



(12) 发明专利

(10) 授权公告号 CN 116402318 B

(45) 授权公告日 2023.12.01

(21) 申请号 202310664211.4

G06Q 10/0637 (2023.01)

(22) 申请日 2023.06.07

G06Q 10/04 (2023.01)

G06Q 50/06 (2012.01)

(65) 同一申请的已公布的文献号

申请公布号 CN 116402318 A

(56) 对比文件

(43) 申请公布日 2023.07.07

CN 115328653 A, 2022.11.11

CN 116033025 A, 2023.04.28

(73) 专利权人 北京智芯微电子科技有限公司

CN 114896065 A, 2022.08.12

地址 100192 北京市海淀区西小口路66号

CN 113157444 A, 2021.07.23

中关村东升科技园A区3号楼

CN 113448721 A, 2021.09.28

(72) 发明人 苑佳楠 霍超 郑利斌 甄岩

CN 114090253 A, 2022.02.25

张港红 高建 罗安琴 谢凡

CN 115129463 A, 2022.09.30

CN 115169810 A, 2022.10.11

(74) 专利代理机构 北京润平知识产权代理有限公司

公司 11283

CN 115202836 A, 2022.10.18

CN 115460216 A, 2022.12.09

专利代理师 谢熠

CN 115580634 A, 2023.01.06

(51) Int. Cl.

审查员 李娜

G06Q 10/0631 (2023.01)

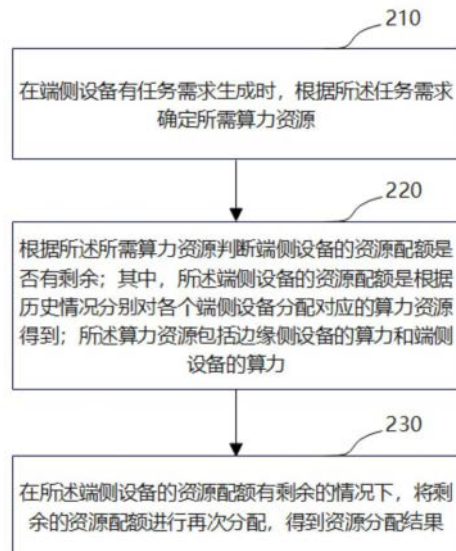
权利要求书2页 说明书10页 附图3页

(54) 发明名称

面向配电网的多级算力资源分配方法、装置及网络架构

(57) 摘要

本发明提供一种面向配电网的多级算力资源分配方法、装置及网络架构,属于配电网技术领域。所述面向配电网的多级算力资源分配方法包括:在端侧设备有任务需求生成时,根据所述任务需求确定所需算力资源;根据所述所需算力资源判断端侧设备的资源配额是否有剩余;在所述端侧设备的资源配额有剩余的情况下,将剩余的资源配额进行再次分配,得到资源分配结果;其中,所述端侧设备的资源配额是根据历史情况分别对各个端侧设备分配对应的算力资源得到;所述算力资源包括边缘侧设备的算力和端侧设备的算力。实现了业务处理需求与算力资源之间的动态匹配,从而提高了配电网算力效率,使得算力资源充分利用。



1. 一种面向配电网的多级算力资源分配方法,其特征在于,包括:
在端侧设备有任务需求生成时,根据所述任务需求确定所需算力资源;
根据所述所需算力资源判断端侧设备的资源配额是否有剩余;
在所述端侧设备的资源配额有剩余的情况下,将剩余的资源配额进行再次分配,得到资源分配结果;

其中,所述端侧设备的资源配额是根据历史情况分别对各个端侧设备分配对应的算力资源得到;所述算力资源包括边缘侧设备的算力和端侧设备的算力;

其中,所述根据所述任务需求确定所需算力资源,包括:
对所述任务需求进行分割,生成多个子任务;
对各个子任务设置优先级;
确定各个子任务对应的算力资源,以得到所需算力资源;包括:
按照各个子任务的优先级顺序,依次在端侧设备的资源配额中确定出各个子任务对应的算力资源。

2. 根据权利要求1所述的面向配电网的多级算力资源分配方法,其特征在于,在确定各个子任务对应的算力资源之后,还包括:

根据所述任务需求建立网络连接,形成算力网络;
将所述各个子任务通过所述算力网络发送至算力资源对应的计算设备中进行处理。

3. 根据权利要求2所述的面向配电网的多级算力资源分配方法,其特征在于,所述将所述各个子任务通过所述算力网络发送至算力资源对应的计算设备中进行处理,包括:

实时获取所述算力网络的网络信息;
根据所述网络信息制定出路由策略;
将所述各个子任务按照所述路由策略发送至算力资源对应的计算设备中进行处理。

4. 根据权利要求1所述的面向配电网的多级算力资源分配方法,其特征在于,还包括:
对各个端侧设备设置优先级;

所述在所述端侧设备的资源配额有剩余的情况下,将剩余的资源配额进行再次分配,得到资源分配结果,包括:

在所述端侧设备的资源配额有剩余的情况下,根据所述各个端侧设备的优先级,将剩余的资源配额进行再次分配,得到资源分配结果。

5. 一种面向配电网的多级算力资源分配的网络架构,用于实现权利要求1-4中任一项所述的面向配电网的多级算力资源分配方法,其特征在于,包括:基础资源层和算力资源层;其中,

所述基础资源层用于在端侧设备有任务需求生成时,根据所述任务需求确定所需算力资源;包括:对所述任务需求进行分割,生成多个子任务;对各个子任务设置优先级;确定各个子任务对应的算力资源,以得到所需算力资源;包括:按照各个子任务的优先级顺序,依次在端侧设备的资源配额中确定出各个子任务对应的算力资源;

所述算力资源层用于根据所述所需算力资源判断端侧设备的资源配额是否有剩余;并在所述端侧设备的资源配额有剩余的情况下,将剩余的资源配额进行再次分配,得到资源分配结果。

6. 一种面向配电网的多级算力资源分配装置,其特征在于,包括:

确定模块,用于在端侧设备有任务需求生成时,根据所述任务需求确定所需算力资源;

判断模块,用于根据所述所需算力资源判断端侧设备的资源配额是否有剩余;其中,所述端侧设备的资源配额是根据历史情况分别对各个端侧设备分配对应的算力资源得到;所述算力资源包括边缘侧设备的算力和端侧设备的算力;

分配模块,用于在所述端侧设备的资源配额有剩余的情况下,将剩余的资源配额进行再次分配,得到资源分配结果;

其中,所述确定模块包括:

分割单元,用于对所述任务需求进行分割,生成多个子任务;

设置模块,用于对各个子任务设置优先级;

资源确定单元,用于确定各个子任务对应的算力资源,以得到所需算力资源;

所述资源确定单元包括:

顺序确定子单元,用于按照各个子任务的优先级顺序,依次在端侧设备的资源配额中确定出各个子任务对应的算力资源。

7.一种处理器,其特征在于,被配置成执行根据权利要求1至4中任一项所述的面向配电网的多级算力资源分配方法。

8.一种机器可读存储介质,该机器可读存储介质上存储有指令,其特征在于,该指令在被处理器执行时使得所述处理器被配置成执行根据权利要求1至4中任一项所述的面向配电网的多级算力资源分配方法。

面向配电网的多级算力资源分配方法、装置及网络架构

技术领域

[0001] 本发明涉及配电网技术领域,具体地涉及一种面向配电网的多级算力资源分配方法、一种面向配电网的多级算力资源分配装置、一种面向配电网的多级算力资源分配的网络架构、一种机器可读存储介质及一种处理器。

背景技术

[0002] 随着配电网物联网的建设,量测体系以及其他环节将会累积大量数据信息,配电网数据信息来源广泛、数据信息量大,不同数据间的关系复杂,而高效的数据信息采集、传输及分析能够为可靠稳定的供电提供重要保障。云计算、边缘计算以及终端的发展驱动整个配电网的算力分配分散和泛在化,即底层设备周围不同距离会散布不同规模的算力。高效利用这些算力,保证云边端算力的无缝协同,同时借助网络使数据与算力得到快速连接、处理。

[0003] 电力物联网新业务生态下的终端、数据和业务不断增加,海量电力数据需要在边缘计算设备上进行分析计算,进而实现高效灵活的业务处理和决策。但是目前业务处理需求与算力资源之间无法动态匹配,导致配电网算力效率低、算力资源利用不充分。

发明内容

[0004] 本发明实施例的目的是提供一种面向配电网的多级算力资源分配方法、一种面向配电网的多级算力资源分配装置、一种面向配电网的多级算力资源分配的网络架构、一种机器可读存储介质及一种处理器,该面向配电网的多级算力资源分配方法实现了业务处理需求与算力资源之间的动态匹配,从而提高了配电网算力效率,使得算力资源充分利用。

[0005] 为了实现上述目的,本申请第一方面提供一种面向配电网的多级算力资源分配方法,包括:

[0006] 在端侧设备有任务需求生成时,根据所述任务需求确定所需算力资源;

[0007] 根据所述所需算力资源判断端侧设备的资源配额是否有剩余;

[0008] 在所述端侧设备的资源配额有剩余的情况下,将剩余的资源配额进行再次分配,得到资源分配结果;

[0009] 其中,所述端侧设备的资源配额是根据历史情况分别对各个端侧设备分配对应的算力资源得到;所述算力资源包括边缘侧设备的算力和端侧设备的算力。

[0010] 在本申请实施例中,所述根据所述任务需求确定所需算力资源,包括:

[0011] 对所述任务需求进行分割,生成多个子任务;

[0012] 确定各个子任务对应的算力资源,以得到所需算力资源。

[0013] 在本申请实施例中,还包括:

[0014] 对各个子任务设置优先级;

[0015] 所述确定各个子任务对应的算力资源,包括:

[0016] 按照各个子任务的优先级顺序,依次在端侧设备的资源配额中确定出各个子任务

对应的算力资源。

[0017] 在本申请实施例中,在确定各个子任务对应的算力资源之后,还包括:

[0018] 根据所述任务需求建立网络连接,形成算力网络;

[0019] 将所述各个子任务通过所述算力网络发送至算力资源对应的计算设备中进行处理。

[0020] 在本申请实施例中,所述将所述各个子任务通过所述算力网络发送至算力资源对应的计算设备中进行处理,包括:

[0021] 实时获取所述算力网络的网络信息;

[0022] 根据所述网络信息制定出路由策略;

[0023] 将所述各个子任务按照所述路由策略发送至算力资源对应的计算设备中进行处理。

[0024] 在本申请实施例中,还包括:

[0025] 对各个端侧设备设置优先级;

[0026] 所述在所述端侧设备的资源配额有剩余的情况下,将剩余的资源配额进行再次分配,得到资源分配结果,包括:

[0027] 在所述端侧设备的资源配额有剩余的情况下,根据所述各个端侧设备的优先级,将剩余的资源配额进行再次分配,得到资源分配结果。

[0028] 本申请第二方面提供一种面向配电网的多级算力资源分配的网络架构,用于实现上述的面向配电网的多级算力资源分配方法,包括:基础资源层和算力资源层;其中,

[0029] 所述基础资源层用于在端侧设备有任务需求生成时,根据所述任务需求确定所需算力资源;

[0030] 所述算力资源层用于根据所述所需算力资源判断端侧设备的资源配额是否有剩余;并在所述端侧设备的资源配额有剩余的情况下,将剩余的资源配额进行再次分配,得到资源分配结果。

[0031] 本申请第三方面提供一种面向配电网的多级算力资源分配装置,包括:

[0032] 确定模块,用于在端侧设备有任务需求生成时,根据所述任务需求确定所需算力资源;

[0033] 判断模块,用于根据所述所需算力资源判断端侧设备的资源配额是否有剩余;其中,所述端侧设备的资源配额是根据历史情况分别对各个端侧设备分配对应的算力资源得到;所述算力资源包括边缘侧设备的算力和端侧设备的算力;

[0034] 分配模块,用于在所述端侧设备的资源配额有剩余的情况下,将剩余的资源配额进行再次分配,得到资源分配结果。

[0035] 在本申请实施例中,所述确定模块包括:

[0036] 分割单元,用于对所述任务需求进行分割,生成多个子任务;

[0037] 资源确定单元,用于确定各个子任务对应的算力资源,以得到所需算力资源。

[0038] 在本申请实施例中,还包括:

[0039] 设置模块,用于对各个子任务设置优先级;

[0040] 所述资源确定单元包括:

[0041] 顺序确定子单元,用于按照各个子任务的优先级顺序,依次在端侧设备的资源配

额中确定出各个子任务对应的算力资源。

[0042] 本申请第四方面提供一种处理器,被配置成执行上述的面向配电网的多级算力资源分配方法。

[0043] 本申请第五方面提供一种机器可读存储介质,该机器可读存储介质上存储有指令,该指令在被处理器执行时使得所述处理器被配置成执行上述的面向配电网的多级算力资源分配方法。

[0044] 通过上述技术方案,通过实时检测业务需求,动态调整算力资源,完成各类任务高效处理和整合输出,并在满足业务需求的前提下实现算力资源合理分配,为端侧设备业务可靠的提供所需算力需求。实现了业务处理需求与算力资源之间的动态匹配,从而提高了配电网算力效率,使得算力资源充分利用。通过边缘侧配置边缘终端及边缘网关,与云主站协同运行,实现不同需求业务的合理配置。

[0045] 本发明实施例的其它特征和优点将在随后的具体实施方式部分予以详细说明。

附图说明

[0046] 附图是用来提供对本发明实施例的进一步理解,并且构成说明书的一部分,与下面的具体实施方式一起用于解释本发明实施例,但并不构成对本发明实施例的限制。在附图中:

[0047] 图1示意性示出了根据本申请实施例的一种面向配电网的多级算力资源分配方法的流程示意图;

[0048] 图2示意性示出了根据本申请实施例的多级算力网络中的算力资源分配方案结构框图;

[0049] 图3示意性示出了根据本申请实施例的面向配电网的多级算力网络架构图;

[0050] 图4示意性示出了根据本申请实施例的面向配电网的多级算力资源分配装置的结构框图;

[0051] 图5示意性示出了根据本申请实施例的计算机设备的内部结构图。

[0052] 附图标记说明

[0053] 410-确定模块;420-判断模块;430-分配模块;A01-处理器;A02-网络接口;A03-内存储器;A04-显示屏;A05-输入装置;A06-非易失性存储介质;B01-操作系统;B02-计算机程序。

具体实施方式

[0054] 以下结合附图对本发明实施例的具体实施方式进行详细说明。应当理解的是,此处所描述的具体实施方式仅用于说明和解释本发明实施例,并不用于限制本发明实施例。

[0055] 请参看图1-图2,图1示意性示出了根据本申请实施例的一种面向配电网的多级算力资源分配方法的流程示意图;图2示意性示出了根据本申请实施例的多级算力网络中的算力资源分配方案结构框图。

[0056] 本实施例提供一种面向配电网的多级算力资源分配方法,包括以下步骤:

[0057] 步骤210:在端侧设备有任务需求生成时,根据所述任务需求确定所需算力资源;

[0058] 在本实施例中,所述端侧设备主要包括智能摄像头、智能电表、传感器等具有一定

算力和通信性能的设备,可进行数据采集、推理、指令执行等一些列功能操作。端侧设备产生任务需求,所述任务需求包括视频、图像、文字等多类型任务,可以通过将任务需求进入需求解析模块,分析业务实时性需求、算力资源需求大小、算法类型等信息,实现对任务类型的感知识别,确定该任务后续所需执行动作。由于不同的任务需求所需的算力资源不同,所以需要确定当前任务需求所需算力资源。

[0059] 在一些实施例中,为了更加准确地确定所需算力资源,所述根据所述任务需求确定所需算力资源包括以下步骤:

[0060] 首先,对所述任务需求进行分割,生成多个子任务;

[0061] 然后,确定各个子任务对应的算力资源,以得到所需算力资源。

[0062] 在本实施例中,可以从实时性和资源需求角度将任务分隔开,以得到多个子任务。还可以是从算力资源需求大小方向进行分割,具体分割方式本实施例不做限定。上述确定各个子任务对应的算力资源可以是根据各个子任务后续所需执行动作确定,在每一个子任务都确定出对应的算力资源后,就可以得到整个任务需求所需算力资源。通过分别确定出各个子任务对应的算力资源,以便于更加准确地得到任务需求所需算力资源。

[0063] 在一些实施例中,考虑到子任务在执行的时候可能存在优先级,因此,在对所述任务需求进行分割时,还可以对各个子任务设置优先级;所述确定各个子任务对应的算力资源的过程包括:按照各个子任务的优先级顺序,依次在端侧设备的资源配额中确定出各个子任务对应的算力资源。

[0064] 在本实施例中,设置优先级就是给各个子任务打上优先级标签,可以是按照预先设置好的优先级顺序进行设置。比如:子任务A为非实时性任务,子任务B为实时性任务,则可以给予子任务A设置优先级为高,给予子任务B设置优先级为低。在设置好优先级后,可以根据优先级顺序来确定出对应的算力资源,这样就可以优先保证优先级较高的子任务有对应的算力资源,从而保证任务的正常执行。

[0065] 例如:匹配需要执行的子任务的算力资源,综合考虑用户的需求,即计算需求、网络需求,从而适应来自不同用户的不同任务需求。对于非实时的视频任务,端侧设备可以优先将数据上传至区域内的边缘计算网关,网关内配置CPU、NPU、GPU等高性能处理器,可以处理复杂指令任务。对于实时性要求较高的任务,优先将数据上传至区域范围内的某一边缘终端上,边缘终端配置CPU、NPU,可以并行地完成计算任务。对于模型建立、仿真及算力预测、分配等高强度计算任务,选择在云主站进行,云主站配置CPU、DPU、TPU、GPU、FPGA等专用处理器,可以快速稳定的进行高强度、高复杂度任务处理。

[0066] 在一些实施例中,在确定各个子任务对应的算力资源之后,还包括以下步骤:

[0067] 首先,根据所述任务需求建立网络连接,形成算力网络;

[0068] 然后,将所述各个子任务通过所述算力网络发送至算力资源对应的计算设备中进行处理。

[0069] 在本实施例中,依据任务需求,在数据和算力资源层之间按照任务需要动态地建立网络连接。算力分配模块和算力调度模块为了实现最优的算力资源分配,会将任务进行任务分割、模型分割,并根据使用的调度优化算法将不同的计算子任务分配算力资源层的不同算力资源中。

[0070] 其中,将所述各个子任务通过所述算力网络发送至算力资源对应的计算设备中进

行处理,包括以下步骤:

[0071] 首先,实时获取所述算力网络的网络信息;

[0072] 然后,根据所述网络信息制定出路由策略;

[0073] 最后,将所述各个子任务按照所述路由策略发送至算力资源对应的计算设备中进行处理。

[0074] 在本实施例中,在算力网络中,网络资源为各种算力任务提供传输路径。在子任务传输的过程中,网络的时延、可靠性、能源的消耗以及资源使用率等信息将反馈到网络管理模块,综合算力服务通告、算力状态通告和算力告知调度信息,制定路由策略。具体的子任务传输到负责计算的设备中,由确定的计算设备进行计算,最后将计算结果传回至端侧设备,完成对端侧设备的遥测、遥信、遥控、遥调,通过应用层配置的APP,可以实现配电网全景感知、故障诊断及定位、故障自动隔离、分布式资源集群调控等功能。

[0075] 上述实现过程中,通过实时获取网络信息,以便于制定出最优的路由策略,以便于将各个子任务快速准确地传输至算力资源对应的计算设备中进行处理。

[0076] 步骤220:根据所述所需算力资源判断端侧设备的资源配额是否有剩余;其中,所述端侧设备的资源配额是根据历史情况分别对各个端侧设备分配对应的算力资源得到;所述算力资源包括边缘侧设备的算力和端侧设备的算力;

[0077] 在本实施例中,上述历史情况是指端侧设备在历史任务需求中占用的算力资源,在端侧设备执行任务前,主站按照历史情况数据对端侧设备进行执行算力及网络资源配额,端侧设备可以按照所分配的算力资源进行任务上传。上述边缘侧设备的算力包括融合终端的算力和边缘网关的算力。

[0078] 步骤230:在所述端侧设备的资源配额有剩余的情况下,将剩余的资源配额进行再次分配,得到资源分配结果;

[0079] 在本实施例中,在端侧设备的资源配额有剩余的情况下,可以将剩余的资源配额进行二次分配,即将剩余的算力资源拿出来共享给其他的端侧设备进行使用。

[0080] 在一些实施例中,还可以对各个端侧设备设置优先级;

[0081] 所述在所述端侧设备的资源配额有剩余的情况下,将剩余的资源配额进行再次分配,得到资源分配结果,包括:在所述端侧设备的资源配额有剩余的情况下,根据所述各个端侧设备的优先级,将剩余的资源配额进行再次分配,得到资源分配结果。

[0082] 在本实施例中,可以考虑端侧设备的优先级,在端侧设备的资源配额有剩余的情况下,优先将剩余的资源配额分配给优先级别高的端侧设备,从而保证优先级别高的端侧设备有足够的算力资源。

[0083] 上述实现过程中,通过在端侧设备有任务需求生成时,根据所述任务需求确定所需算力资源;根据所述所需算力资源判断端侧设备的资源配额是否有剩余;并在所述端侧设备的资源配额有剩余的情况下,将剩余的资源配额进行再次分配,得到资源分配结果。实现了业务处理需求与算力资源之间的动态匹配,从而提高了配电网算力效率,使得算力资源充分利用。通过实时检测业务需求,动态调整算力资源,完成各类任务高效处理和整合输出,并在满足业务需求的前提下实现算力资源合理分配,为端侧设备业务可靠的提供所需算力需求。同时,通过将算力资源、网络资源及优先级预设相结合的方式,为端侧设备提供了可靠的业务需求所需算力需求。通过边缘侧配置边缘终端及边缘网关,与云主站协同运

行,实现不同需求业务的合理配置。

[0084] 图1为一个实施例中面向配电网的多级算力资源分配方法的流程示意图。应该理解的是,虽然图1的流程图中的各个步骤按照箭头的指示依次显示,但是这些步骤并不是必然按照箭头指示的顺序依次执行。除非本文中有明确的说明,这些步骤的执行并没有严格的顺序限制,这些步骤可以以其它的顺序执行。而且,图1中的至少一部分步骤可以包括多个子步骤或者多个阶段,这些子步骤或者阶段并不必然是在同一时刻执行完成,而是可以在不同的时刻执行,这些子步骤或者阶段的执行顺序也不必然是依次进行,而是可以与其它步骤或者其它步骤的子步骤或者阶段的至少一部分轮流或者交替地执行。

[0085] 请参看图3,图3示意性示出了根据本申请实施例的面向配电网的多级算力网络架构图。本实施例提供一种面向配电网的多级算力资源分配的网络架构,用于实现所述的面向配电网的多级算力资源分配方法,包括:基础资源层、算力网络层、算力资源层、应用服务层及算力网络管理编排层;其中,

[0086] 所述基础资源层用于在端侧设备有任务需求生成时,根据所述任务需求确定所需算力资源;

[0087] 所述算力资源层用于根据所述所需算力资源判断端侧设备的资源配额是否有剩余;并在所述端侧设备的资源配额有剩余的情况下,将剩余的资源配额进行再次分配,得到资源分配结果。

[0088] 在本实施例中,基础资源层包括端侧设备及网络基础设施。端侧设备主要包括智能摄像头、智能电表、传感器等具有一定算力和通信性能的设备,可进行数据采集、推理、指令执行等一些列功能操作。网络基础设施主要指连接端侧设备、边缘侧设备、云主站的网络基础设施,包括控制面的SDN 控制器、传统网管以及转发面的网络设备。算力网络层支持对网络、计算、存储等多维资源、服务的感知与通告,以算力服务通告、算力状态通告、网络信息通告等信息为依据,通过“网络+算力”的联合调控,制定算力路由策略,实现业务请求在路由层的按需调度。算力资源层在综合业务算力建模、算力预测、算力标识、算力度量的计算量以及实时性要求,将算力资源需求分配到边缘终端、边缘网关及云主站上,在网络范围内实现灵活可控的算力资源分配调度。计算资源包括CPU、NPU、GPU、DPU、TPU、FPGA等多种计算能力组合。应用服务层通过实现配电网全景监测、故障分析诊断定位自愈、分布式能源集群控制等功能。算力网络管理编排层完成算力路由、算力服务协同编排。具体功能包括对应用管理、算力管理、网络管理以及安全管理等,实现对算力网络的整体管理。

[0089] 上述实现过程中,通过构建面向配电网的多级算力资源分配的网络架构,可以实现业务处理需求与算力资源之间的动态匹配,从而提高了配电网算力效率,使得算力资源充分利用。通过实时检测业务需求,动态调整算力资源,完成各类任务高效处理和整合输出,并在满足业务需求的前提下实现算力资源合理分配,为端侧设备业务可靠的提供所需算力需求。同时,通过将算力资源、网络资源及优先级预设相结合的方式,为端侧设备提供了可靠的业务需求所需算力需求。通过边缘侧配置边缘终端及边缘网关,与云主站协同运行,实现不同需求业务的合理配置。

[0090] 请参看图4,图4示意性示出了根据本申请实施例的面向配电网的多级算力资源分配装置的结构框图。本实施例提供一种面向配电网的多级算力资源分配装置,包括确定模块410、判断模块420和分配模块430,其中:

[0091] 确定模块410,用于在端侧设备有任务需求生成时,根据所述任务需求确定所需算力资源;

[0092] 判断模块420,用于根据所述所需算力资源判断端侧设备的资源配额是否有剩余;其中,所述端侧设备的资源配额是根据历史情况分别对各个端侧设备分配对应的算力资源得到;所述算力资源包括边缘侧设备的算力和端侧设备的算力;

[0093] 分配模块430,用于在所述端侧设备的资源配额有剩余的情况下,将剩余的资源配额进行再次分配,得到资源分配结果。

[0094] 其中,所述确定模块410包括:

[0095] 分割单元,用于对所述任务需求进行分割,生成多个子任务;

[0096] 资源确定单元,用于确定各个子任务对应的算力资源,以得到所需算力资源。

[0097] 其中,还包括:

[0098] 设置模块,用于对各个子任务设置优先级;

[0099] 所述资源确定单元包括:

[0100] 顺序确定子单元,用于按照各个子任务的优先级顺序,依次在端侧设备的资源配额中确定出各个子任务对应的算力资源。

[0101] 上述实现过程中,通过确定模块410在端侧设备有任务需求生成时,根据所述任务需求确定所需算力资源;判断模块420根据所述所需算力资源判断端侧设备的资源配额是否有剩余;分配模块430在所述端侧设备的资源配额有剩余的情况下,将剩余的资源配额进行再次分配,得到资源分配结果。实现了业务处理需求与算力资源之间的动态匹配,从而提高了配电网算力效率,使得算力资源充分利用。通过实时检测业务需求,动态调整算力资源,完成各类任务高效处理和整合输出,并在满足业务需求的前提下实现算力资源合理分配,为端侧设备业务可靠的提供所需算力需求。同时,通过将算力资源、网络资源及优先级预设相结合的方式,为端侧设备提供了可靠的业务需求所需算力需求。通过边缘侧配置

[0102] 所述面向配电网的多级算力资源分配装置包括处理器和存储器,上述确定模块410、判断模块420和分配模块430等均作为程序单元存储在存储器中,由处理器执行存储在存储器中的上述程序单元来实现相应的功能。

[0103] 处理器中包含内核,由内核去存储器中调取相应的程序单元。内核可以设置一个或以上,通过调整内核参数来实现业务处理需求与算力资源之间的动态匹配,提高配电网算力效率,使得算力资源充分利用。

[0104] 存储器可能包括计算机可读介质中的非永久性存储器,随机存取存储器(RAM)和/或非易失性内存等形式,如只读存储器(ROM)或闪存(flash RAM),存储器包括至少一个存储芯片。

[0105] 本发明实施例提供了一种机器可读存储介质,其上存储有程序,该程序被处理器执行时实现所述面向配电网的多级算力资源分配方法。

[0106] 本发明实施例提供了一种处理器,所述处理器用于运行程序,其中,所述程序运行时执行所述面向配电网的多级算力资源分配方法。

[0107] 在一个实施例中,提供了一种计算机设备,该计算机设备可以是终端,其内部结构图可以如图5所示。该计算机设备包括通过系统总线连接的处理器A01、网络接口A02、显示屏A04、输入装置A05和存储器(图中未示出)。其中,该计算机设备的处理器A01用于提供计

算和控制能力。该计算机设备的存储器包括内存储器A03和非易失性存储介质A06。该非易失性存储介质A06存储有操作系统B01和计算机程序B02。该内存储器A03为非易失性存储介质A06中的操作系统B01和计算机程序B02的运行提供环境。该计算机设备的网络接口A02用于与外部的终端通过网络连接通信。该计算机程序被处理器A01执行时以实现一种面向配电网的多级算力资源分配方法。该计算机设备的显示屏A04可以是液晶显示屏或者电子墨水显示屏,该计算机设备的输入装置A05可以是显示屏上覆盖的触摸层,也可以是计算机设备外壳上设置的按键、轨迹球或触控板,还可以是外接的键盘、触控板或鼠标等。

[0108] 本领域技术人员可以理解,图5中示出的结构,仅仅是与本申请方案相关的部分结构的框图,并不构成对本申请方案所应用于其上的计算机设备的限定,具体的计算机设备可以包括比图中所示更多或更少的部件,或者组合某些部件,或者具有不同的部件布置。

[0109] 在一个实施例中,本申请提供的面向配电网的多级算力资源分配装置可以实现为一种计算机程序的形式,计算机程序可在如图5所示的计算机设备上运行。计算机设备的存储器中可存储组成该面向配电网的多级算力资源分配装置的各个程序模块,比如,图4所示的确定模块410、判断模块420和分配模块430。各个程序模块构成的计算机程序使得处理器执行本说明书中描述的本申请各个实施例的文件系统的调用方法中的步骤。

[0110] 图5所示的计算机设备可以通过如图4所示的面向配电网的多级算力资源分配装置中的确定模块410执行步骤210、判断模块420执行步骤220、分配模块430执行步骤230。

[0111] 本申请实施例提供了一种设备,设备包括处理器、存储器及存储在存储器上并可在处理器上运行的程序,处理器执行程序时实现以下步骤:

[0112] 在端侧设备有任务需求生成时,根据所述任务需求确定所需算力资源;

[0113] 根据所述所需算力资源判断端侧设备的资源配额是否有剩余;

[0114] 在所述端侧设备的资源配额有剩余的情况下,将剩余的资源配额进行再次分配,得到资源分配结果;

[0115] 其中,所述端侧设备的资源配额是根据历史情况分别对各个端侧设备分配对应的算力资源得到;所述算力资源包括边缘侧设备的算力和端侧设备的算力。

[0116] 在一个实施例中,所述根据所述任务需求确定所需算力资源,包括:

[0117] 对所述任务需求进行分割,生成多个子任务;

[0118] 确定各个子任务对应的算力资源,以得到所需算力资源。

[0119] 在一个实施例中,还包括:

[0120] 对各个子任务设置优先级;

[0121] 所述确定各个子任务对应的算力资源,包括:

[0122] 按照各个子任务的优先级顺序,依次在端侧设备的资源配额中确定出各个子任务对应的算力资源。

[0123] 在一个实施例中,在确定各个子任务对应的算力资源之后,还包括:

[0124] 根据所述任务需求建立网络连接,形成算力网络;

[0125] 将所述各个子任务通过所述算力网络发送至算力资源对应的计算设备中进行处理。

[0126] 在一个实施例中,所述将所述各个子任务通过所述算力网络发送至算力资源对应的计算设备中进行处理,包括:

- [0127] 实时获取所述算力网络的网络信息；
- [0128] 根据所述网络信息制定出路由策略；
- [0129] 将所述各个子任务按照所述路由策略发送至算力资源对应的计算设备中进行处理。
- [0130] 在一个实施例中,还包括：
- [0131] 对各个端侧设备设置优先级；
- [0132] 所述在所述端侧设备的资源配额有剩余的情况下,将剩余的资源配额进行再次分配,得到资源分配结果,包括：
- [0133] 在所述端侧设备的资源配额有剩余的情况下,根据所述各个端侧设备的优先级,将剩余的资源配额进行再次分配,得到资源分配结果。
- [0134] 本领域内的技术人员应明白,本申请的实施例可提供为方法、系统、或计算机程序产品。因此,本申请可采用完全硬件实施例、完全软件实施例、或结合软件和硬件方面的实施例的形式。而且,本申请可采用在一个或多个其中包含有计算机可用程序代码的计算机可用存储介质(包括但不限于磁盘存储器、CD-ROM、光学存储器等)上实施的计算机程序产品的形式。
- [0135] 本申请是参照根据本申请实施例的方法、设备(系统)、和计算机程序产品的流程图和/或方框图来描述的。应理解可由计算机程序指令实现流程图和/或方框图中的每一流程和/或方框、以及流程图和/或方框图中的流程和/或方框的结合。可提供这些计算机程序指令到通用计算机、专用计算机、嵌入式处理机或其他可编程数据处理设备的处理器以产生一个机器,使得通过计算机或其他可编程数据处理设备的处理器执行的指令产生用于实现在流程图一个流程或多个流程和/或方框图一个方框或多个方框中指定的功能的装置。
- [0136] 这些计算机程序指令也可存储在能引导计算机或其他可编程数据处理设备以特定方式工作的计算机可读存储器中,使得存储在该计算机可读存储器中的指令产生包括指令装置的制造品,该指令装置实现在流程图一个流程或多个流程和/或方框图一个方框或多个方框中指定的功能。
- [0137] 这些计算机程序指令也可装载到计算机或其他可编程数据处理设备上,使得在计算机或其他可编程设备上执行一系列操作步骤以产生计算机实现的处理,从而在计算机或其他可编程设备上执行的指令提供用于实现在流程图一个流程或多个流程和/或方框图一个方框或多个方框中指定的功能的步骤。
- [0138] 在一个典型的配置中,计算设备包括一个或多个处理器(CPU)、输入/输出接口、网络接口和内存。
- [0139] 存储器可能包括计算机可读介质中的非永久性存储器,随机存取存储器(RAM)和/或非易失性内存等形式,如只读存储器(ROM)或闪存(flash RAM)。存储器是计算机可读介质的示例。
- [0140] 计算机可读介质包括永久性和非永久性、可移动和非可移动媒体可以由任何方法或技术来实现信息存储。信息可以是计算机可读指令、数据结构、程序的模块或其他数据。计算机的存储介质的例子包括,但不限于相变内存(PRAM)、静态随机存取存储器(SRAM)、动态随机存取存储器(DRAM)、其他类型的随机存取存储器(RAM)、只读存储器(ROM)、电可擦除可编程只读存储器(EEPROM)、快闪记忆体或其他内存技术、只读光盘只读存储器

(CD-ROM)、数字多功能光盘 (DVD) 或其他光学存储、磁盒式磁带, 磁带磁磁盘存储或其他磁性存储设备或任何其他非传输介质, 可用于存储可以被计算设备访问的信息。按照本文中的界定, 计算机可读介质不包括暂存电脑可读媒体 (transitory media), 如调制的数据信号和载波。

[0141] 还需要说明的是, 术语“包括”、“包含”或者其任何其他变体意在涵盖非排他性的包含, 从而使得包括一系列要素的过程、方法、商品或者设备不仅包括那些要素, 而且还包括没有明确列出的其他要素, 或者是还包括为这种过程、方法、商品或者设备所固有的要素。在没有更多限制的情况下, 由语句“包括一个……”限定的要素, 并不排除在包括要素的过程、方法、商品或者设备中还存在另外的相同要素。

[0142] 以上仅为本申请的实施例而已, 并不用于限制本申请。对于本领域技术人员来说, 本申请可以有各种更改和变化。凡在本申请的精神和原理之内所作的任何修改、等同替换、改进等, 均应包含在本申请的权利要求范围之内。

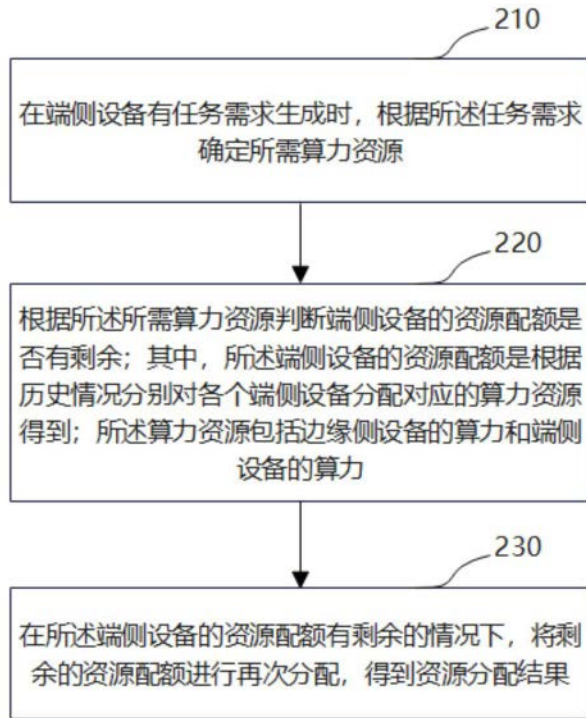


图1

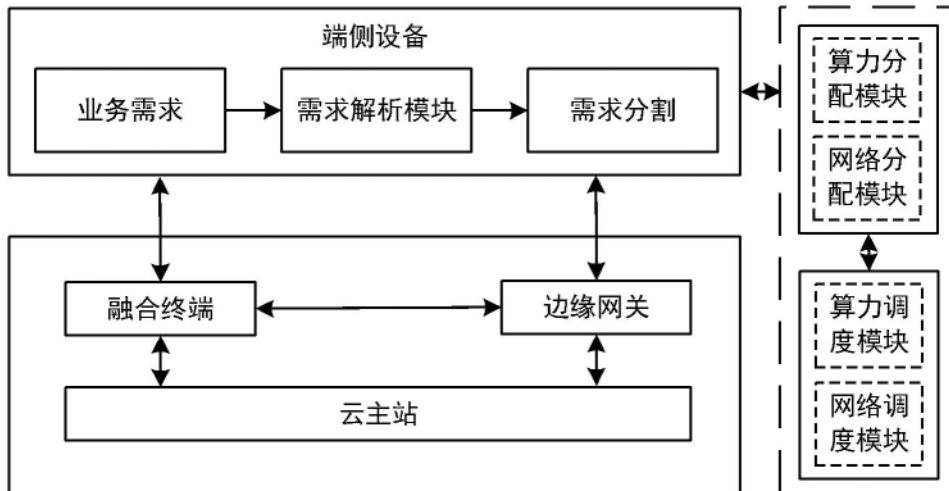


图2

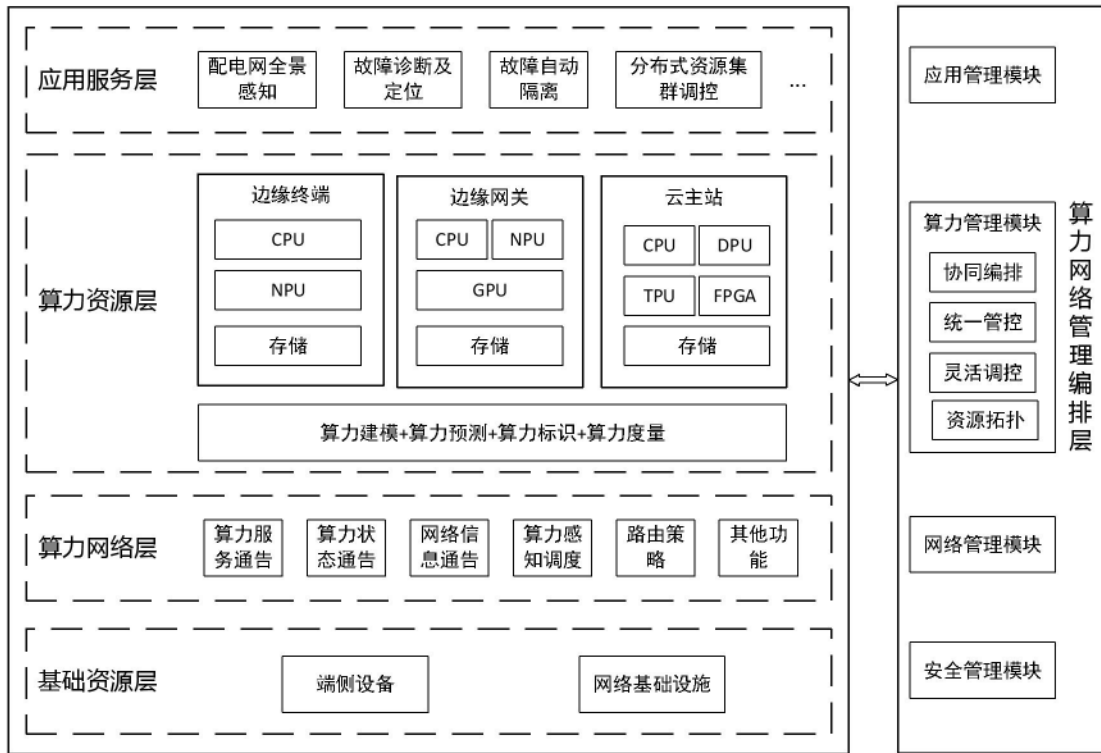


图3

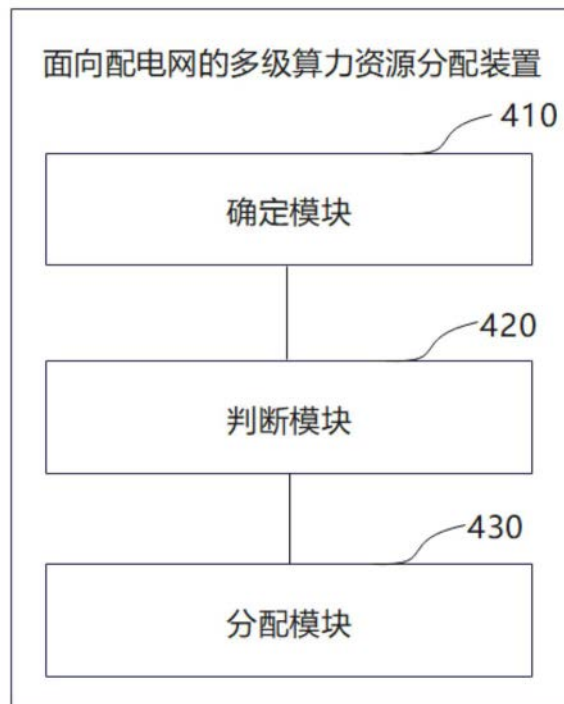


图4

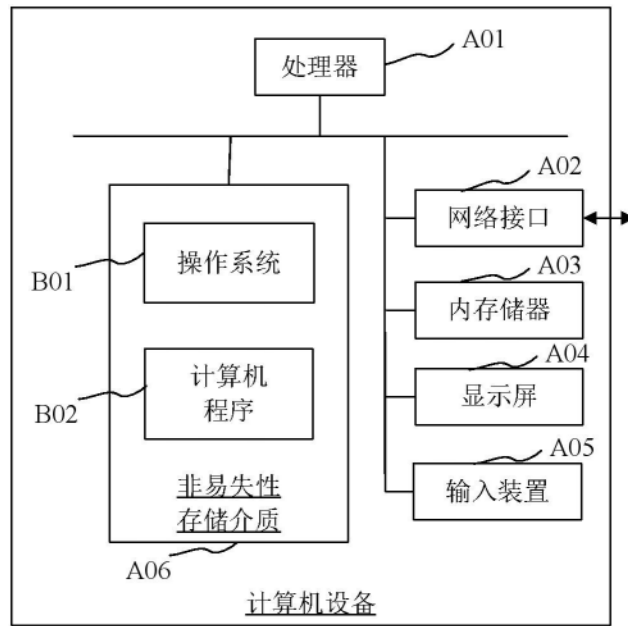


图5