

(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
2017 年 3 月 16 日 (16.03.2017)



(10) 国际公布号
WO 2017/041674 A1

- (51) 国际专利分类号:
G06F 9/445 (2006.01)
- (21) 国际申请号: PCT/CN2016/098009
- (22) 国际申请日: 2016 年 9 月 5 日 (05.09.2016)
- (25) 申请语言: 中文
- (26) 公布语言: 中文
- (30) 优先权:
201510573758.9 2015 年 9 月 10 日 (10.09.2015) 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资本大厦一座四层 847 号邮箱, Grand Cayman (KY)。
- (72) 发明人: 才华 (CAI, Hua); 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)。 王家忙 (WANG, Jiamang); 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: 北京三友知识产权代理有限公司 (BEIJING SANYOU INTELLECTUAL PROPERTY AGENCY LTD.); 中国北京市金融街 35 号国际企业大厦 A 座 16 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, 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,

[见续页]

(54) Title: METHOD, DEVICE AND ELECTRONIC EQUIPMENT FOR STARTING BACKUP TASK

(54) 发明名称: 一种启动备份任务的方法、装置及电子设备

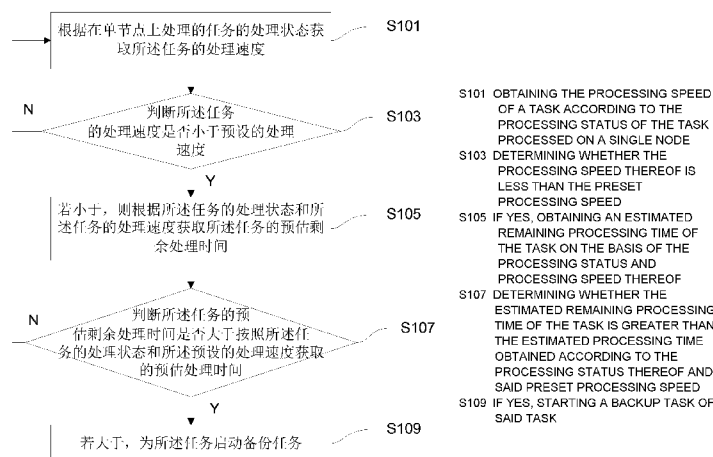


图 1

(57) Abstract: Disclosed are a method, device and electronic equipment for starting a backup task in a distributed data processing system. The method for starting a backup task in a distributed data processing system comprises: obtaining the processing speed of a task according to the processing status of the task processed on a single node (S101), and determining whether the processing speed thereof is less than the preset processing speed (S103); if yes, an estimated remaining processing time of the task is obtained on the basis of the processing status and processing speed thereof (S105), and it is determined whether the estimated remaining processing time of the task is greater than the estimated processing time obtained according to the processing status thereof and said preset processing speed (S107); if yes, a backup task is started for said task (S109). The technical solution solves the problems of the complicated process for preventing long tail delay and unreasonable pre-configuration. Especially during backup task start, the invention can also prevent resource waste in the situation wherein after a backup task has started, the original task is processed before backup task is processed.

(57) 摘要:

[见续页]



WO 2017/041674 A1



RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, **本国际公布:**

CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, — 包括国际检索报告(条约第 21 条(3))。
TG)。

用于分布式数据处理系统中启动备份任务的方法、装置及电子设备。其中，用于分布式数据处理系统中启动备份任务的方法包括：根据在单节点上处理的任务的¹处理状态获取所述任务的¹处理速度（S101），判断所述任务的¹处理速度是否小于预设的¹处理速度（S103），若小于，则根据所述任务的¹处理状态和所述任务的¹处理速度获取所述任务的¹预估剩余处理时间（S105），判断所述任务的¹预估剩余处理时间是否大于按照所述任务的¹处理状态和所述预设的¹处理速度获取的¹预估处理时间（S107），若大于，为所述任务启动备份任务（S109）。所述技术方案解决了避免长尾延迟的过程繁琐，预先配置不合理的问题，特别是在启动备份任务时，也能避免备份任务启动后，原任务早于备份任务处理完毕，资源浪费的问题。

一种启动备份任务的方法、装置及电子设备

本申请要求 2015 年 09 月 10 日递交的申请号为 201510573758.9、发明名称为“一种启动备份任务的方法、装置及电子设备”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

本申请涉及分布式计算技术领域，具体涉及一种用于分布式数据处理系统中启动备份任务的方法以及一种用于分布式数据处理系统中启动备份任务的装置；本申请同时涉及一种电子设备。

10

背景技术

目前互联网应用已经成为人们生活的一部分，这些应用往往要为上亿用户服务，而支撑拥有海量用户互联网应用的各种分布式计算环境也成为不可或缺的基础。在分布式环境下，任务会被划分多个任务切片由许多节点进行分步处理，并且会被部署到多个服务器上。尽管在分布式环境下能够显著提高资源利用率，但同时也会带来任务之间相互干扰的问题。例如在分布式环境下，长尾延迟效应会被进一步放大。

15

在分布式环境下，处理在性能较差的节点上的任务切片的处理时间会明显长于正常机器，变成整个任务处理的瓶颈，大大增加了整个任务的处理时间，这就会出现严重的长尾延迟效应。

20

目前，避免长尾延迟的方案为：给长尾的任务切片建立一个备份任务，处理同样的数据。但是在现有技术下，避免长尾延迟的方案存在如下问题：需要用户来配置参数，用户通过设置参数来决定在何时应该启动备份任务；由于参数是预先设定好的固定的数值，很难做到配置的合理性，即使用户通过大量的实践做了相对合理的配置，但是对于每个任务都要做配置，工作量也是非常巨大的；并且很难全面地考虑到任务处理的实时状态，因为每个任务切片的处理情况跟整个系统的软硬件的实时处理情况具有关联。

25

由此可见，在现有避免长尾延迟的方案下，过程比较繁琐，用户操作起来不方便，预先配置不合理，例如任务数据切片不均匀引入的长尾问题，这种长尾是无法通过预先配置参数启动备份任务来缓解的。特别是如果启动备份任务的时候没有考虑原任务的执行速度，很容易出现备份任务虽然启动了，但是原任务仍然早于备份任务处理完毕，导

致了资源的白白浪费。

发明内容

本申请提供一种用于分布式数据处理系统中启动备份任务的方法以及一种用于分布式数据处理系统中启动备份任务的装置,以解决现有技术中避免长尾延迟的问题。本申请同时涉及一种电子设备。

本申请提供了一种用于分布式数据处理系统中启动备份任务的方法,所述方法包括:
根据在单节点上处理的任务的处理状态获取所述任务的处理速度;

判断所述任务的处理速度是否小于预设的处理速度;

若小于,则根据所述任务的处理状态和所述任务的处理速度获取所述任务的预估剩余处理时间;

判断所述任务的预估剩余处理时间是否大于按照所述任务的处理状态和所述预设的处理速度获取的预估处理时间;

若大于,为所述任务启动备份任务。

可选的,所述根据在单节点上处理的任务的处理状态获取所述任务的处理速度具体是,按照预设的时间间隔,根据在单节点上处理的任务的处理状态获取所述任务的处理速度。

可选的,所述任务的处理状态,包括:所述任务已处理的数据量以及已处理的时间。

可选的,所述根据在单节点上处理的任务的处理状态获取所述任务的处理速度,包括:

获取当前在单节点上处理的任务的已处理的数据量以及已处理的时间;

将所述任务的已处理的数据量和所述任务的已处理的时间的商,作为该任务的处理速度。

可选的,所述根据在单节点上处理的任务的处理状态获取所述任务的处理速度的算法为:

$$\text{AvgInstRate}_N = \text{dataSize}_N / \text{runTime}_N$$

其中,AvgInstRate 代表单节点上处理的任务的处理速度,其数值越大说明任务的处理速度越快;dataSize 代表单节点上处理的任务已处理的数据量,runTime 代表单节点上处理的任务已处理的时间,字母 N 代表某一具体单节点上处理的任务。

可选的,所述预设的处理速度是,当前在所述处理系统中所有节点上处理的任务的

平均处理速度。

可选的，所述预设的处理速度，通过如下方式获取：

计算当前在所述处理系统中所有节点上处理的任务已处理的数据量的和，将所述当前在所述处理系统中所有节点上处理的任务已处理的数据量的和作为所述处理系统中总任务已处理的数据量；

计算当前在所述处理系统中所有节点上处理的任务已处理的时间的和，将所述当前在所述处理系统中所有节点上处理的任务已处理的时间的和作为所述处理系统中总任务已处理的时间；

将所述总任务已处理的数据量和所述总任务已处理的时间的商，作为所述预设的处理速度。

可选的，所述预设的处理速度的算法为：

$$\text{AvgRate} = (\text{dataSize}_1 + \text{dataSize}_2 + \dots + \text{dataSize}_N) / (\text{runTime}_1 + \text{runTime}_2 + \dots + \text{runTime}_N)$$

其中，AvgRate 代表预设的处理速度，其数值越大说明任务的处理速度越快；dataSize 代表单节点上处理的任务已处理的数据量，runTime 代表单节点上处理的任务已处理的时间，该算法中的数字和字母 N 代表某一具体单节点上处理的任务。

可选的，所述任务的处理状态，还包括：所述任务未处理时的原始数据量。

可选的，所述根据所述任务的处理状态和所述任务的处理速度获取所述任务的预估剩余处理时间，包括：

获取所述任务未处理时的原始数据量、所述任务的已处理的数据量以及所述任务的

处理速度；

将所述任务未处理时的原始数据量和所述任务的已处理的数据量的差，作为所述任务的剩余数据量；

将所述任务的剩余数据量和所述任务的处理速度的商，作为所述任务的预估剩余处理时间。

可选的，所述根据所述任务的处理状态和所述任务的处理速度获取所述任务的预估剩余处理时间的算法为：

$$\text{TimeLast}_N = (\text{AllData}_N - \text{dataSize}_N) / \text{AvgInstRate}_N$$

其中，TimeLast 代表单节点上处理的任务的预估剩余处理时间，其数值越大说明处理完任务还需要的时间就越多；dataSize 代表单节点上处理的任务已处理的数据量，AllData 代表单节点上处理的任务未处理时的原始数据量，AvgInstRate 代表单节点上处

理的任务的处理速度，字母 N 代表某一具体单节点上处理的任务。

可选的，所述按照所述任务的处理状态和所述预设的处理速度获取的预估处理时间，包括：

获取所述任务未处理时的原始数据量；

5 获取所述任务在处理该任务的节点上的启动时间；所述启动时间至少包括所述任务申请资源和启动处理该任务的节点的时间；

将所述任务未处理时的原始数据量和所述预设的处理速度的商，作为所述任务的备份任务的初步预估处理时间；

10 将所述任务在处理该任务的节点上的启动时间和所述任务的备份任务的初步预估处理时间的和，作为所述按照所述任务的处理状态和所述预设的处理速度获取的预估处理时间。

可选的，所述按照所述任务的处理状态和所述预设的处理速度获取的预估处理时间的算法为：

$$\text{BackupTime}_N = \text{AllData}_N / \text{AvgRate} + \text{overheadTime}$$

15 其中，BackupTime 代表单节点上处理的任务的预估处理时间，其数值越大说明处理该任务的备份任务时处理时间就越长，AllData 代表单节点上处理的任务未处理时的原始数据量，AvgRate 代表预设的处理速度，字母 N 代表某一具体单节点上处理的任务，overheadTime 代表启动单节点上处理的任务所需要的时间，包括申请资源和节点启动等。

可选的，所述为所述任务启动备份任务，包括：

20 为所述任务，创建该任务的备份任务；

为所述备份任务分配节点；

在所述节点中处理该备份任务。

相应的，本申请还提供了一种用于分布式数据处理系统中启动备份任务的装置，所述装置包括：

25 处理速度获取单元，用于根据在单节点上处理的任务的处理状态获取所述任务的处理速度；

处理速度判断单元，用于判断所述任务的处理速度是否小于预设的处理速度；

预估剩余处理时间获取单元，用于接收所述处理速度判断单元的判断结果，若是，则根据所述任务的处理状态和所述任务的处理速度获取所述任务的预估剩余处理时间；

30 处理时间判断单元，用于判断所述任务的预估剩余处理时间是否大于按照所述任务

的处理状态和所述预设的处理速度获取的预估处理时间；

备份任务启动单元，用于接收所述处理时间判断单元的判断结果，若是，为所述任务启动备份任务。

可选的，所述处理速度获取单元，具体用于按照预设的时间间隔，根据在单节点上处理的任务的已处理的数据量以及已处理的时间，获取所述任务的已处理的数据量以及已处理的时间；

可选的，所述处理速度获取单元，具体用于按照预设的时间间隔，根据在单节点上处理的任务的已处理的数据量以及已处理的时间，获取所述任务的已处理的数据量以及已处理的时间；

可选的，所述处理速度获取单元，包括：

数据量时间获取子单元，用于获取当前在单节点上处理的任务的已处理的数据量以及已处理的时间；

处理速度计算子单元，用于将所述任务的已处理的数据量和所述任务的已处理的时间的商，作为该任务的已处理的数据量；

可选的，所述处理速度判断单元，具体用于判断所述任务的已处理的数据量是否小于当前在所述处理系统中所有节点上处理的任务的平均处理速度。

可选的，所述处理速度判断单元，包括：

节点数据量计算子单元，用于计算当前在所述处理系统中所有节点上处理的任务的已处理的数据量的和，将所述当前在所述处理系统中所有节点上处理的任务的已处理的数据量的和作为所述处理系统中总任务已处理的数据量；

节点时间计算子单元，用于计算当前在所述处理系统中所有节点上处理的任务的已处理的时间的和，将所述当前在所述处理系统中所有节点上处理的任务的已处理的时间的和作为所述处理系统中总任务已处理的时间；

处理速度计算子单元，用于将所述总任务已处理的数据量和所述总任务已处理的时间的商，作为所述预设的处理速度。

可选的，所述处理速度获取单元，具体用于按照预设的时间间隔，根据在单节点上处理的任务的已处理的数据量、已处理的时间以及所述任务未处理时的原始数据量，获取所述任务的已处理的数据量以及已处理的时间；

可选的，所述预估剩余处理时间获取单元，包括：

处理状态获取子单元，用于获取所述任务未处理时的原始数据量、所述任务的已处理的数据量以及所述任务的已处理的时间；

剩余数据量计算子单元，用于将所述任务未处理时的原始数据量和所述任务的已处理的数据量以及所述任务的已处理的时间；

理的数据量的差，作为所述任务的剩余数据量；

剩余处理时间计算子单元，用于将所述任务的剩余数据量和所述任务的处理速度的商，作为所述任务的预估剩余处理时间。

可选的，所述处理时间判断单元，包括：

5 原始数据量获取子单元，用于获取所述任务未处理时的原始数据量；

启动时间获取子单元，用于获取所述任务在处理该任务的节点上的启动时间；所述启动时间至少包括所述任务申请资源和启动处理该任务的节点的时间；

初步预估子单元，用于将所述任务未处理时的原始数据量和所述预设的处理速度的商，作为所述任务的备份任务的初步预估处理时间；

10 处理时间计算子单元，用于将所述任务在处理该任务的节点上的启动时间和所述任务的备份任务的初步预估处理时间的和，作为所述按照所述任务的处理状态和所述预设的处理速度获取的预估处理时间。

可选的，所述备份任务启动单元，包括：

备份任务创建子单元，用于为所述任务，创建该任务的备份任务；

15 节点分配子单元，用于为所述备份任务分配节点；

备份任务处理子单元，用于在所述节点中处理该备份任务。

此外，本申请还提供了一种电子设备，包括：

显示器；

处理器；

20 存储器，用于存储备份任务启动程序，所述程序在被所述处理器读取执行时，执行如下操作：根据在单节点上处理的任务的处理状态获取所述任务的处理速度，判断所述任务的处理速度是否小于预设的处理速度，若小于，则根据所述任务的处理状态和所述任务的
25 处理速度获取所述任务的预估剩余处理时间，判断所述任务的预估剩余处理时间是否大于按照所述任务的处理状态和所述预设的处理速度获取的预估处理时间，若大于，为所述任务启动备份任务。

与现有技术相比，本申请具有以下优点：

本申请提供一种用于分布式数据处理系统中启动备份任务的方法和一种用于分布式数据处理系统中启动备份任务的装置以及电子设备，通过根据在单节点上处理的任务的
30 处理状态获取所述任务的处理速度；判断所述任务的处理速度是否小于预设的处理速度；若小于，则根据所述任务的处理状态和所述任务的处理速度获取所述任务的预估剩

余处理时间；判断所述任务的预估剩余处理时间是否大于按照所述任务的处理状态和所述预设的处理速度获取的预估处理时间；若大于，为所述任务启动备份任务。所述技术方案解决了避免长尾延迟的过程比较繁琐，操作起来不方便，预先配置不合理的问题，特别是在启动备份任务时，也能通过上述方法避免备份任务启动后，原任务仍然早于备份任务处理完毕，导致了资源的白白浪费的问题。

附图说明

为了更清楚地说明本申请实施例或现有技术中的技术方案，下面将对实施例或现有技术描述中所需要使用的附图作简单地介绍，显而易见地，下面描述中的附图仅仅是本申请中记载的一些实施例，对于本领域普通技术人员来讲，还可以根据这些附图获得其他的附图。

图 1 示出了根据本申请的实施例提供的用于分布式数据处理系统中启动备份任务的方法的流程图；

图 2 示出了根据本申请的实施例提供的根据在单节点上处理的任务的处理状态获取所述任务的处理速度的流程图；

图 3 示出了根据本申请的实施例提供的获取预设的处理速度的流程图；

图 4 示出了根据本申请的实施例提供的根据所述任务的处理状态和所述任务的处理速度获取所述任务的预估剩余处理时间的流程图；

图 5 示出了根据本申请的实施例提供的按照所述任务的处理状态和所述预设的处理速度获取的预估处理时间的流程图；

图 6 示出了根据本申请的实施例提供的为所述任务启动备份任务的流程图；

图 7 示出了根据本申请的实施例提供的用于分布式数据处理系统中启动备份任务的装置的示意图；

图 8 示出了根据本申请的实施例提供的电子设备的示意图。

具体实施方式

为了能够更清楚地理解本申请的上述目的、特征和优点，下面结合附图和具体实施方式对本申请进行进一步的详细描述。需要说明的是，在不冲突的情况下，本申请的实施例及实施例中的特征可以相互组合。

在下面的描述中阐述了很多具体细节以便于充分理解本申请。但是，本申请能够以

很多不同于在此描述的其它方式来实施，本领域技术人员可以在不违背本申请内涵的情况下做类似推广，因此，本申请不受下面公开的具体实施的限制。

本申请的实施例提供了一种用于分布式数据处理系统中启动备份任务的方法以及一种用于分布式数据处理系统中启动备份任务的装置;本申请同时涉及一种电子设备。在下
5 面的实施例中逐一进行详细说明。

目前，避免长尾延迟的方案为：给长尾的任务切片建立一个备份任务，处理同样的数据。但是，避免长尾延迟时需要用户来配置参数，用户通过设置参数来决定在何时应该启动备份任务。由此可见，在现有避免长尾延迟的方案下，过程比较繁琐，用户操作起来不方便，预先配置不合理，例如任务数据切片不均匀引入的长尾问题，这种长尾是
10 无法通过预先配置参数启动备份任务来缓解的。特别是如果启动备份任务的时候没有考虑原任务的执行速度，很容易出现备份任务虽然启动了，但是原任务仍然早于备份任务处理完毕，导致了资源的白白浪费。针对这一问题，本申请的技术方案通过判断所述任务的处理速度是否小于预设的处理速度和判断所述任务的预估剩余处理时间是否大于按照所述任务的处理状态和所述预设的处理速度获取的预估处理时间，从而实现了避免长
15 尾延迟的功能。

在详细描述本实施例的具体步骤之前，先对本技术方案涉及的长尾延迟作简要说明。

在分布式系统中，将任务分为有限数量的子任务，并在不同节点上处理，由有限数量的子任务组成的分布式系统执行单元称为分布式作业，但是在大型分布式系统中，处理任务的机器软硬件配置各不相同，还具有外界干扰因素，因此导致每个子任务的处
20 理时间不同，运行在性能较差的节点上的子任务的处理时间会明显长于其他机器处理子任务的处理时间，变成整个任务运行的瓶颈，大大增加了整个任务的运行时间，这就是长尾延迟现象。产生长尾延迟现象的原因有：在机器中运行的子任务较多时导致的系统繁忙，使该子任务无法正常申请到资源而导致子任务运行缓慢；子任务由于依赖其他节点的子任务所产生的数据，而这些节点也可能由于系统繁忙导致影响子任务的运行缓慢；
25 对于数据传输密集型的子任务，由于网络带宽较小，导致子任务传输网络数据的速度缓慢；机器中数据磁盘无法正常读写、系统盘读写缓慢等硬件故障，或者操作系统的某些故障，导致子任务运行缓慢。

本申请的实施例提供了一种用于分布式数据处理系统中启动备份任务的方法。由于分布式系统中，将具体的任务分为有限数量的子任务，并在不同节点上处理，因此在本
30 实施例中重点描述单一子任务在单一节点上启动备份任务的实施方式，为了描述方便，

在下文实施例的描述中对子任务简称为任务；对划分为多个子任务的任务称为总任务。
所述用于分布式数据处理系统中启动备份任务的方法实施例如下：

请参考图 1，其示出了根据本申请的实施例提供的用于分布式数据处理系统中启动备份任务的方法的流程图。

5 所述用于分布式数据处理系统中启动备份任务的方法包括：

步骤 S101，根据在单节点上处理的任务的处理状态获取所述任务的处理速度。

在本实施例中，所述根据在单节点上处理的任务的处理状态获取所述任务的处理速度，可以采用如下方式实现：按照预设的时间间隔，根据在单节点上处理的任务的处理状态获取所述任务的处理速度。所述预设的时间间隔可以是任意时间间隔，但是时间间隔
10 隔过短会使分布式系统浪费额外的资源，时间间隔过长会增加后续处理中的误差，在本实施例中提供了较理想的时间间隔，例如：可以每间隔 15 秒，根据在单节点上处理的任务的处理状态，获取一次所述任务的处理速度。

需要说明的是，所述在单节点上处理的任务的处理状态，包括：所述任务已处理的数据量以及已处理的时间。

15 可以理解的，每间隔 15 秒根据在单节点上处理的任务的当前已处理的数据量以及已处理的时间，获取所述在单节点上处理的任务的处理速度。

在步骤 S101 中，所述根据在单节点上处理的任务的处理状态获取所述任务的处理速度，具体包括步骤 S101-1 至 S101-2，下面结合图 2 作进一步说明。

请参考图 2，其示出了根据本申请的实施例提供的根据在单节点上处理的任务的处
20 理状态获取所述任务的处理速度的流程图。

步骤 S101-1，获取当前在单节点上处理的任务的已处理的数据量以及已处理的时间。

所述获取当前在单节点上处理的任务的已处理的数据量以及已处理的时间，可以采用如下方式实现：根据每间隔 15 秒获取在单节点上处理的任务已处理的数据量以及已处理的时间。例如：当前在单节点上处理的任务的已处理的数据量为 1m，已处理的时间为
25 2s。

步骤 S101-2，将所述任务的已处理的数据量和所述任务的已处理的时间的商，作为该任务的处理速度。

例如：当前在单节点上处理的任务的已处理的数据量为 1m，已处理的时间为 2s，将所述任务的已处理的数据量 1m 除以所述任务已处理的时间 2s，获取的所述任务的处
30 理速度为 0.5m/s。

在具体实施时，所述根据在单节点上处理的任务的处理状态获取所述任务的处理速度的算法为：

$$\text{AvgInstRate}_N = \text{dataSize}_N / \text{runTime}_N$$

其中，AvgInstRate 代表单节点上处理的任务的处理速度，其数值越大说明任务的处
理速度越快；dataSize 代表单节点上处理的任务已处理的数据量，runTime 代表单节点上
处理的任务已处理的时间，字母 N 代表某一具体单节点上处理的任务。

例如：所述任务是在第一个节点上处理的任务，则当前在节点上处理的所述任务的
已处理的数据量 dataSize1 为 1m，已处理的时间 runTime1 为 2s，根据上述处理速度的算
法计算出的所述任务的处理速度 AvgInstRate1 为 0.5m/s。

在步骤 S101 中，获取所述任务的处理速度，可以有两种方式：

一是，根据每间隔 15 秒获取在单节点上处理的任务已处理的数据量以及已处理的时
间，在单节点上根据所述任务的处理速度的算法计算出所述任务的处理速度后，上报所
述任务的处理速度。例如：将所述任务的处理速度上报给 master。

需要说明的是，master 是一种在分布式系统中并行计算结构，master 与 slave 之间保
持心跳，并且互相通信传送任务，最后由 master 汇集结果。

二是，在单节点上，根据每间隔 15 秒获取在单节点上处理的任务已处理的数据量以
及已处理的时间，上报所述任务已处理的数据量以及已处理的时间后，由 master 根据上
报的所述任务已处理的数据量以及已处理的时间，根据所述任务的处理速度的算法计算
出所述任务的处理速度。

需要说明的是，在本步骤 S101 中，若采用第一种方式获取所述任务的处理速度，则
在后续的步骤中，就都在单节点上处理并将处理后的结果上报 master；同样的道理，若
采用第二种方式获取所述任务的处理速度，则在后续的步骤中，就都上报所述任务的处
理状态之后，在 master 中处理计算。在本实施例中，以第二种方式进行举例说明。

步骤 S103，判断所述任务的处理速度是否小于预设的处理速度。

在本实施例中，所述预设的处理速度是指：当前在所述处理系统中所有节点上处理
的任务的平均处理速度。

可以理解的，由于分布式系统中，将任务分为有限数量的子任务，并在不同节点上
处理，因此在每一节点上处理的任务的数据量和处理时间都不相同，则每一节点上处理
的任务的处理速度都不相同。

在步骤 S103 中，所述预设的处理速度可以通过如下步骤获取，具体包括步骤 S102-1

至 S102-3, 下面结合图 3 作进一步说明。

请参考图 3, 其示出了根据本申请的实施例提供的获取预设的处理速度的流程图。

步骤 S102-1, 计算当前在所述处理系统中所有节点上处理的任务已处理的数据量的和, 将所述当前在所述处理系统中所有节点上处理的任务已处理的数据量的和作为所述处理系统中总任务已处理的数据量。

根据上报给 master 的每一节点上处理的任务已处理的数据量, 获取在分布式系统中每一节点上处理的任务已处理的数据量, 并根据已获取的每一节点上处理的任务已处理的数据量计算在当前所有节点上处理的任务已处理的数据量的和。例如: 在分布式系统中, 共有 3 个任务分别在 3 个不同节点上处理, 第一个节点上处理的任务在当前已处理的数据量为 1m, 第二个节点上处理的任务在当前已处理的数据量为 2m, 第三个节点上处理的任务在当前已处理的数据量为 3m, 则当前在所述处理系统中所有节点上处理的任务已处理的数据量的和为 6m, 将 6m 作为所述处理系统中总任务已处理的数据量。

步骤 S102-2, 计算当前在所述处理系统中所有节点上处理的任务已处理的时间的和, 将所述当前在所述处理系统中所有节点上处理的任务已处理的时间的和作为所述处理系统中总任务已处理的时间。

根据上报给 master 的每一节点上处理的任务已处理的时间, 获取在分布式系统中每一节点上处理的任务已处理的时间, 并根据已获取的每一节点上处理的任务已处理的时间计算在当前所有节点上处理的任务已处理的时间的和。例如: 在分布式系统中, 共有 3 个任务分别在 3 个不同节点上处理, 第一个节点上处理的任务在当前已处理的时间为 2s, 第二个节点上处理的任务在当前已处理的时间为 2s, 第三个节点上处理的任务在当前已处理的时间为 1s, 则当前在所述处理系统中所有节点上处理的任务已处理的时间的和为 5s, 将 5s 作为所述处理系统中总任务已处理的时间。

步骤 S102-3, 将所述总任务已处理的数据量和所述总任务已处理的时间的商, 作为所述预设的处理速度。

例如: 在分布式系统中, 总任务已处理的数据量为 6m, 总任务已处理的时间为 5s, 将所述总任务已处理的数据量 6m 除以所述总任务已处理的时间 5s, 获取所述总任务的处理速度为 1.2m/s, 将 1.2m/s 作为所述预设的处理速度。

在具体实施时, 所述预设的处理速度的算法为:

$$\text{AvgRate} = (\text{dataSize}_1 + \text{dataSize}_2 + \cdots + \text{dataSize}_N) / (\text{runTime}_1 + \text{runTime}_2 + \cdots + \text{runTime}_N)$$

其中, AvgRate 代表预设的处理速度, 其数值越大说明任务的处理速度越快; dataSize

代表单节点上处理的任务已处理的数据量，runTime 代表单节点上处理的任务已处理的时间，该算法中的数字和字母 N 代表某一具体单节点上处理的任务。

例如：在分布式系统中，共有 3 个任务分别在 3 个不同节点上处理，第一个节点上处理的任务在当前已处理的数据量 dataSize₁ 为 1m，第二个节点上处理的任务在当前已处理的数据量 dataSize₂ 为 2m，第三个节点上处理的任务在当前已处理的数据量 dataSize₃ 为 3m，则当前在所述处理系统中所有节点上处理的任务已处理的数据量 dataSize₁+dataSize₂+dataSize₃ 的和为 6m，将 6m 作为所述处理系统中总任务已处理的数据量；第一个节点上处理的任务在当前已处理的时间 runTime₁ 为 2s，第二个节点上处理的任务在当前已处理的时间 runTime₂ 为 2s，第三个节点上处理的任务在当前已处理的时间 runTime₃ 为 1s，则当前在所述处理系统中所有节点上处理的任务已处理的时间 runTime₁+runTime₂+runTime₃ 的和为 5s，根据上述预设的处理速度的算法计算出的预设的处理速度 AvgRate 为 1.2 m/s。

在本实施例中，所述判断所述任务的处理速度是否小于预设的处理速度，可以采用如下方式实现：接收步骤 S101 中，根据在单节点上处理的任务的处理状态获取所述任务的 15 处理速度，与所述预设的处理速度进行对比判断步骤 S101 中获取的所述任务的处理速度是否小于所述预设的处理速度，即：判断步骤 S101 中获取的所述任务的处理速度是否小于当前在所述处理系统中所有节点上处理的任务的平均处理速度。

步骤 S105，若小于，则根据所述任务的处理状态和所述任务的处理速度获取所述任务的预估剩余处理时间。

20 如果步骤 S101 中，根据在单节点上处理的任务的处理状态获取所述任务的处理速度小于预设的处理速度，说明所述任务的处理速度小于当前在所述处理系统中所有节点上处理的任务的平均处理速度，这种情况下执行步骤 S105。

在本实施例中，所述根据所述任务的处理状态和所述任务的处理速度获取所述任务的预估剩余处理时间，可以采用如下方式实现：根据当前所述任务的已处理的数据量、 25 所述任务未处理时的原始数据量以及在步骤 S101 中获取的所述任务的处理速度，获取所述任务的预估剩余处理时间。在本实施例中，所述任务的处理状态，还包括：所述任务未处理时的原始数据量。

在步骤 S105 中，所述根据所述任务的处理状态和所述任务的处理速度获取所述任务的预估剩余处理时间，具体包括步骤 S105-1 至 S105-3，下面结合图 4 作进一步说明。

30 请参考图 4，其示出了根据本申请的实施例提供的根据所述任务的处理状态和所述

任务的处理速度获取所述任务的预估剩余处理时间的流程图。

步骤 S105-1, 获取所述任务未处理时的原始数据量、所述任务的已处理的数据量以及所述任务的处理速度。

所述获取所述任务未处理时的原始数据量、所述任务的已处理的数据量以及所述任务的处理速度, 可以采用如下方式实现: 根据每间隔 15 秒获取在单节点上处理的所述任务未处理时的原始数据量、所述任务的已处理的数据量以及所述任务的处理速度。例如: 当前在单节点上处理的任务的已处理的数据量为 1m, 未处理时的原始数据量为 9m, 所述任务的处理速度为 0.5m/s。

步骤 S105-2, 将所述任务未处理时的原始数据量和所述任务的已处理的数据量的差, 作为所述任务的剩余数据量。

例如: 当前在单节点上处理的任务的已处理的数据量为 1m, 未处理时的原始数据量为 9m, 将所述未处理时的原始数据量为 9m 减去所述任务的已处理的数据量为 1m, 获取所述任务的剩余数据量为 8m。

步骤 S105-3, 将所述任务的剩余数据量和所述任务的处理速度的商, 作为所述任务的预估剩余处理时间。

例如: 当前在单节点上处理的任务的剩余数据量为 8m, 根据步骤 S101 中获取到的所述任务的处理速度为 0.5m/s, 将所述任务的剩余数据量为 8m 除以所述任务的处理速度 0.5m/s, 获取所述任务的预估剩余处理时间 16s。

在具体实施时, 所述根据所述任务的处理状态和所述任务的处理速度获取所述任务的预估剩余处理时间的算法为:

$$\text{TimeLast}_N = (\text{AllData}_N - \text{dataSize}_N) / \text{AvgInstRate}_N$$

其中, TimeLast 代表单节点上处理的任务的预估剩余处理时间, 其数值越大说明处理完任务还需要的时间就越多; dataSize 代表单节点上处理的任務已处理的数据量, AllData 代表单节点上处理的任務未处理时的原始数据量, AvgInstRate 代表单节点上处理的任務的处理速度, 字母 N 代表某一具体单节点上处理的任務。

例如: 所述任务是在第一个节点上处理的任務, 则当前在节点上处理的所述任务的已处理的数据量 dataSize_1 为 1m, 未处理时的原始数据量 AllData_1 为 9m, 在步骤 S101 中获取到的所述任务的處理速度为 0.5m/s, 根据上述根据所述任务的處理状态和所述任务的處理速度获取所述任务的预估剩余处理时间的算法计算出的所述任务的预估剩余处理时间 TimeLast_1 为 16s。

步骤 S107, 判断所述任务的预估剩余处理时间是否大于按照所述任务的处理状态和所述预设的处理速度获取的预估处理时间。

在本实施例中, 所述按照所述任务的处理状态和所述预设的处理速度获取的预估处理时间是指: 启动所述任务的备份任务至完成该备份任务时的预估处理时间。

5 可以理解的, 由于分布式系统中, 将任务分为有限数量的子任务, 并在不同节点上处理, 由于节点会受到机器系统繁忙、网络带宽或数据磁盘系统盘读写缓慢的影响, 导致该节点处理任务的处理速度缓慢, 造成长尾延迟现象, 若将所述任务备份到其他机器中的节点上处理时, 就可能避免上述问题的发生。将所述任务备份到其他机器中的节点上处理就是启动备份任务的过程。

10 本步骤就是对所述任务估算启动备份任务时的处理时间, 在步骤 S107 中, 所述按照所述任务的处理状态和所述预设的处理速度获取的预估处理时间, 具体包括步骤 S107-1 至 S107-4, 下面结合图 5 作进一步说明。

请参考图 5, 其示出了根据本申请的实施例提供的按照所述任务的处理状态和所述预设的处理速度获取的预估处理时间的流程图。

15 步骤 S107-1, 获取所述任务未处理时的原始数据量。

所述获取所述任务未处理时的原始数据量, 可以采用如下方式实现: 每间隔 15 秒获取在单节点上处理的所述任务未处理时的原始数据量。例如: 当前在单节点上处理的任

务未处理时的原始数据量为 9m。

20 步骤 S107-2, 获取所述任务在处理该任务的节点上的启动时间; 所述启动时间至少包括所述任务申请资源和启动处理该任务的节点的时间。

在本实施例中, 所述获取所述任务在处理该任务的节点上的启动时间, 可以采用如下方式实现: 将所述任务在启动时申请资源的时间和启动处理所述任务的节点的时间的和, 作为所述任务在处理该任务的节点上的启动时间。例如: 所述任务在启动时申请资源的时间为 2s, 启动处理所述任务的节点的时间为 1s, 则所述任务的启动时间为 3s。

25 步骤 S107-3, 将所述任务未处理时的原始数据量和所述预设的处理速度的商, 作为所述任务的备份任务的初步预估处理时间。

例如: 当前在单节点上处理的任

务未处理时的原始数据量为 9m, 在步骤 S102 中获取的所述预设的处理速度为 1.2m/s, 将所述任务未处理时的原始数据量 9m 除以所述预设的处理速度 1.2m/s, 获取的所述任务的备份任务的初步预估处理时间为 7.5s。

30 步骤 S107-4, 将所述任务在处理该任务的节点上的启动时间和所述任务的备份任务

的初步预估处理时间的和，作为所述按照所述任务的处理状态和所述预设的处理速度获取的预估处理时间。

例如：当前在单节点上处理的任务的启动时间为 3s，所述任务的备份任务的初步预估处理时间为 7.5s，将所述任务的启动时间 3s 加上所述任务的备份任务的初步预估处理时间 7.5s，作为所述按照所述任务的处理状态和所述预设的处理速度获取的预估处理时间，即：启动所述任务的备份任务至完成该备份任务时的预估处理时间为 10.5s。

在具体实施时，所述按照所述任务的处理状态和所述预设的处理速度获取的预估处理时间的算法为：

$$\text{BackupTime}_N = \text{AllData}_N / \text{AvgRate} + \text{overheadTime}$$

其中，BackupTime 代表单节点上处理的任务的预估处理时间，其数值越大说明处理该任务的备份任务时处理时间就越长，AllData 代表单节点上处理的未处理时的原始数据量，AvgRate 代表预设的处理速度，字母 N 代表某一具体单节点上处理的备份任务，overheadTime 代表启动单节点上处理的备份任务所需要的时间，包括申请资源和节点启动等。

例如：所述任务是在第一个节点上处理的备份任务，则当前在节点上处理的所述任务未处理时的原始数据量 AllData₁ 为 9m，在步骤 S102 中获取的所述预设的处理速度 AvgRate 为 1.2m/s，在步骤 S107-2 中获取所述任务在处理该任务的节点上的启动时间 overheadTime 为 3s，根据上述按照所述任务的处理状态和所述预设的处理速度获取的预估处理时间的算法计算出的按照所述任务的处理状态和所述预设的处理速度获取的预估处理时间 BackupTime₁ 为 10.5s，即：启动所述任务的备份任务至完成该备份任务时的预估处理时间为 10.5s

在本实施例中，所述判断所述任务的预估剩余处理时间是否大于按照所述任务的处理状态和所述预设的处理速度获取的预估处理时间，可以采用如下方式实现：接收步骤 S105 中，根据所述任务的处理状态和所述任务的备份任务的备份速度获取所述任务的预估剩余处理时间与按照所述任务的处理状态和所述预设的处理速度获取的预估处理时间进行对比，判断步骤 S105 中获取的所述任务的预估剩余处理时间是否大于按照所述任务的备份任务的备份速度获取的预估处理时间，即：判断步骤 S105 中获取的所述任务的预估剩余处理时间是否大于启动所述任务的备份任务至完成该备份任务时的预估处理时间。

步骤 S109，若大于，为所述任务启动备份任务。

如果步骤 S105 中，根据所述任务的备份任务的备份速度获取所述任务的

预估剩余处理时间大于按照所述任务的处理状态和所述预设的处理速度获取的预估处理时间，说明所述任务的预估剩余处理时间是否大于启动所述任务的备份任务至完成该备份任务时的预估处理时间，这种情况下执行步骤 S109。

在步骤 S109 中，所述为所述任务启动备份任务，具体包括步骤 S109-1 至 S109-3，

5 下面结合图 6 作进一步说明。

请参考图 6，其示出了根据本申请的实施例提供的为所述任务启动备份任务的流程图。

步骤 S109-1，为所述任务，创建该任务的备份任务。

在本实施例中，所述备份任务是指：与原有的所述任务处理的数据相同，且具有相同处理逻辑的任务。

步骤 S109-2，为所述备份任务分配节点。

所述为所述备份任务分配节点，可以采用如下方式实现：选取与处理原有的所述任务不同节点，该节点用于处理所述备份任务。所述节点可以是，在相同机器中未进行处理任务的节点、在分布式系统下的其他机器中的节点。

15 步骤 S109-3，在所述节点中处理该备份任务。

所述在所述节点中处理该备份任务，可以采用如下方式实现：所述备份任务启动所述节点，并在在所述节点中申请该备份任务需要的资源，运行所述备份任务。

通过本申请实施例提供的一种用于分布式数据处理系统中启动备份任务的方法，通过判断所述任务的处理速度是否小于预设的处理速度和判断所述任务的预估剩余处理时间是否大于按照所述任务的处理状态和所述预设的处理速度获取的预估处理时间，简化了用户预先配置的时间，并且考虑了原有任务的处理速度，避免了启动所述任务的备份任务后，原有任务早于备份任务处理完毕从而造成的资源浪费的问题。

在上述的实施例中，提供了一种用于分布式数据处理系统中启动备份任务的方法，与上述用于分布式数据处理系统中启动备份任务的方法相对应的，本申请还提供了一种用于分布式数据处理系统中启动备份任务的装置。由于装置的实施例基本相似于方法的实施例，所以描述得比较简单，相关之处参见方法实施例的部分说明即可。下述描述的装置实施例仅仅是示意性的。所述用于分布式数据处理系统中启动备份任务的装置实施例如下：

请参考图 7，其示出了根据本申请的实施例提供的用于分布式数据处理系统中启动备份任务的装置的示意图。

所述用于分布式数据处理系统中启动备份任务的装置,包括:处理速度获取单元 701、处理速度判断单元 703、预估剩余处理时间获取单元 705、处理时间判断单元 707 以及备份任务启动单元 709;

所述处理速度获取单元 701, 用于根据在单节点上处理的任务的⁵处理状态获取所述任务的⁵处理速度;

所述处理速度判断单元 703, 用于判断所述任务的⁵处理速度是否小于预设的处理速度;

所述预估剩余处理时间获取单元 705, 用于接收所述处理速度判断单元 205 的判断结果, 若是, 则根据所述任务的⁵处理状态和所述任务的⁵处理速度获取所述任务的¹⁰预估剩余处理时间;

所述处理时间判断单元 707, 用于判断所述任务的¹⁰预估剩余处理时间是否大于按照所述任务的¹⁰处理状态和所述预设的¹⁰处理速度获取的¹⁰预估处理时间;

所述备份任务启动单元 709, 用于接收所述处理时间判断单元 209 的判断结果, 若是, 为所述任务启动备份任务。

¹⁵ 可选的, 所述处理速度获取单元 701, 具体用于按照预设的时间间隔, 根据在单节点上处理的任务的¹⁵处理状态获取所述任务的¹⁵处理速度。

可选的, 所述处理速度获取单元 701, 具体用于按照预设的时间间隔, 根据在单节点上处理的任务的¹⁵已处理的数据量以及已处理的时间, 获取所述任务的¹⁵处理速度。

可选的, 所述处理速度获取单元 701, 包括: 数据量时间获取子单元以及处理速度²⁰计算子单元;

所述数据量时间获取子单元, 用于获取当前在单节点上处理的任务的²⁰已处理的数据量以及已处理的时间;

所述处理速度计算子单元, 用于将所述任务的²⁰已处理的数据量和所述任务的²⁰已处理的时间的商, 作为该任务的²⁰处理速度。

²⁵ 可选的, 所述处理速度判断单元 703, 具体用于判断所述任务的²⁵处理速度是否小于当前在所述处理系统中所有节点上处理的任务的²⁵平均处理速度。

可选的, 所述处理速度判断单元 703, 包括: 节点数据量计算子单元、节点时间计算子单元以及处理速度计算子单元;

所述节点数据量计算子单元, 用于计算当前在所述处理系统中所有节点上处理的³⁰任务已处理的数据量的和, 将所述当前在所述处理系统中所有节点上处理的³⁰任务已处理的

数据量的和作为所述处理系统中总任务已处理的数据量；

所述节点时间计算子单元，用于计算当前在所述处理系统中所有节点上处理的任务已处理的时间的和，将所述当前在所述处理系统中所有节点上处理的任务已处理的时间的和作为所述处理系统中总任务已处理的时间；

- 5 所述处理速度计算子单元，用于将所述总任务已处理的数据量和所述总任务已处理的时间的商，作为所述预设的处理速度。

可选的，所述处理速度获取单元 701，具体用于按照预设的时间间隔，根据在单节点上处理的任务的已处理的数据量、已处理的时间以及所述任务未处理时的原始数据量，获取所述任务的处理速度。

- 10 可选的，所述预估剩余处理时间获取单元 705，包括：处理状态获取子单元、剩余数据量计算子单元以及剩余处理时间计算子单元；

所述处理状态获取子单元，用于获取所述任务未处理时的原始数据量、所述任务的已处理的数据量以及所述任务的处理速度；

- 15 所述剩余数据量计算子单元，用于将所述任务未处理时的原始数据量和所述任务的已处理的数据量的差，作为所述任务的剩余数据量；

所述剩余处理时间计算子单元，用于将所述任务的剩余数据量和所述任务的处理速度的商，作为所述任务的预估剩余处理时间。

可选的，所述处理时间判断单元 707，包括：原始数据量获取子单元、启动时间获取子单元、初步预估子单元以及处理时间计算子单元；

- 20 所述原始数据量获取子单元，用于获取所述任务未处理时的原始数据量；

所述启动时间获取子单元，用于获取所述任务在处理该任务的节点上的启动时间；所述启动时间至少包括所述任务申请资源和启动处理该任务的节点的时间；

所述初步预估子单元，用于将所述任务未处理时的原始数据量和所述预设的处理速度的商，作为所述任务的备份任务的初步预估处理时间；

- 25 所述处理时间计算子单元，用于将所述任务在处理该任务的节点上的启动时间和所述任务的备份任务的初步预估处理时间的和，作为所述按照所述任务的处理状态和所述预设的处理速度获取的预估处理时间。

可选的，所述备份任务启动单元 709，包括：

所述备份任务创建子单元，用于为所述任务，创建该任务的备份任务；

- 30 所述节点分配子单元，用于为所述备份任务分配节点；

所述备份任务处理子单元，用于在所述节点中处理该备份任务。

在上述的实施例中，提供了一种用于分布式数据处理系统中启动备份任务的方法以及一种用于分布式数据处理系统中启动备份任务的装置，此外，本申请还提供了一种电子设备；所述电子设备实施例如下：

5 请参考图 8，其示出了根据本申请的实施例提供的电子设备的示意图。

所述电子设备，包括：显示器 801、处理器 803 以及存储器 805；

所述存储器 805，用于存储备份任务启动程序，所述程序在被所述处理器读取执行时，
执行如下操作：根据在单节点上处理的任务的处理状态获取所述任务的处理速度，判断所
述任务的处理速度是否小于预设的处理速度，若小于，则根据所述任务的处理状态和所
10 述任务的处理速度获取所述任务的预估剩余处理时间，判断所述任务的预估剩余处理时
间是否大于按照所述任务的处理状态和所述预设的处理速度获取的预估处理时间，若大
于，为所述任务启动备份任务。

在上述的实施例中，提供了一种用于分布式数据处理系统中启动备份任务的方法、
一种用于分布式数据处理系统中启动备份任务的装置以及一种电子设备，此外，本申请
15 还提供了一种分布式处理系统数据处理的方法；所述分布式处理系统数据处理的方法，
包括：

系统中有至少一个在节点上运行的任务，所述任务采用如上述的用于分布式数据处理
系统中启动备份任务的方法。

基于上述可以理解的是，本申请提供的一种分布式处理系统数据处理的方法，该方
20 法中在分布式系统中包含多个在各节点上处理的任務，每一在各节点上处理的任務在处
理时，都可以参考上述实施例中的用于分布式数据处理系统中启动备份任务的方法中的
说明启动对应任务的备份任务，此处不再赘述。

例如：在分布式系统中共有 2 个任务分别在 2 个不同节点上处理，可以参考上述实
施例中的用于分布式数据处理系统中启动备份任务的方法中的说明，根据在每一节点上
25 处理的任務的处理状态获取对应的每一任務的处理速度。

在一个典型的配置中，计算设备包括一个或多个处理器(CPU)、输入/输出接口、网
络接口和内存。

内存可能包括计算机可读介质中的非永久性存储器，随机存取存储器 (RAM) 和/或
非易失性内存等形式，如只读存储器 (ROM) 或闪存(flash RAM)。内存是计算机可读介
30 质的示例。

1、计算机可读介质包括永久性和非永久性、可移动和非可移动媒体可以由任何方法或技术来实现信息存储。信息可以是计算机可读指令、数据结构、程序的模块或其他数据。计算机的存储介质的例子包括，但不限于相变内存 (PRAM)、静态随机存取存储器 (SRAM)、动态随机存取存储器 (DRAM)、其他类型的随机存取存储器 (RAM)、只读存储器 (ROM)、电可擦除可编程只读存储器 (EEPROM)、快闪记忆体或其他内存技术、只读光盘只读存储器 (CD-ROM)、数字多功能光盘 (DVD) 或其他光学存储、磁盒式磁带，磁带磁磁盘存储或其他磁性存储设备或任何其他非传输介质，可用于存储可以被计算设备访问的信息。按照本文中的界定，计算机可读介质不包括非暂存电脑可读媒体 (transitory media)，如调制的数据信号和载波。

2、本领域技术人员应明白，本申请的实施例可提供为方法、系统或计算机程序产品。因此，本申请可采用完全硬件实施例、完全软件实施例或结合软件和硬件方面的实施例的形式。而且，本申请可采用在一个或多个其中包含有计算机可用程序代码的计算机可用存储介质（包括但不限于磁盘存储器、CD-ROM、光学存储器等）上实施的计算机程序产品的形式。

本申请虽然以较佳实施例公开如上，但其并不是用来限定本申请，任何本领域技术人员在不脱离本申请的精神和范围内，都可以做出可能的变动和修改，因此本申请的保护范围应当以本申请权利要求所界定的范围为准。

权 利 要 求 书

1、一种用于分布式数据处理系统中启动备份任务的方法，其特征在于，包括：

根据在单节点上处理的任务的处理状态获取所述任务的处理速度；

判断所述任务的处理速度是否小于预设的处理速度；

5 若小于，则根据所述任务的处理状态和所述任务的处理速度获取所述任务的预估剩余处理时间；

判断所述任务的预估剩余处理时间是否大于按照所述任务的处理状态和所述预设的处理速度获取的预估处理时间；

若大于，为所述任务启动备份任务。

10 2、根据权利要求 1 所述的用于分布式数据处理系统中启动备份任务的方法，其特征在于，所述根据在单节点上处理的任务的处理状态获取所述任务的处理速度具体是，按照预设的时间间隔，根据在单节点上处理的任务的处理状态获取所述任务的处理速度。

3、根据权利要求 1 所述的用于分布式数据处理系统中启动备份任务的方法，其特征在于，所述任务的处理状态，包括：所述任务已处理的数据量以及已处理的时间。

4、根据权利要求 3 所述的用于分布式数据处理系统中启动备份任务的方法，其特征在于，所述根据在单节点上处理的任务的处理状态获取所述任务的处理速度，包括：

获取当前在单节点上处理的任务的已处理的数据量以及已处理的时间；

20 将所述任务的已处理的数据量和所述任务的已处理的时间的商，作为该任务的处理速度。

5、根据权利要求 4 所述的用于分布式数据处理系统中启动备份任务的方法，其特征在于，所述根据在单节点上处理的任务的处理状态获取所述任务的处理速度的算法为：

$$\text{AvgInstRate}_N = \text{dataSize}_N / \text{runTime}_N$$

25 其中，AvgInstRate 代表单节点上处理的任务的处理速度，其数值越大说明任务的处理速度越快；dataSize 代表单节点上处理的任务已处理的数据量，runTime 代表单节点上处理的任务已处理的时间，字母 N 代表某一具体单节点上处理的任务。

6、根据权利要求 3 所述的用于分布式数据处理系统中启动备份任务的方法，其特征在于，所述预设的处理速度是，当前在所述处理系统中所有节点上处理的任务的平均

处理速度。

7、根据权利要求 6 所述的用于分布式数据处理系统中启动备份任务的方法，其特征在于，所述预设的处理速度，通过如下方式获取：

计算当前在所述处理系统中所有节点上处理的任务已处理的数据量的和，将所述当前在所述处理系统中所有节点上处理的任务已处理的数据量的和作为所述处理系统中总任务已处理的数据量；

计算当前在所述处理系统中所有节点上处理的任务已处理的时间的和，将所述当前在所述处理系统中所有节点上处理的任务已处理的时间的和作为所述处理系统中总任务已处理的时间；

10 将所述总任务已处理的数据量和所述总任务已处理的时间的商，作为所述预设的处理速度。

8、根据权利要求 7 所述的用于分布式数据处理系统中启动备份任务的方法，其特征在于，所述预设的处理速度的算法为：

$$\text{AvgRate} = (\text{dataSize}_1 + \text{dataSize}_2 + \cdots + \text{dataSize}_N) / (\text{runTime}_1 + \text{runTime}_2 + \cdots + \text{runTime}_N)$$

15 其中，AvgRate 代表预设的处理速度，其数值越大说明任务的处理速度越快；dataSize 代表单节点上处理的任务已处理的数据量，runTime 代表单节点上处理的任务已处理的时间，该算法中的数字和字母 N 代表某一具体单节点上处理的任

9、根据权利要求 3 所述的用于分布式数据处理系统中启动备份任务的方法，其特征在于，所述任务的处理状态，还包括：所述任务未处理时的原始数据量。

20 10、根据权利要求 9 所述的用于分布式数据处理系统中启动备份任务的方法，其特征在于，所述根据所述任务的处理状态和所述任务的处理速度获取所述任务的预估剩余处理时间，包括：

获取所述任务未处理时的原始数据量、所述任务的已处理的数据量以及所述任务的

25 将所述任务未处理时的原始数据量和所述任务的已处理的数据量的差，作为所述任务的剩余数据量；

将所述任务的剩余数据量和所述任务的处理速度的商，作为所述任务的预估剩余处理时间。

30 11、根据权利要求 10 所述的用于分布式数据处理系统中启动备份任务的方法，其特征在于，所述根据所述任务的处理状态和所述任务的处理速度获取所述任务的预估剩

余处理时间的算法为：

$$\text{TimeLast}_N = (\text{AllData}_N - \text{dataSize}_N) / \text{AvgInstRate}_N$$

其中，TimeLast 代表单节点上处理的任务的预估剩余处理时间，其数值越大说明处理完任务还需要的时间就越多；dataSize 代表单节点上处理的任務已处理的数据量，AllData 代表单节点上处理的任務未处理时的原始数据量，AvgInstRate 代表单节点上处理的任務的处理速度，字母 N 代表某一具体单节点上处理的任務。

12、根据权利要求 9 所述的用于分布式数据处理系统中启动备份任务的方法，其特征在于，所述按照所述任務的处理状态和所述预设的处理速度获取的预估处理时间，包括：

10 获取所述任務未处理时的原始数据量；

获取所述任務在处理該任務的节点上的启动时间；所述启动时间至少包括所述任務申请资源和启动处理該任務的节点的时间；

将所述任務未处理时的原始数据量和所述预设的处理速度的商，作为所述任務的备份任务的初步预估处理时间；

15 将所述任務在处理該任務的节点上的启动时间和所述任務的备份任务的初步预估处理时间的和，作为所述按照所述任務的处理状态和所述预设的处理速度获取的预估处理时间。

13、根据权利要求 12 所述的用于分布式数据处理系统中启动备份任务的方法，其特征在于，所述按照所述任務的处理状态和所述预设的处理速度获取的预估处理时间的算法为：

$$\text{BackupTime}_N = \text{AllData}_N / \text{AvgRate} + \text{overheadTime}$$

其中，BackupTime 代表单节点上处理的任务的预估处理时间，其数值越大说明处理該任务的备份任务时处理时间就越长，AllData 代表单节点上处理的任務未处理时的原始数据量，AvgRate 代表预设的处理速度，字母 N 代表某一具体单节点上处理的任務，overheadTime 代表启动单节点上处理的任務所需要的时间，包括申请资源和节点启动等。

14、根据权利要求 1 所述的用于分布式数据处理系统中启动备份任务的方法，其特征在于，所述为所述任務启动备份任务，包括：

为所述任務，创建該任务的备份任务；

30 为所述备份任务分配节点；

在所述节点中处理该备份任务。

15、一种用于分布式数据处理系统中启动备份任务的装置，其特征在于，包括：

处理速度获取单元，用于根据在单节点上处理的任务的处理状态获取所述任务的
处理速度；

5 处理速度判断单元，用于判断所述任务的处理速度是否小于预设的处理速度；

预估剩余处理时间获取单元，用于接收所述处理速度判断单元的判断结果，若是，
则根据所述任务的处理状态和所述任务的处理速度获取所述任务的预估剩余处理时间；

处理时间判断单元，用于判断所述任务的预估剩余处理时间是否大于按照所述任务
的处理状态和所述预设的处理速度获取的预估处理时间；

10 备份任务启动单元，用于接收所述处理时间判断单元的判断结果，若是，为所述任
务启动备份任务。

16、根据权利要求 15 所述的用于分布式数据处理系统中启动备份任务的装置，其
特征在于，所述处理速度获取单元，具体用于按照预设的时间间隔，根据在单节点上处
理的任务的处理状态获取所述任务的处理速度。

15 17、根据权利要求 15 所述的用于分布式数据处理系统中启动备份任务的装置，其
特征在于，所述处理速度获取单元，具体用于按照预设的时间间隔，根据在单节点上处
理的任务的已处理的数据量以及已处理的时间，获取所述任务的处理速度。

18、根据权利要求 17 所述的用于分布式数据处理系统中启动备份任务的装置，其
特征在于，所述处理速度获取单元，包括：

20 数据量时间获取子单元，用于获取当前在单节点上处理的任务的已处理的数据量以
及已处理的时间；

处理速度计算子单元，用于将所述任务的已处理的数据量和所述任务的已处理的时
间的商，作为该任务的处理速度。

25 19、根据权利要求 17 所述的用于分布式数据处理系统中启动备份任务的装置，其
特征在于，所述处理速度判断单元，具体用于判断所述任务的处理速度是否小于当前在
所述处理系统中所有节点上处理的任务的平均处理速度。

20、根据权利要求 19 所述的用于分布式数据处理系统中启动备份任务的装置，其
特征在于，所述处理速度判断单元，包括：

30 节点数据量计算子单元，用于计算当前在所述处理系统中所有节点上处理的任务已
处理的数据量的和，将所述当前在所述处理系统中所有节点上处理的任务已处理的数据

量的和作为所述处理系统中总任务已处理的数据量；

节点时间计算子单元，用于计算当前在所述处理系统中所有节点上处理的任务已处理的时间的和，将所述当前在所述处理系统中所有节点上处理的任务已处理的时间的和作为所述处理系统中总任务已处理的时间；

- 5 处理速度计算子单元，用于将所述总任务已处理的数据量和所述总任务已处理的时间的商，作为所述预设的处理速度。

21、根据权利要求 17 所述的用于分布式数据处理系统中启动备份任务的装置，其特征在于，所述处理速度获取单元，具体用于按照预设的时间间隔，根据在单节点上处理的任务的已处理的数据量、已处理的时间以及所述任务未处理时的原始数据量，获取
10 所述任务的处理速度。

22、根据权利要求 21 所述的用于分布式数据处理系统中启动备份任务的装置，其特征在于，所述预估剩余处理时间获取单元，包括：

处理状态获取子单元，用于获取所述任务未处理时的原始数据量、所述任务的已处理的数据量以及所述任务的处理速度；

- 15 剩余数据量计算子单元，用于将所述任务未处理时的原始数据量和所述任务的已处理的数据量的差，作为所述任务的剩余数据量；

剩余处理时间计算子单元，用于将所述任务的剩余数据量和所述任务的处理速度的商，作为所述任务的预估剩余处理时间。

23、根据权利要求 21 所述的用于分布式数据处理系统中启动备份任务的装置，其
20 特征在于，所述处理时间判断单元，包括：

原始数据量获取子单元，用于获取所述任务未处理时的原始数据量；

启动时间获取子单元，用于获取所述任务在处理该任务的节点上的启动时间；所述启动时间至少包括所述任务申请资源和启动处理该任务的节点的时间；

- 25 初步预估子单元，用于将所述任务未处理时的原始数据量和所述预设的处理速度的商，作为所述任务的备份任务的初步预估处理时间；

处理时间计算子单元，用于将所述任务在处理该任务的节点上的启动时间和所述任务的备份任务的初步预估处理时间的和，作为所述按照所述任务的处理状态和所述预设的处理速度获取的预估处理时间。

24、根据权利要求 15 所述的用于分布式数据处理系统中启动备份任务的装置，其
30 特征在于，所述备份任务启动单元，包括：

备份任务创建子单元，用于为所述任务，创建该任务的备份任务；

节点分配子单元，用于为所述备份任务分配节点；

备份任务处理子单元，用于在所述节点中处理该备份任务。

25、一种电子设备，其特征在于，所述电子设备包括：

5 显示器；

处理器；

存储器，用于存储备份任务启动程序，所述程序在被所述处理器读取执行时，执行如下操作：根据在单节点上处理的任务的处理状态获取所述任务的处理速度，判断所述任务的
10 处理速度是否小于预设的处理速度，若小于，则根据所述任务的处理状态和所述任务的
处理速度获取所述任务的预估剩余处理时间，判断所述任务的预估剩余处理时间是否
大于按照所述任务的处理状态和所述预设的处理速度获取的预估处理时间，若大于，为
所述任务启动备份任务。

15

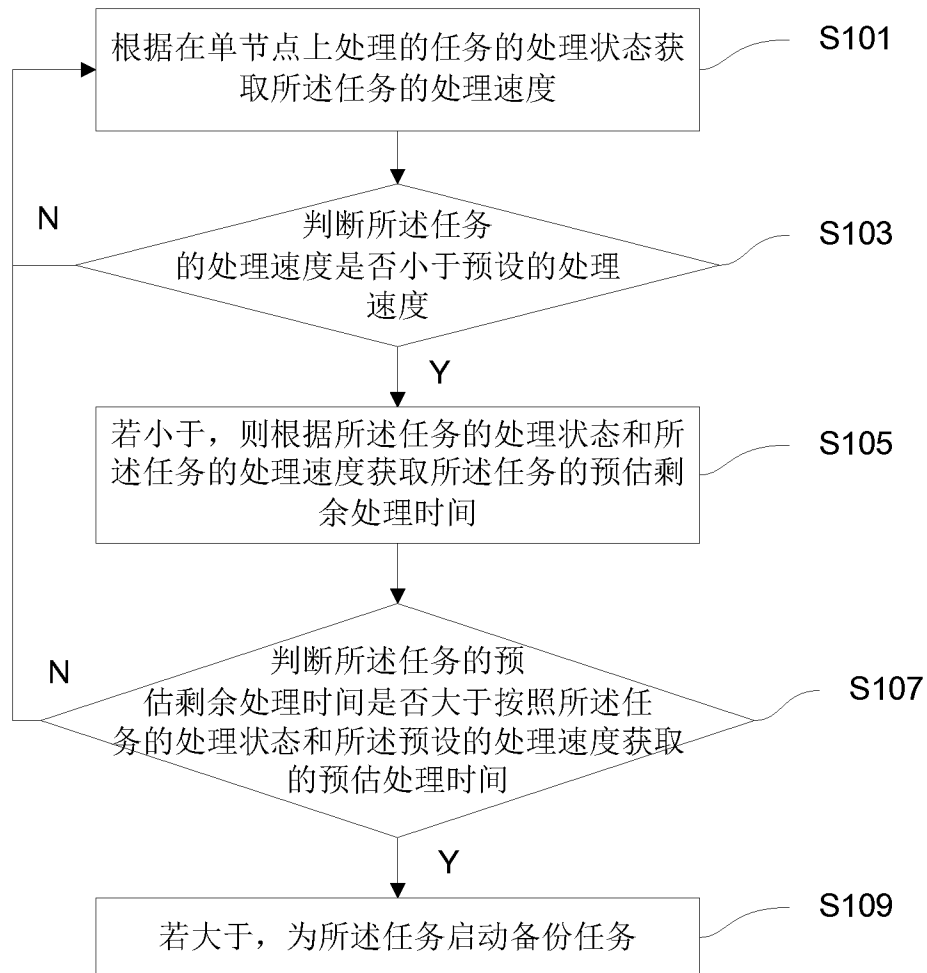


图 1

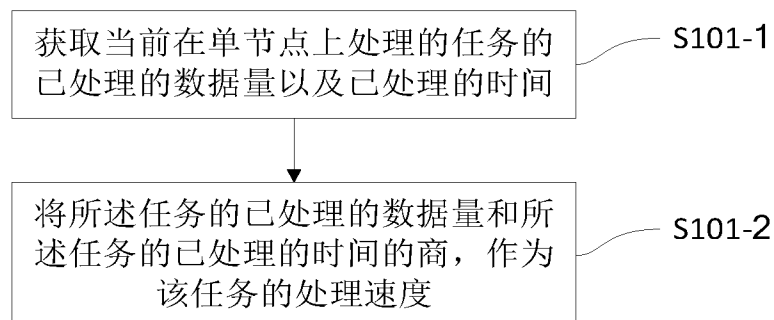


图 2

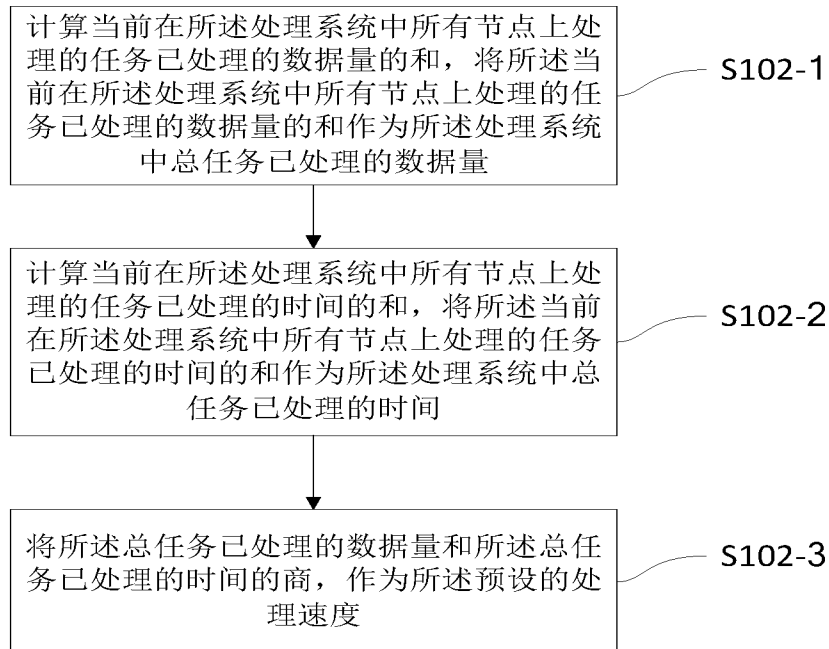


图 3

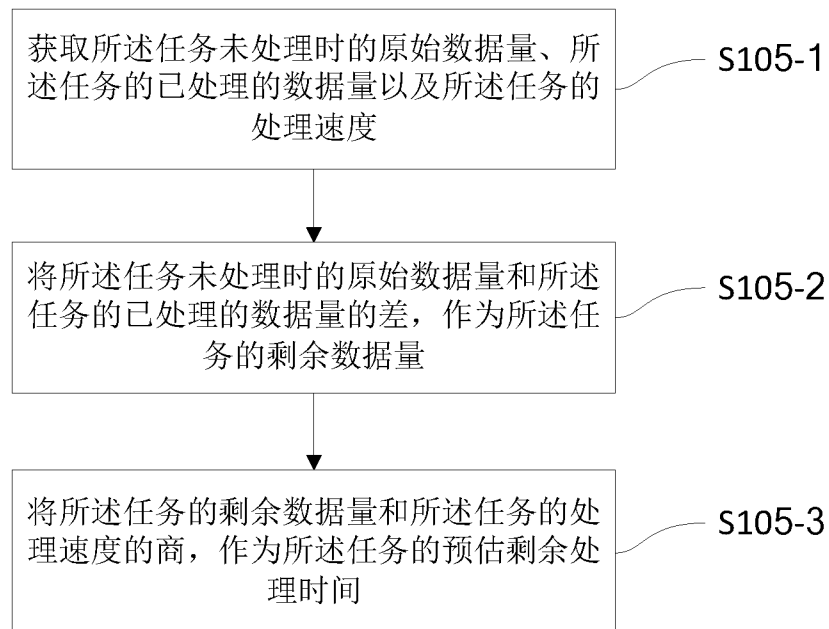


图 4

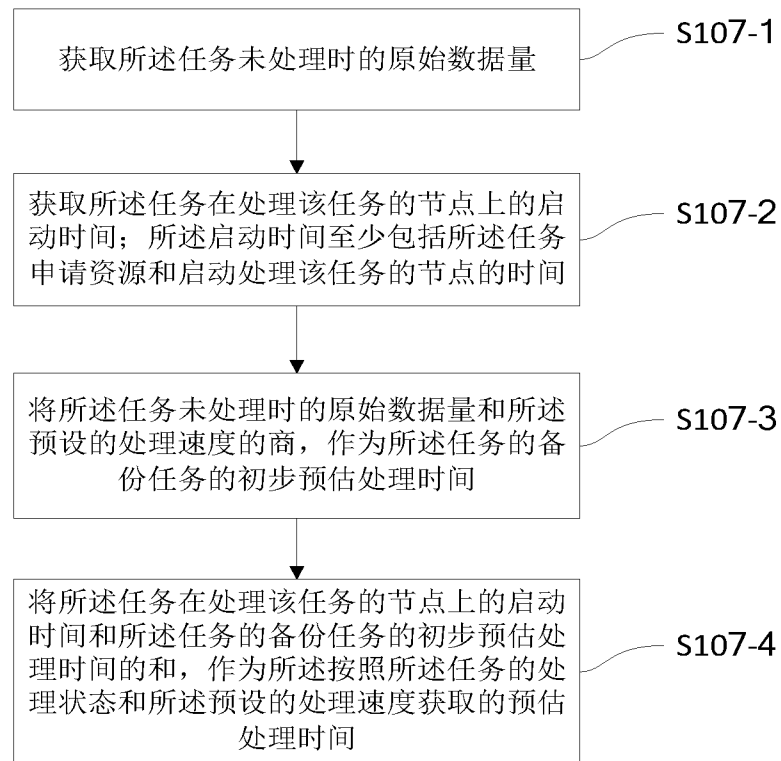


图 5

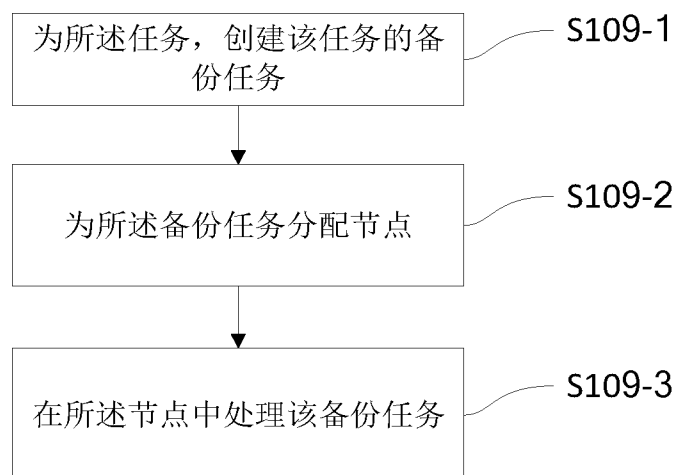


图 6

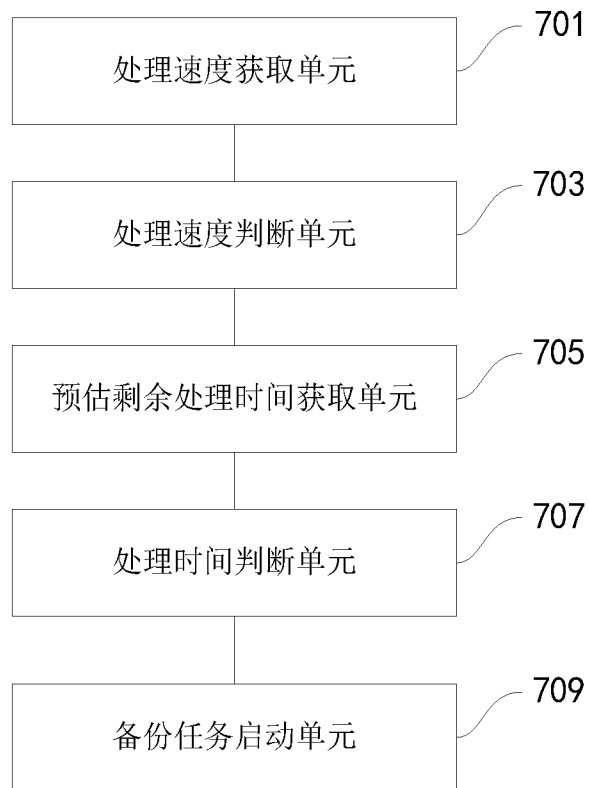


图 7

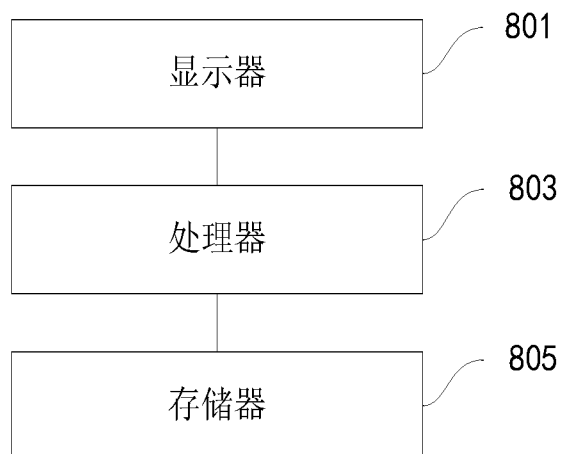


图 8

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2016/098009**A. CLASSIFICATION OF SUBJECT MATTER**

G06F 9/445 (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)

CNKI, CNPAT, WPI, EPODOC: distributed, remaining time, slow task, laggard task, long-tail delay, mapreduce, late, hadoop, samr, task, distribute, start, backup, speed, time, slow, delay

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	CN 104199739 A (INSPUR (BEIJING) ELECTRONIC INFORMATION INDUSTRY CO., LTD.), 10 December 2014 (10.12.2014), description, paragraph [0029], and claims 1-2	1-25
A	CN 102609303 A (HUAWEI TECHNOLOGIES CO., LTD.), 25 July 2012 (25.07.2012), the whole document	1-25
A	CN 1C), 3473345 A (SHENZHEN TEMOBI SCIENCE & TECH DEVELOPMENT CO., LTD.), 25 December 2013 (25.12.2013), the whole document	1-25
A	CN 103440167 A (FUZHOU UNIVERSITY), 11 December 2013 (11.12.2013), the whole document	1-25

☐ 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 13 October 2016 (13.10.2016)	Date of mailing of the international search report 28 October 2016 (28.10.2016)
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 ZOU, Yuting Telephone No.: (86-10) 62414043

International application No.

Patent Documents referred in the Report	Publication Date	Patent Family	Publication Date
CN 104199739 A	10 December 2014	None	
CN 102609303 A	25 July 2012	None	
CN 103473345 A	25 December 2013	None	
CN 103440167 A	11 December 2013	None	

国际检索报告

国际申请号

PCT/CN2016/098009

A. 主题的分类

G06F 9/445(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))

CNKI, CNPAT, WPI, EPODOC: 分布式, 启动, 备份, 任务, 速度, 剩余时间, 慢任务, 落后任务, 长尾延迟, mapreduce, late, hadoop, samr, task, distribute, start, backup, speed, time, slow, delay

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
A	CN 104199739 A (浪潮北京电子信息产业有限公司) 2014年 12月 10日 (2014-12-10) 说明书第[0029]段, 权利要求1-2	1-25
A	CN 102609303 A (华为技术有限公司) 2012年 7月 25日 (2012-07-25) 全文	1-25
A	CN 103473345 A (深圳市融创天下科技股份有限公司) 2013年 12月 25日 (2013-12-25) 全文	1-25
A	CN 103440167 A (福州大学) 2013年 12月 11日 (2013-12-11) 全文	1-25

☐ 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 10月 13日

国际检索报告邮寄日期

2016年 10月 28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

传真号 (86-10)62019451

受权官员

邹予婷

电话号码 (86-10)62414043

表 PCT/ISA/210 (第2页) (2009年7月)

国际申请号	PCT/CN2016/098009
-------	-------------------

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	104199739	A	2014年 12月 10日	无	
CN	102609303	A	2012年 7月 25日	无	
CN	103473345	A	2013年 12月 25日	无	
CN	103440167	A	2013年 12月 11日	无	