

(12) 按照专利合作条约所公布的国际申请

(19) 世界知识产权组织
国际局



(10) 国际公布号

WO 2020/232891 A1

(43) 国际公布日
2020 年 11 月 26 日 (26.11.2020)

(51) 国际专利分类号:
H04L 29/08 (2006.01) H04L 29/12 (2006.01)

(21) 国际申请号: PCT/CN2019/103262

(22) 国际申请日: 2019 年 8 月 29 日 (29.08.2019)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:
201910410793.7 2019 年 5 月 17 日 (17.05.2019) CN

(71) 申请人: 平安科技(深圳)有限公司(PING AN TECHNOLOGY (SHENZHEN) CO., LTD.) [CN/CN];
中国广东省深圳市福田区福田街道福安社区益田路 5033 号平安金融中心 23 楼, Guangdong 518000 (CN)。

(72) 发明人: 黄桂钦(HUANG, Guiqin); 中国广东省深圳市福田区福田街道福安社区益田路 5033 号平安金融中心 23 楼, Guangdong 518000 (CN)。

(74) 代理人: 深圳市精英专利事务所(SHENZHEN TALENT PATENT SERVICE); 中国广东省深圳市福田区深南中路 6009 号绿景广场 B 栋 20 层 B, Guangdong 518000 (CN)。

(81) 指定国(除另有指明, 要求每一种可提供的国家保护): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL,

(54) Title: LOAD BALANCING APPLICATION VIRTUAL IP SWITCHING METHOD, APPARATUS, COMPUTER DEVICE AND STORAGE MEDIUM

(54) 发明名称: 负载均衡应用虚拟 IP 切换方法、装置、计算机设备及存储介质

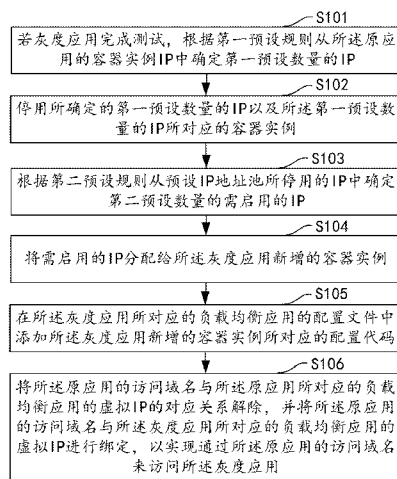


图 1

(57) Abstract: Embodiments of the present application relate to the technical field of cloud computing, and provided thereby are a load balancing application virtual IP switching method, an apparatus, a computer device and a storage medium, which can be applied to a PaaS platform. The method comprises: in embodiments of the present application, disabling a first preset number of container instances of an original application as well as IPs of the first preset number of container instances, then newly adding container instances of a gray scale application, obtaining a second preset number of IPs from among the disabled IPs, and allocating the second preset number of IPs to the newly added container instances of the gray scale application, adding the newly added container instances of the gray scale application into a configuration file of a load balancing application of the gray scale application, and configuring a correspondence relationship between the access domain name of the original application and the virtual IP of the load balancing application corresponding to the gray scale application.

SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG,
US, UZ, VC, VN, ZA, ZM, ZW。

- (84) 指定国 (除另有指明, 要求每一种可提供的地区
保护): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ,
NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), 欧亚 (AM,
AZ, BY, KG, KZ, RU, TJ, TM), 欧洲 (AL, AT, BE, BG,
CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU,
IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT,
RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI,
CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG)。

本国际公布:

- 包括国际检索报告 (条约第21条(3))。

(57) 摘要: 本申请实施例提供一种负载均衡应用虚拟IP切换方法、装置、计算机设备及存储介质, 其涉及云计算技术领域, 可应用于PaaS平台中。所述方法包括: 本申请实施例通过停用原应用的第一预设数量的容器实例, 以及该第一预设数量的容器实例IP, 再新增灰度应用的容器实例, 从所停用的IP中获取第二预设数量的IP, 并将第二预设数量的IP分配给灰度应用新增的容器实例, 在灰度应用的负载均衡应用的配置文件中添加灰度应用新增的容器实例, 并设置原应用的访问域名与灰度应用所对应的负载均衡应用的虚拟IP的对应关系。

负载均衡应用虚拟 IP 切换方法、装置、计算机设备及存储介质

本申请要求于 2019 年 5 月 17 日提交中国专利局、申请号为 201910410793.7、
5 申请名称为“负载均衡应用虚拟 IP 切换方法、装置、计算机设备及存储介质”
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

10 本申请涉及云计算技术领域，尤其涉及一种负载均衡应用虚拟 IP 切换方
法、装置、计算机设备及存储介质。

背景技术

若一个应用在原版本的基础上添加了新功能或者改进了原版本的某些功
能，那么会涉及到灰度发布。将灰度发布所对应的版本的应用称为灰度应用。
15 灰度应用完成测试后，会形成一个稳定的新版本。目前上线新版本应用（灰度
应用）时，需要关闭原应用，以从原应用切换到新版本应用（灰度应用）上，
如此在版本切换时会导致服务的中断，影响了用户的体验。

发明内容

20 本申请实施例提供一种负载均衡应用虚拟 IP 切换方法、装置、计算机设备
及存储介质，可在原应用和灰度应用的版本切换时不会导致服务的中断，提高
了版本切换的效率，提升了用户的体验。

第一方面，本申请实施例提供了一种负载均衡应用虚拟 IP 切换方法，该方法包括：

25 若灰度应用完成测试，根据第一预设规则从原应用的容器实例 IP 中确定第
一预设数量的 IP，其中，所述第一预设数量小于所述原应用的容器实例数量，
所述灰度应用是在所述原应用的版本上添加新功能或者改进功能之后所形成的
应用；停用所确定的第一预设数量的 IP 以及所述第一预设数量的 IP 所对应的
容器实例，其中，所停用的所述第一预设数量的 IP 被回收至预设 IP 地址池；
30 根据第二预设规则从预设 IP 地址池所停用的 IP 中确定第二预设数量的需启用的
IP，其中，所述第二预设数量小于或者等于所停用的 IP 数量；将需启用的
IP 分配给所述灰度应用新增的容器实例；在所述灰度应用所对应的负载均衡应
用的配置文件中添加所述灰度应用新增的容器实例所对应的配置代码；将所述
原应用的访问域名与所述原应用所对应的负载均衡应用的虚拟 IP 的对应关系
35 解除，并将所述原应用的访问域名与所述灰度应用所对应的负载均衡应用的虚
拟 IP 进行绑定，以实现通过所述原应用的访问域名来访问所述灰度应用。

第二方面，本申请实施例提供了一种负载均衡应用虚拟 IP 切换装置，该负
载均衡应用虚拟 IP 切换装置包括：停用 IP 确定单元，用于若灰度应用完成测

试, 根据第一预设规则从原应用的容器实例 IP 中确定第一预设数量的 IP, 其中, 所述第一预设数量小于所述原应用的容器实例数量, 所述灰度应用是在所述原应用的版本上添加新功能或者改进功能之后所形成的应用; 停用单元, 用于停用所确定的第一预设数量的 IP 以及所述第一预设数量的 IP 所对应的容器实例, 其中, 所停用的所述第一预设数量的 IP 被回收到预设 IP 地址池; 启用 IP 确定单元, 用于根据第二预设规则从预设 IP 地址池所停用的 IP 中确定第二预设数量的需启用的 IP, 其中, 所述第二预设数量小于或者等于所停用的 IP 数量; 启用单元, 用于将需启用的 IP 分配给所述灰度应用新增的容器实例; 添加单元, 用于在所述灰度应用所对应的负载均衡应用的配置文件中添加所述灰度应用新增的容器实例所对应的配置代码; 绑定单元, 用于将所述原应用的访问域名与所述原应用所对应的负载均衡应用的虚拟 IP 的对应关系解除, 并将所述原应用的访问域名与所述灰度应用所对应的负载均衡应用的虚拟 IP 进行绑定, 以实现通过所述原应用的访问域名来访问所述灰度应用。

第三方面, 本申请实施例提供了一种计算机设备, 所述计算机设备包括存储器, 以及与所述存储器相连的处理器;

所述存储器用于存储计算机程序, 所述处理器用于运行所述存储器中存储的计算机程序, 以执行上述第一方面所述的方法。

第四方面, 本申请实施例提供了一种计算机可读存储介质, 所述计算机可读存储介质存储有计算机程序, 所述计算机程序被处理器执行时, 实现上述第一方面所述的方法。

附图说明

为了更清楚地说明本申请实施例技术方案, 下面将对实施例描述中所需要使用的附图作简单地介绍, 显而易见地, 下面描述中的附图是本申请的一些实施例, 对于本领域普通技术人员来讲, 在不付出创造性劳动的前提下, 还可以根据这些附图获得其他的附图。

图 1 是本申请实施例提供的负载均衡应用虚拟 IP 切换方法的流程示意图;

图 2 是本申请实施例提供的负载均衡应用虚拟 IP 切换方法的子流程示意图;

图 3 是本申请实施例提供的负载均衡应用虚拟 IP 切换方法的子流程示意图;

图 4 是本申请实施例提供的负载均衡应用虚拟 IP 切换方法的子流程示意图;

图 5 是本申请实施例提供的负载均衡应用虚拟 IP 切换装置的示意性框图;

图 6 是本申请实施例提供的停用单元的示意性框图;

图 7 是本申请实施例提供的启用单元的示意性框图;

图 8 是本申请实施例提供的配置加入单元的示意性框图;

图 9 是本申请实施例提供的计算机设备的示意性框图。

具体实施方式

下面将结合本申请实施例中的附图, 对本申请实施例中的技术方案进行清楚、完整地描述, 显然, 所描述的实施例是本申请一部分实施例, 而不是全部

的实施例。基于本申请中的实施例，本领域普通技术人员在没有做出创造性劳动前提下所获得的所有其他实施例，都属于本申请保护的范围。

图1是本申请实施例提供的负载均衡应用虚拟IP切换方法的流程示意图。如图1所示，该方法包括S101-S106。

5 S101，若灰度应用完成测试，根据第一预设规则从原应用的容器实例IP中确定第一预设数量的IP，其中，所述第一预设数量小于所述原应用的容器实例数量，所述灰度应用是在所述原应用的版本上添加新功能或者改进功能之后所形成的应用。

10 其中，灰度应用是在原应用的版本上添加新功能或者改进功能之后所形成的应用。灰度应用完成测试，理解为灰度应用可以成为一个全量升级的新版本。在本实施例中，灰度应用也可以简单理解为应用的一个新版本。

根据第一预设规则从原应用的容器实例IP中确定第一预设数量的IP，包括：获取所述原应用的容器实例IP和容器实例数量，根据第一预设规则从原应用的容器实例IP中确定第一预设数量的IP。

15 其中，获取原应用的容器实例IP和容器实例数量可通过多种方式。一般情况下，在该PaaS平台的服务器中保存有所有应用所创建的容器实例IP和容器实例数量，如此可从保存应用所创建的容器实例IP和容器实例数量的数据库或者文件中直接获取原应用的容器实例IP和容器实例数量。

20 其中，根据第一预设规则从原应用的容器实例IP中确定第一预设数量的IP，如可以从原应用的容器实例IP中随机确定第一预设数量的容器实例IP，也可以从原应用的容器实例IP中按照容器实例IP分配的先后顺序确定第一预设数量的容器实例IP，还可以从原应用的容器实例IP中优先确定不属于同一个网段中的容器实例IP（如某一个网段中的容器实例IP较多，那么优先从其他网段中获取第一预设数量的IP）、再按照随机或者容器实例IP分配的选后
25 顺序确定第一预设数量的IP，还按照其他的规则选择第一预设数量的IP。其中，第一预设数量小于原应用的容器实例数量，如原应用的容器实例数量为8个，第一预设数量可以为4个等。其中，第一预设数量小于原应用的容器实例数量的目的是，在对原应用的第一预设数量的容器实例进行处理时，原应用仍然有容器实例（除了第一预设数量之外的其他容器实例）是正常运行的，正常运行的
30 容器实例可以提供服务，如此，不会导致服务的中断。

在一实施例中，可开发一个对应的管理面板，以在管理面板上显示应用与应用的容器实例IP的对应关系。如此，通过调用管理面板，获取在管理面板上显示的原应用以及原应用的容器实例IP。从显示的原应用的容器实例IP中按照第一预设规则选择第一预设数量的容器实例IP，并接收按照第一预设规则选择的
35 第一预设数量的容器实例IP作为所确定的第一预设数量的IP。

S102，停用所确定的第一预设数量的IP以及所述第一预设数量的IP所对应的容器实例，其中，所停用的所述第一预设数量的IP被回收至预设IP地址池。

其中，将所述原应用的容器实例IP和灰度应用的容器实例IP作为预设IP

地址池。在一实施例中，在步骤 S102 之前，预先将原应用的容器实例 IP 和灰度应用的容器实例 IP 作为预设 IP 地址池，并设置原应用的容器实例和灰度应用的容器实例共享该预设 IP 地址池。如此，一旦一个 IP 被分配给原应用的容器实例和原应用的灰度应用的容器实例，那么该 IP 加入到共享的预设 IP 地址池中，同时该 IP 不能被其他应用占用，只可供原应用的容器实例和灰度应用的容器实例使用。设置原应用的容器实例 IP 和灰度应用的容器实例 IP 共享预设 IP 地址池是因为，一般情况下，一个应用的容器实例 IP 被停用之后，会被回收收到空闲 IP 地址池，而空闲 IP 地址池的 IP 可以分配给任意一个容器应用使用。如此，若停用原应用的容器实例 IP 或者灰度应用的容器实例 IP 之后，停用的 IP 被回收收到空闲 IP 地址池，不再属于原应用或者灰度应用了。而将原应用的容器实例 IP 和灰度应用的容器实例 IP 作为预设 IP 地址池，并设置原应用的容器实例 IP 和灰度应用的容器实例 IP 共享该预设 IP 地址池，那么预设 IP 地址池中的 IP 只属于原应用的容器实例和灰度应用的容器实例，不会被其他应用占用。

在确定了第一预设数量的容器实例 IP 后，停用所确定的容器实例 IP 以及容器实例 IP 所对应的容器实例。

在一实施例中，如图 2 所示，步骤 S102 包括以下步骤 S1021-S1024。

S1021，设置所确定的第一预设数量的 IP 的使用状态标识为停用状态。

需要注意的是，将第一预设数量的 IP 设置的使用状态设置为停用状态，理解为该 IP 仍然属于该应用，不可以被其他应用使用。

S1022，销毁所述第一预设数量的 IP 所对应的容器实例。

如通过编排工具下发销毁指令到容器实例所对应的宿主机，销毁指令如 `docker stop` 指令，在所对应的宿主机上将对应的容器实例进行销毁。其中，销毁容器实例也可理解为缩容容器实例，销毁容器实例将会释放该容器实例所占用的所有资源，如包括容器实例所对应的环境信息、关闭文件输入输出流、关闭与数据库的连接、修改其他数据库中保存的该容器实例的相关信息等。其中，编排工具如 `marathon` 等。

需要注意的是，首先停用容器实例的 IP 再销毁容器实例，而不是先销毁容器实例再停用容器实例的 IP，是为了以防容器实例的 IP 有效，而容器实例无效，原应用的负载均衡应用将接收到的用户访问请求分配到该容器实例，而实际上该容器实例已经无效了，那么会导致用户访问请求的请求失败。

S1023，回收使用状态标识为停用状态的 IP 到预设 IP 地址池。

需要注意的是，该预设 IP 地址池中的 IP 是原应用与灰度应用可共享的 IP。在该预设 IP 地址池中，使用状态标识为停用的 IP 不能被其他应用占用。

S1024，根据所述第一预设数量修改所述原应用的容器实例数量。

其中，修改后的容器实例数量为原应用的原容器实例数量减去第一预设数量。

需要注意的是，步骤 S1023 与步骤 S1024 的执行顺序并不严格限制，也可以先执行步骤 S1024 再执行步骤 S1023。

S103, 根据第二预设规则从预设 IP 地址池所停用的 IP 中确定第二预设数量的需启用的 IP, 其中, 所述第二预设数量小于或者等于所停用的 IP 数量。

如可以从预设 IP 地址池中的使用状态标识为停用的 IP 中随机确定第二预设数量的容器实例 IP, 也可以从使用状态标识为停用的 IP 中按照 IP 使用状态标识改变为停用状态的先后顺序确定第二预设数量的容器实例 IP 等, 还可以根据其他的规则确定第二预设数量的容器实例 IP。其中, 若第二预设数量小于所停用的 IP 数量, 意味着将所停用的一部分 IP 用来分配给灰度应用新增的容器实例, 意味着灰度应用需要新增容器实例的数量小于原应用销毁的容器实例的数量。若第二预设数量等于所停用的 IP 数量, 意味着将所停用的所有 IP 用来分配给灰度应用新增的容器实例。那么意味着原应用有多少个容器实例销毁, 那么灰度应用就需要新增多少个容器实例。

在一实施例中, 预设 IP 地址池中所停用的 IP 可能大于本次所停用的第一预设数量的 IP。那么可以理解为, 在本次停用原应用的容器实例 IP 之前, 还停用过原应用的容器实例 IP 或者还停用过灰度应用的容器实例 IP, 如此, 导致预设 IP 地址池中所停用的 IP 大于本次所停用的第一预设数量的 IP。

在一实施例中, 若已开发对应的管理面板, 在管理面板上可显示应用与应用的容器实例 IP 的对应关系、以及 IP 的使用状态标识。如此, 获取并在管理面板上显示原应用和灰度应用所对应的容器 IP 以及容器 IP 的使用状态标识; 从使用状态标识为停用的 IP 中选择第二预设数量的需启用的 IP, 接收所选择的需启用的 IP 作为所确定的需启用的 IP。并从应用信息中选择灰度应用, 如根据灰度应用所对应的应用名选择灰度应用, 接收所选择的灰度应用, 以进一步将需启用的 IP 分配给灰度应用。

S104, 将需启用的 IP 分配给所述灰度应用新增的容器实例。

确定了需启用的 IP 后, 将需启用的 IP 分配给新增的灰度应用的容器实例。

在一实施例中, 如图 3 所示, 所述 S104 包括以下步骤 S1041-S1043。

S1041, 根据所述灰度应用以及需启用的 IP 数量确定所述灰度应用需新增的容器实例所在的宿主机。

具体地, 根据灰度应用确定灰度应用需新增的容器实例所在的区域, 根据需启用的 IP 数量从所述区域中确定相同数量的需新增的灰度应用的容器实例的宿主机 (宿主机 IP)。其中, 宿主机的数量与新增的容器实例的数量相同, 新增的容器实例的数量与需启用的 IP 数量相同。

可以理解地, 先确定灰度应用需启用的 IP 的数量, 再根据需启用的 IP 的数量来确定灰度应用新增的容器实例数量, 是因为启用 IP 是从预设 IP 地址池中的所停用的 IP 中获得的, 所停用的 IP 的数量是有限的, 不是无限的。因此, 需要先根据第二预设数量的启用 IP 来确定新增的容器实例的数量。

S1042, 在所确定的宿主机上创建所述灰度应用的容器实例。

具体地, 发送新建命令到编排工具, 以通过编排工具推送新建命令到对应的宿主机, 并通过新建命令在对应的宿主机上新建灰度应用所对应的容器实例。其中, 新建命令中包括有灰度应用的应用镜像等信息。

需要注意的是，不同区域的编排工具是不同的，根据灰度所在的区域确定所需要使用的编排工具。如此，编排工具与灰度应用所在的区域相同，也与新增灰度应用的容器实例的宿主机所在的区域相同，提高了新增容器实例的速度。将新建命令发送到所需要使用的编排工具上，以通过编排工具推送新建命令到对应的宿主机，并在对应的宿主主机上创建灰度应用所对应的容器实例，将所创建的灰度应用的容器实例作为灰度应用新增的容器实例。

S1043，将需启用的 IP 分配给所述灰度应用新创建的容器实例。

具体地，通过运维工具如 `ansible` 工具去连接容器实例，将所创建的容器实例与需启用的 IP 进行绑定。

S105，在所述灰度应用所对应的负载均衡应用的配置文件中添加所述灰度应用新增的容器实例所对应的配置代码。

其中，灰度应用的负载均衡应用的类型包括 F5、LVS、Haproxy、Nginx 等。灰度应用的负载均衡应用的类型可以在 PaaS 平台数据库中获取。

在一实施例中，步骤 S105 包括：检测所述灰度应用所对应的负载均衡应用的类型；根据所述灰度应用所对应的负载均衡应用的类型，通过与所述类型相匹配的方式获取所述负载均衡应用的配置文件并在所述配置文件中加入所述灰度应用新增的容器实例所对应的配置代码。

其中，F5 负载均衡应用是硬负载均衡应用，是在具体的 F5 设备上创建的，而 LVS、Haproxy、Nginx 负载均衡应用是软负载均衡应用，因此 F5 负载均衡应用和 LVS、Haproxy、Nginx 负载均衡应用获取配置文件的方式、对配置文件修改的方式、使配置文件生效的方式以及添加的配置代码等都存在不同。

在一实施例中，如图 4 所示，根据所述灰度应用所对应的负载均衡应用的类型，通过与所述类型相匹配的方式获取所述负载均衡应用的配置文件并在所述配置文件中加入所述灰度应用新增的容器实例所对应的配置代码的步骤，包括以下步骤 S1051-S1058。

S1051，判断所述负载均衡应用的类型是否为 F5。

S1052，若所述负载均衡应用的类型为 F5，根据所述灰度应用确定 F5 负载均衡应用所在的 F5 设备。

S1053，在所述 F5 设备上获取所述 F5 负载均衡应用的配置文件，并在所述配置文件中加入所述灰度应用新增的容器实例所对应的配置代码。

在所述配置文件中加入所述灰度应用新增的容器实例所对应的配置代码，包括：在配置文件中添加 `pool member`（新增的灰度应用的容器实例）的信息；修改 `pool` 信息中的 `member` 数量。其中，添加 `pool member`（新增的灰度应用的容器实例）的信息，包括新增的灰度应用的容器实例的 IP 地址、端口信息；将该容器实例启动，并修改容器实例的监听状态为运行（`running`）状态。其中，端口默认与之前存在的容器实例的端口相同，如 80 端口。其中，修改后的 `member` 数量为修改前的 `member` 数量与新增的灰度应用的容器实例的数量。

S1054，若所述负载均衡应用的类型为非 F5，根据所述灰度应用确定所对应的所述负载均衡应用的容器，其中非 F5 负载均衡应用包括 LVS 负载均衡应用、

Haproxy 负载均衡应用、Nginx 负载均衡应用。

根据所述灰度应用确定所述负载均衡应用的容器，可以理解为，根据灰度应用确定灰度应用所对应的负载均衡应用的负载均衡应用的容器的虚拟 IP。

5 S1055, 从对应的负载均衡应用的容器中获取配置文件，将所获取的配置文件下载到本地。

根据虚拟 IP 从对应容器中获取对应负载均衡应用的配置文件；将所获取的配置文件下载到本地的一个预设目录。需要注意的是，将所获取的配置文件下载到本地，以对配置文件进行修改，而不是直接远程在对应负载均衡应用容器上进行修改，是考虑到若在对配置文件进行修改的过程中，网络出现中断，那么改写的过程也会出现中断，即排除远程改写的过程中出现的意外情况对改写造成的影响，另一方面，在本地改写配置文件比在远程改写配置文件的速度快。即在本
10 地写文件的速度要比远程写文件的速度快，如至少不用通过网络进行传输。

15 S1056, 在所述配置文件中加入所述灰度应用新增的容器实例所对应的配置代码。

具体地，步骤 S1056 包括：获取对应的容器实例配置代码模板；在对应的容器实例配置代码模板中添加新增的灰度应用的容器实例的 IP 和端口信息（其中，端口默认与之前存在的容器实例的端口相同，如 80 端口）以形成新增的容器实例的配置代码；将新增的容器实例的配置代码追加到配置文件中最后一个容器实例的配置代码之后。
20

其中，LVS 负载均衡应用、Haproxy 负载均衡应用、Nginx 应用所对应的容器实例配置代码模板不同。如 LVS 负载均衡应用中容器实例配置代码模板中包括容器实例 IP 信息、容器实例的端口信息（用于提供服务的端口信息）、该容器实例（即负载均衡应用的服务器）所对应的权重信息、该容器实例所对应的健康检查方式信息（包括连接超时时间、重连次数、重连间隔时间等）信息等，将配置代码模板中的容器实例 IP 参数、容器实例的端口参数以新增的灰度应用的容器实例的 IP 和端口进行替换，以形成新增的容器实例的配置代码，将新增的容器实例的配置代码追加到配置文件（该处的配置文件包括 keepalived.conf）中最后一个容器实例的配置代码之后。如 Haproxy 负载均衡应用中，配置代码模板可以是以 server 节点开头的一段代码，如 server 容器实例 IP: 容器实例的端口 check inter 30000 rise 1 fall 2，将配置代码模板中的容器实例 IP 参数、容器实例的端口参数以新增的灰度应用的容器实例的 IP 和端口进行替换，以形成新增的容器实例的配置代码，将新增的容器实例的配置代码追加到配置文件（该处的配置文件包括 Haproxy.conf）中最后一个 server 代码段的后面，其中，可通过定位 Haproxy 负载均衡应用的 option forwardfor 节点中的 server 节点来定位最后一个 server 代码段。如在 Nginx 负载均衡应用中，配置代码模板也是以 server 节点开头的一段代码，在配置代码模板中加入了新增的容器实例的 IP 和容器实例的端口后，加入到配置文件中最后一个容器实例的配置代码。其中，可以通过定位 upstream 节点中的 server
30
35

内容来定位配置文件中最后一个容器实例的配置代码。

S1057, 将加入了所述配置代码的配置文件推送至对应的负载均衡应用的容器中。

5 S1058, 重启对应的负载均衡应用的容器的负载均衡服务, 以使得加入了所述配置代码的配置文件生效。

如可以用 `restart` 命令重启所对应的负载均衡应用的负载均衡服务。

该实施例进一步限定根据不同负载均衡应用的类型来确定不同的方式来实现灰度应用的配置文件中加入灰度应用新增的容器实例所对应的配置代码。该实施例的过程自动完成, 无需人工干预, 提高了速度。

10 S106, 将所述原应用的访问域名与所述原应用所对应的负载均衡应用的虚拟 IP 的对应关系解除, 并将所述原应用的访问域名与所述灰度应用所对应的负载均衡应用的虚拟 IP 进行绑定, 以实现通过所述原应用的访问域名来访问所述灰度应用。

具体地, 在该 PaaS 平台的服务器中保存有原应用的访问域名与负载均衡应用的虚拟 IP 的对应关系。所述原应用的访问域名与所述原应用所对应的负载均衡应用的虚拟 IP 的对应关系解除, 可以理解为删除所保存的原应用的访问域名与原应用所对应的负载均衡应用的虚拟 IP 的对应关系。将所述原应用的访问域名与所述灰度应用所对应的负载均衡应用的虚拟 IP 进行绑定, 可以理解为在数据库中
15 添加原应用的访问域名与所述灰度应用所对应的负载均衡应用的虚拟 IP 的对应关系。
20

在原应用的访问域名与原应用所对应的负载均衡应用的虚拟 IP 的对应关系解除之前, 用户仍可以通过原应用的访问域名来访问原应用, 即访问原应用的部分容器实例 (除了第一预设数量的容器实例), 如此, 原应用所提供的服务不会中断。

25 在原应用的访问域名与所述灰度应用所对应的负载均衡应用的虚拟 IP 进行绑定之后, 用户即可以通过原应用的访问域名来访问灰度应用所对应的容器实例, 如此, 灰度应用的容器实例即开始提供服务。

以下举例对本实施例的方案进行简单的说明。如原应用的容器实例数量为 8 个, 灰度应用的容器实例数量为 1 个 (假设默认创建的灰度应用的容器实例数量为 1)。将原应用的容器实例 IP 和灰度应用的容器实例 IP 作为预设 IP 地址池, 原应用的容器实例和灰度应用的容器实例共享该预设 IP 地址池, 预设 IP 地址池中的 IP 数量为 9 个。若第一预设数量为 5 个, 那么从原应用的容器实例 IP 中确定 5 个 IP 地址, 销毁该 5 个 IP 地址所对应的容器实例, 并将该 5 个 IP 地址回收
30 到预设 IP 地址池, 设置该 5 个 IP 地址的使用状态为停用状态。原应用还剩 3 个容器实例, 在没有修改用户访问域名与原应用所对应的负载均衡应用的虚拟 IP 的对应关系之前, 用户仍可以通过剩下的 3 个容器实例访问原应用。确定第二预设数量 (小于或者等于第一预设数量), 第二预设数量可以为 1、或者 2 或者 3 或者 4 或者 5。若第二预设数量为 5 个, 那么可理解为所有回收的 IP 将被重新利用。新增 5 个灰度应用的容器实例, 将 5 个停用状态的
35

IP 分配给灰度应用新增的 5 个容器实例。在灰度应用所对应的负载均衡应用的配置文件中添加新增的容器实例信息，如此灰度应用所对应的负载均衡应用的配置文件中就会存在 6 个灰度应用的容器实例（默认创建的 1 个容器实例和新增的 5 个容器实例）。将用户访问域名与原应用的负载均衡应用的 VIP 的对应关系解除，在解除之前，用户通过访问域名仍可以访问原应用的另外 3 个容器实例，如此在切换的过程中，服务不会间断，因为还有另外 3 个容器实例存在，这 3 个容器实例可以继续提供服务。在解除之后，将用户访问域名与灰度应用所对应的负载均衡应用的 VIP 绑定。如此，用户访问时不发生任何改变，通过原应用的访问域名即可访问灰度应用的 6 个容器实例。

上述方法实施例通过停用原应用的第一预设数量的容器实例，以及该第一预设数量的容器实例 IP，再新增灰度应用的容器实例，从所停用的 IP 中获取第二预设数量的 IP，并将第二预设数量的 IP 分配给灰度应用新增的容器实例，在灰度应用的负载均衡应用的配置文件中添加灰度应用新增的容器实例，并设置原应用的访问域名与灰度应用所对应的负载均衡应用的虚拟 IP 的对应关系。从而实现了将原应用的虚拟 IP 切换到灰度应用，通过原应用的访问域名来访问灰度应用。上述方法实施例快速地实现了原应用的下线，灰度应用的上线，在原应用和灰度应用的版本切换时不会导致服务的中断，实现了无痕切换，提高了版本切换的效率，提升了用户的体验。同时上述实施例无需修改访问域名，通过同一个访问域名（原应用的访问域名）就可以访问灰度应用。

在一实施例中，在步骤 S101 之前，所述负载均衡应用虚拟 IP 切换方法包括以下步骤：克隆原应用，并在克隆的原应用版本的基础上，添加新功能或者改进功能所对应的代码，以形成灰度应用镜像；根据所述原应用的容器实例的创建区域以及所述灰度应用镜像，在测试环境下所述区域中创建灰度应用的容器实例以及灰度应用所对应的负载均衡应用。

其中，克隆原应用，包括克隆原应用的代码，原应用所对应的配置文件、系统文件、分配的内存和 CPU 个数等等。克隆原版本的应用，以得到一个与原应用完全相同的应用。在该克隆的原应用版本的基础上，添加新功能或者改进功能所对应的代码，以形成灰度应用镜像。

另外，需要注意的是，原应用是可以对外提供服务的，因此原应用已经创建了对应的容器实例以及原应用所对应的负载均衡应用。如此，可获取原应用容器实例的创建区域，其中，创建区域指的是所创建的容器实例的机房位置等。在灰度应用通过简单测试（如开发人员的简单测试）后，将灰度应用推送到测试环境下与原应用容器实例创建区域相同区域下的灰度应用所对应的镜像仓库中。需要注意的是，每个不同的机房中会对应有不同应用环境的镜像仓库，如开发环境（该应用的容器实例创建成功后供内部开发使用）、生产环境（意味着该应用的容器实例创建成功后对接外部环境，或者供外部用户进行访问）、测试环境（意味着该应用的容器实例创建成功后供测试使用）。在创建灰度应用的容器实例时，可根据设置灰度应用容器实例数量（如果不进行设置，那么会使用默认的值，默认值为 1，即默认的灰度应用容器实例数量为 1）并从对

应镜像仓库中选择的灰度应用镜像，在所述区域创建对应数量的容器实例，并为所创建的容器实例分配相同数量的容器实例 IP。

另外，若在创建灰度应用的容器实例的过程中，选择了创建负载均衡应用，那么会在创建灰度应用的容器实例的同时创建灰度应用所对应的负载均衡应用，并设置所创建的灰度应用和所创建的负载均衡应用的关联关系，同时会自动生成所创建的负载均衡应用的配置文件以及负载均衡应用所对应的域名。若在创建灰度应用的容器实例的过程中，没有选择创建负载均衡应用，那么在灰度应用的容器实例创建完成后，会为该灰度应用创建负载均衡应用。创建该灰度应用的负载均衡应用时会自动生成所创建的负载均衡应用的配置文件以及负载均衡应用所对应的域名。

该实施例进一步限定如何创建灰度应用的容器实例。图 5 是本申请实施例提供的负载均衡应用虚拟 IP 切换装置的示意性框图。该装置包括用于执行上述负载均衡应用虚拟 IP 切换方法所对应的单元。具体地，如图 5 所示，该装置 100 包括停用 IP 确定单元 101、停用单元 102、启用 IP 确定单元 103、启用单元 104、添加单元 105 以及绑定单元 106。

停用 IP 确定单元 101，用于若灰度应用完成测试，根据第一预设规则从原应用的容器实例 IP 中确定第一预设数量的 IP，其中，所述第一预设数量小于所述原应用的容器实例数量，所述灰度应用是在所述原应用的版本上添加新功能或者改进功能之后所形成的应用。

停用单元 102，用于停用所确定的第一预设数量的 IP 以及所述第一预设数量的 IP 所对应的容器实例，其中，所停用的所述第一预设数量的 IP 被回收到预设 IP 地址池。

在一实施例中，如图 6 所示，停用单元 102 包括：IP 状态设置单元 1021、销毁单元 1022、回收单元 1023、数量修改单元 1024。其中，IP 状态设置单元 1021，用于设置所确定的第一预设数量的 IP 的使用状态标识为停用状态。销毁单元 1022，用于销毁所述第一预设数量的 IP 所对应的容器实例。回收单元 1023，用于回收使用状态标识为停用状态的 IP 到预设 IP 地址池。数量修改单元 1024，用于根据所述第一预设数量修改所述原应用的容器实例数量。

启用 IP 确定单元 103，用于根据第二预设规则从预设 IP 地址池所停用的 IP 中确定第二预设数量的需启用的 IP，其中，所述第二预设数量小于或者等于所停用的 IP 数量。

启用单元 104，用于将需启用的 IP 分配给所述灰度应用新增的容器实例。

在一实施例中，如图 7 所示，启用单元 104 包括宿主机确定单元 1041、创建单元 1042、分配启用单元 1043。其中，宿主机确定单元 1041，用于根据所述灰度应用以及需启用的 IP 数量确定所述灰度应用需新增的容器实例所在的宿主机。创建单元 1042，用于在所确定的宿主机上创建所述灰度应用的容器实例。分配启用单元 1043，用于将需启用的 IP 分配给所述灰度应用新创建的容器实例。

添加单元 105，用于在所述灰度应用所对应的负载均衡应用的配置文件中

添加所述灰度应用新增的容器实例所对应的配置代码。

在一实施例中，所述添加单元 105 包括类型检测单元、配置加入单元。其中，类型检测单元，用于检测所述灰度应用所对应的负载均衡应用的类型。配置加入单元，用于根据所述灰度应用所对应的负载均衡应用的类型，通过与所述类型相匹配的方式获取所述负载均衡应用的配置文件并在所述配置文件中加入所述灰度应用新增的容器实例所对应的配置代码。

在一实施例中，如图 8 所示，配置加入单元 1050 包括类型判断单元 1051、设备确定单元 1052、配置代码加入单元 1053、容器确定单元 1054、配置下载单元 1055、推送单元 1056 以及重启单元 1057。其中，类型判断单元 1051，用于判断所述负载均衡应用的类型是否为 F5。设备确定单元 1052，用于若所述负载均衡应用的类型为 F5，根据所述灰度应用确定 F5 负载均衡应用所在的 F5 设备。配置代码加入单元 1053，用于在所述 F5 设备上获取所述 F5 负载均衡应用的配置文件，并在所述配置文件中加入所述灰度应用新增的容器实例所对应的配置代码。容器确定单元 1054，用于若所述负载均衡应用的类型为非 F5，根据所述灰度应用确定所述负载均衡应用的容器，其中非 F5 负载均衡应用包括 LVS 负载均衡应用、Haproxy 负载均衡应用、Nginx 负载均衡应用。配置下载单元 1055，用于从对应的负载均衡应用的容器中获取配置文件，将所获取的配置文件下载到本地。配置代码加入单元 1053，还用于在所述配置文件中加入所述灰度应用新增的容器实例所对应的配置代码。推送单元 1056，用于将加入了所述配置代码的配置文件推送至对应的负载均衡应用的容器中。重启单元 1057，用于重启对应的负载均衡应用的容器的负载均衡服务，以使得加入了所述配置代码的配置文件生效。绑定单元 106，用于将所述原应用的访问域名与所述原应用所对应的负载均衡应用的虚拟 IP 的对应关系解除，并将所述原应用的访问域名与所述灰度应用所对应的负载均衡应用的虚拟 IP 进行绑定，以实现通过所述原应用的访问域名来访问所述灰度应用。

在一实施例中，所述负载均衡应用虚拟 IP 切换装置还包括：灰度镜像形成单元、灰度应用创建单元。其中，灰度镜像形成单元，用于克隆原应用，并在克隆的原应用版本的基础上，添加新功能或者改进功能所对应的代码，以形成灰度应用镜像。灰度应用创建单元，用于根据所述原应用的容器实例的创建区域以及所述灰度应用镜像，在测试环境下所述区域中创建灰度应用的容器实例以及灰度应用所对应的负载均衡应用。

需要说明的是，所属领域的技术人员可以清楚地了解到，上述装置和各单元的具体实现过程，可以参考前述方法实施例中的相应描述，为了描述的方便和简洁，在此不再赘述。

上述装置可以实现为一种计算机程序的形式，计算机程序可以在如图 9 所示的计算机设备上运行。图 9 为本申请实施例提供的一种计算机设备的示意性框图。该设备为终端等设备，如 PaaS 平台中的服务器等。该设备 200 包括通过系统总线 201 连接的处理器 202、存储器和网络接口 203，其中，存储器可以包括非易失性存储介质 204 和内存储器 205。

该非易失性存储介质 204 可存储操作系统 2041 和计算机程序 2042。该非易失性存储介质中所存储的计算机程序 2042 被处理器 202 执行时,可实现上述所述的负载均衡应用虚拟IP切换方法。该处理器 202 用于提供计算和控制能力,支撑整个设备 200 的运行。该内存存储器 205 为非易失性存储介质中的计算机程序的运行提供环境,该计算机程序被处理器 202 执行时,可使得处理器 202 执行上述所述的负载均衡应用虚拟IP切换方法。该网络接口 203 用于进行网络通信。本领域技术人员可以理解,图中示出的结构,仅仅是与本申请方案相关的部分结构的框图,并不构成对本申请方案所应用于其上的设备的限定,具体的设备可以包括比图中所示更多或更少的部件,或者组合某些部件,或者具有不同的部件布置。

其中,所述处理器 202 用于运行存储在存储器中的计算机程序,以实现本申请实施例的负载均衡应用虚拟 IP 切换方法。

本领域技术人员可以理解,图 9 中示出的计算机设备的实施例并不构成对计算机设备具体构成的限定,在其他实施例中,计算机设备可以包括比图示更多或更少的部件,或者组合某些部件,或者不同的部件布置。例如,在一些实施例中,计算机设备可以仅包括存储器及处理器,在这样的实施例中,存储器及处理器的结构及功能与图 9 所示实施例一致,在此不再赘述。

应当理解,在本申请实施例中,所称处理器 202 可以是中央处理单元(Central Processing Unit, CPU),该处理器还可以是其他通用处理器、数字信号处理器(Digital Signal Processor, DSP)、专用集成电路(应用程序 Ilication Specific Integrated Circuit, ASIC)、现成可编程门阵列(Field-Programmable Gate Array, FPGA)或者其他可编程逻辑器件、分立门或者晶体管逻辑器件等。通用处理器可以是微处理器或者该处理器也可以是任何常规的处理器等。

在本申请的另一实施例中提供计算机可读存储介质。该计算机可读存储介质可以为非易失性的计算机可读存储介质。该计算机可读存储介质存储有计算机程序,其中计算机程序被处理器执行时实现本申请实施例的负载均衡应用虚拟 IP 切换方法。

所述存储介质为实体的、非瞬时性的存储介质,例如可以是 U 盘、移动硬盘、只读存储器(Read-Only Memory, ROM)、磁碟或者光盘等各种可以存储程序代码的实体存储介质。

所属领域的技术人员可以清楚地了解到,为了描述的方便和简洁,上述描述的设备、装置和单元的具体工作过程,可以参考前述方法实施例中的对应过程,在此不再赘述。

以上所述,仅为本申请的具体实施方式,但本申请的保护范围并不局限于此,任何熟悉本技术领域的技术人员在本申请揭露的技术范围内,可轻易想到各种等效的修改或替换,这些修改或替换都应涵盖在本申请的保护范围之内。因此,本申请的保护范围应以权利要求的保护范围为准。

权 利 要 求 书

1. 一种负载均衡应用虚拟 IP 切换方法, 所述方法包括:

若灰度应用完成测试, 根据第一预设规则从原应用的容器实例 IP 中确定第一预设数量的 IP, 其中, 所述第一预设数量小于所述原应用的容器实例数量, 所述灰度应用是在所述原应用的版本上添加新功能或者改进功能之后所形成的应用;

停用所确定的第一预设数量的 IP 以及所述第一预设数量的 IP 所对应的容器实例, 其中, 所停用的所述第一预设数量的 IP 被回收到预设 IP 地址池;

根据第二预设规则从预设 IP 地址池所停用的 IP 中确定第二预设数量的需启用的 IP, 其中, 所述第二预设数量小于或者等于所停用的 IP 数量;

将需启用的 IP 分配给所述灰度应用新增的容器实例;

在所述灰度应用所对应的负载均衡应用的配置文件中添加所述灰度应用新增的容器实例所对应的配置代码;

将所述原应用的访问域名与所述原应用所对应的负载均衡应用的虚拟 IP 的对应关系解除, 并将所述原应用的访问域名与所述灰度应用所对应的负载均衡应用的虚拟 IP 进行绑定, 以实现通过所述原应用的访问域名来访问所述灰度应用。

2. 根据权利要求 1 所述的方法, 其中, 所述停用所确定的第一预设数量的 IP 以及所述第一预设数量的 IP 所对应的容器实例, 包括:

设置所确定的第一预设数量的 IP 的使用状态标识为停用状态;

销毁所述第一预设数量的 IP 所对应的容器实例;

回收使用状态标识为停用状态的 IP 到预设 IP 地址池;

根据所述第一预设数量修改所述原应用的容器实例数量。

3. 根据权利要求 1 所述的方法, 其中, 所述将需启用的 IP 分配给所述灰度应用新增的容器实例, 包括:

根据所述灰度应用以及需启用的 IP 数量确定所述灰度应用需新增的容器实例所在的宿主机;

在所确定的宿主机上创建所述灰度应用的容器实例;

将需启用的 IP 分配给所述灰度应用新创建的容器实例。

4. 根据权利要求 3 所述的方法, 其中, 所述在所确定的宿主机上创建所述灰度应用的容器实例, 包括:

发送新建命令到编排工具, 以通过所述编排工具推送所述新建命令到所确定的宿主机, 并通过所述新建命令在对应的宿主机上创建所述灰度应用所对应的容器实例, 其中, 所述新建命令中包括灰度应用的镜像信息。

5. 根据权利要求 1 所述的方法, 其中, 在所述灰度应用完成测试之前, 所述方法还包括:

克隆原应用, 并在克隆的原应用版本的基础上, 添加新功能或者改进功能所对应的代码, 以生成灰度应用镜像;

根据所述原应用的容器实例的创建区域以及所述灰度应用镜像, 在测试环

境下所述区域中创建灰度应用的容器实例以及灰度应用所对应的负载均衡应用。

6. 根据权利要求 1 所述的方法, 其中, 所述在所述灰度应用所对应的负载均衡应用的配置文件中添加所述灰度应用新增的容器实例的配置代码, 包括:

5 检测所述灰度应用所对应的负载均衡应用的类型;

根据所述灰度应用所对应的负载均衡应用的类型, 通过与所述类型相匹配的方式获取所述负载均衡应用的配置文件并在所述配置文件中加入所述灰度应用新增的容器实例所对应的配置代码。

7. 根据权利要求 6 所述的方法, 其中, 所述根据所述灰度应用所对应的负载均衡应用的类型, 通过与所述类型相匹配的方式获取所述负载均衡应用的配置文件并在所述配置文件中加入所述灰度应用新增的容器实例所对应的配置代码, 包括:

若所述负载均衡应用的类型为 F5, 根据所述灰度应用确定 F5 负载均衡应用所在的 F5 设备;

15 在所述 F5 设备上获取所述 F5 负载均衡应用的配置文件, 并在所述配置文件中加入所述灰度应用新增的容器实例所对应的配置代码;

若所述负载均衡应用的类型为非 F5, 根据所述灰度应用确定所述负载均衡应用的容器, 其中非 F5 负载均衡应用包括 LVS 负载均衡应用、Haproxy 负载均衡应用、Nginx 负载均衡应用;

20 从对应的负载均衡应用的容器中获取配置文件, 将所获取的配置文件下载到本地;

在所述配置文件中加入所述灰度应用新增的容器实例所对应的配置代码;

将加入了所述配置代码的配置文件推送至对应的负载均衡应用的容器中;

25 重启对应的负载均衡应用的容器的负载均衡服务, 以使得加入了所述配置代码的配置文件生效。

8. 根据权利要求 7 所述的方法, 其中, 所述在所述配置文件中加入所述灰度应用新增的容器实例所对应的配置代码, 包括:

获取对应的容器实例配置代码模板;

30 在对应的容器实例配置代码模板中添加新增的灰度应用的容器实例的 IP 和端口信息以形成新增的容器实例的配置代码;

将新增的容器实例的配置代码追加到配置文件中最后一个容器实例的配置代码之后。

9. 根据权利要求 8 所述的方法, 其中, 所述 LVS 负载均衡应用中容器实例配置代码模板中包括容器实例 IP 信息、容器实例的端口信息、所述容器实例所对应的权重信息、所述容器实例所对应的健康检查方式信息;

所述 Nginx 负载均衡应用和所述 Haproxy 负载均衡应用中容器实例配置代码模板是以 server 节点开头的代码。

10. 一种负载均衡应用虚拟 IP 切换装置, 包括:

停用 IP 确定单元, 用于若灰度应用完成测试, 根据第一预设规则从原应用

的容器实例 IP 中确定第一预设数量的 IP，其中，所述第一预设数量小于所述原应用的容器实例数量，所述灰度应用是在所述原应用的版本上添加新功能或者改进功能之后所形成的应用；

5 停用单元，用于停用所确定的第一预设数量的 IP 以及所述第一预设数量的 IP 所对应的容器实例，其中，所停用的所述第一预设数量的 IP 被回收到预设 IP 地址池；

启用 IP 确定单元，用于根据第二预设规则从预设 IP 地址池所停用的 IP 中确定第二预设数量的需启用的 IP，其中，所述第二预设数量小于或者等于所停用的 IP 数量；

10 启用单元，用于将需启用的 IP 分配给所述灰度应用新增的容器实例；

添加单元，用于在所述灰度应用所对应的负载均衡应用的配置文件中添加所述灰度应用新增的容器实例所对应的配置代码；

15 绑定单元，用于将所述原应用的访问域名与所述原应用所对应的负载均衡应用的虚拟 IP 的对应关系解除，并将所述原应用的访问域名与所述灰度应用所对应的负载均衡应用的虚拟 IP 进行绑定，以实现通过所述原应用的访问域名来访问所述灰度应用。

11. 一种计算机设备，包括存储器、处理器及存储在所述存储器上并可在所述处理器上运行的计算机程序，其中，所述处理器执行所述计算机程序时实现以下步骤：

20 若灰度应用完成测试，根据第一预设规则从原应用的容器实例 IP 中确定第一预设数量的 IP，其中，所述第一预设数量小于所述原应用的容器实例数量，所述灰度应用是在所述原应用的版本上添加新功能或者改进功能之后所形成的应用；

25 停用所确定的第一预设数量的 IP 以及所述第一预设数量的 IP 所对应的容器实例，其中，所停用的所述第一预设数量的 IP 被回收到预设 IP 地址池；

根据第二预设规则从预设 IP 地址池所停用的 IP 中确定第二预设数量的需启用的 IP，其中，所述第二预设数量小于或者等于所停用的 IP 数量；

将需启用的 IP 分配给所述灰度应用新增的容器实例；

30 在所述灰度应用所对应的负载均衡应用的配置文件中添加所述灰度应用新增的容器实例所对应的配置代码；

将所述原应用的访问域名与所述原应用所对应的负载均衡应用的虚拟 IP 的对应关系解除，并将所述原应用的访问域名与所述灰度应用所对应的负载均衡应用的虚拟 IP 进行绑定，以实现通过所述原应用的访问域名来访问所述灰度应用。

35 12. 根据权利要求 11 所述的计算机设备，其中，所述停用所确定的第一预设数量的 IP 以及所述第一预设数量的 IP 所对应的容器实例，包括：

设置所确定的第一预设数量的 IP 的使用状态标识为停用状态；

销毁所述第一预设数量的 IP 所对应的容器实例；

回收使用状态标识为停用状态的 IP 到预设 IP 地址池；

根据所述第一预设数量修改所述原应用的容器实例数量。

13. 根据权利要求 11 所述的计算机设备, 其中, 所述将需启用的 IP 分配给所述灰度应用新增的容器实例, 包括:

5 根据所述灰度应用以及需启用的 IP 数量确定所述灰度应用需新增的容器实例所在的宿主机;

在所确定的宿主机上创建所述灰度应用的容器实例;

将需启用的 IP 分配给所述灰度应用新创建的容器实例。

14. 根据权利要求 13 所述的计算机设备, 其中, 所述在所确定的宿主机上创建所述灰度应用的容器实例, 包括:

10 发送新建命令到编排工具, 以通过所述编排工具推送所述新建命令到所确定的宿主机, 并通过所述新建命令在对应的宿主机上创建所述灰度应用所对应的容器实例, 其中, 所述新建命令中包括灰度应用的镜像信息。

15 15. 根据权利要求 11 所述的计算机设备, 其中, 在所述灰度应用完成测试之前, 所述处理器执行所述计算机程序时还实现以下步骤:

15 克隆原应用, 并在克隆的原应用版本的基础上, 添加新功能或者改进功能所对应的代码, 以生成灰度应用镜像;

根据所述原应用的容器实例的创建区域以及所述灰度应用镜像, 在测试环境下所述区域中创建灰度应用的容器实例以及灰度应用所对应的负载均衡应用。

20 16. 根据权利要求 11 所述的计算机设备, 其中, 所述在所述灰度应用所对应的负载均衡应用的配置文件中添加所述灰度应用新增的容器实例的配置代码, 包括:

检测所述灰度应用所对应的负载均衡应用的类型;

25 根据所述灰度应用所对应的负载均衡应用的类型, 通过与所述类型相匹配的方式获取所述负载均衡应用的配置文件并在所述配置文件中加入所述灰度应用新增的容器实例所对应的配置代码。

30 17. 根据权利要求 16 所述的计算机设备, 其中, 所述根据所述灰度应用所对应的负载均衡应用的类型, 通过与所述类型相匹配的方式获取所述负载均衡应用的配置文件并在所述配置文件中加入所述灰度应用新增的容器实例所对应的配置代码, 包括:

若所述负载均衡应用的类型为 F5, 根据所述灰度应用确定 F5 负载均衡应用所在的 F5 设备;

在所述 F5 设备上获取所述 F5 负载均衡应用的配置文件, 并在所述配置文件中加入所述灰度应用新增的容器实例所对应的配置代码;

35 若所述负载均衡应用的类型为非 F5, 根据所述灰度应用确定所述负载均衡应用的容器, 其中非 F5 负载均衡应用包括 LVS 负载均衡应用、Haproxy 负载均衡应用、Nginx 负载均衡应用;

从对应的负载均衡应用的容器中获取配置文件, 将所获取的配置文件下载到本地;

在所述配置文件中加入所述灰度应用新增的容器实例所对应的配置代码；
将加入了所述配置代码的配置文件推送至对应的负载均衡应用的容器中；
重启对应的负载均衡应用的容器的负载均衡服务，以使得加入了所述配置代码的配置文件生效。

5 18. 根据权利要求 17 所述的计算机设备，其中，所述在所述配置文件中加入所述灰度应用新增的容器实例所对应的配置代码，包括：

获取对应的容器实例配置代码模板；

在对应的容器实例配置代码模板中添加新增的灰度应用的容器实例的 IP 和端口信息以形成新增的容器实例的配置代码；

10 将新增的容器实例的配置代码追加到配置文件中最后一个容器实例的配置代码之后。

19. 根据权利要求 18 所述的计算机设备，其中，所述 LVS 负载均衡应用中容器实例配置代码模板中包括容器实例 IP 信息、容器实例的端口信息、所述容器实例所对应的权重信息、所述容器实例所对应的健康检查方式信息；

15 所述 Nginx 负载均衡应用和所述 Haproxy 负载均衡应用中容器实例配置代码模板是以 server 节点开头的代码。

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程序，所述计算机程序当被处理器执行时使所述处理器执行以下操作：

20 若灰度应用完成测试，根据第一预设规则从原应用的容器实例 IP 中确定第一预设数量的 IP，其中，所述第一预设数量小于所述原应用的容器实例数量，所述灰度应用是在所述原应用的版本上添加新功能或者改进功能之后所形成的应用；

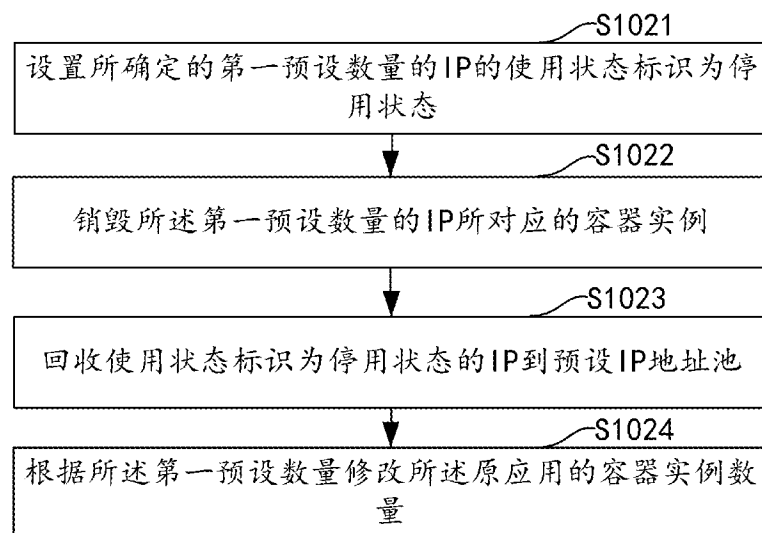
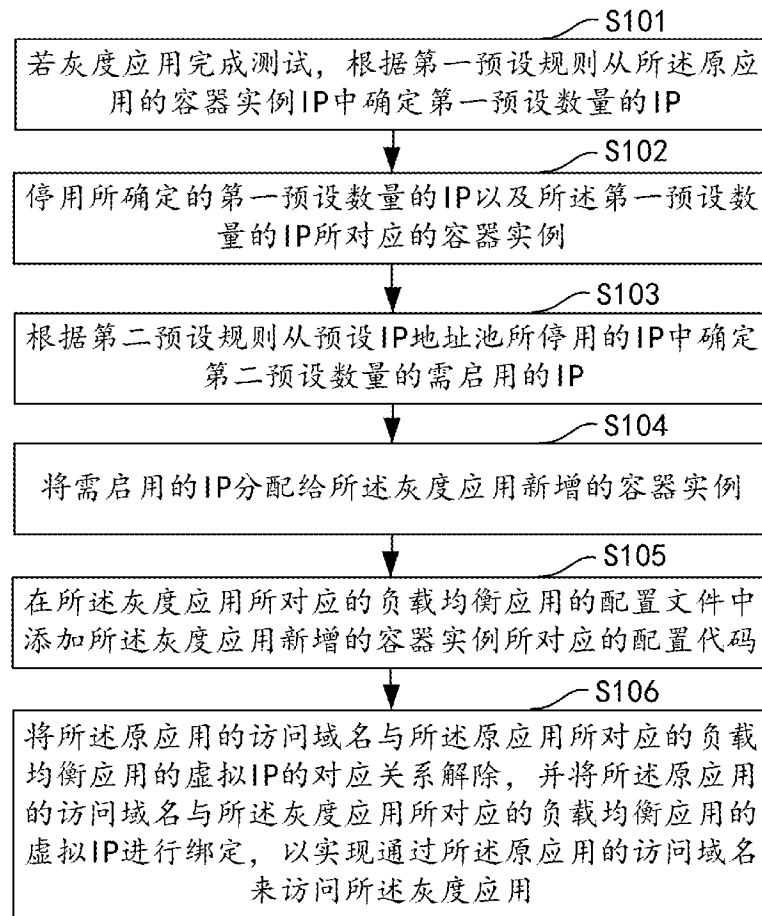
停用所确定的第一预设数量的 IP 以及所述第一预设数量的 IP 所对应的容器实例，其中，所停用的所述第一预设数量的 IP 被回收至预设 IP 地址池；

25 根据第二预设规则从预设 IP 地址池所停用的 IP 中确定第二预设数量的需启用的 IP，其中，所述第二预设数量小于或者等于所停用的 IP 数量；

将需启用的 IP 分配给所述灰度应用新增的容器实例；

在所述灰度应用所对应的负载均衡应用的配置文件中添加所述灰度应用新增的容器实例所对应的配置代码；

30 将所述原应用的访问域名与所述原应用所对应的负载均衡应用的虚拟 IP 的对应关系解除，并将所述原应用的访问域名与所述灰度应用所对应的负载均衡应用的虚拟 IP 进行绑定，以实现通过所述原应用的访问域名来访问所述灰度应用。



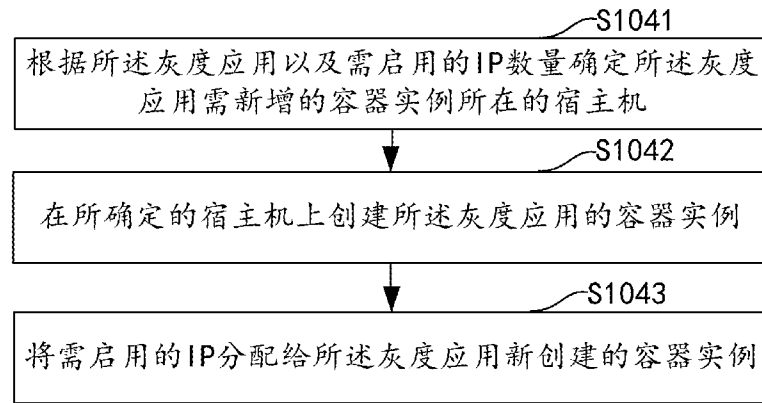


图 3

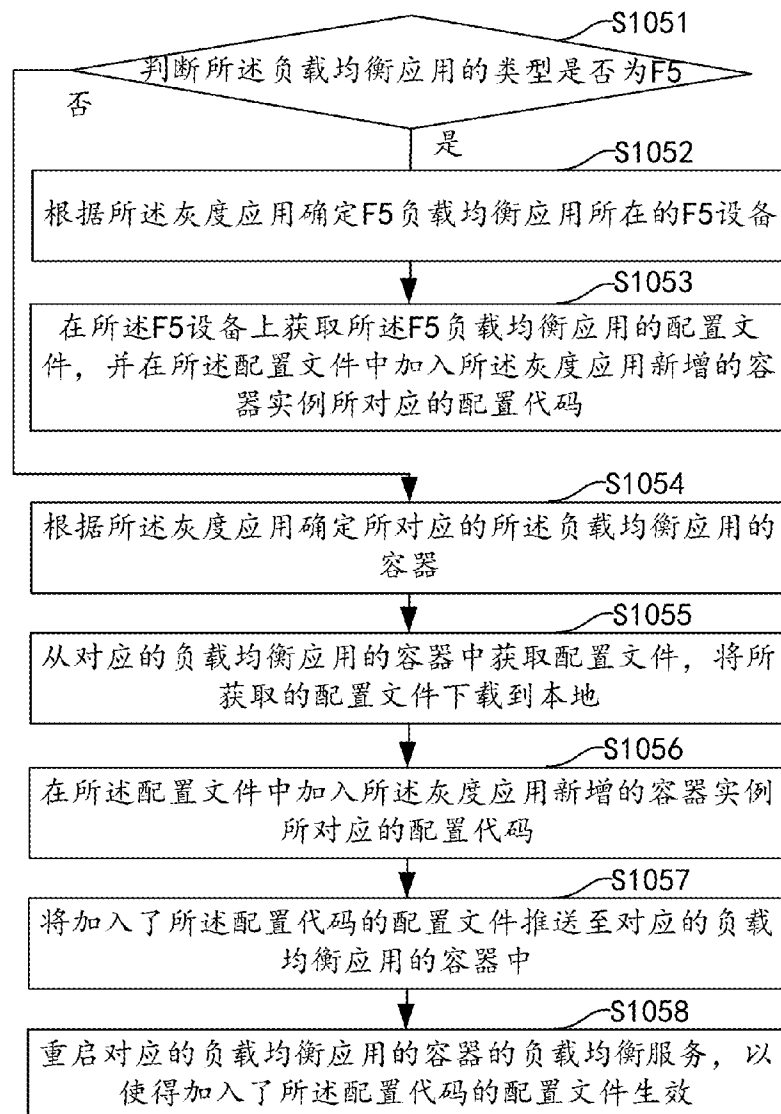


图 4

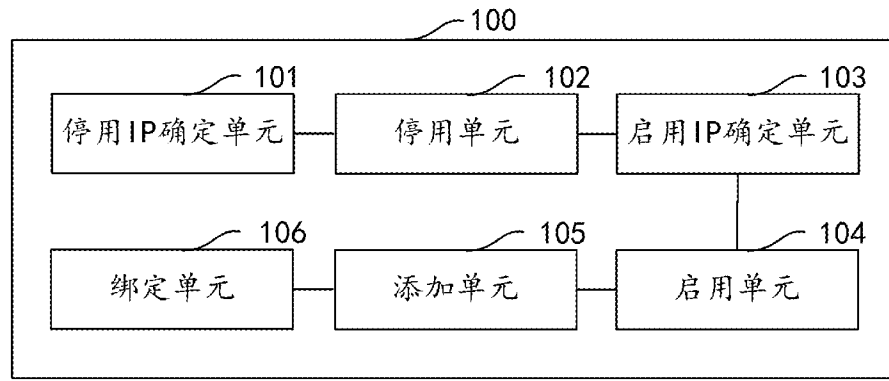


图 5

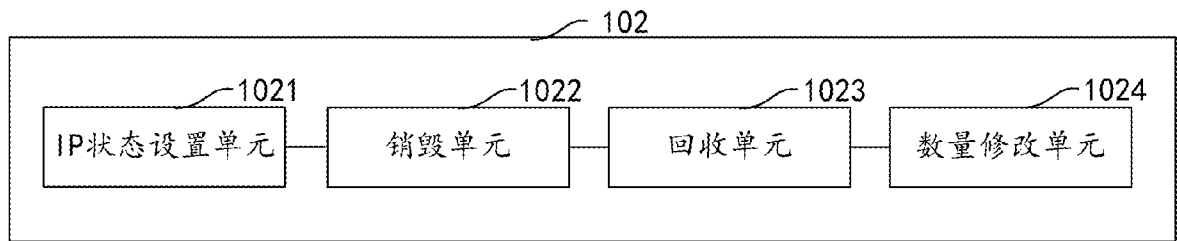


图 6

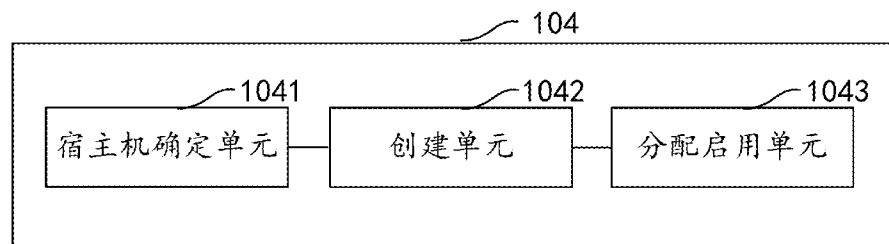


图 7

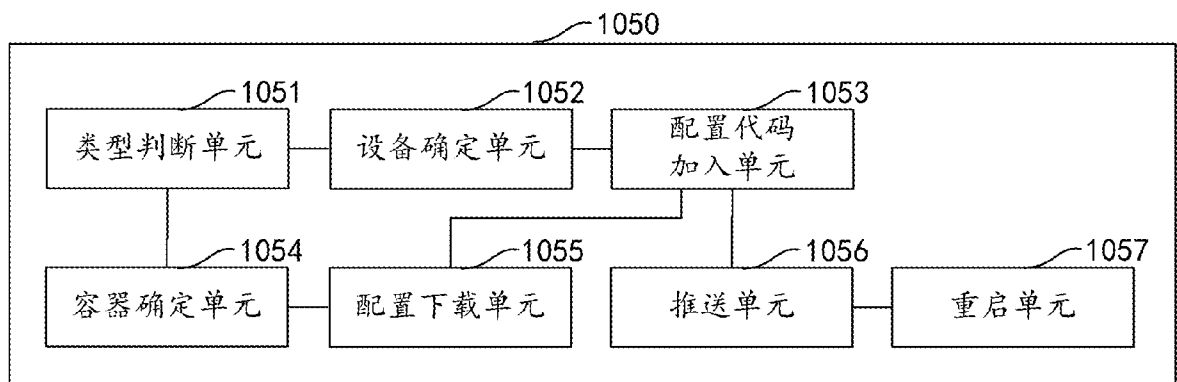


图 8

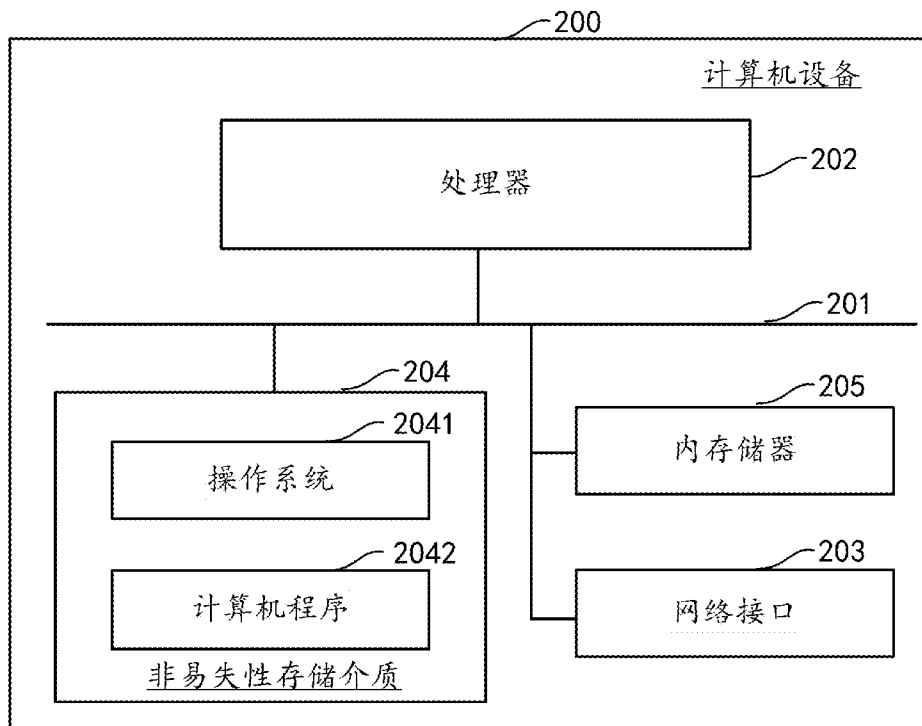


图 9

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2019/103262

A. CLASSIFICATION OF SUBJECT MATTER

H04L 29/08(2006.01)i; H04L 29/12(2006.01)i

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

H04L

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

CNABS; CNTXT; CNKI; VEN; USTXT; EPTXT; WOTXT: 负载均衡, 灰度, 应用, 发布, 切换, 更新, 升级, 版本, 测试, 地址池, 回收, 停用, 绑定, 解绑, 释放, 分配, 域名, load balancing, grayscale, application, switch, IP, address pool, recycling, distribution, allocation, version, update, binding, domain name

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	CN 107666525 A (BEIJING JINGDONG SHANGKE INFORMATION TECHNOLOGY CO., LTD. et al.) 06 February 2018 (2018-02-06) description, paragraphs [0039]-[0119]	1-20
A	CN 101729947 A (ZTE CORPORATION) 09 June 2010 (2010-06-09) entire document	1-20
A	CN 106412146 A (HANGZHOU DPTECH TECHNOLOGIES CO., LTD.) 15 February 2017 (2017-02-15) entire document	1-20
A	US 2007118558 A1 (RESEARCH IN MOTION LIMITED) 24 May 2007 (2007-05-24) entire document	1-20



Further documents are listed in the continuation of Box C.



See patent family annex.

* Special categories of cited documents:

“A” document defining the general state of the art which is not considered to be of particular relevance

“E” earlier application or patent but published on or after the international filing date

“L” document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

“O” document referring to an oral disclosure, use, exhibition or other means

“P” document published prior to the international filing date but later than the priority date claimed

“T” later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

“X” document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

“Y” document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

“&” document member of the same patent family

Date of the actual completion of the international search

07 January 2020

Date of mailing of the international search report

20 February 2020

Name and mailing address of the ISA/CN

China National Intellectual Property Administration (ISA/
CN)
No. 6, Xitucheng Road, Jimenqiao Haidian District, Beijing
100088
China

Facsimile No. (86-10)62019451

Authorized officer

Telephone No.

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/CN2019/103262

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
CN	107666525	A	06 February 2018	None			
CN	101729947	A	09 June 2010	None			
CN	106412146	A	15 February 2017	CN	106412146	B	06 September 2019
US	2007118558	A1	24 May 2007	US	8045958	B2	25 October 2011

国际检索报告

国际申请号

PCT/CN2019/103262

A. 主题的分类

H04L 29/08 (2006.01) i; H04L 29/12 (2006.01) i

按照国际专利分类 (IPC) 或者同时按照国家分类和 IPC 两种分类

B. 检索领域

检索的最低限度文献 (标明分类系统和分类号)

H04L

包含在检索领域中的除最低限度文献以外的检索文献

在国际检索时查阅的电子数据库 (数据库的名称, 和使用的检索词 (如使用))

CNABS; CNTXT; CNKI; VEN; USTXT; EPTXT; WOTXT; 负载均衡, 灰度, 应用, 发布, 切换, 更新, 升级, 版本, 测试, 地址池, 回收, 停用, 绑定, 解绑, 释放, 分配, 域名, load balancing, grayscale, application, switch, IP, address pool, recycling, distribution, allocation, version, update, binding, domain name

C. 相关文件

类 型*	引用文件, 必要时, 指明相关段落	相关的权利要求
A	CN 107666525 A (北京京东尚科信息技术有限公司 等) 2018年 2月 6日 (2018 - 02 - 06) 说明书第[0039]-[0119]段	1-20
A	CN 101729947 A (中兴通讯股份有限公司) 2010年 6月 9日 (2010 - 06 - 09) 全文	1-20
A	CN 106412146 A (杭州迪普科技有限公司) 2017年 2月 15日 (2017 - 02 - 15) 全文	1-20
A	US 2007118558 A1 (RESEARCH IN MOTION LTD) 2007年 5月 24日 (2007 - 05 - 24) 全文	1-20

☐ 其余文件在C栏的续页中列出。☒ 见同族专利附件。

* 引用文件的具体类型:

“A” 认为不特别相关的表示了现有技术一般状态的文件

“E” 在国际申请日的当天或之后公布的在先申请或专利

“L” 可能对优先权要求构成怀疑的文件, 或为确定另一篇引用文件的公布日而引用的或者因其他特殊理由而引用的文件 (如具体说明的)

“O” 涉及口头公开、使用、展览或其他方式公开的文件

“P” 公布日先于国际申请日但迟于所要求的优先权日的文件

“T” 在申请日或优先权日之后公布, 与申请不相抵触, 但为了理解发明之理论或原理的在后文件

“X” 特别相关的文件, 单独考虑该文件, 认定要求保护的发明不是新颖的或不具有创造性

“Y” 特别相关的文件, 当该文件与另一篇或者多篇该类文件结合并且这种结合对于本领域技术人员为显而易见时, 要求保护的发明不具有创造性

“&” 同族专利的文件

国际检索实际完成的日期

2020年 1月 7日

国际检索报告邮寄日期

2020年 2月 20日

ISA/CN的名称和邮寄地址

中国国家知识产权局 (ISA/CN)

中国北京市海淀区蓟门桥西土城路6号 100088

传真号 (86-10)62019451

受权官员

孙慧珠

电话号码 86- (512) -88996093

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2019/103262

检索报告引用的专利文件			公布日 (年/月/日)	同族专利	公布日 (年/月/日)
CN	107666525	A	2018年 2月 6日	无	
CN	101729947	A	2010年 6月 9日	无	
CN	106412146	A	2017年 2月 15日	CN 106412146 B	2019年 9月 6日
US	2007118558	A1	2007年 5月 24日	US 8045958 B2	2011年 10月 25日