

权利要求书

1、一种存储设备，其特征在于，包括处理器、硬盘和人工智能 AI 装置，所述 AI 装置和所述处理器之间通过高速互联网络通信；

所述处理器，用于接收业务数据，并将所述业务数据存储至所述硬盘中；

所述 AI 装置，用于向所述处理器发送数据获取请求以获取所述业务数据，并对所述业务数据进行 AI 计算。

2、根据权利要求 1 所述的存储设备，其特征在于，所述数据获取请求包括第一数据获取请求；

所述处理器，用于响应于所述第一数据获取请求，从所述硬盘中获取所述业务数据，并将所述业务数据发送给所述 AI 装置。

3、根据权利要求 1 所述的存储设备，其特征在于，所述数据获取请求包括第二数据获取请求；

所述处理器，用于响应于所述第二数据获取请求，将所述业务数据的元数据发送给所述 AI 装置，所述元数据用于指示所述业务数据的地址；

所述 AI 装置，用于当所述元数据指示所述业务数据位于所述存储设备本地时，向所述硬盘发送第一数据访问请求，所述第一数据访问请求包括所述元数据；

所述硬盘，用于根据所述元数据获取所述业务数据，并通过直接内存存取 DMA 方式将所述业务数据写入所述 AI 装置。

4、根据权利要求 1 所述的存储设备，其特征在于，所述数据获取请求包括第三数据获取请求；

所述处理器，用于响应于所述第三数据获取请求，将所述业务数据的元数据发送给所述 AI 装置，所述元数据用于指示所述业务数据的地址；

所述 AI 装置，用于当所述元数据指示所述业务数据位于其他存储设备时，向所述其他存储设备发送第二数据访问请求，所述第二数据访问请求包括所述元数据。

5、根据权利要求 1 所述的存储设备，其特征在于，所述存储设备还包括内存，所述处理器，还用于从所述内存中划分一段内存空间预留给所述 AI 装置使用。

6、根据权利要求 5 所述的存储设备，其特征在于，所述 AI 装置包括 AI 处理器和 AI 内存；

所述 AI 处理器用于当所述 AI 内存的可用容量达到预设阈值时向所述处理器发送内存申请请求，所述 AI 内存的可用容量通过设定的批尺寸确定，所述内存申请请求用于请求所述处理器从所述内存中划分一段内存空间预留给所述 AI 装置使用。

7、根据权利要求 1 所述的存储设备，其特征在于，所述 AI 装置包括 AI 处理器；

所述 AI 处理器，用于将计算任务划分为至少两个子任务，将所述至少两个子任务中的第

一子任务发送给所述处理器，所述计算任务是指所述对所述业务数据进行 AI 计算；

所述处理器，还用于执行所述第一子任务，并将计算结果发送给所述 AI 处理器。

8、一种分布式存储系统，其特征在于，包括多个存储设备，所述多个存储设备中的第一存储设备包括第一处理器、第一硬盘和第一人工智能 AI 装置，所述第一 AI 装置和所述第一处理器之间通过高速互连网络通信；

所述第一处理器，用于接收业务数据，并将所述业务数据存储至所述第一硬盘中；

所述第一 AI 装置，用于向所述第一处理器发送数据获取请求以获取所述业务数据，并对所述业务数据进行 AI 计算。

9、根据权利要求 8 所述的系统，其特征在于，所述数据获取请求包括第一数据获取请求；

所述第一处理器，用于响应于所述第一数据获取请求，从所述第一硬盘中获取所述业务数据，并将所述业务数据发送给所述第一 AI 装置。

10、根据权利要求 8 所述的系统，其特征在于，所述数据获取请求包括第二数据获取请求；

所述第一处理器，用于响应于所述第二数据获取请求，将所述业务数据的元数据发送给所述第一 AI 装置，所述元数据用于指示所述业务数据的地址；

所述第一 AI 装置，用于当所述元数据指示所述业务数据位于所述第一存储设备本地时，向所述第一硬盘发送第一数据访问请求，所述第一数据访问请求包括所述元数据；

所述第一硬盘，用于根据所述元数据获取所述业务数据，并通过直接内存存取 DMA 方式将所述业务数据写入所述第一 AI 装置。

11、根据权利要求 8 所述的系统，其特征在于，所述数据获取请求包括第三数据获取请求；

所述第一处理器，用于响应于所述第三数据获取请求，将所述业务数据的元数据发送给所述第一 AI 装置，所述元数据用于指示所述业务数据的地址；

所述第一 AI 装置，用于当所述元数据指示所述业务数据位于所述多个存储设备中的第二存储设备时，向所述第二存储设备发送第二数据访问请求，所述第二数据访问请求包括所述元数据；接收所述第二存储设备响应于所述第二数据访问请求所发送的所述业务数据。

12、根据权利要求 8 所述的系统，其特征在于，

所述多个存储设备中的第二存储设备，用于通过第二网络将所述业务数据传输给所述第一存储设备；

所述第二存储设备，还用于通过第一网络将 AI 参数传输给所述第一存储设备，所述 AI 参数用于对所述业务数据进行 AI 计算。

13、根据权利要求 12 所述的系统，其特征在于，

所述第二存储设备，还用于通过所述第一网络将其他业务数据传输给所述第一存储设备；

所述第二存储设备，还用于通过所述第二网络传输其他 AI 参数，所述其他 AI 参数用于对所述其他业务数据进行 AI 计算。

14、根据权利要求 8 所述的系统，其特征在于，还包括管理装置；

所述管理装置，用于接收第一作业请求，根据所述第一作业请求确定待训练的数据集的分布，所述数据集包括所述业务数据；当确定所述业务数据分布在所述第一存储设备时，向所述第一存储设备发送第一计算请求，所述第一计算请求用于请求所述第一 AI 装置对所述业务数据进行 AI 计算。

15、根据权利要求 14 所述的系统，其特征在于，所述管理装置向所述第一存储设备的第一 AI 装置发送第一计算请求之前，还用于确定所述第一存储设备的运行状态满足设定条件。

16、根据权利要求 14 所述的系统，其特征在于，所述数据集还包括其他业务数据；

所述管理装置，还用于当确定所述其他业务数据分布在所述多个存储设备中的第二存储设备中时，进一步确定所述第二存储设备的运行状态，当所述第二存储设备的运行状态不满足所述设定条件时，向所述第一存储设备发送第二计算请求，所述第一存储设备与所述第二存储设备之间的距离低于设定的距离阈值；

所述第一 AI 装置，还用于基于所述第二计算请求，从所述第二存储设备中获取所述其他业务数据，对所述其他业务数据进行 AI 计算，获得第二计算结果。

17、一种数据处理方法，其特征在于，所述方法应用于分布式存储系统中，所述分布式存储系统包括多个存储设备；所述多个存储设备中的第一存储设备包括第一处理器、第一硬盘和第一 AI 装置，所述第一 AI 装置和所述第一处理器之间通过高速互联网络通信；

所述第一处理器，接收业务数据，并将所述业务数据存储至所述第一硬盘中；

所述第一 AI 装置，向所述第一处理器发送数据获取请求以获取所述业务数据，并对所述业务数据进行 AI 计算。

18、根据权利要求 17 所述的方法，其特征在于，所述数据获取请求包括第一数据获取请求；

所述第一 AI 装置获取所述业务数据包括：所述第一处理器，响应于所述第一数据获取请求，从所述第一硬盘中获取所述业务数据，并将所述业务数据发送给所述第一 AI 装置。

19、根据权利要求 17 所述的方法，其特征在于，所述数据获取请求包括第二数据获取请求；所述第一 AI 装置获取所述业务数据包括：

所述第一处理器，响应于所述第二数据获取请求，将所述业务数据的元数据发送给所述第一 AI 装置，所述元数据用于指示所述业务数据的地址；

所述第一 AI 装置，当所述元数据指示所述业务数据位于所述第一存储设备本地时，向所述第一硬盘发送第一数据访问请求，所述第一数据访问请求包括所述元数据；所述第一硬盘，根据所述元数据获取所述业务数据，并通过 DMA 方式将所述业务数据写入所述第一 AI 装置。

20、根据权利要求 17 所述的方法，其特征在于，所述数据获取请求包括第三数据获取请求；所述第一处理器获取所述业务数据包括：

所述第一处理器，响应于所述第三数据获取请求，将所述业务数据的元数据发送给所述第一 AI 装置，所述元数据用于指示所述业务数据的地址；

所述第一 AI 装置，当所述元数据指示所述业务数据位于所述多个存储设备中的第二存储设备时，向所述第二存储设备发送第二数据访问请求，所述第二数据访问请求包括所述元数据；接收所述第二存储设备响应于所述第二数据访问请求所发送的所述业务数据。

21、根据权利要求 20 所述的方法，其特征在于，所述方法还包括：

所述第二存储设备通过第二网络将所述业务数据传输给所述第一存储设备；

以及通过第一网络将 AI 参数传输给所述第一存储设备，所述 AI 参数用于对所述业务数据进行 AI 计算。

22、根据权利要求 17 所述的方法，其特征在于，所述方法还包括：

所述多个存储设备中的第二存储设备，通过所述第一网络将其他业务数据传输给所述第一存储设备；通过所述第二网络传输其他 AI 参数，所述其他 AI 参数用于对所述其他业务数据进行 AI 计算。

23、根据权利要求 17 所述的方法，其特征在于，所述分布式存储系统还包括管理装置；所述方法还包括：

所述管理装置，接收第一作业请求，根据所述第一作业请求确定待训练的数据集的分布，所述数据集包括所述业务数据；当确定所述业务数据分布在所述第一存储设备时，向所述第一存储设备发送第一计算请求，所述第一计算请求用于请求所述第一 AI 装置对所述业务数据进行 AI 计算。

24、根据权利要求 23 所述的方法，其特征在于，在所述向所述第一存储设备的第一 AI 装置发送第一计算请求之前，所述方法还包括：

所述管理装置确定所述第一存储设备的运行状态满足设定条件。

25、根据权利要求 24 所述的方法，其特征在于，所述数据集还包括其他业务数据；所述方法还包括：

所述管理装置，当确定所述其他业务数据分布在所述多个存储设备中的第二存储设备中时，进一步确定所述第二存储设备的运行状态，当所述第二存储设备的运行状态不满足所述设定条件时，向所述第一存储设备发送第二计算请求，所述第一存储设备与所述第二存储设备之间的距离低于设定的距离阈值；

所述第一 AI 装置，基于所述第二计算请求，从所述第二存储设备中获取所述其他业务数据，对所述其他业务数据进行 AI 计算，获得第二计算结果。