



(12)发明专利

(10)授权公告号 CN 105979009 B

(45)授权公告日 2019.05.17

(21)申请号 201610529884.9

H04L 29/12(2006.01)

(22)申请日 2016.07.06

(56)对比文件

(65)同一申请的已公布的文献号  
申请公布号 CN 105979009 A

CN 105681217 A,2016.06.15,

CN 105468362 A,2016.04.06,

CN 104038540 A,2014.09.10,

US 9195510 B2,2015.11.24,

WO 2013122815 A1,2013.08.22,

(43)申请公布日 2016.09.28

(73)专利权人 乾云数创(山东)信息技术研究院  
有限公司

审查员 袁欣

地址 250002 山东省济南市市中区经七路  
28-1号901室

(72)发明人 吴恒 张辉

(74)专利代理机构 济南圣达知识产权代理有限公司 37221

代理人 张勇

(51)Int.Cl.

H04L 29/08(2006.01)

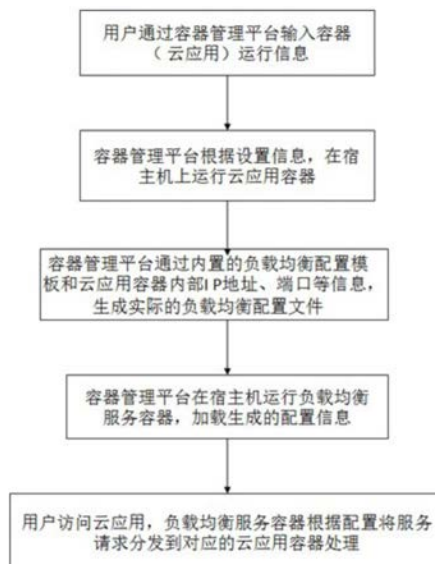
权利要求书2页 说明书4页 附图1页

(54)发明名称

一种针对云应用容器的增加负载自动均衡方法

(57)摘要

本发明公开了一种针对云应用容器的增加负载自动均衡方法,在容器云应用部署时不需要单独外部配置和的负载均衡容器实例或虚拟机节点,并自动根据需要来启用宿主机节点运行的代理实例容器来提供容器云应用负载均衡服务,简化云应用部署步骤,提升部署效率,节约运维人员在生产环境中部署容器云应用的工作量。同时,结合采集到的资源占用情况,可根据不同时间段的访问请求负载情况自动添加、减少运行的云应用容器数量,节约资源使用,减少云应用容器死机概率,提高云应用容器可用性。



1. 一种针对云应用容器的增加负载自动均衡方法,其特征是:包括以下步骤:

(1) 注册宿主机信息,在容器云应用服务平台管理软件注册管理的宿主机节点,通过在宿主机节点运行注册指令,保持宿主机节点和容器云应用服务平台的信息通讯,并在宿主机节点运行容器云应用服务平台的指令,具体包括创建云应用容器、采集宿主机资源占用情况、宿主机健康检查和负载均衡服务;

(2) 创建容器,输入云应用容器运行环境的镜像名称及版本信息,根据需要设置云应用容器运行参数信息;

(3) 根据设置信息,在宿主机上运行云应用容器,自动分配每个云应用容器的内部IP地址信息;

(4) 调用运行环境的内置负载均衡配置模板信息,根据实际的配置要求,修改云应用容器的IP地址和端口信息,生成服务对应需要的具体负载均衡配置文件信息;

(5) 在宿主机上运行负载均衡服务容器,加载生成的负载均衡配置文件信息;

(6) 访问云应用,负载均衡服务容器根据服务请求,将其分发到对应的云应用容器。

2. 如权利要求1所述的一种针对云应用容器的增加负载自动均衡方法,其特征是:所述步骤(1)中,具体步骤包括:注册宿主机信息,在容器云应用服务平台管理软件注册管理的宿主机节点,通过在宿主机节点运行注册指令,在每个宿主机节点上创建容器云应用服务平台网络代理容器,通过网络代理容器,来保持宿主机节点和容器云应用服务平台的信息通讯,接收容器云应用服务平台的指令,并在宿主机上运行相应的指令。

3. 如权利要求1所述的一种针对云应用容器的增加负载自动均衡方法,其特征是:所述步骤(2)中,云应用容器运行参数信息具体包括端口信息、访问协议、运行实例数量、内外端口映射信息、宿主机CPU、内存资源使用信息和环境变量信息。

4. 如权利要求1所述的一种针对云应用容器的增加负载自动均衡方法,其特征是:所述步骤(3)中,根据用户输入的信息,在所在环境的容器宿主机上生成运行相应的应用容器对应的创建命令信息,运行命令并启动应用容器,平台通过内置的网络模块,自动分配每个容器的内部IP地址信息,保证容器之间网络及数据的互联互通。

5. 如权利要求1所述的一种针对云应用容器的增加负载自动均衡方法,其特征是:所述步骤(3)中,根据注册和配置信息,在所管理的宿主机节点通过代理容器实例上创建并启动对应数量的云应用容器实例,并根据平台内置网络IP规则自动分配每个云应用容器的内部IP地址和端口信息。

6. 如权利要求1所述的一种针对云应用容器的增加负载自动均衡方法,其特征是:所述步骤(4)中,容器云应用服务平台依据云应用创建时输入的协议类型、端口信息、访问上下文信息及端口映射信息,在宿主机创建的云应用容器实例的内部IP地址和端口信息,生成LB服务元数据信息,并将信息发送给宿主机节点上运行的代理实例容器。

7. 如权利要求1所述的一种针对云应用容器的增加负载自动均衡方法,其特征是:所述步骤(4)中,宿主机节点的代理实例容器收到指令信息后,调用内置代码生成工具,根据内置的负载均衡算法规则信息和接收到的LB服务元数据信息,通过内置的文件生成工具,创建出云应用服务对应需要的负载均衡配置文件。

8. 如权利要求1所述的一种针对云应用容器的增加负载自动均衡方法,其特征是:所述步骤(5)中,在宿主机上通讯代理实例容器,启动内置基于负载均衡软件的负载均衡服务容

器,加载生成的负载均衡配置文件,在宿主机节点运行负载均衡分发服务。

9. 如权利要求1所述的一种针对云应用容器的增加负载自动均衡方法,其特征是:所述步骤(6)中,用户通过浏览器或客户端访问云应用服务外部服务地址,宿主机节点接收到请求后,由宿主机节点运行负载均衡分发服务,根据负载均衡配置文件中的服务分发规则将请求分发到宿主机节点上运行的云应用容器内部IP和端口,云应用容器接收到请求信息,处理对应的请求,并将处理结果通过负载均衡服务容器返回给用户的浏览器或客户端。

## 一种针对云应用容器的增加负载自动均衡方法

### 技术领域

[0001] 本发明涉及一种针对云应用容器的增加负载自动均衡方法。

### 背景技术

[0002] 如果使用开源的应用容器引擎(Docker)来运行云应用,通常就使用Nginx、HAproxy等Http server做云应用集群的负载均衡。通常是配置完Web应用容器后,再手工配置Nginx容器,来指向配置的云应用,如:

[0003] Docker容器会随机分配IP和端口号,默认情况下,IP和端口号是专用于容器的,不能被外部访问所以从客户端寻找它们会变得异常复杂,当部署一个云应用服务容器,每次重启或增加新的云应用容器实例个数后,云应用容器地址都会发生改变,这时就存在正确的访问服务问题。

[0004] 目前的处理方法是将云应用容器IP地址绑定到容器所在的宿主机上,绑定云应用容器到宿主机端口存在一个问题,不支持在同一台主机运行多个容器。例如,在同一个时间内只能有一个容器可以绑定到80端口。同时这种方案在扩展方面存在缺陷,在云应用新容器实例开始运行之前,原来的云应用容器实例必须停止。

### 发明内容

[0005] 本发明为了解决上述问题,提出了一种针对云应用容器的增加负载自动均衡方法,本方法云应用服务平台根据云应用容器的配置信息自动生成负载均衡配置信息,并运行自动运行负载均衡服务,完成容器云应用的自动负载均衡。

[0006] 为了实现上述目的,本发明采用如下技术方案:

[0007] 一种针对云应用容器的增加负载自动均衡方法,包括以下步骤:

[0008] (1)注册宿主机信息,在容器云应用服务平台管理软件注册管理的宿主机节点,通过在宿主机节点运行注册指令,保持宿主机节点和容器云应用服务平台的信息通讯;

[0009] (2)创建容器,输入云应用容器运行环境的镜像名称及版本信息,根据需要设置云应用容器运行参数信息;

[0010] (3)根据设置信息,在宿主机上运行云应用容器,自动分配每个云应用容器的内部IP地址信息;

[0011] (4)调用运行环境的内置负载均衡配置模板信息,根据实际的配置要求,修改云应用容器的IP地址和端口信息,生成服务对应需要的具体负载均衡配置文件信息;

[0012] (5)在宿主机上运行负载均衡服务容器,加载生成的负载均衡配置文件信息;

[0013] (6)访问云应用,负载均衡服务容器根据服务请求,将其分发到对应的云应用容器。

[0014] 所述步骤(1)中,具体步骤包括:注册宿主机信息,在容器云应用服务平台管理软件注册管理的宿主机节点,通过在宿主机节点运行注册指令,在每个宿主机节点上创建容器云应用服务平台网络代理容器,通过网络代理容器,来保持宿主机节点和容器云应用服

务平台的信息通讯,接收容器云应用服务平台的指令,并在宿主机上运行相应的指令。

[0015] 所述步骤(1)中,容器云应用服务平台的指令包括创建云应用容器、采集宿主机资源占用情况、宿主机健康检查和负载均衡服务。

[0016] 所述步骤(2)中,云应用容器运行参数信息具体包括端口信息、访问协议、运行实例数量、内外端口映射信息、宿主机CPU、内存资源使用信息和环境变量信息。

[0017] 所述步骤(2)中,根据用户输入的信息,在所在环境的容器宿主主机上生成运行相应的应用容器对应的创建命令信息,运行命令并启动应用容器,平台通过内置的网络模块,自动分配每个容器的内部IP地址信息,保证容器之间网络及数据的互联互通。

[0018] 所述步骤(3)中,根据步骤(1)中的注册和配置信息,在所管理的宿主机节点通过代理容器实例上创建并启动对应数量的云应用容器实例,并根据平台内置网络IP规则自动分配每个云应用容器的内部IP地址和端口信息。

[0019] 所述步骤(4)中,容器云应用服务平台依据云应用创建时输入的协议类型、端口信息、访问上下文信息及端口映射信息,在宿主机创建的云应用容器实例的内部IP地址和端口信息,生成LB服务元数据信息,并将信息发送给宿主机节点上运行的代理实例容器。

[0020] 所述步骤(4)中,宿主机节点的代理实例容器收到指令信息后,调用内置代码生成工具,根据内置的负载均衡算法规则信息和接收到的LB服务元数据信息,通过内置的文件生成工具,创建出云应用服务对应需要的负载均衡配置文件。

[0021] 所述步骤(5)中,在宿主机上通讯代理实例容器,启动内置基于负载均衡软件的内置负载模块,加载生成的负载均衡配置文件,在宿主机节点运行负载均衡分发服务。

[0022] 所述步骤(6)中,用户通过浏览器或客户端访问云应用服务外部服务地址,宿主机节点接收到请求后,由宿主机节点运行负载均衡分发服务,根据负载均衡配置文件中的服务分发规则将请求分发到宿主机节点上运行的云应用容器内部IP和端口,云应用容器接收到请求信息,处理对应的请求,并将处理结果通过负载均衡服务模块返回给用户的浏览器或客户端。

[0023] 本发明的有益效果为:

[0024] (1) 本发明在容器云应用部署时不需要单独外部配置和的负载均衡容器实例或虚拟机节点,并自动根据需要来启用宿主机节点运行的代理实例容器来提供容器云应用负载均衡服务,简化云应用部署步骤,提升部署效率,节约运维人员在生产环境中部署容器云应用的工作量;

[0025] (2) 节约开发人员分析生产环境云应用异常的工作量;

[0026] (3) 本发明结合采集到的资源占用情况,可根据不同时间段的访问请求负载情况自动添加、减少运行的云应用容器数量,节约资源使用,减少云应用容器死机概率,提高云应用容器可用性。

## 附图说明

[0027] 图1为本发明的流程示意图。

## 具体实施方式:

[0028] 下面结合附图与实施例对本发明作进一步说明。

[0029] 一种针对云应用容器的增加负载自动均衡方法,包括以下步骤:

[0030] (1) 注册宿主机信息,在容器云应用服务平台管理软件注册管理的宿主机节点,通过在宿主机节点运行注册指令,在每个宿主机节点上创建容器云应用服务平台网络代理容器,通过网络代理容器,来保持宿主机节点和容器云应用服务平台管理软件的信息通讯,接受平台的相关指令,并在宿主机上运行多种指令(如创建云应用容器、采集宿主机资源占用情况,宿主机健康检查、负载均衡服务)。

[0031] (2) 创建容器,在容器云应用服务平台管理软件输入云应用容器运行环境的镜像名称及版本信息,设置云应用容器运行启动参数信息,包括端口信息,访问协议(HTTP, TCP),运行实例数量、内外端口映射信息、宿主机CPU、内存资源使用信息,环境变量信息等;

[0032] (3) 容器云应用服务平台管理软件,根据(1)中的配置信息,通过平台内置的调度管理服务模块,在所管理的宿主机节点通过代理容器实例上创建并启动对应数量的云应用容器实例,并根据平台内置网络IP规则自动分配每个云应用容器的内部IP地址和端口信息;

[0033] (4) 容器云应用服务平台管理软件依据云应用创建时输入的协议类型(支持HTTP、TCP两种类型)、端口信息、访问上下文信息及端口映射信息,在宿主机创建的云应用容器实例的内部IP地址和端口信息,生成LB服务元数据信息,并将信息发送给宿主机节点上运行的代理实例容器,宿主机节点的代理实例容器,收到指令信息后,调用内置代码生成工具,根据内置的负载均衡算法规则信息和接收到的LB元数据信息,通过内置的文件生成工具,创建出云应用服务对应需要的负载均衡配置文件HAProxy.cfg;

[0034] (5) 在宿主机上通讯代理实例容器,启动内置基于HAProxy软件的内置负载模块,加载步骤(4)生成的负载均衡配置文件HAProxy.cfg,在宿主机节点运行负载均衡分发服务;

[0035] (6) 用户通过浏览器或客户端访问云应用服务外部服务地址,宿主机节点接收到请求后,由宿主机节点运行负载均衡分发服务,根据HAProxy.cfg配置中的服务分发规则将请求分发到宿主机节点上运行的云应用容器内部IP和端口,云应用容器接收到请求信息,处理对应的请求,并将处理结果通过负载均衡服务模块返回给用户的浏览器或客户端。

[0036] 本发明要解决的技术问题是根据云应用容器的配置信息自动生成负载均衡配置信息,并运行自动运行负载均衡服务,完成容器云应用的自动负载均衡,包括以下三方面:

[0037] 从容器管理平台获取容器云应用运行的端口配置信息等;

[0038] 自动生成云应用所需要服务负载配置文件信息;

[0039] 根据生成的配置文件信息,构建LB服务,并运行负载均衡服务。

[0040] 如图1所示,云资源池的JEE应用资源弹性调度方法流程图。

[0041] (1) 用户通过容器管理平台输入容器(web应用)运行信息

[0042] 用户通过容器管理平台,创建容器,输入Web应用镜像名称及版本信息,根据业务需要,对Web应用运行的端口,容器运行的节点数,端口映射信息,使用的通讯协议(HTTP、TCP)等信息进行设置。

[0043] (2) 容器管理平台根据前面的设置信息,在宿主机上运行web应用容器

[0044] 容器管理平台根据用户输入的信息,在所在环境的容器宿主机上生成运行web应用容器对应的创建命令信息,运行命令并启动Web应用容器,平台通过内置的网络模块,自

动分配每个容器的内部IP地址信息等,保证容器之间网络及数据的互联互通。

[0045] (3) 容器管理平台负载均衡配置模板和容器的内部IP地址、端口生成对应需要的实际负载均衡配置文件

[0046] 容器管理平台调用平台内置的HA Proxy配置模板信息,根据实际的配置要求,修改其中的IP地址和端口等信息,生成服务对应需要的具体负载均衡配置文件信息。

[0047] (4) 容器管理平台在宿主机运行LB服务容器,加载生成的LB配置信息

[0048] 容器平台在服务所在的宿主机运行平台内置的HA proxy负载均衡器软件,加载平台生成的具体HA proxy负载均衡配置文件信息,提供负载均衡服务。

[0049] (5) 用户访问Web应用,LB服务容器根据LB配置信息,将HTTP请求分发到对应的Web应用容器(根据IP地址和端口),由对应的容器来处理请求,并返回结果

[0050] 用户通过创建服务时输入的IP地址和的端口信息,访问容器中运行的Web应用,LB服务容器根据LB配置信息,将HTTP请求分发到对应的Web应用容器(根据IP地址和端口),由对应的容器来处理请求,并返回处理结果给浏览器。

[0051] 上述虽然结合附图对本发明的具体实施方式进行了描述,但并非对本发明保护范围的限制,所属领域技术人员应该明白,在本发明的技术方案的基础上,本领域技术人员不需要付出创造性劳动即可做出的各种修改或变形仍在本发明的保护范围以内。

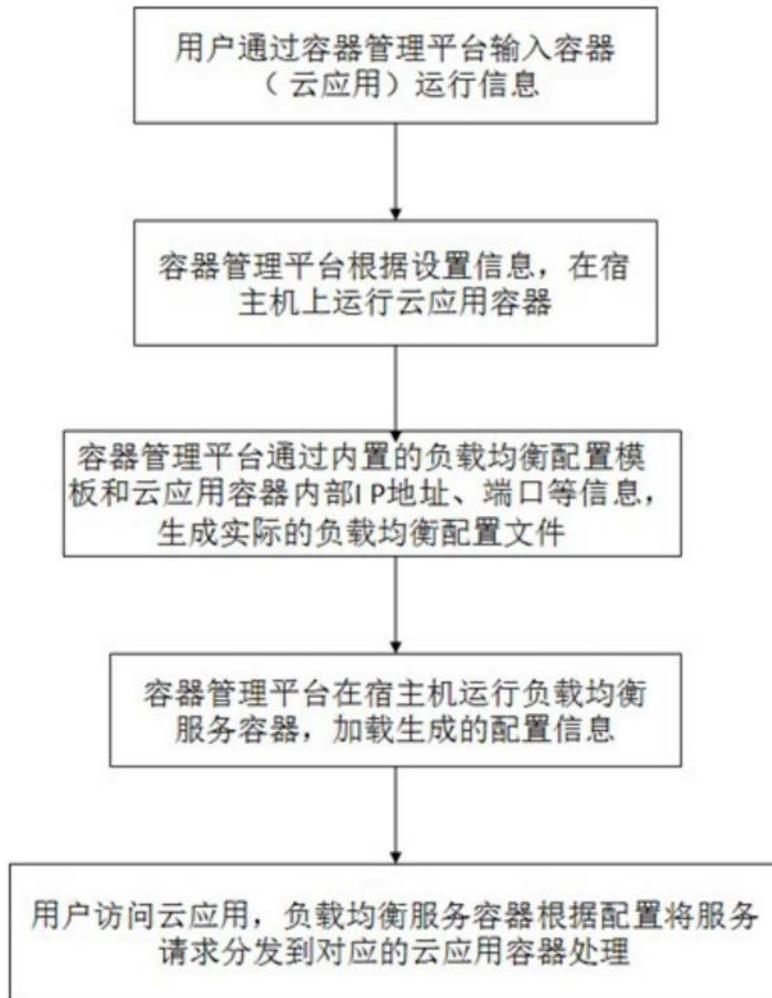


图1