



(21) 申请号 202210203493.3

(22) 申请日 2022.03.02

(71) 申请人 上海交通大学

地址 200240 上海市闵行区东川路800号

(72) 发明人 林圣凯 赵世振 曹培睿 韩昕驰
刘运卓

(74) 专利代理机构 上海汉声知识产权代理有限公司 31236

专利代理师 胡晶

(51) Int. Cl.

G06F 9/50 (2006.01)

G06F 9/46 (2006.01)

G06F 30/27 (2020.01)

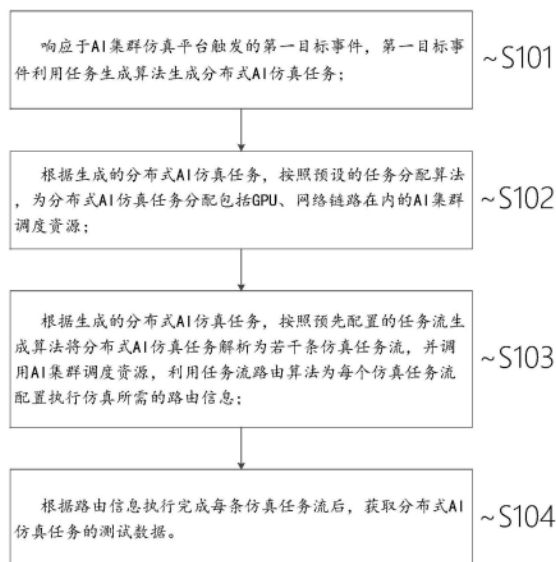
权利要求书1页 说明书5页 附图2页

(54) 发明名称

一种用于分布式AI集群的网络仿真方法及装置

(57) 摘要

本发明公开了一种用于分布式AI集群的网络仿真方法及装置,方法包括:响应于AI集群仿真平台触发生成分布式AI仿真任务;根据生成的分布式AI仿真任务,按照预设的任务分配算法,为分布式AI仿真任务分配AI集群调度资源;根据生成的分布式AI仿真任务,按照预先配置的任务流生成算法将分布式AI仿真任务解析为若干条仿真任务流,并调用AI集群调度资源,利用任务流路由算法为每个仿真任务流配置执行仿真所需的路由信息;根据路由信息执行完成每条仿真任务流后,获取分布式AI仿真任务的测试数据。本发明通过AI集群仿真平台封装通用网络仿真器的底层逻辑框架,进行任意的AI集群仿真任务,提高了AI集群仿真效率,扩展了仿真研究的广度。



1. 一种用于分布式AI集群的网络仿真方法,其特征在于,所述方法包括:

响应于AI集群仿真平台触发的第一目标事件,所述第一目标事件利用任务生成算法生成分布式AI仿真任务;

根据生成的分布式AI仿真任务,按照预设的任务分配算法,为所述分布式AI仿真任务分配包括GPU、网络链路在内的AI集群调度资源;

根据生成的分布式AI仿真任务,按照预先配置的任务流生成算法将所述分布式AI仿真任务解析为若干条仿真任务流,并调用所述AI集群调度资源,利用任务流路由算法为每个仿真任务流配置执行仿真所需的路由信息;

根据路由信息执行完成每条仿真任务流后,获取分布式AI仿真任务的测试数据。

2. 如权利要求1所述的用于分布式AI集群的网络仿真方法,其特征在于,所述AI集群仿真平台配置有若干数据接口,通过各所述数据接口获取包括任务生成算法、任务分配算法、任务流生成算法、任务流路由算法在内的自定义AI集群仿真信息。

3. 如权利要求1所述的用于分布式AI集群的网络仿真方法,其特征在于,所述AI集群调度资源包括GPU在内的计算资源、网络链路在内的网络资源。

4. 如权利要求1所述的用于分布式AI集群的网络仿真方法,其特征在于,所述任务流生成算法包括但不限于ring allreduce算法、butterfly算法、double binary tree算法。

5. 如权利要求1所述的用于分布式AI集群的网络仿真方法,其特征在于,所述任务流路由算法包括但不限于dijkstra算法。

6. 一种用于分布式AI集群的网络仿真装置,采用权利要求1-5任意一项所述的方法,其特征在于,包括:

任务生成模块,配置为响应于AI集群仿真平台触发的第一目标事件,所述第一目标事件利用任务生成算法生成分布式AI仿真任务;

资源调度模块,配置为根据生成的分布式AI仿真任务,按照预设的任务分配算法,为所述分布式AI仿真任务分配包括GPU、网络链路在内的AI集群调度资源;

路由分配模块,配置为根据生成的分布式AI仿真任务,按照预先配置的任务流生成算法将所述分布式AI仿真任务解析为若干条仿真任务流,并调用所述AI集群调度资源,利用任务流路由算法为每个仿真任务流配置执行仿真所需的路由信息;

仿真执行模块,配置根据路由信息执行完成每条仿真任务流后,获取分布式AI仿真任务的测试数据。

一种用于分布式AI集群的网络仿真方法及装置

技术领域

[0001] 本发明涉及仿真技术领域,尤其涉及一种用于分布式AI集群的网络仿真方法及装置。

背景技术

[0002] 随着研究人员对大规模神经网络的训练需求不断提升,单机训练已无法满足训练需求,因此采用多机多卡分布式训练。现采用云服务商提供的大规模AI集群进行分布式神经网络训练,因此对数据中心的多租户AI集群的研究日渐受到关注。此外,随着GPU类专用设备的发展,训练的计算过程被大大提速,分布式神经网络训练的性能瓶颈已经逐渐转移到网络侧,因此对于多租户AI集群的网络侧研究也受到了大量研究者的关注。目前,对于数据中心网络的研究几乎都是在网络仿真平台上进行的,并且在搭建实际系统进行验证之前,也需要通过网络仿真的验证,因此,对于分布式AI集群的研究也不例外。

[0003] 但是,目前网络仿真框架(例如omnet,ns3等)都是包级别或者流级别,均是针对通用用途的网络仿真,比如,对各种网络、各种任务的仿真。可以看出,由于包级别或者流级别的仿真粒度很细,直接仿真分布式AI集群十分复杂、不便。究其本质,目前网络仿真框架只是提供了对于AI集群数据面的仿真,而未涉及对AI集群控制面(比如,AI任务生成、集群资源调度、任务流生成、任务流路由等功能)。但目前很多对于AI集群的研究都需要AI集群控制面的仿真功能,于是如果直接使用上述的网络仿真框架对AI集群进行研究,则需要手动编写整个控制面,因此整个过程会很复杂、不便。

发明内容

[0004] 本申请实施例通过提供一种用于分布式AI集群的网络仿真方法及装置,解决了现有技术中没有专门针对AI集群任务仿真的技术方案,解决了AI集群仿真时过程复杂操作不便的技术问题,实现了简化了分布式AI集群仿真操作流程,提高了研究效率以及扩展了分布式AI集群的仿真广度的效果。

[0005] 第一方面,本申请实施例提供了一种用于分布式AI集群的网络仿真方法,所述方法包括:

[0006] 响应于AI集群仿真平台触发的第一目标事件,所述第一目标事件利用任务生成算法生成分布式AI仿真任务;

[0007] 根据生成的分布式AI仿真任务,按照预设的任务分配算法,为所述分布式AI仿真任务分配包括GPU、网络链路在内的AI集群调度资源;

[0008] 根据生成的分布式AI仿真任务,按照预先配置的任务流生成算法将所述分布式AI仿真任务解析为若干条仿真任务流,并调用所述AI集群调度资源,利用任务流路由算法为每个仿真任务流配置执行仿真所需的路由信息;

[0009] 根据路由信息执行完成每条仿真任务流后,获取分布式AI仿真任务的测试数据。

[0010] 在一种实施例中,所述AI集群仿真平台配置有若干数据接口,通过各所述数据接

口获取包括任务生成算法、任务分配算法、任务流生成算法、任务流路由算法在内的自定义AI集群信息。

[0011] 在一种实施例中,所述AI集群调度资源包括GPU在内的计算资源、网络链路在内的网络资源。

[0012] 在一种实施例中,所述任务流生成算法包括但不限于ring allreduce算法、butterfly算法、double binary tree算法。

[0013] 在一种实施例中,所述任务流路由算法包括但不限于dijkstra算法。

[0014] 第二方面,本申请实施例提供了一种用于分布式AI集群的网络仿真装置,采用第一方面任意一项所述的方法,包括:

[0015] 任务生成模块,配置为响应于AI集群仿真平台触发的第一目标事件,所述第一目标事件利用任务生成算法生成的分布式AI仿真任务;

[0016] 资源调度模块,配置为根据生成的分布式AI仿真任务,获取任务资源调度信息,按照预设的任务分配算法,为所述分布式AI仿真任务分配包括GPU、网络链路在内的AI集群调度资源;

[0017] 路由分配模块,配置为根据生成的分布式AI仿真任务,按照预先配置的任务流生成算法将所述分布式AI仿真任务解析为若干条仿真任务流,并调用所述AI集群调度资源,利用任务流路由算法为每个仿真任务流配置执行仿真所需的路由信息;

[0018] 仿真执行模块,配置根据路由信息执行完成每条仿真任务流后,获取分布式AI仿真任务的测试数据。

[0019] 本申请实施例中提供的技术方案,至少具有如下技术效果:

[0020] 本发明以通用网络仿真框架为基础,进行任意的AI集群仿真,采用多数据接口获取仿真配置,以实现分布式AI集群的仿真,可以确保网络仿真的准确性,提高了分布式AI集群仿真的开发和研究效率。本实施例中在AI集群仿真平台中设置若干数据接口,通过数据接口扩展自定义功能算法,使得仿真程序的灵活性大大提升,拓展了对于AI集群的研究广度。

附图说明

[0021] 图1为本申请实施例一中用于分布式AI集群的网络仿真方法的流程图;

[0022] 图2为本申请实施例二中用于分布式AI集群的网络仿真装置模块图。

具体实施方式

[0023] 为了更好的理解上述技术方案,下面将结合说明书附图以及具体的实施方式对上述技术方案进行详细的说明。

[0024] 实施例一

[0025] 本申请实施例提供了一种用于分布式AI集群的网络仿真方法,该方法包括如下步骤:

[0026] 步骤S101,响应于AI集群仿真平台触发的第一目标事件,所述第一目标事件利用任务生成算法生成分布式AI仿真任务。

[0027] 本实施例中的AI集群仿真平台采用通用网络仿真框架人工配置AI集群仿真环境,

并获取仿真任务执行后的测试数据,比如,任务流的执行时间,执行效率等。

[0028] 进一步地,所述AI集群仿真平台配置有若干数据接口,通过各所述数据接口获取包括任务生成算法、任务分配算法、任务流生成算法、任务流路由算法在内的自定义AI集群仿真信息。

[0029] 步骤S102,根据生成的分布式AI仿真任务,按照预设的任务分配算法,为所述分布式AI仿真任务分配包括GPU、网络链路在内的AI集群调度资源。

[0030] 本步骤中的所述AI集群调度资源包括GPU在内的计算资源、网络链路在内的网络资源。本实施例中的任务分配算法包括但不限于随机分配、顺序分配等算法。

[0031] 基于AI集群仿真平台中可以通过数据接口自定义配置AI集群调度资源。

[0032] 步骤S103,根据生成的分布式AI仿真任务,按照预先配置的任务流生成算法将所述分布式AI仿真任务解析为若干条仿真任务流,并调用所述AI集群调度资源,利用任务流路由算法为每个仿真任务流配置执行仿真所需的路由信息。

[0033] 本步骤中的所述任务流生成算法包括但不限于ring allreduce算法、butterfly算法、double binary tree算法。所述任务流路由算法包括但不限于dijkstra算法。

[0034] 进一步说明,在执行当前的仿真任务流时,可以同时生成新的分布式AI仿真任务,可以设定仿真环境中的使用数据并行、模型并行或混合并行。本实施例中对于数据并行可以设定任意的通信模式,比如,ring allreduce、butterfly、double binary tree等算法。在不冲突的前提下,可以使用任意算法,或者使用随机生成方式生成分布式AI仿真任务。

[0035] 本实施例中的任务分配算法、任务流生成算法可以是AI集群仿真平台预先设定的,也可以是通过数据接口自定义设定的。

[0036] 步骤S104,根据路由信息执行完成每条仿真任务流后,获取分布式AI仿真任务的测试数据。

[0037] 可以看出,本实施例给出的用于分布式AI集群的网络仿真方法,AI集群仿真平台以通用网络仿真框架为基础,可以进行任意的AI集群仿真,在确保网络仿真准确性以及目标仿真环境相同的前提下,提高了分布式AI集群仿真的开发和研究效率。本实施例中在AI集群仿真平台中设置若干数据接口,通过数据接口扩展自定义功能算法,使得仿真程序的灵活性大大提升,拓展了对于AI集群的研究广度。

[0038] 实施例二

[0039] 本申请实施例提了一种用于分布式AI集群的网络仿真装置,采用实施例一种任意一项所述的方法步骤。其包括如下几个模块。

[0040] 任务生成模块101,配置为响应于AI集群仿真平台触发的第一目标事件,所述第一目标事件利用任务生成算法生成分布式AI仿真任务。

[0041] 资源调度模块102,配置为根据生成的分布式AI仿真任务,获取任务资源调度信息,按照预设的任务分配算法,为所述分布式AI仿真任务分配包括GPU、网络链路在内的AI集群调度资源。

[0042] 路由分配模块103,配置为根据生成的分布式AI仿真任务,按照预先配置的任务流生成算法将所述分布式AI仿真任务解析为若干条仿真任务流,并调用所述AI集群调度资源,利用任务流路由算法为每个仿真任务流配置执行仿真所需的路由信息。

[0043] 仿真执行模块104,配置根据路由信息执行完成每条仿真任务流后,获取分布式AI

仿真任务的测试数据。

[0044] 实施例三

[0045] 本申请实施例还提供了一种电子设备,包括:一个或多个处理器;存储器;用于存储一个或多个程序;当一个或多个程序被一个或多个处理器执行时,使得一个或多个处理器实现如实施例一或实施例二中任一的方法。

[0046] 本实施例还提供了一种计算机可读介质,其上存储有计算机程序,计算机程序被处理器执行时实现如实施例一或实施例二中任一的方法。

[0047] 从而使得处理器在执行方法过程中,执行实施例一中的方法时,执行如下步骤:

[0048] 步骤S101,响应于AI集群仿真平台触发的第一目标事件,所述第一目标事件利用任务生成算法生成分布式AI仿真任务。

[0049] 步骤S102,根据生成的分布式AI仿真任务,按照预设的任务分配算法,为所述分布式AI仿真任务分配包括GPU、网络链路在内的AI集群调度资源。

[0050] 步骤S103,根据生成的分布式AI仿真任务,按照预先配置的任务流生成算法将所述分布式AI仿真任务解析为若干条仿真任务流,并调用所述AI集群调度资源,利用任务流路由算法为每个仿真任务流配置执行仿真所需的路由信息。

[0051] 步骤S104,根据路由信息执行完成每条仿真任务流后,获取分布式AI仿真任务的测试数据。

[0052] 本领域内的技术人员应明白,本发明的实施例可提供为方法、系统、或计算机程序产品。因此,本发明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,本发明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。

[0053] 本发明是参照根据本发明实施例的方法、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。

[0054] 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。

[0055] 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。

[0056] 尽管已描述了本发明的优选实施例,但本领域内的技术人员一旦得知了基本创造性概念,则可对这些实施例作出另外的变更和修改。所以,所附权利要求意欲解释为包括优选实施例以及落入本发明范围的所有变更和修改。

[0057] 显然,本领域的技术人员可以对本发明进行各种改动和变型而不脱离本发明的精神和范围。这样,倘若本发明的这些修改和变型属于本发明权利要求及其等同技术的范围之内,则本发明也意图包含这些改动和变型在内。

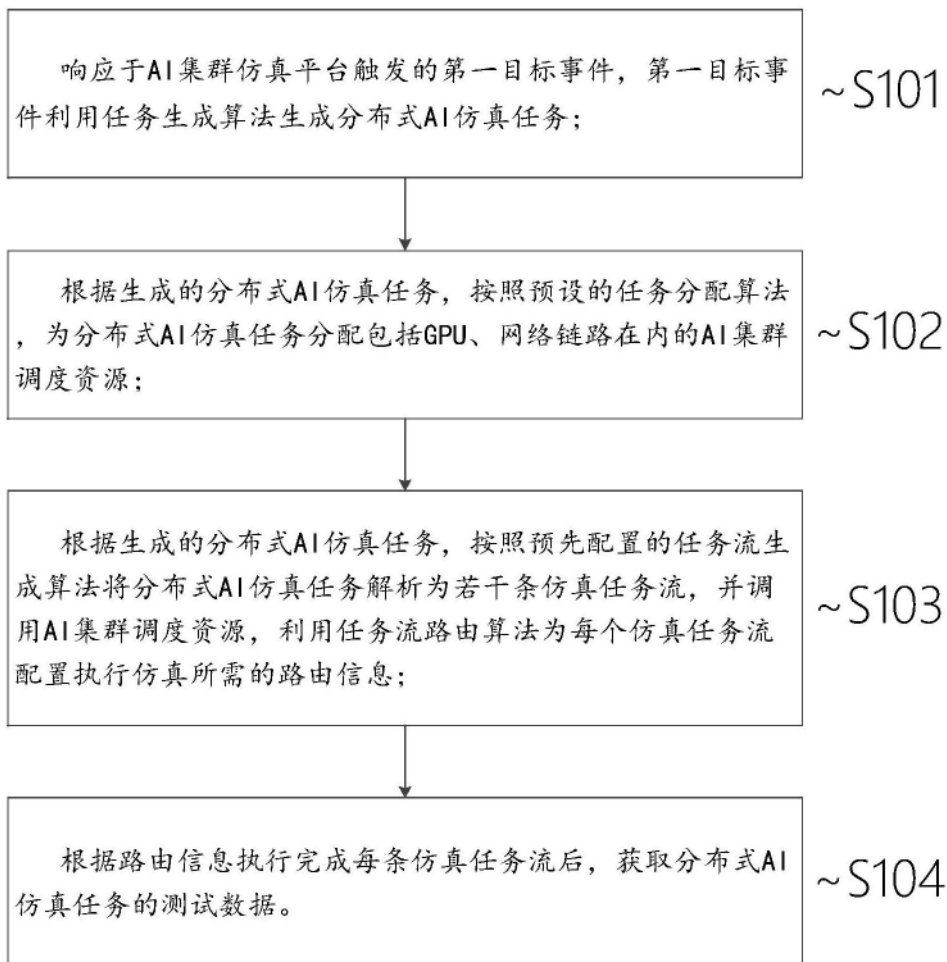


图1

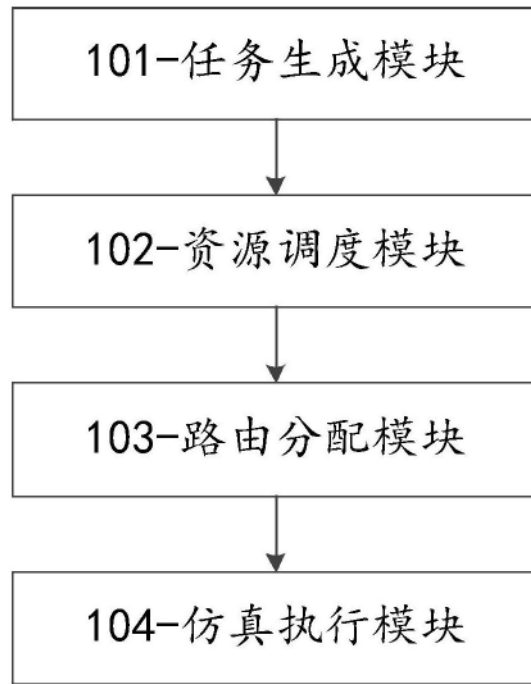


图2