

(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
2014 年 6 月 26 日 (26.06.2014)



(10) 国际公布号
WO 2014/094495 A1

- (51) 国际专利分类号:
H04L 29/08 (2006.01) G06F 9/38 (2006.01)
- (21) 国际申请号: PCT/CN2013/085748
- (22) 国际申请日: 2013 年 10 月 23 日 (23.10.2013)
- (25) 申请语言: 中文
- (26) 公布语言: 中文
- (30) 优先权:
201210563107.8 2012 年 12 月 22 日 (22.12.2012) CN
- (71) 申请人: 深圳先进技术研究院 (SHENZHEN INSTITUTES OF ADVANCED TECHNOLOGY) [CN/CN]; 中国广东省深圳市南山区西丽大学城学苑大道 1068 号, Guangdong 518055 (CN)。
- (72) 发明人: 朱定局 (ZHU, Dingju); 中国广东省深圳市南山区西丽大学城学苑大道 1068 号, Guangdong 518055 (CN)。
- (74) 代理人: 广州华进联合专利商标代理有限公司 (ADVANCE CHINA I.P. LAW OFFICE); 中国广东省广州市先烈中路 69 号东山广场 918-920 室, Guangdong 510095 (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, LT, 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, 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: CLOUD SCHEDULING METHOD AND SYSTEM THEREOF

(54) 发明名称: 云调度方法及其系统

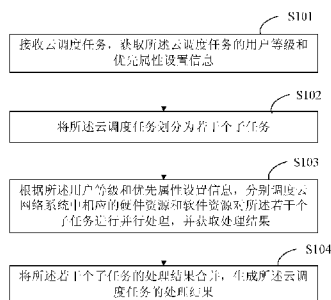


图 1 / FIG.1

S101 Receiving a cloud scheduling task, and acquiring the user grade and priority attribute setting information of the cloud scheduling task
S102 Dividing the cloud scheduling task into a plurality of subtasks
S103 Respectively scheduling, according to the user grade and the priority attribute setting information, the corresponding hardware resources and software resources in a cloud network system to simultaneously process the plurality of subtasks, and acquiring the processing result
S104 Merging the processing results of the plurality of subtasks into the processing result of the cloud scheduling task

(57) Abstract: Provided are a cloud scheduling method and system thereof, the method comprising: receiving a cloud scheduling task, and acquiring the user grade and priority attribute setting information of the cloud scheduling task; dividing the cloud scheduling task into a plurality of subtasks; respectively scheduling, according to the user grade and the priority attribute setting information, the corresponding hardware resources and software resources in a cloud network system to simultaneously process the plurality of subtasks, and acquiring the processing result; merging the processing results of the plurality of subtasks into the processing result of the cloud scheduling task. The present invention divides a cloud scheduling task into a plurality of subtasks for simultaneous processing, and schedules different hardware resources and software resources for different cloud scheduling tasks as required, thus improving the processing efficiency, and satisfying the customized service requirement of different cloud scheduling tasks.

(57) 摘要: 本发明提供一种云调度方法及其系统, 所述方法包括: 接收云调度任务, 获取所述云调度任务的用户等级和优先属性设置信息; 将所述云调度任务划分为若干个子任务; 根据所述用户等级和优先属性设置信息, 分别调度云网络系统中相应的硬件资源和软件资源对所述若干个子任务进行并行处理, 并获取处理结果; 将所述若干个子任务的处理结果合并, 生成所述云调度任务的处理结果。本发明能够将云调度任务划分成为多个子任务并行处理, 对不同的云调度任务根据其需求调度不同的硬件和软件资源, 提高处理效率, 满足不同云调度任务的个性化服务需求。

WO 2014/094495 A1

发明名称：云调度方法及其系统

[1] 【技术领域】

[2] 本发明涉及通过网络的云计算调度的技术领域，特别是涉及一种云调度方法，以及一种云调度系统。

[3] 【背景技术】

[4] 随着计算机网络技术的发展，通过网络的云计算调度技术的应用越来越广泛，目前云计算调度都是仅根据云调度任务的任务量、资源量的信息进行网络资源的云调度。

[5] 但是，仅根据任务量、资源量的信息进行云调度，虽然能够达到充分利用云计算资源、尽快处理云调度任务的目的，但却不能满足不同云调度任务的不同个性化需求，例如有的云处理任务要求高的服务质量，而对成本因素并不要求，而有的云处理任务要求采取低成本的处理方式，但对处理速度等服务质量要求不高，而在现有云调度模式下，对所有的云调度任务都是同等对待的，各个云调度任务获得的云服务调度是相同的，而无法获得与自身要求符合的云服务效果。

[6] 【发明内容】

[7] 针对上述背景技术中存在的问题，本发明的目的在于提供一种云调度方法及其系统，能够对不同的云调度任务根据其需求调度不同的硬件和软件资源进行处理，满足不同云调度任务的服务需求。

[8] 一种云调度方法，包括以下步骤：

[9] 接收云调度任务，获取所述云调度任务的用户等级和优先属性设置信息；

[10] 将所述云调度任务划分为若干个子任务；

[11] 根据所述用户等级和优先属性设置信息，分别调度云网络系统中相应的硬件资源和软件资源对所述若干个子任务进行并行处理，并获取处理结果；

[12] 将所述若干个子任务的处理结果合并，生成所述云调度任务的处理结果。

[13] 一种云调度系统，包括：

- [14] 信息获取模块，用于接收云调度任务，获取所述云调度任务的用户等级和优先属性设置信息；
- [15] 任务划分模块，用于将所述云调度任务划分为若干个子任务；
- [16] 资源调度模块，用于根据所述用户等级和优先属性设置信息，分别调度云网络系统中相应的硬件资源和软件资源对所述若干个子任务进行并行处理，并获取处理结果；
- [17] 结果合并模块，用于将所述若干个子任务的处理结果合并，生成所述云调度任务的处理结果。
- [18] 本发明的云调度方法及其系统通过获取云调度任务的用户等级和优先属性设置信息，将所述云调度任务划分为若干个子任务，根据所述用户等级和优先属性设置信息调度对应的软、硬件资源对各个子任务进行处理，满足不同的云调度任务的处理要求，最后合并子任务的处理结果，得到所述云调度任务的处理结果，提高处理效率，满足不同云调度任务的个性化服务需求。

[19] **【附图说明】**

- [20] 图1是本发明云调度方法的流程示意图；
- [21] 图2是本发明云调度系统的结构示意图。

[22] **【具体实施方式】**

- [23] 请参阅图1，图1是本发明云调度方法的流程示意图。
- [24] 所述云调度方法包括以下步骤：
- [25] S101，接收云调度任务，获取所述云调度任务的用户等级和优先属性设置信息；
- [26] S102，将所述云调度任务划分为若干个子任务；
- [27] S103，根据所述用户等级和优先属性设置信息，分别调度云网络系统中相应的硬件资源和软件资源对所述若干个子任务进行并行处理，并获取处理结果；
- [28] S104，将所述若干个子任务的处理结果合并，生成所述云调度任务的处理结果。
- [29] 本发明能够将云调度任务划分成为多个子任务并行处理，对不同的云调度任务根据其需求调度不同的硬件和软件资源，提高处理效率，满足不同云调度任务

的个性化服务需求。

- [30] 其中，对于步骤S101，所述云调度任务包括各种用户发出的云计算调度任务，包括免费用户、付费用户以及VIP用户发出的调度云资源进行计算处理的视频云服务的调度、机器人云服务的调度、数据审查云服务的调度任务。
- [31] 在接收到所述云调度任务时，获取所述云调度任务的用户等级和优先属性设置信息。所述云调度任务的用户等级包括免费用户、付费用户以及VIP用户。可以通过获取用户的登陆信息、注册信息或者付费信息等用户账户信息来判断发出所述云调度任务的用户的用户等级；
- [32] 所述云调度任务的优先属性设置信息是指所述云调度任务要求的各种云调度服务指标的优先级，包括速度优先、成本优先、质量优先及其组合。所述优先属性设置信息可以在发出所述云调度任务时提示用户输入，即通过用户界面中提供的选项，提供给用户进行选择，例如通过下拉框模式，在所述下拉框中有多个选项，不同的选项中包含有不同的优先属性设置信息，如速度优先、成本优先、质量优先及其组合。所述云调度任务的优先属性设置信息也可以通过对相应的用户的各种行为信息和输入信息的搜集、筛选提取得出。
- [33] 对于上述步骤S102，对所述云调度任务的划分可以根据所述云调度任务的任务量（运算量）来确定，优选地，本发明提供一种所述云调度任务的划分方式如下：
- [34] 首先计算所述云调度任务的默认可用资源量，根据以下公式： $Y=M \times X/N$ ；其中， Y 为所述云调度任务的默认可用资源量， M 为资源总量， N 为任务总量， X 为所述云调度任务的任务量；
- [35] 然后，根据所述默认可用资源量计算所述云调度任务的默认并发度，按照以下公式计算： $p=Y/R$ ，其中， p 为默认并发度， R 为每一个并发进程占用的资源量；
- [36] 最后根据所述默认并发度，设置所述云调度任务划分的子任务个数。
- [37] 在本实施方式中，对所述云调度任务的划分首先考虑其可用资源量，即在所述云网络系统中可以占用的资源份额，如果所述可用资源量则可以划分成更多的子任务，使每一个子任务所占用的处理时间缩短，尽量提高云调度的速度。
- [38] 在获取所述默认并发度之后，可以将所述云调度任务划分的子任务个数统一设

置成等于所述默认并发度。

- [39] 本发明中，优选根据用户个性化需求设置所述云调度任务划分的子任务个数，则考虑到所述用户等级包括免费用户、付费用户以及VIP用户；所述优先属性设置信息包括：速度优先、成本优先以及质量优先；在根据所述默认并发度设置所述云调度任务划分的子任务个数时，首先判断所述云调度任务的用户等级以及优先属性设置信息：
- [40] 当用户等级为免费用户，或者用户等级为付费用户或VIP用户、且优先属性设置信息为成本优先时，设置所述云调度任务划分的子任务个数小于所述默认并发度；
- [41] 当用户等级为付费用户或VIP用户，并且优先属性设置信息为速度优先，或质量优先时，设置所述云调度任务划分的子任务个数大于所述默认并发度。
- [42] 通过上述方式，可以根据用户或者不同云调度任务的个性化需求，具体设置每个云调度任务划分的子任务个数，在所述云调度任务要求的处理速度和处理质量更高时，可以将增加其划分的子任务数，调度更多的资源去处理各个子任务，提高整个云调度任务的处理速度和处理质量。
- [43] 对于上述步骤S103，根据所述用户等级和优先属性设置信息，在云网络系统中调度相应的硬件资源和软件资源对各个所述若干个子任务进行处理。为实现相应的资源的调度，可以在调度端预先保存一个资源调度控制表，所述资源调度控制表中预先保存有各个所述硬件资源和软件资源的资源属性，以及各种所述用户等级以及所述优先属性设置信息所对应的资源属性。
- [44] 在各个硬件资源和软件资源投入云网络系统中使用时，首先将所述硬件资源和软件资源的资源属性录入并保存在所述资源调度控制表中，并通过对所述资源调度控制表的设置录入各种所述用户等级以及所述优先属性设置信息所对应的资源属性；
- [45] 则在实际处理云调度任务时，根据所述云调度任务的用户等级和优先属性设置信息，查询预先保存的所述资源调度控制表，获取对应的资源属性，再根据所述资源属性查询对应的硬件资源和软件资源，根据查询结果调度云网络系统中对应的硬件资源和软件资源对所述若干个子任务进行并行处理。

- [46] 在调度各个所述硬件资源和软件资源时，可以通过调用所述云计算系统提供的调度接口衔接来实现云调度，现有的云计算系统（例如hadoop，一种分布式系统基础架构）都是有调度模块API接口的，通过调度模块API接口就可以调度现有云计算系统中的资源。
- [47] 在所述资源调度控制表中，所述硬件资源和软件资源的资源属性包括：成本、速度、稳定性。
- [48] 考虑到所述用户等级包括：免费用户、付费用户以及VIP用户；所述优先属性设置信息包括：速度优先，成本优先以及质量优先；
- [49] 在查询所述资源调度控制表，根据查询结果调度对应的硬件资源和软件资源时，优选采用与方式进行相应资源的调度：
- [50] 当所述用户等级为免费用户，或者所述用户等级为付费用户或者VIP用户、并且所述优先属性设置信息为成本优先时，调用资源属性中的成本低于默认值的硬件资源和软件资源对所述若干个子任务进行并行处理。亦即，如果是免费用户或选取成本低优先的收费用户和VIP用户，则采用成本低优先的云调度方式，即尽量调用比默认调度方式耗资源更少的、成本更低廉的硬件和软件模块，此时云服务质量先对较低和云服务速度相对较慢；
- [51] 当所述用户等级为付费用户或者VIP用户，并且所述优先属性设置信息为速度优先时，调用资源属性中的速度高于默认值的硬件资源和软件资源对所述若干个子任务进行并行处理。即，如果是收费用户或VIP用户选取速度优先方式，则尽量调用比默认调度方式更高性能的硬件模块、功能更简要的软件模块；
- [52] 当所述用户等级为付费用户或者VIP用户，并且所述优先属性设置信息为质量优先时，调用性资源属性中的稳定性高于默认值的硬件资源和软件资源对所述若干个子任务进行并行处理。即，如果是收费用户或VIP用户选取质量优先方式，则尽量调用比默认调度方式更稳定的硬件模块、功能更完备的软件服务模块；
- [53] 当所述用户等级为付费用户或者VIP用户，并且所述优先属性设置信息为质量优先以及速度优先时，调用资源属性中的速度高于默认值、并且稳定性高于默认值的硬件资源和软件资源对所述若干个子任务进行并行处理。即，如果是收

费用户或VIP用户同时选取速度优先方式和质量优先，则尽量调用比默认调度方式更稳定、更高性能的硬件模块、功能更完备的软件服务模块。

[54] 通常，各种资源属性所对应的默认值可以根据实际需要设定，也以设定为不区分云调度任务的个性化需求，对所有的云调度任务同等对待时调度的资源属性平均值。

[55] 通过上述方式，根据不同的用户等级以及优先属性设置信息来调度不同的硬件资源和软件资源对各个子任务进行处理，满足不同云调度任务的个性化需求。

[56] 根据查询结果调度所述云网络系统中对应的硬件资源和软件资源对所述若干个子任务进行并行处理的方法为：

[57] 根据所述若干个子任务的集合，生成与所述子任务一一对应的若干个并行云处理进程；将所述若干个并行云处理进程分别指派到查询获取的若干个相应的硬件资源上运行，运行的进程内容为查询获取的若干个软件资源。

[58] 由于是通过并行云处理进程对多个子任务进行并行处理，因此大大缩短了对所述云调度任务的处理时间，提高了处理效率。

[59] 在所述资源调度控制表中，所述硬件资源和软件资源的资源属性还可以设置成其他的参数，例如，对于硬件资源，可设置其资源属性为CPU性能范围，memory性能范围和disk性能范围。

[60] 如果在步骤S102中将所述云调度任务 v 划分为 p 个子任务 $\{v_1, v_2, \dots, v_p\}$ ，则在本步骤中，根据所述资源调度控制表的查询结果，从云资源池中选取CPU性能范围满足 (c_1, c_2) ，memory性能范围满足 (m_1, m_2) ，disk性能范围满足 (d_1, d_2) 的 p 个硬件资源 $\{h_1, h_2, \dots, h_p\}$ ，以及对应的 p 个软件资源 $\{s_1, s_2, \dots, s_p\}$ 。

[61] 根据 p 个并发子任务集合，克隆出 p 个与所述子任务集合对应的并行云处理进程；根据 p 个并行云处理进程、 p 个计算机硬件资源、 p 个软件资源，将 p 个并行云处理进程指派到 p 个相应的计算机硬件资源上运行，运行的进程内容为 p 个软件资源。

[62] 对于上述步骤S104，对 p 个并行云处理进程对所述若干个子任务的处理结果集合进行合并，生成所述云调度任务的处理结果，并返回给用户，完成云调度任务的处理过程。

- [63] 本发明的云调度方法将云调度任务划分为若干个子任务，对各个所述子任务进行并行处理后将处理结果合并成所述云调度任务的处理结果，因此，能够大大缩短处理时间。通过云调度任务的用户等级和优先属性设置信息调用不同的软、硬件资源对各个子任务进行处理，能够满足不同的云调度任务的处理要求，使云调度更灵活，提高资源利用率。
- [64] 请参阅图2，图2是本发明云调度系统的结构示意图。
- [65] 所述云调度系统包括：
- [66] 信息获取模块11，用于接收云调度任务，获取所述云调度任务的用户等级和优先属性设置信息；
- [67] 任务划分模块12，用于将所述云调度任务划分为若干个子任务；
- [68] 资源调度模块13，用于根据所述用户等级和优先属性设置信息，分别调度云网络系统中相应的硬件资源和软件资源对所述若干个子任务进行并行处理，并获取处理结果；
- [69] 结果合并模块14，用于将所述若干个子任务的处理结果合并，生成所述云调度任务的处理结果。
- [70] 其中，所述云调度任务包括各种用户发出的云计算调度任务，包括免费用户、付费用户以及VIP用户发出的调度云资源进行计算处理的视频云服务的调度、机器人云服务的调度、数据审查云服务的调度任务。
- [71] 在接收到所述云调度任务时，所述信息获取模块11获取所述云调度任务的用户等级和优先属性设置信息。所述云调度任务的用户等级包括免费用户、付费用户以及VIP用户。所述信息获取模块11可以通过获取用户的登陆信息、注册信息或者付费信息等用户账户信息来判断发出所述云调度任务的用户的用户等级；
- [72] 所述云调度任务的优先属性设置信息是指所述云调度任务要求的各种云调度服务指标的优先级，包括速度优先、成本优先、质量优先及其组合。所述优先属性设置信息可以在发出所述云调度任务时提示用户输入，即通过用户界面中提供的选项，提供给用户进行选择，例如通过下拉框模式，在所述下拉框中有多个选项，不同的选项中包含有不同的优先属性设置信息，如速度优先、成本优先、质量优先及其组合。所述云调度任务的优先属性设置信息也可以通过对相

应的用户的各种行为信息和输入信息的搜集、筛选提取得出。

[73] 所述任务划分模块12对所述云调度任务的划分可以根据所述云调度任务的任务量（运算量）来确定，优选地，所述任务划分模块包括以下子模块：

[74] 可用资源计算模块，用于根据以下公式计算所述云调度任务的默认可用资源量： $Y=M*X/N$ ，其中， Y 为所述云调度任务的默认可用资源量， M 为资源总量， N 为任务总量， X 为所述云调度任务的任务量；

[75] 并发度计算模块，用于根据以下公式计算所述云调度任务的默认并发度： $p=Y/R$ ，其中， p 为默认并发度， R 为每一个并发进程占用的资源量；

[76] 划分模块，用于根据所述默认并发度，设置所述云调度任务划分的子任务个数。

[77] 在本实施方式中，所述任务划分模块12对所述云调度任务的划分首先考虑其可用资源量，即在所述云网络系统中可以占用的资源份额，如果所述可用资源量则可以划分成更多的子任务，使每一个子任务所占用的处理时间缩短，尽量提高云调度的速度。

[78] 在获取所述默认并发度之后，可以将所述云调度任务划分的子任务个数统一设置成等于所述默认并发度。

[79] 本发明中，优选根据用户个性化需求设置所述云调度任务划分的子任务个数，考虑到所述用户等级包括免费用户、付费用户以及VIP用户；所述优先属性设置信息包括：速度优先、成本优先以及质量优先。则所述任务划分模块12在根据所述默认并发度设置所述云调度任务划分的子任务个数时，首先判断所述云调度任务的用户等级以及优先属性设置信息：

[80] 当用户等级为免费用户，或者用户等级为付费用户或VIP用户、且优先属性设置信息为成本优先时，设置所述云调度任务划分的子任务个数小于所述默认并发度；

[81] 当用户等级为付费用户或VIP用户，并且优先属性设置信息为速度优先，或质量优先时，设置所述云调度任务划分的子任务个数大于所述默认并发度。

[82] 通过上述方式，可以根据用户或者不同云调度任务的个性化需求，具体设置每个云调度任务划分的子任务个数，在所述云调度任务要求的处理速度和处理质

量更高时，可以将增加其划分的子任务数，调度更多的资源去处理各个子任务，提高整个云调度任务的处理速度和处理质量。

[83] 所述资源调度模块13用于根据所述用户等级和优先属性设置信息，在云网络系统中调度相应的硬件资源和软件资源对各个所述若干个子任务进行处理。为实现相应的资源的调度，可以在调度端预先保存一个资源调度控制表，所述资源调度控制表中预先保存有各个所述硬件资源和软件资源的资源属性，以及各种所述用户等级以及所述优先属性设置信息所对应的资源属性。

[84] 在各个硬件资源和软件资源投入云网络系统中使用时，首先将所述硬件资源和软件资源的资源属性录入并保存在所述资源调度控制表中，并通过对所述资源调度控制表的设置录入各种所述用户等级以及所述优先属性设置信息所对应的资源属性；

[85] 则在实际处理云调度任务时，根据所述云调度任务的用户等级和优先属性设置信息，查询预先保存的所述资源调度控制表，获取对应的资源属性，再根据所述资源属性查询对应的硬件资源和软件资源，根据查询结果调度云网络系统中对应的硬件资源和软件资源对所述若干个子任务进行并行处理。

[86] 所述资源调度模块13在调度各个所述硬件资源和软件资源时，可以通过调用所述云计算系统提供的调度接口衔接来实现云调度，现有的云计算系统（例如hadoop）都是有调度模块API接口的，通过调度模块API接口就可以调度现有云计算系统中的资源。

[87] 在所述资源调度控制表中，所述硬件资源和软件资源的资源属性包括：成本、速度、稳定性。

[88] 考虑到所述用户等级包括：免费用户、付费用户以及VIP用户；所述优先属性设置信息包括：速度优先，成本优先以及质量优先；

[89] 则，所述资源调度模块13优选包括以下子模块：

[90] 第一调度模块，用于在所述用户等级为免费用户，或者所述用户等级为付费用户或者VIP用户、并且所述优先属性设置信息为成本优先时，调用资源属性中的成本低于默认值的硬件资源和软件资源对所述若干个子任务进行并行处理。亦即，如果是免费用户或选取成本低优先的收费用户和VIP用户，则采用成本低优

先的云调度方式，即尽量调用比默认调度方式耗资源更少的、成本更低廉的硬件和软件模块，此时云服务质量先对较低和云服务速度相对较慢；

[91] 第二调度模块，用于在所述用户等级为付费用户或者VIP用户，并且所述优先属性设置信息为速度优先时，调用资源属性中的速度高于默认值的硬件资源和软件资源对所述若干个子任务进行并行处理。即，如果是收费用户或VIP用户选取速度优先方式，则尽量调用比默认调度方式更高性能的硬件模块、功能更简

要的软件模块；

[92] 第三调度模块，用于在所述用户等级为付费用户或者VIP用户，并且所述优先属性设置信息为质量优先时，调用性资源属性中的稳定性高于默认值的硬件资源和软件资源对所述若干个子任务进行并行处理。即，如果是收费用户或VIP用户选取质量优先方式，则尽量调用比默认调度方式更稳定的硬件模块、功能更完备的软件服务模块；

[93] 第四调度模块，用于在所述用户等级为付费用户或者VIP用户，并且所述优先属性设置信息为质量优先以及速度优先时，调用资源属性中的速度高于默认值、并且稳定性高于默认值的硬件资源和软件资源对所述若干个子任务进行并行处理。即，如果是收费用户或VIP用户同时选取速度优先方式和质量优先，则尽量调用比默认调度方式更稳定、更高性能的硬件模块、功能更完备的软件服务模块。

[94] 通常，各种资源属性所对应的默认值可以根据实际需要设定，也以设定为不区分云调度任务的个性化需求，对所有的云调度任务同等对待时调度的资源属性平均值。

[95] 通过上述方式，根据不同的用户等级以及优先属性设置信息来调度不同的硬件资源和软件资源对各个子任务进行处理，满足不同云调度任务的个性化需求。

[96] 根据查询结果调度所述云网络系统中对应的硬件资源和软件资源对所述若干个子任务进行并行处理的方法为：

[97] 根据所述若干个子任务的集合，生成与所述子任务一一对应的若干个并行云处理进程；将所述若干个并行云处理进程分别指派到查询获取的若干个相应的硬件资源上运行，运行的进程内容为查询获取的若干个软件资源。

- [98] 由于是通过并行云处理进程对多个子任务进行并行处理，因此大大缩短了对所述云调度任务的处理时间，提高了处理效率。
- [99] 在所述资源调度控制表中，所述硬件资源和软件资源的资源属性还可以设置成其他的参数，例如，对于硬件资源，可设置其资源属性为CPU性能范围，memory性能范围和disk性能范围等属性。
- [100] 所述结果合并模块14对p个并行云处理进程对所述若干个子任务的处理结果集合进行合并，生成所述云调度任务的处理结果，并返回给用户。
- [101] 本发明的云调度方法及其系统可以用于所有各种类型应用的云调度，包括但不限于：视频云服务的调度、机器人云服务的调度、数据审查云服务的调度。
- [102] 本领域普通技术人员可以理解实现上述实施方式中的全部或部分流程以及对应的系统，是可以通过计算机程序来指令相关的硬件来完成，所述的程序可存储于一计算机可读取存储介质中，该程序在执行时，可包括如上述各实施方式的流程。其中，所述的存储介质可为磁碟、光盘、只读存储记忆体（Read-Only Memory, ROM）或随机存储记忆体（Random Access Memory, RAM）等。
- [103] 以上所述实施例仅表达了本发明的几种实施方式，其描述较为具体和详细，但并不能因此而理解为对本发明专利范围的限制。应当指出的是，对于本领域的普通技术人员来说，在不脱离本发明构思的前提下，还可以做出若干变形和改进，这些都属于本发明的保护范围。因此，本发明的保护范围应以所附权利要求为准。

权利要求书

[权利要求 1]

一种云调度方法，其特征在于，包括以下步骤：

接收云调度任务，获取所述云调度任务的用户等级和优先属性设置信息；

将所述云调度任务划分为若干个子任务；

根据所述用户等级和优先属性设置信息，分别调度云网络系统中相应的硬件资源和软件资源对所述若干个子任务进行并行处理，并获取处理结果；

将所述若干个子任务的处理结果合并，生成所述云调度任务的处理结果。

[权利要求 2]

如权利要求1所述的云调度方法，其特征在于，将所述云调度任务划分为若干个子任务的步骤包括：

根据以下公式计算所述云调度任务的默认可用资源量： $Y=M*X/N$ ，其中， Y 为所述云调度任务的默认可用资源量， M 为资源总量， N 为任务总量， X 为所述云调度任务的任务量；

根据以下公式计算所述云调度任务的默认并发度： $p=Y/R$ ，其中， p 为默认并发度， R 为每一个并发进程占用的资源量；

根据所述默认并发度，设置所述云调度任务划分的子任务个数。

[权利要求 3]

如权利要求2所述的云调度方法，其特征在于，所述用户等级包括免费用户、付费用户以及VIP用户；所述优先属性设置信息包括：速度优先、成本优先以及质量优先；

根据所述默认并发度，设置所述云调度任务划分的子任务个数的步骤包括：

当用户等级为免费用户，或者用户等级为付费用户或VIP用户、且优先属性设置信息为成本优先时，设置所述云调度任务划分的子任务个数小于所述默认并发度；

当用户等级为付费用户或VIP用户，并且优先属性设置信息为速度优先，或质量优先时，设置所述云调度任务划分的子任务个数大

于所述默认并发度。

[权利要求 4]

如权利要求1所述的云调度方法，其特征在于，根据所述用户等级和优先属性设置信息，分别调度云网络系统中相应的硬件资源和软件资源对所述若干个子任务进行并行处理的步骤包括：

根据所述用户等级和优先属性设置信息，查询预先保存的资源调度控制表，根据查询结果调度所述云网络系统中对应的硬件资源和软件资源对所述若干个子任务进行并行处理；其中，所述资源调度控制表中预先保存有各个所述硬件资源和软件资源的资源属性，以及各种所述用户等级以及所述优先属性设置信息所对应的资源属性。

[权利要求 5]

如权利要求4所述的云调度方法，其特征在于，根据查询结果调度所述云网络系统中对应的硬件资源和软件资源对所述若干个子任务进行并行处理的步骤包括：

根据所述若干个子任务的集合，生成与所述子任务一一对应的若干个并行云处理进程；

将所述若干个并行云处理进程分别指派到查询获取的若干个相应的硬件资源上运行，运行的进程内容为查询获取的若干个软件资源。

[权利要求 6]

如权利要求4所述的云调度方法，其特征在于，所述用户等级包括：免费用户、付费用户以及VIP用户；所述优先属性设置信息包括：速度优先，成本优先以及质量优先；所述资源属性包括：成本、速度、稳定性；

查询预先保存的资源调度控制表，根据查询结果调度所述云网络系统中对应的硬件资源和软件资源对所述若干个子任务进行并行处理的步骤包括：

当所述用户等级为免费用户，或者所述用户等级为付费用户或者VIP用户、并且所述优先属性设置信息为成本优先时，调用资源属性中的成本低于默认值的硬件资源和软件资源对所述若干个子任

务进行并行处理；

当所述用户等级为付费用户或者VIP用户，并且所述优先属性设置信息为速度优先时，调用资源属性中的速度高于默认值的硬件资源和软件资源对所述若干个子任务进行并行处理；

当所述用户等级为付费用户或者VIP用户，并且所述优先属性设置信息为质量优先时，调用性资源属性中的稳定性高于默认值的硬件资源和软件资源对所述若干个子任务进行并行处理；

当所述用户等级为付费用户或者VIP用户，并且所述优先属性设置信息为质量优先以及速度优先时，调用资源属性中的速度高于默认值、并且稳定性高于默认值的硬件资源和软件资源对所述若干个子任务进行并行处理。

[权利要求 7]

一种云调度系统，其特征在于，包括：

信息获取模块，用于接收云调度任务，获取所述云调度任务的用户等级和优先属性设置信息；

任务划分模块，用于将所述云调度任务划分为若干个子任务；

资源调度模块，用于根据所述用户等级和优先属性设置信息，分别调度云网络系统中相应的硬件资源和软件资源对所述若干个子任务进行并行处理，并获取处理结果；

结果合并模块，用于将所述若干个子任务的处理结果合并，生成所述云调度任务的处理结果。

[权利要求 8]

如权利要求7所述的云调度系统，其特征在于，所述任务划分模块包括：

可用资源计算模块，用于根据以下公式计算所述云调度任务的默认可用资源量： $Y=M*X/N$ ，其中，Y为所述云调度任务的默认可用资源量，M为资源总量，N为任务总量，X为所述云调度任务的任务量；

并发度计算模块，用于根据以下公式计算所述云调度任务的默认并发度： $p=Y/R$ ，其中，p为默认并发度，R为每一个并发进程占

用的资源量；

划分模块，用于根据所述默认并发度，设置所述云调度任务划分的子任务个数；在用户等级为免费用户，或者用户等级为付费用户或VIP用户、且优先属性设置信息为成本优先时，设置所述云调度任务划分的子任务个数小于所述默认并发度；在用户等级为付费用户或VIP用户，并且优先属性设置信息为速度优先，或质量优先时，设置所述云调度任务划分的子任务个数大于所述默认并发度。

[权利要求 9]

如权利要求7所述的云调度系统，其特征在于，所述资源调度模块根据所述用户等级和优先属性设置信息，查询预先保存的资源调度控制表，根据查询结果调度云网络系统中对应的硬件资源和软件资源对所述若干个子任务进行并行处理；其中，所述资源调度控制表中预先保存有各个所述硬件资源和软件资源的资源属性，以及各种所述用户等级以及所述优先属性设置信息所对应的资源属性。

[权利要求 10]

如权利要求9所述的云调度系统，其特征在于，所述用户等级包括：免费用户、付费用户以及VIP用户；所述优先属性设置信息包括：速度优先，成本优先以及质量优先；所述资源属性包括：成本、速度、稳定性；

所述资源调度模块包括：

第一调度模块，用于在所述用户等级为免费用户，或者所述用户等级为付费用户或者VIP用户、并且所述优先属性设置信息为成本优先时，调用资源属性中的成本低于默认值的硬件资源和软件资源对所述若干个子任务进行并行处理；

第二调度模块，用于在所述用户等级为付费用户或者VIP用户，并且所述优先属性设置信息为速度优先时，调用资源属性中的速度高于默认值的硬件资源和软件资源对所述若干个子任务进行并行处理；

第三调度模块，用于在所述用户等级为付费用户或者VIP用户，并且所述优先属性设置信息为质量优先时，调用性资源属性中的稳定性高于默认值的硬件资源和软件资源对所述若干个子任务进行并行处理；

第四调度模块，用于在所述用户等级为付费用户或者VIP用户，并且所述优先属性设置信息为质量优先以及速度优先时，调用资源属性中的速度高于默认值、并且稳定性高于默认值的硬件资源和软件资源对所述若干个子任务进行并行处理。

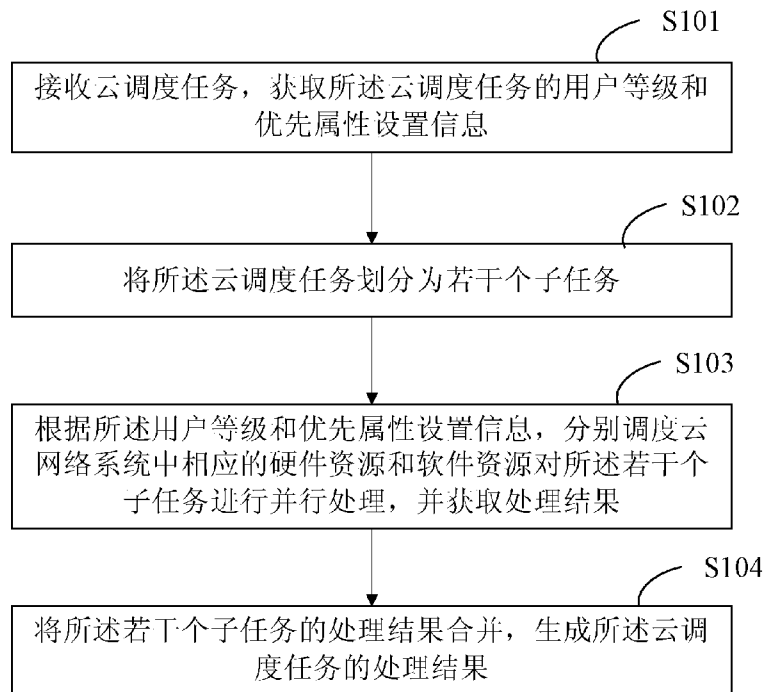


图 1

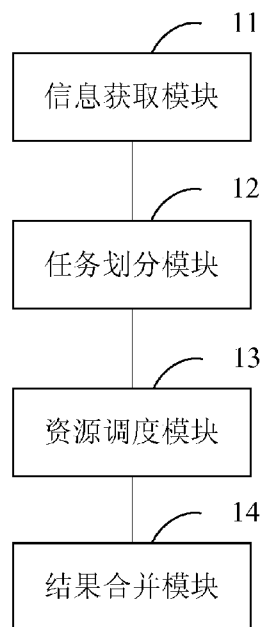


图 2

INTERNATIONAL SEARCH REPORT

International application No.
PCT/CN2013/085748

A. CLASSIFICATION OF SUBJECT MATTER

See the extra sheet

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: H04L; 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)

WPI, EPODOC, CNPAT, CNKI: cloud computing, scheduling, rank, priority, sub-task, parallel, unite

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	CN 102402423 A (BAIDU ON-LINE NETWORK TECHNOLOGY (BEIJING) CO., LTD.) 04 April 2012 (04.04.2012) description, paragraphs [0010]-[0117] and figures 1-7	1, 4-7, 9, 10
A		2, 3, 8
X	CN 102222174 A (BGI SHENZHEN CO., LTD.) 19 October 2011 (19.10.2011) description, paragraphs [0003]-[0067] and figures 1-9	1, 4-7, 9, 10
A		2, 3, 8
A	US 2011/0154350 A1 (INTERNATIONAL BUSINESS MACHINES CORPORATION) 23 June 2011 (23.06.2011) the whole document	1-10
PX	CN 103067468 A (SHENZHEN INSTITUTE OF ADVANCED TECHNOLOGY) 24 April 2013 (24.04.2013) the whole document	1-10

☐ 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 16 December 2013 (16.12.2013)	Date of mailing of the international search report 23 January 2014 (23.01.2014)
Name and mailing address of the ISA 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 WANG, Guogang Telephone No. (86-10) 62413284

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/CN2013/085748

Patent Documents referred in the Report	Publication Date	Patent Family	Publication Date
CN 102402423 A	04.04.2012	None	
CN 102222174 A	19.10.2011	WO 2012113290 A1	30.08.2012
		HK 1162712 A0	31.08.2012
US 2011/0154350 A1	23.06.2011	US 2012192197 A1	26.07.2012
CN 103067468 A	24.04.2013	None	

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2013/085748

CLASSIFICATION OF SUBJECT MATTER:

H04L 29/08 (2006.01) i

G06F 9/38 (2006.01) i

国际检索报告

国际申请号
PCT/CN2013/085748

A. 主题的分类

参见附加页

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: H04L, 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))

WPI, EPODOC, CNPAT, CNKI: 云计算, 调度, 等级, 优先级, 子任务, 并行, 合并, cloud computing, scheduling, rank, priority, sub-task, parallel, unite

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 102402423 A (百度在线网络技术(北京)有限公司) 04.4 月 2012 (04.04.2012) 说明书第[0010]-[0117]段, 附图 1-7	1, 4-7, 9-10
A		2-3, 8
X	CN 102222174 A (深圳华大基因科技有限公司) 19.10 月 2011 (19.10.2011) 说明书第[0003]-[0067]段, 附图 1-9	1, 4-7, 9-10
A		2-3, 8
A	US 2011/0154350 A1 (INTERNATIONAL BUSINESS MACHINES CORPORATION) 23.6 月 2011 (23.06.2011) 全文	1-10
PX	CN 103067468 A (深圳先进技术研究院) 24.4 月 2013 (24.04.2013) 全文	1-10

☐ 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
16.12 月 2013 (16.12.2013)

国际检索报告邮寄日期
23.1 月 2014 (23.01.2014)

ISA/CN 的名称和邮寄地址:
中华人民共和国国家知识产权局
中国北京市海淀区蓟门桥西土城路 6 号 100088
传真号: (86-10)62019451

受权官员

王国纲
电话号码: (86-10) **62413284**

国际检索报告
关于同族专利的信息

国际申请号
PCT/CN2013/085748

检索报告中引用的 专利文件	公布日期	同族专利	公布日期
CN 102402423 A	04.04.2012	无	
CN 102222174 A	19.10.2011	WO 2012113290 A1	30.08.2012
		HK 1162712 A0	31.08.2012
US 2011/0154350 A1	23.06.2011	US 2012192197 A1	26.07.2012
CN 103067468 A	24.04.2013	无	

主题的分类:

H04L 29/08 (2006.01) i

G06F 9/38 (2006.01) i