

镜像加速系统、方法及装置

本申请要求于 2022 年 07 月 28 日提交中国专利局、申请号为
5 202210898260.X、申请名称为“镜像加速系统、方法及装置”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

本申请涉及云计算技术领域，具体涉及镜像加速系统、方法和装置，以及
10 电子设备。

背景技术

分布式无服务器（Serverless）计算平台是一种公有云 FaaS
（Function-as-a-Service，函数即服务或者功能即服务）产品，能使云应用
开发者可以更轻松地运行和管理微服务应用程序。

15 分布式无服务器计算平台支持容器，提供容器镜像服务，允许用户部署
云原生应用的镜像，并支持镜像启动加速，成为云原生应用安全托管及高效
分发平台。一种典型的容器镜像冷启动加速方式是通过镜像按需读取和 P2P
分发，来减轻应用实例批量扩容时对单点镜像仓库带来的总体负载。容器镜
像加速系统通过专用的回源节点（SuperNode）从镜像仓库获取数据，利用树
20 形 P2P 网络对进行镜像数据分发，树形 P2P 网络内的计算节点（容器宿主机）
之间可相互传输数据，应用实例批量扩容时可快速分发数据到所有计算节点，
实现了在秒级延迟启动容器。

然而，在实现本发明过程中，发明人发现上述方案至少存在如下问题：1）
在专用的回源节点因故障或者维护需要而无法使用时，会影响 P2P 网络传输
25 效率，镜像仓库大规模并行分发场景下会产生网络拥堵，导致容器冷启动时
间较长，不支持大规模和大镜像分发加速。具体而言，回源节点由于稳定性
等原因，会出现机器被回收的时间，例如在回源节点进行升级时，会主动进
行机器的回收，因此回源节点的生命周期是不可信的。目前的 P2P 组网结构
无法自适应回源节点的变化，导致回源节点出现问题时没有可替代的回源节
30 点，最终回源节点数量会变少。计算节点要获取镜像数据需经过回源节点，
当回源节点变少时，就会影响数据传输效率。2）专用的回源节点独立于 P2P
网络内的计算节点而额外存在，因此系统开销较大。

发明内容

本申请提供镜像加速系统，以解决现有技术存在的回源节点无法使用情
况下的镜像冷启动高延迟、系统开销较大的问题。本申请另外提供镜像加速
35

方法、装置，电子设备。

本申请提供一种镜像加速系统，包括：

5 P2P 管理装置，部署在管理节点侧，用于获取基于 P2P 网络结构的镜像服务集群内的多个代理节点的系统信息、和出现异常的根节点信息，所述根节点和所述代理节点用于根据镜像数据在虚拟运行环境中启动应用实例，所述根节点还用于从镜像数据源获取镜像数据；根据所述系统信息，从所述多个代理节点中选取节点作为异常根节点的替代节点；从集群组网信息中清除所述出现异常的根节点信息，将集群组网信息中所述替代节点设置为根节点；向所述镜像服务集群内的计算节点提供更新后的集群组网信息；

10 P2P 服务装置，部署在计算节点侧，用于获取计算节点的所述系统信息，向 P2P 管理装置提供所述系统信息；获取 P2P 管理装置提供的所述更新后的集群组网信息；根据所述更新后的集群组网信息，处理镜像数据。

本申请还提供一种镜像加速方法，应用于管理节点，包括：

15 获取基于 P2P 网络结构的镜像服务集群内的多个代理节点的系统信息、和出现异常的根节点信息，所述根节点和所述代理节点用于根据镜像数据在虚拟运行环境中启动应用实例，所述根节点还用于从镜像数据源获取镜像数据；

根据所述系统信息，从所述多个代理节点中选取节点作为异常根节点的替代节点；

20 从集群组网信息中清除所述出现异常的根节点信息，将集群组网信息中所述替代节点设置为根节点；

向所述镜像服务集群内的计算节点提供更新后的集群组网信息。

可选的，所述根据所述系统信息，从所述多个代理节点中选取节点作为异常根节点的替代节点，包括：

25 根据所述代理节点的内存空闲比例、CPU 空闲比例、磁盘使用量和机器启动时间的至少一项数据，确定所述代理节点的得分；

根据所述多个代理节点的得分顺序，从多个代理节点中确定所述替代节点。

可选的，还包括：

30 获取重新加入到所述镜像服务集群的计算节点信息；

在集群组网信息中将重新加入的计算节点设置为代理节点。

可选的，还包括：

获取出现异常的代理节点信息；

从集群组网信息中清除所述出现异常的代理节点信息。

35 可选的，所述向所述镜像服务集群内的计算节点提供更新后的集群组网信息，包括：

通过广播方式，向所述镜像服务集群内的计算节点发送更新后的集群组网信息。

可选的，获取所述出现异常的根节点信息，包括：

若根节点异常检测条件成立，则获取所述出现异常的根节点信息，所述根节点异常检测条件包括但不限于：达到根节点异常检测周期。

可选的，获取所述出现异常的根节点信息，采用如下方式的至少一个：

若根节点的系统信息满足节点异常条件，则判定根节点出现异常；

5 若与根节点上的 P2P 服务装置进行心跳保活失败，则判定根节点出现异常；

若接收到根节点发送的异常信息，则判定根节点出现异常。

可选的，所述方法的执行主体为管理节点，所述管理节点包括：P2P 管理装置，计算节点管理装置，管理服务装置；

10 所述计算节点包括：P2P 服务装置，计算节点代理装置，容器加速装置；所述系统信息采用如下方式获取：

所述 P2P 管理装置通过与所述 P2P 服务装置进行心跳保活，获取所述系统信息。

可选的，还包括：

15 所述 P2P 管理装置若判定计算节点信息获取条件成立，则通过所述计算节点管理装置，获取所述镜像服务集群的计算节点信息，所述计算节点信息获取条件包括但不限于：达到节点信息获取周期；

通过与所述 P2P 服务装置进行心跳保活，获取可用计算节点信息；

20 所述 P2P 管理装置通过与所述 P2P 服务装置进行心跳保活，获取所述系统信息，包括：

所述 P2P 管理装置通过与所述可用计算节点的 P2P 服务装置进行心跳保活，获取所述系统信息。

本申请还提供一种镜像加速方法，应用于计算节点，包括：

25 获取计算节点的系统信息，所述计算节点属于基于 P2P 网络结构的镜像服务集群；

向 P2P 管理装置提供所述系统信息，以使得 P2P 管理装置若确定所述镜像服务集群内存在出现异常的根节点，则根据所述镜像服务集群内的多个代理节点的系统信息，从所述多个代理节点中选取节点作为异常根节点的替代节点；从集群组网信息中清除所述出现异常的根节点信息，将集群组网信息中所述替代节点设置为根节点；

获取 P2P 管理装置提供的更新后的集群组网信息；

根据所述更新后的集群组网信息，处理镜像数据，所述根节点用于从镜像数据源获取镜像数据及根据镜像数据在虚拟运行环境中启动应用实例，所述代理节点用于根据镜像数据在虚拟运行环境中启动应用实例。

35 本申请还提供一种镜像加速方法，包括：

获取基于 P2P 网络结构的镜像服务集群内的多个代理节点的系统信息、和出现异常的根节点信息，所述根节点用于从镜像数据源获取镜像数据及根据镜像数据在虚拟运行环境中启动应用实例，所述代理节点用于根据镜像数

据在虚拟运行环境中启动应用实例；

根据所述系统信息，从所述多个代理节点中选取节点作为异常根节点的替代节点；

从集群组网信息中清除所述出现异常的根节点信息，将集群组网信息中
5 所述替代节点设置为根节点；

根据更新后的集群组网信息，处理镜像数据。

本申请还提供一种容器镜像加速装置，包括：

信息获取单元，用于获取基于 P2P 网络结构的镜像服务集群内的多个代理节点的系统信息、和出现异常的根节点信息，所述根节点用于从镜像数据源获取镜像数据及根据镜像数据在虚拟运行环境中启动应用实例，所述代理节点用于根据镜像数据在虚拟运行环境中启动应用实例；
10

根节点重新指派单元，用于根据所述系统信息，从所述多个代理节点中选取节点作为异常根节点的替代节点；

组网信息更新单元，用于从集群组网信息中清除所述出现异常的根节点信息，将集群组网信息中所述替代节点设置为根节点；
15

镜像数据处理单元，用于根据更新后的集群组网信息，处理镜像数据。

本申请还提供一种电子设备，包括：

处理器和存储器；存储器，用于存储实现上述方法的程序，该设备通电并通过所述处理器运行该方法的程序。

本申请还提供一种计算机可读存储介质，所述计算机可读存储介质中存储有指令，当其在计算机上运行时，使得计算机执行上述各种方法。
20

本申请还提供一种包括指令的计算机程序产品，当其在计算机上运行时，使得计算机执行上述各种方法。

与现有技术相比，本申请具有以下优点：

本申请实施例提供的镜像加速系统，包括可部署在基于 P2P 网络结构的镜像服务集群的管理节点侧的 P2P 管理装置，还包括可部署在 P2P 网络内的计算节点侧的 P2P 服务装置。P2P 管理装置，用于获取镜像服务集群内的多个代理节点的系统信息、和出现异常的根节点信息，所述根节点和所述代理节点用于根据镜像数据在虚拟运行环境中启动应用实例，所述根节点还用于从镜像数据源获取镜像数据；根据所述系统信息，从所述多个代理节点中选取节点作为异常根节点的替代节点；从集群组网信息中清除所述出现异常的根节点信息，将集群组网信息中所述替代节点设置为根节点；向所述镜像服务集群内的计算节点提供更新后的集群组网信息；P2P 服务装置，用于获取计算节点的所述系统信息，向 P2P 管理装置提供所述系统信息；获取 P2P 管理装置提供的所述更新后的集群组网信息；根据所述更新后的集群组网信息，处理镜像数据。采用这种处理方式，使得支持动态的 P2P 组网结构，即使其中某个回源节点（根节点）由于某种原因下线了，P2P 管理装置（P2Pmanager）在感知到这一变化后，就会从 P2P 网络中指派新的回源节点，新的回源节点上会拥有绝大部分所需要的镜像数据，可以近乎做到无缝的回源节点替代，
25
30
35

这样在一段时间内就拥有固定数量的回源节点；因此，可以有效确保某些回源节点不可用情况下的数据传输效率，从而支持大规模和大镜像分发加速。此外，这种处理方式实现了去中心化的回源节点，所有的回源节点选择均为 P2P 网络中的计算节点，根节点和代理节点均为传统意义中 P2P 的对等节点，

5 无需再额外部署专用的回源节点；因此，可以有效降低系统开销。再者，由于去中心化的回源节点利用了 P2P 网络中的计算节点（容器宿主机）的闲置网络和硬盘资源来回源拉取镜像，这类回源节点还可同时根据镜像启动应用实例，因此也不会影响线上应用。

附图说明

- 10 图 1 本申请提供的镜像加速系统的实施例的结构示意图；
图 2 本申请提供的镜像加速系统的实施例的应用场景示意图；
图 3 本申请提供的镜像加速系统的实施例的交互示意图；
图 4a 本申请提供的镜像加速系统的实施例的更新前的 P2P 组网结构示意图；
15 图 4b 本申请提供的镜像加速系统的实施例的更新后的 P2P 组网结构示意图；
图 5 本申请提供的镜像加速系统的实施例的 P2P 网络服务装置的示意图；
图 6 本申请提供的镜像加速方法的管理节点实施例的流程示意图；
图 7 本申请提供的镜像加速方法的计算节点实施例的流程示意图；
20 图 8 本申请提供的镜像加速方法的实施例的流程示意图。

具体实施方式

在下面的描述中阐述了很多具体细节以便于充分理解本申请。但是本申请能够以很多不同于在此描述的其它方式来实施，本领域技术人员可以在不违背本申请内涵的情况下做类似推广，因此本申请不受下面公开的具体实施的

25 限制。

在本申请中，提供了镜像加速系统、方法和装置，以及电子设备。为了能够清楚地说明本申请实施例提供的方案，下面先对容器镜像冷启动加速技术进行说明。然后，在下面的实施例中逐一对本申请实施例提供的各种方案进行详细说明。

30 无服务器计算平台是支撑用户应用系统（如电商平台、即时通信软件等）的基础设施，平台会为用户准备好计算资源，以弹性、可靠的方式运行代码，并可提供日志查询、性能监测、报警等功能，用户无需再管理服务器等基础设施，只需编写代码并上传，可实现应用系统快速上云及 Serverless 架构改造。

35 对于基于无服务器计算平台运行的应用程序，容器镜像（Image）包括打包的应用程序及其依赖关系，以及有关启动时所运行的进程的信息。容器镜像相比于代码包有更好的可移植性和更丰富的工具链生态，但其自带的与应用无关的数据（如容器镜像自带的操作系统、命令工具以及其他与应用本身

不需要的文件) 极易使镜像臃肿。在应用实例启动过程中, 需要从镜像仓库拉取镜像, 大镜像不可避免地加大了需要下载和解压的数据量, 增加了应用实例冷启动时间。对于较大的镜像, 如解压前 1 GB 以上级别的容器镜像会造成分钟级别的镜像拉取延迟, 明显拖慢了应用实例的冷启动性能。如果是大规模启动, 这个问题则会变得更糟糕。镜像仓库会因为镜像并发拉取导致带宽打满或者服务端压力过大而直接崩溃。

为了提升容器镜像冷启动速度, 对于运行环境为自定义容器 (Custom Container) 且使用容器镜像服务的应用程序, 无服务器计算平台为其开启镜像启动加速功能。实现镜像启动加速的使用原理是, 每次创建或更新该类应用程序时, 无服务器计算平台会拉取镜像, 并且转存至计算平台内部的加速镜像缓存服务中。转存结束后, 将大部分的容器镜像读请求拦截在集群内, 提高缓存命中率, 不需要请求进行回源拉取, 节省容器镜像服务节点与网络镜像存储节点的带宽消耗, 消除通过网络拉取镜像数据的性能瓶颈, 应用程序中的自定义容器镜像启动速度将会得到显著提升, 为计算平台带来更快的容器镜像启动速度, 进而提高用户的端到端体验。开启镜像加速可将分钟级的镜像拉取缩短至秒级。

容器镜像加速系统位于镜像存储和应用之间的加速层, 支持镜像按需读取和 P2P 镜像分发, 可减轻计算节点批量扩容时对单点镜像仓库带来的总体负载。P2P 镜像分发网络包括多个计算节点, 这类节点可以是容器宿主机或者虚拟机宿主机。P2P 镜像分发的原理是将镜像分成很多的“块 (block)”, 如果某个计算节点拉取了这个块, 那么其他的节点就可以从这个节点拉数据, 从而避免所有的请求都打到镜像仓库。例如, 在节点上部署了应用容器引擎 Docker 客户端, 如果某个 Docker 客户端拉取了这个块, 那么其他的 Docker 客户端就可以从这个客户端拉数据。

在 P2P 镜像分发的现有技术 (也称为原生 P2P 方案) 中, 由固定不变的回源节点负责存储“块”和客户端 (计算节点) 的关系, 客户端 A 第一次请求会被打到回源节点, 然后回源节点去镜像仓库拉取数据转发给客户端 A, 并且会记录这些块和客户端 A 的对应关系。后续其他客户端 (B, C, ...) 请求这些块的时候, 回源节点会告诉客户端 (B, C, ...) 应该去刚才成功拉取的节点 (客户端 A) 上获取数据, 从而降低镜像仓库的负载。正如申请人在上述背景技术中所阐述的, 在这种原生 P2P 方案的回源节点不可用时, 会影响 P2P 网络传输效率, 因此不支持大规模和大镜像分发加速。并且, 由于额外部署的专用回源节点独立于 P2P 网络而存在, 因此系统开销也比较大。

需要说明的是, 无服务器计算平台向用户提供的镜像服务, 不仅可以是在容器内根据镜像数据启动应用实例, 还可以是在虚拟机内根据镜像数据启动应用实例。容器相比虚拟机的特点是轻量化和快速启动。相比虚拟机动辄十几个 G 的镜像, 容器镜像只包含应用以及应用所需的依赖库, 所以可以做到几百 M 甚至更少。下面的实施例主要以容器镜像启动场景为例进行说明, 但本申请实施例提供的各种方案也可用于虚拟机镜像启动场景。

第一实施例

请参看图 1，其为本申请的镜像加速系统的实施例的结构示意图。在本实施例中，所述系统包括：P2P 管理装置 1 和 P2P 服务装置 2。

本申请实施例提供的系统，采用去中心化的集群组网信息 (NodeList) 的组网逻辑，通过 P2P 管理装置 1 和 P2P 服务装置 2 的配合来实现 P2P 拓扑逻辑的管理，实现了计算节点可复用的容器镜像加速方式。具体而言，所述系统去除了原生 P2P 方案中固定不变的专用回源节点，将从 P2P 网络中的对等节点分为根节点 (Root 节点) 和代理节点 (Agent 节点) 两类，代理节点负责采用容器镜像启动应用实例，根节点不仅负责采用容器镜像启动应用实例，还利用计算节点的闲置网络和硬盘资源负责数据回源的任务。所述系统通过 P2P 管理装置 1，可在发现某些根节点将要下线时，根据 P2P 网络中其它代理节点的设备运行情况，从其它代理节点中选取替代的根节点，由此实现了动态指派回源节点，可保证在一段时间内拥有固定数量的回源节点，因此可支持大规模和大镜像分发加速。

如图 2 所示，在本实施例中，无服务器计算平台包括容器镜像服务器集群，该集群可包括至少一个镜像仓库，镜像仓库可部署在存储节点侧，一个管理节点 (左侧部分)，由多个计算节点 (EE，右侧集群范围部分) 构成的树形 P2P 镜像分发网络。所述系统利用树形 P2P 网络对进行镜像数据分发。P2P 网络中的一部分节点为根节点，根节点既能作为容器宿主机负责采用容器镜像启动应用实例，还可作为回源节点利用节点的闲置网络和硬盘资源回源拉取镜像。根节点从镜像仓库获取镜像数据后，将镜像数据存储至 P2P 网络中，相当于提供了加速镜像缓存服务。P2P 网络中的另一部分节点为代理节点，代理节点为容器的宿主机，代理节点之间可相互传输镜像数据，应用实例批量扩容时可快速分发镜像数据到所有计算节点 (包括根节点和代理节点)，由此可实现在秒级延迟启动容器。

如图 3 所示，P2P 管理装置 1 用于获取基于 P2P 网络结构的镜像服务集群内的多个代理节点的系统信息、和出现异常的根节点信息；根据所述系统信息，从所述多个代理节点中选取节点作为异常根节点的替代节点；从集群组网信息中清除所述出现异常的根节点信息，将集群组网信息中所述替代节点设置为根节点；向所述镜像服务集群内的计算节点提供更新后的集群组网信息；P2P 服务装置 2，用于获取计算节点的所述系统信息，向 P2P 管理装置 1 提供所述系统信息；获取 P2P 管理装置 1 提供的所述更新后的集群组网信息；根据所述更新后的集群组网信息，处理镜像数据。其中，所述根节点和所述代理节点用于根据镜像数据在虚拟运行环境中启动应用实例，所述根节点还用于从镜像数据源获取镜像数据。

所述出现异常的根节点，可以是出现故障的根节点，也可以是要进行系统升级等维护操作的根节点。

所述集群组网信息包括各计算节点的角色信息，角色可以是根节点或者代理节点。如图 4a 所示，集群组网信息 NodeList 包括了各节点的角色信息，

其中 Node-1、6、2 为根节点，Node-3、4、5、8、7 为根节点。在本实施例中，集群组网信息是由管理节点主动维护的集群维度根节点列表。

本申请提出了集群组网信息 (nodelist) 的概念，并提出了通过动态更新集群组网信息，使得计算节点根据集群组网信息寻址到可用的根节点，并通过访问该根节点下载镜像数据的技术构思。在本实施例中，通过 P2P 管理装置和 P2P 服务装置之间的配合，动态更新集群组网信息，可适应更加动态的机器生命周期，且可以自动摘除不可用节点。因此，可以有效提升容错概率，自动运维 P2P 拓扑逻辑状态，从而提升容器镜像冷启动速度。

在本实施例中，所述系统的回源节点并非固定不变，而是可根据 P2P 网络内的计算节点的设备运行情况动态变化。当一台根节点服务器由于故障或者维护需要而无法使用的时候，将由 P2P 管理装置 1 来决策用哪一个代理节点服务器替代该根节点，自动接管该根节点的工作。如图 4b 所示，图 4a 中的根节点 Node-6 因升级而被回收下线时，可指派计算节点 Node-5 作为根节点，替代下线的根节点 Node-6，此时集群组网信息 NodeList 中清除了 Node-6 的角色信息，将 Node-5 标记为根节点。采用这种处理方式，使得在高并发量情况下始终有足够的根节点从镜像仓库获取数据，以确保数据传输效率，从而确保容器镜像冷启动低延迟，实现在秒级延迟启动容器。

具体实施时，P2P 网络中的计算节点可周期性地更新集群组网信息 (NodeList 配置文件) 来获取最新的集群内的 Root 和 Agent 列表。当镜像启动请求达到某个计算节点上，该节点首先会从本机的 NodeList 读取集群组网信息，从而实现镜像启动时可以从近存储拉取镜像数据。

如图 2 所示，具体实施时，P2P 管理装置 1 可通过广播方式，向所述镜像服务集群内的计算节点发送更新后的集群组网信息。采用这种处理方式，可以有效提升各计算节点更新集群组网信息的效率。

所述系统信息，包括但不限于：计算节点的开机时间，存储空间使用信息，内存空间使用信息，CPU 使用信息，网络速度。

在一个示例中，P2P 管理装置 2 具体用于根据所述代理节点的内存空闲比例、CPU 空闲比例、磁盘使用量和机器启动时间的至少一项数据，确定所述代理节点的得分；根据所述多个代理节点的得分顺序，从多个代理节点中确定所述替代节点。

具体实施时，节点得分与节点上的内存空闲比例、CPU 空闲比例、磁盘使用量成正比，与机器启动时间成反比。例如，可采用如下公式计算得分：得分 = 内存空闲比例 * CPU 空闲比例 * 磁盘使用量 / 机器启动时间。采用这种处理方式，使得根节点具有更多闲置网络和硬盘资源负责数据回源的任务；因此，可以有效提升回源拉取镜像的效率。

在一个示例中，P2P 管理装置 1 要获取所述出现异常的根节点信息，可采用如下方式：若根节点异常检测条件成立，则获取所述出现异常的根节点信息，所述根节点异常检测条件包括但不限于：达到根节点异常检测周期，如每分钟检测一次是否存在出现异常的根节点。采用这种处理方式，可以实现

定期检测异常根节点。

在一个示例中，P2P 管理装置 1 要获取所述出现异常的根节点信息，可采用如下方式的至少一个：若根节点的系统信息满足节点异常条件，则判定根节点出现异常，所述节点异常条件可根据应用需求设置；若与根节点上的 P2P 服务装置进行心跳保活失败，则判定根节点出现异常；若接收到根节点发送的异常信息，则判定根节点出现异常。

具体实施时，P2P 管理装置 1 可通过与所述 P2P 服务装置 2 进行心跳保活的方式，获取所述系统信息。

P2P 服务装置 2 负责管理其所在计算节点的集群组网信息 NodeList。在一个示例中，P2P 服务装置 2 可与 P2P 管理装置 1 进行 rpc 通信、周期性地收集该节点上的系统信息，包括但不限于开机时间、可用数据盘空间、可用内存空间等。P2P 服务装置 2 可根据当前本机上的 NodeList 配置文件去相应的节点上拉取所需要的镜像文件。

由图 2 可见，在一个示例中，管理节点包括 P2P 管理装置 1(P2P manager)，管理节点侧还部署了计算节点管理装置 (EEManager) 和管理服务装置 (Admin_service)。EEManager 和 Admin_service 是管理节点的已有装置，其中 EEManager 用于获取集群内的多个计算节点 (EE) 信息，包括计算节点的 IP、状态、类型等信息。

如图 5 所示，P2P 网络中的多个计算节点 (包括根节点和代理节点) 均包括 P2P 服务装置 2 (P2P Daemon)，计算节点侧还部署了计算节点代理装置 (EEAgent) 和容器加速装置 (DADI Agent)。其中，DADI 是指基于块存储的数据按需读取加速格式。采用这种处理方式，使得 P2P 服务装置 2 位于计算节点的旁路控制上，并不对现有的应用调用链路有改变，P2P 管理装置 1 和 P2P 服务装置 2 均以插件形式接入到现有的容器镜像加速系统，实现了对生产环境计算节点和管理节点无侵入的容器镜像加速；因此，可以有效降低容器镜像加速系统的改造成本。

在一个示例中，P2P 管理装置 1 还用于若判定计算节点信息获取条件成立，则通过所述计算节点管理装置，获取所述镜像服务集群的计算节点信息，所述计算节点信息获取条件包括但不限于：达到节点信息获取周期；所述 P2P 管理装置 1 通过与所述 P2P 服务装置 2 进行心跳保活，获取可用计算节点信息；所述 P2P 管理装置 1 通过与所述可用计算节点的 P2P 服务装置 2 进行心跳保活，获取所述系统信息。

P2P 管理装置 1 负责管理 P2P 组网内的计算节点机器，更新每个节点上最新的集群内组网信息、将不健康的机器从组网中踢出等。具体实施时，P2P 管理装置 1 可周期性的与计算节点管理装置通信，获取到当前集群内所有的计算节点机器，从中筛选出正在服务应用的机器。P2P 管理装置 1 还可周期性地通过与计算节点上的 P2P 服务装置 2 进行心跳保活，获取到机器最新的系统信息：包括机器的生命周期、剩余可用的数据盘空间等。在每次收集结束后，P2P 管理装置 1 会根据当前的最新数据重新计算组网结构，进行新的根节

点的指派。

在一个示例中，P2P 管理装置 1 还用于获取重新加入到所述镜像服务集群的计算节点信息；在集群组网信息中将重新加入的计算节点设置为代理节点。P2P 管理装置 1 还可用于获取出现异常的代理节点信息；从集群组网信息中清除所述出现异常的代理节点信息。采用这种处理方式，使得 P2P 组网结构是自愈的，可以自适应计算节点的加入、删除等，在异常节点出现的时候，会被 P2P 管理装置 1 移除组网，可以覆盖镜像启动加速场景下计算节点生命周期不可预期的情况；因此，可以有效确保数据传输效率。

具体实施时，当某一个计算节点出现了异常（如网络不通、内存/CPU 等系统资源吃紧等情况出现），P2P 管理装置 1 在进行决策后会在组网逻辑中主动放弃这个节点。例如，代理节点即将被下线，P2P 管理装置 1 与该节点的 P2P 服务装置 2 心跳保活失败，从组网中踢出此节点并进行广播，其他节点照常进行数据服务。

再例如，根节点即将被下线，P2P 管理装置 1 与该节点的 P2P 服务装置 2 心跳保活失败，会做出 NodeList 的更新广播至所有可用计算节点，并选取新的根节点，通过 P2P 管理装置和 P2P 服务装置 2 之间通信来更新集群组网信息，可保证代理节点可以寻址到新的根节点，不会导致数据传输的停滞。例如，一个镜像启动请求落到某个节点 A 上，该节点 A 会计算出这个请求所对应镜像的每个层信息，并根据层信息找到所在的根节点，然后向根节点进行镜像获取请求，对应的根节点收到镜像获取请求后，会告诉节点 A 其需要的数据源具体是哪个节点；节点 A 则去对应的节点获取数据。根节点在收到的镜像获取请求是冷启动的情况下（镜像数据的第一次请求，缓存未命中）会去数据源读取数据。

从上述实施例可见，本申请实施例提供的镜像加速系统，包括可部署在基于 P2P 网络结构的镜像服务集群的管理节点侧的 P2P 管理装置，还包括可部署在 P2P 网络内的计算节点侧的 P2P 服务装置。P2P 管理装置，用于获取镜像服务集群内的多个代理节点的系统信息、和出现异常的根节点信息，所述根节点和所述代理节点用于根据镜像数据在虚拟运行环境中启动应用实例，所述根节点还用于从镜像数据源获取镜像数据；根据所述系统信息，从所述多个代理节点中选取节点作为异常根节点的替代节点；从集群组网信息中清除所述出现异常的根节点信息，将集群组网信息中所述替代节点设置为根节点；向所述镜像服务集群内的计算节点提供更新后的集群组网信息；P2P 服务装置，用于获取计算节点的所述系统信息，向 P2P 管理装置提供所述系统信息；获取 P2P 管理装置提供的所述更新后的集群组网信息；根据所述更新后的集群组网信息，处理镜像数据。采用这种处理方式，使得支持动态的 P2P 组网结构，即使其中某个回源节点（根节点）由于某种原因下线了，P2P 管理装置（P2Pmanager）在感知到这一变化后，就会从 P2P 网络中指派新的回源节点，新的回源节点上会拥有绝大部分所需要的数据，可以近乎做到无缝的

回源节点替代，这样在一段时间内就拥有固定数量的回源节点；因此，可以有效确保某些回源节点不可用情况下的数据传输效率，从而支持大规模和大镜像分发加速。此外，这种处理方式实现了去中心化的回源节点，所有的回源节点选择均为 P2P 网络中的计算节点，根节点和代理节点均为传统意义中 P2P 的对等节点，无需再额外部署专用的回源节点；因此，可以有效降低系统开销。再者，由于去中心化的回源节点利用了 P2P 网络中的计算节点（容器宿主机）的闲置网络和硬盘资源来回源拉取镜像，这类回源节点还可同时根据镜像启动应用实例，因此也不会影响线上应用。

第二实施例

在上述的实施例中，提供了一种镜像加速系统，与之相对应的，本申请还提供一种镜像加速方法。该方法是与上述系统的实施例相对应。由于方法实施例基本相似于系统实施例，所以描述得比较简单，相关之处参见系统实施例的部分说明即可。下述描述的方法实施例仅仅是示意性的。

请参看图 6，其为本申请的镜像加速方法的实施例的流程示意图。在本实施例中，所述方法包括：

步骤 S601：获取基于 P2P 网络结构的镜像服务集群内的多个代理节点的系统信息、和出现异常的根节点信息。

所述根节点和所述代理节点用于根据镜像数据在虚拟运行环境中启动应用实例，所述根节点还用于从镜像数据源获取镜像数据。所述虚拟运行环境，可以是容器或者虚拟机。

在一个示例中，获取所述出现异常的根节点信息，可采用如下方式实施：若根节点异常检测条件成立，则获取所述出现异常的根节点信息，所述根节点异常检测条件包括但不限于：达到根节点异常检测周期。

在一个示例中，获取所述出现异常的根节点信息，可采用如下方式的至少一个：若根节点的系统信息满足节点异常条件，则判定根节点出现异常；若与根节点上的 P2P 服务装置进行心跳保活失败，则判定根节点出现异常；若接收到根节点发送的异常信息，则判定根节点出现异常。

步骤 S603：根据所述系统信息，从所述多个代理节点中选取节点作为异常根节点的替代节点。

在一个示例中，步骤 S603 可包括如下子步骤：根据所述代理节点的内存空闲比例、CPU 空闲比例、磁盘使用量和机器启动时间的至少一项数据，确定所述代理节点的得分；根据所述多个代理节点的得分顺序，从多个代理节点中确定所述替代节点。

步骤 S605：从集群组网信息中清除所述出现异常的根节点信息，将集群组网信息中所述替代节点设置为根节点。

步骤 S607：向所述镜像服务集群内的计算节点提供更新后的集群组网信息。

在一个示例中，步骤 S607 可采用如下方式实现：通过广播方式，向所述镜像服务集群内的计算节点发送更新后的集群组网信息。

在一个示例中，所述方法还可包括如下步骤：获取重新加入到所述镜像服务集群的计算节点信息；在集群组网信息中将重新加入的计算节点设置为代理节点。

5 在一个示例中，所述方法还可包括如下步骤：获取出现异常的代理节点信息；从集群组网信息中清除所述出现异常的代理节点信息。

10 在一个示例中，所述方法的执行主体为管理节点，所述管理节点包括：P2P 管理装置，计算节点管理装置，管理服务装置；所述计算节点包括：P2P 服务装置，计算节点代理装置，容器加速装置；所述系统信息可采用如下方式获取：所述 P2P 管理装置通过与所述 P2P 服务装置进行心跳保活，获取所述系统信息。

15 在一个示例中，所述方法还可包括如下步骤：所述 P2P 管理装置若判定计算节点信息获取条件成立，则通过所述计算节点管理装置，获取所述镜像服务集群的计算节点信息，所述计算节点信息获取条件包括但不限于：达到节点信息获取周期；通过与所述 P2P 服务装置进行心跳保活，获取可用计算节点信息；所述 P2P 管理装置通过与所述可用计算节点的 P2P 服务装置进行心跳保活，获取所述系统信息。

第三实施例

20 在上述的实施例中，提供了一种镜像加速方法，与之相对应的，本申请还提供一种镜像加速装置。该装置是与上述方法的实施例相对应。由于装置实施例基本相似于方法实施例，所以描述得比较简单，相关之处参见方法实施例的部分说明即可。下述描述的装置实施例仅仅是示意性的。

本申请另外提供一种镜像加速装置，包括：信息获取单元，根节点重新指派单元，组网信息更新单元，更新数据发送单元。

25 信息获取单元，用于获取基于 P2P 网络结构的镜像服务集群内的多个代理节点的系统信息、和出现异常的根节点信息，所述根节点用于从镜像数据源获取镜像数据及根据镜像数据在虚拟运行环境中启动应用实例，所述代理节点用于根据镜像数据在虚拟运行环境中启动应用实例；

30 根节点重新指派单元，用于根据所述系统信息，从所述多个代理节点中选取节点作为异常根节点的替代节点；

组网信息更新单元，用于从集群组网信息中清除所述出现异常的根节点信息，将集群组网信息中所述替代节点设置为根节点；

更新数据发送单元，用于向所述镜像服务集群内的计算节点提供更新后的集群组网信息。

第四实施例

35 在上述的实施例中，提供了一种镜像加速系统，与之相对应的，本申请还提供一种镜像加速方法。该方法是与上述系统的实施例相对应。由于方法实施例基本相似于系统实施例，所以描述得比较简单，相关之处参见系统实施例的部分说明即可。下述描述的方法实施例仅仅是示意性的。

请参看图 7，其为本申请的镜像加速方法的实施例的流程示意图。在本实施例中，所述方法包括：

步骤 S701：获取计算节点的系统信息，所述计算节点属于基于 P2P 网络结构的镜像服务集群。

5 步骤 S703：向 P2P 管理装置提供所述系统信息，以使得 P2P 管理装置若确定所述镜像服务集群内存在出现异常的根节点，则根据所述镜像服务集群内的多个代理节点的系统信息，从所述多个代理节点中选取节点作为异常根节点的替代节点；从集群组网信息中清除所述出现异常的根节点信息，将集群组网信息中所述替代节点设置为根节点。

10 步骤 S704：获取 P2P 管理装置提供的更新后的集群组网信息。

步骤 S705：根据所述更新后的集群组网信息，处理镜像数据。

所述根节点用于从镜像数据源获取镜像数据及根据镜像数据在虚拟运行环境中启动应用实例，所述代理节点用于根据镜像数据在虚拟运行环境中启动应用实例。

15 在一个示例中，所述方法还可包括如下步骤：获取计算节点的异常信息；向 P2P 管理装置提供所述异常信息，以使得 P2P 管理装置在接收到根节点发送的异常信息后，获取多个代理节点的系统信息；P2P 管理装置在接收到代理节点发送的异常信息后，从集群组网信息中清除代理节点信息。

第五实施例

20 在上述的实施例中，提供了一种镜像加速方法，与之相对应的，本申请还提供一种镜像加速装置。该装置是与上述方法的实施例相对应。由于装置实施例基本相似于方法实施例，所以描述得比较简单，相关之处参见方法实施例的部分说明即可。下述描述的装置实施例仅仅是示意性的。

本申请另外提供一种镜像加速装置，包括：系统信息获取单元，系统信息发送单元，更新信息获取单元，镜像数据处理单元。

系统信息获取单元，用于获取计算节点的系统信息，所述计算节点属于基于 P2P 网络结构的镜像服务集群；

30 系统信息发送单元，用于向 P2P 管理装置提供所述系统信息，以使得 P2P 管理装置若确定所述镜像服务集群内存在出现异常的根节点，则根据所述镜像服务集群内的多个代理节点的系统信息，从所述多个代理节点中选取节点作为异常根节点的替代节点；从集群组网信息中清除所述出现异常的根节点信息，将集群组网信息中所述替代节点设置为根节点；

更新信息获取单元，用于获取 P2P 管理装置提供的更新后的集群组网信息；

35 镜像数据处理单元，用于根据所述更新后的集群组网信息，处理镜像数据，所述根节点用于从镜像数据源获取镜像数据及根据镜像数据在虚拟运行环境中启动应用实例，所述代理节点用于根据镜像数据在虚拟运行环境中启动应用实例。

第六实施例

在上述的实施例中，提供了一种镜像加速系统，与之相对应的，本申请还提供一种镜像加速方法。该方法是与上述系统的实施例相对应。由于方法实施例基本相似于系统实施例，所以描述得比较简单，相关之处参见系统实施例的部分说明即可。下述描述的方法实施例仅仅是示意性的。

请参看图 8，其为本申请的镜像加速方法的实施例的流程示意图。在本实施例中，所述方法包括：

步骤 S801：获取基于 P2P 网络结构的镜像服务集群内的多个代理节点的系统信息、和出现异常的根节点信息。

所述根节点用于从镜像数据源获取镜像数据及根据镜像数据在虚拟运行环境中启动应用实例，所述代理节点用于根据镜像数据在虚拟运行环境中启动应用实例。

步骤 S803：根据所述系统信息，从所述多个代理节点中选取节点作为异常根节点的替代节点。

步骤 S805：从集群组网信息中清除所述出现异常的根节点信息，将集群组网信息中所述替代节点设置为根节点。

步骤 S807：根据更新后的集群组网信息，处理镜像数据。

第七实施例

在上述的实施例中，提供了一种镜像加速方法，与之相对应的，本申请还提供一种镜像加速装置。该装置是与上述方法的实施例相对应。由于装置实施例基本相似于方法实施例，所以描述得比较简单，相关之处参见方法实施例的部分说明即可。下述描述的装置实施例仅仅是示意性的。

本申请另外提供一种镜像加速装置，包括：信息获取单元，根节点重新指派单元，组网信息更新单元，镜像数据处理单元。

信息获取单元，用于获取基于 P2P 网络结构的镜像服务集群内的多个代理节点的系统信息、和出现异常的根节点信息，所述根节点用于从镜像数据源获取镜像数据及根据镜像数据在虚拟运行环境中启动应用实例，所述代理节点用于根据镜像数据在虚拟运行环境中启动应用实例；

根节点重新指派单元，用于根据所述系统信息，从所述多个代理节点中选取节点作为异常根节点的替代节点；

组网信息更新单元，用于从集群组网信息中清除所述出现异常的根节点信息，将集群组网信息中所述替代节点设置为根节点；

镜像数据处理单元，用于根据更新后的集群组网信息，处理镜像数据。

第八实施例

在上述的实施例中，提供了一种镜像加速方法，与之相对应的，本申请还提供一种电子设备。该设备是与上述方法的实施例相对应。由于设备实施例基本相似于方法实施例，所以描述得比较简单，相关之处参见方法实施例的部分说明即可。下述描述的设备实施例仅仅是示意性的。

本申请另外提供一种电子设备，包括：处理器和存储器。其中，存储器用于存储实现上述方法实施例提供的镜像加速方法的程序，该终端通电并通过所述处理器运行该方法的程序。

- 5 本申请虽然以较佳实施例公开如上，但其并不是用来限定本申请，任何本领域技术人员在不脱离本申请的精神和范围内，都可以做出可能的变动和修改，因此本申请的保护范围应当以本申请权利要求所界定的范围为准。

 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。

- 10 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。

- 1、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非暂存电脑可读媒体（transitory media），如调制的数据信号和载波。
- 15 2、本领域技术人员应明白，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。
- 20

权 利 要 求 书

- 1、一种镜像加速系统，其特征在于，包括：
P2P 管理装置，部署在管理节点侧，用于获取基于 P2P 网络结构的镜像服务集群内的多个代理节点的系统信息、和出现异常的根节点信息，所述根节点和所述代理节点用于根据镜像数据在虚拟运行环境中启动应用实例，所述根节点还用于从镜像数据源获取镜像数据；根据所述系统信息，从所述多个代理节点中选取节点作为异常根节点的替代节点；从集群组网信息中清除所述出现异常的根节点信息，将集群组网信息中所述替代节点设置为根节点；向所述镜像服务集群内的计算节点提供更新后的集群组网信息；
- 5
10 P2P 服务装置，部署在计算节点侧，用于获取计算节点的所述系统信息，向 P2P 管理装置提供所述系统信息；获取 P2P 管理装置提供的所述更新后的集群组网信息；根据所述更新后的集群组网信息，处理镜像数据。
- 2、一种镜像加速方法，其特征在于，应用于管理节点，包括：
获取基于 P2P 网络结构的镜像服务集群内的多个代理节点的系统信息、
15 和出现异常的根节点信息，所述根节点和所述代理节点用于根据镜像数据在虚拟运行环境中启动应用实例，所述根节点还用于从镜像数据源获取镜像数据；
根据所述系统信息，从所述多个代理节点中选取节点作为异常根节点的替代节点；
20 从集群组网信息中清除所述出现异常的根节点信息，将集群组网信息中所述替代节点设置为根节点；
向所述镜像服务集群内的计算节点提供更新后的集群组网信息。
- 3、根据权利要求 2 的方法，其特征在于，所述根据所述系统信息，从所述多个代理节点中选取节点作为异常根节点的替代节点，包括：
25 根据所述代理节点的内存空闲比例、CPU 空闲比例、磁盘使用量和机器启动时间的至少一项数据，确定所述代理节点的得分；
根据所述多个代理节点的得分顺序，从多个代理节点中确定所述替代节点。
- 4、根据权利要求 2 所述的方法，其特征在于，还包括：
30 获取重新加入到所述镜像服务集群的计算节点信息；
在集群组网信息中将重新加入的计算节点设置为代理节点。
- 5、根据权利要求 2 所述的方法，其特征在于，还包括：
获取出现异常的代理节点信息；
从集群组网信息中清除所述出现异常的代理节点信息。
- 35 6、根据权利要求 2 的方法，其特征在于，所述向所述镜像服务集群内的计算节点提供更新后的集群组网信息，包括：
通过广播方式，向所述镜像服务集群内的计算节点发送更新后的集群组

网信息。

7、根据权利要求2所述的方法，其特征在于，获取所述出现异常的根节点信息，包括：

5 若根节点异常检测条件成立，则获取所述出现异常的根节点信息，所述根节点异常检测条件包括但不限于：达到根节点异常检测周期。

8、根据权利要求2所述的方法，其特征在于，获取所述出现异常的根节点信息，采用如下方式的至少一个：

若根节点的系统信息满足节点异常条件，则判定根节点出现异常；

10 若与根节点上的P2P服务装置进行心跳保活失败，则判定根节点出现异常；

若接收到根节点发送的异常信息，则判定根节点出现异常。

9、根据权利要求2所述的方法，其特征在于，

所述方法的执行主体为管理节点，所述管理节点包括：P2P管理装置，计算节点管理装置，管理服务装置；

15 所述计算节点包括：P2P服务装置，计算节点代理装置，容器加速装置；所述系统信息采用如下方式获取：

所述P2P管理装置通过与所述P2P服务装置进行心跳保活，获取所述系统信息。

10、根据权利要求9所述的方法，其特征在于，还包括：

20 所述P2P管理装置若判定计算节点信息获取条件成立，则通过所述计算节点管理装置，获取所述镜像服务集群的计算节点信息，所述计算节点信息获取条件包括但不限于：达到节点信息获取周期；

通过与所述P2P服务装置进行心跳保活，获取可用计算节点信息；

25 所述P2P管理装置通过与所述P2P服务装置进行心跳保活，获取所述系统信息，包括：

所述P2P管理装置通过与所述可用计算节点的P2P服务装置进行心跳保活，获取所述系统信息。

11、一种镜像加速方法，其特征在于，应用于计算节点，包括：

30 获取计算节点的系统信息，所述计算节点属于基于P2P网络结构的镜像服务集群；

35 向P2P管理装置提供所述系统信息，以使得P2P管理装置若确定所述镜像服务集群内存在出现异常的根节点，则根据所述镜像服务集群内的多个代理节点的系统信息，从所述多个代理节点中选取节点作为异常根节点的替代节点；从集群组网信息中清除所述出现异常的根节点信息，将集群组网信息中所述替代节点设置为根节点；

获取P2P管理装置提供的更新后的集群组网信息；

根据所述更新后的集群组网信息，处理镜像数据，所述根节点用于从镜像数据源获取镜像数据及根据镜像数据在虚拟运行环境中启动应用实例，所述代理节点用于根据镜像数据在虚拟运行环境中启动应用实例。

12、一种镜像加速方法，其特征在于，包括：

- 获取基于 P2P 网络结构的镜像服务集群内的多个代理节点的系统信息、
和出现异常的根节点信息，所述根节点用于从镜像数据源获取镜像数据及根
据镜像数据在虚拟运行环境中启动应用实例，所述代理节点用于根据镜像数
5 据在虚拟运行环境中启动应用实例；
根据所述系统信息，从所述多个代理节点中选取节点作为异常根节点的
替代节点；
从集群组网信息中清除所述出现异常的根节点信息，将集群组网信息中
所述替代节点设置为根节点；
10 根据更新后的集群组网信息，处理镜像数据。

13、一种容器镜像加速装置，其特征在于，包括：

- 信息获取单元，用于获取基于 P2P 网络结构的镜像服务集群内的多个代
理节点的系统信息、和出现异常的根节点信息，所述根节点用于从镜像数据
源获取镜像数据及根据镜像数据在虚拟运行环境中启动应用实例，所述代理
15 节点用于根据镜像数据在虚拟运行环境中启动应用实例；
根节点重新指派单元，用于根据所述系统信息，从所述多个代理节点中
选取节点作为异常根节点的替代节点；
组网信息更新单元，用于从集群组网信息中清除所述出现异常的根节点
信息，将集群组网信息中所述替代节点设置为根节点；
20 镜像数据处理单元，用于根据更新后的集群组网信息，处理镜像数据。

14、一种电子设备，其特征在于，包括：

- 处理器和存储器；
所述存储器用于存储实现根据权利要求 1 至 12 任一项所述方法的程序，
该终端通电并通过所述处理器运行该方法的程序。