

(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
2023 年 10 月 5 日 (05.10.2023)



(10) 国际公布号
WO 2023/185071 A1

(51) 国际专利分类号:
G06F 16/21 (2019.01)

(21) 国际申请号: PCT/CN2022/136753

(22) 国际申请日: 2022 年 12 月 6 日 (06.12.2022)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:
202210331107.9 2022年3月31日 (31.03.2022) CN

(71) 申请人: 北京沃东天骏信息技术有限公司 (BEIJING WODONG TIANJUN INFORMATION TECHNOLOGY CO., LTD.) [CN/CN]; 中国北京市北京经济技术开发区科创十一街18号院2号楼4层A402室, Beijing 100176 (CN)。北京京东世纪贸易有限公司 (BEIJING JINGDONG

CENTURY TRADING CO., LTD.) [CN/CN]; 中国北京市北京经济技术开发区科创十一街18号院C座2层201室, Beijing 100176 (CN)。

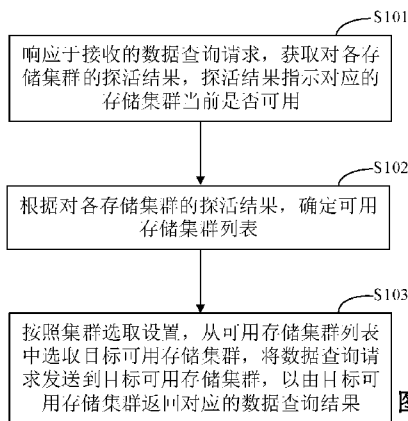
(72) 发明人: 张进 (ZHANG, Jin); 中国北京市北京经济技术开发区科创十一街18号院2号楼4层A402室, Beijing 100176 (CN)。

(74) 代理人: 中原信达知识产权代理有限责任公司 (CHINA SINDA INTELLECTUAL PROPERTY LIMITED); 中国北京市西城区金融街19号富凯大厦B座11层, Beijing 100033 (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, CV, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IQ,

(54) Title: DATA QUERY METHOD, DATA WRITING METHOD, RELATED APPARATUS AND SYSTEM

(54) 发明名称: 数据查询方法、数据写入方法、相关装置和系统



S101 In response to a received data query request, obtain a liveness probe result for each storage cluster, the liveness probe result indicating whether a corresponding storage cluster is currently available
S102 According to the liveness probe result for each storage cluster, determine an available storage cluster list
S103 Select a target available storage cluster from the available storage cluster list according to a cluster selection setting, and send the data query request to the target available storage cluster, so that the target available storage cluster returns a corresponding data query result

图 1

(57) Abstract: Disclosed are a data query method, a data writing method, a related apparatus and a system. A specific embodiment comprises: in response to a received data query request, obtaining a liveness probe result for each storage cluster, the liveness probe result indicating whether a corresponding storage cluster is currently available; according to the liveness probe result for each storage cluster, determining an available storage cluster list; and selecting a target available storage cluster from the available storage cluster list according to a cluster selection setting, and sending the data query request to the target available storage cluster, so that the target available storage cluster returns a corresponding data query result.

(57) 摘要: 本公开公开了一种数据查询方法、数据写入方法、相关装置和系统, 具体实施方式包括: 响应于接收的数据查询请求, 获取对各存储集群的探活结果, 探活结果指示对应的存储集群当前是否可用; 根据对各存储集群的探活结果, 确定可用存储集群列表; 按照集群选取设置, 从可用存储集群列表选取目标可用存储集群, 将数据查询请求发送到目标可用存储集群, 以由目标可用存储集群返回对应的数据查询结果。

IR, IS, IT, JM, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ,
LA, LC, LK, LR, LS, LU, LY, MA, MD, 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, CV, GH, GM, KE, LR, LS, MW,
MZ, NA, RW, SC, 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, ME, 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))。

数据查询方法、数据写入方法、相关装置和系统

相关申请的交叉引用

5 本申请要求享有 2022 年 3 月 31 日提交的名称为“一种数据查询方法、数据写入方法、相关装置和系统”的中国专利申请 No. 202210331107.9 的优先权，在此全文引用上述中国专利申请公开的内容以作为本申请的一部分或全部。

10 技术领域

 本公开涉及计算机技术领域，尤其涉及一种数据查询方法和装置、数据写入方法和装置、以及数据查询系统。

背景技术

15 目前集群部署有两种方案：方案一为基于 ElasticSearch（一种全文检索框架，是一个 java 语言编写的搜索引擎）的集群部署，一个集群局限于同一个机房；方案二为跨机房部署集群，向集群写入数据，或者读取数据时，集群内部需要进行通信。

20 在实现本公开过程中，发明人发现现有技术中至少存在如下问题：
 方案一无法实现机房之间的容灾部署，一旦出现机房级别的故障，将会导致整个集群不可用，最终导致服务不可用；方案二集群内通信的效率降低，在向集群写入数据或者读取数据时，由于集群内部需要进行通信，耗费时间大。

25

发明内容

 有鉴于此，本公开实施例提供一种数据查询方法、数据写入方法、相关装置和系统，能够实现向多个集群写入数据的一致性，可自动切换集群，提高集群部署的容灾能力，解决了在机房级别故障时导致整个集群不可用的问题，同时克服了由于集群内部需要进行通信而耗费

30

时间大的缺陷。

为实现上述目的，根据本公开实施例的一个方面，提供了一种数据查询方法。

5

一种数据查询方法，包括：响应于接收的数据查询请求，获取对各存储集群的探活结果，所述探活结果指示对应的存储集群当前是否可用；根据对各存储集群的探活结果，确定可用存储集群列表；按照集群选取设置，从所述可用存储集群列表中选取目标可用存储集群，
10 将所述数据查询请求发送到所述目标可用存储集群，以由所述目标可用存储集群返回对应的数据查询结果。

在本公开的一个或多个实施例中，所述获取对各存储集群的探活结果，包括：从配置中心或本地缓存中获取探活节点集群对所述各存储集群的探活结果，其中，所述配置中心通过动态监听机制，在监听到所述探活节点集群对所述存储集群的探活结果后，将对应的所述探活结果缓存到所述本地缓存。
15

在本公开的一个或多个实施例中，所述探活节点集群包括一个或多个探活节点，所述探活节点用于向所述各存储集群分别发送健康度请求，以根据所述存储集群对所述健康度请求的响应情况，确定对应的所述存储集群是否可用。
20

在本公开的一个或多个实施例中，所述根据对各存储集群的探活结果，确定可用存储集群列表，包括：对于每一存储集群，判断各所述探活节点对该存储集群的探活结果中，指示该存储集群可用的探活节点数量占探活节点总数的比例是否超过比例阈值，若是，则确定该存储集群可用，并添加到所述可用存储集群列表。
25

在本公开的一个或多个实施例中，所述按照集群选取设置，从所
30

述可用存储集群列表中选取目标可用存储集群，将所述数据查询请求发送到所述目标可用存储集群，以由所述目标可用存储集群返回对应的数据查询结果，包括：以所述可用存储集群列表中的所有可用存储集群作为所述目标可用存储集群，在收到其中一个目标可用存储集群返回的数据查询结果后，输出所述数据查询结果。

在本公开的一个或多个实施例中，所述按照集群选取设置，从所述可用存储集群列表选取目标可用存储集群，将所述数据查询请求发送到所述目标可用存储集群，以由所述目标可用存储集群返回对应的数据查询结果，包括：按照轮询策略从所述可用存储集群列表选出当前轮询到的一个可用存储集群，作为所述目标可用存储集群，在收到所述目标可用存储集群返回的数据查询结果后，输出所述数据查询结果。

在本公开的一个或多个实施例中，所述按照集群选取设置，从所述可用存储集群列表选取目标可用存储集群，将所述数据查询请求发送到所述目标可用存储集群，以由所述目标可用存储集群返回对应的数据查询结果，包括：从所述可用存储集群列表中随机选出一个可用存储集群，作为所述目标可用存储集群，在收到所述目标可用存储集群返回的数据查询结果后，输出所述数据查询结果。

根据本公开实施例的另一方面，提供了一种数据写入方法。

一种数据写入方法，包括：根据存储集群的数量，发起对消息中间件主题的多个监听，每一监听对应一个存储集群，所述消息中间件主题包括待写入数据；响应于对所述消息中间件主题的每一监听，将所述待写入数据写入对应的存储集群。

根据本公开实施例的又一方面，提供了一种数据查询装置。

一种数据查询装置，包括：探活结果获取模块，用于响应于接收的数据查询请求，获取对各存储集群的探活结果，所述探活结果指示对应的存储集群当前是否可用；可用存储集群确定模块，用于根据对各存储集群的探活结果，确定可用存储集群列表；数据查询模块，用于按照集群选取设置，从所述可用存储集群列表中选取目标可用存储集群，将所述数据查询请求发送到所述目标可用存储集群，以由所述目标可用存储集群返回对应的数据查询结果。

根据本公开实施例的又一方面，提供了一种数据写入装置。

一种数据写入装置，包括：监听模块，用于根据存储集群的数量，发起对消息中间件主题的多个监听，每一监听对应一个存储集群，所述消息中间件主题包括待写入数据；数据写入模块，用于响应于对所述消息中间件主题的每一监听，将所述待写入数据写入对应的存储集群。

根据本公开实施例的又一方面，提供了一种数据查询系统。

一种数据查询系统，包括：本公开实施例提供的数据查询装置，还包括各存储集群、配置中心和探活节点集群，其中：所述探活节点集群包括一个或多个探活节点，所述探活节点用于向所述各存储集群分别发送健康度请求，以根据所述存储集群对所述健康度请求的响应情况，确定所述存储集群的探活结果，所述探活结果指示对应的所述存储集群当前是否可用；所述配置中心用于接收各所述探活节点上报的所述探活结果，并向所述数据查询装置提供所述探活结果。

根据本公开实施例的又一方面，提供了一种电子设备。

一种电子设备，包括：一个或多个处理器；存储器，用于存储一个或多个程序，当所述一个或多个程序被所述一个或多个处理器执行

时，使得所述一个或多个处理器实现本公开实施例所提供的数据查询方法或数据写入方法。

根据本公开实施例的又一方面，提供了一种计算机可读介质。

5

一种计算机可读介质，其上存储有计算机程序，所述计算机程序被处理器执行时实现本公开实施例所提供的数据查询方法或数据写入方法。

10

上述的非惯用的可选方式所具有的进一步效果将在下文中结合具体实施方式加以说明。

附图说明

附图用于更好地理解本公开，不构成对本公开的不当限定。其中：

15

图 1 是根据本公开一个或多个实施例中的数据查询方法的主要步骤示意图；

图 2 是根据本公开一个或多个实施例中的数据写入方法的主要步骤示意图；

20

图 3 是根据本公开一个或多个实施例中的数据查询和数据写入的流程示意图；

图 4 是根据本公开一个或多个实施例中的数据查询装置的主要模块示意图；

图 5 是根据本公开一个或多个实施例中的数据写入装置的主要模块示意图；

25

图 6 是根据本公开一个或多个实施例中的数据查询系统的主要架构示意图；

图 7 是本公开一个或多个实施例中可以应用于其中的示例性系统架构图；

30

图 8 是适于用来实现本公开一个或多个实施例中的终端设备或服务器的计算机系统的结构示意图。

具体实施方式

以下结合附图对本公开的示范性实施例做出说明，其中包括本公开实施例的各种细节以助于理解，应当将它们认为仅仅是示范性的。

5 因此，本领域普通技术人员应当认识到，可以对这里描述的实施例做出各种改变和修改，而不会背离本公开的范围和精神。同样，为了清楚和简明，以下的描述中省略了对公知功能和结构的描述。

10 图 1 是根据本公开一个或多个实施例中的数据查询方法的主要步骤示意图。

如图 1 所示，本公开一个或多个实施例中的数据查询方法主要包括如下的步骤 S101 至步骤 S103。

15 步骤 S101：响应于接收的数据查询请求，获取对各存储集群的探活结果，探活结果指示对应的存储集群当前是否可用。

20 获取对各存储集群的探活结果，可以包括：从配置中心或本地缓存中获取探活节点集群对各存储集群的探活结果，其中，配置中心通过动态监听机制，在监听到探活节点集群对存储集群的探活结果后，将对应的探活结果缓存到本地缓存。

25 探活节点集群可以包括一个或多个探活节点，探活节点用于向各存储集群分别发送健康度请求，以根据存储集群对健康度请求的响应情况，确定对应的存储集群是否可用。本公开存储集群的数量为多个，存储集群例如 ES（ElasticSearch）集群，探活节点向各存储集群分别发送健康度请求时，可以利用 ES 集群提供的专用于获取健康度的命令去获取集群的健康度（cluster health）。

30 步骤 S102：根据对各存储集群的探活结果，确定可用存储集群列

表。

根据对各存储集群的探活结果，确定可用存储集群列表，可以包括：对于每一存储集群，判断各探活节点对该存储集群的探活结果中，指示该存储集群可用的探活节点数量占探活节点总数的比例是否超过比例阈值，若是，则确定该存储集群可用，并添加到可用存储集群列表。探活结果可包括 `timeout` 字段，进行健康度请求时，需要规定一个存储集群响应的的时间，比如 1s（秒），存储集群在 1s 之内返回了结果，就认为该存储集群是存活的，探活结果中的 `timeout=false` 表示存储集群存活，如果超时，认为该存储集群已经故障。对于多个探活节点，比例阈值可根据需要设定，例如，如果半数（比例阈值=50%）以上探活节点返回的对应某个存储集群的探活结果都为该存储集群存活，即指示该存储集群可用，则根据各探活结果确定出该存储集群可用。

步骤 S103：按照集群选取设置，从可用存储集群列表中选取目标可用存储集群，将数据查询请求发送到目标可用存储集群，以由目标可用存储集群返回对应的数据查询结果。

在一个实施例中，按照集群选取设置，从可用存储集群列表中选取目标可用存储集群，将数据查询请求发送到目标可用存储集群，以由目标可用存储集群返回对应的数据查询结果，可以包括：以可用存储集群列表中的所有可用存储集群作为目标可用存储集群，在收到其中一个目标可用存储集群返回的数据查询结果后，输出数据查询结果。

在另一个实施例中，按照集群选取设置，从可用存储集群列表中选取目标可用存储集群，将数据查询请求发送到目标可用存储集群，以由目标可用存储集群返回对应的数据查询结果，可以包括：按照轮询策略从可用存储集群列表中选出当前轮询到的一个可用存储集群，作为目标可用存储集群，在收到目标可用存储集群返回的数据查询结果后，输出数据查询结果。

在又一个实施例中，按照集群选取设置，从可用存储集群列表
中选取目标可用存储集群，将数据查询请求发送到目标可用存储集群，
以由目标可用存储集群返回对应的数据查询结果，可以包括：从可用
存储集群列表中随机选出一个可用存储集群，作为目标可用存储集群，
在收到目标可用存储集群返回的数据查询结果后，输出数据查询结果。

图 2 是根据本公开一个或多个实施例中的数据写入方法的主要步
骤示意图。

如图 2 所示，本公开一个实施例的数据写入方法主要包括如下的
步骤 S201 至步骤 S202。

步骤 S201：根据存储集群的数量，发起对消息中间件主题的多个
监听，每一监听对应一个存储集群，消息中间件主题包括待写入数据。

步骤 S202：响应于对消息中间件主题的每一监听，将待写入数据
写入对应的存储集群。

消息中间件例如消息队列（MQ）。

响应于对消息中间件主题的每一监听，将待写入数据写入对应的
存储集群，可以包括：响应于对消息中间件主题的每一监听，通过预
设的写入接口将待写入数据写入对应的存储集群，预设的写入接口的
参数包括各存储集群的标识（ID），其中，若任一存储集群中待写入
数据写入失败，则重试对该存储集群的待写入数据写入的操作。

本公开实施例在业务需要将数据写入 ES 的时候，将数据发送到一个
消息队列主题中，通过监听此消息队列主题，以将数据写入到 ES 中，
这种方式将写入操作和最终写入动作解耦开来，多个存储集群写入则

多次监听此消息队列主题，若某个存储集群写入失败只需该存储集群重试即可，目前大部分的消息队列中间件都可以实现最终一致性，应用此特性来达到多存储集群数据的最终一致性，并大大增强了扩展性。

5 图 3 是根据本公开一个或多个实施例中的数据查询和数据写入的流程示意图。其中，客户端可以为用于数据查询和/或用于数据写入的装置。

10 如图 3 所示，在一个实施例中，配置中心通过动态监听机制，在监听到探活节点集群对存储集群的探活结果后，将对应的探活结果缓存到本地缓存，客户端响应于接收的数据查询请求，从配置中心或本地缓存中获取探活节点集群对各存储集群的探活结果。

15 具体地，存储集群可以为 ES (ElasticSearch) 集群，在多个机房分别部署 ES 集群 (简称集群)，每个集群命名可以为 cluster1、cluster2...，所有集群可以总称为 es-cluster；搭建一个配置中心，配置中心可以为 zookeeper (一种高可用分布式配置中心)；搭建探活节点集群，探活节点集群可包括至少 3 个探活节点，探活节点可以分别命名为 node1、node2、node3... (如图 3 以三个节点为例，即节点 1、节点 2、节点 3)，
20 每个探活节点连接到每个 ES 集群；搭建一个用于数据查询的客户端，建立客户端对每个 ES 集群的连接以及对配置中心的连接。启动探活程序，探活节点集群中的每个探活节点定时对每个 ES 集群发起健康度请求，ES 集群响应健康度请求生成探活结果并存放到 es-cluster 目录中。配置中心动态监听 es-cluster 目录，在该目录发生变化时，配置中心修改客户端的本地缓存，将探活结果保存至客户端的本地缓存中。客户端从配置中心或者本地缓存中获取探活结果。其中，本地缓存可以通过 Map (将键映射到值的对象) 实现。
25

30 在一个实施例中，对于每一存储集群，判断各探活节点对该存储集群的探活结果中，指示该存储集群可用的探活节点数量占探活节点

总数的比例是否超过比例阈值，若是，则确定该存储集群可用，并添加到可用存储集群列表。例如，存在两个存储集群（a，b），探活节点的数量为3个，3个探活节点对存储集群的探活结果分别为（1，1）、（0，1）、（0，0），其中1表示可用，0表示不可用，探活结果中的第一个数字为第一个集群的可用状态，第二个数字为第二个集群的可用状态，那么有一个探活节点的探活结果指示a集群可用，有两个探活节点的探活结果指示b集群可用。若将比例阈值设置为0.5，则对于a集群，指示该集群可用的探活节点数量占探活节点总数的比例为1/3，没有超过比例阈值0.5，则a集群不可用，不将a集群添加到可用存储集群列表；对于b集群，指示该集群可用的探活节点数量占探活节点总数的比例为2/3，超过了比例阈值0.5，则b集群可用，将b集群添加到可用存储集群列表。对于不可用的存储集群，可以进行上报处理，以进行修复。

在一个实施例中，按照集群选取设置，从可用存储集群列表中选择目标可用存储集群，将数据查询请求发送到目标可用存储集群，以由目标可用存储集群返回对应的数据查询结果。具体地，根据集群的负载大小确定负载均衡策略，从而进行集群选取设置，负载均衡策略可以包括全部请求策略、轮询策略和随机策略。其中，全部请求策略，即以可用存储集群列表中的所有可用存储集群作为目标可用存储集群，在收到其中一个目标可用存储集群返回的数据查询结果后，输出数据查询结果，该策略可以削平网络的波动和某一个存储集群的抖动对整个集群的影响，只要有一个存储集群正常，就可以正常提供服务，适用于集群整体压力不是很大的情况。轮询策略，即按照轮询策略从可用存储集群列表中选出当前轮询到的一个可用存储集群，作为目标可用存储集群，在收到目标可用存储集群返回的数据查询结果后，输出数据查询结果。随机策略，即每次使用随机的算法从可用存储集群列表中随机选出一个可用存储集群，作为目标可用存储集群，在收到目标可用存储集群返回的数据查询结果后，输出数据查询结果。

在一个实施例中，根据存储集群的数量，发起对消息中间件主题的多个监听，每一监听对应一个存储集群，响应于对消息中间件主题的每一监听，通过预设的写入接口将待写入数据写入对应的存储集群，预设的写入接口的参数包括各存储集群的标识。具体地，在有写入动作时，将待写入数据发送到消息中间件主题，按照写入接口参数中包括的待写入的各存储集群的标识，通过建立映射的方式，待写入的各存储集群订阅消息中间件主题，实现对消息中间件主题的监听，其中，订阅的数量与待写入的存储集群的数量相同。各存储集群通过多次监听此消息中间件主题，将待写入数据写入到各存储集群中。若某个存储集群写入失败，则只需该存储集群重试即可。本公开实施例基于消息中间件可以实现各存储集群的最终一致性，同时大大增强了扩展性。

图 4 是根据本公开一个或多个实施例中的数据查询装置的主要模块示意图。

如图 4 所示，本公开一个实施例的数据查询装置 400 主要包括：探活结果获取模块 401、可用存储集群确定模块 402、数据查询模块 403。

探活结果获取模块 401，用于响应于接收的数据查询请求，获取对各存储集群的探活结果，探活结果指示对应的存储集群当前是否可用。

可用存储集群确定模块 402，用于根据对各存储集群的探活结果，确定可用存储集群列表。

数据查询模块 403，用于按照集群选取设置，从可用存储集群列表中选择目标可用存储集群，将数据查询请求发送到目标可用存储集群，以由目标可用存储集群返回对应的数据查询结果。

在一个实施例中，探活结果获取模块 401 具体用于：从配置中心或本地缓存中获取探活节点集群对各存储集群的探活结果，其中，配

置中心通过动态监听机制，在监听到探活节点集群对存储集群的探活结果后，将对应的探活结果缓存到本地缓存。

5 在一个实施例中，可用存储集群确定模块 402 具体用于：对于每一存储集群，判断各探活节点对该存储集群的探活结果中，指示该存储集群可用的探活节点数量占探活节点总数的比例是否超过比例阈值，若是，则确定该存储集群可用，并添加到可用存储集群列表。

10 在一个实施例中，数据查询模块 403 具体用于：以可用存储集群列表中的所有可用存储集群作为目标可用存储集群，在收到其中一个目标可用存储集群返回的数据查询结果后，输出数据查询结果。

15 在一个实施例中，数据查询模块 403 具体用于：按照轮询策略从可用存储集群列表中选出当前轮询到的一个可用存储集群，作为目标可用存储集群，在收到目标可用存储集群返回的数据查询结果后，输出数据查询结果。

20 在一个实施例中，数据查询模块 403 具体用于：从可用存储集群列表中随机选出一个可用存储集群，作为目标可用存储集群，在收到目标可用存储集群返回的数据查询结果后，输出数据查询结果。

另外，在本公开实施例中数据查询装置的具体实施内容，在上面数据查询方法中已经详细说明了，故在此重复内容不再说明。

25 图 5 是根据本公开一个或多个实施例中的数据写入装置的主要模块示意图。

如图 5 所示，本公开一个实施例的数据写入装置 500 主要包括：监听模块 501、数据写入模块 502。

30

监听模块 501，用于根据存储集群的数量，发起对消息中间件主题的多个监听，每一监听对应一个存储集群，消息中间件主题包括待写入数据。

5 数据写入模块 502，用于响应于对消息中间件主题的每一监听，将待写入数据写入对应的存储集群。

10 在一个实施例中，数据写入模块 502 具体用于：响应于对消息中间件主题的每一监听，通过预设的写入接口将待写入数据写入对应的存储集群，预设的写入接口的参数包括各存储集群的标识，其中，若任一存储集群中待写入数据写入失败，则重试对该存储集群的待写入数据写入的操作。

15 另外，在本公开实施例中数据写入装置的具体实施内容，在上面数据写入方法中已经详细说明了，故在此重复内容不再说明。

图 6 是根据本公开一个或多个实施例中的数据查询系统的主要架构示意图。

20 如图 6 所示，本公开一个实施例的数据查询系统 600 主要包括：数据查询装置 601，还包括存储集群 602、配置中心 603 和探活节点集群 604。存储集群 602 的数量为多个。

25 数据查询装置 601 与上文实施例中的数据查询装置 400 所含的功能模块相同。

30 探活节点集群 604 包括一个或多个探活节点，探活节点用于向各存储集群分别发送健康度请求，以根据存储集群对健康度请求的响应情况，确定存储集群的探活结果，探活结果指示对应的存储集群当前是否可用。

配置中心 603 用于接收各探活节点上报的探活结果，并向数据查询装置提供探活结果。

5 在一个实施例中，配置中心 603 具体用于：响应于数据查询装置 601 的探活结果获取请求，向数据查询装置 601 提供探活结果。

 在一个实施例中，配置中心 603 具体用于：通过动态监听机制，
10 在监听到探活节点集群对存储集群的探活结果后，将对应的探活结果
 缓存到数据查询装置 601 对应的本地缓存。

 对于上文实施例已经介绍过的内容，本实施例不再赘述。

 图 7 示出了可以应用本公开一个或多个实施例中的数据查询方法、
15 数据写入方法、数据查询装置、数据写入装置或数据查询系统的示例
 性系统架构 700。

 如图 7 所示，系统架构 700 可以包括终端设备 701、702、703，网
20 络 704 和服务器 705。网络 704 用以在终端设备 701、702、703 和服务
 器 705 之间提供通信链路的介质。网络 704 可以包括各种连接类型，
 例如有线、无线通信链路或者光纤电缆等等。

 用户可以使用终端设备 701、702、703 通过网络 704 与服务器 705
25 交互，以接收或发送消息等。终端设备 701、702、703 上可以安装有
 各种通讯客户端应用，例如购物类应用、网页浏览器应用、搜索类应
 用、即时通信工具、邮箱客户端、社交平台软件等（仅为示例）。

 终端设备 701、702、703 可以是具有显示屏并且支持网页浏览的
30 各种电子设备，包括但不限于智能手机、平板电脑、膝上型便携计算
 机和台式计算机等等。

服务器 705 可以是提供各种服务的服务器，例如对用户利用终端设备 701、702、703 所浏览的购物类网站提供支持的后台管理服务器（仅为示例）。后台管理服务器可以对接收到的产品信息查询请求等数据进行分析等处理，并将处理结果（例如目标推送信息、产品信息--

5 仅为示例）反馈给终端设备。

需要说明的是，在一个实施例中，本公开实施例所提供的数据查询方法一般由服务器 705 或终端设备 701、702、703 执行，相应地，

10 数据查询装置一般设置于服务器 705 或终端设备 701、702、703 中。

在另一个实施例中，本公开实施例所提供的数据写入方法一般由服务器 705 或终端设备 701、702、703 执行，相应地，数据写入装置一般设置于服务器 705 或终端设备 701、702、703 中。

15

应该理解，图 7 中的终端设备、网络和服务器的数目仅仅是示意性的。根据实现需要，可以具有任意数目的终端设备、网络和服务器。

下面参考图 8，其示出了适于用来实现本公开一个或多个实施例中的终端设备或服务器的计算机系统 800 的结构示意图。图 8 示出的终端设备或服务器仅仅是一个示例，不应对本公开实施例的功能和使用范围带来任何限制。

20

如图 8 所示，计算机系统 800 包括中央处理单元（CPU）801，其可以根据存储在只读存储器（ROM）802 中的程序或者从存储部分 808 加载到随机访问存储器（RAM）803 中的程序而执行各种适当的动作和处理。在 RAM 803 中，还存储有系统 800 操作所需的各种程序和数据。CPU 801、ROM 802 以及 RAM 803 通过总线 804 彼此相连。输入/输出（I/O）接口 805 也连接至总线 804。

25

30

以下部件连接至 I/O 接口 805：包括键盘、鼠标等的输入部分 806；包括诸如阴极射线管（CRT）、液晶显示器（LCD）等以及扬声器等的输出部分 807；包括硬盘等的存储部分 808；以及包括诸如 LAN 卡、调制解调器等网络接口卡的通信部分 809。通信部分 809 经由诸如因特网的网络执行通信处理。驱动器 810 也根据需要连接至 I/O 接口 805。可拆卸介质 811，诸如磁盘、光盘、磁光盘、半导体存储器等等，根据需要安装在驱动器 810 上，以便于从其上读出的计算机程序根据需要被安装入存储部分 808。

特别地，根据本公开的实施例，上文参考流程图描述的过程可以被实现为计算机软件程序。例如，本公开的实施例包括一种计算机程序产品，其包括承载在计算机可读介质上的计算机程序，该计算机程序包含用于执行流程图所示的方法的程序代码。在这样的实施例中，该计算机程序可以通过通信部分 809 从网络上被下载和安装，和/或从可拆卸介质 811 被安装。在该计算机程序被中央处理单元（CPU）801 执行时，执行本公开的系统中限定的上述功能。

需要说明的是，本公开所示的计算机可读介质可以是计算机可读信号介质或者计算机可读存储介质或者是上述两者的任意组合。计算机可读存储介质例如可以是一——但不限于——电、磁、光、电磁、红外线、或半导体的系统、装置或器件，或者任意以上的组合。计算机可读存储介质的更具体的例子可以包括但不限于：具有一个或多个导线的电连接、便携式计算机磁盘、硬盘、随机访问存储器（RAM）、只读存储器（ROM）、可擦式可编程只读存储器（EPROM 或闪存）、光纤、便携式紧凑磁盘只读存储器（CD-ROM）、光存储器件、磁存储器件、或者上述的任意合适的组合。在本公开中，计算机可读存储介质可以是任何包含或存储程序的有形介质，该程序可以被指令执行系统、装置或者器件使用或者与其结合使用。而在本公开中，计算机可读的信号介质可以包括在基带中或者作为载波一部分传播的数据信号，其中承载了计算机可读的程序代码。这种传播的数据信号可以采

用多种形式，包括但不限于电磁信号、光信号或上述的任意合适的组合。计算机可读的信号介质还可以是计算机可读存储介质以外的任何计算机可读介质，该计算机可读介质可以发送、传播或者传输用于由指令执行系统、装置或者器件使用或者与其结合使用的程序。计算机可读介质上包含的程序代码可以用任何适当的介质传输，包括但不限于：无线、电线、光缆、RF 等等，或者上述的任意合适的组合。

附图中的流程图和框图，图示了按照本公开各种实施例的系统、方法和计算机程序产品的可能实现的体系架构、功能和操作。在这点上，流程图或框图中的每个方框可以代表一个模块、程序段、或代码的一部分，上述模块、程序段、或代码的一部分包含一个或多个用于实现规定的逻辑功能的可执行指令。也应当注意，在有些作为替换的实现中，方框中所标注的功能也可以以不同于附图中所标注的顺序发生。例如，两个接连地表示的方框实际上可以基本并行地执行，它们有时也可以按相反的顺序执行，这依所涉及的功能而定。也要注意的，框图或流程图中的每个方框、以及框图或流程图中的方框的组合，可以用执行规定的功能或操作的专用的基于硬件的系统来实现，或者可以用专用硬件与计算机指令的组合来实现。

描述于本公开实施例中所涉及到的模块可以通过软件的方式实现，也可以通过硬件的方式来实现。所描述的模块也可以设置在处理器中，例如，可以描述为：一种处理器包括探活结果获取模块、可用存储集群确定模块、数据查询模块。其中，这些模块的名称在某种情况下并不构成对该模块本身的限定，例如，探活结果获取模块还可以被描述为“用于响应于接收的数据查询请求，获取对各存储集群的探活结果的模块”。

作为另一方面，本公开还提供了一种计算机可读介质，该计算机可读介质可以是上述实施例中描述的设备中所包含的；也可以是单独存在，而未装配入该设备中。上述计算机可读介质承载有一个或者多

个程序，当上述一个或者多个程序被一个该设备执行时，使得该设备包括：响应于接收的数据查询请求，获取对各存储集群的探活结果，探活结果指示对应的存储集群当前是否可用；根据对各存储集群的探活结果，确定可用存储集群列表；按照集群选取设置，从可用存储集群列表中选取目标可用存储集群，将数据查询请求发送到目标可用存储集群，以由目标可用存储集群返回对应的数据查询结果。

根据本公开实施例的技术方案，响应于接收的数据查询请求，获取对各存储集群的探活结果，探活结果指示对应的存储集群当前是否可用；根据对各存储集群的探活结果，确定可用存储集群列表；按照集群选取设置，从可用存储集群列表中选取目标可用存储集群，将数据查询请求发送到目标可用存储集群，以由目标可用存储集群返回对应的数据查询结果。能够实现向多个集群写入数据的一致性，可自动切换集群，提高集群部署的容灾能力，解决了在机房级别故障时导致整个集群不可用的问题，同时克服了由于集群内部需要进行通信而耗费时间大的缺陷。

上述具体实施方式，并不构成对本公开保护范围的限制。本领域技术人员应该明白的是，取决于设计要求和因素，可以发生各种各样的修改、组合、子组合和替代。任何在本公开的精神和原则之内所作的修改、等同替换和改进等，均应包含在本公开保护范围之内。

权利要求

1. 一种数据查询方法，包括：

5 响应于接收的数据查询请求，获取对各存储集群的探活结果，所述探活结果指示对应的存储集群当前是否可用；

 根据对各存储集群的探活结果，确定可用存储集群列表；

 按照集群选取设置，从所述可用存储集群列表选取目标可用存储集群，将所述数据查询请求发送到所述目标可用存储集群，以由所述目标可用存储集群返回对应的数据查询结果。

10

2. 根据权利要求 1 所述的方法，其中，所述获取对各存储集群的探活结果，包括：

 从配置中心或本地缓存中获取探活节点集群对所述各存储集群的探活结果，其中，所述配置中心通过动态监听机制，在监听到所述探活节点集群对所述存储集群的探活结果后，将对应的所述探活结果缓存到所述本地缓存。

15

3. 根据权利要求 2 所述的方法，其中，所述探活节点集群包括一个或多个探活节点，所述探活节点用于向所述各存储集群分别发送健康度请求，以根据所述存储集群对所述健康度请求的响应情况，确定对应的所述存储集群是否可用。

20

4. 根据权利要求 2 所述的方法，其中，所述根据对各存储集群的探活结果，确定可用存储集群列表，包括：

25

 对于每一存储集群，判断各所述探活节点对该存储集群的探活结果中，指示该存储集群可用的探活节点数量占探活节点总数的比例是否超过比例阈值，若是，则确定该存储集群可用，并添加到所述可用存储集群列表。

30

5. 根据权利要求 1 所述的方法，其中，所述按照集群选取设置，从所述可用存储集群列表中选取目标可用存储集群，将所述数据查询请求发送到所述目标可用存储集群，以由所述目标可用存储集群返回对应的数据查询结果，包括：

5 以所述可用存储集群列表中的所有可用存储集群作为所述目标可用存储集群，在收到其中一个目标可用存储集群返回的数据查询结果后，输出所述数据查询结果。

6. 根据权利要求 1 所述的方法，其中，所述按照集群选取设置，
10 从所述可用存储集群列表中选取目标可用存储集群，将所述数据查询请求发送到所述目标可用存储集群，以由所述目标可用存储集群返回对应的数据查询结果，包括：

 按照轮询策略从所述可用存储集群列表中选出当前轮询到的一个可用存储集群，作为所述目标可用存储集群，在收到所述目标可用存储
15 集群返回的数据查询结果后，输出所述数据查询结果。

7. 根据权利要求 1 所述的方法，其中，所述按照集群选取设置，从所述可用存储集群列表中选取目标可用存储集群，将所述数据查询请求发送到所述目标可用存储集群，以由所述目标可用存储集群返回
20 对应的数据查询结果，包括：

 从所述可用存储集群列表中随机选出一个可用存储集群，作为所述目标可用存储集群，在收到所述目标可用存储集群返回的数据查询结果后，输出所述数据查询结果。

25 8. 一种数据写入方法，包括：

 根据存储集群的数量，发起对消息中间件主题的多个监听，每一监听对应一个存储集群，所述消息中间件主题包括待写入数据；

 响应于对所述消息中间件主题的每一监听，将所述待写入数据写入对应的存储集群。

30

9. 根据权利要求 8 所述的方法，其中，所述响应于对所述消息中间件主题的每一监听，将所述待写入数据写入对应的存储集群，包括：

响应于对所述消息中间件主题的每一监听，通过预设的写入接口将所述待写入数据写入对应的存储集群，所述预设的写入接口的参数包括各存储集群的标识，其中，若任一存储集群中所述待写入数据写入失败，则重试对该存储集群的所述待写入数据写入的操作。

10. 一种数据查询装置，包括：

探活结果获取模块，用于响应于接收的数据查询请求，获取对各存储集群的探活结果，所述探活结果指示对应的存储集群当前是否可用；

可用存储集群确定模块，用于根据对各存储集群的探活结果，确定可用存储集群列表；

数据查询模块，用于按照集群选取设置，从所述可用存储集群列表选取目标可用存储集群，将所述数据查询请求发送到所述目标可用存储集群，以由所述目标可用存储集群返回对应的数据查询结果。

11. 根据权利要求 10 所述的装置，其中，所述探活结果获取模块还用于：

从配置中心或本地缓存中获取探活节点集群对所述各存储集群的探活结果，其中，所述配置中心通过动态监听机制，在监听到所述探活节点集群对所述存储集群的探活结果后，将对应的所述探活结果缓存到所述本地缓存。

12. 根据权利要求 11 所述的装置，其中，所述可用存储集群确定模块还用于：

对于每一存储集群，判断各所述探活节点对该存储集群的探活结果中，指示该存储集群可用的探活节点数量占探活节点总数的比例是否超过比例阈值，若是，则确定该存储集群可用，并添加到所述可用存储集群列表。

13. 根据权利要求 10 所述的装置，其中，所述数据查询模块还用于：

5 以所述可用存储集群列表中的所有可用存储集群作为所述目标可用存储集群，在收到其中一个目标可用存储集群返回的数据查询结果后，输出所述数据查询结果。

14. 根据权利要求 10 所述的装置，其中，所述数据查询模块还用于：

10 按照轮询策略从所述可用存储集群列表中选出当前轮询到的一个可用存储集群，作为所述目标可用存储集群，在收到所述目标可用存储集群返回的数据查询结果后，输出所述数据查询结果。

15 15. 根据权利要求 10 所述的装置，其中，所述数据查询模块还用于：

从所述可用存储集群列表中随机选出一个可用存储集群，作为所述目标可用存储集群，在收到所述目标可用存储集群返回的数据查询结果后，输出所述数据查询结果。

20 16. 一种数据写入装置，包括：

监听模块，用于根据存储集群的数量，发起对消息中间件主题的多个监听，每一监听对应一个存储集群，所述消息中间件主题包括待写入数据；

25 数据写入模块，用于响应于对所述消息中间件主题的每一监听，将所述待写入数据写入对应的存储集群。

17. 根据权利要求 16 所述的装置，其中，所述数据写入模块还用于：

30 响应于对所述消息中间件主题的每一监听，通过预设的写入接口将所述待写入数据写入对应的存储集群，所述预设的写入接口的参数

包括各存储集群的标识，其中，若任一存储集群中所述待写入数据写入失败，则重试对该存储集群的所述待写入数据写入的操作。

5 18. 一种数据查询系统，包括：如权利要求 10 至 15 中任一项所述的数据查询装置，还包括各存储集群、配置中心和探活节点集群，其中：

所述探活节点集群包括一个或多个探活节点，所述探活节点用于向所述各存储集群分别发送健康度请求，以根据所述存储集群对所述健康度请求的响应情况，确定所述存储集群的探活结果，所述探活结果指示对应的所述存储集群当前是否可用；

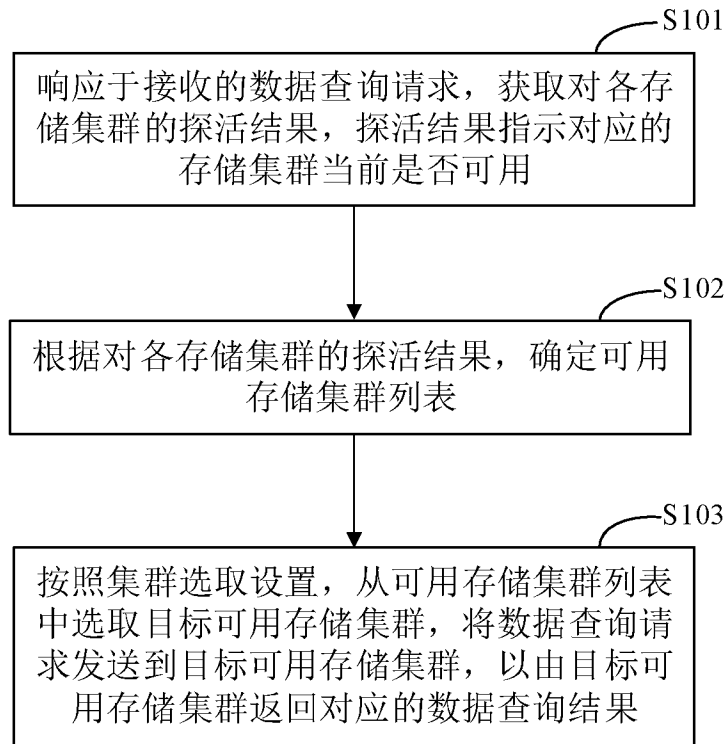
所述配置中心用于接收各所述探活节点上报的所述探活结果，并向所述数据查询装置提供所述探活结果。

15 19. 根据权利要求 18 所述的系统，其中，所述配置中心还用于：响应于所述数据查询装置的探活结果获取请求，向所述数据查询装置提供所述探活结果。

20 20. 根据权利要求 18 所述的系统，其中，所述配置中心还用于：通过动态监听机制，在监听到所述探活节点集群对所述存储集群的探活结果后，将对应的所述探活结果缓存到所述数据查询装置对应的本地缓存。

21. 一种电子设备，包括：
一个或多个处理器；
25 存储装置，用于存储一个或多个程序，
当所述一个或多个程序被所述一个或多个处理器执行，使得所述一个或多个处理器实现如权利要求 1-9 中任一所述的方法。

30 22. 一种计算机可读介质，其上存储有计算机程序，所述计算机程序被处理器执行时实现如权利要求 1-9 中任一所述的方法。

**图 1**

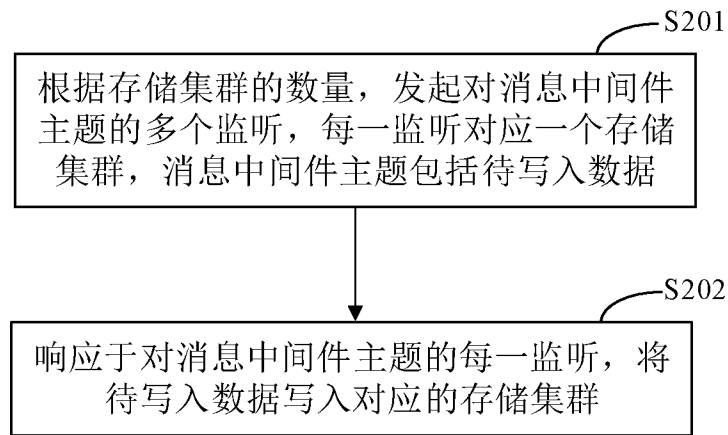


图 2

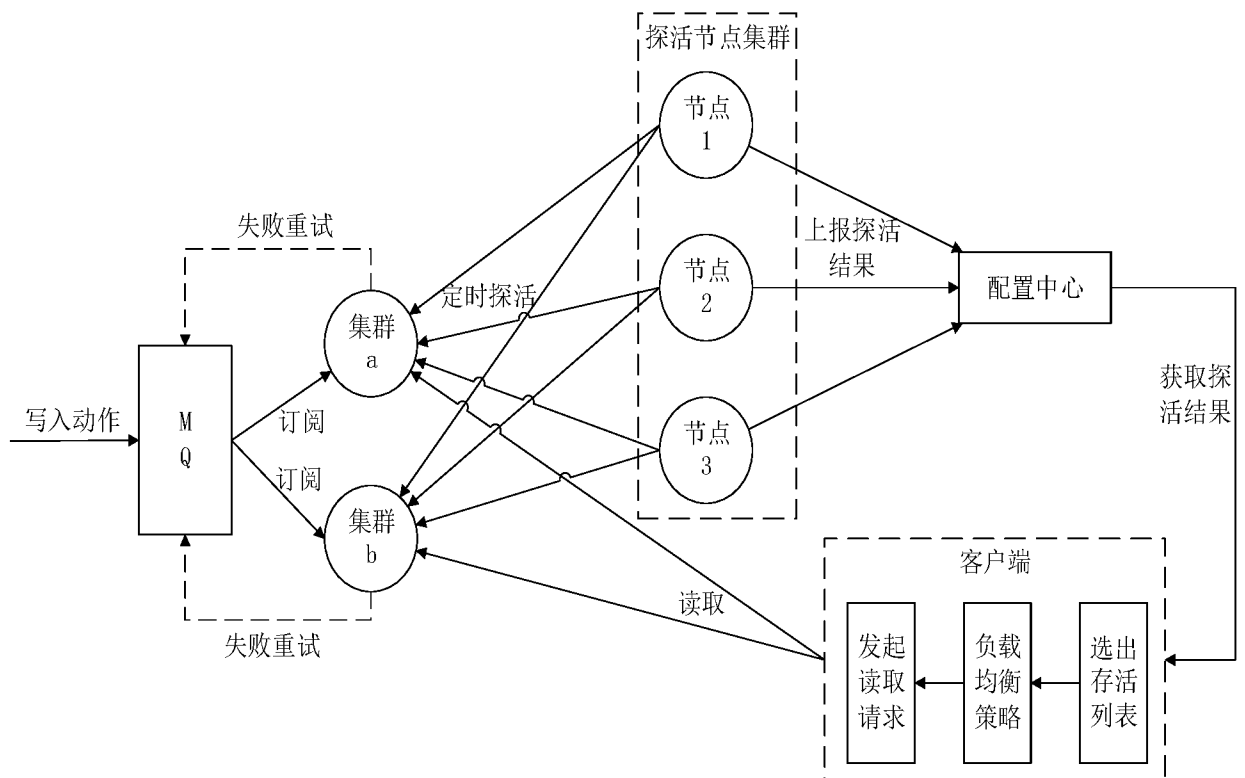


图 3

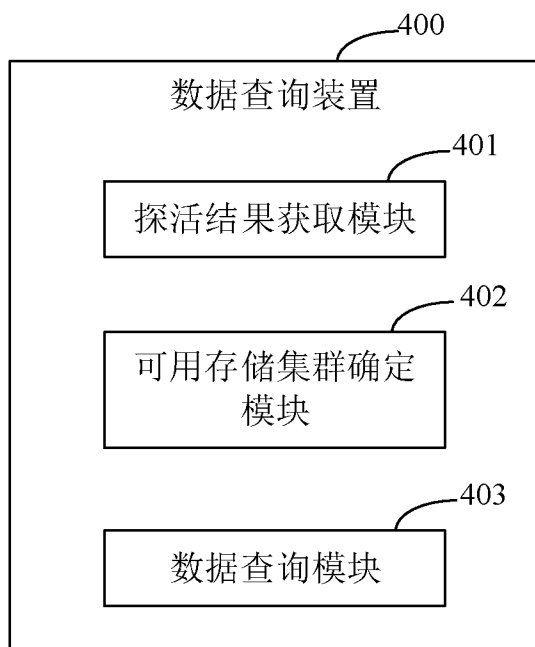


图 4

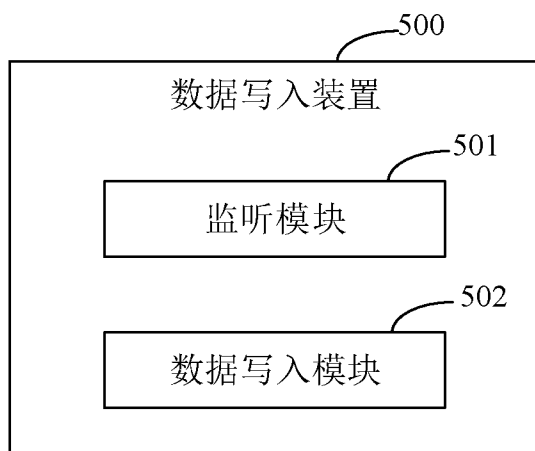


图 5

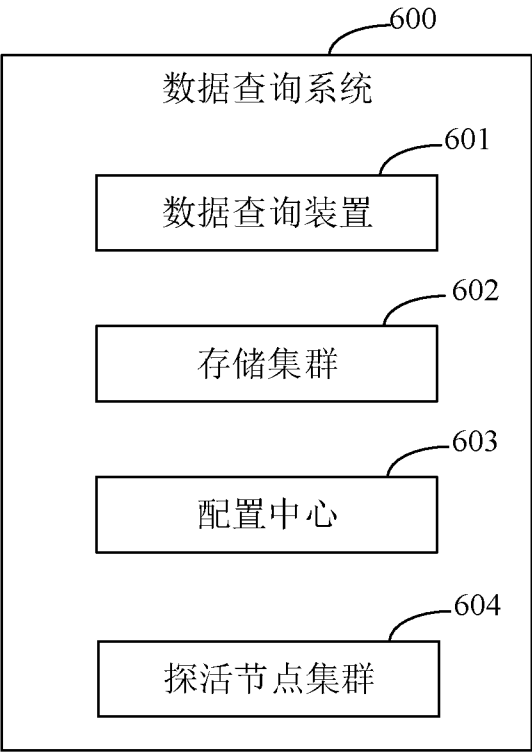


图 6

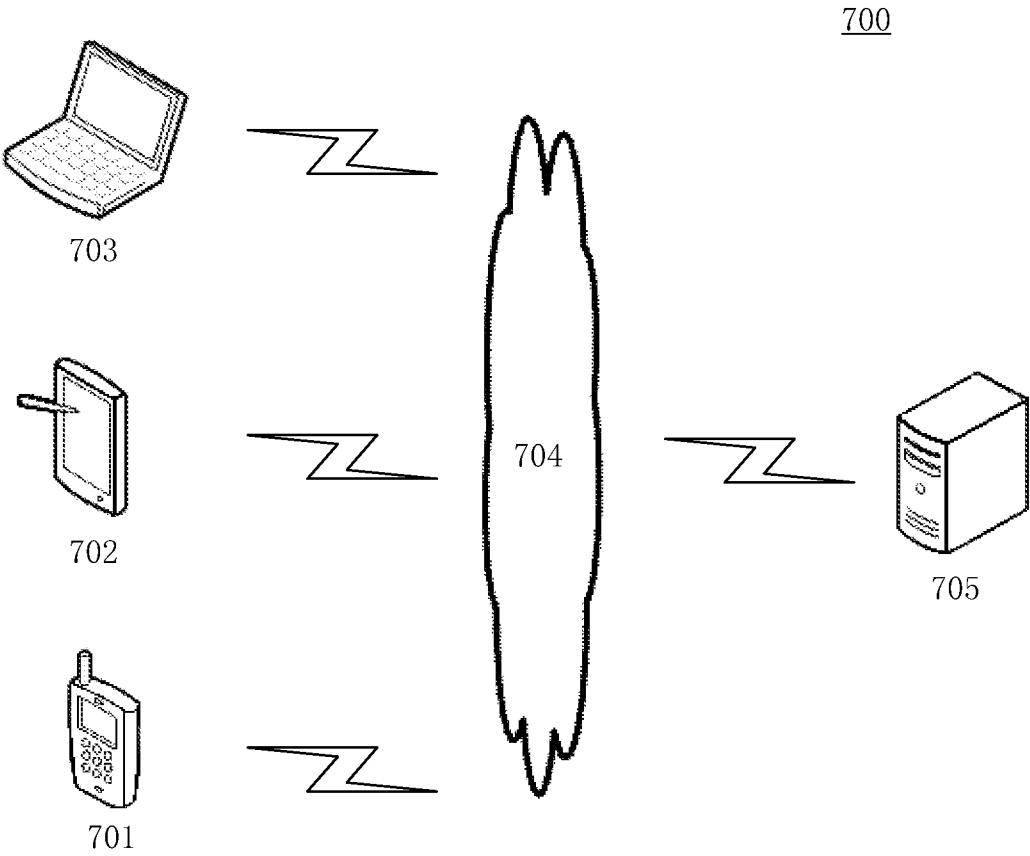


图 7

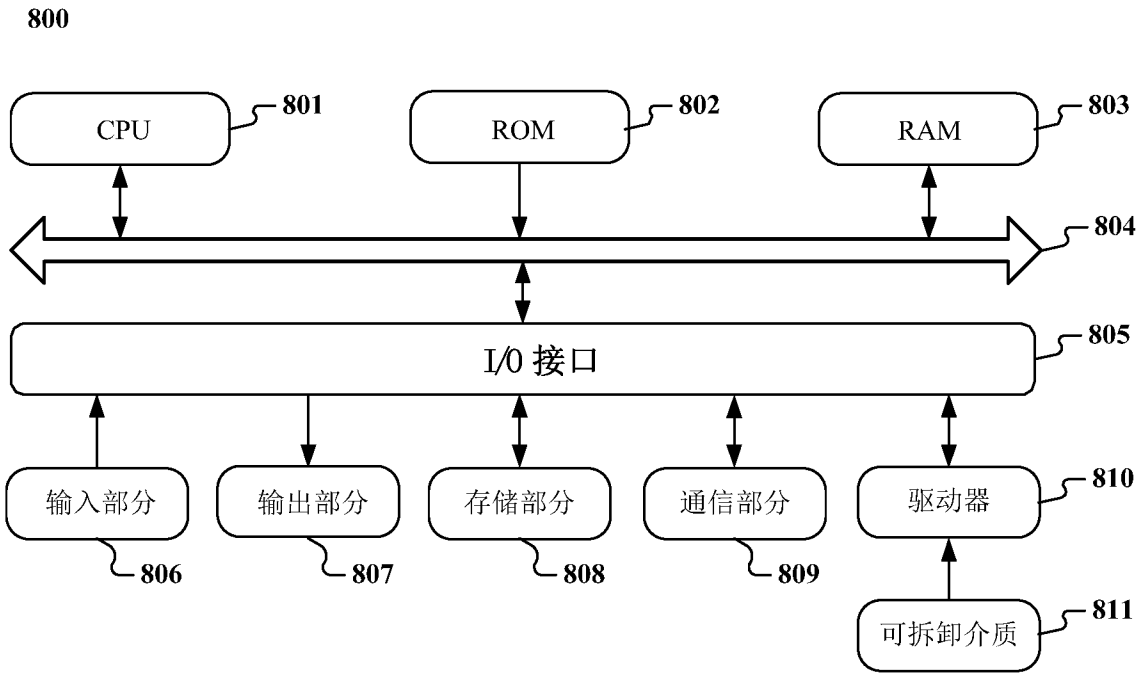


图 8

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2022/136753

A. CLASSIFICATION OF SUBJECT MATTER

G06F16/21(2019.01)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)

IPC: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)

CNTXT, ENTXT, ENTXTC, DWPL, CNKI: 存储, 集群, 检测, 监听, 健康, 可用, 探活, storage, cluster, detect, monitor, health, available, active

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	CN 108241469 A (BEIJING KINGSOFT CLOUD NETWORK TECHNIQUE CO., LTD.; BEIJING KSYUN TECHNOLOGY CO., LTD.) 03 July 2018 (2018-07-03) description, paragraphs 6-128	1-22
A	CN 112583760 A (CHINA MOBILE SUZHOU SOFTWARE TECH CO., LTD.; CHINA MOBILE COMMUNICATIONS GROUP CO., LTD.) 30 March 2021 (2021-03-30) entire document	1-22
A	US 9645628 B1 (EMC IP HOLDING COMPANY L.L.C.) 09 May 2017 (2017-05-09) entire document	1-22



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

“D” document cited by the applicant in the international application

“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 February 2023

Date of mailing of the international search report

09 February 2023

Name and mailing address of the ISA/CN

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

Facsimile No. (86-10)62019451

Authorized officer

Telephone No.

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/CN2022/136753

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
CN	108241469	A	03 July 2018	CN	108241469	B	30 April 2021
CN	112583760	A	30 March 2021	CN	112583760	B	31 May 2022
US	9645628	B1	09 May 2017	None			

国际检索报告

国际申请号

PCT/CN2022/136753

A. 主题的分类

G06F16/21 (2019.01) 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)

IPC: 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))

CNTXT, ENTXT, ENTXTC, DWPI, CNKI: 存储, 集群, 检测, 监听, 健康, 可用, 探活, storage, cluster, detect, monitor, health, available, active

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
X	CN 108241469 A (北京金山云网络技术有限公司 北京金山云科技有限公司) 2018年7月3日 (2018 - 07 - 03) 说明书第6-128段	1-22
A	CN 112583760 A (中移(苏州)软件技术有限公司 中国移动通信集团有限公司) 2021年3月30日 (2021 - 03 - 30) 全文	1-22
A	US 9645628 B1 (EMC IP HOLDING CO LLC) 2017年5月9日 (2017 - 05 - 09) 全文	1-22

☐ 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

“D” 申请人在国际申请中引证的文件

“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

2023年2月7日

国际检索报告邮寄日期

2023年2月9日

ISA/CN的名称和邮寄地址

中国国家知识产权局

中国北京市海淀区蓟门桥西土城路6号 100088

传真号 (86-10)62019451

授权官员

王雪莲

电话号码 (+86) 62411747

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2022/136753

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	108241469	A	2018年7月3日	CN	108241469 B	2021年4月30日
CN	112583760	A	2021年3月30日	CN	112583760 B	2022年5月31日
US	9645628	B1	2017年5月9日	无		