



(12)发明专利申请

(10)申请公布号 CN 110908671 A

(43)申请公布日 2020.03.24

(21)申请号 201811083312.8

(22)申请日 2018.09.18

(71)申请人 北京京东尚科信息技术有限公司  
地址 100195 北京市海淀区杏石口路65号  
西杉创意园四区11号楼东段1-4层西  
段1-4层  
申请人 北京京东世纪贸易有限公司

(72)发明人 李海龙

(74)专利代理机构 中国国际贸易促进委员会专  
利商标事务所 11038  
代理人 姜雍 许蓓

(51)Int.Cl.  
G06F 8/61(2018.01)

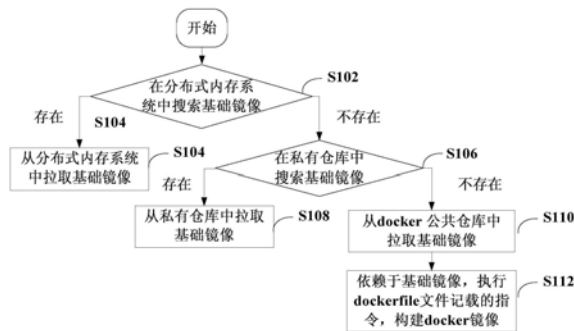
权利要求书3页 说明书8页 附图2页

(54)发明名称

构建docker镜像的方法、装置及计算机可读  
存储介质

(57)摘要

本公开提供了一种构建docker镜像的方法、  
系统、装置及计算机可读存储介质,涉及计算机  
技术领域。其中的构建docker镜像的方法包括:  
在分布式内存系统中搜索dockerfile文件记载  
的基础镜像;在分布式内存系统存有基础镜像的  
情况下,从分布式内存系统中拉取基础镜像;在  
分布式内存系统不存有基础镜像的情况下,在私  
有仓库中搜索基础镜像;在私有仓库存有基础镜  
像的情况下,从私有仓库中拉取基础镜像;在私  
有仓库不存有基础镜像的情况下,从docker公共  
仓库中拉取基础镜像;依赖于基础镜像,执行  
dockerfile文件记载的指令,构建docker镜像。  
本公开能够缩短docker镜像的构建时间,提高  
docker镜像的构建速度和构建效率。



1. 一种构建docker镜像的方法,包括:

在分布式内存系统中搜索dockerfile文件记载的基础镜像;

在分布式内存系统存有所述基础镜像的情况下,从分布式内存系统中拉取所述基础镜像;在分布式内存系统不存有所述基础镜像的情况下,在私有仓库中搜索所述基础镜像;

在私有仓库存有所述基础镜像的情况下,从私有仓库中拉取所述基础镜像;在私有仓库不存有所述基础镜像的情况下,从docker公共仓库中拉取所述基础镜像;

依赖于所述基础镜像,执行dockerfile文件记载的指令,构建docker镜像。

2. 如权利要求1所述的方法,其中,

私有仓库中的基础镜像,是复制docker公共仓库中拉取频率高于预设值的基础镜像得到的;

分布式内存系统中的基础镜像,是复制私有仓库中拉取频率排前N位的基础镜像得到的,N为分布式内存系统能存储基础镜像的最大数量。

3. 如权利要求1所述的方法,所述方法还包括:

将dockerfile文件中各个代码块的公用指令提取至各个代码块之前,并删除重复的公用指令,以便docker引擎优先执行dockerfile文件中的公用指令,后执行dockerfile文件中各个代码块的非公用指令。

4. 如权利要求1所述的方法,所述方法还包括:

为dockerfile文件记载的基础镜像或软件包标注版本号,以便docker引擎使用与版本号对应的基础镜像或软件包构建docker镜像。

5. 如权利要求1所述的方法,所述方法还包括:

从dockerfile文件中提取目标软件安装指令的软件名;

利用所述软件名查询预先定义的软件名与软件功能的映射关系,确定所述软件名对应的软件功能;

若dockerfile文件中已存在其它软件安装指令,且所述其它软件安装指令的软件名与所述软件功能相对应,则删除目标软件安装指令。

6. 如权利要求1所述的方法,所述方法还包括:

在dockerfile文件中,利用指令串联符号,将删除软件包的指令串联在其它操作所述软件包的指令之后,以便docker引擎在相同的镜像层对所述软件包执行其它操作之后删除所述软件包。

7. 如权利要求1所述的方法,所述方法还包括:

利用预先定义的软件功能分类表,确定属于相同软件功能类别的软件;

利用指令串联符号,将dockerfile文件操作相同软件功能类别的软件的指令进行串联,得到串联指令;

采用相同的run指令执行所述串联指令,以使得相同软件功能类别的软件位于所构建docker镜像的相同镜像层。

8. 如权利要求1所述的方法,所述执行dockerfile文件记载的指令,构建docker镜像包括:

在构建每层docker镜像结束前,删除预设文件目录中的安装文件和临时文件。

9. 如权利要求1所述的方法,所述执行dockerfile文件记载的指令,构建docker镜像包

括：

加载过滤文件，所述过滤文件包含构建上下文目录中需要被排除的文件和目录；  
打包构建上下文目录时，根据所述过滤文件排除所述文件和目录。

10. 一种构建docker镜像的装置，包括：

基础镜像搜索模块，被配置为：在分布式内存系统中搜索dockerfile文件记载的基础镜像；

基础镜像拉取模块，被配置为：在分布式内存系统存有所述基础镜像的情况下，从分布式内存系统中拉取所述基础镜像；在分布式内存系统不存有所述基础镜像的情况下，在私有仓库中搜索所述基础镜像；在私有仓库存有所述基础镜像的情况下，从私有仓库中拉取所述基础镜像；在私有仓库不存有所述基础镜像的情况下，从docker公共仓库中拉取所述基础镜像；

docker镜像构建模块，被配置为：依赖于所述基础镜像，执行dockerfile文件记载的指令，构建docker镜像。

11. 如权利要求10所述的装置，其中，

私有仓库中的基础镜像，是复制docker公共仓库中拉取频率高于预设值的基础镜像得到的；

分布式内存系统中的基础镜像，是复制私有仓库中拉取频率排前N位的基础镜像得到的，N为分布式内存系统能存储基础镜像的最大数量。

12. 如权利要求10所述的装置，所述装置还包括公用指令优化模块，被配置为：

将dockerfile文件中各个代码块的公用指令提取至各个代码块之前，并删除重复的公用指令，以便docker引擎优先执行dockerfile文件中的公用指令，后执行dockerfile文件中各个代码块的非公用指令。

13. 如权利要求10所述的装置，所述装置还包括版本号标注模块，被配置为：为dockerfile文件记载的基础镜像或软件包标注版本号，以便docker引擎使用与版本号对应的基础镜像或软件包构建docker镜像。

14. 如权利要求10所述的装置，所述装置还包括指令删除模块，被配置为：

从dockerfile文件中提取目标软件安装指令的软件名；

利用所述软件名查询预先定义的软件名与软件功能的映射关系，确定所述软件名对应的软件功能；

若dockerfile文件中已存在其它软件安装指令，且所述其它软件安装指令的软件名与所述软件功能相对应，则删除目标软件安装指令。

15. 如权利要求10所述的装置，所述装置还包括第一指令串联模块，被配置为：

在dockerfile文件中，利用指令串联符号，将删除软件包的指令串联在其它操作所述软件包的指令之后，以便docker引擎在相同的镜像层对所述软件包执行其它操作之后删除所述软件包。

16. 如权利要求10所述的装置，所述装置还包括第二指令串联模块，被配置为：

利用预先定义的软件功能分类表，确定属于相同软件功能类别的软件；

利用指令串联符号，将dockerfile文件操作相同软件功能类别的软件的指令进行串联，得到串联指令；

采用相同的run指令执行所述串联指令,以使得相同软件功能类别的软件位于所构建docker镜像的相同镜像层。

17. 如权利要求10所述的装置,所述docker镜像构建模块被配置为:在构建每层docker镜像结束前,删除预设文件目录中的安装文件和临时文件。

18. 如权利要求10所述的方法,所述docker镜像构建模块被配置为:

加载过滤文件,所述过滤文件包含构建上下文目录中需要被排除的文件和目录;

打包构建上下文目录时,根据所述过滤文件排除所述文件和目录。

19. 一种构建docker镜像的装置,包括:

存储器;以及

耦接至所述存储器的处理器,所述处理器被配置为基于存储在所述存储器中的指令,执行如权利要求1至9中任一项所述的构建docker镜像的方法。

20. 一种计算机可读存储介质,其中,所述计算机可读存储介质存储有计算机指令,所述指令被处理器执行时实现如权利要求1至9中任一项所述的构建docker镜像的方法。

## 构建docker镜像的方法、装置及计算机可读存储介质

### 技术领域

[0001] 本公开涉及计算机技术领域,特别涉及一种构建docker镜像的方法、装置及计算机可读存储介质。

### 背景技术

[0002] Docker是一个开源的应用容器引擎,让开发者可以将应用及依赖包打包至一个可移植的容器中,然后发布到linux机器上。基于linux平台上的多项开源技术,docker能够提供高效、敏捷、轻量级的容器,对应用的封装、分发、部署、运行的生命周期进行管理。

[0003] Docker镜像(image)是docker的核心概念,是用来启动容器的只读模版,是启动容器所需要的rootfs(root file system,根文件系统)。docker镜像由多个只读层文件系统叠加而成,启动一个容器的时候,docker会加载镜像层并在镜像层上添加一个可写层。容器上所做的任何更改都记录在可写层上。

[0004] Docker镜像的构建主要通过编写dockerfile文件来定制,在dockerfile中每运行一个run指令就会增加一个镜像层,并生成新的镜像文件。

### 发明内容

[0005] 本公开解决的一个技术问题是,如何提高docker镜像的构建速度。

[0006] 根据本公开实施例的一个方面,提供了一种构建docker镜像的方法,包括:在分布式内存系统中搜索dockerfile文件记载的基础镜像;在分布式内存系统存有基础镜像的情况下,从分布式内存系统中拉取基础镜像;在分布式内存系统不存有基础镜像的情况下,在私有仓库中搜索基础镜像;在私有仓库存有基础镜像的情况下,从私有仓库中拉取基础镜像;在私有仓库不存有基础镜像的情况下,从docker公共仓库中拉取基础镜像;依赖于基础镜像,执行dockerfile文件记载的指令,构建docker镜像。

[0007] 在一些实施例中,私有仓库中的基础镜像,是复制docker公共仓库中拉取频率高于预设值的基础镜像得到的;分布式内存系统中的基础镜像,是复制私有仓库中拉取频率排前N位的基础镜像得到的,N为分布式内存系统能存储基础镜像的最大数量。

[0008] 在一些实施例中,该方法还包括:将dockerfile文件中各个代码块的公用指令提取至各个代码块之前,并删除重复的公用指令,以便docker引擎优先执行dockerfile文件中的公用指令,后执行dockerfile文件中各个代码块的非公用指令。

[0009] 在一些实施例中,该方法还包括:为dockerfile文件记载的基础镜像或软件包标注版本号,以便docker引擎使用与版本号对应的基础镜像或软件包构建docker镜像。

[0010] 在一些实施例中,该方法还包括:从dockerfile文件中提取目标软件安装指令的软件名;利用软件名查询预先定义的软件名与软件功能的映射关系,确定软件名对应的软件功能;若dockerfile文件中已存在其它软件安装指令,且其它软件安装指令的软件名与软件功能相对应,则删除目标软件安装指令。

[0011] 在一些实施例中,该方法还包括:在dockerfile文件中,利用指令串联符号,将删

除软件包的指令串联在其它操作软件包的指令之后,以便docker引擎在相同的镜像层对软件包执行其它操作之后删除软件包。

[0012] 在一些实施例中,该方法还包括:利用预先定义的软件功能分类表,确定属于相同软件功能类别的软件;利用指令串联符号,将dockerfile文件操作相同软件功能类别的软件的指令进行串联,得到串联指令;采用相同的run指令执行串联指令,以使得相同软件功能类别的软件位于所构建docker镜像的相同镜像层。

[0013] 在一些实施例中,执行dockerfile文件记载的指令,构建docker镜像包括:在构建每层docker镜像结束前,删除预设文件目录中的安装文件和临时文件。

[0014] 在一些实施例中,执行dockerfile文件记载的指令,构建docker镜像包括:加载过滤文件,过滤文件包含构建上下文目录中需要被排除的文件和目录;打包构建上下文目录时,根据过滤文件排除文件和目录。

[0015] 根据本公开实施例的一个方面,提供了一种构建docker镜像的装置,包括:基础镜像搜索模块,被配置为:在分布式内存系统中搜索dockerfile文件记载的基础镜像;基础镜像拉取模块,被配置为:在分布式内存系统存有基础镜像的情况下,从分布式内存系统中拉取基础镜像;在分布式内存系统不存有基础镜像的情况下,在私有仓库中搜索基础镜像;在私有仓库存有基础镜像的情况下,从私有仓库中拉取基础镜像;在私有仓库不存有基础镜像的情况下,从docker公共仓库中拉取基础镜像;docker镜像构建模块,被配置为:依赖于基础镜像,执行dockerfile文件记载的指令,构建docker镜像。

[0016] 在一些实施例中,私有仓库中的基础镜像,是复制docker公共仓库中拉取频率高于预设值的基础镜像得到的;分布式内存系统中的基础镜像,是复制私有仓库中拉取频率排前N位的基础镜像得到的,N为分布式内存系统能存储基础镜像的最大数量。

[0017] 在一些实施例中,该装置还包括公用指令优化模块,被配置为:将dockerfile文件中各个代码块的公用指令提取至各个代码块之前,并删除重复的公用指令,以便docker引擎优先执行dockerfile文件中的公用指令,后执行dockerfile文件中各个代码块的非公用指令。

[0018] 在一些实施例中,该装置还包括版本号标注模块,被配置为:为dockerfile文件记载的基础镜像或软件包标注版本号,以便docker引擎使用与版本号对应的基础镜像或软件包构建docker镜像。

[0019] 在一些实施例中,该装置还包括指令删除模块,被配置为:从dockerfile文件中提取目标软件安装指令的软件名;利用软件名查询预先定义的软件名与软件功能的映射关系,确定软件名对应的软件功能;若dockerfile文件中已存在其它软件安装指令,且其它软件安装指令的软件名与软件功能相对应,则删除目标软件安装指令。

[0020] 在一些实施例中,该装置还包括第一指令串联模块,被配置为:在dockerfile文件中,利用指令串联符号,将删除软件包的指令串联在其它操作软件包的指令之后,以便docker引擎在相同的镜像层对软件包执行其它操作之后删除软件包。

[0021] 在一些实施例中,该装置还包括第二指令串联模块,被配置为:利用预先定义的软件功能分类表,确定属于相同软件功能类别的软件;利用指令串联符号,将dockerfile文件操作相同软件功能类别的软件的指令进行串联,得到串联指令;采用相同的run指令执行串联指令,以使得相同软件功能类别的软件位于所构建docker镜像的相同镜像层。

[0022] 在一些实施例中,docker镜像构建模块被配置为:在构建每层docker镜像结束前,删除预设文件目录中的安装文件和临时文件。

[0023] 在一些实施例中,docker镜像构建模块被配置为:加载过滤文件,过滤文件包含构建上下文目录中需要被排除的文件和目录;打包构建上下文目录时,根据过滤文件排除文件和目录。

[0024] 根据本公开实施例的又一个方面,提供了一种构建docker镜像的装置,包括:存储器;以及耦接至存储器的处理器,处理器被配置为基于存储在存储器中的指令,执行前述的构建docker镜像的方法。

[0025] 根据本公开实施例的再一个方面,提供了一种计算机可读存储介质,其中,计算机可读存储介质存储有计算机指令,指令被处理器执行时实现前述的构建docker镜像的方法。

[0026] 本公开能够缩短docker镜像的构建时间,提高docker镜像的构建速度和构建效率。

[0027] 通过以下参照附图对本公开的示例性实施例的详细描述,本公开的其它特征及其优点将会变得清楚。

#### 附图说明

[0028] 为了更清楚地说明本公开实施例或现有技术中的技术方案,下面将对实施例或现有技术描述中所需要使用的附图作简单地介绍,显而易见地,下面描述中的附图仅仅是本公开的一些实施例,对于本领域普通技术人员来讲,在不付出创造性劳动性的前提下,还可以根据这些附图获得其他的附图。

[0029] 图1示出了本公开一些实施例的构建docker镜像的方法的流程示意图。

[0030] 图2示出了本公开一些实施例的构建docker镜像的装置的结构示意图。

[0031] 图3示出了本公开另一些实施例的构建docker镜像的装置的结构示意图。

[0032] 图4示出了本公开又一些实施例的构建docker镜像的装置的结构示意图。

#### 具体实施方式

[0033] 下面将结合本公开实施例中的附图,对本公开实施例中的技术方案进行清楚、完整地描述,显然,所描述的实施例仅仅是本公开一部分实施例,而不是全部的实施例。以下对至少一个示例性实施例的描述实际上仅仅是说明性的,决不作为对本公开及其应用或使用的任何限制。基于本公开中的实施例,本领域普通技术人员在没有作出创造性劳动前提下所获得的所有其它实施例,都属于本公开保护的范围。

[0034] 发明人研究发现,构建docker镜像时需要复用现有的基础镜像,因此构建docker镜像的传统方法默认在docker hub中搜索并拉取需要依赖的基础镜像.docker hub是docker官方维护的公共仓库,存有现有的基础镜像。由于docker hub属于国外的网站,因此构建docker镜像的传统方法依赖于访问docker hub的网络速度和网络稳定性。同时,软件尺寸较大的镜像需要更长的镜像拉取时间。以上原因造成了docker镜像的构建时间长,构建速度慢,构建效率低下。

[0035] 针对以上问题,本公开提供了一种构建docker镜像的方法,能够缩短docker镜像

的构建时间,提高docker镜像的构建速度和构建效率。

[0036] 首先结合图1描述本公开构建docker镜像的方法的一些实施例。

[0037] 图1示出了本公开一些实施例的构建docker镜像的方法的流程示意图。如图1所示,本实施例包括步骤S102~步骤S110。

[0038] 在步骤S102中,在分布式内存系统中搜索dockerfile文件记载的基础镜像。

[0039] 其中,分布式内存系统中的基础镜像,是复制私有仓库中拉取频率排前N位(例如前100位)的基础镜像得到的,N为分布式内存系统能存储基础镜像的最大数量。私有仓库中的基础镜像,是复制docker公共仓库(docker hub)中拉取频率高于预设值(例如高于5%)的基础镜像得到的。本领域技术人员应理解,由于私有仓库中基础镜像的拉取频率是动态更新的,因此分布式内存系统中的基础镜像也是动态更新的,分布式内存系统中LRU(Least Recently Used,最近最久未使用)的基础镜像会被替换掉,移出分布式内存系统。

[0040] Dockerfile是一个文本格式的配置文​​件,用户可以使用dockerfile来快速构建自定义的镜像。Dockerfile由一行行命令行语句组成,主要包括:基础镜像信息、维护者信息、镜像操作指令和容器启动时执行指令。Dockerfile中首先描述所依赖的基础镜像名称,接下来是说明维护者信息,然后是镜像操作指令,(例如run指令,用于对镜像执行跟随的命令),最后是cmd指令(用来指定运行容器时的操作命令)。由于dockerfile文件会记载所依赖的基础镜像名称,通过依赖的基础镜像名称可以搜索基础镜像。搜索时,判断分布式内存系统是否存有该基础镜像。若是,则执行步骤S104;若否,则执行步骤S106。

[0041] 在步骤S104中,从分布式内存系统中拉取基础镜像。

[0042] 分布式内存系统例如可以采用tachyon等分布式内存系统。从分布式内存系统中搜索和拉取镜像速度快(秒级),且访问分布式内存系统比访问docker公共仓库或私有仓库更加稳定,极大地缩短了镜像的搜索时间及拉取时间,进而缩短镜像构建时间。

[0043] 在步骤S106中,在私有仓库中搜索基础镜像。

[0044] 搜索时,判断私有仓库是否存有该基础镜像。若是,则执行步骤S108;若否,则执行步骤S110。

[0045] 在步骤S108中,从私有仓库中拉取基础镜像。

[0046] 在步骤S110中,从docker公共仓库中拉取基础镜像。

[0047] 在步骤S112中,依赖于基础镜像,执行dockerfile文件记载的指令,构建docker镜像。

[0048] 例如,docker守护进程依赖于基础镜像,逐层解析dockerfile文件,构建出新的docker镜像。

[0049] 上述实施例中,提前将docker hub中的已有镜像拉取至私有仓库,并提前将私有仓库中拉取频率较高的已有镜像进一步拉取至内存中。在需要复用已有镜像时,按照既定的先后顺序,优先选择从内存中搜索和拉取已有镜像,次优选择从私有仓库中搜索和拉取已有镜像,从而大幅缩短docker镜像的构建时长,提高docker镜像的构建速度和构建效率。在访问docker hub时网络不稳定的情况下,或者在已有镜像的软件尺寸较大的情况下,上述实施例构建docker镜像的速度提升效果尤为显著。

[0050] 在一些实施例中,步骤S112可以包括:在构建每层docker镜像结束前,删除预设文件目录中的安装文件和临时文件。



[0051] 在构建镜像的过程中会产生大量的临时文件、安装文件或临时目录,例如保存在包管理器中的临时文件或者在dockerfile中需要安装的软件源代码。发明人发现,如果对这些文件或目录不做处理,会大幅增加镜像的尺寸。有鉴于此,在构建每层docker镜像结束前,及时自动清理/var/lib/apt/list/目录中的安装文件以及预设的临时文件目录中的临时文件,可以大幅减少镜像尺寸。

[0052] 在一些实施例中,步骤S112中还包括:加载过滤文件,过滤文件包含构建上下文目录中需要被排除的文件和目录;打包构建上下文目录时,根据过滤文件排除文件和目录。

[0053] 构建docker镜像时,客户端发起构建指令传递给docker引擎(守护进程)之后,docker会将指定的构建上下文目录打包传递给docker引擎,该构建上下文目录中并非所有文件都会在dockerfile中使用到。对不会被使用的文件或目录进行打包,会严重影响了镜像构建的性能,降低构建镜像的速度。

[0054] 有鉴于此,本实施例配置了过滤文件,过滤文件包含构建上下文目录中需要被排除的文件和目录。在生成docker镜像过程中打包构建上下文目录时,docker引擎加载过滤文件,根据过滤文件排除文件和目录(例如临时目录、临时文件、readme文件等等),从而精简了打包的内容,减小所构建镜像的尺寸,并缩短了镜像构建时间。

[0055] 发明人还发现,传统方法构建docker镜像需要花费加多的时间,还与dockerfile中指令的执行效率低下有关。Dockerfile中指令的执行顺序和执行方式,不仅影响镜像的构建速度,还影响镜像占用的空间。编写高质量的dockerfile文件要求非常有经验的docker开发人员。因此,发明人提供了dockerfile文件指令优化的功能,对初学者开发的dockerfile文件指令进行以下几个方面的优化。

[0056] (一)提取公用指令:将dockerfile文件中各个代码块的公用指令提取至各个代码块之前,并删除重复的公用指令,以便docker引擎优先执行dockerfile文件中的公用指令,后执行dockerfile文件中各个代码块的非公用指令。

[0057] Dockerfile文件中,各个代码块中很多指令都是公用的,例如数字认证指令、文件拷贝指令、目录创建指令等等。将这些公用指令提取至各个代码块之前执行,并删除重复的公用指令,实现了对指令的执行顺序以及执行方式的优化。本实施例提供的方法不仅能够提升dockerfile文件的指令执行效率,缩减构建docker镜像的时间,还能够减小构建镜像的尺寸。此外,本实施例提供的方法利于dockerfile文件的指令维护。

[0058] (二)固化软件版本:为dockerfile文件记载的基础镜像或软件包标注版本号,以便docker引擎使用与版本号对应的基础镜像或软件包构建docker镜像。

[0059] 如果dockerfile文件没有记载基础镜像或软件包的版本号,docker默认会使用最新的镜像版本或软件版本进行处理。最新版本的基础镜像或软件包,可能删除构建docker镜像时需要的某些功能,从而造成构建失败。即使最新版本的基础镜像或软件包没有删除构建docker镜像时需要的功能,也会造成构建镜像尺寸的急剧膨胀,从而带来很多不确定性。

[0060] 为dockerfile文件记载的基础镜像或软件包标注版本号,例如,使用软件包版本号为ubuntu16.04或ubuntu14.04等等,能够使构建生成的镜像占用相对稳定的空间,从而减小构建镜像的尺寸。

[0061] (三)精简软件指令:首先,从dockerfile文件中提取目标软件安装指令的软件名。

然后,利用软件名查询预先定义的软件名与软件功能的映射关系,确定软件名对应的软件功能。若dockerfile文件中已存在其它软件安装指令,且其它软件安装指令的软件名与软件功能相对应,则删除目标软件安装指令。

[0062] 例如,构建镜像时已经选择使用了centos软件,centos软件附带了ssh服务和mysql数据库功能,即预先定义的软件名与软件功能的映射关系中Centos对应SSH服务和mysql数据库功能。那么,在dockerfile文件中的后续软件安装指令中,可以删除涉及ssh服务或mysql数据库的指令,不再对ssh服务或mysql数据库进行单独安装。

[0063] 通过精简指令,可以减小构建镜像的软件尺寸,进一步缩减编译和部署docker应用的时间,从而缩短docker镜像的构建时长,提高docker镜像的构建速度和构建效率。

[0064] (四) 删除软件包:在dockerfile文件中,利用指令串联符号,将删除软件包的指令串联在其它操作软件包的指令之后,以便docker引擎在相同的镜像层对软件包执行其它操作之后删除软件包。

[0065] 在构建镜像时,对软件的相关指令编译完成后,这个软件的安装包就没有作用了。如果单独执行删除软件包指令,根据docker镜像的分层原理,这个软件包是删除不掉的,导致构建docker镜像的软件尺寸较大。

[0066] 有鉴于此,根据dockerfile编写规范,操作相同软件进行的指令通常位于相同的代码块。扫描这个代码块中的指令,将这些指令按照原有的指令顺序串联在一起。Docker引擎在相同的镜像层对软件包执行其它的操作指令之后在执行删除指令,才能真正将软件包删除,从而减小构建镜像的软件尺寸,同时便于对指令的维护。

[0067] (五) 合并镜像层级:首先,利用预先定义的软件功能分类表,确定属于相同软件功能类别的软件。然后,利用指令串联符号,将dockerfile文件操作相同软件功能类别的软件的指令进行串联,得到串联指令。最后,采用相同的run指令执行串联指令,以使得相同软件功能类别的软件位于所构建docker镜像的相同镜像层。

[0068] 构建镜像时,dockerfile文件中每增加一个run指令就会增加一个镜像层。镜像层的最大值是127,超过127层时docker会合并镜像层。由于docker在合并镜像层时需要对每一层与上一层进行对比,镜像层级越多耗时越多,使得构建镜像的效率非常低。同时,如果dockerfile中run指令太多,例如一个简单的拷贝指令或删除指令就会生成一层镜像层,这样的docker镜像很难复用。

[0069] 针对这种情况,发明人提供了镜像层级合并的功能。将功能相近软件的操作指令串联在一起,采用相同的run指令执行。在构建docker镜像时,这些软件就会融合在相同的镜像层,从而减少了镜像层级的数量,在大幅提高镜像构建的速度同时,显著减小镜像的尺寸。不仅能够提升docker引擎的工作性能,还能够节约存储资源。

[0070] 下面结合图2描述本公开一些实施例的构建docker镜像的装置。

[0071] 图2示出了本公开一些实施例的构建docker镜像的装置的结构示意图。如图2所示,本实施例中的构建docker镜像的装置20包括:

[0072] 基础镜像搜索模块202,被配置为:在分布式内存系统中搜索dockerfile文件记载的基础镜像;基础镜像拉取模块204,被配置为:在分布式内存系统存有基础镜像的情况下,从分布式内存系统中拉取基础镜像;在分布式内存系统不存有基础镜像的情况下,在私有仓库中搜索基础镜像;在私有仓库存有基础镜像的情况下,从私有仓库中拉取基础镜像;在

私有仓库不存有基础镜像的情况下,从docker公共仓库中拉取基础镜像;docker镜像构建模块206,被配置为:依赖于基础镜像,执行dockerfile文件记载的指令,构建docker镜像。

[0073] 在一些实施例中,私有仓库中的基础镜像,是复制docker公共仓库中拉取频率高于预设值的基础镜像得到的;分布式内存系统中的基础镜像,是复制私有仓库中拉取频率排前N位的基础镜像得到的,N为分布式内存系统能存储基础镜像的最大数量。

[0074] 在一些实施例中,docker镜像构建模块206被配置为:在构建每层docker镜像结束前,删除预设文件目录中的安装文件和临时文件。

[0075] 在一些实施例中,docker镜像构建模块206被配置为:加载过滤文件,过滤文件包含构建上下文目录中需要被排除的文件和目录;打包构建上下文目录时,根据过滤文件排除文件和目录。

[0076] 下面结合图3描述本公开另一些实施例的构建docker镜像的装置。

[0077] 图3示出了本公开另一些实施例的构建docker镜像的装置的结构示意图。

[0078] 在一些实施例中,构建docker镜像的装置30还包括:公用指令优化模块3051,被配置为:将dockerfile文件中各个代码块的公用指令提取至各个代码块之前,并删除重复的公用指令,以便docker引擎优先执行dockerfile文件中的公用指令,后执行dockerfile文件中各个代码块的非公用指令。

[0079] 在一些实施例中,该装置30还包括版本号标注模块3053,被配置为:为dockerfile文件记载的基础镜像或软件包标注版本号,以便docker引擎使用与版本号对应的基础镜像或软件包构建docker镜像。

[0080] 在一些实施例中,该装置30还包括指令删除模块3055,被配置为:从dockerfile文件中提取目标软件安装指令的软件名;利用软件名查询预先定义的软件名与软件功能的映射关系,确定软件名对应的软件功能;若dockerfile文件中已存在其它软件安装指令,且其它软件安装指令的软件名与软件功能相对应,则删除目标软件安装指令。

[0081] 在一些实施例中,该装置30还包括第一指令串联模块3057,被配置为:在dockerfile文件中,利用指令串联符号,将删除软件包的指令串联在其它操作软件包的指令之后,以便docker引擎在相同的镜像层对软件包执行其它操作之后删除软件包。

[0082] 在一些实施例中,该装置30还包括第二指令串联模块3059,被配置为:利用预先定义的软件功能分类表,确定属于相同软件功能类别的软件;利用指令串联符号,将dockerfile文件操作相同软件功能类别的软件的指令进行串联,得到串联指令;采用相同的run指令执行串联指令,以使得相同软件功能类别的软件位于所构建docker镜像的相同镜像层。

[0083] 图4示出了本公开又一些实施例的构建docker镜像的装置的结构示意图。如图4所示,该实施例的构建docker镜像的装置40包括:存储器410以及耦接至该存储器410的处理器420,处理器420被配置为基于存储在存储器410中的指令,执行前述任意一些实施例中的构建docker镜像的方法。

[0084] 其中,存储器410例如可以包括系统存储器、固定非易失性存储介质等。系统存储器例如存储有操作系统、应用程序、引导装载程序(Boot Loader)以及其他程序等。

[0085] 构建docker镜像的装置40还可以包括输入输出接口430、网络接口440、存储接口450等。这些接口430、440、450以及存储器410和处理器420之间例如可以通过总线460连接。

其中,输入输出接口430为显示器、鼠标、键盘、触摸屏等输入输出设备提供连接接口。网络接口440为各种联网设备提供连接接口。存储接口450为SD卡、U盘等外置存储设备提供连接接口。

[0086] 本公开还包括一种计算机可读存储介质,其上存储有计算机指令,该指令被处理器执行时实现前述任意一些实施例中的构建docker镜像的方法。

[0087] 本领域内的技术人员应明白,本公开的实施例可提供为方法、系统、或计算机程序产品。因此,本公开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,本公开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非瞬时性存储介质(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。

[0088] 本公开是参照根据本公开实施例的方法、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。

[0089] 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。

[0090] 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。

[0091] 以上所述仅为本公开的较佳实施例,并不用以限制本公开,凡在本公开的精神和原则之内,所作的任何修改、等同替换、改进等,均应包含在本公开的保护范围之内。

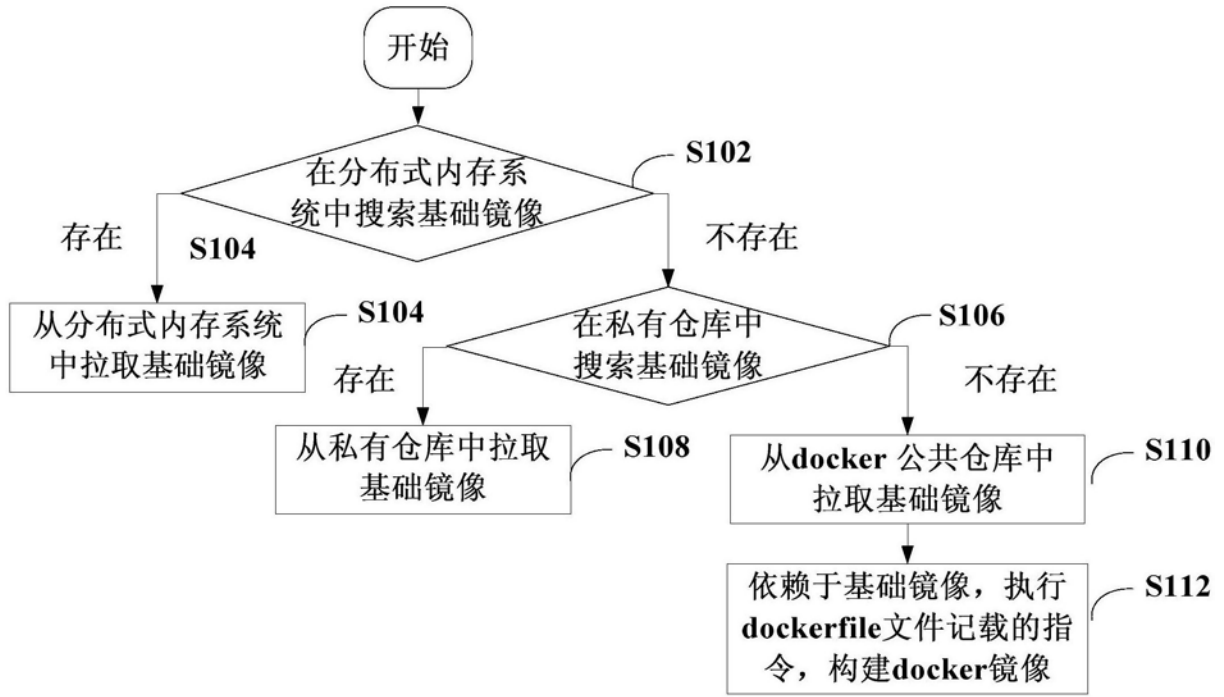


图1

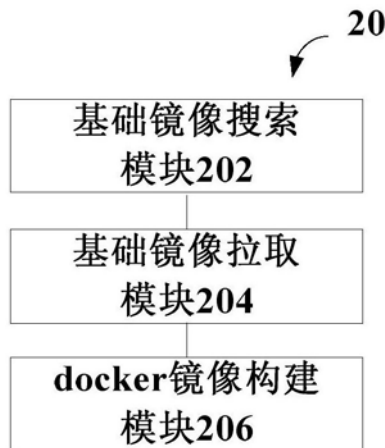


图2

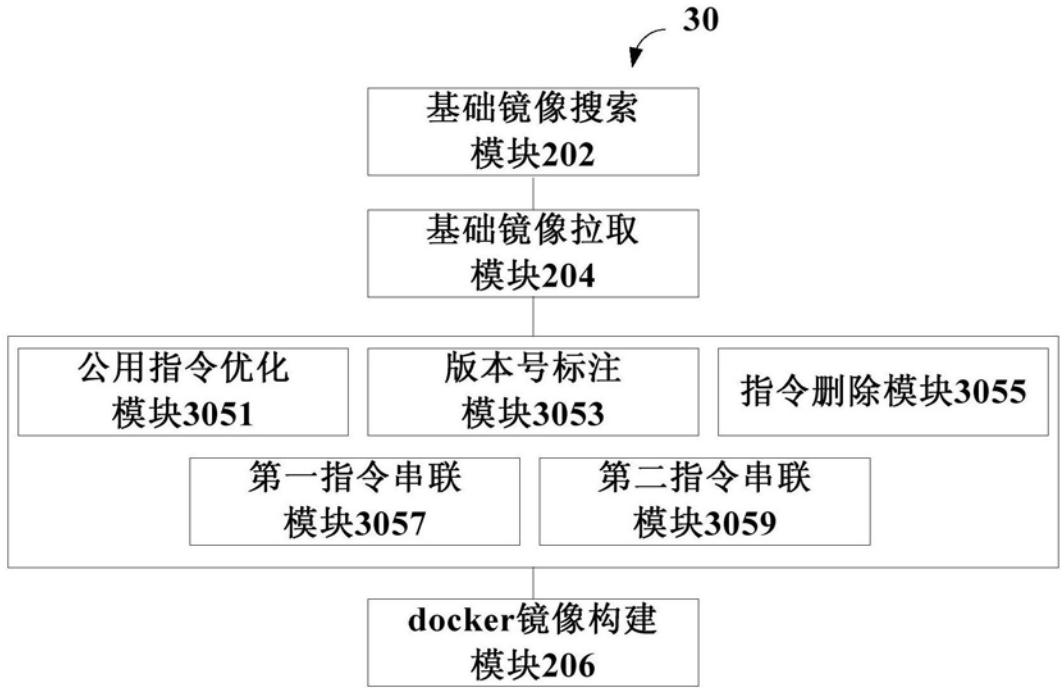


图3

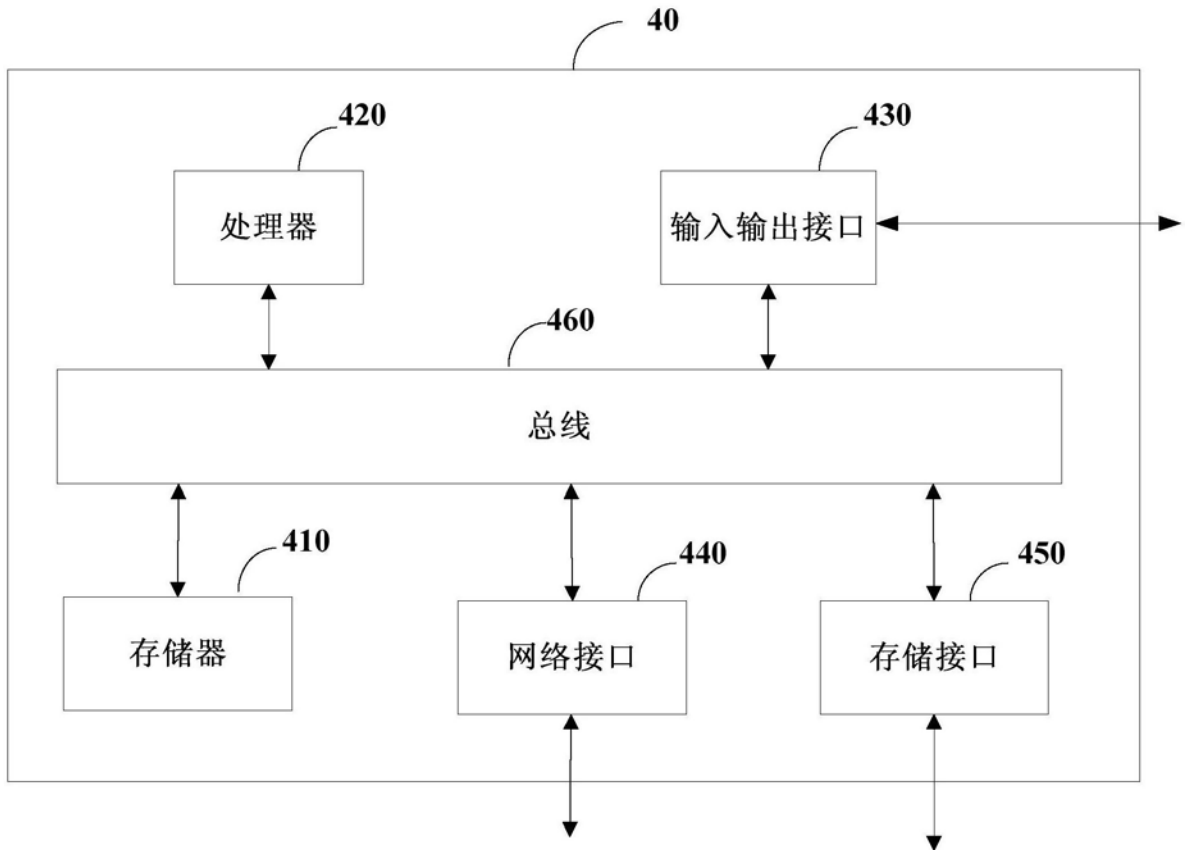


图4