于基于调度策略实现对于算力资源、网络资源以及存储资源的调度，同时为了实现平台能力下沉，在调度层实现三大能力集，即：计算能力集、网络控制能力集和机器学习能力集；

(3) **编排层**：将调度层的能力以服务化的方式提供服务注册、服务发现、服务路由等功能，并且按照最新的服务网格的方式提供扁平化的服务编排方式；

(4) **应用层**：基于容器化调度机制面向应用提供统一的封装和事务性功能，包括：应用商店、统一的镜像仓库、鉴权管理、函数服务以及其他事务性处理等功能；

(5) **自服务门户**：为普通用户提供一站式的应用服务选择和一键部署等算

力网络能力部署能力，以及通过用户登录和注册分配相关的用户权限；

(6) **开发者门户**：为开发者提供算力网络开放的 API 调用接口和 Devops 开发流程，从而可以尽可能开放算力网络底层能力，从而为第三方开发者可以结合场景开发专属的应用产品功能；

(7) **监控层**：负责监控整个算力网络编排调度平台的运行情况，实现资源、服务、事务、安全等方面的监控，确保整体平台的安全、平稳运行。

3.3.3 统一调度核心流程

满足多租户场景下的算力配额、多用户优先级（作业选择策略框架）、多任务 CPU、NPU、GPU、ARM、TPU（资源调度策略框架）算力需求。

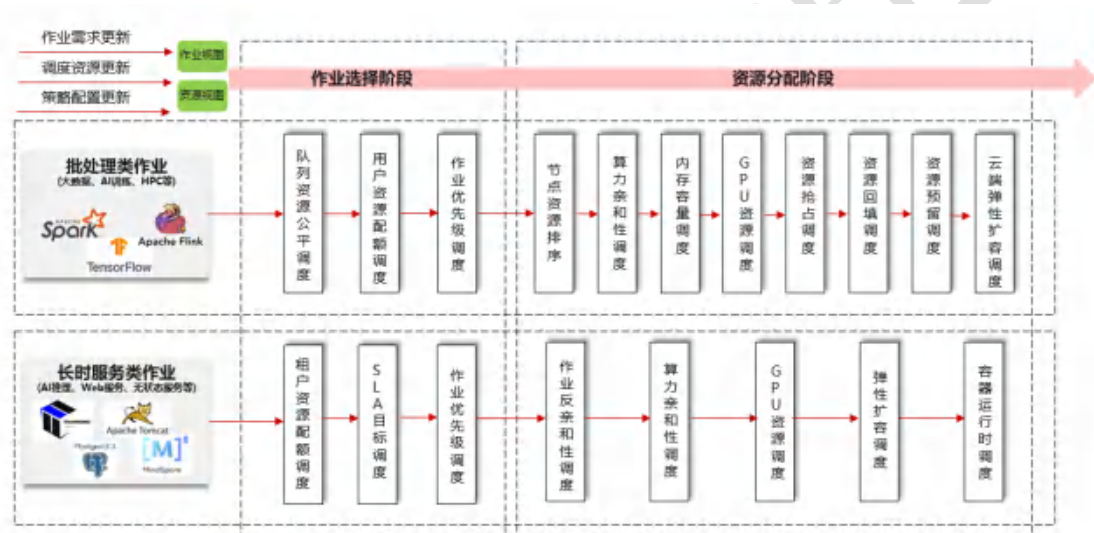


图 3-5 统一调度核心流程图

如上图 3-5 所示，作业调度流程主要分为作业选择和资源分配两个阶段，其中作业选择阶段决定了作业调度的优先级，资源分配阶段决定了作业需求资源的最终分配方式。以批处理类作业和长时服务类作业为例，不同负载类型对应不同的核心调度流程和策略：

- 队列资源公平调度

基于不同资源池、不同资源类型的使用优先级，划分不同的作业资源请求队列，依据队列使用资源情况，对作业进行优先级排序，保证队列资源使用的公平性。

- 用户资源配额调度

基于集群多用户场景，以树状形式描述不同组织、用户的资源使用策略，依据使用资源情况，对作业进行优先级排序，保证多用户间资源使用的公平性。

- 作业优先级调度

依据用户作业递交时指定的作业优先级、递交时间、资源请求量、用户资源使用量等维度综合考虑进行排序，优先级更高的作业优先进行调度。

- 节点资源排序

不同的场景需要采取不同的节点资源排序策略，如在集群负载均衡场景下，优先选择资源使用量最少的节点；在集群节能场景下，优先选择未休眠的节点，避免不必要的唤醒操作和能源消耗。

- 算力亲和性调度

不同作业负载存在多样化的算力需求，如 AI 推理作业对于张量计算需求高，亲和 TPU、NPU 等计算资源、图像处理作业对于矢量计算需求高，亲和 GPU 等计算资源。通过算力亲和调度，实现多样化算力需求任务的最优资源匹配分发，可实现算力资源的最大化利用。

- 内存容量、GPU 资源调度

异构算力资源调度，此外还支持自定义扩展资源调度。

- 资源抢占调度

统一调度支持用户、队列间的资源借用，用以提高集群资源利用率，同时涉及高优先级作业针对被借用资源的紧急回收问题，即资源抢占调度。

- 资源预留调度

通过资源预留的方式，解决大作业和小作业混合负载的场景下，小作业资源请求频繁，导致大作业资源需求持续无法被满足的问题。

- 资源回填调度

资源回填调度是基于资源预留调度的前提下，通过感知作业结束时间，将已预留但又暂时无法满足作业需求的资源，在一定时间范围内进行重新调度分发，提高集群资源利用率。

- 云端弹性扩容调度

集群负载通常不均衡，具有突发激增性特点。统一调度支持本地集群与云端资源进行对接，在高负载场景下，可将本地作业分发到云端执行，以降低本

地集群运营成本。

- SLA 目标调度

SLA (Service Level Agreement) 服务等级协议指服务提供商与受服务用户之间具体达成承诺的服务指标。调度器支持根据具体的 SLA 目标如作业调度平均等待时间、吞吐量、平均故障恢复时间、服务响应时间等，进行作业排序选择。

- 作业反亲和性调度

长时服务作业通常会占据部分计算资源并保持长期运行，并且不同的长时服务作业存在异构资源需求，如果分配到相同计算节点，存在资源竞争的可能，如网络带宽、IO 带宽。通过反亲和调度可以将存在资源竞争的任务隔离分发，实现作业性能的提升。

- 弹性扩容调度

多实例、实例数量动态变化是长时服务的典型特征之一，支持弹性扩容调度策略，典型应用场景如用户自定义弹性伸缩策略，通过算力负载压力感知，动态伸缩实例数量，保证服务质量。

- 容器运行时调度

长时服务考虑到持续集成、可移植性、隔离性等原因，通常以容器负载形态进行呈现，统一调度支持容器运行时调度。

3.3.4 异构算力调度智能运营

面向异构算力的统一调度的过程中，在实现异构算力的统一纳管的同时，面向智能化运维体系的建立将会变得尤为重要。积极构建面向算力网络的异构算力统一调度智能化运营体系需要从以下三个方面来考虑：

- 异构算力的统一监控

面向“云、边、端”的泛在算力统一调度过程需要统一的运营监控体系来实现，统一的监控系统不仅仅能够实现算力资源的统一运行情况统一管理，并且也为算力的调度和扩容提供技术保障；

- 异构算力的统一适配

面向异厂家的计算单元以及多云资源的统一接入和纳管，需要构建统一的适

配来实现算力联盟内异构算力的接入，并且通过统一适配层进一步屏蔽掉底层算力的差异化，为上层提供统一的调度环境。而随着多厂家接入的增加，为了满足异厂家算力快速接入的需要，统一适配层具备满足灵活的横向扩展能力；

● 异构算力溯源

面向算力处理的数据需求下，大多数场景需要保证数据能够在可信的计算环境下进行，从而算力的溯源一方面保障用户的数据能够在可掌握的环境中处理和运行；另一方面当计算过程中出现异常，包括数据泄露等情况时，能够实时追溯到算力来源进行确认。

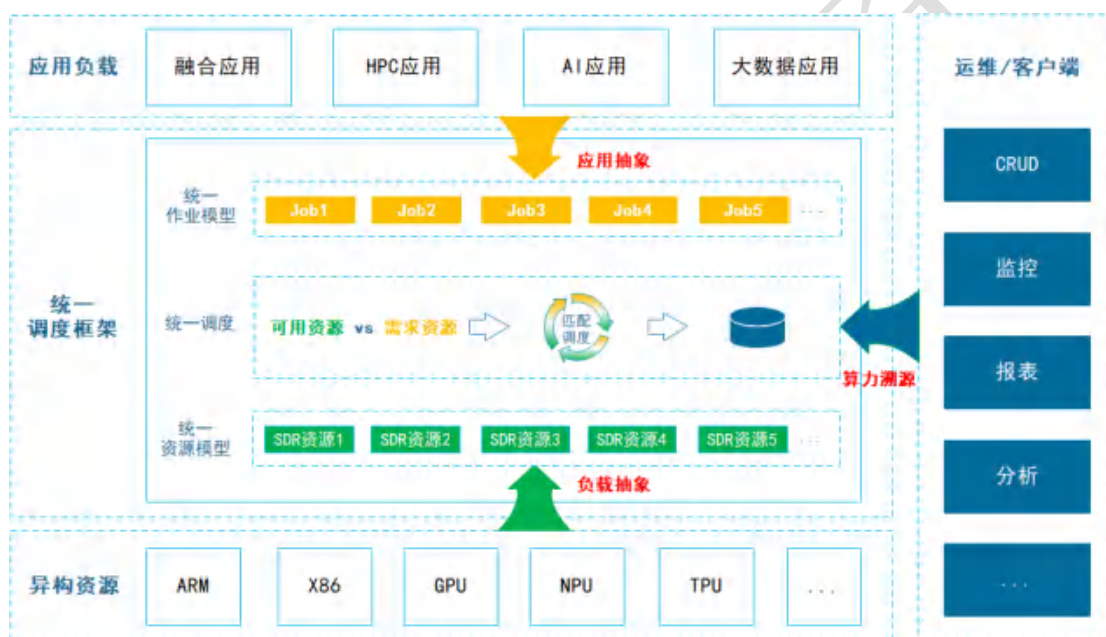


图 3-6 异构算力调度智能运营系统框图

如图 3-6 中国联通异构算力调度智能运营系统框图所示：

● 应用负载

支持多样化应用负载，包括 HPC 应用（MPI 并行作业、Array 作业、WorkFlow 作业、交互式作业等）、大数据应用（Spark、Flink 等）、AI 应用（TensorFlow、MindSpore 等）、Web 应用（Tomcat、Nginx 等）以及新兴的多场景业务融合应用等。

● 统一调度框架

统一调度框架通过负载适配层和资源适配层分别对多样化应用负载和多样化异构资源进行抽象，并且通过统一调度层实现需求面和资源面的管控、匹配、决策生成、决策下发、过程数据持久化等，为算力溯源提供数据支持。

- **异构资源**

支持多样化异厂家算力资源，可涵盖现有数据中心适用的异构算力资源，典型算力架构包括：CPU、GPU、NPU、TPU、FPGA 等。

- **统一运维**

统一运维旨在实现算力资源运行情况的统一管理以及外部系统的对接，支持过程数据 CRUD（增、删、改、查），接入算力统一监控，报表数据导出，数据分析策略集成、对接等功能。

4 总结与展望

国民经济数字化转型的稳步推进，使得数据中心的建设需求进一步扩大。2020 年 12 月 28 日，国家发改委联合工信部等部委围绕“优化数据中心建设布局，推动算力、算法、数据、应用资源集约化和服务化创新”发布了《关于加快构建全国一体化大数据中心协同创新体系的指导意见》，其中将算力资源作为核心的生产要素在整个国民经济和数据中心向更高效和智能化的方向转化发挥重要作用。

4.1 异构算力的融合发展

在市场需求的驱动和推动下，算力的发展一方面呈现多样性，打破传统的 X86 单一架构的算力形态，实现了异构算力以应对不同场景下的数据处理应用；另一方面又呈现出异构算力下的能力开放和统一管理，从算力生态的需求来看，不论是芯片厂商还是平台厂商目前都围绕自身的产品系统能够将底层的异构算力能力进行协同和融合，能够吸引更多的产业链上下游企业能够共同打造生态环境。因此彼此都是相向而行，逐步向融合方向进行推进。

4.2 算力能力进一步开放

在算力能力开放方面，为了推动算力在最大化的应用场景中发挥作用，同时也为了进一步实现算力的流通和开放，目前已经在行业内普遍达成共识。并且在能力开放方面主要分为三个层面的能力开放：

- **算力芯片能力开放**，芯片厂商进一步通过 SDK、BSP 等方面能够实现硬件

芯片能力的开放；

- **算力平台能力开放**，不论是平台厂商还是芯片厂商实现异构算力的接入，通过云计算或者云原生等方式实现算力虚拟化以实现算力资源共享，并且为了吸引更多的上层应用开发和部署，会构建自身平台级的能力开放平台以期能够实现更多应用的孵化；
- **算力交易和流通机制**，一方面将算力基础能力平台下沉，抽象成共性基础平台，打造应用商店以实现算力的交易和流通；另一方面，积极建立算力的度量和评价体系，从而实现算力和应用需求的合理匹配；

4.3 算力网络促进算力合理化布局

虽然算力作为国民经济发展过程中重要的基础设施资源，得到了大力的发展和建设，但是由于缺少前期的规划、供需关系匹配以及能源、税收等建设因素等问题，数据中心建设目前还处于无序状态，从目前全国的数据中心布局来看，存在算力需求的“东密西疏”和算力产能的“热点分散”之间的不匹配。而算力的协同计算和数据的大容量传输，需要新型的网络进行连接，算力网络是近年来比较热点的研究领域，从原有的云网融合到目前的算网融合，算力网络着重以新型的网络形态能够促进算力的协同发展和合理布局。

以 CFN、确定性网络、SRv6 作为算力网络研究重点的新型网络技术和框架，也在改变传统意义上的以数据转发为第一要因的网络传输模式，结合算力感知、业务感知来实现不同的应用场景需求来调度不同的算力来进行匹配，从而使得网络更加具有智能化，而网络自身的能力开放和服务化等因素进一步作为一种算力服务能力提供给上层用户进行调用和使用，从而可以使得网络应用更具有互联网化的思维模式。

4.4 异构算力的标准化建议

异构算力的融合发展和产业生态的良好构建离不开顶层设计和标准化体系的建立，异构算力的标准化从底层算力基础设施到上层的算力商店以及可信算力环境的构建都需要通过标准化体系来保障，从而使整个产业实现良性发展。从整个算力联盟本身来看，能够实现算力的安全可靠的计算环境，以及能够实

现基于网络的算力调度和流通，更需要通过标准化的方式来实现和保障。因此，针对异构算力的标准化方面，结合算力网络结合的视角，给出如下几方面的建议：

（1）异构算力统一标识和网络标识映射解析标准

异构算力统一标识和资源抽象结合网络标识建立对应的映射关系，可以促进异构算力在网络连接中的唯一属性，将传统的以五元组为网络转发机制逐步结合算力标识和业务标识来实现算力调度转发和路由。目前国内网络 5.0 协会的团体标准体系正在开展相关网络标识的标准编制和研究工作，而在国际上中国电信和中国联通等国内运营商也在积极推进算力网络标准体系框架，其中将算力抽象成底层资源层来实现上层的调度和管理，相关标准和提案也在稳步推进。

（2）算力应用商店相关标准

算力应用商店结合现有的云原生和云计算的应用商店来实现算力应用的流通和一键化部署能力，能够进一步打通不同算力平台和云计算平台之间的算力能力之间的流通。目前围绕云技术的应用商店（Service Catalog）等通用接口，边缘计算的 MEP 平台接口互通性等纷纷开展相关标准化的研究工作，而随着未来算力应用服务能力的不断提升和多样性，更加需要促进不同平台的应用商店之间的互通，需要不断推进相关标准的研究和制定。

（3）构建可信算力网络标准体系

算力所谓算力网络中能够切实为用户提供稳定、多样的可执行环境，需要从用户的角度考虑如何为用户数据提供一个可信的计算环境以及安全的网络环境。而这种可信的算力网络环境需要由一个独立的安全保障体系来实现，目前可信计算的相关标准化研究面向异构算力环境能够实现独立的算力安全认证和算力可溯源的保障体系，从而保障用户能够将数据放在安全有效的环境中进行计算和传输，从而兼顾打破数据孤岛和保障数据安全的统一。

5 参考文献

1. 网络 5.0 技术白皮书[R]. 网络 5.0 产业和技术创新联盟, 2019.
2. 中国联通算力网络白皮书[R]. 2019.
3. 中国移动算力感知网络技术白皮书[R]. 2019.
4. 李铭轩, 曹畅, 唐雄燕, 何涛, 李建飞, 刘秋妍. 面向算力网络的边缘资源调度解决方案研究[J]. 数据与计算发展前沿, 2020, 2(04):80-91.
5. 智能计算中心规划建设指南[R]. 国家信息中心信息化和产业发展部, 2020.
6. 运营商边缘计算网络技术白皮书[R]. 边缘计算网络产业联盟 (ECNI), 2019.
7. 工业互联网标识解析标准化白皮书[R]. 工业互联网产业联盟, 2020.
8. 云端设计, 与时间赛跑 [R]. 阿里云研究中心, 2020.
9. 2020 年全球算力指数评估报告[R]. 浪潮信息&IDC, 2021.
10. 2018-2019 年中国人工智能算力发展评估报告[R]. 浪潮信息&IDC, 2021.
11. 何涛, 曹畅, 唐雄燕等. 面向 6G 需求的算力网络技术[J]. 移动通信, 2020(6).
12. ITU-T. Draft Recommendation ITU-T Y. CPN-arch: framework and architecture of computing power network [R].
13. ITU-T. Draft Recommendation ITU-T Q. CPN: Signaling requirement of computing power network [R].
14. 算力网络需求与架构 (征求意见稿) CCSA TC3 WG1 制定[R].
15. 面向业务体验的算力需求量化与建模研究 (征求意见稿) CCSA TC1 WG2 制定[R].
16. Framework of Compute First Networking (CFN) draft-li-rtgwg-cfn-framework-00, IETF [R].
17. 泛在计算服务白皮书[R]. 云计算开源产业联盟, 2020.
18. 鲲鹏计算产业发展白皮书[R]. 绿色计算产业联盟, 2019.
19. 昇腾计算产业发展白皮书[R]. 绿色计算产业联盟, 2019.
20. 服务器系列报告(计算篇)-CPU 平台[R]. 安信证券, 2020.

6 缩略语

缩略语	英文全称	中文释义
AI	Artificial Intelligence	人工智能
AR	Augmented Reality	增强现实
BBF	Broadband Forum	宽带论坛
C/U	Control/User	控制/用户
CFN	Compute First Network	计算优先网络
CO	Central Office	中心局
CPU	Central Processing Unit	中央处理器
CT	Computed Tomography	电子计算机断层扫描
CUBE	Cloud-oriented Ubiquitous-Broadband Elastic Network	面向云服务的泛在宽带弹性网
DC	Data Center	数据中心
DCI	Data Center Interconnection	数据中心互联
DCN	Data Communication Network	数据通信网络
DNS	Domain Name System	域名系统
ECC	Edge Computing Consortium	边缘计算产业联盟
ETSI	European Telecommunications Standards Institute	欧洲电信标准化协会
GPU	Graphics Processing Unit	图形处理器
GSMA	Global System for Mobile Communications Association	全球移动通信系统协会
IB	InfiniBand	无限带宽
ICT	Information and Communication Technology	信息和通信技术
IDN	Internationalized Domain Names	国际化域名
IEC	International Electrotechnical Commission	国际电工委员会
IETF	The Internet Engineering Task Force	国际互联网工程任务组
IoT	Internet of Things	物联网
IT	Internet Technology	互联网技术
ITU	International Telecommunication Union	国际电信联盟
M2M	Machine-To-Machine	机器对机器
MBB	Mobile Broadband	移动宽带业务
MEC	Multi-access Edge Computing	多接入边缘计算
MR	Mix Reality	混合现实
NFV	Network Function Virtualization	网络功能虚拟化
OTT	Over The Top	互联网公司越过运营商
P2P	Peer to Peer	对等网络
PC	Personal Computer	个人计算机
PSTN	Public Switched Telephone Network	公共交换电话网络
RAN	Radio Access Network	无线接入网
RDMA	Remote Direct Memory Access	远程直接数据存取
SDN	Software Defined Network	软件定义网络
SNAI	Sensor Network Artificial Intelligence	传感器、网络和人工智能
SRv6	Segment Routing IPv6	IPv6 分段路由

TBC	Trusted Block Chain	可信区块链
TCP	Transmission Control Protocol	传输控制协议
TLS	Transport Layer Security	传输层安全性协议
ToB	To Business	面向企业
ToC	To Customer	面向用户
URLLC	Ultra-reliable and Low-Latency Communication	超高可靠与低时延通信
VM	Virtual Machine	虚拟机
VR	Virtual Reality	虚拟现实
VXLAN	Virtual Extensible Local Area Network	虚拟扩展局域网
WLAN	Wireless Local Area Network	无线局域网