

(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
2020 年 7 月 30 日 (30.07.2020)

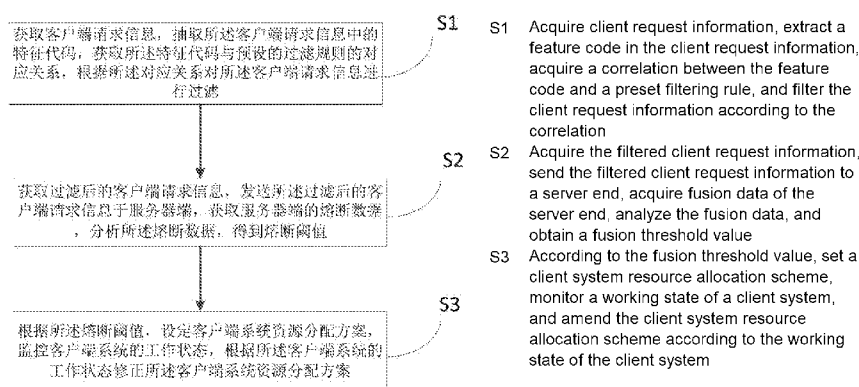


(10) 国际公布号
WO 2020/151358 A1

- (51) 国际专利分类号:
G06F 9/50 (2006.01)
- (21) 国际申请号: PCT/CN2019/120525
- (22) 国际申请日: 2019 年 11 月 25 日 (25.11.2019)
- (25) 申请语言: 中文
- (26) 公布语言: 中文
- (30) 优先权:
201910061291.8 2019 年 1 月 23 日 (23.01.2019) CN
- (71) 申请人: 深圳壹账通智能科技有限公司(ONE CONNECT SMART TECHNOLOGY CO., LTD. (SHENZHEN)) [CN/CN]; 中国广东省深圳市前海深港合作区前湾一路 1 号 A 栋 201 室, Guangdong 518052 (CN)。
- (72) 发明人: 何刘兴 (HE, Liuxing); 中国广东省深圳市前海深港合作区前湾一路 1 号 A 栋 201 室, Guangdong 518052 (CN)。
- (74) 代理人: 北京市京大律师事务所(BEIJING JINGDA LAW FIRM); 中国北京市海淀区海淀中街 16 号 7 层 2 单元 703, Beijing 100080 (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, 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,

(54) **Title:** SYSTEM RESOURCE ALLOCATION METHOD AND APPARATUS, COMPUTER DEVICE, AND STORAGE MEDIUM

(54) 发明名称: 系统资源分配方法、装置、计算机设备和存储介质



(57) **Abstract:** The present application relates to the technical field of software monitoring, and in particular to a system resource allocation method and apparatus, a computer device and a storage medium. The method comprises: acquiring client request information, extracting a feature code in the client request information, acquiring a correlation between the feature code and a preset filtering rule, and filtering the client request information according to the correlation; acquiring the filtered client request information, sending the filtered client request information to a server end, acquiring fusion data of the server end, analyzing the fusion data, and obtaining a fusion threshold value; and according to the fusion threshold value, setting a client system resource allocation scheme, monitoring a working state of a client system, and amending the client system resource allocation scheme according to the working state of the client system. According to the present application, by means of effective monitoring of a fusion process, system capacity can be timely expanded or reduced, thus guaranteeing normal operation of the system under high availability.

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：本申请涉及软件监控技术领域，尤其涉及一种系统资源分配方法、装置、计算机设备和存储介质。获取客户端请求信息，抽取所述客户端请求信息中的特征代码，获取所述特征代码与预设的过滤规则的对应关系，根据所述对应关系对所述客户端请求信息进行过滤；获取过滤后的客户端请求信息，发送所述过滤后的客户端请求信息于服务器端，获取服务器端的熔断数据，分析所述熔断数据，得到熔断阈值；根据所述熔断阈值，设定客户端系统资源分配方案，监控客户端系统的工作状态，根据所述客户端系统的工作状态修正所述客户端系统资源分配方案。本申请通过熔断过程的有效监控，能够及时扩充系统容量或者减少系统容量，从而保证系统在高可用性下的正常运行。

系统资源分配方法、装置、计算机设备和存储介质

本申请要求于2019年1月23日提交中国专利局、申请号为201910061291.8、发明名称为“系统资源分配方法、装置、计算机设备和存储介质”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。

5

技术领域

本申请涉及软件监控技术领域，尤其涉及一种系统资源分配方法、装置、计算机设备和存储介质。

背景技术

10 在系统使用过程中存在由于软件或者硬件导致的雪崩，雪崩是指平时正常调用和被调用的A系统和B系统，突然A系统对B系统的访问超出了B系统的承受能力则会造成B系统崩溃。

目前，在应对雪崩时，通常采用扩容的方法，业内普遍做法是硬件解决方案，即增加机器节点来抗压，但成本过高且浪费。纯软件解决方案较为经济，
15 无需增加机器节点，在流量过大时屏蔽掉部分请求，避免整个系统雪崩，使系统仍能正常运行。大部分软件解决方案要求修改现有系统，成本也较高，而开源熔断器hystrix提供了一种低成本零侵入的软件方案，其主要利用熔断机制。所谓熔断机制由于某些原因使得系统出现了过载现象，为防止造成整个系统故障，从而采用的一种保护措施，所以很多地方把熔断亦称为过载保护。

20 但是hystrix的可用性策略不够灵活和自动化，而且和所有纯软件方案一样，要通过丢弃部分请求来换取可用性，这在高可用性要求的场景无法适用。

发明内容

25 基于此，有必要针对系统在熔断的过程中存在着降级，导致无法适用于高可用性的场景的问题，提供一种系统资源分配方法、装置、计算机设备和存储介质。

一种系统资源分配方法，包括如下步骤：

获取客户端请求信息，抽取所述客户端请求信息中的特征代码，获取所述特征代码与预设的过滤规则的对应关系，根据所述对应关系对所述客户端请求信息进行过滤；

获取过滤后的客户端请求信息，发送所述过滤后的客户端请求信息于服务器端，获取服务器端的熔断数据，分析所述熔断数据，得到熔断阈值；

根据所述熔断阈值，设定客户端系统资源分配方案，监控客户端系统的工作状态，根据所述客户端系统的工作状态修正所述客户端系统资源分配方案。

5 一种系统资源分配装置，包括如下模块：

请求信息过滤模块，设置为获取客户端请求信息，抽取所述客户端请求信息中的特征代码，获取所述特征代码与预设的过滤规则的对应关系，根据所述对应关系对所述客户端请求信息进行过滤；

10 熔断阈值生成模块，设置为获取过滤后的客户端请求信息，发送所述过滤后的客户端请求信息于服务器端，获取服务器端的熔断数据，分析所述熔断数据，得到熔断阈值；

系统运行监控模块，设置为根据所述熔断阈值，设定客户端系统资源分配方案，监控客户端系统的工作状态，根据所述客户端系统的工作状态修正所述客户端系统资源分配方案。

15 一种计算机设备，包括存储器和处理器，所述存储器中存储有计算机可读指令，所述计算机可读指令被所述处理器执行时，使得所述处理器执行上述系统资源分配方法的步骤。

一种存储有计算机可读指令的存储介质，所述计算机可读指令被一个或多个处理器执行时，使得一个或多个处理器执行上述系统资源分配方法的步骤。

20 附图说明

通过阅读下文优选实施方式的详细描述，各种其他的优点和益处对于本领域普通技术人员将变得清楚明了。附图仅用于示出优选实施方式的目的，而并不认为是对本申请的限制。

图 1 为本申请在一个实施例中的一种系统资源分配方法的整体流程图；

25 图 2 为本申请在一个实施例中的一种系统资源分配方法中的请求信息过滤过程示意图；

图 3 为本申请在一个实施例中的一种系统资源分配方法中的系统运行监控过程示意图；

图 4 为本申请在一个实施例中的一种系统资源分配装置的结构图。

30

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仅用以解释本申请，并不用于限定本申请。

本技术领域技术人员可以理解，除非特意声明，这里使用的单数形式“一”、“一个”、“所述”和“该”也可包括复数形式。应该进一步理解的是，本申请的说明书中使用的措辞“包括”是指存在所述特征、整数、步骤、操作、元件和/或组件，但是并不排除存在或添加一个或多个其他特征、整数、步骤、操作、元件、组件和/或它们的组。

图 1 为本申请在一个实施例中的一种系统资源分配方法的整体流程图，如图 1 所示，一种系统资源分配方法，包括以下步骤：

S1，获取客户端请求信息，抽取所述客户端请求信息中的特征代码，获取所述特征代码与预设的过滤规则的对应关系，根据所述对应关系对所述客户端请求信息进行过滤；

具体的，在获取所述客户端请求信息后对所述客户端的权限要进行核验，即从所述客户端请求信息中抽取出客户端的身份识别符，将所述身份识别符与存储在服务器中的客户端信息进行比对，比对一致后则接受所述客户端发出的请求信息，否则不接受。

本步骤中的过滤规则主要包含有整体过滤、分段过滤、间隔过滤等，每一种过滤规则对应一个编号，将获取的特征代码与编号进行比较，若所述特征代码的首字符与所述编号一致，则选用所述编号对应的过滤规则对客户端请求信息进行过滤。

S2，获取过滤后的客户端请求信息，发送所述过滤后的客户端请求信息于服务器端，获取服务器端的熔断数据，分析所述熔断数据，得到熔断阈值；

具体的，所述服务器端的熔断数据是指在服务器端的模拟器中进行负载模拟测试，当数个客户端请求信息进入到模拟器后，模拟器的容量会达到上限，此时停止发送客户端请求信息到模拟器中，即此时服务器端发生熔断，记录此时模拟器中的数据量和发生熔断的时间得到所述熔断阈值。

S3，根据所述熔断阈值，设定客户端系统资源分配方案，监控客户端系统的工作状态，根据所述客户端系统的工作状态修正所述客户端系统资源分配方案。

具体的，在设定好客户端资源分配方案后，根据客户端系统发生熔断的历史数据，设置数个监测客户端工作状态的时间节点，当所述时间节点到来时，

采集所述客户端的工作数据，若出现过载熔断的情况则对所述客户端系统资源分配方案，否则不修正。

本实施例，通过对熔断数据分析后设置熔断阈值，并生成客户端资源分配方案从而实现了系统资源的合理分配。

5 图 2 为本申请在一个实施例中的一种系统资源分配方法中的请求信息过滤过程示意图，如图所示，所述 S1，获取客户端请求信息，抽取所述客户端请求信息中的特征代码，获取所述特征代码与预设的过滤规则的对应关系，根据所述对应关系对所述客户端请求信息进行过滤，包括：

10 S101、获取发出所述客户端请求信息的客户端 IP 地址，将所述客户端 IP 地址进行 DNS 解析得到 DNS 解析代码；

具体的，客户端的 IP 地址可以是动态 IP 地址，在对所述客户端 IP 地址进行 DNS 解析时可以采集客户端 IP 地址的历史数据，根据历史数据获取所述 DNS 解析代码所对应的各个 IP 地址。

15 S102、根据所述 DNS 解析代码与预设的 DNS 解析代码表之间的映射关系，获得所述客户端 IP 地址对应的客户端性能参数表；

具体的，在 DNS 解析代码表中设置有索引目录，将所述 DNS 解析代码的前 3 个字符与所述索引目录中的索引项进行比对一致后得到所述 DNS 解析代码对应的索引项，根据所述索引项获知所述 DNS 解析代码在所述预设的解析代码表中的哪一行。在解析代码表中，每一个 DNS 解析代码所在行都链接有所述
20 DNS 解析代码对应的性能参数表。因此，只需获取客户端的 DNS 解析代码就可以得到客户端性能参数表。其中，所述预设的 DNS 解析代码表存储在数据库中。

S103、按照所述客户端性能参数表中的各性能参数的位置，建立客户端性能参数矩阵，根据客户端各性能参数的类型将所述客户端参数矩阵中的元素进行重新排列，使每一列的参数类型一致，得到修正后的客户端参数矩阵，其中，
25 所述客户端参数矩阵的元素为所述性能参数；

具体的，在客户端性能参数表中，各个性能参数的位置是根据性能参数的首字符进行排列的，在建立客户端性能参数矩阵时，可以预先设置为 m 行、n 列，其中 m 和 n 为大于 1 的自然数，然后根据实际性能参数的数量对 m 和 n 进行调节以保证客户端性能参数矩阵中每一个位置非空。

30 其中，参数的类型主要有工作温度、硬盘使用率等等。

S104、对所述修正后的客户端参数矩阵二值化处理后得到二值化的客户端参数矩阵；

具体的，在对修正后的客户端参数矩阵二值化处理时，可以先对所述修正后的客户端矩阵进行分块，分割成数个子矩阵，然后对每一个子矩阵进行二值化处理后再进行汇总。

5 S105、抽取所述二值化的客户端参数矩阵中的首数字为“1”的元素，作为所述客户端请求信息的特征代码，获取所述过滤规则中的特征值，建立所述特征代码与所述特征值的对应关系；

10 具体的，将所述过滤规则中的特征值进行二值化转换后得到二值化的过滤规则特征值，将客户端请求信息的特征代码和二值化的过滤规则特征值进行比较，若相似度大于 90%，则认为所述客户端请求信息的特征代码对应所述二值化的过滤规则特征值。

S106、根据所述特征代码与所述特征值的对应关系对所述客户端请求信息进行过滤。

本实施例，通过对客户端请求信息对应的过滤规则进行有效识别，从而增强了对客户端系统资源分配的准确性。

15 在一个实施例中，所述 S2，获取过滤后的客户端请求信息，发送所述过滤后的客户端请求信息于服务器端，获取服务器端的熔断数据，分析所述熔断数据，得到熔断阈值，包括：

获取过滤后的客户端请求信息，赋予所述过滤后的客户端请求信息以负载系数；

20 具体的，在对过滤后的客户端请求信息赋予负载系数时，主要是根据所述客户端请求信息中所包含的数据量进行赋予，即客户端请求信息中的数据量大则负载系数大，信息量小则负载系数小。

根据所述负载系数的大小，将所述过滤后的客户端请求信息按负载系数由大到小的顺序排序；

25 将负载系数最大的所述客户端请求信息发送到所述服务器端的模拟器中进行模拟熔断实验；

其中，服务器端的模拟器在进行熔断实验时，模拟器根据客户端熔断的历史数据设置一个熔断的上限，当接收到的数据量大于所述上限时则发生熔断。

30 若所述负载系数最大的客户端请求信息在发送到所述模拟器中后没有发生熔断，则依次投入负载系数次最大的客户端请求信息于所述模拟器中，直到发生熔断；

获取熔断发生的起始时间节点，将所述起始时间节点和熔断对应的熔断类型打包形成熔断结果；

其中，熔断类型包括硬件熔断和软件熔断。

抽取所述熔断结果中的不同熔断类型发生的次数，根据所述不同熔断类型发生的次数设定不同的熔断阈值。

5 本实施例，通过对熔断情况进行分析得到不同类型熔断的熔断阈值，从而能够更好的分析系统资源分配情况。

图 3 为本申请在一个实施例中的一种系统资源分配方法中的系统运行监控过程示意图，如图所示，所述 S3，根据所述熔断阈值，设定客户端系统资源分配方案，监控客户端系统的工作状态，根据所述客户端系统的工作状态修正所述客户端系统资源分配方案，包括：

10 S301、获取所述熔断阈值，建立所述熔断阈值与客户端系统组件的映射关系；

具体的，获取所有的所述熔断阈值，建立熔断阈值矩阵，熔断阈值矩阵中的元素为熔断阈值；获取所述熔断阈值矩阵中每一个元素对应的错误发生次数最多的类型，查询所述错误发生次数最多的类型所对应的系统组件；从而建立
15 所述熔断阈值与所述系统组件的一一映射关系。

S302、根据所述熔断阈值的数值大小，设定客户端系统资源的分配方案，监控客户端系统的工作状态，获取熔断发生的次数，根据所述熔断发生的次数，修正所述客户端系统资源分配方案。

具体的，若所述熔断阈值数值大则相应的增加系统的容量，反之减少系统的容量；若熔断的次数超出根据历史数据得到的次数阈值，则对系统资源分配
20 方案进行调整。

本实施例，通过对客户端系统工作状态进行监控，从而调整资源分配方案使其更符合实际工作状态。

在一个实施例中，所述 S106、根据所述特征代码与所述特征值的对应关系
25 对所述客户端请求信息进行过滤，包括：

将所述客户端请求信息进行向量分段，获得 k 个待测向量；

具体的，在对所述请求信息进行向量分段时，可以根据所述客户端请求信息中的特殊字符进行分割，特征字符可以是“、”，“；”等标点，其中 k 的取值范围为 2~n，n 为大于 2 的正整数。

30 将所述 k 个待测向量分别与预设的 k 个检测向量中相应的检测向量进行内积运算，获得 k 个运算结果；

获取所述 k 个运算结果中的值最大的运算结果和所述值最大的运算结果对应的所述检测向量的标识；

获取与所述标识对应的预设的信息过滤器，应用所述信息过滤器对所述客户端请求信息进行过滤。

具体的，在服务器端设置有数个信息过滤器，每一个信息过滤器是根据信息的类型进行设置的，不同类型的信息过滤器具有不同的编号，将所述标识与
5 所述编号进行一致性比对，比对一致后，应用所述信息过滤器对所述客户端请求信息进行过滤。

本实施例，通过客户端请求信息进行有效过滤，从而增强了系统资源分配的准确度。

在一个实施例中，所述获取过滤后的客户端请求信息，赋予所述过滤后的
10 客户端请求信息以负载系数，包括：

获取预设的信息抓取阈值，当所述过滤后的客户端请求信息的数据量大于所述信息抓取阈值时，对所述过滤后的客户端请求信息进行抓取；

对抓取到的客户端请求信息进行特征字符查询，根据预设的所述特征字符与信息类别的对应关系，确定所述客户端请求信息的信息类别，其中所述特征
15 字符与所述信息类别的对应关系存储在数据库中；

具体的，所述特征字符主要是指客户端请求信息中的数字，根据所述数字查询信息类别列表，从信息类别列表中抽取出具有所述数字的信息类别。

根据所述信息类别对应的负载数值，赋予所述过滤后的客户端请求信息以不同的负载系数，负载系数的范围为 0.5~1。

具体的，在信息类别列表中查询出所述信息类别的所在行，在所述所在行中记录有所述信息类别对应的负载数值；即每一种信息类别都对应一个不同的负载数值。

本实施例，通过对不同类别的请求信息赋予不同的负载系数，从而能够更好的对熔断过程进行分析。

在一个实施例中，所述 S302、根据所述熔断阈值的数值大小，设定客户端系统资源的分配方案，监控客户端系统的工作状态，获取熔断发生的次数，根据所述熔断发生的次数，修正所述客户端系统资源分配方案，包括：

获取所述熔断阈值与所述客户端系统组件的映射关系，根据所述熔断阈值的大小，赋予所述客户端系统组件以不同的资源配置；

具体的，不同类型的熔断阈值对应着一个客户端组件，熔断阈值大则对应的客户端系统组件分配到的资源多，熔断阈值小则对应的客户端系统组件分配到的资源少。

获取预设的监控区间，在所述监控区间内，获取所述客户端系统发生熔断的次数；

具体的，预设的监控区间是根据熔断发生的历史数据设定的，比如某一客户端发生熔断的历史频率为 3 次/小时，则可以设定监控区间为 1 小时或者 2 小时。

若熔断发生的次数大于预设的次数阈值，则对所述熔断阈值对应的客户端系统组件进行容量扩展；

若熔断发生的次数小于预设的次数阈值，则对所述熔断阈值对应的客户端系统组件进行容量缩减。

本实施例，通过客户端实际发生熔断的情况对客户端系统组件进行合理分配，使得客户端系统资源分配方案更加优化。

在一个实施例中，提出了一种系统资源分配装置，如图 4 所示，包括如下模块：

请求信息过滤模块 41，设置为获取客户端请求信息，抽取所述客户端请求信息中的特征代码，获取所述特征代码与预设的过滤规则的对应关系，根据所述对应关系对所述客户端请求信息进行过滤；

熔断阈值生成模块 42，设置为获取过滤后的客户端请求信息，发送所述过滤后的客户端请求信息于服务器端，获取服务器端的熔断数据，分析所述熔断数据，得到熔断阈值；

系统运行监控模块 43，设置为根据所述熔断阈值，设定客户端系统资源分配方案，监控客户端系统的工作状态，根据所述客户端系统的工作状态修正所述客户端系统资源分配方案。

一种计算机设备，包括存储器和处理器，所述存储器中存储有计算机可读指令，所述计算机可读指令被所述处理器执行时，使得所述处理器执行上述各实施例中所述系统资源分配方法的步骤。

一种存储有计算机可读指令的存储介质，所述计算机可读指令被一个或多个处理器执行时，使得一个或多个处理器执行上述各实施例中所述系统资源分配方法的步骤。所述存储介质可以为非易失性存储介质，也可以为易失性存储介质，具体本申请不做限定。

以上所述实施例仅表达了本申请一些示例性实施例，其中描述较为具体和

详细，但并不能因此而理解为对本申请专利范围的限制。应当指出的是，对于本领域的普通技术人员来说，在不脱离本申请构思的前提下，还可以做出若干变形和改进，这些都属于本申请的保护范围。因此，本申请专利的保护范围应以所附权利要求为准。

1、一种系统资源分配方法，其中，包括：

获取客户端请求信息，抽取所述客户端请求信息中的特征代码，获取所述特征代码与预设的过滤规则的对应关系，根据所述对应关系对所述客户端请求信息进行过滤；

5 获取过滤后的客户端请求信息，发送所述过滤后的客户端请求信息于服务器端，获取服务器端的熔断数据，分析所述熔断数据，得到熔断阈值；

根据所述熔断阈值，设定客户端系统资源分配方案，监控客户端系统的工作状态，根据所述客户端系统的工作状态修正所述客户端系统资源分配方案。

2、根据权利要求 1 所述的系统资源分配方法，其中，所述获取客户端请求
10 信息，抽取所述客户端请求信息中的特征代码，获取所述特征代码与预设的过滤规则的对应关系，根据所述对应关系对所述客户端请求信息进行过滤，包括：

获取发出所述客户端请求信息的客户端 IP 地址，将所述客户端 IP 地址进行 DNS 解析得到 DNS 解析代码；

15 根据所述 DNS 解析代码与预设的 DNS 解析代码表之间的映射关系，获得所述客户端 IP 地址对应的客户端性能参数表；

按照所述客户端性能参数表中的各性能参数的位置，建立客户端性能参数矩阵，根据客户端各性能参数的类型将所述客户端参数矩阵中的元素进行重新排列，使每一列的参数类型一致，得到修正后的客户端参数矩阵，其中，所述客户端参数矩阵的元素为所述性能参数；

20 对所述修正后的客户端参数矩阵二值化处理后得到二值化的客户端参数矩阵；

抽取所述二值化的客户端参数矩阵中的首数字为“1”的元素，作为所述客户端请求信息的特征代码，获取所述过滤规则中的特征值，建立所述特征代码与所述特征值的对应关系；

25 根据所述特征代码与所述特征值的对应关系对所述客户端请求信息进行过滤。

3、根据权利要求 1 所述的系统资源分配方法，其中，所述获取过滤后的客户端请求信息，发送所述过滤后的客户端请求信息于服务器端，获取服务器端的熔断数据，分析所述熔断数据，得到熔断阈值，包括：

30 获取过滤后的客户端请求信息，赋予所述过滤后的客户端请求信息以负载系数；

根据所述负载系数的大小，将所述过滤后的客户端请求信息按负载系数由大到小的顺序排序；

将负载系数最大的所述客户端请求信息发送到所述服务器端的模拟器中进行模拟熔断实验；

若所述负载系数最大的客户端请求信息在发送到所述模拟器中后没有发生熔断，则依次投入负载系数次最大的客户端请求信息于所述模拟器中，直到发生熔断；

获取熔断发生的起始时间节点，将所述起始时间节点和熔断对应的熔断类型打包形成熔断结果；

抽取所述熔断结果中的不同熔断类型发生的次数，根据所述不同熔断类型发生的次数设定不同的熔断阈值。

4、根据权利要求 1 所述的系统资源分配方法，其中，所述根据所述熔断阈值，设定客户端系统资源分配方案，监控客户端系统的工作状态，根据所述客户端系统的工作状态修正所述客户端系统资源分配方案，包括：

获取所述熔断阈值，建立所述熔断阈值与客户端系统组件的映射关系；

根据所述熔断阈值的数值大小，设定客户端系统资源的分配方案，监控客户端系统的工作状态，获取熔断发生的次数，根据所述熔断发生的次数，修正所述客户端系统资源分配方案。

5、根据权利要求 2 所述的系统资源分配方法，其中，所述根据所述特征代码与所述特征值的对应关系对所述客户端请求信息进行过滤，包括：

将所述客户端请求信息进行向量分段，获得 k 个待测向量；

将所述 k 个待测向量分别与预设的 k 个检测向量中相应的检测向量进行内积运算，获得 k 个运算结果；

获取所述 k 个运算结果中的值最大的运算结果和所述值最大的运算结果对应的所述检测向量的标识；

获取与所述标识对应的预设的信息过滤器，应用所述信息过滤器对所述客户端请求信息进行过滤。

6、根据权利要求 3 所述的系统资源分配方法，其中，所述获取过滤后的客户端请求信息，赋予所述过滤后的客户端请求信息以负载系数，包括：

获取预设的信息抓取阈值，当所述过滤后的客户端请求信息的数据量大于所述信息抓取阈值时，对所述过滤后的客户端请求信息进行抓取；

对抓取到的客户端请求信息进行特征字符查询，根据预设的所述特征字符与信息类别的对应关系，确定所述客户端请求信息的信息类别，其中所述特征字符与所述信息类别的对应关系存储在数据库中；

根据所述信息类别对应的负载数值，赋予所述过滤后的客户端请求信息以不同的负载系数，负载系数的范围为 0.5~1。

7、根据权利要求 4 所述的系统资源分配方法，其中，所述根据所述熔断阈值的数值大小，设定客户端系统资源的分配方案，监控客户端系统的工作状态，
5 获取熔断发生的次数，根据所述熔断发生的次数，修正所述客户端系统资源分配方案，包括：

获取所述熔断阈值与所述客户端系统组件的映射关系，根据所述熔断阈值的大小，赋予所述客户端系统组件以不同的资源配置；

10 获取预设的监控区间，在所述监控区间内，获取所述客户端系统发生熔断的次数；

若熔断发生的次数大于预设的次数阈值，则对所述熔断阈值对应的客户端系统组件进行容量扩展；

若熔断发生的次数小于预设的次数阈值，则对所述熔断阈值对应的客户端系统组件进行容量缩减。

15 8、一种系统资源分配装置，其中，包括：

请求信息过滤模块，设置为获取客户端请求信息，抽取所述客户端请求信息中的特征代码，获取所述特征代码与预设的过滤规则的对应关系，根据所述对应关系对所述客户端请求信息进行过滤；

20 熔断阈值生成模块，设置为获取过滤后的客户端请求信息，发送所述过滤后的客户端请求信息于服务器端，获取服务器端的熔断数据，分析所述熔断数据，得到熔断阈值；

系统运行监控模块，设置为根据所述熔断阈值，设定客户端系统资源分配方案，监控客户端系统的工作状态，根据所述客户端系统的工作状态修正所述客户端系统资源分配方案。

25 9、根据权利要求 8 所述的系统资源分配装置，其中，所述请求信息过滤模块还设置为：

获取发出所述客户端请求信息的客户端 IP 地址，将所述客户端 IP 地址进行 DNS 解析得到 DNS 解析代码；

30 根据所述 DNS 解析代码与预设的 DNS 解析代码表之间的映射关系，获得所述客户端 IP 地址对应的客户端性能参数表；

按照所述客户端性能参数表中的各性能参数的位置，建立客户端性能参数矩阵，根据客户端各性能参数的类型将所述客户端参数矩阵中的元素进行重新

排列，使每一列的参数类型一致，得到修正后的客户端参数矩阵，其中，所述客户端参数矩阵的元素为所述性能参数；

对所述修正后的客户端参数矩阵二值化处理后得到二值化的客户端参数矩阵；

- 5 抽取所述二值化的客户端参数矩阵中的首数字为“1”的元素，作为所述客户端请求信息的特征代码，获取所述过滤规则中的特征值，建立所述特征代码与所述特征值的对应关系；

根据所述特征代码与所述特征值的对应关系对所述客户端请求信息进行过滤。

- 10 10、根据权利要求8所述的系统资源分配装置，其中，所述熔断阈值生成模块还设置为：

获取过滤后的客户端请求信息，赋予所述过滤后的客户端请求信息以负载系数；

- 15 根据所述负载系数的大小，将所述过滤后的客户端请求信息按负载系数由大到小的顺序排序；

将负载系数最大的所述客户端请求信息发送到所述服务器端的模拟器中进行模拟熔断实验；

- 20 若所述负载系数最大的客户端请求信息在发送到所述模拟器中后没有发生熔断，则依次投入负载系数次最大的客户端请求信息于所述模拟器中，直到发生熔断；

获取熔断发生的起始时间节点，将所述起始时间节点和熔断对应的熔断类型打包形成熔断结果；

抽取所述熔断结果中的不同熔断类型发生的次数，根据所述不同熔断类型发生的次数设定不同的熔断阈值。

- 25 11、根据权利要求8所述的系统资源分配装置，其中，所述系统运行监控模块还设置为：

获取所述熔断阈值，建立所述熔断阈值与客户端系统组件的映射关系；

- 30 根据所述熔断阈值的数值大小，设定客户端系统资源的分配方案，监控客户端系统的工作状态，获取熔断发生的次数，根据所述熔断发生的次数，修正所述客户端系统资源分配方案。

12、根据权利要求9所述的系统资源分配装置，其中，所述请求信息过滤模块还设置为：

将所述客户端请求信息进行向量分段，获得k个待测向量；

将所述 k 个待测向量分别与预设的 k 个检测向量中相应的检测向量进行内积运算，获得 k 个运算结果；

获取所述 k 个运算结果中的值最大的运算结果和所述值最大的运算结果对应的所述检测向量的标识；

- 5 获取与所述标识对应的预设的信息过滤器，应用所述信息过滤器对所述客户端请求信息进行过滤。

13、根据权利要求 10 所述的系统资源分配装置，其中，所述熔断阈值生成模块还设置为：

- 10 获取预设的信息抓取阈值，当所述过滤后的客户端请求信息的数据量大于所述信息抓取阈值时，对所述过滤后的客户端请求信息进行抓取；

对抓取到的客户端请求信息进行特征字符查询，根据预设的所述特征字符与信息类别的对应关系，确定所述客户端请求信息的信息类别，其中所述特征字符与所述信息类别的对应关系存储在数据库中；

- 15 根据所述信息类别对应的负载数值，赋予所述过滤后的客户端请求信息以不同的负载系数，负载系数的范围为 0.5~1。

14、根据权利要求 11 所述的系统资源分配装置，其中，所述系统运行监控模块还设置为：

获取所述熔断阈值与所述客户端系统组件的映射关系，根据所述熔断阈值的大小，赋予所述客户端系统组件以不同的资源配置；

- 20 获取预设的监控区间，在所述监控区间内，获取所述客户端系统发生熔断的次数；

若熔断发生的次数大于预设的次数阈值，则对所述熔断阈值对应的客户端系统组件进行容量扩展；

- 25 若熔断发生的次数小于预设的次数阈值，则对所述熔断阈值对应的客户端系统组件进行容量缩减。

15、一种计算机设备，包括存储器和处理器，所述存储器中存储有计算机可读指令，所述计算机可读指令被所述处理器执行时，使得所述处理器执行以下步骤：

- 30 获取客户端请求信息，抽取所述客户端请求信息中的特征代码，获取所述特征代码与预设的过滤规则的对应关系，根据所述对应关系对所述客户端请求信息进行过滤；

获取过滤后的客户端请求信息，发送所述过滤后的客户端请求信息于服务器端，获取服务器端的熔断数据，分析所述熔断数据，得到熔断阈值；

根据所述熔断阈值，设定客户端系统资源分配方案，监控客户端系统的工作状态，根据所述客户端系统的工作状态修正所述客户端系统资源分配方案。

16、一种存储有计算机可读指令的存储介质，所述计算机可读指令被一个或多个处理器执行时，使得一个或多个处理器执行以下步骤：

5 获取客户端请求信息，抽取所述客户端请求信息中的特征代码，获取所述特征代码与预设的过滤规则的对应关系，根据所述对应关系对所述客户端请求信息进行过滤；

获取过滤后的客户端请求信息，发送所述过滤后的客户端请求信息于服务器端，获取服务器端的熔断数据，分析所述熔断数据，得到熔断阈值；

10 根据所述熔断阈值，设定客户端系统资源分配方案，监控客户端系统的工作状态，根据所述客户端系统的工作状态修正所述客户端系统资源分配方案。

17、根据权利要求 16 所述的存储有计算机可读指令的存储介质，其中，所述计算机可读指令被一个或多个处理器执行时，使得一个或多个处理器执行所述获取客户端请求信息，抽取所述客户端请求信息中的特征代码，获取所述特征代码与预设的过滤规则的对应关系，根据所述对应关系对所述客户端请求信息进行过滤的步骤时，还执行如下步骤：

获取发出所述客户端请求信息的客户端 IP 地址，将所述客户端 IP 地址进行 DNS 解析得到 DNS 解析代码；

20 根据所述 DNS 解析代码与预设的 DNS 解析代码表之间的映射关系，获得所述客户端 IP 地址对应的客户端性能参数表；

按照所述客户端性能参数表中的各性能参数的位置，建立客户端性能参数矩阵，根据客户端各性能参数的类型将所述客户端参数矩阵中的元素进行重新排列，使每一列的参数类型一致，得到修正后的客户端参数矩阵，其中，所述客户端参数矩阵的元素为所述性能参数；

25 对所述修正后的客户端参数矩阵二值化处理后得到二值化的客户端参数矩阵；

抽取所述二值化的客户端参数矩阵中的首数字为“1”的元素，作为所述客户端请求信息的特征代码，获取所述过滤规则中的特征值，建立所述特征代码与所述特征值的对应关系；

30 根据所述特征代码与所述特征值的对应关系对所述客户端请求信息进行过滤。

18、根据权利要求 16 所述的存储有计算机可读指令的存储介质，其中，所述计算机可读指令被一个或多个处理器执行时，使得一个或多个处理器执行所

述获取过滤后的客户端请求信息，发送所述过滤后的客户端请求信息于服务器端，获取服务器端的熔断数据，分析所述熔断数据，得到熔断阈值的步骤时，还执行如下步骤：

获取过滤后的客户端请求信息，赋予所述过滤后的客户端请求信息以负载系数；

根据所述负载系数的大小，将所述过滤后的客户端请求信息按负载系数由大到小的顺序排序；

将负载系数最大的所述客户端请求信息发送到所述服务器端的模拟器中进行模拟熔断实验；

若所述负载系数最大的客户端请求信息在发送到所述模拟器中后没有发生熔断，则依次投入负载系数次最大的客户端请求信息于所述模拟器中，直到发生熔断；

获取熔断发生的起始时间节点，将所述起始时间节点和熔断对应的熔断类型打包形成熔断结果；

抽取所述熔断结果中的不同熔断类型发生的次数，根据所述不同熔断类型发生的次数设定不同的熔断阈值。

19、根据权利要求 16 所述的存储有计算机可读指令的存储介质，其中，所述计算机可读指令被一个或多个处理器执行时，使得一个或多个处理器执行所述根据所述熔断阈值，设定客户端系统资源分配方案，监控客户端系统的工作状态，根据所述客户端系统的工作状态修正所述客户端系统资源分配方案的步骤时，还执行如下步骤：

获取所述熔断阈值，建立所述熔断阈值与客户端系统组件的映射关系；

根据所述熔断阈值的数值大小，设定客户端系统资源的分配方案，监控客户端系统的工作状态，获取熔断发生的次数，根据所述熔断发生的次数，修正所述客户端系统资源分配方案。

20、根据权利要求 17 所述的存储有计算机可读指令的存储介质，其中，所述计算机可读指令被一个或多个处理器执行时，使得一个或多个处理器执行所述根据所述特征代码与所述特征值的对应关系对所述客户端请求信息进行过滤的步骤时，还执行如下步骤：

将所述客户端请求信息进行向量分段，获得 k 个待测向量；

将所述 k 个待测向量分别与预设的 k 个检测向量中相应的检测向量进行内积运算，获得 k 个运算结果；

获取所述 k 个运算结果中的值最大的运算结果和所述值最大的运算结果对应的所述检测向量的标识；

获取与所述标识对应的预设的信息过滤器，应用所述信息过滤器对所述客户端请求信息进行过滤。

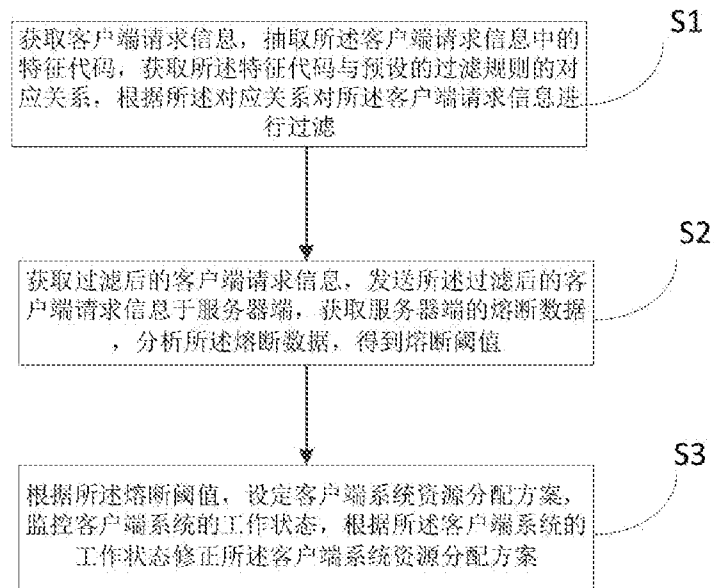


图 1

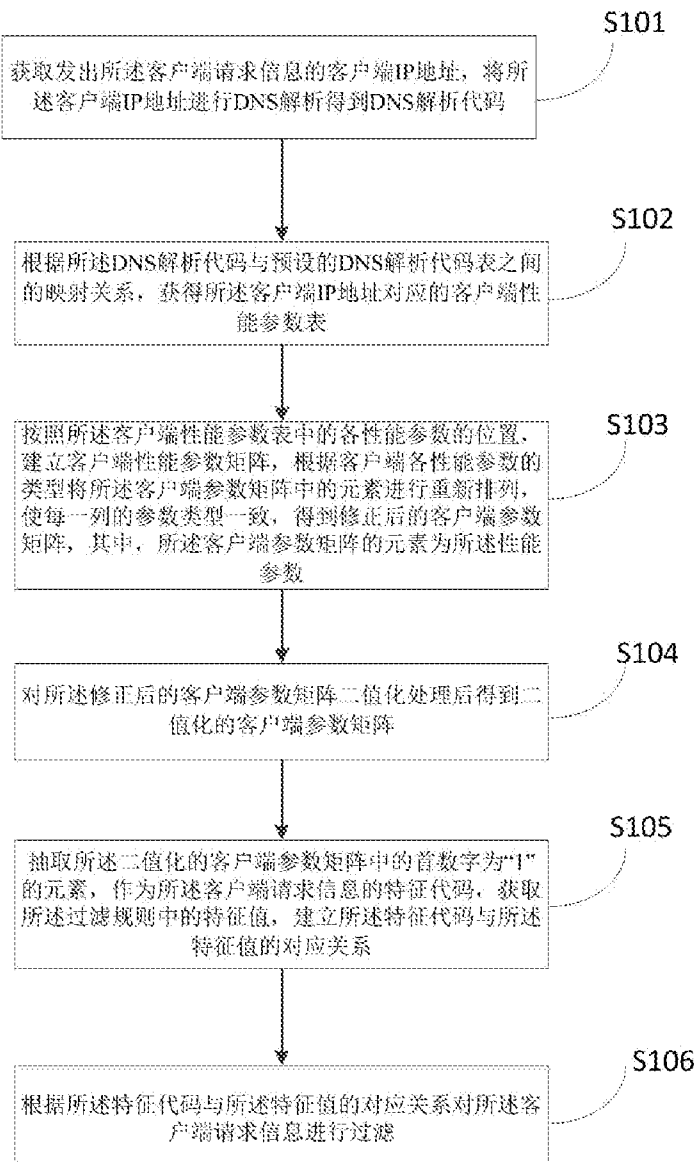


图 2

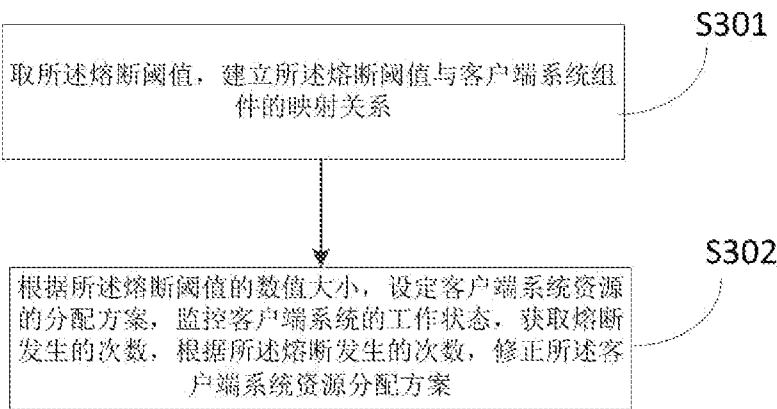


图 3

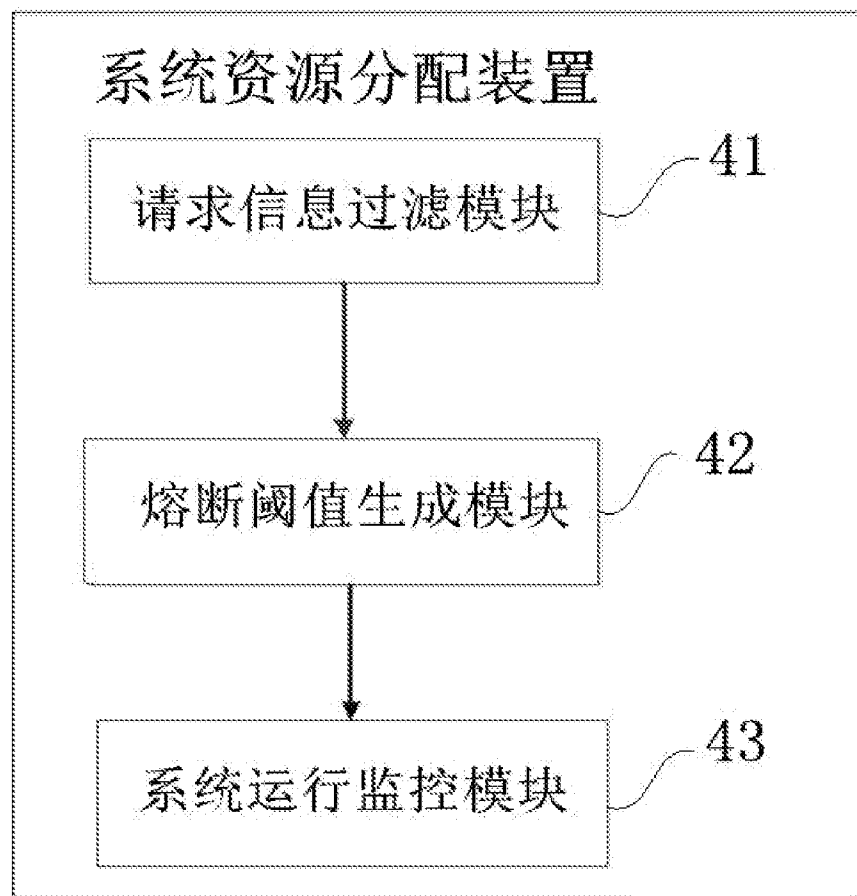


图 4

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2019/120525

A. CLASSIFICATION OF SUBJECT MATTER

G06F 9/50(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)

CNTXT; CNABS; CNKI; TWTXT; TWABS; USTXT; VEN; EPTXT; WOTXT: 资源, 调度, 分配, 熔断, 监控, 过载, 负载, 阈值, resource, schedule, allocation, fuse, monitored, overload, threshold

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
PX	CN 109918196 A (SHENZHEN ONECONNECT TECHNOLOGY CO., LTD.) 21 June 2019 (2019-06-21) entire document	1-20
Y	CN 107171828 A (SI-TECH INFORMATION TECHNOLOGY LTD.) 15 September 2017 (2017-09-15) description, paragraphs [0033]-[0052], and figure 1	1-20
Y	CN 108664321 A (GUANGDONG SHENMA SEARCH TECHNOLOGY CO., LTD.) 16 October 2018 (2018-10-16) description, paragraphs [0024]-[0048], and figure 1	1-20
A	CN 108667859 A (ZTE CORPORATION) 16 October 2018 (2018-10-16) entire document	1-20
A	CN 105808342 A (CHINA TELECOM CORPORATION LIMITED) 27 July 2016 (2016-07-27) entire document	1-20
A	US 2014007128 A1 (SCHROTH, Albrecht et al.) 02 January 2014 (2014-01-02) 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

20 February 2020

Date of mailing of the international search report

26 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

Authorized officer

Facsimile No. (86-10)62019451

Telephone No.

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/CN2019/120525

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
CN	109918196	A	21 June 2019	None			
CN	107171828	A	15 September 2017	None			
CN	108664321	A	16 October 2018	None			
CN	108667859	A	16 October 2018	WO	2018177042	A1	04 October 2018
CN	105808342	A	27 July 2016	CN	105808342	B	13 December 2019
US	2014007128	A1	02 January 2014	WO	2012141677	A1	18 October 2012
				US	9465660	B2	11 October 2016

国际检索报告

国际申请号

PCT/CN2019/120525

A. 主题的分类

G06F 9/50 (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))

CNTXT; CNABS; CNKI; TWTXT; TWABS; USTXT; VEN; EPTXT; WOTXT; 资源, 调度, 分配, 熔断, 监控, 过载, 负载, 阈值, resource, schedule, allocation, fuse, monitored, overload, threshold

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
PX	CN 109918196 A (深圳壹账通智能科技有限公司) 2019年 6月 21日 (2019 - 06 - 21) 全文	1-20
Y	CN 107171828 A (北京思特奇信息技术股份有限公司) 2017年 9月 15日 (2017 - 09 - 15) 说明书第33-52段, 图1	1-20
Y	CN 108664321 A (广东神马搜索科技有限公司) 2018年 10月 16日 (2018 - 10 - 16) 说明书第24-48段, 图1	1-20
A	CN 108667859 A (中兴通讯股份有限公司) 2018年 10月 16日 (2018 - 10 - 16) 全文	1-20
A	CN 105808342 A (中国电信股份有限公司) 2016年 7月 27日 (2016 - 07 - 27) 全文	1-20
A	US 2014007128 A1 (SCHROTH, Albrecht 等) 2014年 1月 2日 (2014 - 01 - 02) 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 2月 20日

国际检索报告邮寄日期

2020年 2月 26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

古志春

电话号码 86-(20)-28950386

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2019/120525

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	109918196	A	2019年 6月 21日	无			
CN	107171828	A	2017年 9月 15日	无			
CN	108664321	A	2018年 10月 16日	无			
CN	108667859	A	2018年 10月 16日	WO	2018177042	A1	2018年 10月 4日
CN	105808342	A	2016年 7月 27日	CN	105808342	B	2019年 12月 13日
US	2014007128	A1	2014年 1月 2日	WO	2012141677	A1	2012年 10月 18日
				US	9465660	B2	2016年 10月 11日