



(12)发明专利申请

(10)申请公布号 CN 109032809 A

(43)申请公布日 2018.12.18

(21)申请号 201810918745.4

(22)申请日 2018.08.13

(71)申请人 华东计算技术研究所(中国电子科技集团公司第三十二研究所)

地址 201800 上海市嘉定区嘉罗路1485号

(72)发明人 黄子君 谢彬 王敬平 沈晨
杜真真 许光泞 褚少鹤 周洁

(74)专利代理机构 上海段和段律师事务所
31334

代理人 李佳俊 郭国中

(51)Int.Cl.

G06F 9/50(2006.01)

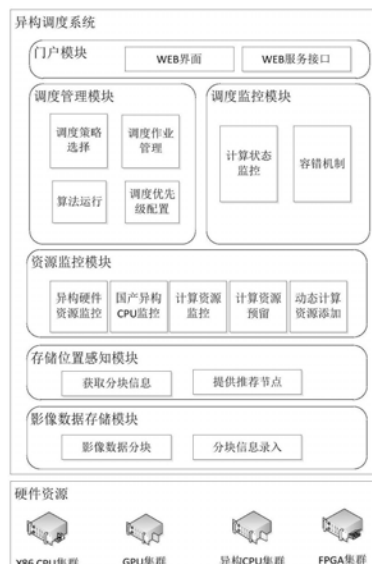
权利要求书2页 说明书8页 附图1页

(54)发明名称

基于遥感影像存储位置的异构并行调度系统

(57)摘要

本发明提供基于遥感影像存储位置的异构调度系统,影像数据存储对原始遥感影像分块处理并存储到分布式存储节点;存储位置感知采集影像信息块存储后的存放位置;资源监控对集群内各计算机硬件资源、操作系统、CPU使用情况、内存使用情况的监控;调度管理调度作业的发起操作、挂起操作、终止操作,发起操作依据调度策略选择影像处理算法的计算节点;调度监控计算节点上的调度作业的执行情况进行监控;门户提供可视化的Web界面和REST service接口。本发明实现在多机集群模式下,基于影像存储的机器节点位置的遥感算法调度,减少数据源的迁移,降低对网络IO的使用;支持在多种异构芯片环境下的并行调度处理,提高影像处理的效率。



1. 一种基于遥感影像存储位置的异构调度系统,其特征在于,包括影像数据存储模块、存储位置感知模块、资源监控模块、调度管理模块、调度监控模块、门户模块;

影像数据存储模块:对原始遥感影像进行分块处理,获得遥感影像分块信息,将携带所述遥感影像分块信息的物理分块存储到分布式存储节点上,所述物理分块记为影像信息块;

存储位置感知模块:采集分布式存储系统中,所述影像信息块存储后的存放位置的记录,获得存储位置感知结果;

资源监控模块:对集群内各计算机硬件资源、操作系统、CPU使用情况、内存使用情况的监控,以及调度备选机器节点的维护;

调度管理模块:调度作业的发起操作、挂起操作、终止操作,所述调度作业的发起操作依据调度策略选择影像处理算法的计算节点;

调度监控模块:监控计算节点上的调度作业的执行情况;

门户模块:提供可视化的Web界面,以及供外部系统使用的REST service接口。

2. 根据权利要求1所述的基于遥感影像存储位置的异构调度系统,其特征在于,所述影像数据存储模块包括影像数据分块模块、分块信息录入模块;

影像数据分块模块:对原始遥感影像的数据进行切分,得到遥感影像分块信息,同时保留遥感影像分块信息的影像源数据信息;

分块信息录入模块:对所述遥感影像分块信息进行存储,在选择存储节点时,能够优先选择有加速硬件的机器,或者计算资源充足的机器。

3. 根据权利要求1所述的基于遥感影像存储位置的异构调度系统,其特征在于,所述存储位置感知模块包括获取分块信息模块、提供推荐节点模块;

获取分块信息模块:根据原始遥感影像的标识,获取对应的影像信息块的存放位置信息,所述存放位置信息包括存储节点IP、影像信息块的标识、影像信息块的实际存放路径、机器节点信息;

提供推荐节点模块:调度节点优选,根据获取的存放位置信息中的机器节点信息,提供调度节点备选机器的集合,并根据所述集合中的机器的资源情况,筛选出适合执行影像处理任务的机器。

4. 根据权利要求1所述的基于遥感影像存储位置的异构调度系统,其特征在于,所述资源监控模块包括异构硬件资源监控模块、异构CPU监控模块、计算资源监控模块、计算资源预留模块、动态计算资源添加模块;

异构硬件资源监控模块:对异构硬件的监测,获取第一检测结果,所述第一检测结果能够用于调度管理,依据遥感影像处理算法的特性选择异构硬件实现硬件加速;

异构CPU监控模块:监测处理器基础信息,所述处理器基础信息包括CPU类型、CUP资源使用情况;

计算资源监控模块:监控计算机器的资源情况,所述资源情况包括CUP资源情况、内存使用情况;

计算资源预留模块:定义计算节点的最小资源剩余,将所述最小资源剩余作为预留空间,用于计算节点所在的机器本身运行使用;

动态计算资源添加模块:动态添加计算节点到调度备选机器节点中。

5. 根据权利要求1所述的基于遥感影像存储位置的异构调度系统,其特征在于,所述调度管理模块包括调度策略选择模块、调度作业管理模块、算法运行模块、调度优先级配置模块;

调度策略选择模块:依据调度策略选择影像处理算法的计算节点,所述调度策略包括基于存储位置优先、基于硬件情况优先、基于资源情况优先;

调度作业管理模块:调度遥感影像处理作业的发起操作、挂起操作、终止操作,所述发起操作包括选择影像信息块、获取存储位置感知结果、选择调度策略、选择计算节点、执行影像处理算法,所述挂起操作包括暂定正在运行的调度作业、记录运行阶段、持久化中间数据,所述终止操作包括终止调度作业、释放计算节点的资源;

算法运行模块:遥感影像处理算法的并行运行,记为并行运行机制;

调度优先级配置模块:根据优先等级将遥感影像处理作业分为高优先级作业、低优先级作业,所述高优先级作业拥有资源使用优先,当高优先权作业发起时,低优先权作业被挂起。

6. 根据权利要求1所述的基于遥感影像存储位置的异构调度系统,其特征在于,所述调度监控模块包括计算状态监控模块、容错机制模块;

计算状态监控模块:监控计算节点上的影像处理算法的执行情况,并记录算法执行日志信息、运行结果;

容错机制模块:调度作业运行异常时的错误处理方式,所述错误处理方式包括等待重试、重试、终止。

7. 根据权利要求1所述的基于遥感影像存储位置的异构调度系统,其特征在于,所述门户模块包括Web界面模块、Web服务接口模块;

Web界面模块:提供调度作业管理与监控的Web界面展示,以及集群资源的可视化监控;

Web服务接口模块:提供作业管理与监控的REST service接口,以及支持外部系统的调用。

8. 根据权利要求5所述的基于遥感影像存储位置的异构调度系统,其特征在于,所述并行运行机制是在获取存储位置感知结果,检测影像信息块的存储节点的计算资源是否满足遥感影像处理算法的运行条件,若满足,则将遥感影像处理算法发送到影像信息块的存储节点上,若不满足,则将影像信息块的存储数据迁移到可用的计算节点上。

9. 根据权利要求6所述的基于遥感影像存储位置的异构调度系统,其特征在于,所述容错机制模块还能够用于集群下的机器容错处理,当集群下的第一主服务器机器发生故障,则选取一台从服务器机器作为主服务器机器,当第一主服务器机器故障接触后,则将第一主服务器机器作为从服务器机器。

10. 根据权利要求1所述的基于遥感影像存储位置的异构调度系统,其特征在于,所述资源监控模块包括特殊进程占用机制,主要是检测客户端特殊进程的运行情况,对存在设定的进程运行的客户端,所述客户端所在的计算节点将被纳入服务器端的黑名单,对所述客户端已结束特殊进程运行时,所述客户端所在的计算节点将被从服务器端的黑名单中移除,所述黑名单中的计算节点避免被调度管理模块使用。

基于遥感影像存储位置的异构并行调度系统

技术领域

[0001] 本发明涉及遥感影像处理领域,具体地,涉及一种基于遥感影像存储位置的异构并行调度系统。

背景技术

[0002] 遥感影像处理工程存在数据量庞大,计算资源消耗大,计算周期长的特点,提高遥感影像处理效率是业内关注的问题。在传统的遥感影像处理软件中,采用的是集中式存储、单节点计算的架构模式。随着分布式存储和分布式计算技术的日益完善与广泛使用,出现了遥感影像处理领域的分布式计算平台。但这种计算平台存在以下问题:遥感影像数据量大,在分布式系统中进行分布式计算时,难免会因数据迁移和频繁读写,造成网络IO和磁盘IO的压力;一般分布式计算集群下,只支持CPU计算,纯CPU的计算环境不足以支持遥感影像领域中的高性能处理计算;传统的遥感影像处理中,算法将整幅影像读取到内存后再进行处理,这种形式对计算机的内存配置要求较高,而且单机运算的计算效率会存在瓶颈,无法体现集群模式下的性能优势。

发明内容

[0003] 针对现有技术中的缺陷,本发明的目的是提供一种基于遥感影像存储位置的异构并行调度系统。

[0004] 根据本发明提供的一种基于遥感影像存储位置的异构调度系统,包括影像数据存储模块、存储位置感知模块、资源监控模块、调度管理模块、调度监控模块、门户模块;

[0005] 影像数据存储模块:对原始遥感影像进行分块处理,获得遥感影像分块信息,将携带所述遥感影像分块信息的物理分块存储到分布式存储节点上,所述物理分块记为影像信息块;

[0006] 存储位置感知模块:采集分布式存储系统中,所述影像信息块存储后的存放位置的记录,获得存储位置感知结果;

[0007] 资源监控模块:对集群内各计算机硬件资源、操作系统、CPU使用情况、内存使用情况的监控,以及调度备选机器节点的维护;

[0008] 调度管理模块:调度作业的发起操作、挂起操作、终止操作,所述调度作业的发起操作依据调度策略选择影像处理算法的计算节点;

[0009] 调度监控模块:监控计算节点上的调度作业的执行情况;

[0010] 门户模块:提供可视化的Web界面,以及供外部系统使用的REST service接口。

[0011] 优选地,所述影像数据存储模块包括影像数据分块模块、分块信息录入模块;

[0012] 影像数据分块模块:对原始遥感影像的数据进行切分,得到遥感影像分块信息,同时保留遥感影像分块信息的影像源数据信息;

[0013] 分块信息录入模块:对所述遥感影像分块信息进行存储,在选择存储节点时,能够优先选择有加速硬件的机器,或者计算资源充足的机器。

- [0014] 优选地,所述存储位置感知模块包括获取分块信息模块、提供推荐节点模块;
- [0015] 获取分块信息模块:根据原始遥感影像的标识,获取对应的影像信息块的存放位置信息,所述存放位置信息包括存储节点IP、影像信息块的标识、影像信息块的实际存放路径、机器节点信息;
- [0016] 提供推荐节点模块:调度节点优选,根据获取的存放位置信息中的机器节点信息,提供调度节点备选机器的集合,并根据所述集合中的机器的资源情况,筛选出适合执行影像处理任务的机器。
- [0017] 优选地,所述资源监控模块包括异构硬件资源监控模块、异构CPU监控模块、计算资源监控模块、计算资源预留模块、动态计算资源添加模块;
- [0018] 异构硬件资源监控模块:对异构硬件的监测,获取第一检测结果,所述第一检测结果能够用于调度管理,依据遥感影像处理算法的特性选择异构硬件实现硬件加速;
- [0019] 异构CPU监控模块:监测处理器基础信息,所述处理器基础信息包括CPU类型、CUP资源使用情况;
- [0020] 计算资源监控模块:监控计算机器的资源情况,所述资源情况包括CUP资源情况、内存使用情况;
- [0021] 计算资源预留模块:定义计算节点的最小资源剩余,将所述最小资源剩余作为预留空间,用于计算节点所在的机器本身运行使用;
- [0022] 动态计算资源添加模块:动态添加计算节点到调度备选机器节点中。
- [0023] 优选地,所述调度管理模块包括调度策略选择模块、调度作业管理模块、算法运行模块、调度优先级配置模块;
- [0024] 调度策略选择模块:依据调度策略选择影像处理算法的计算节点,所述调度策略包括基于存储位置优先、基于硬件情况优先、基于资源情况优先;
- [0025] 调度作业管理模块:调度遥感影像处理作业的发起操作、挂起操作、终止操作,所述发起操作包括选择影像信息块、获取存储位置感知结果、选择调度策略、选择计算节点、执行影像处理算法,所述挂起操作包括暂定正在运行的调度作业、记录运行阶段、持久化中间数据,所述终止操作包括终止调度作业、释放计算节点的资源;
- [0026] 算法运行模块:遥感影像处理算法的并行运行,记为并行运行机制;
- [0027] 调度优先级配置模块:根据优先等级将遥感影像处理作业分为高优先级作业、低优先级作业,所述高优先级作业拥有资源使用优先,当高优先权作业发起时,低优先权作业被挂起。
- [0028] 优选地,所述调度监控模块包括计算状态监控模块、容错机制模块;
- [0029] 计算状态监控模块:监控计算节点上的影像处理算法的执行情况,并记录算法执行日志信息、运行结果;
- [0030] 容错机制模块:调度作业运行异常时的错误处理方式,所述错误处理方式包括等待重试、重试、终止。
- [0031] 优选地,所述门户模块包括Web界面模块、Web服务接口模块;
- [0032] Web界面模块:提供调度作业管理与监控的Web界面展示,以及集群资源的可视化监控;
- [0033] Web服务接口模块:提供作业管理与监控的REST service接口,以及支持外部系统

的调用。

[0034] 优选地,所述并行运行机制是在获取存储位置感知结果,检测影像信息块的存储节点的计算资源是否满足遥感影像处理算法的运行条件,若满足,则将遥感影像处理算法发送到影像信息块的存储节点上,若不满足,则将影像信息块的存储数据迁移到可用的计算节点上。

[0035] 优选地,所述容错机制模块还能够用于集群下的机器容错处理,当集群下的第一主服务器机器发生故障,则选取一台从服务器机器作为主服务器机器,当第一主服务器机器故障接触后,则将第一主服务器机器作为从服务器机器。

[0036] 优选地,所述资源监控模块包括特殊进程占用机制,主要是检测客户端特殊进程的运行情况,对存在设定的进程运行的客户端,所述客户端所在的计算节点将被纳入服务器端的黑名单,对所述客户端已结束特殊进程运行时,所述客户端所在的计算节点将被从服务器端的黑名单中移除,所述黑名单中的计算节点避免被调度管理模块使用。

[0037] 与现有技术相比,本发明具有如下的有益效果:

[0038] 1、在多机集群模式下,基于影像存储的机器节点位置的遥感算法调度,减少数据源的迁移,降低对网络IO的使用;

[0039] 2、支持在多种异构芯片环境下的并行调度处理,提高影像处理的效率;

[0040] 3、允许使用硬件加速的方式,进行遥感影像加速处理,提高处理算法的执行效率,优先使用硬件加速的策略能尽可能发挥加速硬件的性能优势;

[0041] 4、使用基于影像分块位置的方式进行处理算法的调度,能降低跨机器移动影像分块造成的网络IO压力,同时降低频发读写数据造成的磁盘IO压力;

[0042] 5、使用分块并行处理方式,能提高计算节点的使用率,提高遥感影像处理的整体效率;

[0043] 6、对客户端计算节点采用特殊进程占用机制,保障客户端机器中重要进程的成功运行。

附图说明

[0044] 通过阅读参照以下附图对非限制性实施例所作的详细描述,本发明的其它特征、目的和优点将会变得更明显:

[0045] 图1为本发明的系统架构图;

[0046] 图2为本发明的系统部署图。

具体实施方式

[0047] 下面结合具体实施例对本发明进行详细说明。以下实施例将有助于本领域的技术人员进一步理解本发明,但不以任何形式限制本发明。应当指出的是,对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,还可以做出若干变化和改进。这些都属于本发明的保护范围。

[0048] 本发明实现在多机集群模式下,基于影像存储的机器节点位置的遥感算法调度,减少数据源的迁移,降低对网络IO的使用;同时,支持在多种异构芯片环境下的并行调度处理,提高影像处理的效率。通过遥感影像存储位置的感知,主要解决传统的分布式系统中,

迁移影像元数据所造成的网络IO压力、磁盘IO压力。本系统将遥感影像数据进行物理分块处理,通过对遥感影像分块位置的感知,选择合适的计算机器,将影像处理算法调度到该计算机器并进行影像处理,最大程度减少影像数据源的跨节点迁移。通过遥感影像处理算法的异构调度,主要解决异构计算环境下遥感影像处理算法的调度与计算。本系统支持不同型号的CPU、GPU、FPGA等异构芯片的混合运行环境,根据处理算法以及影像的特性,选择合适的硬件环境进行影像处理计算,实现影像处理算法处理的硬件加速。通过遥感影像分块并行处理,主要解决单机处理巨幅影像时存在的效率与硬件评价问题。在基于遥感影像存储位置感知的调度下,将影像处理算法发送到各影像分块所在的计算节点上,利用各计算节点的计算资源,分别对影像分块进行同一算法处理。支持基于MPI、OpenMP、CUDA等高性能并行计算架构开发的处理算法的调度运行,使用高性能计算框架开发的算法要求运行环境具备相应的依赖支持,调度系统需识别各机器的运算环境,在下发处理算法时选择具有运行环境的机器进行处理。采用特殊进程占用机制,遥感影像的处理软件大多体量庞大,占用资源多,而在实际生产中如果操作人员在用某些特殊软件处理紧急任务时,此时这台计算节点的资源应该尽可能用于紧急任务,调度系统应避免使用这台计算节点进行影像的调度处理。本发明建立特殊进程占用机制,由各客户端监听特殊进程的运行情况,对于存在特殊进程的机器节点,服务端将该节点纳入黑名单,并尽可能减少对该节点的使用;当客户端的特殊进程结束后,该节点从黑名单中移除。特殊进程占用机制能保障客户端重要进程的成功运行。

[0049] 为了解决遥感影像处理的效率问题,本发明采用基于存储位置感知的调度技术以及异构硬件加速技术。在数据源录入系统进行分布式存储前,先对遥感影像数据源进行分块处理,将GB级别的影像数据进行分块切割,同时保留影像数据的元数据信息,保证每个分块可作为一个独立的影像数据源使用;完成切分后,将分块数据源存储到分布式文件系统中。系统采用存算一体的结构,客户端既是计算节点也是存储节点。各客户端需采集本机器的硬件信息、CPU使用情况、内存使用情况、网络信息、操作系统信息等,并推送至服务端进行汇总监控。系统中集成多种常用的遥感影像处理算法,根据算法的功能特点进行归类,如适用于GPU的算法、适用于FPGA的算法等。

[0050] 以下对本发明做进一步的阐述。

[0051] 根据本发明公开的一种基于遥感影像存储位置的异构调度系统,包括影像数据存储模块、存储位置感知模块、资源监控模块、调度管理模块、调度监控模块、门户模块;

[0052] 影像数据存储模块用于对原始遥感影像进行分块处理,获得遥感影像分块信息,将携带所述遥感影像分块信息的物理分块存储到分布式存储节点上,所述物理分块记为影像信息块;

[0053] 存储位置感知模块用于采集分布式存储系统中,所述影像信息块存储后的存放位置的记录,获得存储位置感知结果;

[0054] 资源监控模块用于对集群内各计算机硬件资源、操作系统、CPU使用情况、内存使用情况的监控,以及调度备选机器节点的维护;

[0055] 调度管理模块用于调度作业的发起操作、挂起操作、终止操作,所述调度作业的发起操作依据调度策略选择影像处理算法的计算节点;

[0056] 调度监控模块用于监控计算节点上的调度作业的执行情况;

[0057] 门户模块用于提供可视化的Web界面,以及供外部系统使用的REST service接口。

[0058] 具体地,所述影像数据存储模块包括影像数据分块模块、分块信息录入模块;影像数据存储模块主要用于对原始遥感影像进行分块处理,并将物理分块存储到分布式存储节点上;影像数据分块模块用于对原始遥感影像的数据进行切分,得到遥感影像分块信息,同时保留遥感影像分块信息的影像源数据信息,将遥感影像的图像数据进行切分,同时保留各分块的影像源数据信息;分块信息录入模块用于对所述遥感影像分块信息进行存储,在选择存储节点时,能够优先选择有加速硬件的机器,或者计算资源充足的机器。在选择存储节点的时候,根据数据的不同特性,优先选择有加速硬件的机器,或是计算资源充足的机器。

[0059] 具体地,所述存储位置感知模块包括获取分块信息模块、提供推荐节点模块;存储位置感知模块主要用于采集分布式存储系统中,数据分块导入后的存放位置的记录,数据位置感知的结果作为调度选择计算机器的参考。感知模块采用客户端主动推送的形式获取。客户端在计算机接入调度系统时,向服务器端上报当前计算节点的硬件信息、资源信息,并在接收到影像分块后,将分块的存储信息上报到服务器端;服务器端提供一个守护进程,监听并收集客户端推送的硬件信息、资源使用信息、影像分块存储信息。这种主动上报的模式,提高了集群模式下数据采集的灵活性。

[0060] 获取分块信息模块用于根据原始遥感影像的标识,获取对应的影像信息块的存放位置信息,所述存放位置信息包括存储节点IP、影像信息块的标识、影像信息块的实际存放路径、机器节点信息;获取分块信息模块在分块位置推送汇总时,根据处理源影像的标识,获取集群中该影像分块的信息,包括分块对应的机器节点IP,分块数据的标识,分块数据实际存放的路径,各计算节点的客户端将数据推送到服务器端,服务器端更新客户端推送数据。

[0061] 提供推荐节点模块用于调度节点优选,根据获取的存放位置信息中的机器节点信息,提供调度节点备选机器的集合,并根据所述集合中的机器的资源情况,筛选出适合执行影像处理任务的机器;在调度节点优选时,根据获取到的影像源分块的机器节点信息,提供调度节点备选机器集合,并根据当前这些机器的资源情况,筛选出适合执行影像处理任务的机器。

[0062] 具体地,所述资源监控模块包括异构硬件资源监控模块、国产异构CPU监控模块、计算资源监控模块、计算资源预留模块、动态计算资源添加模块;资源监控模块主要实现对集群内各计算机硬件资源、操作系统、CPU使用情况、内存使用情况的监控。

[0063] 异构硬件资源监控模块用于对异构硬件的监测,获取第一检测结果,所述第一检测结果能够用于调度管理,依据遥感影像处理算法的特性选择异构硬件实现硬件加速;支持对GPU、FPGA等异构硬件的监测,可针对不同的遥感影像处理算法,选择合适的加速硬件。

[0064] 国产异构CPU监控模块用于监测处理器基础信息,所述处理器基础信息包括CPU类型、CUP资源使用情况;支持对国产CPU的调度,监测包括CPU类型、CPU资源使用情况在内的处理器基础信息,国产CPU包括:申威、飞腾、龙芯。

[0065] 计算资源监控模块用于监控计算机器的资源情况,所述资源情况包括CUP资源情况、内存使用情况;支持对传统X86CPU资源情况、内存使用情况的监测,作为选择计算节点时的参考。

[0066] 计算资源预留模块用于定义计算节点的最小资源剩余,将所述最小资源剩余作为预留空间,用于计算节点所在的机器本身运行使用;支持计算资源预留,定义各计算节点最小的资源剩余,避免由于调度作业占用过多资源,影响计算机原本的使用。

[0067] 动态计算资源添加模块用于动态添加计算节点到调度备选机器节点中;支持动态添加计算资源,在热部署环境下,动态添加计算机节点到调度备选机器节点中。

[0068] 具体地,所述调度管理模块包括调度策略选择模块、调度作业管理模块、算法运行模块、调度优先级配置模块;调度管理模块主要用于对调度作业进行发起、挂起、终止等操作,其中在作业发起时需要指定相关的调度策略,选择合适的影像处理算法执行节点。

[0069] 调度策略选择模块用于依据调度策略选择影像处理算法的计算节点,所述调度策略包括基于存储位置优先、基于硬件情况优先、基于资源情况优先;调度策略选择,调度作业可基于不同的策略进行节点选择:基于存储位置优先,选择分块最多的若干台机器进行计算,但选择的机器必须具有进行影像处理的最小资源剩余;基于硬件情况优先,选择具有加速硬件的计算机进行处理算法的计算;基于资源情况优先,选择具有充足CPU剩余资源、内存剩余资源的机器进行处理算法的计算。

[0070] 调度作业管理模块用于调度遥感影像处理作业的发起操作、挂起操作、终止操作,调度作业可通过接口、界面等方式进行发起、挂起、终止等操作。所述发起操作包括选择影像信息块、获取存储位置感知结果、选择调度策略、选择计算节点、执行影像处理算法,发起操作选择需要处理的遥感影像数据源,获取数据源位置信息,并根据选择的调度策略,优选计算节点进行影像处理算法的执行;所述挂起操作包括暂定正在运行的调度作业、并记录运行阶段、持久化中间数据;所述终止操作包括终止调度作业、释放计算节点的资源;

[0071] 算法运行模块用于遥感影像处理算法的并行运行,记为并行运行机制;算法并行运行,不同的遥感影像处理算法会对应不同格式的输入输出数据,其运行参数、运行命令也各有不一,系统支持不同处理算法在不同操作系统、不同硬件环境下的运行;系统获取到各影像分块的位置后,检测分块所在存储节点的计算资源是否满足处理算法的运行条件,将算法及相关依赖文件发送到满足资源要求的节点,并将不满足要求的节点上的数据迁移到可用的计算机上。各计算节点上的算法可分别对本节点的影像数据分块进行并行运算;

[0072] 调度优先级配置模块用于根据优先等级将遥感影像处理作业分为高优先级作业、低优先级作业,所述高优先级作业拥有资源使用优先,当高优先权作业发起时,低优先权作业被挂起;调度优先级配置,根据遥感影像处理作业的优先级等级,判断当前作业是否应该抢占计算资源进行优先处理。对于被抢占资源的作业,应挂起当前作业并持久化当前作业状态,在优先作业完成后,重新恢复被挂起的作业。

[0073] 具体地,所述调度监控模块包括计算状态监控模块、容错机制模块,主要用于对各计算节点上的调度作业执行情况的监控;

[0074] 计算状态监控模块用于监控计算节点上的影像处理算法的执行情况,并输出记录算法执行日志信息、运行结果;

[0075] 容错机制模块用于调度作业运行异常时的错误处理方式,所述错误处理方式包括等待重试、重试、终止。作业发起并调度到计算节点后,计算节点客户端将向服务端返回调度确认标记;当服务器端经过一定时间仍未接收到确认标记,则认为该作业启动失败,抛出客户端异常。当影像处理算法抛出异常后,系统应返回终端错误信息,并提供等待重试、重

试、终止等错误处理方式。其中,等待重试是计算节点隔一定时间后,重新执行影像处理的计算,重试若干次后仍失败,则终止该调度作业的任务;重试是重新选择计算运行的节点,并执行影像处理计算;终止是终止当前运行的影像处理计算。

[0076] 具体地,所述门户模块包括Web界面模块、Web服务接口模块,提供可视化的web界面以及供外部系统使用的REST service接口;

[0077] Web界面模块用于提供调度作业管理与监控的Web界面展示,以及集群资源的可视化监控;

[0078] Web服务接口模块用于提供作业管理与监控的REST service接口,以及支持外部系统的调用。

[0079] 具体地,所述并行运行机制是在获取存储位置感知结果,检测影像信息块的存储节点的计算资源是否满足遥感影像处理算法的运行条件,若满足,则将遥感影像处理算法发送到影像信息块的存储节点上,若不满足,则将影像信息块的存储数据迁移到可用的计算节点上。

[0080] 具体地,所述容错机制模块还能够用于集群下的机器容错处理,当集群下的第一主服务器机器发生故障,则选取一台从服务器机器作为主服务器机器,当第一主服务器机器故障接触后,则将第一主服务器机器作为从服务器机器。服务器端的管理节点采用集群模式部署。其中一台机器作为主服务器,其他为从服务器。当主服务器发生故障,将在从服务器中选取一台作为新的主服务器;原服务器功能恢复后,可作为从服务器加入集群。

[0081] 具体地,所述资源监控模块包括特殊进程占用机制,主要是检测客户端特殊进程的运行情况,对存在设定的进程运行的客户端,所述客户端所在的计算节点将被纳入服务器端的黑名单,对所述客户端已结束特殊进程运行时,所述客户端所在的计算节点将被从服务器端的黑名单中移除,所述黑名单中的计算节点避免被调度管理模块使用,优选地,在黑名单中的计算节点将会被尽可能减少被调度系统分配影像处理任务。

[0082] 在本发明的实际应用中,采用在存算一体的计算机中,安装系统的监控客户端,用于采集机器的硬件资源、计算资源情况、操作系统信息、调度作业进程信息、调度作业结果信息等。客户端与服务器端进行通信,服务器收集各计算节点的资源情况,在发起调度任务时,根据收集的资源情况进行算法推送调度。当存储节点的计算资源不足以支撑处理算法的运行时,系统将该存储节点中的数据迁移到资源充足的计算节点中。在进行影像数据源分块后,服务器优先选择资源充足、具有加速硬件的计算节点作为数据存储的载体。在推送分块影像的同时,记录各分块的存储位置、存储节点IP、分块标识、原数据源的标识等信息,用于对数据源分块的查找。通过浏览器访问调度系统WEB界面,进行调度作业的发起与监控。调度引擎根据作业的输入数据、机器节点资源情况,选择合适的计算节点,并将要执行的处理算法推送到计算节点进行计算。

[0083] 本领域技术人员知道,除了以纯计算机可读程序代码方式实现本发明提供的系统、装置及其各个模块以外,完全可以通过将方法步骤进行逻辑编程来使得本发明提供的系统、装置及其各个模块以逻辑门、开关、专用集成电路、可编程逻辑控制器以及嵌入式微控制器等的形式来实现相同程序。所以,本发明提供的系统、装置及其各个模块可以被认为是一种硬件部件,而对其内包括的用于实现各种程序的模块也可以视为硬件部件内的结构;也可以将用于实现各种功能的模块视为既可以是实现方法的软件程序又可以是硬件部

件内的结构。

[0084] 以上对本发明的具体实施例进行了描述。需要理解的是,本发明并不局限于上述特定实施方式,本领域技术人员可以在权利要求的范围内做出各种变化或修改,这并不影响本发明的实质内容。在不冲突的情况下,本申请的实施例和实施例中的特征可以任意相互组合。

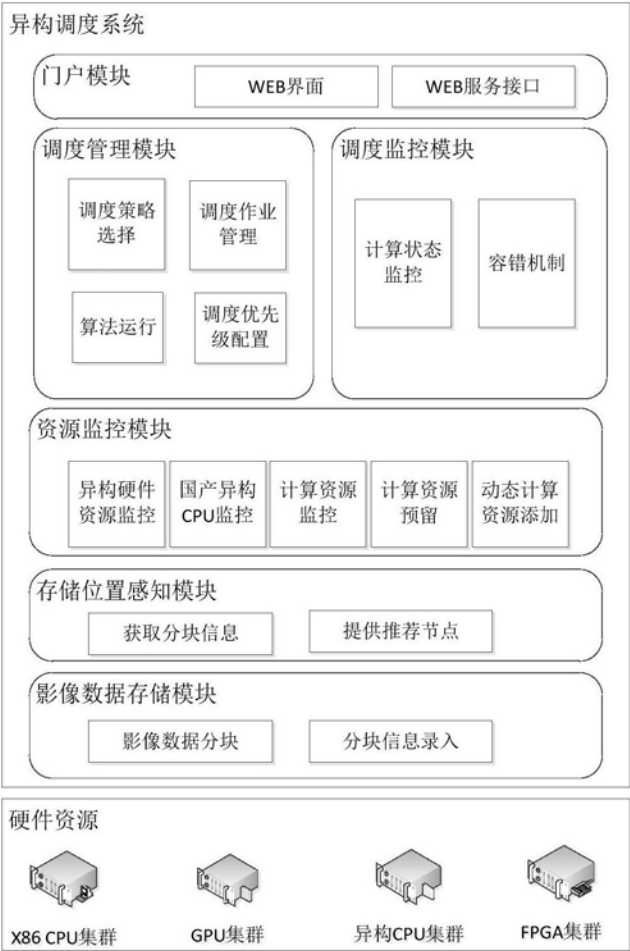


图1

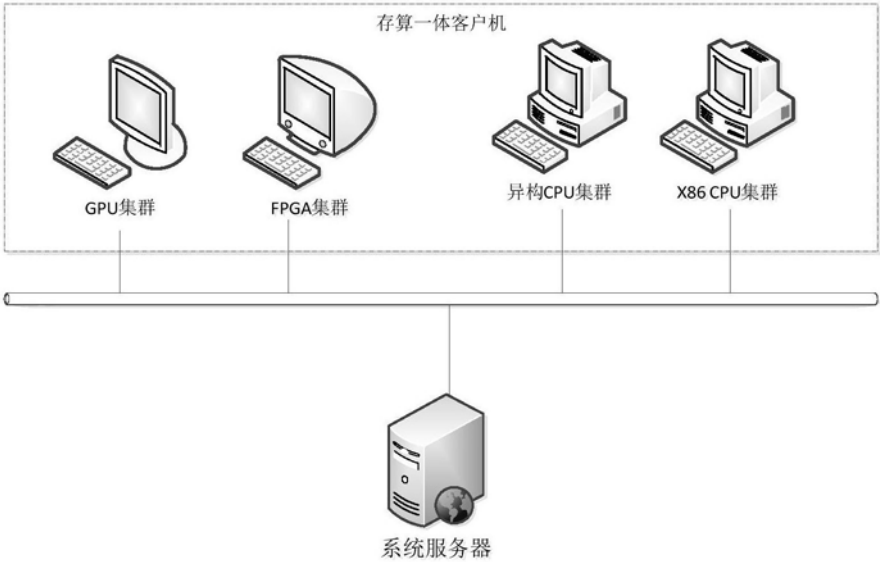


图2