

计算集群及其数据采集方法、设备及存储介质

5 本申请要求于 2022 年 05 月 17 日提交中国专利局、申请号为 202210541467.1、申请名称为“计算集群及其数据采集方法、设备及存储介质”的中国专利申请的优先权，其全部内容通过引用结合在本申请中。

技术领域

10 本申请涉及云计算技术领域，尤其涉及一种计算集群及其数据采集方法、设备及存储介质。

背景技术

15 高性能计算（High Performance Computing，缩写 HPC）指通常使用很多处理器（作为单个机器的一部分）或者某一集群中组织的几台计算机（作为单个计算资源操作）的计算系统和环境。有许多类型的 HPC 系统，其范围从标准计算机的大型集群，到高度专用的硬件。大多数基于集群的 HPC 系统使用高性能网络将计算机进行互连。

其中，性能监测和性能分析是构建 HPC 系统不可缺少的一部分。对 HPC 系统来说，性能监测和性能分析时，性能指标数据的处理开销是 HPC 系统面临的一个挑战，包括数据传输开销、数据处理开销、数据存储开销和数据分析开销。

20 为了降低性能指标数据的处理开销，通常做法是使用较低的采集频率，来减少性能指标数据的数据量，然而较低的采集频率可能造成性能指标数据的缺失或失真，影响性能分析结果的准确性，导致基于性能分析结果的决策失误。

发明内容

25 本申请的多个方面提供一种计算集群及其数据采集方法、设备及存储介质，用以解决性能指标数据的处理开销与性能分析结果准确性之间的矛盾问题，保证性能指标的采集精度，保障性能分析的准确性，同时减少性能指标数据的采集和处理开销。

30 本申请实施例提供一种计算集群，包括：管控节点和多个计算节点，每个计算节点上部署有多个采集器，不同采集器用于采集不同的性能指标；管控节点，用于在多个计算节点中的至少两个计算节点上部署同一作业任务，并控制至少两个计算节点执行作业任务；每个计算节点，用于在执行作业任务过程中，启动与作业任务相关的至少两个目标采集器，以使所述至少两个目标采集器以当前采集频率采集其所在计算节点的至少两种性能指标数据；在确定自身是至少两个计算节点中的主节点的情况下，根据至少两种性能指标数据的变化信息，调整至少两个目标采集器的采集频率；通知至少两个计算节点中其它计算节点调整其上至少两个目标采集器的采集频率，以使其它计算节点上的至少两个目标采集器以调整后的采集频率继续采集其所在计算节点的至少两种性能指标数据。

本申请实施例还提供一种用于计算集群的数据采集方法，包括：在执行作业任务过程中，启动与作业任务相关的至少两个目标采集器，以使所述至少两个目标采集器以当前采集频率采集其所在计算节点的至少两种性能指标数据，作业任务被部署在计算集群中的至少两个计算节点上；在确定自身是至少两个计算节点中的主节点的情况下，根据至少两种性能指标数据的变化信息，调整至少两个目标采集器的采集频率；通知至少两个计算节点中的其它计算节点调整其上部署的至少两个目标采集器的采集频率，以使其它计算节点上的至少两个目标采集器以调整后的采集频率继续采集其所在计算节点的至少两种性能指标数据。

本申请实施例还提供另一种用于计算集群的数据采集方法，包括：在执行作业任务过程中，启动与作业任务相关的至少两个目标采集器，以使所述至少两个目标采集器以当前采集频率采集其所在计算节点的至少两种性能指标数据；根据至少两个目标采集器负责采集的性能指标的关联性，将至少两个目标采集器划分至至少两个关联采集器分组中；根据每个关联采集器分组内目标采集器采集到的性能指标数据的变化信息，分别调整每个关联采集器分组内目标采集器的采集频率，以使目标采集器以调整后的采集频率继续采集其所在计算节点的至少两种性能指标数据。

本申请实施例还提供一种数据采集装置，应用于计算集群中的任一计算节点，包括：启动模块，用于在执行作业任务过程中，启动与作业任务相关的至少两个目标采集器，以使所述至少两个目标采集器以当前采集频率采集其所在计算节点的至少两种性能指标数据，作业任务被部署在计算集群中的至少两个计算节点上；调整模块，用于在确定自身是至少两个计算节点中的主节点的情况下，根据至少两种性能指标数据的变化信息，调整至少两个目标采集器的采集频率；通知模块，用于通知至少两个计算节点中的其它计算节点调整其上部署的至少两个目标采集器的采集频率，以使其它计算节点上的至少两个目标采集器以调整后的采集频率继续采集其所在计算节点的至少两种性能指标数据。

本申请实施例还提供一种计算节点，可应用于计算集群，计算节点包括：存储器和处理器；存储器，用于存储计算机程序；处理器，与存储器耦合，用于执行计算机程序，以用于执行以上所述方法中的步骤。

本申请实施例还提供一种存储有计算机程序的计算机可读存储介质，当计算机程序被处理器执行时，致使处理器能够实现以上所述方法中的步骤。

在本申请实施例中，在计算集群场景中，根据性能指标数据的变化信息，自适应的改变性能指标数据的采集频率，既能保证采集精度，保障基于性能指标数据的性能分析和基于分析结果的决策准确性，又能减少性能指标数据的采集和处理开销；在自适应改变采集频率过程中，针对执行同一作业任务的至少两个计算节点，由至少两个计算节点中的主节点负责采集频率的适应性变化处理并在需要改变的情况下同步给其它计算节点，其它计算节点无需负责采集频率的自适应变化处理，可减轻其它计算节点的处理负担。

35

附图说明

此处所说明的附图用来提供对本申请的进一步理解，构成本申请的一部分，本申请的示意性实施例及其说明用于解释本申请，并不构成对本申请的不当限定。在附图中：

图 1a 为本申请一示例性实施例提供的计算集群的结构示意图；

图 1b 为本申请一示例性实施例提供的多个频率组的示意图；

图 2 为本申请又一示例性实施例提供的计算集群的数据采集方法的流程示意图；

图 3 为本申请又一示例性实施例提供的计算集群的数据采集方法的流程示意图；

5 图 4 为本申请又一示例性实施例提供的数据采集装置的结构示意图；

图 5 为本申请又一示例性实施例提供的数据采集装置的结构示意图；

图 6 为本申请又一示例性实施例提供的计算节点的结构示意图。

具体实施方式

10 为使本申请的目的、技术方案和优点更加清楚，下面将结合本申请具体实施例及相应的附图对本申请技术方案进行清楚、完整地描述。显然，所描述的实施例仅是本申请一部分实施例，而不是全部的实施例。基于本申请中的实施例，本领域普通技术人员在没有做出创造性劳动前提下所获得的所有其他实施例，都属于本申请保护的范围。

以下结合附图，详细说明本申请各实施例提供的技术方案。

15 图 1a 为本申请一示例性实施例提供的计算集群 100 的结构示意图。本实施例的计算集群 100 可以实现为规模较大的计算平台，或者是 HPC 系统，或者是一个或多个机房，互联网数据中心（IDC），或者是云计算系统等，本实施例对计算集群 100 的具体实现形态不做限定。如图 1a 所示，该计算集群 100 包括：管控节点 101 和多个计算节点 102。管控节点 101 与多个计算节点 102 之间，以及多个计算节点 102 之间可以实现通信连接。

20 在本实施例中，上述通信连接可以是有线或无线通信连接。可选地，在无线通信连接的情况下，各节点之间可以通过移动网络实现通信连接，相应地，移动网络的网络制式可以为 2G（GSM）、2.5G（GPRS）、3G（WCDMA、TD-SCDMA、CDMA2000、UTMS）、4G（LTE）、4G+（LTE+）、5G、WiMax 或者未来即将出现的新网络制式等中的任意一种。可选地，各节点也可以位于同一局域网内，则在无线通信连接的情况下，各节点也可以通过蓝牙、WiFi、红外、zigbee 或 NFC 等方式实现通信连接。

25 在本实施例中，并不限定管控节点 101 和计算节点 102 的实现形态。其中，管控节点 101 的实现形态有多种，例如可以部署在虚拟机、云服务器、云主机或物理机上。可选地，管控节点 101 可以集中部署在一台物理机或一台虚拟机上，也可以分布式部署在多台物理机或多台虚拟机上，并不限于此。相应地，计算节点 102 可以是任何具有一定计算能力和通信能力的设备形态，例如可以是虚拟机、物理机（如服务器、计算机设备）、云服务器、云主机、虚拟中心、服务器阵列或数据库等。

30 其中，管控节点 101 一方面可以面向用户提供人机交互接口，通过人机交互接口接收用户提交的作业任务，另一方面可以对计算集群 100 进行各种管控，例如可以在多个计算节点 102 上部署作业任务，控制各计算节点 102 执行作业任务，管理各计算节点 102 的任务执行状态等。在实际应用中，一个作业任务可以被部署在一个计算节点 102 上，或者也可以被部署在至少两个计算节点 102 上，具体是作业任务的类型和性能要求而定，对于计算量较大或者对计算效率要求较高的作业任务，可以同时被部署在至少两个计算节点 102

上，由至少两个计算节点 102 并行执行，提高计算效率。基于此，管控节点 101 具体可用于在多个计算节点 102 中的至少两个计算节点 102 上部署同一作业任务，并控制至少两个计算节点 102 执行该作业任务。其中，在计算节点 102 上部署作业任务的方式可以是但不限于：将作业任务相关的数据下发至计算节点 102，或者，向计算节点 102 下发任务指令，在任务指令中携带作业任务的标识信息，由计算节点 102 根据作业任务的标识信息从任务数据库中获取作业任务相关的数据。其中，控制计算节点 102 执行作业任务的方式可以是但不限于：向计算节点 102 发送启动指令，指示计算节点 102 开始执行作业任务；或者，可以向计算节点 102 下发作业指示参数，该作业指示参数包括执行作业任务的时间，例如 10 分钟后启动作业任务，或者在指定时间 xxx 时 xxx 分启动作业任务，等等。

在本实施例中，每个计算节点 102 作为任务执行节点，可接收管控节点 101 部署的作业任务，并根据管控节点 101 的控制执行作业任务。除此之外，每个计算节点 102 上部署有多个采集器，每个采集器负责采集一种性能指标数据，不同采集器负责采集不同的性能指标数据。采集器可以是一种具有数据采集功能的程序代码，在实现形态上可以是依托于主程序的插件、SDK，或者也可以是独立的软件功能模块，对此不做限定。每个采集器可按照一定采集频率进行性能指标数据的采集操作，性能指标数据的数量与采集频率的大小有直接关系；其中，采集频率越高，采集到的性能指标数据越多，基于性能指标数据的性能分析和性能监测的精度越高，当然，数据传输、存储和计算开销也就越大；采集频率越低，采集到的性能指标数据越少，基于性能指标数据的性能分析和性能监测的精度相对较低，当然，数据传输、存储和计算开销也相对越小。

基于上述，本实施例中的每个计算节点 102 除了根据管控节点 101 的控制执行作业任务之外，还可以在作业任务过程中，启动与作业任务相关的至少两个采集器，以使被启动的至少两个采集器以当前采集频率采集至少两种性能指标数据。为了便于描述和区分，将计算节点 102 在执行作业任务过程中所启动的与该作业任务相关的至少两个采集器称为目标采集器，目标采集器的数量为至少两个，对于不同作业任务所启动的目标采集器会有所不同，具体视作业任务对性能指标的要求而定。

本实施例中的性能指标数据包括但不限于：计算节点 102 的 CPU 利用率、内存利用率、内存剩余量、网络带宽的大小、作业任务占用的 CPU 资源量、作业任务占用的内存资源量、作业任务消耗的带宽资源量，等等。根据这些性能指标数据，可以从计算节点 102 和/或作业任务的维度进行性能分析和性能监测。例如，根据这些性能指标数据，可以分析或监测计算节点 102 的性能属性，可分析或可监测的性能属性包括但不限于：任务负载情况、网络状态，以及当前可用资源量等，当前可用资源量至少包括：计算节点 102 的 CPU 或内存的剩余量；进一步，管控节点 101 可以获取计算节点 102 的这些性能属性，并可根据这些性能属性确定是否可以继续为该计算节点 102 分配新的作业任务，以及是否需要动态调整该计算节点 102 的资源量，例如增加 CPU 资源或者网络带宽资源等。例如，根据这些性能指标数据，可以分析或监测作业任务的运行状态、资源消耗情况，以及作业任务对应的服务质量 (QoS) 等；进一步，管控节点 101 可以获取作业任务的运行状态、资源消耗情况，以及作业任务对应的 QoS，并可根据这些信息确定是否需要为作业任务增加或减少计算节点 102，以便在保证作业任务的运行状态和 QoS 的情况下，尽量做到节点资源的合理利用。

基于上述分析，如图 1a 所示，本实施例的计算机集群还包括：性能分析节点 104，性

能分析节点 104 与管控节点 101 及多个计算节点之间实现通信连接。通信连接方式可参见上文描述，在此不再赘述。在本实施例中，性能分析节点 104 负责接收各计算节点 102 上报的性能指标数据，根据各计算节点 102 上报的性能指标数据对计算节点 102 和/或作业任务进行性能分析和性能监测。性能分析和监测是针对计算集群 100 进行防护性维护的必要部分，方便运维人员了解整个集群的运行情况，观察集群资源使用效率等。

具体地，性能分析节点 104 可以根据各计算节点 102 上报的性能指标数据，例如各计算节点 102 的 CPU 利用率、内存利用率、内存剩余量、网络带宽的大小等，可以分析或监测计算节点 102 的性能属性，可分析或可监测的性能属性包括但不限于：任务负载情况、网络状态，以及当前可用资源量等，并将各计算节点 102 的性能属性提供给管控节点 101，以供管控节点 101 做进一步决策。和/或，性能分析节点 104 可以根据各计算节点 102 上报的性能指标数据，例如各计算节点 102 的 CPU 利用率、内存利用率、作业任务占用的 CPU 资源量、内存资源量以及消耗的带宽资源，可以分析或监测作业任务的运行状态、资源消耗情况，以及作业任务对应的 QoS 等性能数据，并将作业任务的各种性能数据提供给管控节点 101，以供管控节点 101 做进一步决策。例如，管控节点 101 可以分析作业任务的运行时的行为特征；根据作业任务的运行时的行为特征，为该作业任务适配较佳的资源配置，这些资源配置至少包括计算节点 102 的数量、每个计算节点 102 上的 CPU、内存、网络等资源。

基于上述，对于在至少两个计算节点 102 上部署同一作业任务的情况，性能分析节点 104 具体可用于根据上述至少两个计算节点 102 分别上报的至少两种性能指标数据，分析得到至少两个计算节点 102 的最新性能属性，并提供给管控节点 101，以供管控节点 101 做进一步决策，使整个计算机集群在性能管控方面形成一个闭环。

在本申请实施例中，每个计算节点 102 可以执行作业任务，并可以在执行作业任务过程中启动与作业任务相关的至少两个目标采集器，以使至少两个目标采集器以当前采集频率采集其所在计算节点 102 的至少两种性能指标数据。除此之外，在至少两个目标采集器采集性能指标数据的过程中，每个计算节点 102 还可以根据性能指标数据的变化信息，自适应的改变其上至少两个目标采集器采集性能指标数据使用的采集频率，以便于实现性能指标数据的变频采集，既能保证采集精度，保障基于性能指标数据的性能分析和决策的准确性，又能减少性能指标数据的采集和处理开销。

在本实施例中，由于每个计算节点 102 上的至少两个目标采集器在采集的过程中可能都需要做变频处理，若每个计算节点 102 针对各自的至少两个目标采集器的采集频率逐个进行计算调整，数据处理量太大，既费时又费力，尤其是在超算场景下，计算节点 102 以及目标采集器的数量较多，计算量会比较大，影响计算节点 102 的整体性能。为了减少动态调整采集频率引起的数据处理量并提高对目标采集器的频率的调整效率，在本申请实施例中，对于部署同一作业任务的至少两个计算节点 102，可以从中选定主节点 103，由主节点 103 根据其上至少两个目标采集器采集到的主节点 103 的至少两种性能指标数据的变化信息，调整至少两个目标采集器的采集频率，并通知至少两个计算节点 102 中其它计算节点 102 调整其上至少两个目标采集器的采集频率，这样各计算节点 102 上的至少两个目标采集器就都能够以调整后的采集频率继续采集至少两种性能指标数据。需要说明的是，每个计算节点上的至少两个目标采集器负责采集其所在计算节点的至少两种性能指标数

据。在该过程中，只有主节点 103 负责执行集频率调整相关的数据处理操作，而其它计算节点 102 无需执行采集频率调整相关的数据处理操作，可以根据主节点 103 的通知直接调整其上至少两个目标采集器的采集频率即可，可减少动态调整采集频率引起的数据处理量，还可以提高对目标采集器的采集频率的调整效率。

5 在本本申请上述或下述实施例中，每个计算节点 102 在执行任务过程中，启动与作业任务相关的至少两个目标采集器，一种具体的实现方式如下：根据与自身所执行任务相关的至少两个性能指标，确定与该至少两个性能指标对应的至少两个目标采集器；之后根据本地存储的上述至少两个目标采集器的标识信息，将启动指令发送至相应的至少两个目标采集器；至少两个目标采集器在接收到启动指令后，开始运行并采集相应的性能指标数据。

10 进一步，至少两个采集器将采集到的性能指标数据发送给所属计算节点 102，所属计算节点 102 会根据至少两个目标采集器采集的至少两种性能指标数据的变化信息，调整至少两个目标采集器的采集频率。由于在执行同一作业任务中，各目标采集器的采集频率可能是相同或相近的，那么为了减少各计算节点 102 对各目标采集器的采集频率调整的任务量，可以从至少两个计算节点 102 中选择一个计算节点 102 作为主节点 103，由该主节点
15 103 根据其上至少两个目标采集器采集到的至少两种性能指标数据的变化信息，确定调整后的采集频率，一方面调整本地至少两个目标采集器的采集频率，另一方面通知给其它计算节点 102，其它计算节点 102 根据通知直接调整其上至少两个目标采集器使用的采集频率。基于此，对执行同一作业任务的至少两个计算节点 102 而言，每个计算节点 102 还需要确定自身是否属于主节点 103，在确定自己是主节点 103 的情况下，根据本地至少两个
20 目标采集器采集到的至少两种性能指标数据的变化信息，调整至少两个目标采集器的采集频率；通知至少两个计算节点 102 中其它计算节点 102 调整其上至少两个目标采集器的采集频率。进一步，对于每个计算节点 102，在确定自己不是主节点 103 的情况下，可以等待主节点 103 的通知，在接收到主节点 103 的通知之前，至少两个目标采集器以当前采集频率采集其所在计算节点的至少两种性能指标数据，在接收到主节点 103 的通知之后，调
25 整本地至少两个目标采集器的采集频率，至少两个目标采集器以调整后的采集频率继续采集其所在计算节点的至少两种性能指标数据。

在本申请实施例中，并不限定主节点 103 的选择方式，可以采用但不限于以下几种：

方式 A1：主节点 103 由管控节点 101 负责选择。具体地，管控节点 101 还用于：根据至少两个计算节点 102 的属性信息，从至少两个计算节点 102 中选择主节点 103，并向主
30 节点 103 发送通知消息。对执行同一作业任务的至少两个计算节点 102 而言，每个计算节点 102 可根据是否接收到管控节点 101 发送的通知消息，如果接收到通知消息，则确定自身是主节点 103，如果未接收到通知消息，则确定自身不是主节点 103。

在本实施例中，至少两个计算节点 102 可能执行一个或多个作业任务，不同的作业任务的工作负载的大小、所需的网络带宽及使用的资源量都可能是不同的。基于此，管控节点 101 根据至少两个计算节点 102 的属性信息，从至少两个计算节点 102 中选择主节点 103
35 的一种可选实施方式如下：根据至少两个计算节点 102 的任务负载、网络状态和可用资源量中的至少一种性能属性，从至少两个计算节点 102 中选择主节点 103。其中，任务负载表示计算节点 102 所执行任务的负载的大小；网络状态表示计算节点 102 在执行任务时的网络带宽的大小；可用资源量表示计算节点 102 当前的 CPU、内存等剩余量。

在一可选实施例中，从至少两个计算节点 102 中选择主节点 103 时几种具体的实施方式如下：根据至少两个计算节点 102 的任务负载，从至少两个计算节点 102 中选择主节点 103；或者，根据至少两个计算节点 102 的网络状态，从至少两个计算节点 102 中选择主节点 103；或者，根据至少两个计算节点 102 的可用资源量，从至少两个计算节点 102 中选择主节点 103；或者，根据至少两个计算节点 102 的任务负载和网络状态，从至少两个计算节点 102 中选择主节点 103；或者，从至少两个计算节点 102 的任务负载和可用资源量，从至少两个计算节点 102 中选择主节点 103；或者，根据至少两个计算节点 102 的网络状态和可用资源量，从至少两个计算节点 102 中选择主节点 103；或者，根据至少两个计算节点 102 的任务负载、网络状态和可用资源量，从至少两个计算节点 102 中选择主节点 103。

进一步，接续于上述可选实施例，在根据至少两个计算节点 102 的任务负载，从至少两个计算节点 102 中选择主节点 103 时，将任务负载小的计算节点 102 作为主节点 103；在根据至少两个主节点 103 的网络状态，从至少两个计算节点 102 中选择主节点 103 时，将网络状态较佳的计算节点 102 作为主节点 103；在根据至少两个计算节点 102 的可用资源量，从至少两个计算节点 102 中选择主节点 103 时，将可用资源量较多的计算节点 102 作为主节点 103；在根据至少两个计算节点 102 的任务负载和网络状态，从至少两个计算节点 102 中选择主节点 103 时，将任务负载较小且网络状态较佳的计算节点 102 作为主节点 103；在根据至少两个计算节点 102 的任务负载和可用资源量，从至少两个计算节点 102 中选择主节点 103 时，将任务负载较小且可用资源量较多的计算节点 102 作为主节点 103；在根据至少两个计算节点 102 的网络状态和可用资源量，从至少两个计算节点 102 中选择主节点 103 时，将网络状态较佳且可用资源量较多的计算节点 102 作为主节点 103；或者，在根据至少两个计算节点 102 的任务负载、网络状态和可用资源量，从至少两个计算节点 102 中选择主节点 103 时，将任务负载较小、网络状态较佳和可用资源量较多的计算节点 102 作为主节点 103。

需要说明的是，上述从至少两个计算节点 102 确定主节点 103 的方式仅为示例性说明，但不限于此。

进一步，在本实施例中，考虑到主节点 103 的性能属性可能是动态变化的，为了提高动态调整采集频率的执行效率，可以动态更换新的主节点 103。基于此，性能分析节点 104 可以根据至少两个计算节点 102 分别上报的至少两种性能指标数据，分析得到至少两个计算节点 102 的最新性能属性，并提供给管控节点 101；管控节点 101 还用于：根据至少两个计算节点 102 的最新性能属性，重新从至少两个计算节点 102 中选择新的主节点 103，并向新的主节点 103 发送通知消息。进一步，当新的主节点 103 接收到通知消息后，可以自动成为新的主节点 103。进一步，管控节点 101 还可以向原有主节点 103 发送转为非主节点 103 的指示信息，原有主节点 103 在接收到指示信息时，关闭主节点 103 的功能。需要说明的是，在图 1a 中示出了管控节点 101 选择主节点 103 并动态更新主节点 103 的方式。

方式 A2：除了管控节点 101 从至少两个计算节点 102 选择主节点 103 之外，计算节点 102 之间还可以按照设定的主节点 103 选择方式自行协商出主节点 103。一种具体的实施方式如下：对于执行同一作业任务的每个计算节点 102，可以根据至少两个计算节点 102

(即自身和其它计算节点 102) 的指定属性信息, 结合预先设定的根据指定属性信息选择主节点 103 应该满足的条件, 确定自身是否为主节点 103。其中, 指定属性信息可以是计算节点 102 的设备编号或 IP 地址, 根据指定属性信息选择主节点 103 应该满足的条件可以是设备编号或 IP 地址最大的节点作为主节点 103, 或者是设备编号或 IP 地址最大的节点作为主节点 103。基于此, 上述根据至少两个计算节点 102 的设备编号或 IP 地址, 结合预先设定的根据指定属性信息选择主节点 103 应该满足的条件, 确定自身是否为主节点 103 的一种方式, 包括: 每个计算节点 102 将自身的设备编号或 IP 地址与其它计算设备的设备编号或 IP 地址进行比较; 若自身的设备编号或 IP 地址是最大的, 则确定自身为主节点 103, 反之确定自身不是主节点 103。在另一种可选实施例中, 每个计算节点 102 将自身的设备编号或 IP 地址与其它计算设备的设备编号或 IP 地址进行比较; 若自身的设备编号或 IP 地址是最小的, 则确定自身为主节点 103, 反之确定自身不是主节点 103。需要说明的是, 在图 1a 中同时示出了计算节点 102 自主协商主节点 103 的方式。

在本申请上述或下述实施例中, 主节点 103 可以根据本地至少两个目标采集器采集到的至少两种性能指标数据的变化信息, 调整至少两个目标采集器的采集频率。其具体实施方式如下: 首先, 根据至少两个目标采集器负责采集的性能指标的关联性, 将至少两个目标采集器划分至至少两个关联采集器分组中; 以关联采集器分组为单位, 根据每个关联采集器分组内目标采集器采集到的性能指标数据的变化信息, 分别调整每个关联采集器分组内目标采集器的采集频率。在本实施例中, 对采集器进行分组, 并以分组为单位对关联性较强的采集器统一进行采集频率的调整, 也就是说, 对同一关联采集器分组内的目标采集器而言, 调整后的采集频率是相同的, 有利于进一步简化调整采集频率所消耗的计算资源, 提高采集频率的整体调整效率。

在一可选实施例中, 可以将至少两个目标采集器负责采集的性能指标的关联性大于预设阈值的目标采集器划分至至少两个关联采集器分组中。例如, 可以将用于采集 CPU 利用率的目标采集器和用于采集 CPU 浮点运行效率的目标采集器划分至同一关联采集器分组中, 将用于采集内存利用率的目标采集器和用于采集读写带宽资源的目标采集器划分至同一关联采集器分组中, 将用于采集网络收/发带宽资源的目标采集器和用于采集网络收/发速率的目标采集器划分至同一关联器分组中, 但不限于此。

进一步, 在一可选实施例中, 针对每个关联采集器分组, 在调整每个关联采集器分组内目标采集器的采集频率之前, 还可以判断该关联采集器分组内各目标采集器的当前采集频率是否相同, 若不相同, 将该关联采集器分组内各目标采集器的当前采集频率调整为相同的采集频率。当前采集频率是指各目标采集器当前使用的采集频率。

接续于上述可选实施例, 在关联采集器分组内各目标采集器的当前采集频率不同的情况下, 将该关联采集器分组内各目标采集器的当前采集频率调整为相同的采集频率, 可以采用以下几种可选的实施方式: 将该关联采集器分组内各目标采集器的当前采集频率均调整为该关联采集器分组内各目标采集器的当前采集频率的平均值; 或者, 将该关联采集器分组内各目标采集器的当前采集频率均调整为该关联采集器分组内各目标采集器的当前采集频率中的最大采集频率; 或者, 将该关联目标采集器内各目标采集器的当前采集频率均调整为该关联采集器分组内各目标采集器的当前采集频率中的最小采集频率。上述对关联采集器分组内各目标采集器的当前采集频率调整为相同的采集频率的方式仅为示例性

说明，并不限于此。

同理，接续于上述可选实施例，在将关联采集器分组内各目标采集器的当前采集频率调整为相同的采集频率后，根据每个关联采集器分组内目标采集器采集到的性能指标数据的变化信息，分别调整每个关联采集器分组内目标采集器的采集频率，其具体实施方式如下：针对每个关联采集器分组，根据该关联采集器分组内目标采集器采集到的性能指标数据的变化信息，确定该关联采集器分组对应的变频方向；之后将该关联采集器分组内目标采集器当前使用的采集频率，调整为多个预设频率中在该变频方向上最接近的预设频率，其中，多个预设频率由小到大。在本实施例中，变频方向包括增大频率，保持不变和减小频率三种，但不限于此，也可以将增大频率、减小频率的变频粒度细化，从而得到更多种变频方向。在本实施例中，预先设定多个频率，且多个预设频率由小到大，并不相同。假设多个预设频率从小到大依次为 f1、f2、f3、f4、f5，假设当前采集频率为 f2，变频方向为增大频率，则多个预设频率中在该变频方向上最接近的预设频率是指频率 f3；同理，在变频方向为减小频率的情况下，多个预设频率中在该变频方向上最接近的预设频率是指频率 f1。

在上述实施例中，预先设定多个频率分组，每个频率分组对应一个预设频率，这些预设频率由小到大，并按照采集频率从大到小有序地排列，如图 1b 所示，包括以下频率组，例如高频组 m-高频组 1，基频组，以及低频组 1-低频组 n；其中，m 和 n 是正整数。对主节点 103 而言，在开始时，可以将至少两个目标采集器初始化为基频，并将这些目标采集器统一添加至基频组中进行管理；然后，根据目标采集器负责采集的性能指标的关联性，将至少两个目标采集器划分到不同的关联采集器分组中，以关联采集器分组为单位，对每个关联采集器分组，根据该关联采集器分组内目标采集器采集到的性能指标数据的变化信息，确定该关联采集器分组对应的变频方向；根据该变频方向，将该关联采集器分组内的目标采集器从当前所在频率组调整到在该变频方向上最接近的频率组中。

具体地，对于变频方向为增大频率的关联采集器分组，将该关联采集器分组内的目标采集器从基频分组中移动到高频组 1 中；对于变频方向为减小频率的关联采集器分组，将该关联采集器分组内的目标采集器从基频分组中移动到低频组 1 中；对于变频方向为保持不变的关联采集器分组，将该关联采集器分组内的目标采集器继续保留在基频组中。随着时间的推移，可以根据类似变频方式，不断调整各关联采集器分组内的目标采集器所在的频率组，对于位于某个频率组中的目标采集器使用其所在频率组对应的预设频率进行性能指标数据的采集。

在一可选实施例中，针对每个关联采集器分组，根据该关联采集器分组内目标采集器采集到的性能指标数据的变化信息，确定该关联采集器分组对应的变频方向，一种具体的实施方式如下：首先针对每个关联采集器分组，获取该关联采集器分组内关键采集器采集到的关键性能指标数据，所述关键采集器是负责采集关键性能指标的目标采集器，关键指标数据是该关联采集器分组中各目标采集器能够采集到的部分指标数据，例如关键性能指标数据可以是重要性排序靠前的一个或多个指标数据；之后根据设定的统计间隔，统计各关键性能指标数据的变化率，并根据各关键性能指标数据的变化率生成全局变化率；进一步，根据所述全局变化率，确定该关联采集器分组对应的变频方向，该关联采集器分组对应的变频方向可以是增大频率、降低频率和保持不变中的任一种。

其中，在统计各关键性能指标数据的变化率时，可以根据设定的统计间隔，统计各关键性能指标数据在统计间隔之内的变化率，本实施例对统计间隔不做限定，可以将目标采集器当前采集频率对应的采集周期作为统计间隔，例如当前采集频率对应的采集周期为 1s，则统计间隔为 1s，也就是每 1s 采集一个关键性能指标数据，并计算相邻两次采集到的关键性能指标数据的变化率；或者，也可以将多个采集周期作为统计间隔，例如，可以将 10 个采集周期作为统计间隔，即统计间隔为 10s，则每 10s 计算一次关键性能指标数据的变化率，可以根据在这 10s 内采集到的 10 次关键性能指标数据，计算关键性能指标数据在这 10s 内的变化率。

为了便于描述，使用 P_i 表示各关键性能指标， $\Delta(P_i)$ 表示关键指标 P_i 相邻两个统计间隔的性能取值的变化量，并为不同的关键性能指标赋予不同的权重，用 $(W_1, W_2 \dots W_n)$ 表示，其中 W_i 代表第 i 个关键性能指标的权重，那么根据各关键性能指标数据的变化率生成的全局变化率可以表示为： $\Delta = \sum_{i=1}^n W_i \Delta(P_i)$ 。需要说明的是，本实施例会设定每个关键性能指标的统计阈值，统计阈值用 $(PT_1, PT_2, \dots PT_n)$ 表示，其中 PT_i 表示第 i 个关键性能指标最低变化阈值，基于此，可以通过判断在一个判断周期内，各关键性能指标的变化量是否超过该阈值，若变化量超过该阈值，统计各关键性能指标数据的变化率，否则若两个相邻周期的关键性能指标数据的变化量未超过对应的阈值，说明关键性能指标数据的变化比较小，可以直接将变化方向设置为 0，表示不需要调频，即变频方向为保持不变。

在统计出各关键性能指标数据的变化率的情况下，根据上述加权求和公式，生成全局变化率。进一步，根据全局变化率，以利用变频策略，确定该关联采集器分组对应的变频方向，其一种具体的实施方式如下：将全局变化率作为变频策略的输入，并根据输出值判定该关联采集器分组对应的变频方向。其中，在输出值为 0 时，表示不改变目标采集器的频率；在输出值为 1 时，表示需要提高目标采集器的采集频率；在输出值为 -1 时，表示需要降低目标采集器的采集频率。进一步可选地，利用变频策略，确定输出值，一种具体的实施方式如下：计算该关联采集器分组中关键指标数据的加权变化率，为了便于描述，使用 $KeyDelta$ 表示关键指标数据的加权变化率，变化率阈值的上下限为 (B_1, B_2) ，其中 $B_1 \leq B_2$ ，当 $KeyDelta > B_2$ 时，需要提高采集频率，输出值为 1；当 $KeyDelta < B_1$ 时，需要降低采集频率，输出值为 -1；当 $B_1 \leq KeyDelta \leq B_2$ ，保持原有采集频率，输出值为 0。

在一可选实施例中，为了进一步提高变频方向的精确性，针对每个关联采集器分组，根据该关联采集器分组内目标采集器采集到的性能指标数据的变化信息，确定该关联采集器分组对应的变频方向，另一种具体的实施方式如下：针对每个关联采集器分组，根据该关联采集器分组内目标采集器采集到的性能指标数据的变化信息，以及最近一次根据至少两个目标采集器采集到的性能指标数据得到的性能分析结果，确定该关联采集器分组对应的变频方向。进一步可选地，针对每个关联采集器分组，根据该关联采集器分组内目标采集器采集到的性能指标数据的变化信息，确定该关联采集器分组对应的第一变频方向，以及根据最近一次根据至少两个目标采集器采集到的性能指标数据得到的分析结果，确定该关联采集器分组对应的第二变频方向，若第一变频方向和第二变频方向相同，则确定第一变频方向为该关联采集器分组对应的变频方向；若第一变频方向和第二变频方向不相同，则可以暂时不调整当前采集频率，可以连续统计多个统计间隔内的第一变频方向，根据连续多个统计间隔内的第一变频方向，最终确定该关联采集器分组对应的变频方向。

需要说明的是，针对每个关联采集器分组，在调整该关联采集器分组内目标采集器的采集频率的过程中，若该关联采集器分组内目标采集器的当前采集频率为最大的预设频率，且变频方向为增大频率的，保持当前采集频率不变；若该关联采集器分组内各目标采集器的当前采集频率为最小的预设频率，且变频方向为减小频率，保持当前采集频率不变。

- 5 在此说明，在本实施例中，每个计算节点 102 都可以对本地至少两个目标采集器进行分组，且各计算节点 102 按照同样的标准对目标采集器进行分组。这样，主节点 103 可以以关联采集器分组为单位，当某个关联采集器分组的变频方向为增大频率或减小频率的情况下，可以向其它计算节点 102 发送通知消息，在该通知消息中携带增大频率或减小频率的指示信息，这样其它计算节点 102 根据该指示信息，可以将对应关联采集器分组中的目标采集器的采样频率进行增大或减小，具体可以增大或减小为该变频方向上最接近的预设频率。进一步可选地，各计算节点 102 在将至少两个目标采集器划分为至少两个关联采集器分组之后，也可以按照同样的方式，将每个关联采集器分组中各目标采集器的当前采集频率进行统一化，这样不同计算节点 102 上的同样关联采集器分组的当前采集频率是一样的，从而保证各计算节点 102 上同样关联采集器分组中的目标采集器可以同时增大或者同时减小，确保使用同样的采集频率，确保不同计算节点 102 上相同采集器的采集频率的一致性，方便后续对同一性能指标数据进行分析处理。

- 本申请上述各实施例中，在计算集群 100 场景中，根据性能指标数据的变化信息，自适应的改变性能指标数据的采集频率，既能保证采集精度，保障基于性能指标数据的性能分析和基于分析结果的决策准确性，又能减少采集开销；在自适应改变采集频率过程中，针对执行同一作业任务的至少两个计算节点 102，由至少两个计算节点 102 中的主节点 103 负责采集频率的适应性变化处理并在需要改变的情况下同步给其它计算节点 102，其它计算节点 102 无需负责采集频率的自适应变化处理，可减轻其它计算节点 102 的采集开销，进一步降低整体的采集开销。

- 在上述实施例中，描述了同一作业任务被部署在至少两个计算节点 102 上的情况下，作为主节点 103 的计算节点 102 如何对进行采集频率调整的功能实现。除此之外，同一作业任务也可能被部署在一个计算节点 102 上，在该情况下，可以不用选择主节点 103，则执行作业任务的计算节点 102 可以采用下述方式自行调节本地目标采集器的采集频率，具体包括：在执行作业任务过程中，启动与作业任务相关的至少两个目标采集器，以使至少两个目标采集器以当前采集频率采集其所在计算节点的至少两种性能指标数据；根据至少两个目标采集器负责采集的性能指标的关联性，将至少两个目标采集器划分至至少两个关联采集器分组中；根据每个关联采集器分组内目标采集器采集到的性能指标数据的变化信息，分别调整每个关联采集器分组内目标采集器的采集频率，以使每个关联采集器分组内的目标采集器以调整后的采集频率继续采集其所在计算节点的至少两种性能指标数据。各步骤的详细实现方式，可参见前述实施例的描述，在此不再赘述。

图 2 为本申请一示例性实施例提供的用于计算集群 100 的数据采集方法的流程示意图。如图 2 所示，该方法包括：

201、在执行作业任务过程中，启动与所述作业任务相关的至少两个目标采集器，以使所述至少两个目标采集器以当前采集频率采集其所在计算节点的至少两种性能指标

数据, 所述作业任务被部署在计算集群 100 中的至少两个计算节点上;

202、在确定自身是所述至少两个计算节点中的主节点的情况下, 根据所述至少两种性能指标数据的变化信息, 调整所述至少两个目标采集器的采集频率;

203、通知所述至少两个计算节点中的其它计算节点调整其上部署的所述至少两个目标采集器的采集频率, 以使其它计算节点上的至少两个目标采集器以调整后的采集频率继续采集其所在计算节点的至少两种性能指标数据。

在本实施例中, 根据所述至少两种性能指标数据的变化信息, 调整所述至少两个目标采集器的采集频率, 包括: 根据所述至少两个目标采集器负责采集的性能指标的关联性, 将所述至少两个目标采集器划分至至少两个关联采集器分组中; 根据每个关联采集器分组内目标采集器采集到的性能指标数据的变化信息, 分别调整每个关联采集器分组内目标采集器的采集频率。

在一可选实施例中, 根据每个关联采集器分组内目标采集器采集到的性能指标数据的变化信息, 分别调整每个关联采集器分组内目标采集器的采集频率, 包括: 针对每个关联采集器分组, 根据该关联采集器分组内目标采集器采集到的性能指标数据的变化信息, 确定该关联采集器分组对应的变频方向; 将该关联采集器分组内目标采集器当前使用的采集频率, 调整为多个预设频率中在该变频方向上最接近的预设频率, 其中, 多个预设频率由小到大。

进一步可选地, 针对每个关联采集器分组, 在调整该关联采集器分组内目标采集器的采集频率的过程中, 还包括: 若该关联采集器分组内目标采集器的当前采集频率为最大的预设频率, 且变频方向为增大频率的, 保持当前采集频率不变; 若该关联采集器分组内各目标采集器的当前采集频率为最小的预设频率, 且变频方向为减小频率, 保持当前采集频率不变。

进一步可选地, 针对每个关联目标采集器, 根据该关联采集器分组内目标采集器采集到的性能指标数据的变化信息, 确定该关联采集器分组对应的变频方向之前, 还包括: 针对每个关联采集器分组, 若该关联采集器分组内各目标采集器的当前采集频率不相同, 将该关联采集器分组内各目标采集器的当前采集频率调整为相同的采集频率。

在一可选实施例中, 将该关联采集器分组内各目标采集器的当前采集频率调整为相同的采集频率, 包括: 将该关联采集器分组内各目标采集器的当前采集频率均调整为该关联采集器分组内各目标采集器的当前采集频率的平均值; 或者, 将该关联采集器分组内各目标采集器的当前采集频率均调整为该关联采集器分组内各目标采集器的当前采集频率中的最大采集频率; 或者, 将该关联目标采集器内各目标采集器的当前采集频率均调整为该关联采集器分组内各目标采集器的当前采集频率中的最小采集频率。

在一可选实施例中, 针对每个关联采集器分组, 根据该关联采集器分组内目标采集器采集到的性能指标数据的变化信息, 确定该关联采集器分组对应的变频方向, 包括: 针对每个关联采集器分组, 获取该关联采集器分组内关键采集器采集到的关键性能指标数据, 所述关键采集器是负责采集关键性能指标的目标采集器; 根据设定的统计间隔, 统计各关键性能指标数据的变化率, 并根据各关键性能指标数据的变化率生成全局变化率; 根据所述全局变化率, 确定该关联采集器分组对应的变频方向, 所述变频方向包括增大频率、降低频率和保持不变中的任一种。

在一可选实施例中，针对每个关联采集器分组，根据该关联采集器分组内目标采集器采集到的性能指标数据的变化信息，确定该关联采集器分组对应的变频方向，包括：针对每个关联采集器分组，根据该关联采集器分组内目标采集器采集到的性能指标数据的变化信息，以及最近一次根据所述至少两个目标采集器采集到的性能指标数据得到的性能分析结果，确定该关联采集器分组对应的变频方向。

这里需要说明的是：本实施例提供的用于计算集群 100 的数据采集方法中的各步骤的具体实现的原理可参见上述计算集群 100 实施例中的相应内容，此处不再赘述。

图 3 为本申请一示例性实施例提供的另一用于计算集群 100 的数据采集方法的流程图。如图 3 所示，该方法包括：

301、在执行作业任务过程中，启动与所述作业任务相关的至少两个目标采集器，以使所述至少两个目标采集器以当前采集频率采集其所在计算节点的至少两种性能指标数据；

302、根据所述至少两个目标采集器负责采集的性能指标的关联性，将所述至少两个目标采集器划分至至少两个关联采集器分组中；

303、根据每个关联采集器分组内目标采集器采集到的性能指标数据的变化信息，分别调整每个关联采集器分组内目标采集器的采集频率，以使目标采集器以调整后的采集频率继续采集其所在计算节点的至少两种性能指标数据。

这里需要说明的是：本实施例提供的用于计算集群 100 的数据采集方法中的各步骤的具体实现的原理可参见上述计算集群 100 实施例中的相应内容，此处不再赘述。

图 4 为本申请一示例性实施例提供的数据采集装置的结构示意图。如图 4 所示，该装置包括：

启动模块 41，用于在执行作业任务过程中，启动与所述作业任务相关的至少两个目标采集器，以使所述至少两个目标采集器以当前采集频率采集其所在计算节点的至少两种性能指标数据，所述作业任务被部署在计算集群 100 中的至少两个计算节点上；

调整模块 42，用于在确定自身是所述至少两个计算节点中的主节点的情况下，根据所述至少两种性能指标数据的变化信息，调整所述至少两个目标采集器的采集频率；

通知模块 43，用于通知所述至少两个计算节点中的其它计算节点调整其上部署的所述至少两个目标采集器的采集频率，以使其它计算节点上的至少两个目标采集器以调整后的采集频率继续采集其所在计算节点的至少两种性能指标数据。

在一可选实施例中，调整模块 42 在用于根据所述至少两种性能指标数据的变化信息，调整所述至少两个目标采集器的采集频率时，具体用于：根据所述至少两个目标采集器负责采集的性能指标的关联性，将所述至少两个目标采集器划分至至少两个关联采集器分组中；根据每个关联采集器分组内目标采集器采集到的性能指标数据的变化信息，分别调整每个关联采集器分组内目标采集器的采集频率。

在一可选实施例中，调整模块 42 在用于根据每个关联采集器分组内目标采集器采集到的性能指标数据的变化信息，分别调整每个关联采集器分组内目标采集器的采集频率时，具体用于：针对每个关联采集器分组，根据该关联采集器分组内目标采集器采集到的性能指标数据的变化信息，确定该关联采集器分组对应的变频方向；将该关联采集器分组内目标采集器当前使用的采集频率，调整为多个预设频率中在该变频方向上最接近的预设频率，

其中，多个预设频率由小到大。

进一步可选地，调整模块 42 在用于针对每个关联采集器分组，在调整该关联采集器分组内目标采集器的采集频率的过程中时，还用于：若该关联采集器分组内目标采集器的当前采集频率为最大的预设频率，且变频方向为增大频率的，保持当前采集频率不变；若该关联采集器分组内各目标采集器的当前采集频率为最小的预设频率，且变频方向为减小频率，保持当前采集频率不变。

进一步可选地，调整模块 42 在用于针对每个关联目标采集器，根据该关联采集器分组内目标采集器采集到的性能指标数据的变化信息，确定该关联采集器分组对应的变频方向之前时，还用于：针对每个关联采集器分组，若该关联采集器分组内各目标采集器的当前采集频率不相同，将该关联采集器分组内各目标采集器的当前采集频率调整为相同的采集频率。

在一可选实施例中，调整模块 42 在用于将该关联采集器分组内各目标采集器的当前采集频率调整为相同的采集频率时，具体用于：将该关联采集器分组内各目标采集器的当前采集频率均调整为该关联采集器分组内各目标采集器的当前采集频率的平均值；或者，将该关联采集器分组内各目标采集器的当前采集频率均调整为该关联采集器分组内各目标采集器的当前采集频率中的最大采集频率；或者，将该关联目标采集器内各目标采集器的当前采集频率均调整为该关联采集器分组内各目标采集器的当前采集频率中的最小采集频率。

在一可选实施例中，调整模块 42 在用于针对每个关联采集器分组，根据该关联采集器分组内目标采集器采集到的性能指标数据的变化信息，确定该关联采集器分组对应的变频方向时，具体用于：针对每个关联采集器分组，获取该关联采集器分组内关键采集器采集到的关键性能指标数据，所述关键采集器是负责采集关键性能指标的目标采集器；根据设定的统计间隔，统计各关键性能指标数据的变化率，并根据各关键性能指标数据的变化率生成全局变化率；根据所述全局变化率，确定该关联采集器分组对应的变频方向，所述变频方向包括增大频率、降低频率和保持不变中的任一种。

在一可选实施例中，调整模块 42 在用于针对每个关联采集器分组，根据该关联采集器分组内目标采集器采集到的性能指标数据的变化信息，确定该关联采集器分组对应的变频方向时，具体用于：针对每个关联采集器分组，根据该关联采集器分组内目标采集器采集到的性能指标数据的变化信息，以及最近一次根据所述至少两个目标采集器采集到的性能指标数据得到的性能分析结果，确定该关联采集器分组对应的变频方向。

这里需要说明的是：本实施例提供的数据采集装置可实现上述图 2 方法实施例中描述的技术方案，上述各模块或单元具体实现的原理可参见上述图 1a 所示计算集群 100 实施例及图 2 方法实施例中的相应内容，此处不再赘述。

图 5 为本申请一示例性实施例提供的另一数据采集装置的结构示意图。如图 5 所示，该装置包括：

启动模块 51，用于在执行作业任务过程中，启动与所述作业任务相关的至少两个目标采集器，以使所述至少两个目标采集器以当前采集频率采集其所在计算节点的至少两种性能指标数据；

分组模块 52，用于根据所述至少两个目标采集器负责采集的性能指标的关联性，将所

述至少两个目标采集器划分至至少两个关联采集器分组中；

调整模块 53，用于根据每个关联采集器分组内目标采集器采集到的性能指标数据的变化信息，分别调整每个关联采集器分组内目标采集器的采集频率，以使目标采集器以调整后的采集频率继续采集其所在计算节点的至少两种性能指标数据。

- 5 这里需要说明的是：本实施例提供的数据采集装置可实现上述图 3 方法实施例中描述的技术方案，上述各模块或单元具体实现的原理可参见上述图 1a 所示计算集群 100 实施例及图 3 方法实施例中的相应内容，此处不再赘述。

图 6 为本申请一示例性实施例提供的计算节点的结构示意图。该计算节点可作为上述计算集群 100 中的任一计算节点，如图 6 所示，该计算节点包括：存储器 60a 和处理器 60b；
10 所述存储器 60a，用于存储计算机程序；所述处理器 60b，与所述存储器 60a 耦合，用于执行所述计算机程序，以用于：

在执行作业任务过程中，启动与作业任务相关的至少两个目标采集器，以使所述至少两个目标采集器以当前采集频率采集其所在计算节点的至少两种性能指标数据，所述作业任务被部署在计算集群 100 中的至少两个计算节点上；

- 15 在确定自身是所述至少两个计算节点中的主节点的情况下，根据所述至少两种性能指标数据的变化信息，调整所述至少两个目标采集器的采集频率；

通知所述至少两个计算节点中的其它计算节点调整其上部署的所述至少两个目标采集器的采集频率，以使其它计算节点上的至少两个目标采集器以调整后的采集频率继续采集其所在计算节点的至少两种性能指标数据。

- 20 在一可选实施例中，处理器 60b 根据所述至少两种性能指标数据的变化信息，调整所述至少两个目标采集器的采集频率时，具体用于：根据所述至少两个目标采集器负责采集的性能指标的关联性，将所述至少两个目标采集器划分至至少两个关联采集器分组中；根据每个关联采集器分组内目标采集器采集到的性能指标数据的变化信息，分别调整每个关联采集器分组内目标采集器的采集频率。

- 25 在一可选实施例中，处理器 60b 根据每个关联采集器分组内目标采集器采集到的性能指标数据的变化信息，分别调整每个关联采集器分组内目标采集器的采集频率时，具体用于：针对每个关联采集器分组，根据该关联采集器分组内目标采集器采集到的性能指标数据的变化信息，确定该关联采集器分组对应的变频方向；将该关联采集器分组内目标采集器当前使用的采集频率，调整为多个预设频率中在该变频方向上最接近的预设频率，其中，多个预设频率由小到大。
30

- 在一可选实施例中，处理器 60b 在针对每个关联采集器分组，在调整该关联采集器分组内目标采集器的采集频率的过程中，还用于：若该关联采集器分组内目标采集器的当前采集频率为最大的预设频率，且变频方向为增大频率的，保持当前采集频率不变；若该关联采集器分组内各目标采集器的当前采集频率为最小的预设频率，且变频方向为减小频率，保持当前采集频率不变。
35

在一可选实施例中，处理器 60b 在针对每个关联目标采集器，根据该关联采集器分组内目标采集器采集到的性能指标数据的变化信息，确定该关联采集器分组对应的变频方向之前，还用于：针对每个关联采集器分组，若该关联采集器分组内各目标采集器的当前采集频率不相同，将该关联采集器分组内各目标采集器的当前采集频率调

整为相同的采集频率。

在一可选实施例中，处理器 60b 在将该关联采集器分组内各目标采集器的当前采集频率调整为相同的采集频率时，具体用于：将该关联采集器分组内各目标采集器的当前采集频率均调整为该关联采集器分组内各目标采集器的当前采集频率的平均值；

- 5 或者，将该关联采集器分组内各目标采集器的当前采集频率均调整为该关联采集器分组内各目标采集器的当前采集频率中的最大采集频率；或者，将该关联目标采集器内各目标采集器的当前采集频率均调整为该关联采集器分组内各目标采集器的当前采集频率中的最小采集频率。

- 10 在一可选实施例中，处理器 60b 在针对每个关联采集器分组，根据该关联采集器分组内目标采集器采集到的性能指标数据的变化信息，确定该关联采集器分组对应的变频方向时，具体用于：针对每个关联采集器分组，获取该关联采集器分组内关键采集器采集到的关键性能指标数据，所述关键采集器是负责采集关键性能指标的目标采集器；根据设定的统计间隔，统计各关键性能指标数据的变化率，并根据各关键性能指标数据的变化率生成全局变化率；根据所述全局变化率，确定该关联采集器分组对
- 15 应的变频方向，所述变频方向包括增大频率、降低频率和保持不变中的任一种。

- 在一可选实施例中，处理器 60b 在针对每个关联采集器分组，根据该关联采集器分组内目标采集器采集到的性能指标数据的变化信息，确定该关联采集器分组对应的变频方向时，具体用于：针对每个关联采集器分组，根据该关联采集器分组内目标采集器采集到的性能指标数据的变化信息，以及最近一次根据所述至少两个目标采集器
- 20 采集到的性能指标数据得到的性能分析结果，确定该关联采集器分组对应的变频方向。

在上述实施例中，描述了同一作业任务被部署在至少两个计算节点上的情况下，作为主节点的计算节点如何对进行采集频率调整的功能实现。除此之外，同一作业任务也可能被部署在一个计算节点上，在该情况下，可以不用选择主节点，执行作业任务的计算节点的处理器 60b 还可以执行以下操作：

- 25 在执行作业任务过程中，启动与所述作业任务相关的至少两个目标采集器，以使所述至少两个目标采集器以当前采集频率采集其所在计算节点的至少两种性能指标数据；

根据所述至少两个目标采集器负责采集的性能指标的关联性，将所述至少两个目标采集器划分至至少两个关联采集器分组中；

- 30 根据每个关联采集器分组内目标采集器采集到的性能指标数据的变化信息，分别调整每个关联采集器分组内目标采集器的采集频率，以使目标采集器以调整后的采集频率继续采集其所在计算节点的至少两种性能指标数据。

进一步地，如图 6 所示，该电子设备还包括：通信组件 60c、电源组件 60d、等其它组件。图 6 中仅示意性给出部分组件，并不意味着电子设备只包括图 6 所示组件。

- 35 这里需要说明的是：本实施例提供的计算节点可实现上述图 2 或图 3 方法实施例中描述的技术方案，上述各模块或单元具体实现的原理可参见上述图 1a 所示计算集群 100 实施例中的相应内容，此处不再赘述。

本申请一示例性实施例提供存储有计算机程序/指令的计算机可读存储介质，当计算机程序/指令被处理器执行时，致使处理器能够实现以上所述方法中的步骤，此处不再赘述。

本申请一示例性实施例提供一种计算机程序产品，包括计算机程序/指令，当计算机程序/指令被处理器执行时，致使处理器能够实现以上所述方法中的步骤，此处不再赘述。

上述实施例中的通信组件被配置为便于通信组件所在节点和其他节点之间有线或无线方式的通信。通信组件所在节点可以接入基于通信标准的无线网络，如 WiFi，2G、3G、4G/LTE、5G 等移动通信网络，或它们的组合。在一个示例性实施例中，通信组件经由广播信道接收来自外部广播管理系统的广播信号或广播相关信息。在一个示例性实施例中，所述通信组件还包括近场通信（NFC）模块，以促进短程通信。例如，在 NFC 模块可基于射频识别（RFID）技术，红外数据协会（IrDA）技术，超宽带（UWB）技术，蓝牙（BT）技术和其他技术来实现。

上述实施例中的显示器包括屏幕，其屏幕可以包括液晶显示器（LCD）和触摸面板（TP）。如果屏幕包括触摸面板，屏幕可以被实现为触摸屏，以接收来自用户的输入信号。触摸面板包括一个或多个触摸传感器以感测触摸、滑动和触摸面板上的手势。所述触摸传感器可以不仅感测触摸或滑动动作的边界，而且还检测与所述触摸或滑动操作相关的持续时间和压力。

上述实施例中的电源组件，为电源组件所在节点的各种组件提供电力。电源组件可以包括电源管理系统，一个或多个电源，及其他与为电源组件所在节点生成、管理和分配电力相关联的组件。

上述实施例中的音频组件，可被配置为输出和/或输入音频信号。例如，音频组件包括一个麦克风（MIC），当音频组件所在节点处于操作模式，如呼叫模式、记录模式和语音识别模式时，麦克风被配置为接收外部音频信号。所接收的音频信号可以被进一步存储在存储器或经由通信组件发送。在一些实施例中，音频组件还包括一个扬声器，用于输出音频信号。

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。

本申请是参照根据本申请实施例的方法、节点（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节点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节点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。

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节点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。

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节点上，使得在计算机

或其他可编程节点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节点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。

在一个典型的配置中，计算节点包括一个或多个处理器 (CPU)、输入/输出接口、网络接口和内存。

内存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)。内存是计算机可读介质的示例。

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，磁带磁盘存储或其他磁性存储节点或任何其他非传输介质，可用于存储可以被计算节点访问的信息。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。

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、商品或者节点不仅包括那些要素，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、商品或者节点所固有的要素。在没有更多限制的情况下，由语句“包括一个……”限定的要素，并不排除在包括所述要素的过程、方法、商品或者节点中还存在另外的相同要素。

以上所述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。

权 利 要 求 书

1、一种计算集群，其特征在于，包括：管控节点和多个计算节点，每个计算节点上部署有多个采集器，不同采集器用于采集不同的性能指标；

所述管控节点，用于在所述多个计算节点中的至少两个计算节点上部署同一作业任务，并控制所述至少两个计算节点执行所述作业任务；

每个计算节点，用于在执行所述作业任务过程中，启动与所述作业任务相关的至少两个目标采集器，以使所述至少两个目标采集器以当前采集频率采集其所在计算节点的至少两种性能指标数据；

在确定自身是所述至少两个计算节点中的主节点的情况下，根据所述至少两种性能指标数据的变化信息，调整所述至少两个目标采集器的采集频率；

通知所述至少两个计算节点中其它计算节点调整其上至少两个目标采集器的采集频率，以使其它计算节点上的至少两个目标采集器以调整后的采集频率继续采集其在计算节点的至少两种性能指标数据。

2、一种用于计算集群的数据采集方法，应用于计算集群中的任一计算节点，其特征在于，所述方法包括：

在执行作业任务过程中，启动与所述作业任务相关的至少两个目标采集器，以使所述至少两个目标采集器以当前采集频率采集其所在计算节点的至少两种性能指标数据，所述作业任务被部署在计算集群中的至少两个计算节点上；

在确定自身是所述至少两个计算节点中的主节点的情况下，根据所述至少两种性能指标数据的变化信息，调整所述至少两个目标采集器的采集频率；

通知所述至少两个计算节点中的其它计算节点调整其上部署的所述至少两个目标采集器的采集频率，以使其它计算节点上的至少两个目标采集器以调整后的采集频率继续采集其所在计算节点的至少两种性能指标数据。

3、根据权利要求2所述的方法，其特征在于，根据所述至少两种性能指标数据的变化信息，调整所述至少两个目标采集器的采集频率，包括：

根据所述至少两个目标采集器负责采集的性能指标的关联性，将所述至少两个目标采集器划分至至少两个关联采集器分组中；

根据每个关联采集器分组内目标采集器采集到的性能指标数据的变化信息，分别调整每个关联采集器分组内目标采集器的采集频率。

4、根据权利要求3所述的方法，其特征在于，根据每个关联采集器分组内目标采集器采集到的性能指标数据的变化信息，分别调整每个关联采集器分组内目标采集器的采集频率，包括：

针对每个关联采集器分组，根据该关联采集器分组内目标采集器采集到的性能指标数据的变化信息，确定该关联采集器分组对应的变频方向；

将该关联采集器分组内目标采集器当前使用的采集频率，调整为多个预设频率中在该变频方向上最接近的预设频率，其中，多个预设频率由小到大。

5、根据权利要求4所述的方法，其特征在于，针对每个关联采集器分组，在调整该关联采集器分组内目标采集器的采集频率的过程中，还包括：

若该关联采集器分组内目标采集器的当前采集频率为最大的预设频率，且变频方

向为增大频率的，保持当前采集频率不变；

若该关联采集器分组内各目标采集器的当前采集频率为最小的预设频率，且变频方向为减小频率，保持当前采集频率不变。

- 6、根据权利要求4所述的方法，其特征在于，针对每个关联目标采集器，根据该关联采集器分组内目标采集器采集到的性能指标数据的变化信息，确定该关联采集器分组对应的变频方向之前，还包括：

针对每个关联采集器分组，若该关联采集器分组内各目标采集器的当前采集频率不相同，将该关联采集器分组内各目标采集器的当前采集频率调整为相同的采集频率。

- 7、根据权利要求6所述的方法，其特征在于，将该关联采集器分组内各目标采集器的当前采集频率调整为相同的采集频率，包括：

将该关联采集器分组内各目标采集器的当前采集频率均调整为该关联采集器分组内各目标采集器的当前采集频率的平均值；

或者

- 将该关联采集器分组内各目标采集器的当前采集频率均调整为该关联采集器分组内各目标采集器的当前采集频率中的最大采集频率；

或者

将该关联目标采集器内各目标采集器的当前采集频率均调整为该关联采集器分组内各目标采集器的当前采集频率中的最小采集频率。

- 8、根据权利要求4所述的方法，其特征在于，针对每个关联采集器分组，根据该关联采集器分组内目标采集器采集到的性能指标数据的变化信息，确定该关联采集器分组对应的变频方向，包括：

针对每个关联采集器分组，获取该关联采集器分组内关键采集器采集到的关键性能指标数据，所述关键采集器是负责采集关键性能指标的目标采集器；

- 根据设定的统计间隔，统计各关键性能指标数据的变化率，并根据各关键性能指标数据的变化率生成全局变化率；

根据所述全局变化率，确定该关联采集器分组对应的变频方向，所述变频方向包括增大频率、降低频率和保持不变中的任一种。

- 9、根据权利要求4-8任一项所述的方法，其特征在于，针对每个关联采集器分组，根据该关联采集器分组内目标采集器采集到的性能指标数据的变化信息，确定该关联采集器分组对应的变频方向，包括：

针对每个关联采集器分组，根据该关联采集器分组内目标采集器采集到的性能指标数据的变化信息，以及最近一次根据所述至少两个目标采集器采集到的性能指标数据得到的性能分析结果，确定该关联采集器分组对应的变频方向。

- 10、一种用于计算集群的数据采集方法，应用于计算集群中的任一计算节点，其特征在于，所述方法包括：

在执行作业任务过程中，启动与所述作业任务相关的至少两个目标采集器，以使所述至少两个目标采集器以当前采集频率采集其所在计算节点的至少两种性能指标数据；

根据所述至少两个目标采集器负责采集的性能指标的关联性，将所述至少两个目

标采集器划分至至少两个关联采集器分组中；

根据每个关联采集器分组内目标采集器采集到的性能指标数据的变化信息，分别调整每个关联采集器分组内目标采集器的采集频率，以使目标采集器以调整后的采集频率继续采集其所在计算节点的至少两种性能指标数据。

5 11、一种数据采集装置，应用于计算集群中的任一计算节点，其特征在于，所述装置包括：

启动模块，用于在执行作业任务过程中，启动与所述作业任务相关的至少两个目标采集器，以使所述至少两个目标采集器以当前采集频率采集其所在计算节点的至少两种性能指标数据，所述作业任务被部署在计算集群中的至少两个计算节点上；

10 调整模块，用于在确定自身是所述至少两个计算节点中的主节点的情况下，根据所述至少两种性能指标数据的变化信息，调整所述至少两个目标采集器的采集频率；

通知模块，用于通知所述至少两个计算节点中的其它计算节点调整其上部署的所述至少两个目标采集器的采集频率，以使其它计算节点上的至少两个目标采集器以调整后的采集频率继续采集其所在计算节点的至少两种性能指标数据。

15 12、一种数据采集装置，应用于计算集群中的任一计算节点，其特征在于，所述装置包括：

启动模块，用于在执行作业任务过程中，启动与所述作业任务相关的至少两个目标采集器，以使所述至少两个目标采集器以当前采集频率采集其所在计算节点的至少两种性能指标数据；

20 分组模块，用于根据所述至少两个目标采集器负责采集的性能指标的关联性，将所述至少两个目标采集器划分至至少两个关联采集器分组中；

调整模块，用于根据每个关联采集器分组内目标采集器采集到的性能指标数据的变化信息，分别调整每个关联采集器分组内目标采集器的采集频率，以使目标采集器以调整后的采集频率继续采集其所在计算节点的至少两种性能指标数据。

25 13、一种计算节点，可应用于计算集群，其特征在于，所述计算节点包括：存储器和处理器；所述存储器，用于存储计算机程序；所述处理器，与所述存储器耦合，用于执行所述计算机程序，以用于执行权利要求 2-10 任一项所述方法中的步骤。

30 14、一种存储有计算机程序的计算机可读存储介质，其特征在于，当所述计算机程序被处理器执行时，致使所述处理器能够实现权利要求 2-10 任一项所述方法中的步骤。