



(12) 发明专利申请

(10) 申请公布号 CN 115834603 A

(43) 申请公布日 2023. 03. 21

(21) 申请号 202211466310.3

(22) 申请日 2022.11.22

(71) 申请人 山石网科通信技术股份有限公司
地址 215163 江苏省苏州市高新区景润路
181号

(72) 发明人 余祖愿 万腾飞 代应浪 沈亚琪
卢梦梦

(74) 专利代理机构 北京康信知识产权代理有限
责任公司 11240
专利代理师 黄海英

(51) Int.Cl.
H04L 67/1095 (2022.01)
H04L 67/63 (2022.01)
H04L 67/1042 (2022.01)
H04L 41/0668 (2022.01)

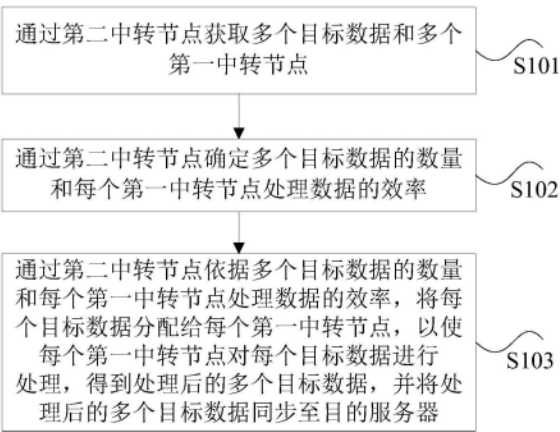
权利要求书3页 说明书20页 附图4页

(54) 发明名称

数据的同步方法及装置、存储介质和处理器

(57) 摘要

本申请公开了一种数据的同步方法及装置、存储介质和处理器。该方法包括：通过第二中转节点获取多个目标数据和多个第一中转节点；通过第二中转节点确定多个目标数据的数量和每个第一中转节点处理数据的效率；通过第二中转节点依据多个目标数据的数量和每个第一中转节点处理数据的效率，将每个目标数据分配给每个第一中转节点，以使每个第一中转节点对每个目标数据进行处理，得到处理后的多个目标数据，并将处理后的多个目标数据同步至目的服务器。通过本申请，解决了相关技术中采用集群环境中的源服务器与用于处理数据的中转节点静态绑定的方法，将集群环境中源服务器的数据同步给目的服务器，导致数据同步的效果较差的问题。



1. 一种数据的同步方法,其特征在于,所述方法应用在数据同步系统中,所述数据同步系统中至少包括多个第一中转节点和第二中转节点,其中所述第一中转节点用于处理数据,所述第二中转节点用于调度所述多个第一中转节点;若所述多个第一中转节点中存在故障节点,则采用所述第一中转节点中除故障节点之外的中转节点代替所述故障节点;若所述第二中转节点发生故障,则采用所述多个第一中转节点中的任意一个中转节点代替所述第二中转节点;该同步方法具体包括:

通过所述第二中转节点获取多个目标数据和所述多个第一中转节点,其中,每个目标数据为集群环境中多个源服务器待同步至目的服务器的数据;

通过所述第二中转节点确定所述多个目标数据的数量和每个第一中转节点处理数据的效率;

通过所述第二中转节点依据所述多个目标数据的数量和每个第一中转节点处理数据的效率,将每个目标数据分配给每个第一中转节点,以使每个第一中转节点对每个目标数据进行处理,得到处理后的多个目标数据,并将所述处理后的多个目标数据同步至所述目的服务器。

2. 根据权利要求1所述的方法,其特征在于,通过所述第二中转节点依据所述多个目标数据的数量和每个第一中转节点处理数据的效率,将每个目标数据分配给每个第一中转节点包括:

依据所述多个目标数据的数量和每个第一中转节点处理数据的效率,计算得到分配给每个第一中转节点的数据量;

基于分配给每个第一中转节点的数据量,将每个目标数据分配给每个第一中转节点。

3. 根据权利要求1所述的方法,其特征在于,在通过所述第二中转节点确定所述多个目标数据的数量和每个第一中转节点处理数据的效率之前,所述方法还包括:

确定多个目标中转节点,其中,所述多个目标中转节点中至少包括所述多个第一中转节点;

从所述多个目标中转节点中确定所述第二中转节点,以通过所述第二中转节点确定所述多个目标数据的数量和每个第一中转节点处理数据的效率,其中,所述第二中转节点为所述多个目标中转节点中除所述多个第一中转节点之外的中转节点。

4. 根据权利要求1所述的方法,其特征在于,在通过所述第二中转节点依据所述多个目标数据的数量和每个第一中转节点处理数据的效率,将每个目标数据分配给每个第一中转节点之后,所述方法还包括:

判断所述多个第一中转节点中是否存在故障节点;

若所述多个第一中转节点中存在故障节点,则从所述多个第一中转节点中确定除故障节点之外的多个第三中转节点,其中,所述第三中转节点用于代替所述第一中转节点对所述多个目标数据进行处理;

获取每个第三中转节点处理数据的效率;

依据所述多个目标数据的数量和每个第三中转节点处理数据的效率,将每个目标数据分配给每个第三中转节点。

5. 根据权利要求3所述的方法,其特征在于,在从所述多个目标中转节点中确定所述第二中转节点之后,所述方法还包括:

判断所述第二中转节点是否发生故障；

若所述第二中转节点发生故障，则从所述多个第一中转节点中确定第四中转节点，其中，所述第四中转节点用于代替所述第二中转节点调度所述多个第一中转节点中除所述第四中转节点之外的中转节点。

6. 根据权利要求1所述的方法，其特征在于，通过所述第二中转节点依据所述多个目标数据的数量和每个第一中转节点处理数据的效率，将每个目标数据分配给每个第一中转节点包括：

判断所述多个目标数据的数据量是否大于预设数据量；

若所述多个目标数据的数据量大于所述预设数据量，则通过所述第二中转节点依据所述多个目标数据的数量和每个第一中转节点处理数据的效率，将每个目标数据分配给每个第一中转节点，以使所述第一中转节点从所述多个源服务器中获取所述多个目标数据，并对所述多个目标数据进行处理；

若所述多个目标数据的数据量不大于所述预设数据量，则通过所述第二中转节点从所述多个源服务器中获取所述多个目标数据；

依据所述多个目标数据的数量和每个第一中转节点处理数据的效率，将从所述多个源服务器中获取的每个目标数据分配给每个第一中转节点。

7. 根据权利要求3所述的方法，其特征在于，从所述多个目标中转节点中确定所述第二中转节点包括：

获取每个目标中转节点对应的编号信息；

基于每个目标中转节点对应的编号信息，确定所述多个目标中转节点中编号最小的中转节点；

将所述多个目标中转节点中编号最小的中转节点作为所述第二中转节点。

8. 一种数据的同步装置，其特征在于，所述装置应用在数据同步系统中，所述数据同步系统中至少包括：多个第一中转节点和第二中转节点，所述第一中转节点用于处理数据，所述第二中转节点用于调度所述多个第一中转节点，若所述多个第一中转节点中存在故障节点，则采用所述第一中转节点中除故障节点之外的中转节点代替所述故障节点，若所述第二中转节点发生故障，则采用所述多个第一中转节点中的任意一个中转节点代替所述第二中转节点，包括：

第一获取模块，用于通过所述第二中转节点获取多个目标数据和所述多个第一中转节点，其中，每个目标数据为集群环境中多个源服务器待同步至目的服务器的数据；

第一确定模块，用于通过所述第二中转节点确定所述多个目标数据的数量和每个第一中转节点处理数据的效率；

第一处理模块，用于通过所述第二中转节点依据所述多个目标数据的数量和每个第一中转节点处理数据的效率，将每个目标数据分配给每个第一中转节点，以使每个第一中转节点对每个目标数据进行处理，得到处理后的多个目标数据，并将所述处理后的多个目标数据同步至所述目的服务器。

9. 一种计算机可读存储介质，其特征在于，所述存储介质存储程序，其中，所述程序执行权利要求1至7中任意一项所述的数据的同步方法。

10. 一种用于数据的同步方法的处理器，其特征在于，所述处理器用于运行程序，其中，

所述程序运行时执行权利要求1至7中任意一项所述的数据的同步方法。

数据的同步方法及装置、存储介质和处理器

技术领域

[0001] 本申请涉及计算机技术领域,具体而言,涉及一种数据的同步方法及装置、存储介质和处理器。

背景技术

[0002] 集群(cluster)技术是一种较新的技术,通过集群技术,可以在付出较低成本的情况下获得在性能、可靠性、灵活性方面的相对较高的收益,其任务调度则是集群系统中的核心技术。

[0003] 而且,一些集群部署的服务器,通常需要向目标网络同步数据,由于数据源服务器与目的服务器所提供的接口不兼容,或者需要的数据形式不一致,需要中转节点将数据进行处理后再进行同步。这里的中转节点既可以是大型服务器,也可以是PC(Personal Computer,个人计算机)上安装的数据处理软件。以下列举三种应用场景:

[0004] (1)在网络安全的审计和控制中,安全厂商常常通过在内部网络中部署代理软件,从集群部署的认证服务器中同步用户,代理软件将用户数据转发至网络安全设备,以帮助网络安全设备完成流量的实名审计和控制。

[0005] (2)在复杂审计系统中,各种设备的日志通常各不相同,为了将日志汇总到日志平台进行统一管理,可以设置中转节点对日志进行统一处理和格式化。

[0006] (3)在小型网络中,集群部署的服务器1存储着可供用户访问的资源,服务器2用于反向代理用户请求,为用户提供资源,服务器1和服务器2之间设置中转节点,进行数据解码、分类、格式化等处理操作。

[0007] 但是,上述的中转节点可能处理数据的能力有限,稳定性也相对有限。为了保证同步效率和同步的稳定性,通常我们需要进行负载均衡和冗余部署,使用多个中转节点对接集群环境中的多个源服务器,将数据处理并转发至目的服务器。

[0008] 相关技术中,采用一对多的方式同步数据,即使用一个中转节点对接集群环境中的多个源服务器,将源服务器的数据同步至目标网络。但是,此方法的缺点在于,单一中转节点处理能力不足,或不具备稳定性,可能造成数据同步节点的拥塞、甚至彻底瘫痪。

[0009] 另外,相关技术中,还采用简单的N对N的方式同步数据,即存在多个中转节点,每个中转节点绑定一个源服务器,将源服务器的数据同步至目标网络。但是,此方法的缺点在于,由于存在绑定关系,因此不方便进行冗余部署,为了防止某个中转节点瘫痪,必须为其设定一个备份的冗余节点,当瘫痪的节点数量大于冗余的节点数量时,同步将会中断或部分中断,此外,由于每个源服务器需要同步的数据量,每个中转节点的实际承载能力都是不同的,静态的绑定关系会使得同步资源分布不均匀,使得部分中转节点过于拥塞,部分中转节点过于空闲。

[0010] 针对相关技术中采用集群环境中的源服务器与用于处理数据的中转节点静态绑定的方法,将集群环境中源服务器的数据同步给目的服务器,导致数据同步的效果较差的问题,目前尚未提出有效的解决方案。

发明内容

[0011] 本申请的主要目的在于提供一种数据的同步方法及装置、存储介质和处理器,以解决相关技术中采用集群环境中的源服务器与用于处理数据的中转节点静态绑定的方法,将集群环境中源服务器的数据同步给目的服务器,导致数据同步的效果较差的问题。

[0012] 为了实现上述目的,根据本申请的一个方面,提供了一种数据的同步方法。该方法应用在数据同步系统中,所述数据同步系统中至少包括:多个第一中转节点和第二中转节点,所述第一中转节点用于处理数据,所述第二中转节点用于调度所述多个第一中转节点,若所述多个第一中转节点中存在故障节点,则采用所述第一中转节点中除故障节点之外的中转节点代替所述故障节点,若所述第二中转节点发生故障,则采用所述多个第一中转节点中的任意一个中转节点代替所述第二中转节点,包括:通过所述第二中转节点获取多个目标数据和所述多个第一中转节点,其中,每个目标数据为集群环境中多个源服务器待同步至目的服务器的数据;通过所述第二中转节点确定所述多个目标数据的数量和每个第一中转节点处理数据的效率;通过所述第二中转节点依据所述多个目标数据的数量和每个第一中转节点处理数据的效率,将每个目标数据分配给每个第一中转节点,以使每个第一中转节点对每个目标数据进行处理,得到处理后的多个目标数据,并将所述处理后的多个目标数据同步至所述目的服务器。

[0013] 进一步地,通过所述第二中转节点依据所述多个目标数据的数量和每个第一中转节点处理数据的效率,将每个目标数据分配给每个第一中转节点包括:依据所述多个目标数据的数量和每个第一中转节点处理数据的效率,计算得到分配给每个第一中转节点的数据量;基于分配给每个第一中转节点的数据量,将每个目标数据分配给每个第一中转节点。

[0014] 进一步地,在通过所述第二中转节点确定所述多个目标数据的数量和每个第一中转节点处理数据的效率之前,所述方法还包括:确定多个目标中转节点,其中,所述多个目标中转节点中至少包括所述多个第一中转节点;从所述多个目标中转节点中确定所述第二中转节点,以通过所述第二中转节点确定所述多个目标数据的数量和每个第一中转节点处理数据的效率,其中,所述第二中转节点为所述多个目标中转节点中除所述多个第一中转节点之外的中转节点。

[0015] 进一步地,在通过所述第二中转节点依据所述多个目标数据的数量和每个第一中转节点处理数据的效率,将每个目标数据分配给每个第一中转节点之后,所述方法还包括:判断所述多个第一中转节点中是否存在故障节点;若所述多个第一中转节点中存在故障节点,则从所述多个第一中转节点中确定除故障节点之外的多个第三中转节点,其中,所述第三中转节点用于代替所述第一中转节点对所述多个目标数据进行处理;获取每个第三中转节点处理数据的效率;依据所述多个目标数据的数量和每个第三中转节点处理数据的效率,将每个目标数据分配给每个第三中转节点。

[0016] 进一步地,在从所述多个目标中转节点中确定所述第二中转节点之后,所述方法还包括:判断所述第二中转节点是否发生故障;若所述第二中转节点发生故障,则从所述多个第一中转节点中确定第四中转节点,其中,所述第四中转节点用于代替所述第二中转节点调度所述多个第一中转节点中除所述第四中转节点之外的中转节点。

[0017] 进一步地,通过所述第二中转节点依据所述多个目标数据的数量和每个第一中转节点处理数据的效率,将每个目标数据分配给每个第一中转节点包括:判断所述多个目标

数据的数据量是否大于预设数据量;若所述多个目标数据的数据量大于所述预设数据量,则通过所述第二中转节点依据所述多个目标数据的数量和每个第一中转节点处理数据的效率,将每个目标数据分配给每个第一中转节点,以使所述第一中转节点从所述多个源服务器中获取所述多个目标数据,并对所述多个目标数据进行处理;若所述多个目标数据的数据量不大于所述预设数据量,则通过所述第二中转节点从所述多个源服务器中获取所述多个目标数据;依据所述多个目标数据的数量和每个第一中转节点处理数据的效率,将从所述多个源服务器中获取的每个目标数据分配给每个第一中转节点。

[0018] 进一步地,从所述多个目标中转节点中确定所述第二中转节点包括:获取每个目标中转节点对应的编号信息;基于每个目标中转节点对应的编号信息,确定所述多个目标中转节点中编号最小的中转节点;将所述多个目标中转节点中编号最小的中转节点作为所述第二中转节点。

[0019] 进一步地,所述方法还包括:获取每个源服务器在单位时间内待同步至目的服务器的数据量;依据每个源服务器在单位时间内待同步至目的服务器的数据量,计算得到每个第一中转节点的平均排队长度,其中,所述平均排队长度用于表示在第一时间段内所述多个目标数据的堆积情况;确定每个第一中转节点的实时排队长度,其中,所述实时排队长度用于表示在第二时间段内每个第一中转节点的负载压力情况,所述第二时间段的长度小于所述第一时间段的长度;基于每个第一中转节点的平均排队长度和实时排队长度,向目标对象发送提醒信息,其中,所述提醒信息用于提醒所述目标对象调整将所述多个目标数据同步至所述目的服务器的策略。

[0020] 进一步地,在获取多个目标数据和多个第一中转节点之后,所述方法还包括:判断所述多个第一中转节点中的第五中转节点是否处于空闲状态;若所述多个第一中转节点中的第五中转节点处于空闲状态,则将多个目标数据中的第一数据分配给所述第五中转节点,以使所述第五中转节点对所述第一数据进行处理,得到处理后的第一数据,并将所述处理后的第一数据同步至所述目的服务器;若所述多个第一中转节点中的第五中转节点不处于空闲状态,则判断所述多个第一中转节点中的第六中转节点是否处于空闲状态;若所述多个第一中转节点中的第六中转节点处于空闲状态,则将多个目标数据中的第一数据分配给所述第六中转节点,以使所述第六中转节点对所述第一数据进行处理,得到处理后的第一数据,并将所述处理后的第一数据同步至所述目的服务器。

[0021] 为了实现上述目的,根据本申请的另一方面,提供了一种数据的同步装置。该装置应用在数据同步系统中,所述数据同步系统中至少包括:多个第一中转节点和第二中转节点,所述第一中转节点用于处理数据,所述第二中转节点用于调度所述多个第一中转节点,若所述多个第一中转节点中存在故障节点,则采用所述第一中转节点中除故障节点之外的中转节点代替所述故障节点,若所述第二中转节点发生故障,则采用所述多个第一中转节点中的任意一个中转节点代替所述第二中转节点,包括:第一获取模块,用于通过所述第二中转节点获取多个目标数据和所述多个第一中转节点,其中,每个目标数据为集群环境中多个源服务器待同步至目的服务器的数据;第一确定模块,用于通过所述第二中转节点确定所述多个目标数据的数量和每个第一中转节点处理数据的效率;第一处理模块,用于通过所述第二中转节点依据所述多个目标数据的数量和每个第一中转节点处理数据的效率,将每个目标数据分配给每个第一中转节点,以使每个第一中转节点对每个目标数据进行处

理,得到处理后的多个目标数据,并将所述处理后的多个目标数据同步至所述目的服务器。

[0022] 进一步地,所述第一处理模块包括:第一计算单元,用于依据所述多个目标数据的数量和每个第一中转节点处理数据的效率,计算得到分配给每个第一中转节点的数据量;第一分配单元,用于基于分配给每个第一中转节点的数据量,将每个目标数据分配给每个第一中转节点。

[0023] 进一步地,所述装置还包括:第二确定模块,用于在通过所述第二中转节点确定所述多个目标数据的数量和每个第一中转节点处理数据的效率之前,确定多个目标中转节点,其中,所述多个目标中转节点中至少包括所述多个第一中转节点;第三确定模块,用于从所述多个目标中转节点中确定所述第二中转节点,以通过所述第二中转节点确定所述多个目标数据的数量和每个第一中转节点处理数据的效率,其中,所述第二中转节点为所述多个目标中转节点中除所述多个第一中转节点之外的中转节点。

[0024] 进一步地,所述装置还包括:第一判断模块,用于在通过所述第二中转节点依据所述多个目标数据的数量和每个第一中转节点处理数据的效率,将每个目标数据分配给每个第一中转节点之后,判断所述多个第一中转节点中是否存在故障节点;第四确定模块,用于若所述多个第一中转节点中存在故障节点,则从所述多个第一中转节点中确定除故障节点之外的多个第三中转节点,其中,所述第三中转节点用于代替所述第一中转节点对所述多个目标数据进行处理;第二获取模块,用于获取每个第三中转节点处理数据的效率;第一分配模块,用于依据所述多个目标数据的数量和每个第三中转节点处理数据的效率,将每个目标数据分配给每个第三中转节点。

[0025] 进一步地,所述装置还包括:第二判断模块,用于在从所述多个目标中转节点中确定所述第二中转节点之后,判断所述第二中转节点是否发生故障;第五确定模块,用于若所述第二中转节点发生故障,则从所述多个第一中转节点中确定第四中转节点,其中,所述第四中转节点用于代替所述第二中转节点调度所述多个第一中转节点中除所述第四中转节点之外的中转节点。

[0026] 进一步地,所述第一处理模块包括:第一判断单元,用于判断所述多个目标数据的数据量是否大于预设数据量;第二分配单元,用于若所述多个目标数据的数据量大于所述预设数据量,则通过所述第二中转节点依据所述多个目标数据的数量和每个第一中转节点处理数据的效率,将每个目标数据分配给每个第一中转节点,以使所述第一中转节点从所述多个源服务器中获取所述多个目标数据,并对所述多个目标数据进行处理;第一获取单元,用于若所述多个目标数据的数据量不大于所述预设数据量,则通过所述第二中转节点从所述多个源服务器中获取所述多个目标数据;第三分配单元,用于依据所述多个目标数据的数量和每个第一中转节点处理数据的效率,将从所述多个源服务器中获取的每个目标数据分配给每个第一中转节点。

[0027] 进一步地,所述第三确定模块包括:第二获取单元,用于获取每个目标中转节点对应的编号信息;第一确定单元,用于基于每个目标中转节点对应的编号信息,确定所述多个目标中转节点中编号最小的中转节点;第二确定单元,用于将所述多个目标中转节点中编号最小的中转节点作为所述第二中转节点。

[0028] 进一步地,所述装置还包括:第三获取模块,用于获取每个源服务器在单位时间内待同步至目的服务器的数据量;第一计算模块,用于依据每个源服务器在单位时间内待同

步至目的服务器的数据量,计算得到每个第一中转节点的平均排队长度,其中,所述平均排队长度用于表示在第一时间段内所述多个目标数据的堆积情况;第六确定模块,用于确定每个第一中转节点的实时排队长度,其中,所述实时排队长度用于表示在第二时间段内每个第一中转节点的负载压力情况,所述第二时间段的长度小于所述第一时间段的长度;第一发送模块,用于基于每个第一中转节点的平均排队长度和实时排队长度,向目标对象发送提醒信息,其中,所述提醒信息用于提醒所述目标对象调整将所述多个目标数据同步至所述目的服务器的策略。

[0029] 进一步地,所述装置还包括:第三判断模块,用于在获取多个目标数据和多个第一中转节点之后,判断所述多个第一中转节点中的第五中转节点是否处于空闲状态;第二分配模块,用于若所述多个第一中转节点中的第五中转节点处于空闲状态,则将多个目标数据中的第一数据分配给所述第五中转节点,以使所述第五中转节点对所述第一数据进行处理,得到处理后的第一数据,并将所述处理后的第一数据同步至所述目的服务器;第四判断模块,用于若所述多个第一中转节点中的第五中转节点不处于空闲状态,则判断所述多个第一中转节点中的第六中转节点是否处于空闲状态;第三分配模块,用于若所述多个第一中转节点中的第六中转节点处于空闲状态,则将多个目标数据中的第一数据分配给所述第六中转节点,以使所述第六中转节点对所述第一数据进行处理,得到处理后的第一数据,并将所述处理后的第一数据同步至所述目的服务器。

[0030] 为了实现上述目的,根据本申请的另一方面,提供了一种计算机可读存储介质,所述存储介质存储程序,其中,所述程序执行上述的任意一项所述的数据的同步方法。

[0031] 为了实现上述目的,根据本申请的另一方面,提供了一种处理器,所述处理器用于运行程序,其中,所述程序运行时执行上述的任意一项所述的数据的同步方法。

[0032] 通过本申请,采用以下步骤:通过第二中转节点获取多个目标数据和多个第一中转节点,其中,每个目标数据为集群环境中多个源服务器待同步至目的服务器的数据;通过第二中转节点确定多个目标数据的数量和每个第一中转节点处理数据的效率;通过第二中转节点依据多个目标数据的数量和每个第一中转节点处理数据的效率,将每个目标数据分配给每个第一中转节点,以使每个第一中转节点对每个目标数据进行处理,得到处理后的多个目标数据,并将处理后的多个目标数据同步至目的服务器,解决了相关技术中采用集群环境中的源服务器与用于处理数据的中转节点静态绑定的方法,将集群环境中源服务器的数据同步给目的服务器,导致数据同步的效果较差的问题。通过第二中转节点依据确定的多个目标数据的数量和每个第一中转节点处理数据的效率,将每个目标数据分配给每个第一中转节点,以使每个第一中转节点对每个目标数据进行处理,得到处理后的多个目标数据,并将处理后的多个目标数据同步至目的服务器,从而提升了数据同步的高效性和稳定性,进而提升了数据同步的效果。

附图说明

[0033] 构成本申请的一部分的附图用来提供对本申请的进一步理解,本申请的示意性实施例及其说明用于解释本申请,并不构成对本申请的不当限定。在附图中:

[0034] 图1是根据本申请实施例提供的数据的同步方法的流程图;

[0035] 图2是根据本申请实施例提供的部署数据同步装置的示意图;

[0036] 图3是本申请实施例中当需要处理的数据占据大量带宽时调度节点作任务数据转发的示意图；

[0037] 图4是本申请实施例中当需要处理的数据无需占据大量带宽时调度节点作任务数据转发的示意图；

[0038] 图5是根据本申请实施例提供的数据的同步方法的流程图一；

[0039] 图6是根据本申请实施例提供的可选的数据的同步方法的流程图；

[0040] 图7是根据本申请实施例提供的数据的同步装置的示意图。

具体实施方式

[0041] 需要说明的是，在不冲突的情况下，本申请中的实施例及实施例中的特征可以相互组合。下面将参考附图并结合实施例来详细说明本申请。

[0042] 为了使本技术领域的人员更好地理解本申请方案，下面将结合本申请实施例中的附图，对本申请实施例中的技术方案进行清楚、完整地描述，显然，所描述的实施例仅仅是本申请一部分的实施例，而不是全部的实施例。基于本申请中的实施例，本领域普通技术人员在没有做出创造性劳动前提下所获得的所有其他实施例，都应当属于本申请保护的范围。

[0043] 需要说明的是，本申请的说明书和权利要求书及上述附图中的术语“第一”、“第二”等是用于区别类似的对象，而不必用于描述特定的顺序或先后次序。应该理解这样使用的数据在适当情况下可以互换，以便这里描述的本申请的实施例。此外，术语“包括”和“具有”以及他们的任何变形，意图在于覆盖不排他的包含，例如，包含了一系列步骤或单元的过程、方法、系统、产品或设备不必限于清楚地列出的那些步骤或单元，而是可包括没有清楚地列出的或对于这些过程、方法、产品或设备固有的其它步骤或单元。

[0044] 实施例1

[0045] 下面结合优选的实施步骤对本发明进行说明，图1是根据本申请实施例提供的数据的同步方法的流程图，如图1所示，该方法应用在数据同步系统中，所述数据同步系统中至少包括：多个第一中转节点和第二中转节点，所述第一中转节点用于处理数据，所述第二中转节点用于调度所述多个第一中转节点，若所述多个第一中转节点中存在故障节点，则采用所述第一中转节点中除故障节点之外的中转节点代替所述故障节点，若所述第二中转节点发生故障，则采用所述多个第一中转节点中的任意一个中转节点代替所述第二中转节点，并包括如下步骤：

[0046] 步骤S101，通过第二中转节点获取多个目标数据和多个第一中转节点，其中，每个目标数据为集群环境中多个源服务器待同步至目的服务器的数据。

[0047] 例如，图2是根据本申请实施例提供的部署数据同步装置的示意图，如图2所示，数据同步装置中目的服务器的数量可以为多个，由于本申请实施例不关注目的服务器与各个中转节点的绑定关系，因此可以将目的服务器视为一个。

[0048] 比如，假定存在 m 个源服务器， n 个中转节点，中转节点从源服务器同步数据，经过处理，并发送给目的服务器。

[0049] 另外，所有中转节点中，将其中一个中转节点视作调度节点（上述的第二中转节点），调度节点负责所有中转节点的调度安排，其它节点负责数据处理，也称作数据处理节

点(上述的第一中转节点)。且对于任何一个中转节点,其内部结构都由网络模块和多进程的数据处理模块构成,前者负责数据的收发,后者负责数据的处理。调度节点与数据处理节点没有区别,只是在实际运行时被选举出来,用于调度其它节点,如有必要,任何一个中转节点都可以作为调度节点。也即,若一个或多个数据处理节点发生故障时,调度节点可以记录故障的数据处理节点,并选取无故障的数据处理节点接替故障的数据处理节点的工作。若当前调度节点发生故障时,可以按照编号顺序从数据处理节点中选举新的调度节点,例如若编号为1的调度节点故障,则编号为2的节点为新的调度节点,若编号为2的调度节点故障,则编号为3的节点为新的调度节点。

[0050] 例如,由调度节点确定需要从集群环境中的多个源服务器同步至目的服务器的数据(上述的多个目标数据),并确定数据同步装置中存在哪些数据处理节点。

[0051] 步骤S102,通过第二中转节点确定多个目标数据的数量和每个第一中转节点处理数据的效率。

[0052] 例如,通过调度节点确定需要从集群环境中的多个源服务器同步至目的服务器的数据的数量,并计算每个数据处理节点单位时间内处理的数据量,即平均处理率(上述的每个第一中转节点处理数据的效率)。

[0053] 步骤S103,通过第二中转节点依据多个目标数据的数量和每个第一中转节点处理数据的效率,将每个目标数据分配给每个第一中转节点,以使每个第一中转节点对每个目标数据进行处理,得到处理后的多个目标数据,并将处理后的多个目标数据同步至目的服务器。

[0054] 例如,调度节点根据需要从集群环境中的多个源服务器同步至目的服务器的数据的数量和每个数据处理节点单位时间内处理的数据量(平均处理率),将待同步的数据分配给每个数据处理节点进行处理。然后每个数据处理节点接收到分配给其的数据之后,对自己分配到的数据进行处理,得到处理后的数据,再将处理后的数据同步给目的服务器。

[0055] 通过上述的步骤S101至S103,通过第二中转节点依据确定的多个目标数据的数量和每个第一中转节点处理数据的效率,将每个目标数据分配给每个第一中转节点,以使每个第一中转节点对每个目标数据进行处理,得到处理后的多个目标数据,并将处理后的多个目标数据同步至目的服务器,从而提升了数据同步的高效性和稳定性,进而提升了数据同步的效果。

[0056] 为了快速准确的确定第二中转节点,在本申请实施例提供的数据的同步方法中,还可以通过以下步骤确定第二中转节点:确定多个目标中转节点,其中,多个目标中转节点中至少包括多个第一中转节点;从多个目标中转节点中确定第二中转节点,以通过第二中转节点确定多个目标数据的数量和每个第一中转节点处理数据的效率,其中,第二中转节点为多个目标中转节点中除多个第一中转节点之外的中转节点。

[0057] 例如,如图2所示,数据同步装置中可以存在多个中转节点(即对应图2中的中转节点1、中转节点2、中转节点3、中转节点4等),且图2中的中转节点1、中转节点2、中转节点3、中转节点4等可以为上述的目标中转节点。然后可以从中转节点1、中转节点2、中转节点3、中转节点4等这些中转节点中选取任意一个中转节点作为调度节点(上述的第二中转节点),再由调度节点确定需要从集群环境中的多个源服务器同步至目的服务器的数据的数量,并计算每个数据处理节点单位时间内处理的数据量,即平均处理率(上述的每个第一中

转节点处理数据的效率)。另外,如果将中转节点1、中转节点2、中转节点3、中转节点4等这些中转节点中的中转节点1作为调度节点,则将剩余的中转节点2、中转节点3、中转节点4等这些中转节点作为数据处理节点(上述的多个第一中转节点)。

[0058] 综上所述,设计多节点的同步装置,并选取一个节点为调度节点,利用调度节点统一对接集群的各个服务器,处理节点基于负载均衡完成处理工作,易于部署,同时能够避免单一节点负载不足。

[0059] 为了快速准确的确定第二中转节点,在本申请实施例提供的数据的同步方法中,还可以通过以下步骤确定第二中转节点:获取每个目标中转节点对应的编号信息;基于每个目标中转节点对应的编号信息,确定多个目标中转节点中编号最小的中转节点;将多个目标中转节点中编号最小的中转节点作为第二中转节点。

[0060] 例如,可以按照加入系统的时间,给中转节点自动编号,并可以将编号为1的中转节点自动选举为调度节点,调度节点负责节点的调度和任务的分配,不负责数据处理。

[0061] 通过上述的方案,可以快速准确的从多个中转节点中确定出调度节点。

[0062] 为了快速准确的将每个目标数据分配给每个第一中转节点,在本申请实施例提供的数据的同步方法中,还可以通过以下步骤将每个目标数据分配给每个第一中转节点:判断多个目标数据的数据量是否大于预设数据量;若多个目标数据的数据量大于预设数据量,则通过第二中转节点依据多个目标数据的数量和每个第一中转节点处理数据的效率,将每个目标数据分配给每个第一中转节点,以使第一中转节点从多个源服务器中获取多个目标数据,并对多个目标数据进行处理;若多个目标数据的数据量不大于预设数据量,则通过第二中转节点从多个源服务器中获取多个目标数据;依据多个目标数据的数量和每个第一中转节点处理数据的效率,将从多个源服务器中获取的每个目标数据分配给每个第一中转节点。

[0063] 例如,当需要处理的数据占据大量带宽,而调度节点传输能力有限时,调度节点作任务分配时,其过程为:调度节点从源服务器获取任务相关信息,如数据的编号等,不实际同步数据。当需要将数据同步工作分配给其它中转节点时,发送消息告知其它中转节点,任务对应的源服务器、数据的编号、数量等信息,由中转节点自行从源服务器获取数据。且当需要处理的数据占据大量带宽时,调度节点作任务数据转发的示意图,如图3所示。

[0064] 例如,当需要处理的数据无需占据大量带宽,而调度节点传输能力充足时,调度节点作任务分配时,其过程为:调度节点从源服务器获取任务相关信息以及数据本身。当需要将数据同步工作分配给其它中转节点时,发送消息将数据转发给其它中转节点,此时无需其它中转节点与源服务器建立连接。且当需要处理的数据无需占据大量带宽时,调度节点作任务数据转发的示意图,如图4所示。

[0065] 综上所述,可以根据同步的数据类型,设置两种部署方式,当数据占据大量带宽时,调度节点作任务指派,当数据不占据大量带宽时,调度节点作实际的任务数据分发。

[0066] 为了快速准确的得到处理后的第一数据,在本申请实施例提供的数据的同步方法中,还可以通过以下步骤得到处理后的第一数据:判断多个第一中转节点中的第五中转节点是否处于空闲状态;若多个第一中转节点中的第五中转节点处于空闲状态,则将多个目标数据中的第一数据分配给第五中转节点,以使第五中转节点对第一数据进行处理,得到处理后的第一数据,并将处理后的第一数据同步至目的服务器;若多个第一中转节点中的

第五中转节点不处于空闲状态,则判断多个第一中转节点中的第六中转节点是否处于空闲状态;若多个第一中转节点中的第六中转节点处于空闲状态,则将多个目标数据中的第一数据分配给第六中转节点,以使第六中转节点对第一数据进行处理,得到处理后的第一数据,并将处理后的第一数据同步至目的服务器。

[0067] 例如,系统初始运行时,采用相对简单的轮询负载均衡方法,其过程如下:调度节点从各个源服务器获取任务,每次遍历其它中转节点,将固定量的任务分配给各个中转节点。分配前查询目标中转节点是否空闲,若空闲则将任务分配给该中转节点,若不空闲则查询下一个中转节点。

[0068] 通过上述的方案,初始时调度节点可以按照轮询负载均衡法,快速的将待同步的数据分配给每个源服务器进行处理。

[0069] 图5是根据本申请实施例提供的数据的同步方法的流程图一,如图5所示,在本申请实施例提供的数据的同步方法中,通过第二中转节点依据多个目标数据的数量和每个第一中转节点处理数据的效率,将每个目标数据分配给每个第一中转节点包括:

[0070] 步骤S501,依据多个目标数据的数量和每个第一中转节点处理数据的效率,计算得到分配给每个第一中转节点的数据量;

[0071] 步骤S502,基于分配给每个第一中转节点的数据量,将每个目标数据分配给每个第一中转节点。

[0072] 例如,在进行指标监测时,可以先计算空闲时间,具体为,对于任一个中转节点,每间隔一段时间接受调度节点分配任务。由中转节点接受调度节点分配的任务,接受任务后进行处理,处理完毕后等待下一次任务的到来,期间调度节点可能会多次询问该节点是否空闲,因为当前任务未处理完,调度节点不会为其分配新的任务。现假定某次接受任务的时间为 t_1 ,本次任务处理完毕的时间为 t_2 ,下次接受任务的时间为 t_3 。 t_2 介于 t_1 与 t_3 之间,在 t_1 到 t_3 时间内,其空闲时间 $T_{idle}=t_3-t_2$ 。

[0073] 然后计算平均处理率,且设定 μ 为中转节点单位时间处理的数据量,即平均处理率。并且,不同类型的数据通常采用不同的计量单位,例如,本申请实施例通过中转节点处理源服务器日志时,可以将日志条数作为计量单位,其计算公式为:

[0074]
$$\mu = \frac{p}{T - T_{idle}}$$

[0075] 其中, p 为统计时间内该中转节点处理的数据量总和, T_{idle} 为统计时间内该中转节点空闲时间总和(即实际不处理数据的时间段), T 为统计时间。

[0076] 例如,调度节点从各个源服务器获取任务,以 μ 作为权值加权分配任务。假定调度节点一次获取各个源服务器的 n 条数据,各个处理节点的 μ 值分别为 $\mu_1, \mu_2, \dots, \mu_{n-1}$ 。

[0077] 则分配给中转节点 k 的数据量为:

[0078]
$$\frac{\mu_k}{\sum_{i=1}^{n-1} \mu_i} \cdot n$$

[0079] 然后,调度节点可以根据计算得到的分配给每个数据处理中转节点的数据量,将待同步至目的服务器的数据分配给每个数据处理中转节点。

[0080] 综上所述,基于统计指标进行动态负载均衡,能够根据节点负载能力分摊同步任

务,尽可能降低各个节点的负载。

[0081] 为了快速准确的将多个目标数据同步至目的服务器,在本申请实施例提供的数据的同步方法中,还可以通过以下步骤将多个目标数据同步至目的服务器:获取每个源服务器在单位时间内待同步至目的服务器的数据量;依据每个源服务器在单位时间内待同步至目的服务器的数据量,计算得到每个第一中转节点的平均排队长度,其中,平均排队长度用于表示在第一时间段内多个目标数据的堆积情况;确定每个第一中转节点的实时排队长度,其中,实时排队长度用于表示在第二时间段内每个第一中转节点的负载压力情况,第二时间段的长度小于第一时间段的长度;基于每个第一中转节点的平均排队长度和实时排队长度,向目标对象发送提醒信息,其中,提醒信息用于提醒目标对象调整将多个目标数据同步至目的服务器的策略。

[0082] 例如,调度节点可以记录如下指标: λ ,为所有源服务器单位时间需要同步的数据量,即平均到达率,其计量单位与 μ 相同。 λ 的计算公式为:

$$[0083] \quad \lambda = \frac{q}{T}$$

[0084] 其中,T为统计时间,q为统计时间内所有源服务器所需要同步的数据量总和。

[0085] 然后,调度节点可以关注当前负载情况是否过重,计算相关指标并提供给用户。其计算指标包括实时排队长度L1和平均排队长度L2,实时排队长度L1是调度节点记录的当前未能完成分配的数据量,实时排队长度反映了当前短时间内中转节点负载压力情况。

[0086] 平均排队长度L2,反映较长时间里的数据堆积情况,反映了中转节点设置是否合理,若该值较大,建议增加中转节点。其计算公式如下:

$$[0087] \quad L_2 = \frac{\lambda^2}{\mu(\mu - \lambda)}$$

[0088] 式子中, λ :所有源服务器单位时间需要同步的数据量, μ :中转节点单位时间处理

的数据量总和,即 $\mu = \sum_{i=1}^{i=n-1} \mu_i$ 。

[0089] 通过上述的方案,在将源服务器中的数据同步至目的服务器时,可以及时的提醒用户对数据同步的方法进行调整,从而提升数据同步的稳定性和高效性。

[0090] 为了快速准确的对中转节点进行故障处理,在本申请实施例提供的数据的同步方法中,还可以通过以下步骤对中转节点进行故障处理:判断第二中转节点是否发生故障;若第二中转节点发生故障,则从多个第一中转节点中确定第四中转节点,其中,第四中转节点用于代替第二中转节点调度多个第一中转节点中除第四中转节点之外的中转节点。

[0091] 例如,本申请实施例所设计的装置除了可以根据排队模型实现动态负载均衡之外,还可以具备故障顶替功能。

[0092] 即当装置中的任一中转节点发生故障,该装置能够实现迅速的故障顶替。

[0093] 若当前调度节点发生故障时,装置执行如下过程:

[0094] (1) 按照编号顺序选举新的调度节点,例如若编号为1的调度节点故障,则编号为2的节点为新的调度节点,若编号为2的调度节点故障,则编号为3的节点为新的调度节点;

[0095] (2) 重新执行初始化部署过程,即重新部署数据同步装置,采用新的中转节点作为

调度节点；

[0096] (3) 重新执行指标监测过程,即初始化完毕后,系统运行并开始观测,每个处理节点记录 μ ,为中转节点单位时间处理的数据量,即平均处理率,并且,调度节点记录 λ ,为所有源服务器单位时间需要同步的数据量,即平均到达率;

[0097] (4) 重新执行动态调整过程,即调度节点根据指标监测过程中监测到的指标变化,对分配给每个源服务器的数据进行动态调整。

[0098] 综上所述,调度节点与统计节点结构完全相同,可以相互替代,二者任一发生故障,都可以进行快速顶替。

[0099] 为了快速准确的对中转节点进行故障处理,在本申请实施例提供的数据的同步方法中,还可以通过以下步骤对中转节点进行故障处理:判断多个第一中转节点中是否存在故障节点;若多个第一中转节点中存在故障节点,则从多个第一中转节点中确定除故障节点之外的多个第三中转节点,其中,第三中转节点用于代替第一中转节点对多个目标数据进行处理;获取每个第三中转节点处理数据的效率;依据多个目标数据的数量和每个第三中转节点处理数据的效率,将每个目标数据分配给每个第三中转节点。

[0100] 例如,若一个或多个处理节点发生故障时,装置执行如下过程:

[0101] 调度节点记录故障的处理节点,继续执行上述的指标监测过程,并将上述的动态调整过程更新如下:

[0102] 分配给某一中转节点k(中转节点为非故障节点)的数据量变更为:

$$[0103] \quad \frac{\mu_k}{\sum_{i=1}^{i=n-1} \mu_i - \sum_{j=1}^{j=b-1} \mu_j} \cdot n$$

[0104] 式中b表示故障节点的个数。

[0105] 另外,无论是调度节点发生故障,还是处理节点发生故障,当发生节点故障后,实时排队长度L1和平均排队长度L2将可能增长,此时用户可以据此增加中转节点数量,将新的中转节点加入装置,新加入的装置自动作为处理节点,其加入过程为:

[0106] (1) 重新执行初始化部署过程,即重新部署数据同步装置,采用新的中转节点作为调度节点或者数据处理中转节点;

[0107] (2) 重新执行指标监测过程,即初始化完毕后,系统运行并开始观测,每个处理节点记录 μ ,为中转节点单位时间处理的数据量,即平均处理率,并且,调度节点记录 λ ,为所有源服务器单位时间需要同步的数据量,即平均到达率;

[0108] (3) 重新执行动态调整过程,即调度节点根据指标监测过程中监测到的指标变化,对分配给每个源服务器的数据进行动态调整,且上述公式中的n值将包含新加入的中转节点。

[0109] 综上所述,调度节点与统计节点可以相互替代,二者任一发生故障,都可以进行快速顶替。

[0110] 例如,图6是根据本申请实施例提供的可选的数据的同步方法的流程图,如图6所示,可选的数据的同步方法包括如下步骤:

[0111] 1. 部署装置

[0112] 本申请实施例做作如下部署,其中目的服务器的数量可以为多个,本申请实施例

方案不关注目的服务器与各个中转节点的绑定关系,因此将目的服务器视为一个。

[0113] 假定存在m个源服务器,n个中转节点,中转节点从源服务器同步数据,经过处理,并发送给目的服务器。

[0114] 所有中转节点中,将其中一个中转节点视作调度节点,调度节点负责所有中转节点的调度安排,其它节点负责数据处理,也称作数据处理节点。

[0115] 对于任何一个中转节点,其内部结构都由网络模块和多进程的数据处理模块构成,前者负责数据的收发,后者负责数据的处理。调度节点与数据处理节点没有区别,只是在实际运行时被选举出来,用于调度其它节点,如有必要,任何一个中转节点都可以作为调度节点。

[0116] 并且,按照加入系统的时间,给中转节点自动编号,将编号为1的中转节点自动选举为调度节点,调度节点负责节点的调度和任务的分配,不负责数据处理。

[0117] 当需要处理的数据占据大量带宽,而调度节点传输能力有限时,调度节点作任务分配时,其过程为:调度节点从源服务器获取任务相关信息,如数据的编号等,不实际同步数据。当需要将数据同步工作分配给其它中转节点时,发送消息告知其它中转节点,任务对应的源服务器、数据的编号、数量等信息,由中转节点自行从源服务器获取数据。

[0118] 当需要处理的数据无需占据大量带宽,而调度节点传输能力充足时,调度节点作任务分配时,其过程为:调度节点从源服务器获取任务相关信息以及数据本身。当需要将数据同步工作分配给其它中转节点时,发送消息将数据转发给其它中转节点,此时无需其它中转节点与源服务器建立连接。

[0119] 系统初始运行时,采用相对简单的轮询负载均衡方法,其过程如下:调度节点从各个源服务器获取任务,每次遍历其它中转节点,将固定量的任务分配给各个中转节点。分配前查询目标中转节点是否空闲,若空闲则将任务分配给该中转节点,若不空闲则查询下一个中转节点。

[0120] 2. 指标监测

[0121] (1) 空闲时间的计算

[0122] 对于任一个中转节点,每间隔一段时间接受调度节点分配任务。

[0123] 由中转节点接受调度节点分配的任务,接受任务后进行处理,处理完毕后等待下一次任务的到来,期间调度节点可能会多次询问该节点是否空闲,因为当前任务未处理完,调度节点不会为其分配新的任务。现假定某次接受任务的时间为t1,本次任务处理完毕的时间为t2,下次接受任务的时间为t3。t2介于t1与t3之间,在t1到t3时间内,其空闲时间Tidle=t3-t2。

[0124] (2) 平均处理率与平均到达率的计算

[0125] 在第1节初始化完毕后,系统运行并开始观测,每个处理节点记录如下指标:

[0126] μ : 中转节点单位时间处理的数据量,即平均处理率,不同类型的数据通常采用不同的计量单位,例如本申请实施例通过中转节点处理源服务器日志时,可以将日志条数作为计量单位。其计算公式为:

$$[0127] \quad \mu = \frac{p}{T - T_{idle}}$$

[0128] 其中,p为统计时间内该中转节点处理的数据量总和,Tidle为统计时间内该中转

节点空闲时间总和(即实际不处理数据的时间段),T为统计时间。

[0129] 调度节点记录如下指标:

[0130] λ :所有源服务器单位时间需要同步的数据量,即平均到达率,其计量单位与 μ 相同。 λ 的计算公式为:

$$[0131] \quad \lambda = \frac{q}{T}$$

[0132] 其中,T为统计时间,q为统计时间内所有源服务器所需要同步的数据量总和。

[0133] 3. 动态调整

[0134] 根据检测指标进行动态调整的目的在于按照节点负载能力分配任务,减小每个节点的负担,同时减小数据平均处理时间,避免数据滞留。当系统需要处理的数据量过大,各个中转节点不足以负担时,调度节点将给出预警。

[0135] 若一直沿用第1节的轮询法,很有可能造成某个节点长时间负载过重的情况。因此需要按照第2节的监测指标进行动态调整。

[0136] 调度节点从各个源服务器获取任务,以 μ 作为权值加权分配任务。假定调度节点一次获取各个源服务器的n条数据,各个处理节点的 μ 值分别为 $\mu_1, \mu_2, \dots, \mu_{n-1}$ 。

[0137] 则分配给中转节点k的数据量为:

$$[0138] \quad \frac{\mu_k}{\sum_{i=1}^{i=n-1} \mu_i} \cdot n$$

[0139] 中转节点以一定周期统计 μ 和 λ ,当 μ 和 λ 经过t检验有显著差异时,更新二者的值。

[0140] 调度节点关注当前负载情况是否过重,计算相关指标并提供给用户。其计算指标包括实时排队长度L1和平均排队长度L2,实时排队长度L1是调度节点记录的当前未能完成分配的数据量,实时排队长度反映了当前短时间内中转节点负载压力情况。

[0141] 平均排队长度L2,反映较长时间里的数据堆积情况,反映了中转节点设置是否合理,若该值较大,建议增加中转节点。其计算公式如下:

$$[0142] \quad L_2 = \frac{\lambda^2}{\mu(\mu - \lambda)}$$

[0143] 式子中, λ :所有源服务器单位时间需要同步的数据量

[0144] μ :中转节点单位时间处理的数据量总和,即 $\mu = \sum_{i=1}^{i=n-1} \mu_i$ 。

[0145] 4. 故障顶替

[0146] 本申请实施例所设计的装置除了根据排队模型实现动态负载均衡之外,还具备故障顶替功能。

[0147] 即当装置中的任一中转节点发生故障,该装置能够实现迅速的故障顶替。

[0148] 若当前调度节点发生故障时,装置执行如下过程:

[0149] (1) 按照编号顺序选举新的调度节点,例如若编号为1的调度节点故障,则编号为2的节点为新的调度节点,若编号为2的调度节点故障,则编号为3的节点为新的调度节点。

[0150] (2) 重新执行第1节所述的初始化过程。

[0151] (3) 重新执行第2节所述的指标监测过程。

[0152] (4) 重新执行第3节所述的动态调整过程。

[0153] 若一个或多个处理节点发生故障时,装置执行如下过程:

[0154] 调度节点记录故障的处理节点,继续执行第2节所述的指标监测过程,将第3节所述的动态调整过程更新如下:

[0155] 分配给某一中转节点k(中转节点为非故障节点)的数据量变更为:

$$[0156] \quad \frac{\mu_k}{\sum_{i=1}^{i=n-1} \mu_i - \sum_{j=1}^{j=b-1} \mu_j} \cdot n$$

[0157] 式中b表示故障节点的个数。

[0158] 无论是调度节点发生故障,还是处理节点发生故障,当发生节点故障后,实时排队长度L1和平均排队长度L2将可能增长,此时用户可以据此增加中转节点数量,将新的中转节点加入装置,新加入的装置自动作为处理节点,其加入过程为:

[0159] (1) 重新执行第1节所述的初始化过程。

[0160] (2) 重新执行第2节所述的指标监测过程。

[0161] (3) 重新执行第3节所述的动态调整过程,上述公式中的n值将包含新加入的中转节点。

[0162] 比如,在网络安全审计和控制中,安全厂商常常通过在内部网络中部署代理软件,从集群部署的认证服务器中同步用户,代理软件将数据转发至网络安全设备,以帮助网络安全设备完成流量的实名审计和控制。在网络安全审计中,不同网络安全设备的日志通常各不相同,为了将日志汇总到日志平台进行统一管理,可以设置中转节点对日志进行统一处理和格式化。在小型网络中,集群部署的服务器1存储着可供用户访问的资源,服务器2用于反向代理用户请求,为用户提供资源,服务器1和服务器2之间设置中转节点,进行数据解码、分类、格式化等处理操作。

[0163] 因此,通过本申请实施例提供的方法,可以使用多个中转节点动态对接集群环境中的多个源服务器,将数据处理并转发至目的服务器。且基于中转节点的处理效率和源服务器的数据量动态调整二者的绑定关系,可以提升数据同步的高效性和稳定性。

[0164] 综上,本申请实施例提供的数据的同步方法,通过第二中转节点获取多个目标数据和多个第一中转节点,其中,每个目标数据为集群环境中多个源服务器待同步至目的服务器的数据;通过第二中转节点确定多个目标数据的数量和每个第一中转节点处理数据的效率;通过第二中转节点依据多个目标数据的数量和每个第一中转节点处理数据的效率,将每个目标数据分配给每个第一中转节点,以使每个第一中转节点对每个目标数据进行处理,得到处理后的多个目标数据,并将处理后的多个目标数据同步至目的服务器,解决了相关技术中采用集群环境中的源服务器与用于处理数据的中转节点静态绑定的方法,将集群环境中源服务器的数据同步给目的服务器,导致数据同步的效果较差的问题。通过第二中转节点依据确定的多个目标数据的数量和每个第一中转节点处理数据的效率,将每个目标数据分配给每个第一中转节点,以使每个第一中转节点对每个目标数据进行处理,得到处理后的多个目标数据,并将处理后的多个目标数据同步至目的服务器,从而提升了数据同步的高效性和稳定性,进而提升了数据同步的效果。

[0165] 需要说明的是,在附图的流程图示出的步骤可以在诸如一组计算机可执行指令的

计算机系统中执行,并且,虽然在流程图中示出了逻辑顺序,但是在某些情况下,可以以不同于此处的顺序执行所示出或描述的步骤。

[0166] 实施例2

[0167] 本申请实施例还提供了一种数据的同步装置,需要说明的是,本申请实施例的数据的同步装置可以用于执行本申请实施例所提供的用于数据的同步方法。以下对本申请实施例提供的数据的同步装置进行介绍。

[0168] 图7是根据本申请实施例的数据的同步装置的示意图。如图7所示,该装置应用在数据同步系统中,数据同步系统中至少包括:多个第一中转节点和第二中转节点,第一中转节点用于处理数据,第二中转节点用于调度多个第一中转节点,若多个第一中转节点中存在故障节点,则采用第一中转节点中除故障节点之外的中转节点代替故障节点,若第二中转节点发生故障,则采用多个第一中转节点中的任意一个中转节点代替第二中转节点,且该装置包括:第一获取模块701、第一确定模块702和第一处理模块703。

[0169] 具体地,第一获取模块701,用于通过第二中转节点获取多个目标数据和多个第一中转节点,其中,每个目标数据为集群环境中多个源服务器待同步至目的服务器的数据;

[0170] 第一确定模块702,用于通过第二中转节点确定多个目标数据的数量和每个第一中转节点处理数据的效率;

[0171] 第一处理模块703,用于通过第二中转节点依据多个目标数据的数量和每个第一中转节点处理数据的效率,将每个目标数据分配给每个第一中转节点,以使每个第一中转节点对每个目标数据进行处理,得到处理后的多个目标数据,并将处理后的多个目标数据同步至目的服务器。

[0172] 综上,本申请实施例提供的数据的同步装置,第一获取模块701通过第二中转节点获取多个目标数据和多个第一中转节点,其中,每个目标数据为集群环境中多个源服务器待同步至目的服务器的数据;第一确定模块702通过第二中转节点确定多个目标数据的数量和每个第一中转节点处理数据的效率;第一处理模块703通过第二中转节点依据多个目标数据的数量和每个第一中转节点处理数据的效率,将每个目标数据分配给每个第一中转节点,以使每个第一中转节点对每个目标数据进行处理,得到处理后的多个目标数据,并将处理后的多个目标数据同步至目的服务器,解决了相关技术中采用集群环境中的源服务器与用于处理数据的中转节点静态绑定的方法,将集群环境中源服务器的数据同步给目的服务器,导致数据同步的效果较差的问题。通过第二中转节点依据确定的多个目标数据的数量和每个第一中转节点处理数据的效率,将每个目标数据分配给每个第一中转节点,以使每个第一中转节点对每个目标数据进行处理,得到处理后的多个目标数据,并将处理后的多个目标数据同步至目的服务器,从而提升了数据同步的高效性和稳定性,进而提升了数据同步的效果。

[0173] 可选地,在本申请实施例提供的数据的同步装置中,第一处理模块包括:第一计算单元,用于依据多个目标数据的数量和每个第一中转节点处理数据的效率,计算得到分配给每个第一中转节点的数据量;第一分配单元,用于基于分配给每个第一中转节点的数据量,将每个目标数据分配给每个第一中转节点。

[0174] 可选地,在本申请实施例提供的数据的同步装置中,该装置还包括:第二确定模块,用于在通过第二中转节点确定多个目标数据的数量和每个第一中转节点处理数据的效

率之前,确定多个目标中转节点,其中,多个目标中转节点中至少包括多个第一中转节点;第三确定模块,用于从多个目标中转节点中确定第二中转节点,以通过第二中转节点确定多个目标数据的数量和每个第一中转节点处理数据的效率,其中,第二中转节点为多个目标中转节点中除多个第一中转节点之外的中转节点。

[0175] 可选地,在本申请实施例提供的数据的同步装置中,该装置还包括:第一判断模块,用于在通过第二中转节点依据多个目标数据的数量和每个第一中转节点处理数据的效率,将每个目标数据分配给每个第一中转节点之后,判断多个第一中转节点中是否存在故障节点;第四确定模块,用于若多个第一中转节点中存在故障节点,则从多个第一中转节点中确定除故障节点之外的多个第三中转节点,其中,第三中转节点用于代替第一中转节点对多个目标数据进行处理;第二获取模块,用于获取每个第三中转节点处理数据的效率;第一分配模块,用于依据多个目标数据的数量和每个第三中转节点处理数据的效率,将每个目标数据分配给每个第三中转节点。

[0176] 可选地,在本申请实施例提供的数据的同步装置中,该装置还包括:第二判断模块,用于在从多个目标中转节点中确定第二中转节点之后,判断第二中转节点是否发生故障;第五确定模块,用于若第二中转节点发生故障,则从多个第一中转节点中确定第四中转节点,其中,第四中转节点用于代替第二中转节点调度多个第一中转节点中除第四中转节点之外的中转节点。

[0177] 可选地,在本申请实施例提供的数据的同步装置中,第一处理模块包括:第一判断单元,用于判断多个目标数据的数据量是否大于预设数据量;第二分配单元,用于若多个目标数据的数据量大于预设数据量,则通过第二中转节点依据多个目标数据的数量和每个第一中转节点处理数据的效率,将每个目标数据分配给每个第一中转节点,以使第一中转节点从多个源服务器中获取多个目标数据,并对多个目标数据进行处理;第一获取单元,用于若多个目标数据的数据量不大于预设数据量,则通过第二中转节点从多个源服务器中获取多个目标数据;第三分配单元,用于依据多个目标数据的数量和每个第一中转节点处理数据的效率,将从多个源服务器中获取的每个目标数据分配给每个第一中转节点。

[0178] 可选地,在本申请实施例提供的数据的同步装置中,第三确定模块包括:第二获取单元,用于获取每个目标中转节点对应的编号信息;第一确定单元,用于基于每个目标中转节点对应的编号信息,确定多个目标中转节点中编号最小的中转节点;第二确定单元,用于将多个目标中转节点中编号最小的中转节点作为第二中转节点。

[0179] 可选地,在本申请实施例提供的数据的同步装置中,该装置还包括:第三获取模块,用于获取每个源服务器在单位时间内待同步至目的服务器的数据量;第一计算模块,用于依据每个源服务器在单位时间内待同步至目的服务器的数据量,计算得到每个第一中转节点的平均排队长度,其中,平均排队长度用于表示在第一时间段内多个目标数据的堆积情况;第六确定模块,用于确定每个第一中转节点的实时排队长度,其中,实时排队长度用于表示在第二时间段内每个第一中转节点的负载压力情况,第二时间段的长度小于第一时间段的长度;第一发送模块,用于基于每个第一中转节点的平均排队长度和实时排队长度,向目标对象发送提醒信息,其中,提醒信息用于提醒目标对象调整将多个目标数据同步至目的服务器的策略。

[0180] 可选地,在本申请实施例提供的数据的同步装置中,该装置还包括:第三判断模

块,用于在获取多个目标数据和多个第一中转节点之后,判断多个第一中转节点中的第五中转节点是否处于空闲状态;第二分配模块,用于若多个第一中转节点中的第五中转节点处于空闲状态,则将多个目标数据中的第一数据分配给第五中转节点,以使第五中转节点对第一数据进行处理,得到处理后的第一数据,并将处理后的第一数据同步至目的服务器;第四判断模块,用于若多个第一中转节点中的第五中转节点不处于空闲状态,则判断多个第一中转节点中的第六中转节点是否处于空闲状态;第三分配模块,用于若多个第一中转节点中的第六中转节点处于空闲状态,则将多个目标数据中的第一数据分配给第六中转节点,以使第六中转节点对第一数据进行处理,得到处理后的第一数据,并将处理后的第一数据同步至目的服务器。

[0181] 数据的同步装置包括处理器和存储器,上述第一获取模块701、第一确定模块702和第一处理模块703等均作为程序单元存储在存储器中,由处理器执行存储在存储器中的上述程序单元来实现相应的功能。

[0182] 处理器中包含内核,由内核去存储器中调取相应的程序单元。内核可以设置一个或以上,通过调整内核参数来提升数据同步的效果。

[0183] 存储器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),存储器包括至少一个存储芯片。

[0184] 本发明实施例提供了一种计算机可读存储介质,其上存储有程序,该程序被处理器执行时实现所述数据的同步方法。

[0185] 本发明实施例提供了一种处理器,所述处理器用于运行程序,其中,所述程序运行时执行所述数据的同步方法。

[0186] 本发明实施例提供了一种电子设备,设备包括处理器、存储器及存储在存储器上并可在处理器上运行的程序,处理器执行程序时实现以下步骤:通过第二中转节点获取多个目标数据和所述多个第一中转节点,其中,每个目标数据为集群环境中多个源服务器待同步至目的服务器的数据;通过所述第二中转节点确定所述多个目标数据的数量和每个第一中转节点处理数据的效率;通过所述第二中转节点依据所述多个目标数据的数量和每个第一中转节点处理数据的效率,将每个目标数据分配给每个第一中转节点,以使每个第一中转节点对每个目标数据进行处理,得到处理后的多个目标数据,并将所述处理后的多个目标数据同步至所述目的服务器。

[0187] 处理器执行程序时还实现以下步骤:通过所述第二中转节点依据所述多个目标数据的数量和每个第一中转节点处理数据的效率,将每个目标数据分配给每个第一中转节点包括:依据所述多个目标数据的数量和每个第一中转节点处理数据的效率,计算得到分配给每个第一中转节点的数据量;基于分配给每个第一中转节点的数据量,将每个目标数据分配给每个第一中转节点。

[0188] 处理器执行程序时还实现以下步骤:在通过所述第二中转节点确定所述多个目标数据的数量和每个第一中转节点处理数据的效率之前,所述方法还包括:确定多个目标中转节点,其中,所述多个目标中转节点中至少包括所述多个第一中转节点;从所述多个目标中转节点中确定所述第二中转节点,以通过所述第二中转节点确定所述多个目标数据的数量和每个第一中转节点处理数据的效率,其中,所述第二中转节点为所述多个目标中转节

点中除所述多个第一中转节点之外的中转节点。

[0189] 处理器执行程序时还实现以下步骤:在通过所述第二中转节点依据所述多个目标数据的数量和每个第一中转节点处理数据的效率,将每个目标数据分配给每个第一中转节点之后,所述方法还包括:判断所述多个第一中转节点中是否存在故障节点;若所述多个第一中转节点中存在故障节点,则从所述多个第一中转节点中确定除故障节点之外的多个第三中转节点,其中,所述第三中转节点用于代替所述第一中转节点对所述多个目标数据进行处理;获取每个第三中转节点处理数据的效率;依据所述多个目标数据的数量和每个第三中转节点处理数据的效率,将每个目标数据分配给每个第三中转节点。

[0190] 处理器执行程序时还实现以下步骤:在从所述多个目标中转节点中确定所述第二中转节点之后,所述方法还包括:判断所述第二中转节点是否发生故障;若所述第二中转节点发生故障,则从所述多个第一中转节点中确定第四中转节点,其中,所述第四中转节点用于代替所述第二中转节点调度所述多个第一中转节点中除所述第四中转节点之外的中转节点。

[0191] 处理器执行程序时还实现以下步骤:通过所述第二中转节点依据所述多个目标数据的数量和每个第一中转节点处理数据的效率,将每个目标数据分配给每个第一中转节点包括:判断所述多个目标数据的数据量是否大于预设数据量;若所述多个目标数据的数据量大于所述预设数据量,则通过所述第二中转节点依据所述多个目标数据的数量和每个第一中转节点处理数据的效率,将每个目标数据分配给每个第一中转节点,以使所述第一中转节点从所述多个源服务器中获取所述多个目标数据,并对所述多个目标数据进行处理;若所述多个目标数据的数据量不大于所述预设数据量,则通过所述第二中转节点从所述多个源服务器中获取所述多个目标数据;依据所述多个目标数据的数量和每个第一中转节点处理数据的效率,将从所述多个源服务器中获取的每个目标数据分配给每个第一中转节点。

[0192] 处理器执行程序时还实现以下步骤:从所述多个目标中转节点中确定所述第二中转节点包括:获取每个目标中转节点对应的编号信息;基于每个目标中转节点对应的编号信息,确定所述多个目标中转节点中编号最小的中转节点;将所述多个目标中转节点中编号最小的中转节点作为所述第二中转节点。

[0193] 处理器执行程序时还实现以下步骤:所述方法还包括:获取每个源服务器在单位时间内待同步至目的服务器的数据量;依据每个源服务器在单位时间内待同步至目的服务器的数据量,计算得到每个第一中转节点的平均排队长度,其中,所述平均排队长度用于表示在第一时间段内所述多个目标数据的堆积情况;确定每个第一中转节点的实时排队长度,其中,所述实时排队长度用于表示在第二时间段内每个第一中转节点的负载压力情况,所述第二时间段的长度小于所述第一时间段的长度;基于每个第一中转节点的平均排队长度和实时排队长度,向目标对象发送提醒信息,其中,所述提醒信息用于提醒所述目标对象调整将所述多个目标数据同步至所述目的服务器的策略。

[0194] 处理器执行程序时还实现以下步骤:在获取多个目标数据和多个第一中转节点之后,所述方法还包括:判断所述多个第一中转节点中的第五中转节点是否处于空闲状态;若所述多个第一中转节点中的第五中转节点处于空闲状态,则将多个目标数据中的第一数据分配给所述第五中转节点,以使所述第五中转节点对所述第一数据进行处理,得到处理后

的第一数据,并将所述处理后的第一数据同步至所述目的服务器;若所述多个第一中转节点中的第五中转节点不处于空闲状态,则判断所述多个第一中转节点中的第六中转节点是否处于空闲状态;若所述多个第一中转节点中的第六中转节点处于空闲状态,则将多个目标数据中的第一数据分配给所述第六中转节点,以使所述第六中转节点对所述第一数据进行处理,得到处理后的第一数据,并将所述处理后的第一数据同步至所述目的服务器。

[0195] 本文中的设备可以是服务器、PC、PAD、手机等。

[0196] 本领域内的技术人员应明白,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。

[0197] 本申请是参照根据本申请实施例的方法、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设备的处理器以产生一个机器,使得通过计算机或其他可编程数据处理设备的处理器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。

[0198] 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。

[0199] 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设备上,使得在计算机或其他可编程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。

[0200] 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。

[0201] 存储器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)。存储器是计算机可读介质的示例。

[0202] 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暂存电脑可读媒体(transitory media),如调制的数据信号和载波。

[0203] 还需要说明的是,术语“包括”、“包含”或者其任何其他变体意在涵盖非排他性的

包含,从而使得包括一系列要素的过程、方法、商品或者设备不仅包括那些要素,而且还包括没有明确列出的其他要素,或者是还包括为这种过程、方法、商品或者设备所固有的要素。在没有更多限制的情况下,由语句“包括一个……”限定的要素,并不排除在包括要素的过程、方法、商品或者设备中还存在另外的相同要素。

[0204] 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。

[0205] 以上仅为本申请的实施例而已,并不用于限制本申请。对于本领域技术人员来说,本申请可以有各种更改和变化。凡在本申请的精神和原理之内所作的任何修改、等同替换、改进等,均应包含在本申请的权利要求范围之内。

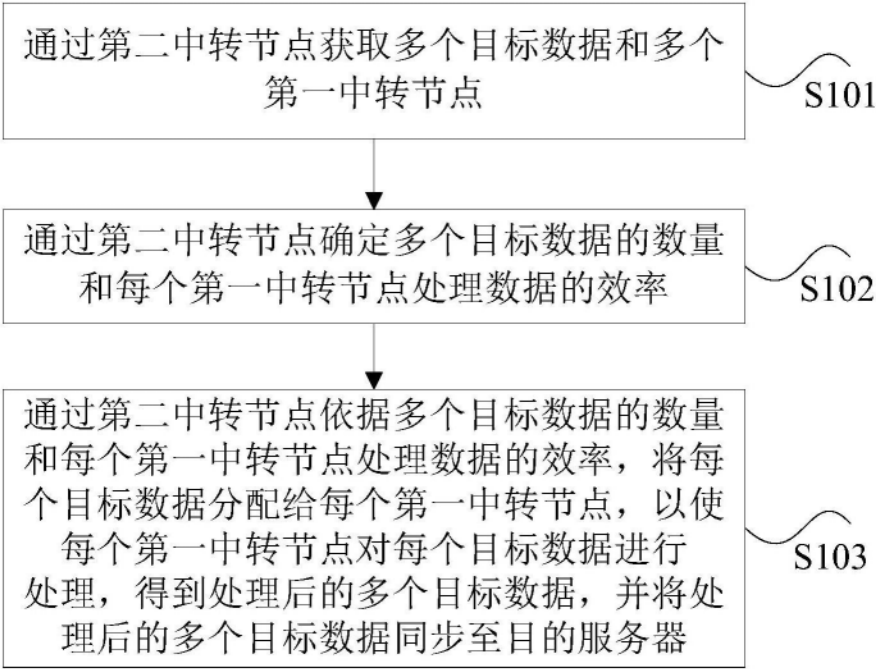


图1



图2

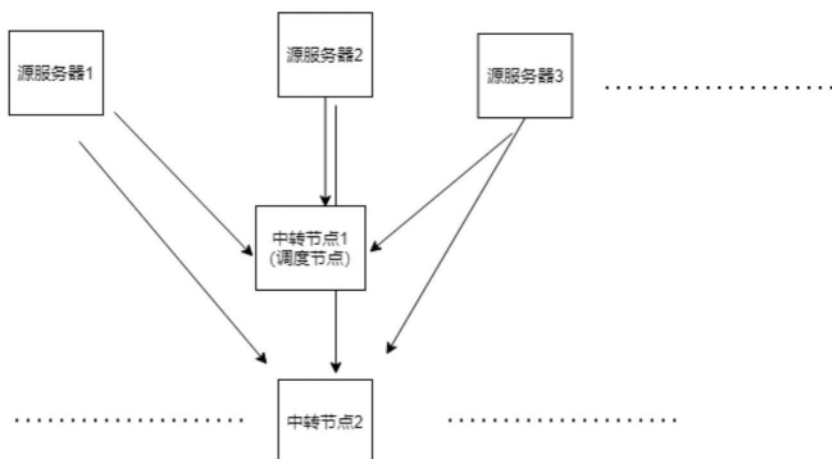


图3

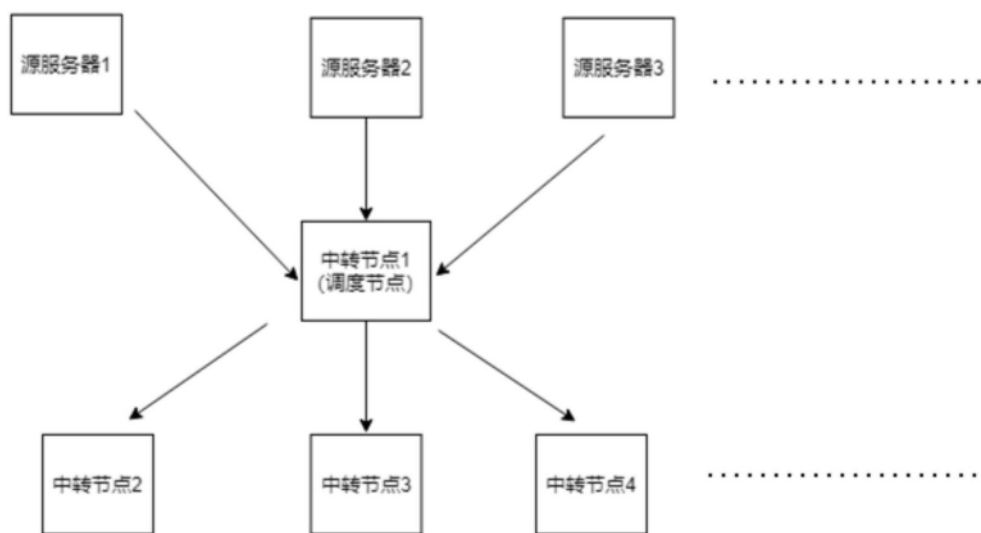


图4

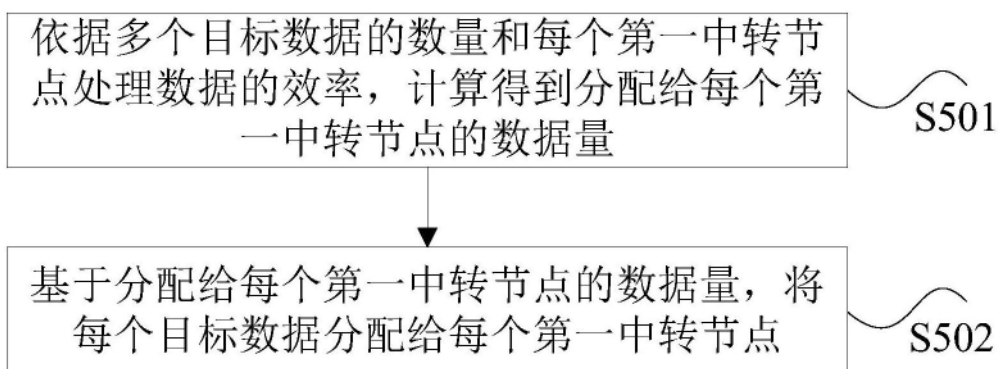


图5

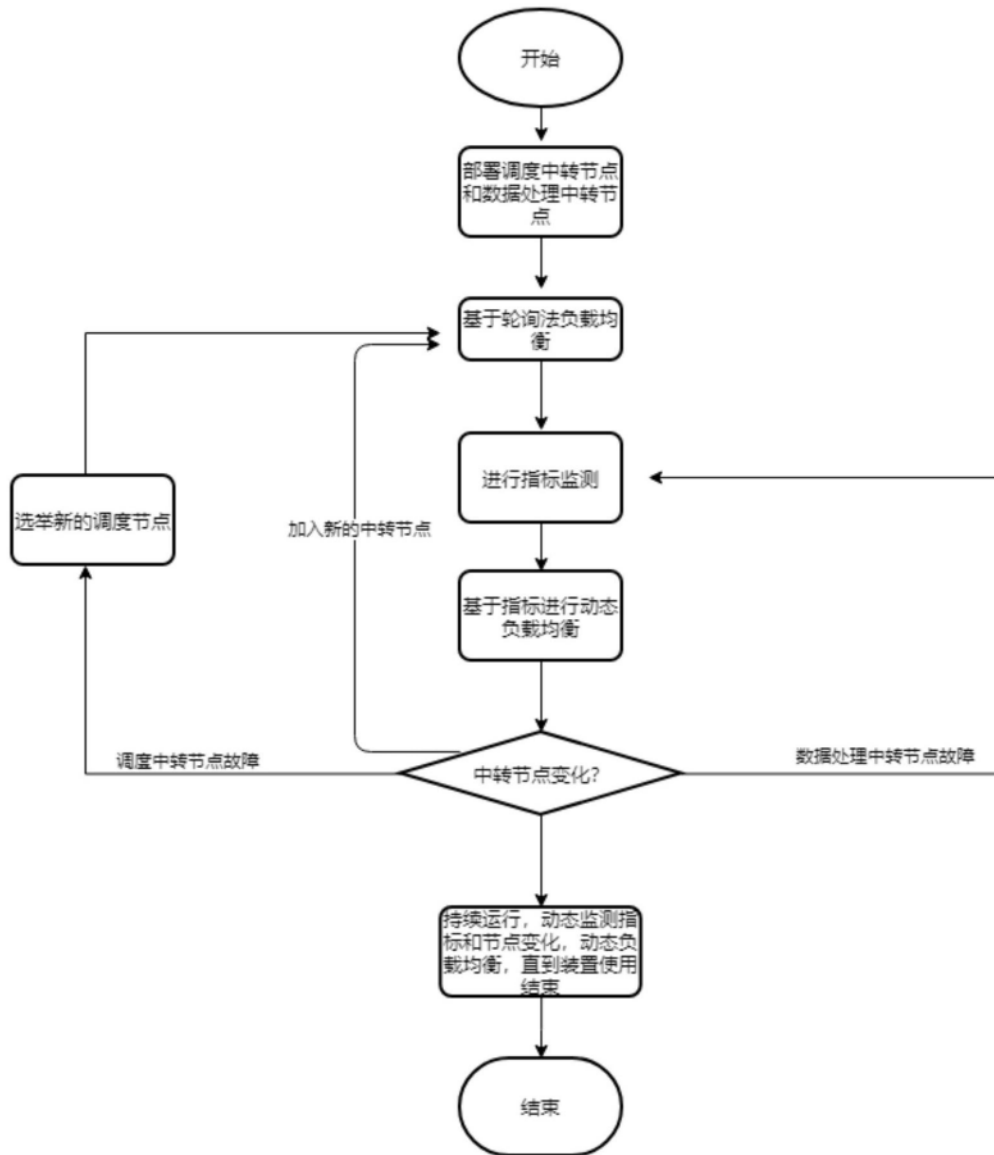


图6



图7