



(12) 发明专利

(10) 授权公告号 CN 118012917 B

(45) 授权公告日 2024.06.11

(21) 申请号 202410417403.X

(22) 申请日 2024.04.09

(65) 同一申请的已公布的文献号  
申请公布号 CN 118012917 A

(43) 申请公布日 2024.05.10

(73) 专利权人 无码科技(杭州)有限公司  
地址 310000 浙江省杭州市萧山区明星路  
371号1幢1208室

(72) 发明人 袁佳奇 冯大辉

(74) 专利代理机构 杭州裕阳联合专利代理有限  
公司 33289  
专利代理师 田金霞

(51) Int. Cl.  
G06F 16/2455 (2019.01)  
G06F 16/27 (2019.01)

(56) 对比文件

CN 102521232 A, 2012.06.27

CN 108268319 A, 2018.07.10

CN 115827280 A, 2023.03.21

CN 116050179 A, 2023.05.02

US 2019114350 A1, 2019.04.18

审查员 赵会玲

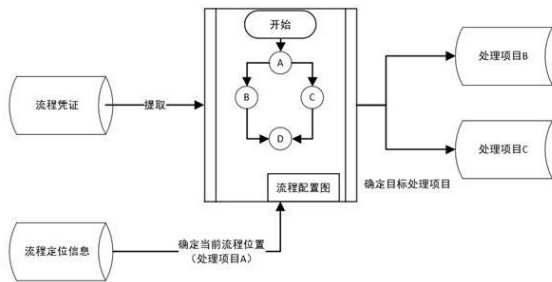
权利要求书3页 说明书10页 附图3页

(54) 发明名称

数据流处理方法、调度控制器、分布式数据流处理系统

(57) 摘要

本发明公开一种数据流处理方法、调度控制器、分布式数据流处理系统,涉及数据处理领域,其中数据流处理方法包括基于调度消息确定目标处理项目的步骤,所述调度消息包括流程凭证和流程定位信息,具体步骤为:基于所述流程凭证提取相应的流程配置图,所述流程配置图为有向无环图,用于指示对应数据流处理任务的处理流程,所述流程配置图的每个节点为一处理项目,节点间的关系为数据流处理的顺序;基于所述流程定位信息于所述流程配置图中确定当前处理位置,将当前处理位置的下一级节点所对应的处理项目作为目标处理项目。本发明在易维护、易扩展的前提下,可适用于数据处理流程不统一的复杂业务场景。



1. 一种数据流处理方法,其特征在于包括以下步骤:

获取调度消息;

基于所述调度消息确定目标处理项目,生成与所述目标处理项目一一对应的执行消息;

基于处理项目,将所述执行消息发送至对应的处理器,所述处理器与所述处理项目一一对应;

其中,所述调度消息包括流程凭证和流程定位信息,基于所述调度消息确定目标处理项目的具体步骤为:

基于所述流程凭证提取相应的流程配置图,所述流程配置图为有向无环图,用于指示对应数据流处理任务的处理流程,所述流程配置图的每个节点为一处理项目,节点间的关系为数据流处理的顺序;

基于所述流程定位信息于所述流程配置图中确定当前处理位置,将当前处理位置的下一级节点所对应的处理项目作为目标处理项目。

2. 根据权利要求1所述的一种数据流处理方法,其特征在于:

所述流程定位信息包括处理类型和处理进度,所述处理类型包括入口类型和回调类型;

调度消息的类型包括任务发起消息和处理反馈消息;

所述任务发起消息为发起数据流处理任务时所生成的消息,所述任务发起消息的处理类型为入口类型,处理进度为空;

所述处理反馈消息为处理器基于执行消息完成相应的处理项目时所生成的消息,所述处理反馈消息的处理类型为回调类型,处理进度用于指示当前完成的处理项目。

3. 根据权利要求2所述的一种数据流处理方法,其特征在于:

属于所述处理反馈消息的调度消息还包括至少一个结果参数;

流程配置图的节点中包含对应处理项目的业务参数,以及与所述业务参数相对应的参数配置信息,所述参数配置信息包括对应业务参数的名称以及名称映射表;

当调度消息的处理类型为回调类型时,为目标处理项目生成相对应的执行消息的具体步骤为:

获取所述调度消息中的结果参数;

从所述流程配置图中提取所述目标处理项目所对应的参数配置信息;

当所述结果参数与相应的业务参数的名称相同时,将所述结果参数赋值于所述业务参数;

当缺少与所述结果参数与名称相同的业务参数时,将所述结果参数基于所述名称映射表将所述结果参数映射至相应的业务参数;

基于所得业务参数生成相应的执行消息。

4. 根据权利要求1至3任一所述的一种数据流处理方法,其特征在于:

流程配置图的节点中包含对应处理项目的并发控制配置,所述并发控制配置包括信号key和动态参数key,所述信号key用于指示对应处理项目并发控制的基础服务,所述动态参数key用于指定进行并发控制的业务参数;

执行消息包含与目标处理项目相对应的并发控制参数;

获取所述并发控制参数的具体步骤为：

从所述流程配置图中提取所述目标处理项目所对应的并发控制配置；

当所述并发控制配置中指定所述动态参数key时，基于所述动态参数key获取相应的业务参数作为所述并发控制参数；

当所述并发控制配置中未指定所述动态参数key时，提取所述信号key作为所述并发控制参数。

5. 根据权利要求4所述的一种数据流处理方法，其特征在于，基于处理项目，将所述执行消息发送至对应的处理器前还包括并发判断步骤，具体步骤为：

基于所述执行消息中的并发控制参数进行并发判断；

当并发判断结果为执行，则基于处理项目将所述执行消息发送至对应的处理器中，并基于所述并发控制参数进行上锁。

6. 根据权利要求5所述的一种数据流处理方法，其特征在于，基于所述执行消息中的并发控制参数进行并发判断的具体步骤为：

基于所述并发控制参数进行本地锁对象查询；

当未查询到所述本地锁对象时，基于所述并发控制参数进行分布式锁对象查询；

当未查询到所述分布式锁对象时，判定并发判断结果为执行，基于处理项目将所述执行消息发送至对应的处理器中，并基于所述并发控制参数生成相对应的本地锁对象及分布式锁对象。

7. 根据权利要求1至3任一所述的一种数据流处理方法，其特征在于：

当基于调度消息生成相应的执行消息失败时，将所述调度消息作为调度重试消息；

基于预设的时间间隔获取并处理所述调度重试消息。

8. 一种调度控制器，其特征在于，包括：

获取模块，用于获取调度消息，所述调度消息包括流程凭证和流程定位信息；

分析处理模块，基于所述调度消息确定目标处理项目，生成与所述目标处理项目一一对应的执行消息；

发送模块，用于基于处理项目向相应的处理器发送所述执行消息，所述处理器与所述处理项目一一对应；

所述分析处理模块，包括：

提取单元，用于基于所述流程凭证提取相应的流程配置图，所述流程配置图为有向无环图，用于指示对应数据流处理任务的处理流程，所述流程配置图的每个节点为一处理项目，节点间的关系为数据流处理的顺序；

目标获取单元，用于基于所述流程定位信息于所述流程配置图中确定当前处理位置，将当前处理位置的下一级节点所对应的处理项目作为目标处理项目。

9. 根据权利要求8所述的一种调度控制器，其特征在于：

所述流程配置图的节点中包含对应处理项目的并发控制配置，所述并发控制配置包括信号key和动态参数key，所述信号key用于指示对应处理项目并发控制的基础服务，所述动态参数key用于指定进行并发控制的业务参数；

所述执行消息包含与目标处理项目相对应的并发控制参数；

所述分析处理模块还包括并发控制配置单元，所述并发控制配置单元：

用于从所述流程配置图中提取所述目标处理项目所对应的并发控制配置；

还用于当所述并发控制配置中指定所述动态参数key时,基于所述动态参数key获取相应的业务参数作为所述并发控制参数；

还用于当所述并发控制配置中未指定所述动态参数key时,提取所述信号key作为所述并发控制参数。

10.一种分布式数据流处理系统,其特征在于,包括:

调度队列,用于存放调度消息,所述调度消息包括流程凭证和流程定位信息;

若干执行队列,所述执行队列与处理项目一一对应,用于存放与所述处理项目相对应的执行消息;

若干个处理器,所述处理器与处理项目一一对应,所述处理器用于处理相应的执行消息;

调度控制器,用于从所述调度队列中获取调度消息,基于所述调度消息确定目标处理项目,生成与所述目标处理项目一一对应的执行消息,并基于处理项目向对应的执行队列发送所述执行消息;所述调度控制器为权利要求8或9任意一项所述的调度控制器;

执行控制器,用于从各执行队列中获取执行消息,并对所述执行消息进行并发判断,并基于并发判断结果将所述执行消息分发至相应的处理器。

## 数据流处理方法、调度控制器、分布式数据流处理系统

### 技术领域

[0001] 本发明涉及数据处理领域,尤其涉及一种数据流处理方法、调度控制器、分布式数据流处理系统。

### 背景技术

[0002] 在实际应用中,针对不同的数据处理业务场景,各类待处理数据的数据类型不尽相同,且进行数据处理的流程也存在区别;

[0003] 现今往往以顺序编码的方式进行数据处理,以获得需要的数据,即,针对各业务场景所对应的数据处理逻辑进行过程式编码,此方案会导致重复编码,存在维护难,扩展难的缺陷。

[0004] 针对上述问题:

[0005] 公开号为CN113065029A的专利文件中公开一种将不同业务场景的内容源数据流处理流程进行单元化、组件化得到统一处理逻辑框架,利用统一处理逻辑框架处理不同业务场景的内容源数据流的数据流处理方案;但是统一处理逻辑框架为固定流程框架,该方案仅适用于数据源不同,但数据处理流程相同的场景;

[0006] 公开号为CN111597058A的专利文件中公开一种由各个处理节点根据自身的拓扑参数自发组建拓扑架构,按照所述拓扑架构进行数据流处理的方案;该拓扑架构亦为固定的流程框架,故该方案亦仅适用于单一流程的数据流处理场景。

### 发明内容

[0007] 本发明针对现有技术仅适用于单一数据处理流程的缺点,提供了一种数据流处理方法及系统、还提供了一种调度控制器。

[0008] 为了解决上述技术问题,本发明通过下述技术方案得以解决:

[0009] 第一方面,提供一种数据流处理方法,包括以下步骤:

[0010] 获取调度消息;

[0011] 基于所述调度消息确定目标处理项目,生成与所述目标处理项目一一对应的执行消息;

[0012] 基于处理项目,将所述执行消息发送至对应的处理器,所述处理器与所述处理项目一一对应;

[0013] 其中,所述调度消息包括流程凭证和流程定位信息,基于所述调度消息确定目标处理项目的具体步骤为:

[0014] 基于所述流程凭证提取相应的流程配置图,所述流程配置图为有向无环图,用于指示对应数据流处理任务的处理流程,所述流程配置图的每个节点为一处理项目,节点间的关系为数据流处理的顺序;

[0015] 基于所述流程定位信息于所述流程配置图中确定当前处理位置,将当前处理位置的下一级节点所对应的处理项目作为目标处理项目。

- [0016] 第二方面,提供一种调度控制器,包括:
- [0017] 获取模块,用于获取调度消息,所述调度消息包括流程凭证和流程定位信息;
- [0018] 分析处理模块,基于所述调度消息确定目标处理项目,生成与所述目标处理项目一一对应的执行消息;
- [0019] 发送模块,用于基于处理项目向相应的处理器发送所述执行消息,所述处理器与所述处理项目一一对应;
- [0020] 所述分析处理模块,包括:
- [0021] 提取单元,用于基于所述流程凭证提取相应的流程配置图,所述流程配置图为有向无环图,用于指示对应数据流处理任务的处理流程,所述流程配置图的每个节点为一处理项目,节点间的关系为数据流处理的顺序;
- [0022] 目标获取单元,用于基于所述流程定位信息于所述流程配置图中确定当前处理位置,将当前处理位置的下一级节点所对应的处理项目作为目标处理项目。
- [0023] 第三方面,提供一种分布式数据流处理系统,包括:
- [0024] 调度队列,用于存放调度消息,所述调度消息包括流程凭证和流程定位信息;
- [0025] 若干执行队列,所述执行队列与处理项目一一对应,用于存放与所述处理项目相对应的执行消息;
- [0026] 若干个处理器,所述处理器与处理项目一一对应,所述处理器用于处理相应的执行消息;
- [0027] 调度控制器,用于从所述调度队列中获取调度消息,基于所述调度消息确定目标处理项目,生成与所述目标处理项目一一对应的执行消息,并基于处理项目向对应的执行队列发送所述执行消息;
- [0028] 执行控制器,用于从各执行队列中获取执行消息,并对所述执行消息进行并发判断,并基于并发判断结果将所述执行消息分发至相应的处理器。
- [0029] 本发明公开的实施例提供的技术方案至少带来以下有益效果:
- [0030] 通过对流程配置图和流程定位信息的设计,无需中心化即可确定下一步数据流向,且处理器可复用于不同数据流程的数据流处理任务,在易维护、易扩展的前提下,适用于数据处理流程不统一的复杂业务场景。

## 附图说明

[0031] 为了更清楚地说明本发明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发明的一些实施例,对于本领域普通技术人员来讲,在不付出创造性劳动的前提下,还可以根据这些附图获得其他的附图。

- [0032] 图1是实施例1中基于所述调度消息确定目标处理项目的流程示意图;
- [0033] 图2是本发明一种调度控制器的模块连接示意图;
- [0034] 图3是图2中分析处理模块的模块连接示意图;
- [0035] 图4是本发明一种分布式数据流处理系统的模块连接示意图。

## 具体实施方式

[0036] 下面结合实施例对本发明做进一步的详细说明,以下实施例是对本发明的解释而本发明并不局限于以下实施例。

[0037] 实施例1、一种数据流处理方法,包括调度控制方法。

[0038] 所述调度控制方法包括以下步骤:

[0039] S100. 获取调度消息。

[0040] 本实施例中调度消息包括流程凭证和流程定位信息。

[0041] 流程凭证:

[0042] 流程凭证为流程配置图的唯一性识别码,所述流程配置图用于指示对应数据流处理任务的处理流程。

[0043] 本实施例中流程配置图为有向无环图,所述流程配置图的每个节点为一处理项目,节点间的关系为数据流处理的顺序,即,每个节点与一处理器对应,节点间的连线指示处理器的执行顺序;

[0044] 所述处理项目为可独立完成的一特定数据处理步骤,本领域技术人员可根据实际需要进行设定,例如数据调用、数据通知均可作为一处理项目,背景技术中指出的文件中所述的统一处理逻辑框架与拓扑架构亦可作为一处理项目,本说明书不对其进行详细限定。

[0045] 在实际应用中,工作人员可根据实际需要配置流程配置图,并将该流程配置图的唯一性识别码作为流程凭证,在数据流处理任务发起时,生成包含该流程凭证的调度消息。

[0046] 流程定位信息:

[0047] 流程定位信息用于在流程配置图中指示当前处理位置。

[0048] 所述流程定位信息包括处理类型和处理进度,所述处理类型包括入口类型和回调类型,入口类型即指示当前处理位置为流程入口,回调类型即指示当前处理位置为具体的处理进度;

[0049] 调度消息的类型分为任务发起消息和处理反馈消息;

[0050] 所述任务发起消息为发起数据流处理任务时所生成的消息,所述任务发起消息的处理类型为入口类型,处理进度为空;

[0051] 所述处理反馈消息为处理器基于执行消息完成相应的处理项目时所生成的消息,所述处理反馈消息的处理类型为回调类型,处理进度用于指示当前完成的处理项目。

[0052] 本领域技术人员可根据实际情况自行设定流程定位信息,如可仅将处理进度作为流程定位信息,本说明书不对其做详细限定。

[0053] 本实施例从消息队列中拉取调度消息。

[0054] S200. 基于所述调度消息确定目标处理项目,生成与所述目标处理项目一一对应的执行消息。

[0055] 当未获得目标处理项目时,说明对应数据流处理任务已完成,无须再处理。

[0056] S300. 基于处理项目,将所述执行消息发送至对应的处理器,所述处理器与所述处理项目一一对应。

[0057] 本实施例中设有与所述处理项目一一对应的消息队列,根据处理项目将执行消息发送至相对应的消息队列。

[0058] 参照图1,步骤S200的具体步骤包括:

- [0059] S210.基于所述流程凭证提取相应的流程配置图。
- [0060] S220.基于所述流程定位信息于所述流程配置图中确定当前处理位置,将当前处理位置的下一级节点所对应的处理项目作为目标处理项目。
- [0061] 下一级节点为直属于当前处理位置的节点。
- [0062] 如图1所示,当前处理位置为处理项目A时,从流程配置图中获取所对应的节点A的位置,以及直属于节点A的节点B和节点C,从而将节点B对应的处理项目B和节点C对应的处理项目C作为目标处理节点。
- [0063] S230.依次为各目标处理项目生成相对应的执行消息。
- [0064] 对于单一数据处理流程的场景,数据流仅需沿着所构建的处理架构进行流转即可,但对于多数据处理流程的处理场景中,往往需要中心化的节点对各处理器进行调度,但中心化的设计与易扩展的目的相悖;
- [0065] 本实施例通过对流程配置图和流程定位信息的设计,无需中心化即可确定下一步数据流向,且处理器可复用于不同数据流程的数据流处理任务,在易维护、易扩展的前提下,适用于数据处理流程不统一的复杂业务场景。
- [0066] 每个处理项目都有相应的业务参数和/或结果参数,业务参数即该处理项目所对应的待处理参数,结果参数即完成该处理项目获得的结果;
- [0067] 进一步地:
- [0068] 属于所述处理反馈消息的调度消息还包括至少一个结果参数;
- [0069] 流程配置图的节点中包含对应处理项目的业务参数,以及与所述业务参数相对应的参数配置信息,所述参数配置信息包括对应业务参数的名称以及名称映射表;
- [0070] 当调度消息的处理类型为回调类型时,步骤S230中为目标处理项目生成相对应的执行消息的具体步骤为:
- [0071] 获取所述调度消息中的结果参数;
- [0072] 从所述流程配置图中提取所述目标处理项目所对应的参数配置信息;
- [0073] 当所述结果参数与相应的业务参数的名称相同时,将所述结果参数赋值于所述业务参数;
- [0074] 当缺少与所述结果参数与名称相同的业务参数时,将所述结果参数基于所述名称映射表将所述结果参数映射至相应的业务参数;
- [0075] 基于所得业务参数生成相应的执行消息。
- [0076] 本实施例中参数配置信息采用JSON 格式,本领域技术人员可根据实际需要自行确定参数配置信息的格式,如XML格式。
- [0077] 在实际应用中,如果下一处理项目的参数配置信息为空,则直接将当前处理项目的结果参数作为其业务参数。
- [0078] 参数配置信息是在流程配置图中特定的处理器下设置的参数映射关系,本实施例通过对参数配置信息的设计,将前序处理项目的结果参数转换为下一处理项目的业务参数,从而消除处理器间的耦合关系。
- [0079] 进一步地:
- [0080] 流程配置图的节点中还包含对应处理项目的状态属性;
- [0081] 所述状态属性用于指示对应处理项目启动或跳过;



- [0082] 当目标处理项目的状态属性为启动时,为所述目标处理项目生成执行消息;
- [0083] 当目标处理项目的状态属性为跳过时,将目标处理项目视为已完成,基于流程配置图,将目标处理项目的下一级节点所对应的处理项目作为新的目标处理项目。
- [0084] 在实际使用中,工作人员可在已有流程配置图的基础上进行调整,以获得所需的流程配置图,即,工作人员可根据实际需要,对现有的流程配置图的状态属性进行设置,以快速生成满足对应数据流处理任务需要的流程配置图。
- [0085] 例如:参照图1中的流程配置图,数据处理流程依次为{A}、{B,C}、{D},此时数据流向分为(A、B、D)与(A、C、D)两个分支,此处D是对处理器的复用;
- [0086] 当处理项目A完成后生成相应的调度消息;
- [0087] 此时基于该调度消息确定目标处理项目为处理项目B和处理项目C;
- [0088] 当处理项目B的启动属性为跳过,处理项目C的启动属性为启动时,将目标处理项目由处理项目B和处理项目C更新为处理项目D和处理项目C;
- [0089] 此时对于(A、B、D)的数据流向分支,继续判断处理项目D的启动属性,在处理项目D的启动属性为启动时,将处理项目A的结果参数映射获得处理项目D的业务参数,生成处理项目D所对应的执行消息;
- [0090] 而对于(A、C、D)的数据流向分支,将处理项目A的结果参数映射获得处理项目C的业务参数,生成处理项目C所对应的执行消息,后续则基于处理项目C的结果参数映射获得处理项目D的业务参数。
- [0091] 进一步地:
- [0092] 当目标处理项目的状态属性为跳过时,将当前所有目标处理项目视为已完成,基于流程配置图,将各目标处理项目的下一级节点所对应的处理项目作为新的目标处理项目。
- [0093] 以上述数据处理流程依次为{A}、{B,C}、{D}为例,当发现处理项目B的启动属性为跳过时,此时直接将目标处理项目由处理项目B和处理项目C更新为处理项目D;
- [0094] 处理项目B和处理项目C的上一级节点均为处理项目A,故,将处理项目B和处理项目C所对应的节点视为同级节点,在实际应用中,同一数据流处理任务中同级节点所对应的启动属性都相同,为减少重复判断,故当一处理项目的启动属性为跳过时,即判定跳过与该处理项目同级的所有处理项目。
- [0095] 进一步地:
- [0096] 当基于调度消息生成相应的执行消息失败时,将所述调度消息作为调度重试消息;
- [0097] 基于预设的时间间隔获取并处理所述调度重试消息,调度重试消息的处理步骤同调度消息。
- [0098] 本实施例设有调度快速队列和调度慢速队列;
- [0099] 所述调度快速队列用于放置调度消息;
- [0100] 所述调度慢速队列用于放置调度重试消息;
- [0101] 在实际应用中实时从调度快速队列中消费调度消息,用于快速消费事件,保证高吞吐量;并按照预设的时间间隔从调度慢速队列中消费所述调度重试消息,用于稳定处理失败事件,保证数据不丢失。

[0102] 进一步地:

[0103] 当处理器处理相应的执行消息成功时,生成相应的处理反馈消息作为调度消息放入调度快速队列中,以快速进行下一处理项目;

[0104] 当处理器处理相应的执行消息失败时,判断其是否属于预设的可忽视的失败类型;

[0105] 如果属于预设的可忽视的失败类型,则将当前业务参数直接作为结果参数生成相应的处理反馈消息作为调度消息放入调度快速队列中;

[0106] 如果不属于预设的可忽视的失败类型,则将当前执行消息作为执行重试消息,基于预设的时间间隔获取并处理所述执行重试消息,执行重试消息的处理步骤同执行消息。

[0107] 本实施例设有若干执行队列,所述执行队列与处理项目一一对应;

[0108] 每组执行队列包括执行快速队列和执行慢速队列;

[0109] 所述执行快速队列用于放置执行消息;

[0110] 所述调度慢速队列用于放置执行重试消息;

[0111] 在实际应用中实时从执行快速队列中消费执行消息,保证高吞吐量;并按照预设的时间间隔从执行慢速队列中消费所述执行重试消息,用于稳定处理失败事件,保证数据不丢失。

[0112] 进一步地:

[0113] 流程配置图的节点中包含对应处理项目的并发控制配置,所述并发控制配置包括信号key和动态参数key,所述信号key用于指示对应处理项目并发控制的基础服务,所述动态参数key用于指定进行并发控制的业务参数;

[0114] 执行消息包含与目标处理项目相对应的并发控制参数。

[0115] 获取所述并发控制参数的具体步骤为:

[0116] 从所述流程配置图中提取所述目标处理项目所对应的并发控制配置;

[0117] 当所述并发控制配置中指定所述动态参数key时,基于所述动态参数key获取相应的业务参数作为所述并发控制参数;

[0118] 当所述并发控制配置中未指定所述动态参数key时,提取所述信号key作为所述并发控制参数。

[0119] 例如:

[0120] 处理项目A为新闻内容获取,信号key为「NEWS\_CONTENT」,动态参数key值为「参数中的 url」,url为统一资源定位符;

[0121] 处理项目B为基于大模型生成摘要,信号key为「BIG\_MODEL」,不设置动态参数key;

[0122] 处理项目C为基于大模型提取公司实体,其采用的大模型与处理项目B相同,其信号key为「BIG\_MODEL」,不设置动态参数key。

[0123] 处理项目A的并发控制参数为相应的url,实现了对相同数据源的并发控制。

[0124] 处理项目B和处理项目C所对应的并发控制参数为「BIG\_MODEL」,实现了对同一基础服务的并发控制。

[0125] 本实施例通过对并发控制参数的设计,能够实现对数据流处理任务的并发处理,对动态参数key的设计,能够对相同源数据进行等待执行,避免重复处理导致数据的异常覆

盖,解决脏读、幻读;通过对信号key的设计,能够对稀有资源进行全局控制,降低系统性崩溃风险,提高数据处理质量及成功率。

[0126] 实施例2、一种数据流处理方法,包括并发控制方法,于实施例1中步骤S300前执行。

[0127] 并发控制方法包括以下步骤:

[0128] S400.基于执行消息或执行重试消息中的并发控制参数进行并发判断;

[0129] S500.当并发判断结果为执行,则基于处理项目将所述执行消息发送至对应的处理器中,并基于所述并发控制参数进行上锁。

[0130] 具体为:

[0131] S510.基于所述并发控制参数进行本地锁对象查询;

[0132] 本地锁对象存储到本地运行时的环境中,故查询速度快,当查询到本地锁对象时直接令该执行消息等待执行。

[0133] S520.当未查询到所述本地锁对象时,基于所述并发控制参数进行分布式锁对象查询;

[0134] 分布式锁对象存储到分布式缓存服务中,分布式缓存服务例如可采用Redis。

[0135] S530.当未查询到所述分布式锁对象时,判定并发判断结果为执行,基于处理项目将所述执行消息发送至对应的处理器中,并基于所述并发控制参数生成相对应的本地锁对象及分布式锁对象。

[0136] 本实施例利用本地锁和分布式锁控制数据处理并发,控制对稀缺资源的竞争,基于此特性以支持分布式部署及横向扩展,即,一方面利用队列和消息机制,根据实际需要处理项目,及与处理项目相对应的执行队列和处理器进行增加,达到分布式扩展;另一方面,通过分布式锁,对处于分布式架构下的各个处理器进行并发控制。

[0137] 实施例3、一种调度控制器,如图2所示,包括:

[0138] 获取模块100,用于获取调度消息,所述调度消息包括流程凭证和流程定位信息;

[0139] 分析处理模块200,基于所述调度消息确定目标处理项目,生成与所述目标处理项目一一对应的执行消息;

[0140] 发送模块300,用于基于处理项目向相应的处理器发送所述执行消息,所述处理器与所述处理项目一一对应;

[0141] 参照图3,所述分析处理模块200,包括:

[0142] 提取单元210,用于基于所述流程凭证提取相应的流程配置图,所述流程配置图有为向无环图,用于指示对应数据流处理任务的处理流程,所述流程配置图的每个节点为一处理项目,节点间的关系为数据流处理的顺序;

[0143] 目标获取单元220,用于基于所述流程定位信息于所述流程配置图中确定当前处理位置,将当前处理位置的下一级节点所对应的处理项目作为目标处理项目;

[0144] 生成单元230,依次为各目标处理项目生成相对应的执行消息。

[0145] 进一步地:

[0146] 所述流程配置图的节点中包含对应处理项目的并发控制配置,所述并发控制配置包括信号key和动态参数key,所述信号key用于指示对应处理项目并发控制的基础服务,所述动态参数key用于指定进行并发控制的业务参数;

- [0147] 所述执行消息包含与目标处理项目相对应的并发控制参数；
- [0148] 参照图3,所述分析处理模块200还包括并发控制配置单元240,所述并发控制配置单元240:
- [0149] 用于从所述流程配置图中提取所述目标处理项目所对应的并发控制配置；
- [0150] 还用于当所述并发控制配置中指定所述动态参数key时,基于所述动态参数key获取相应的业务参数作为所述并发控制参数；
- [0151] 还用于当所述并发控制配置中未指定所述动态参数key时,提取所述信号key作为所述并发控制参数。
- [0152] 实施例4、一种分布式数据流处理系统,参照图4,包括:
- [0153] 调度队列10,用于存放调度消息,所述调度消息包括流程凭证和流程定位信息；
- [0154] 若干执行队列20,所述执行队列20与处理项目一一对应,用于存放与所述处理项目相对应的执行消息；
- [0155] 若干个处理器30,所述处理器30与处理项目一一对应,所述处理器30用于处理相应的执行消息；
- [0156] 调度控制器40,用于从所述调度队列10中获取调度消息,基于所述调度消息确定目标处理项目,生成与所述目标处理项目一一对应的执行消息,并基于处理项目向对应的执行队列20发送所述执行消息;所述调度控制器40为实施例3所述的调度控制器40；
- [0157] 执行控制器50,用于从各执行队列20中获取执行消息,并对所述执行消息进行并发判断,并基于并发判断结果将所述执行消息分发至相应的处理器30。
- [0158] 进一步地:
- [0159] 所述调度队列10包括调度快速队列11和调度慢速队列；
- [0160] 所述调度快速队列11用于放置调度消息；
- [0161] 所述调度慢速队列12用于放置调度重试消息。
- [0162] 所述调度控制器40:
- [0163] 用于实时从所述调度快速队列11中获取调度消息并处理；
- [0164] 还用于当处理调度消息或调度重试消息失败时,将其作为调度重试消息放入调度慢速队列12；
- [0165] 还用于基于预设的间隔时间从所述调度慢速队列12中获取调度重试消息并处理。
- [0166] 进一步地:
- [0167] 所述处理器30用于:
- [0168] 当处理相应的执行消息成功时,生成相应的处理反馈消息作为调度消息放入调度快速队列11中,以快速进行下一处理项目；
- [0169] 当处理相应的执行消息失败时,判断其是否属于预设的可忽视的失败类型；
- [0170] 如果属于预设的可忽视的失败类型,则将当前业务参数直接作为结果参数生成相应的处理反馈消息作为调度消息放入调度快速队列11中；
- [0171] 如果不属于预设的可忽视的失败类型,则将当前执行消息作为执行重试消息放入执行队列20中。
- [0172] 进一步地:
- [0173] 所述执行队列20包括执行快速队列21和执行慢速队列22；

- [0174] 所述执行快速队列21用于放置执行消息；
- [0175] 所述执行慢速队列22用于放置执行重试消息。
- [0176] 所述执行控制器50：
- [0177] 用于实时从所述执行快速队列21中获取执行消息并处理；
- [0178] 还用于当并发判断结果为等待时，将相应的执行消息或执行重试消息回退至对应消息队列；
- [0179] 还用于基于预设的间隔时间从所述执行慢速队列22中获取执行重试消息并处理。
- [0180] 对于装置实施例而言，由于其与方法实施例基本相似，所以描述的比较简单，相关之处参见方法实施例的部分说明即可。
- [0181] 本说明书中的各个实施例均采用递进的方式描述，每个实施例重点说明的都是与其他实施例的不同之处，各个实施例之间相同相似的部分互相参见即可。
- [0182] 本领域内的技术人员应明白，本发明的实施例可提供为方法、装置、或计算机程序产品。因此，本发明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，本发明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。
- [0183] 本发明是参照根据本发明的方法、终端设备（系统）、和计算机程序产品的流程图和/或方框图来描述的。应理解可由计算机程序指令实现流程图和/或方框图中的每一流程和/或方框、以及流程图和/或方框图中的流程和/或方框的结合。可提供这些计算机程序指令到通用计算机、专用计算机、嵌入式处理机或其他可编程数据处理终端设备的处理器30以产生一个机器，使得通过计算机或其他可编程数据处理终端设备的处理器30执行的指令产生用于实现在流程图一个流程或多个流程和/或方框图一个方框或多个方框中指定的功能的装置。
- [0184] 这些计算机程序指令也可存储在能引导计算机或其他可编程数据处理终端设备以特定方式工作的计算机可读存储器中，使得存储在该计算机可读存储器中的指令产生包括指令装置的制造品，该指令装置实现在流程图一个流程或多个流程和/或方框图一个方框或多个方框中指定的功能。
- [0185] 这些计算机程序指令也可装载到计算机或其他可编程数据处理终端设备上，使得在计算机或其他可编程终端设备上执行一系列操作步骤以产生计算机实现的处理，从而在计算机或其他可编程终端设备上执行的指令提供用于实现在流程图一个流程或多个流程和/或方框图一个方框或多个方框中指定的功能的步骤。
- [0186] 需要说明的是：
- [0187] 说明书中提到的“一个实施例”或“实施例”意指结合实施例描述的特定特征、结构或特性包括在本发明的至少一个实施例中。因此，说明书通篇各个地方出现的短语“一个实施例”或“实施例”并不一定均指同一个实施例。
- [0188] 尽管已描述了本发明的优选实施例，但本领域内的技术人员一旦得知了基本创造性概念，则可对这些实施例做出另外的变更和修改。所以，所附权利要求意欲解释为包括优选实施例以及落入本发明范围的所有变更和修改。
- [0189] 此外，需要说明的是，本说明书中所描述的具体实施例，其零、部件的形状、所取名

称等可以不同。凡依本发明专利构思所述的构造、特征及原理所做的等效或简单变化,均包括于本发明专利的保护范围内。本发明所属技术领域的技术人员可以对所描述的具体实施例做各种各样的修改或补充或采用类似的方式替代,只要不偏离本发明的结构或者超越本权利要求书所定义的范围,均应属于本发明的保护范围。

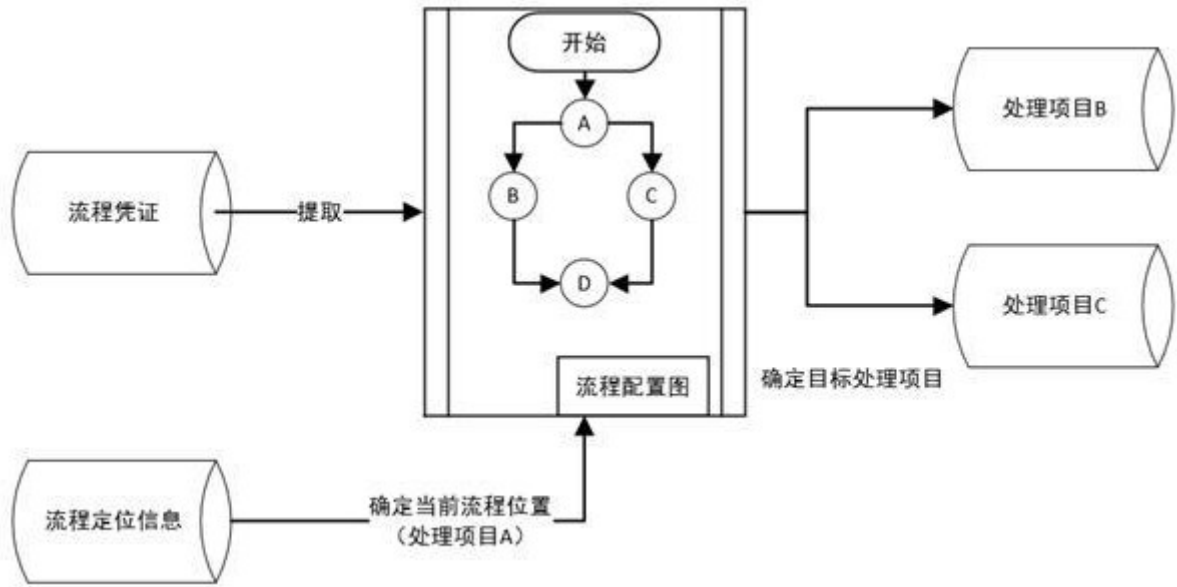


图 1

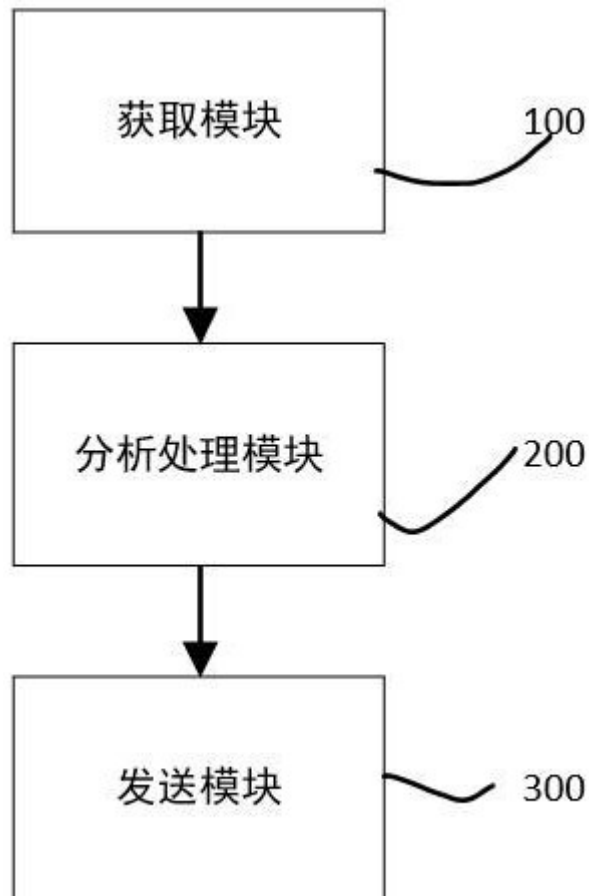


图 2

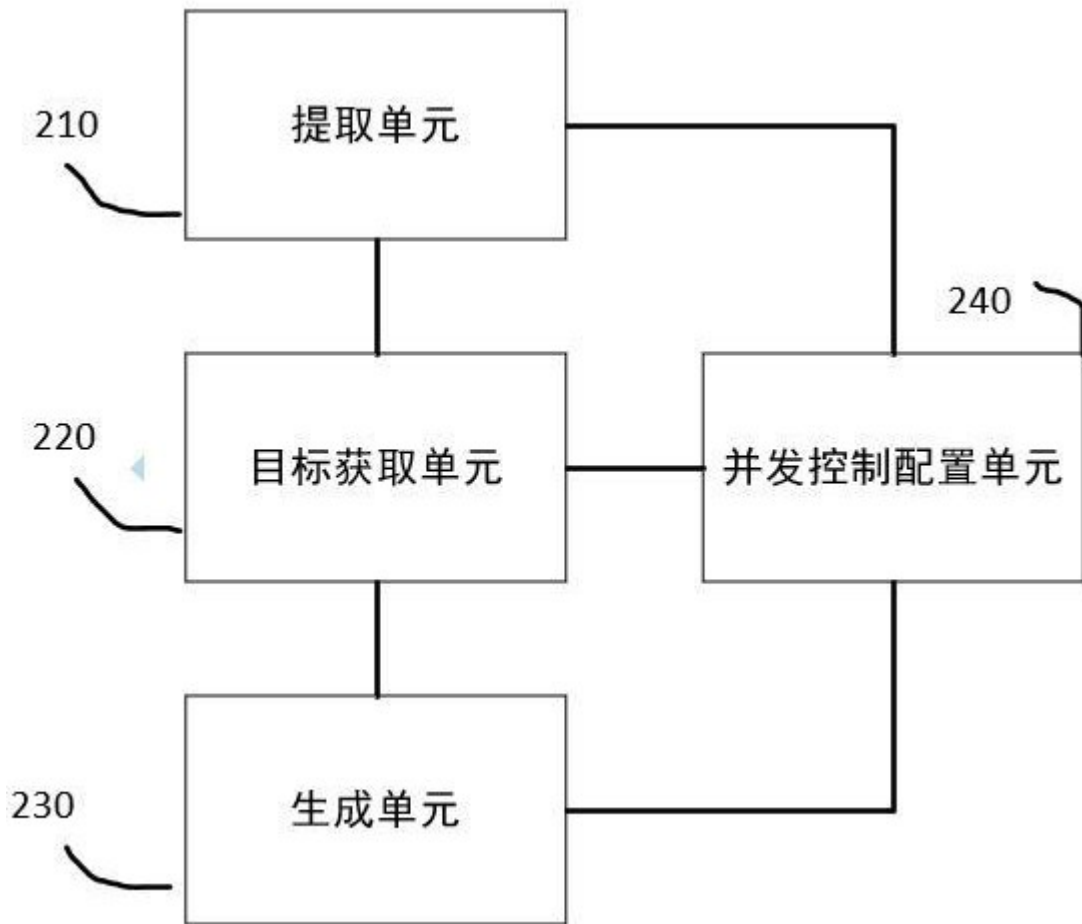


图 3



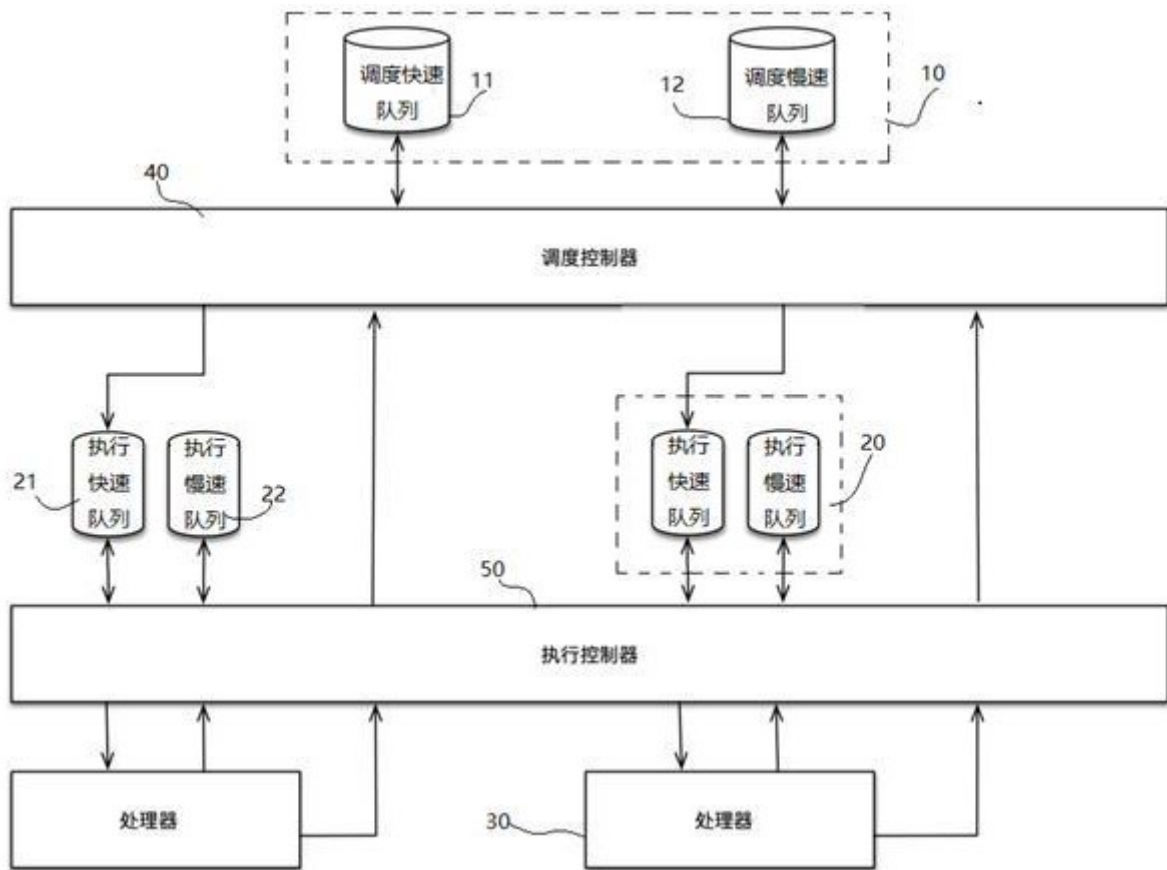


图 4