

(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
2016 年 11 月 10 日 (10.11.2016)

WIPO | PCT

(10) 国际公布号
WO 2016/177258 A1

- (51) 国际专利分类号:
G06F 15/16 (2006.01)
- (21) 国际申请号: PCT/CN2016/079157
- (22) 国际申请日: 2016 年 4 月 13 日 (13.04.2016)
- (25) 申请语言: 中文
- (26) 公布语言: 中文
- (30) 优先权:
201510226967.6 2015 年 5 月 6 日 (06.05.2015) CN
- (71) 申请人: 阿里巴巴集团控股有限公司 (ALIBABA GROUP HOLDING LIMITED) [—/CN]; 英属开曼群岛大开曼乔治城资本大厦一座四层 847 号邮箱, Grand Cayman (KY)。
- (72) 发明人: 及
- (71) 申请人 (仅对美国): 李静 (LI, Jing) [CN/CN]; 中国浙江省杭州市余杭区文一西路 969 号 3 号楼 5 楼阿里巴巴集团法务部, Zhejiang 311121 (CN)。李炉阳 (LI, Luyang) [CN/CN]; 中国浙江省杭州市余杭区文一西路 969 号 3 号楼 5 楼阿里巴巴集团法务部, Zhejiang 311121 (CN)。王丰金 (WANG, Fengjin) [CN/CN]; 中国浙江省杭州市余杭区文一西路 969 号 3 号楼 5 楼阿里巴巴集团法务部, Zhejiang 311121 (CN)。

- (74) 代理人: 北京国昊天诚知识产权代理有限公司 (CO-HORIZON INTELLECTUAL PROPERTY INC.); 中国北京市朝阳区小关北里甲 2 号渔阳置业大厦 B 座 605, Beijing 100029 (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, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, KN, KP, KR, 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, 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))。

(54) Title: COMPUTING RESOURCE MANAGEMENT METHOD AND APPARATUS

(54) 发明名称: 计算资源管理方法和装置

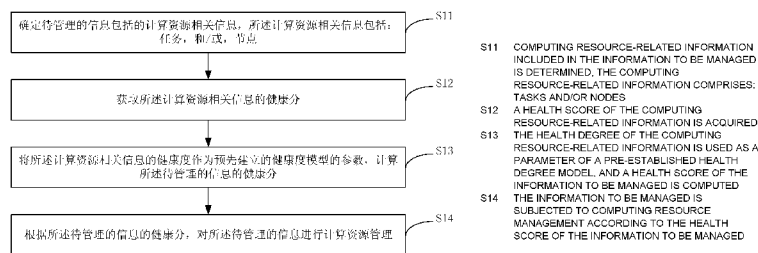


图 1

(57) Abstract: A computing resource management method and apparatus. The computing resource management method comprises: determining the computing resource-related information included in the information to be managed, the computing resource-related information comprising: tasks and/or nodes (S11); acquiring a health score of the computing resource-related information (S12); using the health degree of the computing resource-related information as a parameter of a pre-established health degree model, and computing a health score of the information to be managed (S13); and subjecting the information to be managed to computing resource management according to the health score of the information to be managed (S14). The method can realize a long-term optimization mechanism without confirmations from men and is capable of optimizing not only large tasks, but also other tasks.

(57) 摘要: 一种计算资源管理方法和装置, 该计算资源管理方法包括确定待管理的信息包括的计算资源相关信息, 所述计算资源相关信息包括: 任务, 和/或, 节点(S11); 获取所述计算资源相关信息的健康分(S12); 将所述计算资源相关信息的健康度作为预先建立的健康度模型的参数, 计算所述待管理的信息的健康分(S13); 根据所述待管理的信息的健康分, 对所述待管理的信息进行计算资源管理(S14)。该方法能够实现的长期优化机制, 不需要人为确认, 以及不仅可以优化大任务还可以优化其他任务。



WO 2016/177258 A1

计算资源管理方法和装置

技术领域

5 本申请涉及数据处理技术领域，尤其涉及一种计算资源管理方法和装置。

背景技术

 随着大数据时代的到来，各企业都会遇到一个问题：计算规模以井喷模式快速增长，而计算资源的成本压力也会随之快速增长。并且随着计算量和业务
10 复杂度越来越高，计算资源管理的难度也越来越大。

 现有技术中，一般是整个集群遇到计算资源瓶颈了，然后新加机器又有困难的时候，才会发起一次清理，一般的清理都是先找到计算资源消耗较大的 N 个大任务，然后从大任务着手，看看能不能进行逻辑优化以节省计算资源或者
对无用的计算任务直接清理掉。危机过去以后，就又任其照原样发展了。

15 但是，这种方式存在如下问题：不可能形成长期有效的优化机制；每次要花费大量的人力物力去确认；清理不可能彻底，计算资源消耗较大的 N 个大任务只能解决短期的危机，还有大量存在其他不合理情况的任务无法得到有效的治理。

20 发明内容

 本申请旨在至少在一定程度上解决相关技术中的技术问题之一。

 为此，本申请的一个目的在于提出一种计算资源管理方法，该方法可以实现的长期优化机制，不需要人为确认，以及不仅可以优化大任务还可以优化其他任务。

25 本申请的另一个目的在于提出一种计算资源管理装置。

 为达到上述目的，本申请第一方面实施例提出的计算资源管理方法，包括：确定待管理的信息包括的计算资源相关信息，所述计算资源相关信息包括：任务，和/或，节点；获取所述计算资源相关信息的健康分；将所述计算资源相关信息的健康度作为预先建立的健康度模型的参数，计算所述待管理的信息的健康分；根据所述待管理的信息的健康分，对所述待管理的信息进行计算资源
30 管理。

 本申请第一方面实施例提出的计算资源管理方法，通过采用健康度模型对

待管理的信息进行计算资源管理，可以在需要管理大规模计算资源时，有一个统一的标准，从而可以形成长期有效的机制，不需要确认，并且可以不仅优化大任务，还可以优化其他任务，因此可以实现可行且有效的计算资源管理，可以应用到大规模数据管理场景下。

5 为达到上述目的，本申请第二方面实施例提出的计算资源管理装置，包括：确定模块，用于确定待管理的信息包括的计算资源相关信息，所述计算资源相关信息包括：任务，和/或，节点；获取模块，用于获取所述计算资源相关信息的健康分；计算模块，用于将所述计算资源相关信息的健康度作为预先建立的健康度模型的参数，计算所述待管理的信息的健康分；管理模块，用于根据
10 所述待管理的信息的健康分，对所述待管理的信息进行计算资源管理。

本申请第二方面实施例提出的计算资源管理装置，通过采用健康度模型对待管理的信息进行计算资源管理，可以在需要管理大规模计算资源时，有一个统一的标准，从而可以形成长期有效的机制，不需要确认，并且可以不仅优化大任务，还可以优化其他任务，因此可以实现可行且有效的计算资源管理，可
15 以应用到大规模数据管理场景下。

本申请附加的方面和优点将在下面的描述中部分给出，部分将从下面的描述中变得明显，或通过本申请的实践了解到。

附图说明

20 本申请上述的和/或附加的方面和优点从下面结合附图对实施例的描述中将变得明显和容易理解，其中：

图 1 是本申请一实施例提出的计算资源管理方法的流程示意图；

图 2 是本申请另一实施例提出的计算资源管理装置的结构示意图。

25 具体实施方式

下面详细描述本申请的实施例，所述实施例的示例在附图中示出，其中自始至终相同或类似的标号表示相同或类似的模块或具有相同或类似功能的模块。下面通过参考附图描述的实施例是示例性的，仅用于解释本申请，而不能理解为对本申请的限制。相反，本申请的实施例包括落入所附加权利要求书的精神和内涵范围内的所有变化、修改和等同物。
30

图 1 是本申请一实施例提出的计算资源管理方法的流程示意图，该方法包括：

S11: 确定待管理的信息包括的计算资源相关信息, 所述计算资源相关信息包括: 任务, 和/或, 节点。

其中, 待管理的信息包括但不限于: 数据开发者, 应用程序, 业务单元 (Business Unit, BU)。

5 每个待管理的信息可以包括: 一个或多个任务; 或者, 一个或多个节点; 或者, 一个或多个任务, 以及, 一个或多个节点。其中, 多个是指至少两个。

S12: 获取所述计算资源相关信息的健康分。

其中, 本实施例中, 健康分包括: 对应任务的任务级健康分, 以及, 对应节点的节点级健康分。任务级可以称为开放数据处理服务 (Open Data Processing Service, ODPS) 任务级, 节点级可以称为天网节点级, 天网是一种调度系统。

本实施例以扣分制为例, 当出现一个扣分项时, 从默认的数值开始扣除扣分项对应的数值, 最少扣为 0。本实施例中以默认值是 100 为例。

一个实施例中, 当所述计算资源相关信息是任务时, 根据如下扣分项中的至少一项获取每种任务的健康分:

倾斜度数值, 缺责任人数值, 简单加工数值, 暴力扫描数值, 重复计算数值, 产出为空数值, 映射 (map) 参数设置不合理数值。

具体的, (1) 倾斜度数值 a_1 的计算公式是:

$a_1 = \text{odps_task 倾斜度} = 100 - \text{sum}(\text{fuxi_task 倾斜度}) \in [0, \infty)$;

20 其中, 每个任务 (odps_task) 包括至少一个 fuxi_task, sum() 表示求和运算,

$\text{fuxi_task 倾斜度} = \max(\text{runtime}) / \text{avg}(\text{runtime}) - 3$;

其中, 每个 fuxi_task 包括至少一个 fuxi_instance, runtime 是每个 fuxi_instance 的运行时间, max() 是最大值运算, avg() 是均值运算。

25 由于 fuxi_instance 同时满足如下的两个条件才算数据倾斜:

$\max(\text{runtime}) > 5 \text{ min}$

$\max(\text{runtime}) > 3 * \text{avg}(\text{runtime})$

因此, 如果一个 fuxi_task 的 instance 数小于等于 3 时, 不构成数据倾斜。倾斜度的最大值和 instance 数有关。

30 (2) 缺责任人数值 a_2 的计算公式是:

$a_2 = 0$ 或者 $1 \in \{0, 1\}$, 其中, 当任务有责任人 (owner) 时, $a_2 = 0$, 无责任人时, $a_1 = 0$ 。

通常来讲，当任务有责任人时，就会有 biz_id。但是，本实施例中有责任人不限于有 biz_id 的情况，还包括有责任人无 biz_id 的情况。

(3) 简单加工数值 a3 的计算公式是：

$a3=0$ 或者 $1 \in \{0, 1\}$ ，其中，当是简单加工时， $a3=1$ ，否则 $a3=0$ 。

5 其中，当同时满足如下三点表明是简单加工：

结果表落地，也就是任务对应的操作是写操作，写操作例如：insert
overwrite, creat;

纯 select，也就是任务对应的操作没有 join，没有函数之类复杂的 sql；
where 里只有分区字段，其中，where 是任务对应的 sql 中的一个参数。

10 (4) 暴力扫描数值 a4 的计算公式是：

$a4 = \text{扫描的天数} - 31 \in [0, \infty)$;

其中，计算暴力扫描数值的任务同时满足如下条件：

扫描某个表大于一个月的分区；

$\text{avg}(\text{天分区大小}) > 500\text{M}$ ，或者， $\text{sum}(\text{分区大小}) > 100\text{G}$ ；

15 月任务不计暴力扫描。

其中，一个月采用 30 天，在计算 a4 时，二级及以上分区先汇总到天分区。

(5) 重复计算数值 a5 的计算公式是：

$a5 = \text{column 相似度} * \text{分区记录数相似度} * \text{输入表的相似度} \in [0, 1)$;

20 其中，column 相似度是两个字符串的余弦相似性，两个字符串是分别由
两个输入表的列组成的，两个输入表是当前比较的两个任务分别对应的取数据的表；

分区记录数相似度是 7-61 天的分区记录数在大于 1000 时，这些记录数的协方差；

25 输入表的相似度是两个输入表的余弦相似度，两个输入表是当前比较的两个任务分别对应的取数据的表。

(6) 产出为空数值 a6 的计算公式是：

$a6=0$ 或者 $1 \in \{0, 1\}$ ，其中，当任务产出的表或分区的记录数为 0，则 $a6=1$ ，
否则 $a6=0$ 。

30 需要说明的是，当一个任务为多路输出时，所有表或分区为空才算空，因此，上述的任务产出的表或分区的记录数是指任务产出的表或分区的所有记录数。

(7) 映射参数设置不合理数值 a7 的计算公式是：

$a7=100-\text{平均运行时长} \in [0, 100]$;

其中, 平均运行时长是指一个 fuxi_task 包括的所有 fuxi_instance 的运行时间 (runtime) 的平均值, 该 fuxi_task 需要同时满足如下条件:

仅对无依赖的第一层 fuxi_task 统计 (应该都是 map task);

5 仅统计 fuxi_instance 数大于 100 的 fuxi_task (这里的 100 也可以是小于 100 的值)。

因此, 按照上述计算公式, 如果平均运行时长小于 100s, 就认为 map 参数设置不合理, 属于 split size 设小了。

需要说明的是, 上述的运行时长是实际时间, 而不是 master 时间。另外,
10 如果一个 odps_task 有多个同时满足上述条件的 fuxi_task, 则取分数最低的作为这里的 $a7$ 。

以对应任务的上述各项均包括为例, 单个任务的健康分 C 的计算公式是:

$C = \text{greatest} (0, (100 - \text{least}(a1, 20) - 30 * a2 - 10 * a3 - \text{least}(a4/15, 15) - 10 * a5 - 10 * a6 - a7))^4 / 1000000$;

15 其中, greatest 为取最大值运算, least 是取最小值运算。

另一个实施例中, 当所述计算资源相关信息是任务时, 根据如下扣分项中的至少一项获取每种任务的健康分:

节点破基线数值, 依赖合理性数值, 责任人合理性数值, 任务出错天数数值, 产出表没有被读数值, 同源导入数值, 产出为空数值, 导入持续一致数值。

20 具体的, (1) 节点破基线数值 $b1$ 的计算公式是:

$b1 = \text{最近 10 天破基线的天数} \in [0, 10]$ 。

需要说明的是, 只与本节点对应基线的时间做比较, 不做上下游依赖的推导。比如: 基线时间 5 点的 B 任务依赖了基线时间 8 点的 A 任务, 如果 A 任务在 7 点完成, 则 A 任务不算破基线, 但由于 B 的产出时间必然晚于 7 点,
25 所以 B 任务算破基线。另外, 不计小时任务。

(2) 依赖合理性数值 $b2$ 的计算公式是:

$b2 = 0$ 或者 $1 \in \{0, 1\}$, 其中, 当依赖不合理时取值为 1, 否则为 0。

其中, 同时满足如下条件属于依赖不合理:

上游表由天网产出, 下游表通过血缘解析有依赖, 但是没配置, 且读到的
30 分区为 0。

(3) 责任人 (owner) 合理性数值 $b3$ 的计算公式是:

$b3 = 0, 1, 2$ 或者 $3 \in \{0, 1, 2, 3\}$, 其中,

和表 owner 一致, 任务 owner 且未离职: 0

和表 owner 不完全一致, 任务 owner 未离职: 1

和表 owner 一致, 任务 owner 但已离职: 2

和表 owner 不完全一致, 任务 owner 且已离职: 3

5 (4) 任务出错天数数值 b4 的计算公式是:

b4=近 10 天的任务出错天数, 具体取值:

天任务=近 10 天出错天数*1;

周任务=近 10 天出错天数*3;

月任务=近 10 天出错天数*5。

10 需要说明的是, 有出错记录就算 (修复了也要算、生产冒烟也算)。一天内出错多次, 只取一次。小时任务一天内只算一次。

(5) 产出表没有被读数值 b5 的计算公式是:

b5=0 或者 $1 \in \{0, 1\}$, 其中, 最近一个月产出表没有被读取记录则为 1, 否则为 0。

15 需要说明的是, 只考虑该任务自身, 不考虑其下游。新上线的任务给 10 天的豁免期, 即节点创建时间需要是当前时间 10 天前的才统计。

(6) 同源导入数值 b6 的计算公式是:

b6=0 或者 $1 \in \{0, 1\}$, 其中, 如果该节点与其他节点的导入源相同, 则取值 1, 否则为 0。

20 需要说明的是, 仅针对 datax 节点。

(7) 产出为空数值 b7 的计算公式是:

b7=0 或者 $1 \in \{0, 1\}$, 其中, 如果该节点导入的数据量为 0, 则取值 1, 否则为 0。

需要说明的是, 仅针对 datax 节点。

25 (8) 导入持续一致数值 b8 的计算公式是:

b8=0 或者 $1 \in \{0, 1\}$, 其中, 如果持续 7 天导入的记录数一致且大于 1000 行 (具体行数也可以设置), 则取值 1, 否则为 0。

需要说明的是, 仅针对 datax 节点。

以上述对应节点的各项均包括为例, 单个节点的健康分的计算公式是:

30
$$D = \text{greatest}(0, (100 - \text{least}(b1 * (p+1) * 10, 100) - 30 * b2 - 10 * b3 - 10 * b4 - 50 * b5 - 50 * b6 - 50 * b7 - 50 * b8))^4 / 1000000;$$

其中, greatest 为取最大值运算, least 是取最小值运算;

p 是预设的天网优先级 $\in [0, 8]$ 。

S13: 将所述计算资源相关信息的健康度作为预先建立的健康度模型的参数, 计算所述待管理的信息的健康分。

具体的, 所述将所述计算资源相关信息的健康度作为预先建立的健康度模型的参数, 计算所述待管理的信息的健康分, 包括:

采用如下计算公式, 计算所述待管理的信息的健康分 E:

$$E=y/(x+y)*((\sum(C*n)/\sum n)^4/1000000)+x/(x+y)*((\sum(D*m)/\sum m)^4/1000000);$$

其中, n 是任务级别的权重, 具体可以用如下公式计算得到: $\log(\text{任务运行时间}*\text{plan_cpu_cost}+1)$, plan_cpu_cost 由底层数据库抓取;

m 是节点级别的权重, 具体可以用如下公式计算得到: 节点的优先级+1, 其中, 节点的优先级可以由用户预设;

x 是有节点的任务数, y 是无节点的任务数, 其中, 通常来讲, 有节点的任务是指生产任务, 无节点的任务是指开发任务。由于有些人开发任务比较多比如分析师等, 有些人生产任务比较多比如专业的数据开发工程师等, 为避免
15 失衡所以把两者进行加权平均。

S14: 根据所述待管理的信息的健康分, 对所述待管理的信息进行计算资源管理。

其中, 可以根据预设的策略, 对待管理的信息进行计算资源管理。例如, 可以预先设置一个阈值, 当某个待管理的信息的健康分小于该阈值时, 确定该
20 待管理的信息是需要优化的信息, 对应需要优化的信息可以展示其包括的任务和节点的每个扣分项, 从而有针对性的进行优化。或者, 当待管理的信息是数据开发者时, 如果一个数据开发者的健康分大于阈值, 则可以奖励该数据开发者, 反之, 当一个数据开发者的健康分小于阈值时, 对该数据开发者进行惩罚, 从而可以调动数据开发者的优化积极性。

25 本实施例应用的集群包括但不限于: Hadoop 集群, odps 集群。

本实施例中, 通过采用健康度模型对待管理的信息进行计算资源管理, 可以在需要管理大规模计算资源时, 有一个统一的标准, 从而可以形成长期有效的机制, 不需要人为确认, 并且可以不仅优化大任务, 还可以优化其他任务, 因此可以实现可行且有效的计算资源管理, 可以应用到大规模数据管理场景
30 下。本实施例通过选择上述各种具体的任务级别和节点级别的扣分项, 可以非常适用于大规模的计算资源管理。

图 2 是本申请另一实施例提出的计算资源管理装置的结构示意图, 该

装置 20 包括：确定模块 21，获取模块 22，计算模块 23 和管理模块 24。

确定模块 21，用于确定待管理的信息包括的计算资源相关信息，所述计算资源相关信息包括：任务，和/或，节点；

其中，待管理的信息包括但不限于：数据开发者，应用程序，业务单元
5 (Business Unit, BU)。

每个待管理的信息可以包括：一个或多个任务；或者，一个或多个节点；
或者，一个或多个任务，以及，一个或多个节点。其中，多个是指至少两个。

获取模块 22，用于获取所述计算资源相关信息的健康分；

其中，本实施例中，健康分包括：对应任务的任务级健康分，以及，对应
10 节点的节点级健康分。任务级可以称为开放数据处理服务 (Open Data Processing Service, ODPS) 任务级，节点级可以称为天网节点级，天网是一种调度系统。

本实施例以扣分制为例，当出现一个扣分项时，从默认的数值开始扣除扣分项对应的数值，最少扣为 0。本实施例中以默认值是 100 为例。

15 一个实施例中，所述获取模块 22 具体用于：

当所述计算资源相关信息是任务时，根据如下扣分项中的至少一项获取每种任务的健康分：

倾斜度数值，缺责任人数值，简单加工数值，暴力扫描数值，重复计算数值，产出为空数值，映射参数设置不合理数值。

20 相应的，如果单个任务与上述各扣分项均相关，所述获取模块 22 具体用于采用如下计算公式计算单个任务的健康分：

$$C = \text{greatest} (0, (100 - \text{least}(a1, 20) - 30 * a2 - 10 * a3 - \text{least}(a4/15, 15) - 10 * a5 - 10 * a6 - a7))^4 / 10000000;$$

其中，greatest 为取最大值运算，least 是取最小值运算；

25 C 是单个任务的健康分，a1 是倾斜度数值，a2 是缺责任人数值，a3 是简单加工数值，a4 是暴力扫描数值，a5 是重复计算数值，a6 是产出为空数值，a7 是映射参数设置不合理数值。

其中，上述各数值的计算公式可以参见方法实施例中的相关描述，在此不再赘述。

30 另一个实施例中，所述获取模块 22 具体用于：

当所述计算资源相关信息是节点时，根据如下扣分项中的至少一项获取每种节点的健康分：

节点破基线数值, 依赖合理性数值, 责任人合理性数值, 任务出错天数数值, 产出表没有被读数值, 同源导入数值, 产出为空数值, 导入持续一致数值。

相应的, 如果单个任务与上述各扣分项均相关, 所述获取模块 22 具体用于采用如下计算公式计算单个节点的健康分:

5
$$D = \text{greatest}(0, (100 - \text{least}(b1 * (p+1) * 10, 100) - 30 * b2 - 10 * b3 - 10 * b4 - 50 * b5 - 50 * b6 - 50 * b7 - 50 * b8))^{4/10000000};$$

其中, greatest 为取最大值运算, least 是取最小值运算;

D 是单个节点的健康分, $b1$ 是节点破基线数值, $b2$ 是依赖合理性数值, $b3$ 是责任人合理性数值, $b4$ 是任务出错天数数值, $b5$ 是产出表没有被读数值, 10 $b6$ 是同源导入数值, $b7$ 是产出为空数值, $b8$ 是导入持续一致数值。

其中, 上述各数值的计算公式可以参见方法实施例中的相关描述, 在此不再赘述。

计算模块 23, 用于将所述计算资源相关信息的健康度作为预先建立的健康度模型的参数, 计算所述待管理的信息的健康分;

15 具体的, 所述计算模块 23 具体用于:

采用如下计算公式, 计算所述待管理的信息的健康分:

$$E = y / (x + y) * ((\sum(C * n) / \sum n)^{4/10000000}) + x / (x + y) * ((\sum(D * m) / \sum m)^{4/10000000});$$

其中, E 是待管理的信息的健康分;

n 是任务级别的权重, 具体可以用如下公式计算得到: $\log(\text{任务运行时间} * \text{plan_cpu_cost} + 1)$, plan_cpu_cost 由底层数据库抓取;

m 是节点级别的权重, 具体可以用如下公式计算得到: 节点的优先级+1, 其中, 节点的优先级可以由用户预设;

x 是有节点的任务数, y 是无节点的任务数。

25 管理模块 24, 用于根据所述待管理的信息的健康分, 对所述待管理的的信息进行计算资源管理。

其中, 可以根据预设的策略, 对待管理的信息进行计算资源管理。例如, 可以预先设置一个阈值, 当某个待管理的信息的健康分小于该阈值时, 确定该待管理的信息是需要优化的信息, 对应需要优化的信息可以展示其包括的任务和节点的每个扣分项, 从而有针对性的进行优化。或者, 当待管理的信息是数据开发者时, 如果一个数据开发者的健康分大于阈值, 则可以奖励该数据开发 30 者, 反之, 当一个数据开发者的健康分小于阈值时, 对该数据开发者进行惩罚, 从而可以调动数据开发者的优化积极性。

本实施例应用的集群包括但不限于：Hadoop 集群，odps 集群。

本实施例中，通过采用健康度模型对待管理的信息进行计算资源管理，可以在需要管理大规模计算资源时，有一个统一的标准，从而可以形成长期有效的机制，不需要人为确认，并且可以不仅优化大任务，还可以优化其他任务，因此可以实现可行且有效的计算资源管理，可以应用到大规模数据管理场景下。本实施例通过选择上述各种具体的任务级别和节点级别的扣分项，可以非常适用于大规模的计算资源管理。

需要说明的是，在本申请的描述中，术语“第一”、“第二”等仅用于描述目的，而不能理解为指示或暗示相对重要性。此外，在本申请的描述中，除非另有说明，“多个”的含义是指至少两个。

流程图中或在此以其他方式描述的任何过程或方法描述可以被理解为，表示包括一个或更多个用于实现特定逻辑功能或过程的步骤的可执行指令的代码的模块、片段或部分，并且本申请的优选实施方式的范围包括另外的实现，其中可以不按所示出或讨论的顺序，包括根据所涉及的功能按基本同时的方式或按相反的顺序，来执行功能，这应被本申请的实施例所属技术领域的技术人员所理解。

应当理解，本申请的各部分可以用硬件、软件、固件或它们的组合来实现。在上述实施方式中，多个步骤或方法可以用存储在存储器中且由合适的指令执行系统执行的软件或固件来实现。例如，如果用硬件来实现，和在另一实施方式中一样，可用本领域公知的下列技术中的任一项或他们的组合来实现：具有用于对数据信号实现逻辑功能的逻辑门电路的离散逻辑电路，具有合适的组合逻辑门电路的专用集成电路，可编程门阵列(PGA)，现场可编程门阵列(FPGA)等。

本技术领域的普通技术人员可以理解实现上述实施例方法携带的全部或部分步骤是可以通过程序来指令相关的硬件完成，所述的程序可以存储于一种计算机可读存储介质中，该程序在执行时，包括方法实施例的步骤之一或其组合。

此外，在本申请各个实施例中的各功能单元可以集成在一个处理模块中，也可以是各个单元单独物理存在，也可以两个或两个以上单元集成在一个模块中。上述集成的模块既可以采用硬件的形式实现，也可以采用软件功能模块的形式实现。所述集成的模块如果以软件功能模块的形式实现并作为独立的产品销售或使用，也可以存储在一个计算机可读取存储介质中。

上述提到的存储介质可以是只读存储器，磁盘或光盘等。

在本说明书的描述中，参考术语“一个实施例”、“一些实施例”、“示例”、“具体示例”、或“一些示例”等的描述意指结合该实施例或示例描述的具体特征、结构、材料或者特点包含于本申请的至少一个实施例或示例中。在本说明书5 书中，对上述术语的示意性表述不一定指的是相同的实施例或示例。而且，描述的具体特征、结构、材料或者特点可以在任何的一个或多个实施例或示例中以合适的方式结合。

尽管上面已经示出和描述了本申请的实施例，可以理解的是，上述实施例是示例性的，不能理解为对本申请的限制，本领域的普通技术人员在本申请的10 范围内可以对上述实施例进行变化、修改、替换和变型。

权 利 要 求 书

1、一种计算资源管理方法，其特征在于，包括：

5 确定待管理的信息包括的计算资源相关信息，所述计算资源相关信息包括：任务，和/或，节点；

获取所述计算资源相关信息的健康分；

将所述计算资源相关信息的健康度作为预先建立的健康度模型的参数，计算所述待管理的信息的健康分；

10 根据所述待管理的信息的健康分，对所述待管理的信息进行计算资源管理。

2、根据权利要求 1 所述的方法，其特征在于，所述获取所述计算资源相关信息的健康分，包括：

当所述计算资源相关信息是任务时，根据如下扣分项中的至少一项获取每种任务的健康分：

15 倾斜度数值，缺责任人数值，简单加工数值，暴力扫描数值，重复计算数值，产出为空数值，映射参数设置不合理数值。

3、根据权利要求 1 所述的方法，其特征在于，所述获取所述计算资源相关信息的健康分，包括：

20 当所述计算资源相关信息是节点时，根据如下扣分项中的至少一项获取每种节点的健康分：

节点破基线数值，依赖合理性数值，责任人合理性数值，任务出错天数数值，产出表没有被读数值，同源导入数值，产出为空数值，导入持续一致数值。

4、根据权利要求 2 所述的方法，其特征在于，如果单个任务与上述各扣分项均相关，所述获取所述计算资源相关信息的健康分，包括：

25 单个任务的健康分的计算公式是：

$$C = \text{greatest} (0, (100 - \text{least}(a1, 20) - 30 * a2 - 10 * a3 - \text{least}(a4/15, 15) - 10 * a5 - 10 * a6 - a7))^4 / 1000000;$$

其中，greatest 为取最大值运算，least 是取最小值运算；

30 C 是单个任务的健康分，a1 是倾斜度数值，a2 是缺责任人数值，a3 是简单加工数值，a4 是暴力扫描数值，a5 是重复计算数值，a6 是产出为空数值，a7 是映射参数设置不合理数值。

5、根据权利要求 3 所述的方法，其特征在于，如果单个节点与上述各扣

分项均相关，所述获取所述计算资源相关信息的健康分，包括：

单个节点的健康分的计算公式是：

$$D = \text{greatest}(0, (100 - \text{least}(b1 * (p+1) * 10, 100) - 30 * b2 - 10 * b3 - 10 * b4 - 50 * b5 - 50 * b6 - 50 * b7 - 50 * b8))^4 / 10000000;$$

5 其中，greatest 为取最大值运算，least 是取最小值运算；

D 是单个节点的健康分，b1 是节点破基线数值，b2 是依赖合理性数值，b3 是责任人合理性数值，b4 是任务出错天数数值，b5 是产出表没有被读数值，b6 是同源导入数值，b7 是产出为空数值，b8 是导入持续一致数值。

6、根据权利要求 1-5 任一项所述的方法，其特征在于，所述将所述计算
10 资源相关信息的健康度作为预先建立的健康度模型的参数，计算所述待管理的
信息健康分，包括：

采用如下计算公式，计算所述待管理的信息的健康分：

$$E = y / (x + y) * ((\sum(C * n) / \sum n)^4 / 10000000) + x / (x + y) * ((\sum(D * m) / \sum m)^4 / 10000000);$$

其中，E 是待管理的信息的健康分；

15 n 是任务级别的权重，具体可以用如下公式计算得到： $\log(\text{任务运行时间} * \text{plan_cpu_cost} + 1)$ ，plan_cpu_cost 由底层数据库抓取；

m 是节点级别的权重，具体可以用如下公式计算得到：节点的优先级+1，
其中，节点的优先级可以由用户预设；

x 是有节点的任务数，y 是无节点的任务数。

20 7、一种计算资源管理装置，其特征在于，包括：

确定模块，用于确定待管理的信息包括的计算资源相关信息，所述计算资源
相关信息包括：任务，和/或，节点；

获取模块，用于获取所述计算资源相关信息的健康分；

25 计算模块，用于将所述计算资源相关信息的健康度作为预先建立的健康度
模型的参数，计算所述待管理的信息的健康分；

管理模块，用于根据所述待管理的信息的健康分，对所述待管理的信息进行
计算资源管理。

8、根据权利要求 7 所述的装置，其特征在于，所述获取模块具体用于：

30 当所述计算资源相关信息是任务时，根据如下扣分项中的至少一项获取每
种任务的健康分：

倾斜度数值，缺责任人数值，简单加工数值，暴力扫描数值，重复计算数
值，产出为空数值，映射参数设置不合理数值。

9、根据权利要求7所述的装置，其特征在于，所述获取模块具体用于：

当所述计算资源相关信息是节点时，根据如下扣分项中的至少一项获取每种节点的健康分：

节点破基线数值，依赖合理性数值，责任人合理性数值，任务出错天数数值，产出表没有被读数值，同源导入数值，产出为空数值，导入持续一致数值。

10、根据权利要求8所述的装置，其特征在于，如果单个任务与上述各扣分项均相关，所述获取模块具体用于采用如下计算公式计算单个任务的健康分：

$$C = \text{greatest}(0, (100 - \text{least}(a1, 20) - 30 * a2 - 10 * a3 - \text{least}(a4/15, 15) - 10 * a5 - 10 * a6 - a7))^4 / 1000000;$$

其中，greatest为取最大值运算，least是取最小值运算；

C是单个任务的健康分，a1是倾斜度数值，a2是缺责任人数值，a3是简单加工数值，a4是暴力扫描数值，a5是重复计算数值，a6是产出为空数值，a7是映射参数设置不合理数值。

11、根据权利要求9所述的装置，其特征在于，如果单个任务与上述各扣分项均相关，所述获取模块具体用于采用如下计算公式计算单个节点的健康分：

$$D = \text{greatest}(0, (100 - \text{least}(b1 * (p+1) * 10, 100) - 30 * b2 - 10 * b3 - 10 * b4 - 50 * b5 - 50 * b6 - 50 * b7 - 50 * b8))^4 / 1000000;$$

其中，greatest为取最大值运算，least是取最小值运算；

D是单个节点的健康分，b1是节点破基线数值，b2是依赖合理性数值，b3是责任人合理性数值，b4是任务出错天数数值，b5是产出表没有被读数值，b6是同源导入数值，b7是产出为空数值，b8是导入持续一致数值。

12、根据权利要求7-11任一项所述的装置，其特征在于，所述计算模块具体用于：

采用如下计算公式，计算所述待管理的信息的健康分：

$$E = y / (x + y) * ((\sum(C * n) / \sum n)^4 / 1000000) + x / (x + y) * ((\sum(D * m) / \sum m)^4 / 1000000);$$

其中，E是待管理的信息的健康分；

n是任务级别的权重，具体可以用如下公式计算得到： $\log(\text{任务运行时间} * \text{plan_cpu_cost} + 1)$ ，plan_cpu_cost由底层数据库抓取；

m是节点级别的权重，具体可以用如下公式计算得到：节点的优先级+1，其中，节点的优先级可以由用户预设；

x 是有节点的任务数, y 是无节点的任务数。

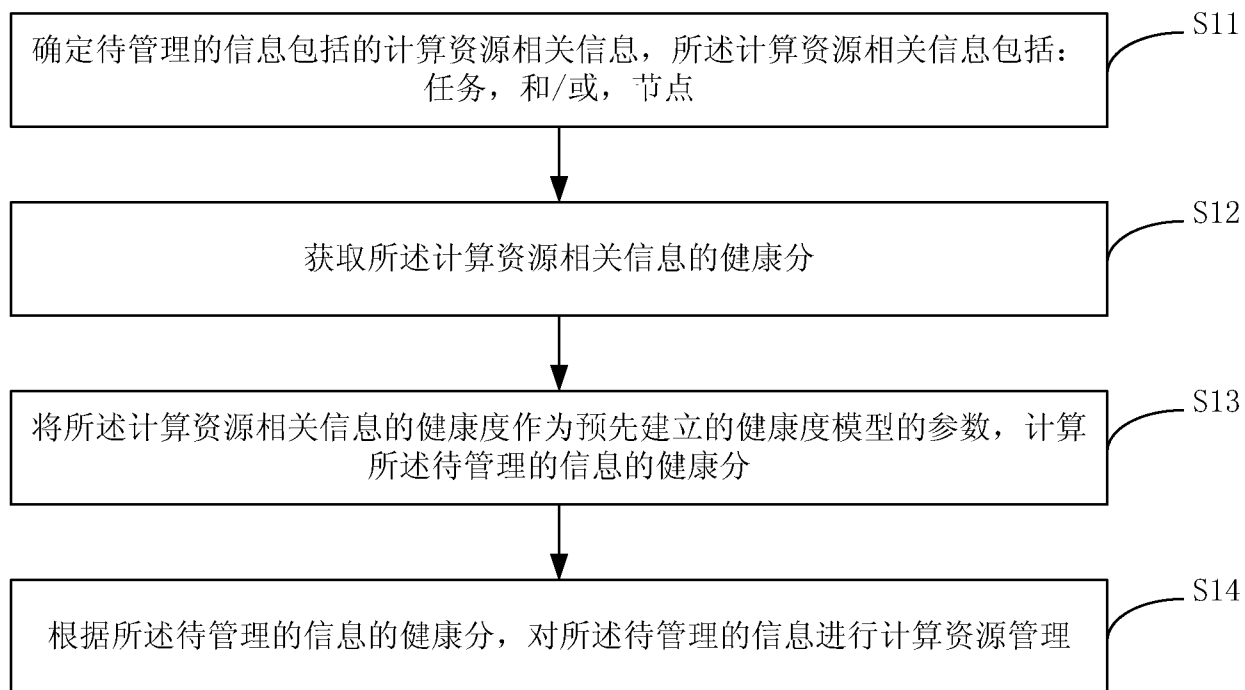


图 1

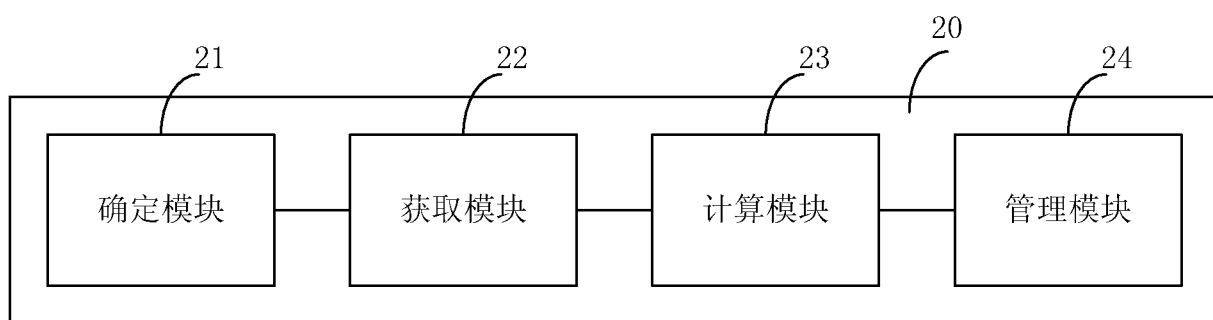


图 2

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2016/079157

A. CLASSIFICATION OF SUBJECT MATTER

G06F 15/16 (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, H04L

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)

CNPAT, CNKI, WPI, EPODOC, IEEE: schedule, node, task, mission, calling, load, balanc+, health, score, degree, resource, weight, parameter

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 2014280899 A1 (BREWSTER, JR.; HERMAN, D. et al.), 18 September 2014 (18.09.2014), description, paragraphs [0010]-[0028], and figures 2-14	1, 7
X	US 2014025414 A1 (WORDEN, B.D. et al.), 23 January 2014 (23.01.2014), description, paragraphs [0026]-[0046]	1, 7
X	CN 104380277 A (INTEL CORPORATION), 25 February 2015 (25.02.2015), description, pages 7-12, and figures 1-6	1, 7
X	US 20140173092 A1 (ULUDERYA, G. et al.), 19 June 2014 (19.06.2014), description, paragraphs [0019]-[0057], and figures 1-8	1, 7
A	CN 101604264 A (SHENZHEN INSTITUTES OF ADVANCED TECHNOLOGY), 16 December 2009 (16.12.2009), the whole document	1-12

☐ Further documents are listed in the continuation of Box C.

☒ See patent family annex.

* Special categories of cited documents:	“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
“A” document defining the general state of the art which is not considered to be of particular relevance	“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
“E” earlier application or patent but published on or after the international filing date	“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
“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)	“&” document member of the same patent family
“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	

Date of the actual completion of the international search 23 June 2016 (23.06.2016)	Date of mailing of the international search report 14 July 2016 (14.07.2016)
Name and mailing address of the ISA/CN: State Intellectual Property Office of the P. R. China No. 6, Xitucheng Road, Jimenqiao Haidian District, Beijing 100088, China Facsimile No.: (86-10) 62019451	Authorized officer LI, Nan Telephone No.: (86-10) 62413690

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/CN2016/079157

Patent Documents referred in the Report	Publication Date	Patent Family	Publication Date
US 2014280899 A1	18 September 2014	None	
US 2014025414 A1	23 January 2014	AU 2013100910 A4	01 August 2013
		BR 102013018540 A2	21 July 2015
		US 2016062356 A1	03 March 2016
CN 104380277 A	25 February 2015	DE 112013003180 T5	07 May 2015
		WO 2014004318 A1	03 January 2014
		US 2014006597 A1	02 January 2014
		US 2016164963 A1	09 June 2016
US 20140173092 A1	19 June 2014	WO 2014099906 A1	26 June 2014
		EP 2932693 A1	21 October 2015
		CN 105122772 A	02 December 2015
CN 101604264 A	16 December 2009	None	

国际检索报告

国际申请号

PCT/CN2016/079157

A. 主题的分类

G06F 15/16(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, H04L

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))

CNPAT, CNKI, WPI, EPDOC, IEEE: 节点, 任务, 调度, 负载, 均衡, 健康, 分, 度, 资源, 权重, 参数, node, task, mission, calling, load, balanc+, health, score, degree, resource, weight, parameter

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	US 2014280899 A1 (BREWSTER, JR., HERMAN DEAN 等) 2014年 9月 18日 (2014 - 09 - 18) 说明书第[0010]-[0028]段, 附图2-14	1, 7
X	US 2014025414 A1 (WORDEN, BRET DWAYNE 等) 2014年 1月 23日 (2014 - 01 - 23) 说明书第[0026]-[0046]段	1, 7
X	CN 104380277 A (英特尔公司) 2015年 2月 25日 (2015 - 02 - 25) 说明书第7-12页, 附图1-6	1, 7
X	US 20140173092 A1 (ULUDERYA, GOKHAN 等) 2014年 6月 19日 (2014 - 06 - 19) 说明书第[0019]-[0057]段, 附图1-8	1, 7
A	CN 101604264 A (深圳先进技术研究院) 2009年 12月 16日 (2009 - 12 - 16) 全文	1-12

☐ 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

2016年 6月 23日

国际检索报告邮寄日期

2016年 7月 14日

ISA/CN的名称和邮寄地址

中华人民共和国国家知识产权局(ISA/CN)
中国北京市海淀区蓟门桥西土城路6号 100088

授权官员

李楠

传真号 (86-10)62019451

电话号码 (86-10)62413690

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2016/079157

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)
US	2014280899	A1	2014年 9月 18日	无			
US	2014025414	A1	2014年 1月 23日	AU	2013100910	A4	2013年 8月 1日
				BR	102013018540	A2	2015年 7月 21日
				US	2016062356	A1	2016年 3月 3日
CN	104380277	A	2015年 2月 25日	DE	112013003180	T5	2015年 5月 7日
				WO	2014004318	A1	2014年 1月 3日
				US	2014006597	A1	2014年 1月 2日
				US	2016164963	A1	2016年 6月 9日
US	20140173092	A1	2014年 6月 19日	WO	2014099906	A1	2014年 6月 26日
				EP	2932693	A1	2015年 10月 21日
				CN	105122772	A	2015年 12月 2日
CN	101604264	A	2009年 12月 16日	无			

表 PCT/ISA/210 (同族专利附件) (2009年7月)