(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

(43) 国际公布日
2022 年 5 月 5 日 (05.05.2022)

WIPO | PCT

(10) 国际公布号
WO 2022/089281 A1

(51) 国际专利分类号:
G06F 9/455 (2006.01)

(21) 国际申请号: PCT/CN2021/125158

(22) 国际申请日: 2021 年 10 月 21 日 (21.10.2021)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:
202011194163.X 2020年10月30日 (30.10.2020) CN

(71) 申请人: 阿里巴巴集团控股有限公司 (ALIBABA GROUP HOLDING LIMITED) [—/CN]; 开曼群岛大开曼资本大厦一座四层 847 号邮箱, Grand Cayman (KY)。

(72) 发明人: 史明伟 (SHI, Mingwei); 中国浙江省杭州市余杭区文一西路 969 号 3 号楼, Zhejiang 311121 (CN)。 周新宇 (ZHOU, Xinyu); 中国浙江省杭州市余杭区文一西路 969 号 3 号楼, Zhejiang 311121 (CN)。 许晓斌 (XU, Xiaobin); 中国浙江省杭州市余杭区文一西路969号3号楼, Zhejiang 311121 (CN)。 聂诗超 (NIE, Shichao); 中国浙江省杭州市余杭区文一西路969号3号楼, Zhejiang 311121 (CN)。 詹洲翔 (ZHAN, Zhouxiang); 中国浙江省杭州市余杭区文一西路969号3号楼, Zhejiang 311121 (CN)。 王川 (WANG, Chuan); 中国浙江省杭州市余杭区文一西路969号3号楼, Zhejiang 311121 (CN)。

(74) 代理人: 北京三友知识产权代理有限公司 (BEIJING SANYOU INTELLECTUAL PROPERTY AGENCY LTD.); 中国北京市金融街 35 号国际企业大厦 A 座 16 层, Beijing 100033 (CN)。

(54) **Title:** CONTAINER-BASED APPLICATION MANAGEMENT METHOD AND APPARATUS

(54) **发明名称:** 基于容器的应用管理方法和装置

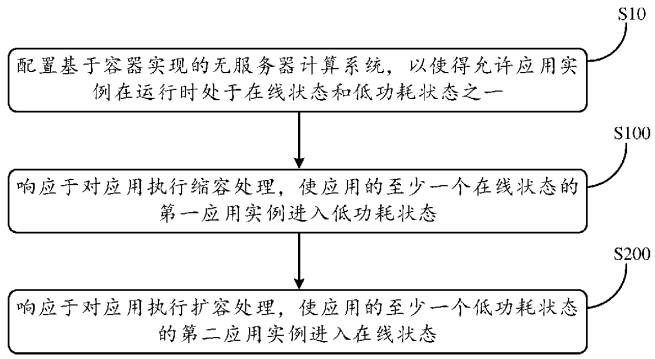

图 4

S10 CONFIGURE A SERVERLESS COMPUTING SYSTEM ACHIEVED ON THE BASIS OF A CONTAINER TO ALLOW APPLICATION INSTANCES TO BE IN ONE OF AN ONLINE STATE AND A LOW POWER CONSUMPTION STATE AT RUNTIME

S100 IN RESPONSE TO PERFORMING A CAPACITY REDUCTION PROCESS ON AN APPLICATION, BRING AT LEAST ONE FIRST APPLICATION INSTANCE OF THE APPLICATION IN THE ONLINE STATE INTO THE LOW POWER CONSUMPTION STATE

S200 IN RESPONSE TO PERFORMING A CAPACITY EXPANSION PROCESS ON THE APPLICATION, BRING AT LEAST ONE SECOND APPLICATION INSTANCE OF THE APPLICATION IN THE LOW POWER CONSUMPTION STATE INTO THE ONLINE STATE

(57) **Abstract:** Disclosed are a container-based application management method and apparatus. A serverless computing system achieved on the basis of a container is configured to allow application instances to be in one of an online state and a low power consumption state at runtime. In response to performing a capacity reduction process on an application, at least one first application instance of the application in the online state is brought into the low power consumption state; and in response to performing a capacity expansion process on the application, at least one second application instance of the application in the low power consumption state is brought into the online state. Thus, application instances can be rapidly and elastically scaled while the costs of the application instances are reduced.

WO 2022/089281 A1

[见续页]

(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, IT, 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, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, 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：公开了一种基于容器的应用管理方法和装置。配置基于容器实现的无服务器计算系统，以使得允许应用实例在运行时处于在线状态和低功耗状态之一。响应于对应用执行缩容处理，使应用的至少一个在线状态的第一应用实例进入低功耗状态；以及响应于对应用执行扩容处理，使应用的至少一个低功耗状态的第二应用实例进入在线状态。由此，能够在降低应用实例成本的同时，对应用实例进行快速弹性伸缩。