

(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 1 月 27 日 (27.01.2022)



(10) 国际公布号
WO 2022/016808 A1

(51) 国际专利分类号:
G06F 9/455 (2006.01)

(21) 国际申请号: PCT/CN2020/140019

(22) 国际申请日: 2020 年 12 月 28 日 (28.12.2020)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:
202010709708.X 2020 年 7 月 22 日 (22.07.2020) CN

(71) 申请人: 中国科学院深圳先进技术研究院 (SHENZHEN INSTITUTES OF ADVANCED TECHNOLOGY CHINESE ACADEMY OF SCIENCES) [CN/CN]; 中国广东省深圳市南山区深圳大学城学苑大道 1068 号, Guangdong 518055 (CN)。

(72) 发明人: 杨磊(YANG, Lei); 中国广东省深圳市南山区深圳大学城学苑大道 1068 号, Guangdong 518055 (CN)。 王洋(WANG, Yang); 中国广东省深圳市南山区深圳大学城学苑大道 1068 号, Guangdong

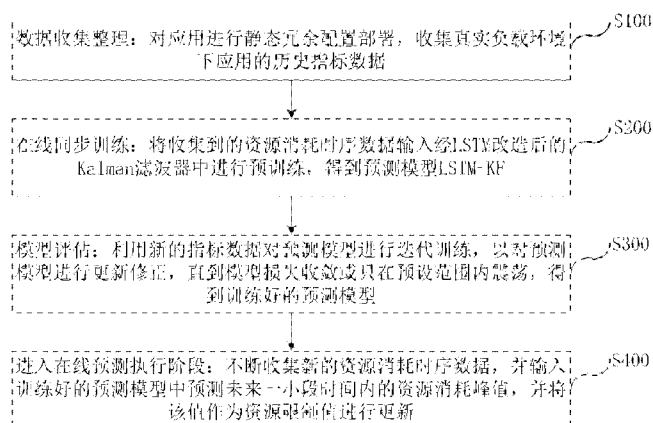
518055 (CN)。 须成忠(XU, Chengzhong); 中国广东省深圳市南山区深圳大学城学苑大道 1068 号, Guangdong 518055 (CN)。

(74) 代理人: 北京中巡通大知识产权代理有限公司 (BEIJING ZHONG XUN TONG DA INTELLECTUAL PROPERTY AGENCY CO., LTD.); 中国北京市海淀区杏石口路 50 号中间建筑 1 区 148 号, Beijing 100195 (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, 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。

(54) Title: KUBERNETES CLUSTER RESOURCE DYNAMIC ADJUSTMENT METHOD AND ELECTRONIC DEVICE

(54) 发明名称: 一种KUBERNETES集群资源动态调整方法及电子设备



S100 Data collection and organisation: implement static redundancy configuration deployment on an application to collect historical index data of the application in a real load environment

S200 Online synchronous training: inputting collected resource consumption timing data into an LSTM-transformed Kalman filter for pre-training to obtain a prediction model LSTM-KF

S300 Model evaluation: using new index data to iteratively train the prediction model in order to update and correct the prediction model until the model loss converges or only oscillates within a preset range to obtain a trained prediction model

S400 Enter an online prediction execution stage: continuously collecting new resource consumption timing data and inputting same into the trained prediction model to predict a resource consumption peak value in a future short period of time, and use the value as a resource limit value for resource updating

图 1

(57) Abstract: The present application relates to a Kubernetes cluster resource dynamic adjustment method and an electronic device. The method comprises: collecting historical monitoring index data in an application and, on the basis of the monitoring index data, obtaining monitoring resource consumption timing data; constructing an LSTM-Kalman filter, and inputting the resource consumption timing data into the LSTM-Kalman for pre-training to obtain a prediction model; collecting new resource consumption timing data, inputting the new resource consumption timing data into the prediction model to predict a resource consumption peak value in a future preset time period, and using the resource consumption peak value as a resource limit value for resource updating. The present application ensures the precision of the data and effectively increases the rate of utilisation of system resources, solving the problem of resource fragmentation, and also has a certain capacity for self-adaption.

(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: 本申请涉及一种Kubernetes集群资源动态调整方法及电子设备。所述方法包括: 收集应用中的历史监控指标数据, 根据所述监控指标数据得到资源消耗时序数据; 构建LSTM-Kalman滤波器, 将所述资源消耗时序数据输入所述LSTM-Kalman进行预训练, 得到预测模型; 收集新的资源消耗时序数据, 将所述新的资源消耗时序数据输入所述预测模型中预测未来预设时间段内的资源消耗峰值, 将所述资源消耗峰值作为资源限制值进行资源更新。本申请保证了数据的精确性, 有效的提高了系统资源利用率, 解决了资源碎片问题, 同时具备一定的自适应的能力。

一种Kubernetes集群资源动态调整方法及电子设备

技术领域

本申请属于计算机领域，特别涉及一种 Kubernetes 集群资源动态调整方法及电子设备。

背景技术

随着以Docker为支撑的容器即服务技术兴起，Kubernetes（一种容器集群管理技术，可以实现容器的编排管理）因其强大的容器编排技术受到业界追捧，已然成为容器集群管理的“事实标准”，更是被誉为“云时代的Linux”。为避免应用过多消耗资源，从而影响其他应用及节点性能，需要通过Kubernetes的资源限制机制以及Docker的Cgroup技术实现应用资源控制隔离。通常在部署应用的时候，只能通过配置文件进行静态部署，需要人为的估算Cpu、Memory等资源需求量。显然当预留资源不足会导致容器发生Cpu throttle或者Out of memory kill，影响服务质量，需要新建Pod（是Kubernetes最基本操作单元，包含一个或多个密切相关的业务逻辑容器，是集群调度的最小单位）再次部署。当出现冗余配置时，会使得集群资源利用率低下，影响节点可分配资源，导致其他Pod无法被调度到该节点上。

通常，如果不对Pod进行资源限制，Pod将会无限制的使用完节点上的资源，出现资源竞争，对服务的稳定性产生巨大影响。在实际生产环境中，流

量会发生规律性变化，因此资源使用率随之发生变化，静态部署的方式会带来资源不足或冗余。为了在资源不断波动情况下，能够动态调整资源，开源社区提供 HPA(Pod 横向自动扩容，<https://kubernetes.io/docs/tasks/run-application/horizontal-pod-autoscale/>)、VPA(垂直自动扩容，<https://github.com/kubernetes/community/blob/master/contributors/design-proposals/autoscaling/vertical-pod-autoscaler.md>)。HPA 的原理是通过 Kubernetes 的 Controller Manager 组件运行一个 Controller，周期性监听 Pod 资源使用情况，当高于设定的阈值时，会自动增加 Pod 的数量，当低于某个阈值时，会自动减少 Pod 数量，通过调整集群中 Pod 的数量来动态改变资源。VPA 与 HPA 相似，但是调整的是单个 Pod 的资源请求值，通过监控组件的数据，计算 Pod 的资源推荐值，允许在节点上进行适当调度，为每个 Pod 提供适当资源，当被 VPA 接管的 Pod 资源实际消耗值与推荐值相差较大时，会对 Pod 进行驱逐并重新建立 Pod 再次部署到集群中，从而达到动态资源改变的目的。

为了服务的稳定性，工业界也会进行全链路的压力测试，利用集成在 Kubernetes 中的 cAdvisor 监控系统，提前预判 Pod 需要的副本数和资源量，在无法准确预算资源时，采用冗余配置的方式。除此之外，学术界一般通过实验得到性能随资源变化的曲线，确定一个合适的资源预留值，提前预留好相应的资源。如文献[Xu, G., Xu, C.-Z MEER: Online Estimation of Optimal Reservations for Long Lived Containers in In-Memory Cluster Computing. ICDCS 2019]提出一种面向长生命周期的容器敏感内存资源预留机制，允许容器可以在比理想内存少很多的情况下运行，只需付出一定的性能损失代价即可。文献采用离线训练，得到模型后，在线预测获得 Pod 资源限额。

如上所述，若采用开源社区的HPA或VPA解决方案，对时间,Cpu,Memory,网络资源等敏感的在线或实时计算应用来说，实时性不够。通常部署的应用都是重量级Java框架或者相关技术栈的Web应用，利用HPA进行水平扩容来调整Pod数量，可以在较短时间能增加Pod数量，但是从触发扩容到经过Kube-Scheduler（Kubernetes的集群调度器，负责将Pod调度到合适的节点上）调度到合适的节点再通过Docker client拉取镜像，开启容器开启服务需要较长时间。至于VPA提供的垂直扩容，需要从节点驱逐旧Pod，重新建立满足资源需求的Pod，再通过API Server（Kubernetes资源对象的唯一操作入口，其他组件都必须通过它提供的API来操作资源数据），Kube-Scheduler重新部署，无法做到原地升级，显然并不满足实时性需求。并且VPA还处于alpha版，并没有合并到官方的Kubernetes release中，而且HPA和VPA目前不兼容，只能选择一个使用，否则两者之间会产生干扰。

对于全链路的离线、在线压测提前预判的方式，或者例如文献[谢文舟, 孙艳霞.基于Kubernetes负载特征的资源预测模型研究[J].网络安全技术与应用,2018(04):27-28.]通过在某几种应用类型上大量实验，不断改变资源预留值，测量应用的性能指标，从而建立资源和性能变化曲线，最后折中得到一个资源预留值。然而在生产环境的异构集群中存在成千上万的应用，各种应用版本发布更新频繁，应用链路复杂，各节点物理配置不尽相同。穷举所有应用进行压测找到最优或次优的资源调整方案，开销过大，耗时耗力，并不现实。即便可以，往往旧版本的测试数据可能并不适合新版本，或者相同的数据在不同的机器上，由于配置不同，表现出的性能往往也存在差异性，并且没有做到动态调整的效果。

发明内容

本申请提供了一种Kubernetes集群资源动态调整方法及电子设备，旨在至少在一定程度上解决现有技术中的上述技术问题之一。

为了解决上述问题，本申请提供了如下技术方案：

一种Kubernetes集群资源动态调整方法，包括以下步骤：

步骤a：收集应用中的历史监控指标数据，根据所述监控指标数据得到资源消耗时序数据；

步骤b：构建LSTM-Kalman滤波器，将所述资源消耗时序数据输入所述LSTM-Kalman进行预训练，得到预测模型；其中，所述LSTM-Kalman滤波器为利用LSTM改造后的Kalman滤波器；

步骤c：收集新的资源消耗时序数据，将所述新的资源消耗时序数据输入所述预测模型中预测未来预设时间段内的资源消耗峰值，将所述资源消耗峰值作为资源限制值进行资源更新。

本申请实施例采取的技术方案还包括：所述步骤a中，所述收集应用中的历史监控指标数据包括：

对运行在Pod中的容器进行监控并收集各种监控指标数据；收集的监控指标数据包括cpu利用率及限额、文件系统读/写利用率及限额、网络报文发送/接收/丢弃率数据中的至少一种或一种以上的组合；

将收集到的监控指标数据存入时序数据库。

本申请实施例采取的技术方案还包括：在所述步骤a中，所述根据所述监控指标数据得到资源消耗时序数据包括：

定时查询所述时序数据库得到感兴趣的指标数据，并对所述感兴趣的指标数据进行缓存及整理，形成可以直接用于预测的资源消耗时序数据。

本申请实施例采取的技术方案还包括：所述对所述感兴趣的指标数据进行缓存及整理包括：

使用滑动窗口依次滑动每一条指标数据，并选择窗口中的最大值作为新的数据集。

本申请实施例采取的技术方案还包括：在所述步骤b中，所述构建LSTM-Kalman滤波器包括：

假设观测值是系统状态值的有噪估计，令观测矩阵 $H=I$ ：

$$X(k) = f(X(k-1)) + W(k), \quad W \sim N(0, Q)$$

$$Z(k) = X(k) + V(k), \quad V \sim N(0, R)$$

上式中， $X(k)$ 为 k 时刻的系统状态， $Z(k)$ 为 k 时刻的观测值， $W(k)$ 和 $V(k)$ 分别表示过程和测量噪声，它们的协方差分别为 Q ， R ； f 为由LSTM_f模型生成的函数模型；

预测步骤为：

$$X(k|k-1) = f(X(k-1|k-1))$$

$$P(k|k-1) = F P(k-1|k-1) F^T + Q(k)$$

上式中， F 代表 f 关于 $X(k-1|k-1)$ 的雅可比矩阵， $Q(k)$ 由LSTM_Q给出；

更新步骤为：

$$K_g(k) = P(k|k-1) ((P(k|k-1) + R(k))^{-1})$$

$$X(k|k) = X(k|k-1) + K_g(k)(Z(k) - X(k|k-1))$$

$$P(k|k) = (I - K_g(k))P(k|k-1)$$

上式中, $R(k)$ 由 $LSTM_r$ 得到, $Z(k)$ 代表 k 时刻的观测值;

将配置文件中的request值作为 $X(k-1|k-1)$, 当前时刻 k 采集到的指标数据作为观测值 $Z(k|k)$, 分别经过所述 $LSTM_f$ 和 $LSTM_r$ 网络得到 k 时刻的预测值 $X(k|k-1)$ 和观测噪声 $R(k)$; $X(k|k-1)$ 经过 $LSTM_Q$ 得到 k 时刻的状态噪声 $Q(k)$; 利用 $R(k)$ 、 $Q(k)$ 、 $X(k-1|k-1)$ 对所述Kalman滤波器进行递推更新。

本申请实施例采取的技术方案还包括: 在所述步骤c中, 所述将所述资源消耗时序数据输入所述LSTM-Kalman进行预训练, 得到预测模型还包括:

模型评估: 利用新的监控指标数据对所述预测模型进行迭代训练, 直到所述预测模型损失收敛或只在预设范围内震荡时停止迭代训练。

本申请实施例采取的技术方案还包括: 所述模型评估还包括:

获取预设时间段内的新的监控指标数据, 将所述预测模型的预测结果与所述新的监控指标数据进行比较, 判断是否需要再次进行模型训练: 如果大于设定阈值的时间消耗值在所述预测模型的预测结果以下则不再更新模型, 反之则使用所述新的监控指标数据更新所述预测模型的参数, 对所述预测模型进行修正。

本申请实施例采取的技术方案还包括: 在所述步骤c中, 所述预测未来预设时间段内的资源消耗峰值还包括:

待所述 $LSTM_f$ 、 $LSTM_r$ 、 $LSTM_Q$ 三个子循环网络损失收敛或者降低到预设阈值, 开始执行递推估计, 通过度量收集器上报数据给预测模型, 所述预测模型对未来预设时间段内的资源消耗峰值进行最优估计, 使用最优估计值对Pod中容器资源进行更新操作。

本申请实施例采取的又一技术方案为: 一种电子设备, 包括:

至少一个处理器；以及

与所述至少一个处理器通信连接的存储器；其中，

所述存储器存储有可被所述一个处理器执行的指令，所述指令被所述至少一个处理器执行，以使所述至少一个处理器能够执行上述的Kubernetes集群资源动态调整方法的以下操作：

步骤a：收集应用中的历史监控指标数据，根据所述监控指标数据得到资源消耗时序数据；

步骤b：构建LSTM-Kalman滤波器，将所述资源消耗时序数据输入所述LSTM-Kalman进行预训练，得到预测模型；其中，所述LSTM-Kalman滤波器为利用LSTM改造后的Kalman滤波器；

步骤c：收集新的资源消耗时序数据，将所述新的资源消耗时序数据输入所述预测模型中预测未来预设时间段内的资源消耗峰值，将所述资源消耗峰值作为资源限制值进行资源更新。

相对于现有技术，本申请实施例产生的有益效果在于：本申请实施例的Kubernetes集群资源动态调整方法及电子设备首先对应用进行静态冗余配置部署，收集真实负载环境下应用的历史监控指标数据，在线进行预训练得到较为可靠的预测模型；然后对预测模型进行迭代训练，直到模型损失收敛或只在预设范围内震荡后，不断采集新的资源消耗时序数据输入预测模型中预测未来一小段时间内资源消耗的峰值，将该值作为资源限制值进行更新，实现动态的Pod资源限制调整。相对于现有技术，本申请至少具有以下有益效果：

1、面对流量“洪峰”短时间内可以起到良好的缓冲作用，保证服务的实时性，稳定性；面对流量“低谷”，又可将空闲的资源归还内核，有效的提高了系统资源利用率，解决了资源碎片问题，同时具备一定的自适应的能力。

2、在预测的同时进行滤噪处理，保证数据的精确性。

3、规避了应用版本差异、集群异构性导致模型不适用的问题，无需进行大量的压测试验，建立性能和资源关系曲线。

4、通过预测未来一段时间资源消耗峰值作为资源限制值，对资源进行动态限制，采用局部最大值代替全局最大值的配置思想，提高了资源利用率，起到了自适应的效果。

附图说明

图1是本申请实施例的Kubernetes集群资源动态调整方法的流程图；

图2为本申请实施例的Kubernetes集群资源动态调整系统结构示意图；

图3为经LSTM改造后的Kalman 滤波器示意图；

图4是本申请实施例提供的Kubernetes集群资源动态调整方法的硬件设备结构示意图。

具体实施方式

为了使本申请的目的、技术方案及优点更加清楚明白，以下结合附图及实施例，对本申请进行进一步详细说明。应当理解，此处所描述的具体实施例仅用以解释本申请，并不用于限定本申请。

针对现有技术的不足，本申请实施例的Kubernetes集群资源动态调整方法提出一种Kubernetes Pod资源在线预测，动态调整解决方案。整个方案包括数据收集整理、在线同步训练、在线预测执行三个阶段。首先对应用进行静态冗余配置部署，收集真实负载环境下应用的Cpu、Memory等历史数据，在线进行预训练得到较为可靠的预测模型；并对预测模型进行迭代训练，直到模型损失收敛或只在预设范围内震荡，则停止训练，并进入在线预测执行阶段，不断采集新的资源消耗时序数据，并输入预测模型中预测未来一小段时间内资源消耗的峰值，将该值作为资源限制值进行更新，实现动态的Pod资源限制调整。

具体的，请参阅图1，是本申请实施例的Kubernetes集群资源动态调整方法的流程图。本申请实施例的Kubernetes集群资源动态调整方法包括以下步骤：

步骤100：数据收集整理：首先对应用进行静态冗余配置部署，收集真实负载环境下应用的Cpu、Memory等历史监控指标数据；

步骤100中，数据收集整理具体包括：

步骤101：通过集成在Kubelet中的cAdvisor对运行在Pod中的容器进行监控并收集各种监控指标数据；

本步骤中，Kubelet为运行在集群每个节点上的进程，用于处理Master节点下发到本节点的任务并负责管理Pod及Pod中的容器。cAdvisor集成于Kubelet组件内，Kubelet从cAdvisor获取单独的容器使用统计信息，然后通过REST API暴露这些聚合后的Pod资源使用的统计信息。cAdvisor是一个开源的分析容器资源使用率以及性能特性的代理工具，它会自动查找所在节点上

的容器，收集每个容器的相关指标数据。本申请实施例中，cAdvisor收集的指标数据包括cpu利用率及限额、文件系统读/写利用率及限额、网络报文发送/接收/丢弃率等数据中的至少一种或一种以上的组合。具体请一并参阅图2，是本申请实施例的Kubernetes集群资源动态调整系统结构图。

步骤102：通过Prometheus将收集到的各种监控指标数据存入TSDB(时序数据库)；

本步骤中，Prometheus是一个开源的服务监控系统 and 时序数据库，提供了通用的数据模型和便捷的数据采集、存储和查询接口。Prometheus的核心组件Prometheus Server（普罗米修斯服务器）定期从静态配置的监控目标或基于服务发现机制自动配置的目标中拉取监控指标数据。在本申请实施例中，Prometheus Server通过cAdvisor提供的metrics接口周期性的进行数据抓取操作，然后将抓取的数据以时间序列的方式存储到服务端内存缓存区或持久化到存储设备中。

步骤103：通过自定义组件Metrics Collector（度量收集器）定时从Prometheus中收集感兴趣的监控指标数据做二级缓存，并对监控指标数据进行整理，形成可以直接用于预测的资源消耗时序数据；

本步骤中，自定义组件Metrics Collector通过查询语言PromQL定时查询Prometheus Server的时序数据库得到感兴趣的指标数据，并对指标进行缓存及整理，形成可以直接用于预测的资源消耗时序数据，减轻TSDB的压力，同时也保证时序数据的实时性，便于Kalman滤波器进行模型训练和估计。

本申请实施例中，在对时序数据进行处理时，使用滑动窗口依次滑动每一条数据，并选择窗口中的最大值作为新的数据集，使得滤波器预测出的值在未来一段时间内，资源消耗都不会超过预测值。

步骤200：在线同步训练：构建LSTM-Kalman滤波器，将Metrics Collector收集到的资源消耗时序数据输入LSTM-Kalman滤波器中进行预训练，得到预测模型LSTM-KF；

步骤200中，Kalman滤波器是一个最优化自回归数据处理算法，Kalman利用目标的动态信息，设法去掉噪声的影响，得到一个关于目标最优估计。考虑到系统的振荡，应用间的相互干扰等因素造成采集的数据不准，因此利用Kalman滤波器规避掉这些不稳定因素，得到较好的估计值。同时模型仅仅通过5步递推就可以实现对新值的最优估计，但是现有的Kalman滤波器在实际应用中需要依赖动力学模型。

Kalman的空间状态方程可以使用一个线性随机差分方程描述：

$$X(k) = AX(k-1) + W(k), \quad W \sim N(0, Q)(1)$$

$$Z(k) = HX(k) + V(k), \quad V \sim N(0, R)(2)$$

式(1)代表状态方程，式(2)代表测量方程。其中 $X(k)$ 为 k 时刻的系统状态， $Z(k)$ 为 k 时刻的观测值， A 为状态转移矩阵， H 为观测矩阵， $W(k)$ 和 $V(k)$ 分别表示过程和测量噪声，假设服从高斯白噪声，它们的协方差分别为 Q ， R 。

现有的Kalman滤波算法如下：

利用系统的当前状态模型预测系统的下一状态；假设系统的当前状态 k ，基于系统的上一状态预测出当前状态：

$$X(k|k-1) = A X(k-1|k-1) \quad (3)$$

式(3)中, $X(k|k-1)$ 是利用上一状态预测的结果, $X(k-1|k-1)$ 是上一状态最优的估计。对应于 $X(k|k-1)$ 的协方差:

$$P(k|k-1) = A P(k-1|k-1) A^T + Q \quad (4)$$

式(3)、(4)对应了系统预测过程, 得到k状态下的预测值, 然后结合测量值 $Z(t)$, 得到k状态下的最优估计值 $X(k|k)$:

$$X(k|k) = X(k|k-1) + K_g(k)(Z(k) - HX(k|k-1)) \quad (5)$$

式(5)中, K_g 为Kalman Gain(增益):

$$K_g(k) = P(k|k-1)H^T / ((H^T P(k|k-1)H + R)) \quad (6)$$

为了使得滤波器不断迭代递推下去, 还需要更新 $X(k|k)$ 状态下的协方差:

$$P(k|k) = (I - K_g(k)H)P(k|k-1) \quad (7)$$

式(7)中, I 是全1的矩阵。当进入 $k+1$ 状态时, $P(k|k)$ 即为式(4)中的 $P(k|k-1)$ 。

上述算法非常依赖动力学模型, 然而对于实际问题, 动力学特性并不知道, 无法先验的给出状态转移矩阵 A 和观测矩阵 H , 以及噪声分也是常常假设服从高斯分布。这些参数的假设存在很多人为主观因素, 使得模型和实际数据不能很好的匹配, 导致较大的预测误差。

因此, 本申请通过使用LSTM对Kalman滤波器进行改造, 使其摆脱对动力学模型的过分依赖, 使得模型参数不需要先验的给出。具体改造过程如下:

给出新的状态空间方程:

假设观测值是系统状态值的有噪估计，令观测矩阵 $H=I$ ：

$$X(k) = f(X(k-1)) + W(k), \quad W \sim N(0, Q) \quad (8)$$

$$Z(k) = X(k) + V(k), \quad V \sim N(0, R) \quad (9)$$

式（9）中， f 为由 $LSTM_f$ 模型生成的函数模型。

新的预测步骤为：

$$X(k|k-1) = f(X(k-1|k-1)) \quad (10)$$

$$P(k|k-1) = F P(k-1|k-1) F^T + Q(k) \quad (11)$$

式（11）中， F 代表 f 关于 $X(k-1|k-1)$ 的雅可比矩阵， $Q(k)$ 由 $LSTM_Q$ 给出。

新的更新步骤为：

$$K_g(k) = P(k|k-1) ((P(k|k-1) + R(k))^{-1}) \quad (12)$$

$$X(k|k) = X(k|k-1) + K_g(k)(Z(k) - X(k|k-1)) \quad (13)$$

$$P(k|k) = (I - K_g(k))P(k|k-1) \quad (14)$$

式（12）、（13）中， $R(k)$ 由 $LSTM_r$ 得到， $Z(k)$ 代表 k 时刻的观测值。

如图3所示，为经LSTM改造后的Kalman滤波器示意图。初始化时，可以将配置文件中的request值作为 $X(k-1|k-1)$ ，当前时刻 k 采集到的指标数据作为观测值 $Z(k|k)$ ，分别经过 $LSTM_f$ 和 $LSTM_r$ 网络得到 k 时刻的预测值 $X(k|k-1)$ 和观测噪声 $R(k)$ 。 $X(k|k-1)$ 经过 $LSTM_Q$ 得到 k 时刻的状态噪声 $Q(k)$ 。利用 $R(k)$ 、 $Q(k)$ 、 $X(k-1|k-1)$ ，通过式(12)、(13)、(14)对Kalman滤波器进行递推更新。

基于上述，经过LSTM改造后的Kalman滤波器兼顾预测和滤噪的同时，摆脱了对动力学模型的过分依赖，模型参数不用先验的给出，可以从数据中

学习出来，充分发挥时间序列前后相关性；并且增强了长程预测能力，而不仅仅依赖于上一个时刻值，预测精度可以显著提高。

步骤300：模型评估：利用新的监控指标数据对预测模型进行迭代训练，以对预测模型进行更新修正，直到模型损失收敛或只在预设范围内震荡，得到训练好的预测模型LSTM-KF；

步骤300中，模型评估过程具体为：通过Controller拉取Metrics Collector设定时间段内的监控指标数据，并对预测模型的预测结果和从Metrics Collector上报的监控指标数据进行比较，判断是否需要再次进行模型训练：如果设定阈值的时间消耗值在预测模型的最优估计值以下则不再更新模型，反之则再次使用新的监控指标数据更新预测模型参数，对模型进行修正。Controller中还备份有一个历史数据峰值，以备滤波器再次训练时，临时作为容器资源限制使用。

步骤400：进入在线预测执行阶段：不断收集新的资源消耗时序数据，并输入训练好的预测模型中预测未来预设时间段内的资源消耗峰值，并将该值作为资源限制值进行更新；

步骤400中，待LSTM_f、LSTM_r、LSTM_Q三个子循环网络损失收敛或者降低到预设阈值，开始执行递推估计，Metrics Collector上报数据给预测模型LSTM-KF，预测模型LSTM-KF对接下来一小段时间的资源消耗峰值进行最优估计，Executor直接使用最优估计值对Pod中容器所在的Container Cgroup的限制值进行更新操作。本申请采用最优估计值代表未来一小段时间峰值，即使用局部峰值代替全局峰值，即可避免冗余配置，又可以保证服务的稳定性，提高了资源利用率，起到了自适应的效果。

Cgroup是Linux内核中用于实现资源使用和统计的模块， Docker利用Cgroup进行资源隔离及限制， Kubelet作为Kubernetes的节点代理， 所有对Cgroup的操作都是由内部的Container Manager模块实现， Container Manager会通过Cgroup对资源使用层层限制。当通过yaml等配置文件对容器资源请求和限制量指定时， Docker会为容器设置运行时的cpu.share、cpu.quota、cpu.period、mem.limit等指标， Executor通过修改这些指标文件进行资源更新， 从而实现动态的Pod资源限制调整。

本申请实施例中， 由于各组件之间存在频繁的调用关系， 为了部署方便， 可以借助Sidecar（一种将应用功能从应用本身剥离出来作为单独进程的方式。做到无侵入的向应用添加多种功能）思想将组件容器化， 部署在同一个Pod内， 便于运维管理。

图4是本申请实施例提供的Kubernetes集群资源动态调整方法的硬件设备结构示意图。如图4所示， 该设备包括一个或多个处理器以及存储器。以一个处理器为例， 该设备还可以包括： 输入系统和输出系统。

处理器、存储器、输入系统和输出系统可以通过总线或者其他方式连接， 图4中以通过总线连接为例。

存储器作为一种非暂态计算机可读存储介质， 可用于存储非暂态软件程序、非暂态计算机可执行程序以及模块。处理器通过运行存储在存储器中的非暂态软件程序、指令以及模块， 从而执行电子设备的各种功能应用以及数据处理， 即实现上述方法实施例的处理方法。

存储器可以包括存储程序区和存储数据区， 其中， 存储程序区可存储操作系统、至少一个功能所需要的应用程序；存储数据区可存储数据等。此外， 存储器可以包括高速随机存取存储器， 还可以包括非暂态存储器， 例如

至少一个磁盘存储器件、闪存器件、或其他非暂态固态存储器件。在一些实施例中，存储器可选包括相对于处理器远程设置的存储器，这些远程存储器可以通过网络连接至处理系统。上述网络的实例包括但不限于互联网、企业内部网、局域网、移动通信网及其组合。

输入系统可接收输入的数字或字符信息，以及产生信号输入。输出系统可包括显示屏等显示设备。

所述一个或者多个模块存储在所述存储器中，当被所述一个或者多个处理器执行时，执行上述任一方法实施例的以下操作：

步骤a：收集应用中的历史监控指标数据，根据所述监控指标数据得到资源消耗时序数据；

步骤b：构建LSTM-Kalman滤波器，将所述资源消耗时序数据输入所述LSTM-Kalman进行预训练，得到预测模型；其中，所述LSTM-Kalman滤波器为利用LSTM改造后的Kalman滤波器；

步骤c：收集新的资源消耗时序数据，将所述新的资源消耗时序数据输入所述预测模型中预测未来预设时间段内的资源消耗峰值，将所述资源消耗峰值作为资源限制值进行资源更新。

上述产品可执行本申请实施例所提供的方法，具备执行方法相应的功能模块和有益效果。未在本实施例中详尽描述的技术细节，可参见本申请实施例提供的方法。

本申请实施例提供了一种非暂态（非易失性）计算机存储介质，所述计算机存储介质存储有计算机可执行指令，该计算机可执行指令可执行以下操作：

步骤a: 收集应用中的历史监控指标数据, 根据所述监控指标数据得到资源消耗时序数据;

步骤b: 构建LSTM-Kalman滤波器, 将所述资源消耗时序数据输入所述LSTM-Kalman进行预训练, 得到预测模型; 其中, 所述LSTM-Kalman滤波器为利用LSTM改造后的Kalman 滤波器;

步骤c: 收集新的资源消耗时序数据, 将所述新的资源消耗时序数据输入所述预测模型中预测未来预设时间段内的资源消耗峰值, 将所述资源消耗峰值作为资源限制值进行资源更新。

本申请实施例提供了一种计算机程序产品, 所述计算机程序产品包括存储在非暂态计算机可读存储介质上的计算机程序, 所述计算机程序包括程序指令, 当所述程序指令被计算机执行时, 使所述计算机执行以下操作:

步骤a: 收集应用中的历史监控指标数据, 根据所述监控指标数据得到资源消耗时序数据;

步骤b: 构建LSTM-Kalman滤波器, 将所述资源消耗时序数据输入所述LSTM-Kalman进行预训练, 得到预测模型; 其中, 所述LSTM-Kalman滤波器为利用LSTM改造后的Kalman 滤波器;

步骤c: 收集新的资源消耗时序数据, 将所述新的资源消耗时序数据输入所述预测模型中预测未来预设时间段内的资源消耗峰值, 将所述资源消耗峰值作为资源限制值进行资源更新。

本申请实施例的Kubernetes集群资源动态调整方法及电子设备首先对应用进行静态冗余配置部署, 收集真实负载环境下应用的历史监控指标数据, 在线进行预训练得到较为可靠的预测模型; 然后对预测模型进行迭代训练, 直

到模型损失收敛或只在预设范围内震荡后，不断采集新的资源消耗时序数据输入预测模型中预测未来一小段时间内资源消耗的峰值，将该值作为资源限制值进行更新，实现动态的Pod资源限制调整。相对于现有技术，本申请至少具有以下有益效果：

1、面对流量“洪峰”短时间内可以起到良好的缓冲作用，保证服务的实时性，稳定性；面对流量“低谷”，又可将空闲的资源归还内核，有效的提高了系统资源利用率，解决了资源碎片问题，同时具备一定的自适应的能力。

2、在预测的同时进行滤噪处理，保证数据的精确性。

3、规避了应用版本差异、集群异构性导致模型不适用的问题，无需进行大量的压测试验，建立性能和资源关系曲线。

4、通过预测未来一段时间资源消耗峰值作为资源限制值，对资源进行动态限制，采用局部最大值代替全局最大值的配置思想，提高了资源利用率，起到了自适应的效果。

对所公开的实施例的上述说明，使本领域专业技术人员能够实现或使用本申请。对这些实施例的多种修改对本领域的专业技术人员来说将是显而易见的，本申请中所定义的一般原理可以在不脱离本申请的精神或范围的情况下，在其它实施例中实现。因此，本申请将不会被限制于本申请所示的这些实施例，而是要符合与本申请所公开的原理和新颖特点相一致的最宽的范围。

权利要求书

1、一种Kubernetes集群资源动态调整方法，其特征在于，包括以下步骤：

步骤a：收集应用中的历史监控指标数据，根据所述监控指标数据得到资源消耗时序数据；

步骤b：构建LSTM-Kalman滤波器，将所述资源消耗时序数据输入所述LSTM-Kalman进行预训练，得到预测模型；其中，所述LSTM-Kalman滤波器为利用LSTM改造后的Kalman滤波器；

步骤c：收集新的资源消耗时序数据，将所述新的资源消耗时序数据输入所述预测模型中预测未来预设时间段内的资源消耗峰值，将所述资源消耗峰值作为资源限制值进行资源更新。

2、根据权利要求1所述的Kubernetes集群资源动态调整方法，其特征在于，所述步骤a中，所述收集应用中的历史监控指标数据包括：

对运行在Pod中的容器进行监控并收集各种监控指标数据；收集的监控指标数据包括cpu利用率及限额、文件系统读/写利用率及限额、网络报文发送/接收/丢弃率数据中的至少一种或一种以上的组合；

将收集到的监控指标数据存入时序数据库。

3、根据权利要求2所述的Kubernetes集群资源动态调整方法，其特征在于，在所述步骤a中，所述根据所述监控指标数据得到资源消耗时序数据包括：

定时查询所述时序数据库得到感兴趣的指标数据，并对所述感兴趣的指标数据进行缓存及整理，形成可以直接用于预测的资源消耗时序数据。

4、根据权利要求3所述的Kubernetes集群资源动态调整方法，其特征在于，所述对所述感兴趣的指标数据进行缓存及整理包括：

使用滑动窗口依次滑动每一条指标数据，并选择窗口中的最大值作为新的数据集。

5、根据权利要求1所述的Kubernetes集群资源动态调整方法，其特征在于，在所述步骤b中，所述构建LSTM-Kalman滤波器包括：

假设观测值是系统状态值的有噪估计，令观测矩阵 $H=I$ ：

$$X(k) = f(X(k-1)) + W(k), \quad W \sim N(0, Q)$$

$$Z(k) = X(k) + V(k), \quad V \sim N(0, R)$$

上式中， $X(k)$ 为 k 时刻的系统状态， $Z(k)$ 为 k 时刻的观测值， $W(k)$ 和 $V(k)$ 分别表示过程和测量噪声，它们的协方差分别为 Q ， R ； f 为由 $LSTM_f$ 模型生成的函数模型；

预测步骤为：

$$X(k|k-1) = f(X(k-1|k-1))$$

$$P(k|k-1) = F P(k-1|k-1) F^T + Q(k)$$

上式中， F 代表 f 关于 $X(k-1|k-1)$ 的雅可比矩阵， $Q(k)$ 由 $LSTM_Q$ 给出；

更新步骤为：

$$K_g(k) = P(k|k-1) ((P(k|k-1) + R(k))^{-1})$$

$$X(k|k) = X(k|k-1) + K_g(k)(Z(k) - X(k|k-1))$$

$$P(k|k) = (I - K_g(k))P(k|k-1)$$

上式中， $R(k)$ 由 $LSTM_r$ 得到， $Z(k)$ 代表 k 时刻的观测值；

将配置文件中的request值作为 $X(k-1|k-1)$ ，当前时刻 k 采集到的指标数据作为观测值 $Z(k|k)$ ，分别经过所述 $LSTM_f$ 和 $LSTM_r$ 网络得到 k 时刻的预测值 $X(k|k-1)$ 和观测噪声 $R(k)$ ； $X(k|k-1)$ 经过 $LSTM_Q$ 得到 k 时刻的状态噪声 $Q(k)$ ；利用 $R(k)$ 、 $Q(k)$ 、 $X(k-1|k-1)$ 对所述Kalman滤波器进行递推更新。

6、根据权利要求1至5任一项所述的Kubernetes集群资源动态调整方法，其特征在于，在所述步骤c中，所述预将所述资源消耗时序数据输入所述LSTM-Kalman进行预训练，得到预测模型还包括：

模型评估：利用新的监控指标数据对所述预测模型进行迭代训练，直到所述预测模型损失收敛或只在预设范围内震荡时停止迭代训练。

7、根据权利要求6所述的Kubernetes集群资源动态调整方法，其特征在于，所述模型评估还包括：

获取设定时间段内的新的监控指标数据，将所述预测模型的预测结果与所述新的监控指标数据进行比较，判断是否需要再次进行模型训练：如果大于设定阈值的时间消耗值在所述预测模型的预测结果以下则不再更新模型，反之则使用所述新的的监控指标数据更新所述预测模型的参数，对所述预测模型进行修正。

8、根据权利要求5所述的Kubernetes集群资源动态调整方法，其特征在于，在所述步骤c中，所述预测未来预设时间段内的资源消耗峰值还包括：

待所述 $LSTM_f$ 、 $LSTM_r$ 、 $LSTM_Q$ 三个子循环网络损失收敛或者降低到预设阈值，开始执行递推估计，通过度量收集器上报数据给预测模型，所述预测模型对未来预设时间段内的资源消耗峰值进行最优估计，使用最优估计值对Pod中容器资源进行更新操作。

9、一种电子设备，包括：

至少一个处理器；以及

与所述至少一个处理器通信连接的存储器；其中，

所述存储器存储有可被所述一个处理器执行的指令，所述指令被所述至少一个处理器执行，以使所述至少一个处理器能够执行上述1至8任一项所述的Kubernetes集群资源动态调整方法的以下操作：

步骤a：收集应用中的历史监控指标数据，根据所述监控指标数据得到资源消耗时序数据；

步骤b：构建LSTM-Kalman滤波器，将所述资源消耗时序数据输入所述LSTM-Kalman进行预训练，得到预测模型；其中，所述LSTM-Kalman滤波器为利用LSTM改造后的Kalman滤波器；

步骤c：收集新的资源消耗时序数据，将所述新的资源消耗时序数据输入所述预测模型中预测未来预设时间段内的资源消耗峰值，将所述资源消耗峰值作为资源限制值进行资源更新。

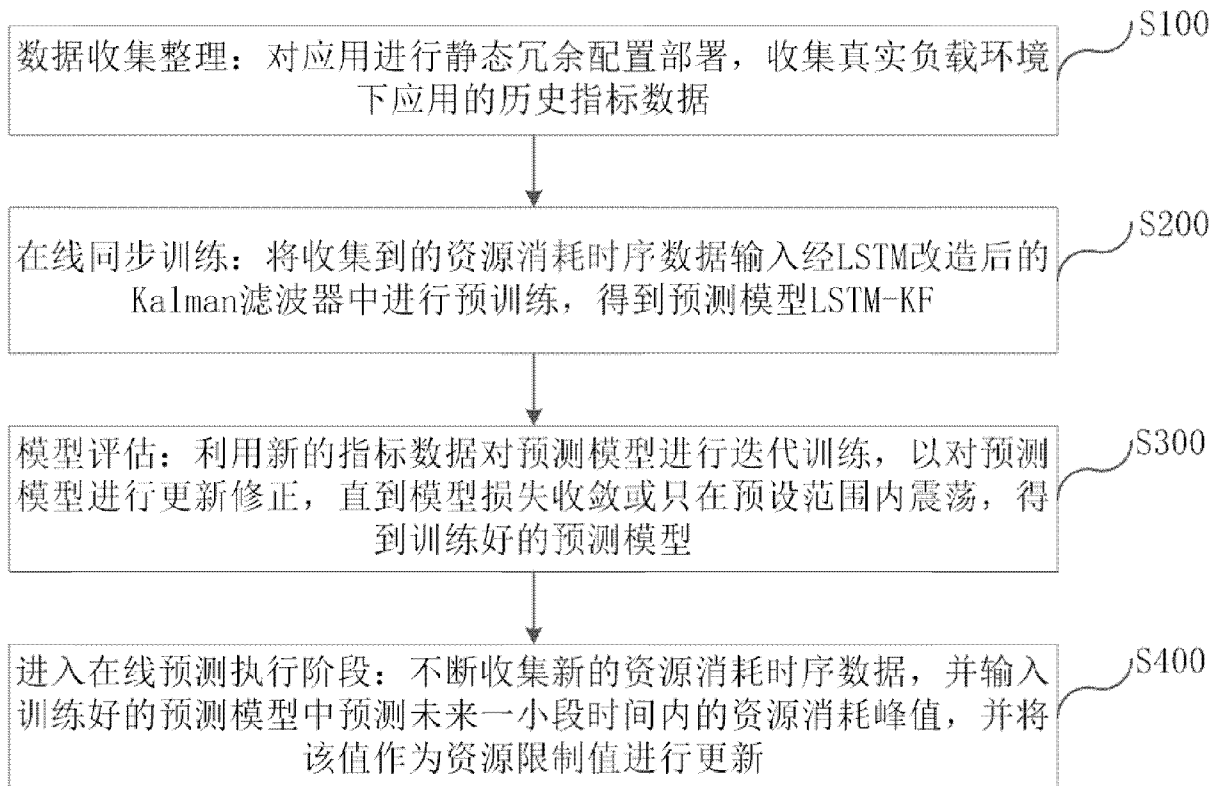


图1

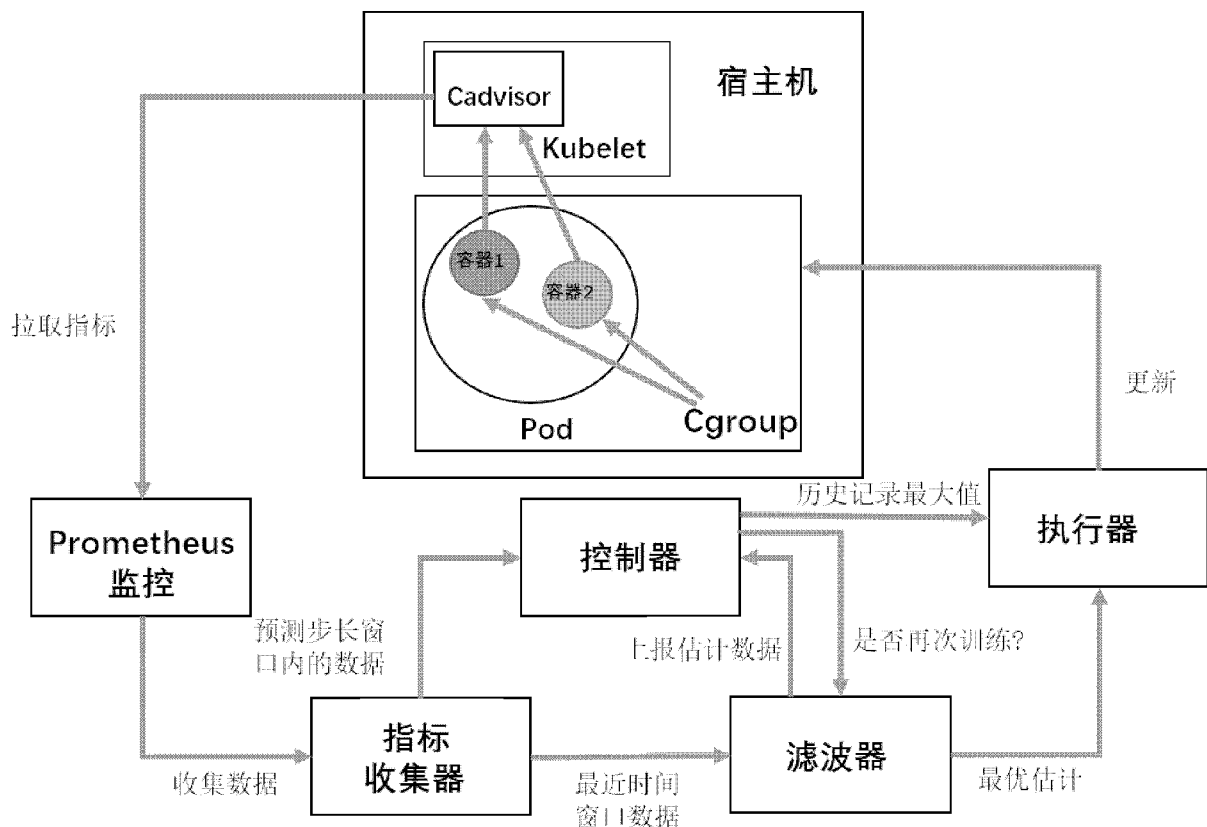


图2

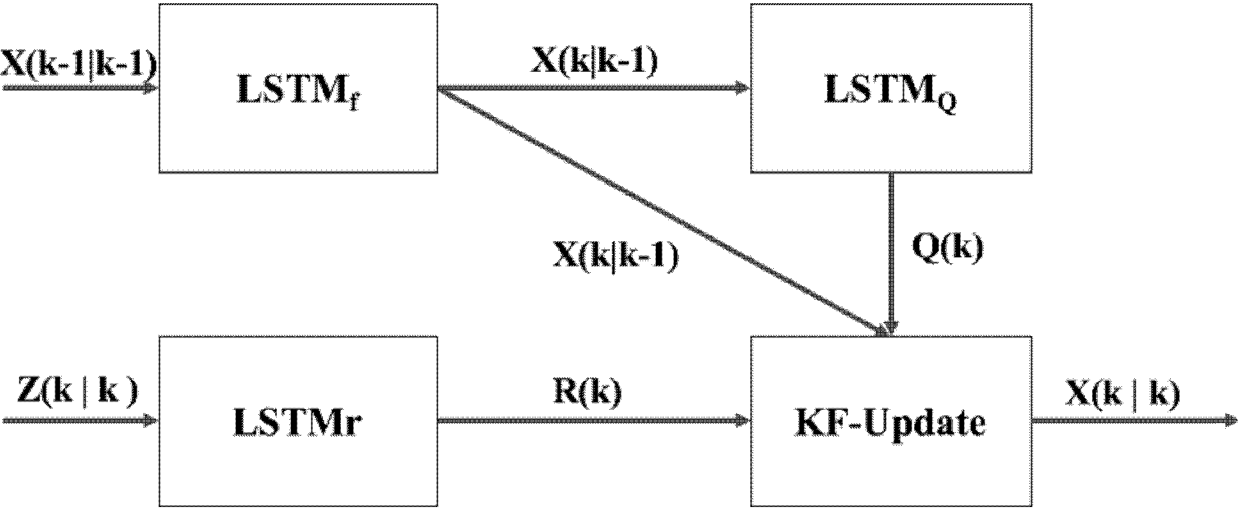


图3

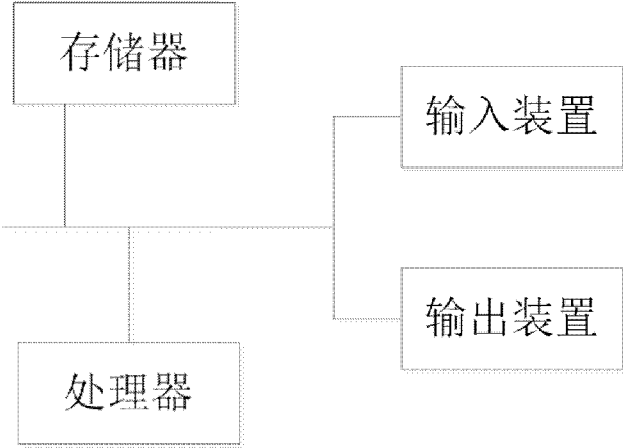


图4

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2020/140019

A. CLASSIFICATION OF SUBJECT MATTER

G06F 9/455(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)

G06F

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)

DWPI; EPTXT; USTXT; VEN; WOTXT; CNABS; CNTXT; CNKI: kubernetes, 容器, 资源, 长短期记忆, 卡尔曼, 预测, docker, container, resource, long short-term memory, LSTM, Kalman, predict, forecast

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
PY	CN 112069039 A (SUZHOU INSPUR INTELLIGENT TECHNOLOGY CO., LTD.) 11 December 2020 (2020-12-11) description, paragraphs 2-56	1-9
Y	田野 (TIAN, Ye). "面向容器的云计算资源自动伸缩问题研究 (Study on Auto Scaling Framework for Container in Cloud Computing)" 中国优秀硕士学位论文全文数据库 信息科技辑 (Chinese Master's Theses Full-text Database, Information Science and Technology), No. 01, 15 January 2020 (2020-01-15), ISSN: 1674-0246, full text, pages 16-37	1-9
Y	CN 108874542 A (GUILIN UNIVERSITY OF ELECTRONIC TECHNOLOGY) 23 November 2018 (2018-11-23) description, paragraphs 12-28, and figures 1-6	1-9
Y	CN 111124689 A (CHINA ACADEMIC OF ELECTRONICS AND INFORMATION TECHNOLOGY) 08 May 2020 (2020-05-08) description paragraphs 19-72	1-9
A	CN 108228347 A (SHANGHAI DIANJI UNIVERSITY) 29 June 2018 (2018-06-29) entire document	1-9



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

19 April 2021

Date of mailing of the international search report

27 April 2021

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

International application No.

PCT/CN2020/140019**C. DOCUMENTS CONSIDERED TO BE RELEVANT**

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	CN 110535894 A (SHENZHEN INSTITUTES OF ADVANCED TECHNOLOGY, CHINESE ACADEMY OF SCIENCES) 03 December 2019 (2019-12-03) entire document	1-9
A	CN 108920153 A (SOUTH CHINA UNIVERSITY OF TECHNOLOGY) 30 November 2018 (2018-11-30) entire document	1-9
A	US 2018349168 A1 (MAGALIX CORPORATION) 06 December 2018 (2018-12-06) entire document	1-9

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/CN2020/140019

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)	Publication date (day/month/year)
CN	112069039	A	11 December 2020	None	
CN	108874542	A	23 November 2018	None	
CN	111124689	A	08 May 2020	None	
CN	108228347	A	29 June 2018	None	
CN	110535894	A	03 December 2019	None	
CN	108920153	A	30 November 2018	None	
US	2018349168	A1	06 December 2018	US 10747568 B2	18 August 2020

国际检索报告

国际申请号

PCT/CN2020/140019

A. 主题的分类

G06F 9/455 (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)

G06F

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))

DWPI; EPTXT; USTXT; VEN; WOTXT; CNABS; CNTXT; CNKI: kubernetes, 容器, 资源, 长短期记忆, 卡尔曼, 预测, docker, container, resource, long short-term memory, LSTM, Kalman, predict, forecast

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
PY	CN 112069039 A (苏州浪潮智能科技有限公司) 2020年 12月 11日 (2020 - 12 - 11) 说明书第2-56段	1-9
Y	田野. “面向容器的云计算资源自动伸缩问题研究” 《中国优秀硕士学位论文全文数据库 信息科技辑》, 第01期, 2020年 1月 15日 (2020 - 01 - 15), ISSN: 1674-0246, 正文第16-37页	1-9
Y	CN 108874542 A (桂林电子科技大学) 2018年 11月 23日 (2018 - 11 - 23) 说明书第12-28段, 图1-6	1-9
Y	CN 111124689 A (中国电子科技集团公司信息科学研究院) 2020年 5月 8日 (2020 - 05 - 08) 说明书第19-72段	1-9
A	CN 108228347 A (上海电机学院) 2018年 6月 29日 (2018 - 06 - 29) 全文	1-9
A	CN 110535894 A (深圳先进技术研究院) 2019年 12月 3日 (2019 - 12 - 03) 全文	1-9

☒ 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

2021年 4月 19日

国际检索报告邮寄日期

2021年 4月 27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

柏娟花

电话号码 86-(20)-28950389

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 108920153 A (华南理工大学) 2018年 11月 30日 (2018 - 11 - 30) 全文	1-9
A	US 2018349168 A1 (MAGALIX CORPORATION) 2018年 12月 6日 (2018 - 12 - 06) 全文	1-9

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2020/140019

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	112069039	A	2020年 12月 11日	无	
CN	108874542	A	2018年 11月 23日	无	
CN	111124689	A	2020年 5月 8日	无	
CN	108228347	A	2018年 6月 29日	无	
CN	110535894	A	2019年 12月 3日	无	
CN	108920153	A	2018年 11月 30日	无	
US	2018349168	A1	2018年 12月 6日	US 10747568 B2	2020年 8月 18日