



(12) 发明专利

(10) 授权公告号 CN 114398400 B

(45) 授权公告日 2022.06.03

(21) 申请号 202210295299.2

(22) 申请日 2022.03.24

(65) 同一申请的已公布的文献号
申请公布号 CN 114398400 A

(43) 申请公布日 2022.04.26

(73) 专利权人 环球数科集团有限公司
地址 518063 广东省深圳市南山区粤海街
道高新南九道10号深圳湾科技生态园
10栋B座17层01-03号

(72) 发明人 张卫平 丁焯 隋银雪 米小武
张伟

(74) 专利代理机构 北京清控智云知识产权代理
事务所(特殊普通合伙)
11919

专利代理人 马肃

(51) Int.Cl.

G06F 16/2453 (2019.01)

G06F 11/30 (2006.01)

(56) 对比文件

CN 106209975 A, 2016.12.07

CN 110569060 A, 2019.12.13

CN 111339066 A, 2020.06.26

CN 113064879 A, 2021.07.02

US 2012089664 A1, 2012.04.12

US 2017286488 A1, 2017.10.05

US 2022043822 A1, 2022.02.10

史英杰等.《云数据管理系统中查询技术研究综述》.《计算机学报》.2013,第36卷(第2期),
孟令玺等.《云计算下的资源负载均衡性调度仿真》.《计算机仿真》.2018,第35卷(第4期),

审查员 罗晓彤

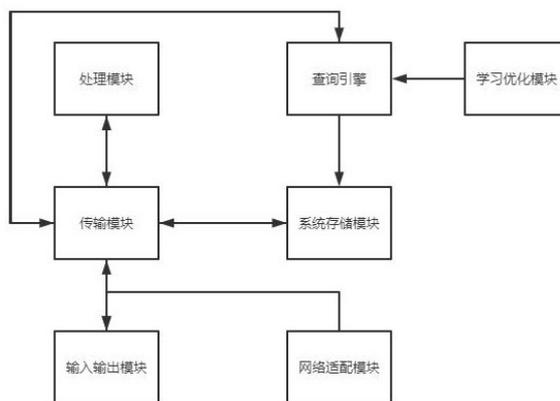
权利要求书2页 说明书7页 附图4页

(54) 发明名称

一种基于主动学习的Serverless资源池系统

(57) 摘要

本发明提供了一种基于主动学习的Serverless资源池系统,包括处理模块、系统存储模块、传输模块、输入输出模块、网络适配模块、查询引擎和学习优化模块,所述处理模块用于资源的执行,所述系统存储模块用于存储资源,所述传输模块用于资源与信息在系统内的传输,所述输入输出模块用于与外部设备通信,所述网络适配模块用于提供网络环境,所述查询引擎用于从所述系统存储模块中查询得到合适的资源,所述学习优化模块用于对查询引擎的查询配置进行优化;该系统能够对查询资源池的查询引擎进行主动学习,使得查询引擎的查询效果不断改进,提高用户的使用体验。



1. 一种基于主动学习的Serverless资源池系统,其特征在于,包括处理模块、系统存储模块、传输模块、输入输出模块、网络适配模块、查询引擎和学习优化模块,所述系统存储模块用于存储服务资源,所述处理模块用于执行服务资源,所述传输模块用于传输服务资源和信息,所述输入输出模块用于与外部设备交互,所述网络适配模块用于提供稳定网络服务,所述查询引擎用于查询服务资源,所述学习优化模块用于对所述查询引擎的查询进行优化;

所述查询引擎根据查询类别和配置参数值查询得到服务资源,在查询得到该服务资源后产生性能数据,所述查询类别、配置参数值和性能数据称为元数据,所述学习优化模块处理得到所述元数据的相关性,再依据相关性对配置参数值进行优化,优化后的配置参数值用于用户的服务查询;

所述学习优化模块计算得到 $P(B, j, i)$ 表示B类查询服务中第j个配置参数值与第i个性能数据的相关性,并建立优化目标函数 Q :

$$Q = \sum_{i=1}^n \alpha_i \cdot (A'(i) - \bar{A}_i)^2;$$

其中, α_i 为第i权重系数, \bar{A}_i 为第i标准数, $A'(i)$ 为第i优化目标, n为性能数据的种类数;

所述第i优化目标与配置参数值相关,并根据历史数据计算得到;

所述学习优化模块根据 $P(B, j, i)$ 值的分布计算出比例分布函数 $W(j)$, 满足下述等式:

$$\begin{cases} \frac{P(B,1, i)}{W(1)} = \frac{P(B,2, i)}{W(2)} = \dots = \frac{P(B, m, i)}{W(m)} \\ \sum_{j=1}^m W(j) = 1 \end{cases};$$

其中, m为配置参数值的种类数;

所述学习优化模块用 $w(i, j)$ 表示对第i类性能数据的分布函数,将分布函数按照下述公式累加得到综合分布函数 $U(j)$:

$$U(j) = \sum_{i=1}^n W(i, j), j \in [1, m];$$

所述学习优化模块根据 $U(j)$ 的分布对配置参数值进行调整,并选取使 Q 最小的一个方案为最优化方案。

2. 如权利要求1所述的一种基于主动学习的Serverless资源池系统,其特征在于,所述学习优化模块根据历史数据构建输入矩阵X和输出矩阵Y,将矩阵X和矩阵Y处理得到矩阵Z,矩阵Z中的元素 $z(j, i)$ 的计算公式为:

$$z(j, i) = \frac{1}{k} \cdot \sum_{k'=1}^k \left(\frac{a(j, k')}{c(k', i)} \right);$$

其中, $a(j, k')$ 为矩阵Y中的元素, $c(k', i)$ 为矩阵X中的元素, k 为历史数据中查询服务资源的次数。

3. 如权利要求2所述的一种基于主动学习的Serverless资源池系统, 其特征在于, 相关性 $P(B, j, i)$ 的计算公式为:

$$P(B, j, i) = \frac{(m-1) \cdot \left(\frac{A(i)}{C(j)} - z(j, i) \right)^2}{\sum_{j' \neq j} z^2(j', i)};$$

其中, $A(i)$ 表示当前查询中元数据的第 i 个性能数据值, $C(j)$ 表示当前查询中元数据的第 j 个配置参数值。

4. 如权利要求3所述的一种基于主动学习的Serverless资源池系统, 其特征在于, 所述配置参数值包括存储器大小、缓冲区大小、序列化选项、压缩参数值、网络参数值、调度特定值和执行选项值。

5. 如权利要求4所述的一种基于主动学习的Serverless资源池系统, 其特征在于, 所述查询引擎设有监控单元用于监控查询引擎的使用状态, 当查询引擎未被大量用于用户的服务查询时, 所述学习优化模块将使用查询引擎对配置参数值进行优化, 当查询引擎被大量使用于用户的服务查询时, 所述学习优化模块将处于休眠状态。

一种基于主动学习的Serverless资源池系统

技术领域

[0001] 本发明大体上涉及云服务系统,且更明确地说涉及一种基于主动学习的Serverless资源池系统。

背景技术

[0002] Serverless云服务是指通过互联网从云计算提供商的服务器按需提供给用户的任何服务,而不是从公司自己的本地服务器提供的服务,云服务旨在提供对应用程序、资源和服务的简单、可扩展的访问,并完全由云服务提供商管理,云存储是一种计算机数据存储模型,其中数字数据存储存储在逻辑池中,物理存储跨越多个服务器,物理环境通常由托管公司拥有和管理,这些云存储提供商负责保持数据的可用性和可访问性,以及保护和运行物理环境,随着用户数量的增多,需要对用户查询服务资源的过程进行改进。

[0003] 现在已经开发出了很多资源池系统,经过我们大量的检索与参考,发现现有的授权系统有如公开号为KR101640231B1, KR101765725B1、CN112256439B和KR1020160032881A所公开的系统,资源纳管模块,用于通过API接口来获取云计算资源池中各服务的服务参数;服务参数收敛模块,用于将所述服务参数划分成业务类参数和资源类参数;参数赋值模块,用于根据云计算资源池中各服务的使用规范对所述资源类参数进行赋值;服务目录生成模块,用于根据所述业务类参数和资源类参数生成服务目录。该该系统在用户获取资源时,不会对这个过程进行改进,导致用户数量增多后容易出现稳定性问题。

发明内容

[0004] 本发明的目的在于,针对所存在的不足,提出了一种基于主动学习的Serverless资源池系统。

[0005] 本发明采用如下技术方案:

[0006] 一种基于主动学习的Serverless资源池系统,包括处理模块、系统存储模块、传输模块、输入输出模块、网络适配模块、查询引擎和学习优化模块,所述系统存储模块用于存储服务资源,所述处理模块用于之执行服务资源,所述传输模块用于传输资源和信息,所述输入输出模块用于与外部设备交互,所述网络适配模块用于提供稳定网络服务,所述查询引擎用于查询服务资源,所述学习优化模块用于对所述查询引擎的查询进行优化;

[0007] 所述查询引擎根据查询类别和配置参数值查询得到服务资源,在查寻得到该服务资源后产生性能数据,所述查询类别、配置参数值和性能数据称为元数据,所述学习优化模块处理得到所述元数据的相关性,再依据相关性对配置参数值进行优化,优化后的配置参数值用于正式的服务查询;

[0008] 所述学习优化模块计算得到 $P(B, j, i)$ 表示B类查询服务中第j个配置参数值与第i个性能数据的相关性,并建立优化目标函数 Q :

$$[0009] \quad Q = \sum_{i=1}^n \alpha_i \cdot (A'(i) - \bar{A}_i)^2;$$

[0010] 其中, α_i 为第 i 权重系数, \bar{A}_i 为第 i 标准数, $A'(i)$ 为第 i 优化目标, n 为性能数

[0011] 据的种类数;

[0012] 所述第 i 优化目标与配置参数值相关, 并根据历史数据计算得到;

[0013] 所述学习优化模块根据 $P(B, j, i)$ 值的分布计算出比例分布函数 $W(j)$, 满足下述等式:

$$[0014] \quad \begin{cases} \frac{P(B,1, i)}{W(1)} = \frac{P(B,2, i)}{W(2)} = \dots = \frac{P(B, m, i)}{W(m)} \\ \sum_{j=1}^m W(j) = 1 \end{cases};$$

[0015] 其中, m 为配置参数值的种类数;

[0016] 所述学习优化模块用 $W(i, j)$ 表示对第 i 类性能数据的分布函数 $W(j)$, 将分布函数按照下述公式累加得到综合分布函数 $U(j)$:

$$[0017] \quad U(j) = \sum_{i=1}^n W(i, j), \quad j \in [1, m];$$

[0018] 所述第三确定单元根据 $U(j)$ 的分布对配置参数值进行调整, 并选取使 Q 最小的一个方案最为优化方案;

[0019] 进一步的, 所述学习优化模块根据历史数据构建输入矩阵 X 和输出矩阵 Y , 将矩阵 X 和矩阵 Y 处理得到矩阵 Z , 矩阵 Z 中的元素 $z(j, i)$ 的计算公式为:

$$[0020] \quad z(j, i) = \frac{1}{k} \cdot \sum_{k'=1}^k \left(\frac{a(j, k')}{c(k', i)} \right);$$

[0021] 其中, $a(j, k')$ 为矩阵 Y 中的元素, $c(k', i)$ 为矩阵 X 中的元素, k 为历史数据中查询的次数;

[0022] 进一步的, 相关性 $P(B, j, i)$ 的计算公式为:

$$[0023] \quad P(B, j, i) = \frac{(m-1) \cdot \left(\frac{A(i)}{C(j)} - z(j, i) \right)^2}{\sum_{j' \neq j} z^2(j', i)};$$

[0024] 其中, $A(i)$ 表示当前查询中元数据的第 i 个性能数据值, $C(j)$ 表示当前查询中元数据的第 j 个配置参数值;

[0025] 进一步的, 所述配置参数值包括存储器大小、缓冲区大小、序列化选项、压缩参数值、网络参数值、调度特定值和执行选项值;

[0026] 进一步的, 所述查询引擎设有监控单元用于监控查询引擎的使用状态, 当查询引擎未被大量用于用户的服务查询时, 所述学习优化模块将使用查询引擎对配置参数值进行

优化,当查询引擎被大量使用于用户的服务查询时,所述学习优化模块将处于休眠状态。

[0027] 本发明所取得的有益效果是:

[0028] 本系统在查询引擎空闲时通过学习优化模块进行主动学习,改良查询引擎的查询效果,使用户能够更快的获取更适合的服务资源,在优化过程中,将配置参数值作为输入数据,性能数据作为输出数据,将历史查询过程中的输入数据和输出数据建立相关性联系,并根据相关性对配置参数值进行不断优化,从而改善查询引擎的使用效果。

[0029] 为能使更进一步了解本发明的特征及技术内容,请参阅以下有关本发明的详细说明与附图,然而所提供的附图仅用于提供参考与说明,并非用来对本发明加以限制。

附图说明

[0030] 图1为本发明整体结构框架示意图;

[0031] 图2为本发明学习优化模块进行优化的流程示意图;

[0032] 图3为本发明服务查询进行优化的框架示意图;

[0033] 图4为本发明系统存储模块的构成示意图;

[0034] 图5为本发明学习优化模块的构成示意图。

具体实施方式

[0035] 以下是通过特定的具体实施例来说明本发明的实施方式,本领域技术人员可由本说明书所公开的内容了解本发明的优点与效果。本发明可通过其他不同的具体实施例加以施行或应用,本说明书中的各项细节也可基于不同观点与应用,在不悖离本发明的精神下进行各种修饰与变更。另外,本发明的附图仅为简单示意说明,并非依实际尺寸的描绘,事先声明。以下的实施方式将进一步详细说明本发明的相关技术内容,但所公开的内容并非用以限制本发明的保护范围。

[0036] 实施例一。

[0037] 本实施例提供了一种基于主动学习的Serverless资源池系统,结合图1,包括处理模块、系统存储模块、传输模块、输入输出模块、网络适配模块、查询引擎和学习优化模块,所述系统存储模块用于存储服务资源,所述处理模块用于之执行服务资源,所述传输模块用于传输资源和信息,所述输入输出模块用于与外部设备交互,所述网络适配模块用于提供稳定网络服务,所述查询引擎用于查询服务资源,所述学习优化模块用于对所述查询引擎的查询进行优化;

[0038] 所述查询引擎根据查询类别和配置参数值查询得到服务资源,在查寻得到该服务资源后产生性能数据,所述查询类别、配置参数值和性能数据称为元数据,所述学习优化模块处理得到所述元数据的相关性,再依据相关性对配置参数值进行优化,优化后的配置参数值用于正式的服务查询;

[0039] 所述学习优化模块计算得到 $P(B, j, i)$ 表示B类查询服务中第j个配置参数值与第i个性能数据的相关性,并建立优化目标函数Q:

$$[0040] \quad Q = \sum_{i=1}^n \alpha_i \cdot (A'(i) - \bar{A}_i)^2;$$

[0041] 其中, α_i 为第i权重系数, \bar{A}_i 为第i标准数, $A'(i)$ 为第i优化目标, n为性能数

[0042] 据的种类数;

[0043] 所述第i优化目标与配置参数值相关, 并根据历史数据计算得到;

[0044] 所述学习优化模块根据 $P(B, j, i)$ 值的分布计算出比例分布函数 $W(j)$, 满足下述等式:

$$[0045] \quad \begin{cases} \frac{P(B,1, i)}{W(1)} = \frac{P(B,2, i)}{W(2)} = \dots = \frac{P(B, m, i)}{W(m)} \\ \sum_{j=1}^m W(j) = 1 \end{cases};$$

[0046] 其中, m为配置参数值的种类数;

[0047] 所述学习优化模块用 $W(i, j)$ 表示对第i类性能数据的分布函数 $W(j)$, 将分布函数按照下述公式累加得到综合分布函数 $U(j)$:

$$[0048] \quad U(j) = \sum_{i=1}^n W(i, j), \quad j \in [1, m];$$

[0049] 所述第三确定单元根据 $U(j)$ 的分布对配置参数值进行调整, 并选取使 Q 最小的一个方案最为优化方案;

[0050] 所述学习优化模块根据历史数据构建输入矩阵 X 和输出矩阵 Y , 将矩阵 X 和矩阵 Y 处理得到矩阵 Z , 矩阵 Z 中的元素 $z(j, i)$ 的计算公式为:

$$[0051] \quad z(j, i) = \frac{1}{k} \cdot \sum_{k'=1}^k \left(\frac{a(j, k')}{c(k', i)} \right);$$

[0052] 其中, $a(j, k')$ 为矩阵 Y 中的元素, $c(k', i)$ 为矩阵 X 中的元素, k为历史数据中查询的次数;

[0053] 相关性 $P(B, j, i)$ 的计算公式为:

$$[0054] \quad P(B, j, i) = \frac{(m-1) \cdot \left(\frac{A(i)}{C(j)} - z(j, i) \right)^2}{\sum_{j' \neq j} z^2(j', i)};$$

[0055] 其中, $A(i)$ 表示当前查询中元数据的第i个性能数据值, $C(j)$ 表示当前查询中元数据的第j个配置参数值;

[0056] 所述配置参数值包括存储器大小、缓冲区大小、序列化选项、压缩参数值、网络参数值、调度特定值和执行选项值;

[0057] 所述查询引擎设有监控单元用于监控查询引擎的使用状态, 当查询引擎未被大量用于用户的服务查询时, 所述学习优化模块将使用查询引擎对配置参数值进行优化, 当查询引擎被大量使用于用户的服务查询时, 所述学习优化模块将处于休眠状态。

[0058] 实施例二。

[0059] 本实施例包含了实施例一的全部内容,提供了一种基于主动学习的Serverless资源池系统,包括处理模块、系统存储模块、传输模块、输入输出模块、网络适配模块、查询引擎和学习优化模块,所述处理模块用于资源的执行,所述系统存储模块用于存储资源,所述传输模块用于资源与信息在系统内的传输,所述输入输出模块用于与外部设备通信,所述网络适配模块用于提供网络环境,所述查询引擎用于从所述系统存储模块中查询得到合适的资源,所述学习优化模块用于对查询引擎的查询配置进行优化;

[0060] 结合图4,所述系统存储模块包括随机存取存储器、高速缓冲存储器、驱动器和程序单元,所述随机存取存储器和所述高速缓冲存储器用于存储本系统中的数据,所述驱动器用于驱动数据在所述随机存取存储器和高速缓冲存储器中读取或写入,所述程序单元包括若干个被配置为执行特定功能的代码封装,属于供用户查询调用的资源;

[0061] 结合图5,所述学习优化模块包括第一确定单元、第二确定单元、第三确定单元、发送器、记录器和应用单元,所述第一确定单元用于确定需要查询的服务类别,所述发送器用于将查询指令发送至查询引擎,所述记录器用于记录在查询引擎上执行的服务查询的元数据,所述元数据包括性能数据、查询类别和配置参数值,所述第二确定单元用于确定所述元数据之间的相关性,所述第三确定单元根据相关性确定最优配置参数值,所述应用单元用于将所述最优配置参数值应用于查询引擎;

[0062] 所述学习优化模块的上述各单元通过所述传输模块进行数据通信;

[0063] 结合图2,所述学习优化模块进行服务查询的优化过程包括如下步骤:

[0064] S1、所述第一确定单元确定需要查询的服务的类别;

[0065] S2、所述发送器将该服务类别的服务查询发送至所述查询引擎;

[0066] S3、所述查询引擎根据接收到的服务查询的元数据进行查询;

[0067] S4、所述第二确定单元对元数据的相关性进行计算确定;

[0068] S5、所述第三确定单元根据相关性对配置参数值进行优化调整;

[0069] S6、所述应用单元将优化调整后的配置参数值应用于查询引擎上的同一类别的扩展服务查询集以验证配置;

[0070] S7、当得到验证的肯定结果时,使用优化后的配置参数值来执行相同类别的服务查询;

[0071] 结合图3,一组同类别的服务查询被发送至查询引擎,该服务查询与配置参数值相关,查询执行的元数据和配置参数值被一起记录在存储区域中并进行关联,激活单元将元数据激活来确定查询类别、选择的配置参数值和结果性能数据,并且基于相关性分析得到配置参数值的优化集合,然后将该优化集合和同一类别下的其他一组服务查询相关联,并发送至查询引擎进行循环优化,当同一类别的服务查询均被优化处理后,将最后得到的优化集合输出成为正式配置;

[0072] 所述查询引擎设有监控单元用于监控查询引擎的使用状态,当查询引擎未被大量用于用户的服务查询时,所述学习优化模块将使用查询引擎对配置参数值进行优化,当查询引擎被大量使用于用户的服务查询时,所述学习优化模块将处于休眠状态;

[0073] 在步骤S4中,对元数据进行相关性计算的过程如下:

[0074] 将此次查询中元数据的性能数据用 $A(i)$ 表示, $i = \{1, 2, 3, \dots, n\}$, n 为性能的种类数,

查询类别用B表示,配置参数值用C(j)表示, $j = \{1, 2, 3, \dots, m\}$, m为配置参数值的种类数;

[0075] 所述第二确定单元根据历史数据以及此次的查询数据计算出 $P(B, j, i)$ 用于表示在B类服务查询类别中第j个配置参数值与第i个性能的相关度;

[0076] 所述第二确定单元根据历史数据构建输入矩阵X:

$$[0077] \quad X = \begin{Bmatrix} c(1,1), c(1,2), \dots, c(1, k) \\ c(2,1), c(2,2), \dots, c(2, k) \\ \dots & \dots & \dots \\ c(m,1), c(m,2), \dots, c(m, k) \end{Bmatrix};$$

[0078] 其中, $c(i, j)$ 表示B类查询服务第j次历史数据中第i个配置参数值的数据值, 而k为历史查询的次数;

[0079] 所述第二确定单元根据历史数据构建输出矩阵Y:

$$[0080] \quad Y = \begin{Bmatrix} a(1,1), a(1,2), \dots, a(1, n) \\ a(2,1), a(2,2), \dots, a(2, n) \\ \dots & \dots & \dots \\ a(k,1), a(k,2), \dots, a(k, n) \end{Bmatrix};$$

[0081] 其中, $a(j, i)$ 表示B类查询服务第j次历史数据中第i个性能的数据值;

[0082] 所述第二确定单元将输入矩阵与输出矩阵处理得到相关矩阵Z:

$$Z = \begin{Bmatrix} z(1,1), z(1,2), \dots, z(1, n) \\ z(2,1), z(2,2), \dots, z(2, n) \\ \dots & \dots & \dots \\ z(m,1), z(m,2), \dots, z(m, n) \end{Bmatrix};$$

[0083] 其中, 元素 $z(j, i)$ 为:

$$[0084] \quad z(j, i) = \frac{1}{k} \cdot \sum_{k'=1}^k \left(\frac{a(j, k')}{c(k', i)} \right);$$

[0085] 所述第二确定单元根据下式计算出 $P(B, j, i)$:

$$[0086] \quad P(B, j, i) = \frac{(m-1) \cdot \left(\frac{A(i)}{C(j)} - z(j, i) \right)^2}{\sum_{j'=j} z^2(j', i)};$$

[0087] 在步骤S5中, 所述第三确定单元建立优化目标函数Q:

$$[0088] \quad Q = \sum_{i=1}^n \alpha_i \cdot (A'(i) - \bar{A}_i)^2;$$

[0089] 其中, α_i 为第i权重系数, \bar{A}_i 为第i标准数, $A'(i)$ 为第i优化目标;

[0090] α_i 和 \bar{A}_i 由工作人员根据服务查询的类别人为设置;

[0091] 所述第三确定单元根据所述第二确定单元计算得到的相关性数据对配置参数值

C(i) 进行调整使得预期得到的优化目标函数值最小；

[0092] 所述第三确定单元根据 $P(B, j, i)$ 值的分布计算出比例分布函数 $w(j)$ ，满足下述等式：

$$[0093] \quad \begin{cases} \frac{P(B,1, i)}{W(1)} = \frac{P(B,2, i)}{W(2)} = \dots = \frac{P(B, m, i)}{W(m)} \\ \sum_{j=1}^m W(j) = 1 \end{cases} ;$$

[0094] 对于一类性能数据能够得到一组分布函数 $w(j)$ ，而对于 n 类性能数据能够得到 n 组分布函数 $w(j)$ ，用 $W(i, j)$ 表示对第 i 类性能数据的分布函数 $w(j)$ ，所述第三确定单元将分布函数按照下述公式累加得到综合分布函数 $U(j)$ ：

$$[0095] \quad U(j) = \sum_{i=1}^n W(i, j) \quad j \in [1, m];$$

[0096] 所述第三确定单元根据 $U(j)$ 的分布对配置参数值 $C(i)$ 进行调整，并选取使 Q 最小的一个方案最为优化方案。

[0097] 以上所公开的内容仅为本发明的优选可行实施例，并非因此局限本发明的保护范围，所以凡是运用本发明说明书及附图内容所做的等效技术变化，均包含于本发明的保护范围内，此外，随着技术发展其中的元素可以更新的。

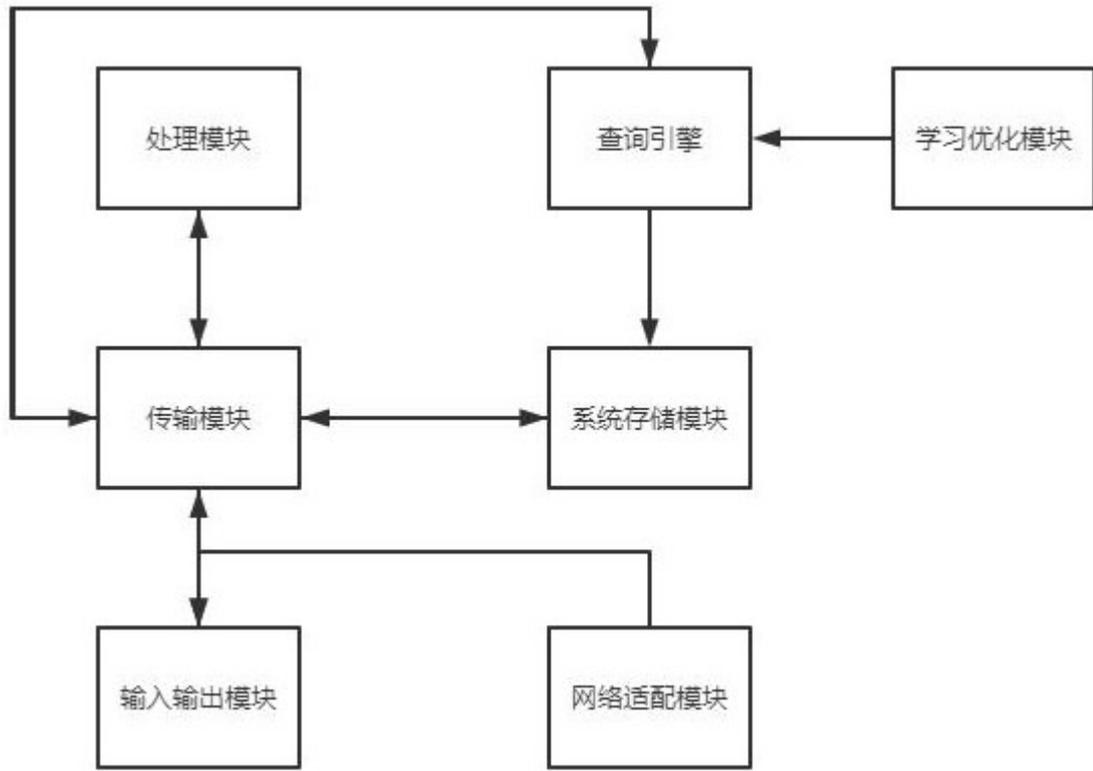


图1

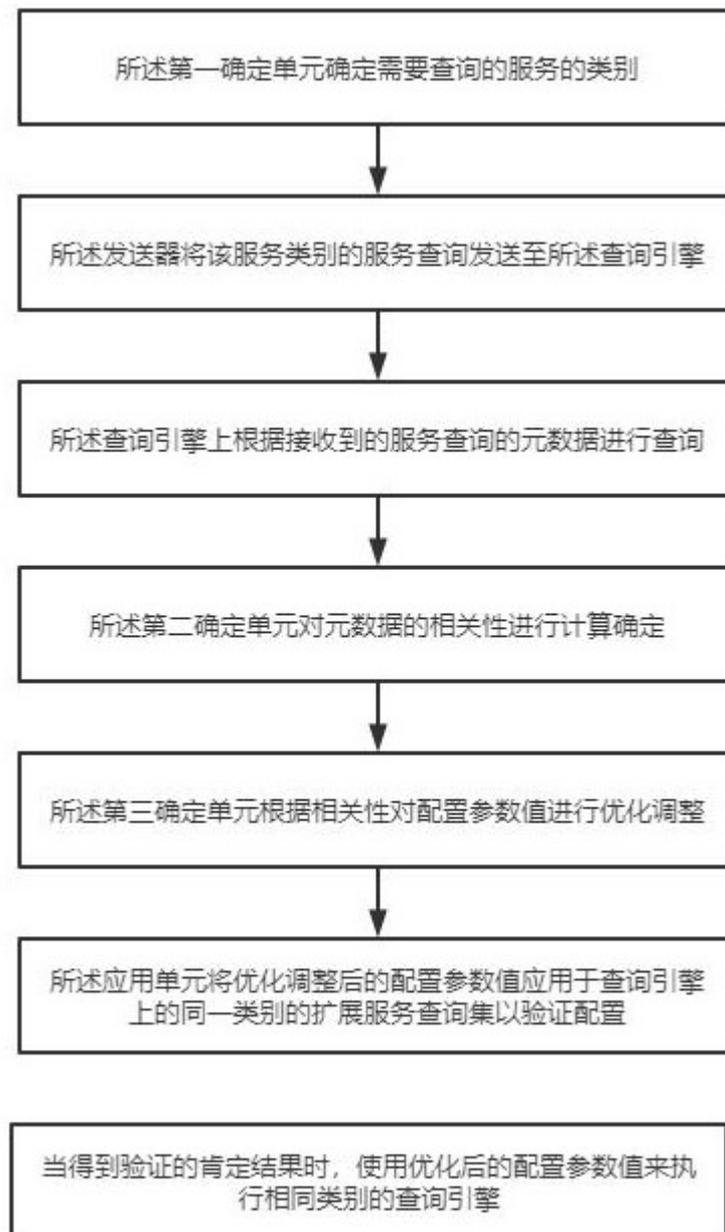


图2

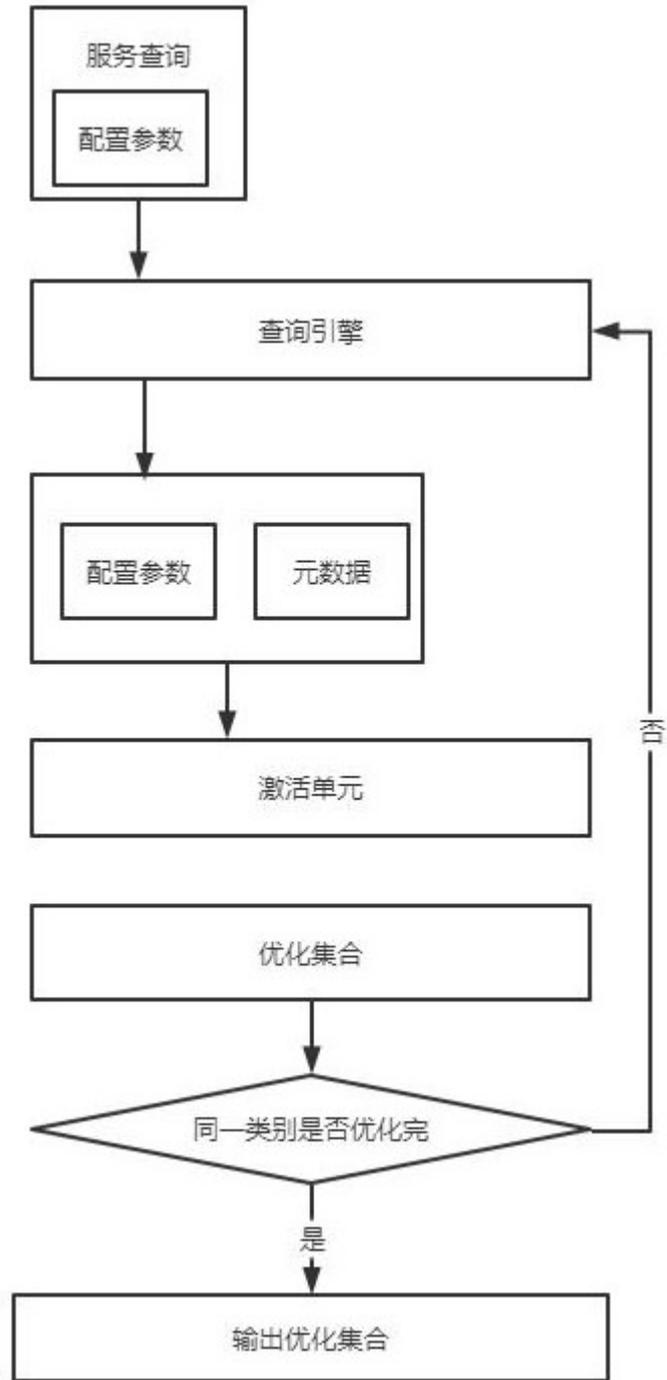


图3

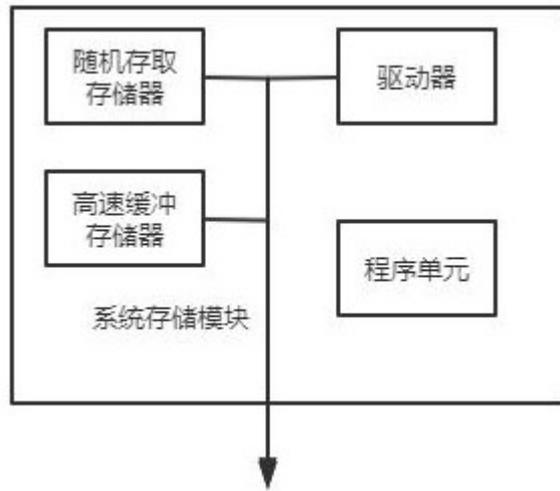


图4

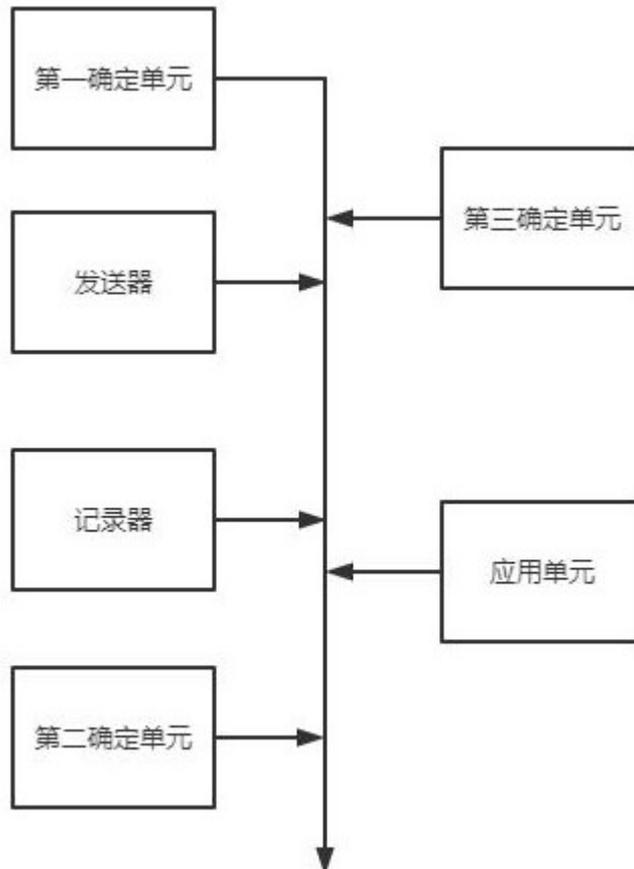


图5