

霍庆, 何文春, 高峰, 等. 基于数据感知的气象算法调度框架设计与应用. 应用气象学报, 2024, 35(4): 502-512.
DOI: 10.11898/1001-7313.20240410

基于数据感知的气象算法调度框架设计与应用

霍庆 何文春* 高峰 陈士旺 徐拥军 王琦

(国家气象信息中心, 北京 100081)

摘要

气象产品加工任务往往由不同算法组合而成,其状态对数据源具有高依赖性。高效感知数据状态,并及时合理启动调度,对提升气象产品加工的时效性至关重要。设计并建设了基于数据感知技术的气象算法调度框架,对比测试结果分析表明:在实时分析数据源到报状态的基础上,数据感知调度实现数据到报即算法启动运行,任务调度延迟较定时调度显著降低,由 3784 ms 降至 11 ms。数据感知调度作为气象大数据云平台加工流水线的核心能力,已在国省部署并实时运行。目前在国家级业务中支撑 19 种核心业务算法的高效调度,日调度次数约为 6.67×10^5 ,平均调度延迟为 31 ms;在省级业务中支撑 14 种算法的高效调度,日调度次数约为 8×10^4 ,平均调度延迟为 156 ms。此外,数据感知调度实现了气象业务上下游算法的无缝串联,可为解决气象业务间的脱节问题、提升气象业务间的协同性提供解决方案。

关键词: 数据感知技术; 任务调度; 气象算法; 气象大数据云平台; 加工流水线

引言

随着气象业务的发展,用户对算法运行效率及其产品生成时效提出了更高要求。产品生成时效既依赖于算法的计算效率,也取决于算法的启动时间。在计算效率不变情况下,若能降低算法启动延迟,便可降低产品生成时间。算法启动延迟依赖于任务调度能力,实际应用中应使算法在启动条件满足后的第一时间启动运行。目前,针对任务调度的研究比较丰富,任务调度按任务类型可分为独立任务调度和 workflow 调度,按调度算法特点可分为启发式、元启发式和混合式任务调度等^[1-4]。在气象业务中有两类典型的任务调度实现方式:一是运行在高性能计算机上的任务,其任务调度主要基于作业调度系统 (Slurm 或 LoadLeveler) 和流程管理与监控软件 (ECFlow) 实现^[5-10],任务以定时方式提交;二是运行在普通计算机上的任务,其任务调度常以定时任

务调度为主,部分任务以常驻内存的方式运行^[11-19]。

气象业务中的数据产品加工算法一般为数据驱动,即数据到报后立即启动算法,以加快数据产品的生成。因此,如何便捷、高效地获取数据到报状态,是实现数据驱动的核心。定时调度通过扫描数据文件目录或数据库方式获取数据到报状态,因存在扫描间隙或因数据文件较多导致较高的扫描耗时,从而无法及时发现数据,降低算法启动时间。尤其在气象观测台站密度和数据观测频率显著提升后,数据量也显著增加,逐小时天气雷达基数据记录数由 500 条增长到超 30000 条,针对海量气象数据的加工处理,目录扫描的高延迟性凸显。算法常驻内存的运行方式可通过监听数据消息,及时获取数据到报状态,但长期占用内存造成资源浪费。因此,以数据驱动为主的气象业务亟需一种高效的任务调度框架,实现数据到报即算法启动运行的目标,提升气象数据产品的生成时效。

中国气象局国家气象信息中心牵头建设的气象

2024-02-26 收到, 2024-05-30 收到再改稿。

资助项目: 中国气象局重点创新团队 (CMA2024ZD05), 中国国家铁路集团有限公司科研计划课题 (N2022T011), 国家气象信息中心创新团队 (NMIC-2024-ZD11)

* 通信作者, 邮箱: hewc@cma.gov.cn

大数据云平台(简称天擎)于 2021 年 12 月开始全国业务运行。其中,加工流水线作为天擎的核心功能实现气象算法的统一管理与集约调度运行^[20]。加工流水线建立了多种任务调度能力,包括定时调度、顺序调度、数据感知调度、人工调度等。其中,基于数据到报状态的数据感知调度实现了数据到报后算法立即启动,极大降低了气象算法的启动延迟和气象数据产品的生成时间。本文详细介绍了数据感知调度的系统架构、功能设计、关键技术等,并对调度性能和业务应用效益进行了分析。

1 系统设计

1.1 功能设计

数据感知调度的核心功能包括数据状态感知组件、任务调度执行组件、任务调度后处理组件及配置管理等。其中,数据状态感知组件实现天擎中各气象数据到报状态的实时分析,并当数据到报率达到调度要求时,发送调度消息至任务调度执行组件。任务调度执行组件结合算法所需资源信息、计算节

点资源信息进行计算节点的优选后,生成任务调度指令,并实现算法启动执行。任务调度后处理组件实现算法执行状态、计算节点状态的采集与更新及算法执行异常的告警信息发送。配置管理支持用户在前端页面进行数据感知调度相关参数的配置。

1.2 系统架构设计

数据感知调度由调度服务端、调度客户端和缓存中间件组成,并基于天擎消息总线进行消息通信。调度服务端基于气象业务综合监控系统(简称天镜)的数据详细监视信息(detail information, DI)^[21-22]实现数据到报状态实时分析,在数据到报状态满足算法启动条件时,及时生成调度指令并下发至调度客户端。其中,DI 标识数据的到报状态和完整性。调度客户端部署在加工流水线的各计算节点上,通过接收调度指令实现算法的启动运行,同时采集并生成算法执行状态消息、计算节点状态消息等,反馈至调度服务端。数据感知调度的配置信息及调度所需的算法信息、任务信息等通过缓存技术,减少调度服务端与数据库交互次数,提升调度时效。数据感知调度的系统架构见图 1。

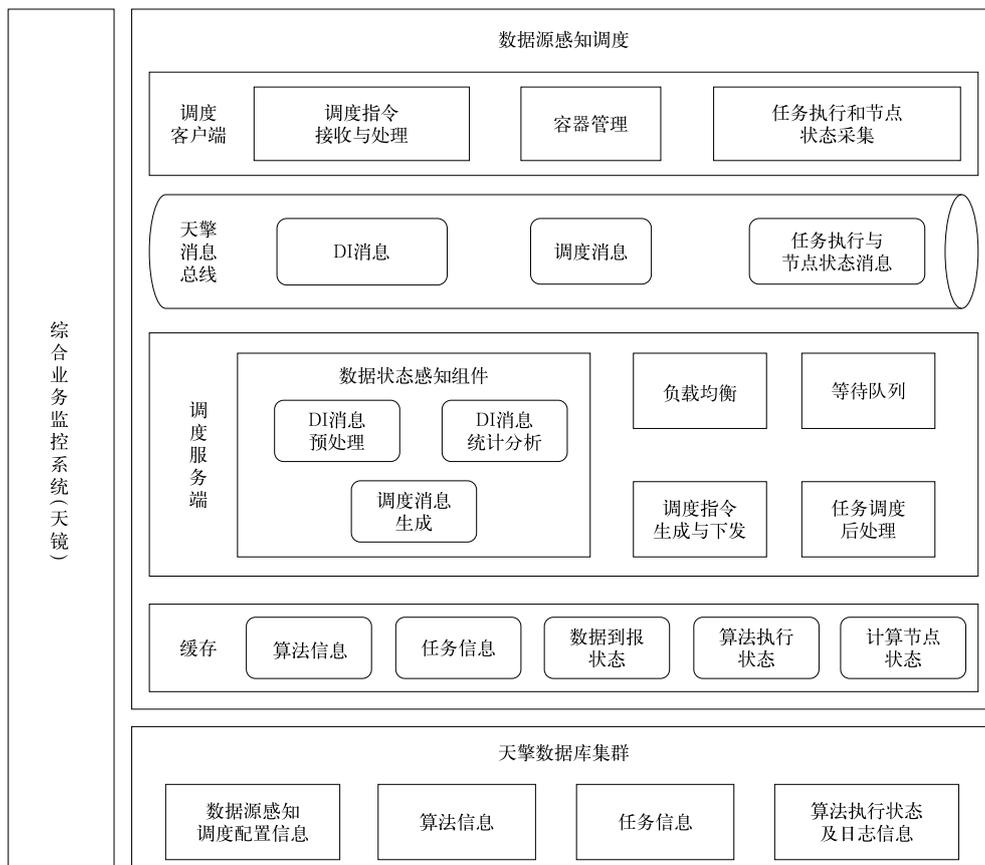


图 1 数据感知调度的系统架构
Fig. 1 Framework of data-arrival scheduling

1.3 业务流程设计

天擎加工流水线完成任务定义并启动任务后,数据状态感知组件开始从天擎消息总线中实时获取 DI 消息,并进行数据到报状态统计,当数据到报或到报率达到一定阈值(如单个雷达基数据文件到报或雷达基数据到报站数达 95% 等)后生成调度消息。调度消息经负载均衡选定计算节点后,生成任务调度指令并下发至调度客户端,客户端针对该计

算节点再次进行计算资源判断。当计算资源充足时,通过调用容器管理平台 Kubernetes(K8s)相关组件完成算法启动运行;当资源不足时,调度消息进行调度重试。调度客户端实时采集算法运行状态和计算节点状态,调度服务端基于状态消息进行算法运行状态更新、计算节点状态更新、监控信息发送等调度后处理。数据状态感知调度的业务流程见图 2。

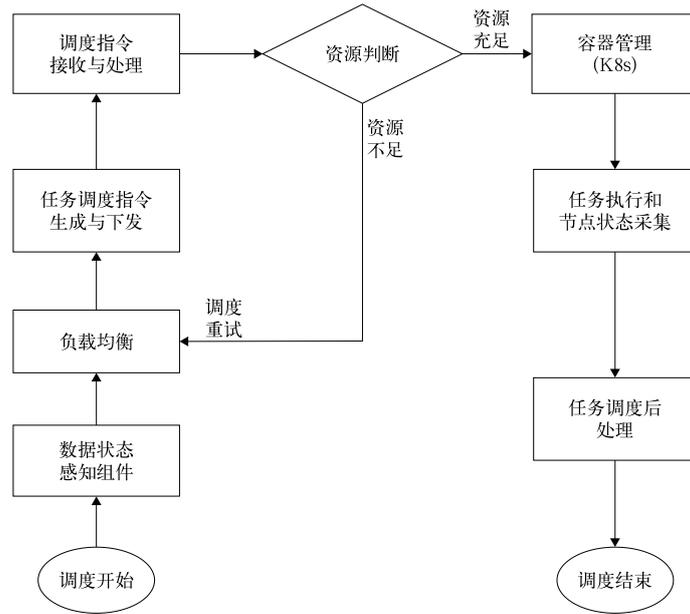


图 2 数据感知调度的业务处理逻辑

Fig. 2 Flowchart of data-arrival scheduling

1.4 功能实现

1.4.1 数据状态感知

数据状态感知组件是数据感知调度的基础,通过监听天擎的 DI 消息对数据到报状态进行实时统计,同时结合任务定义时配置的调度阈值等信息,判断数据到报状态是否达到算法启动条件。在进行数据到报状态统计分析时,首先对 DI 消息进行解析,读取四级编码、数据时间、站号或文件名、文件存储路径等信息,然后按四级编码和数据时间进行分类统计,实时统计每种数据在每个时间的到报状态,并将数据到报状态与调度阈值进行实时比对。当数据到报状态达到算法启动条件时,数据状态感知组件生成调度消息,消息内容包括算法标识、算法运行所需资源(CPU、内存)、算法运行参数等。数据状态感知组件支持将 DI 消息中的信息作为参数,在算法调度时传递给算法,如文件名及文件路径等。因此

数据感知调度可实现在海量数据中快速定位至需处理的数据文件,避免目录扫描导致的时间延迟与资源开销。

1.4.2 任务调度执行

基于调度消息,负载均衡组件进行计算节点优选后,由任务调度执行组件生成任务调度指令。计算节点优选主要基于算法运行所需资源信息及计算节点的负载信息(包括 CPU 使用量、内存使用量、正在运行的任务数量等),筛选出资源最优的计算节点作为本次算法运行的节点。任务调度执行组件结合计算节点优选信息、算法与任务信息等生成任务调度指令,并通过远程调用协议(gRPC)或超文本传输协议(http)下发至调度客户端。

1.4.3 任务调度后处理

部署于计算节点的任务执行与节点状态采集组件实时监听算法的运行状态,在算法运行结束时,将

算法运行状态(成功或失败)以消息的方式发送至调度服务端。其中任务调度后处理组件解析该状态消息,更新算法运行状态,并释放该算法所占用的计算资源。任务调度后处理组件具备异常处理与告警发送能力,当算法运行失败时,支持算法自动重新调用,若算法重新调用后仍然失败,则向天镜发送告警信息,通知业务人员进行人工干预。

1.4.4 配置管理

用户在加工流水线的任务定义环节中进行数据感知调度的参数配置。首先通过选定数据名称确定触发数据源,并指定触发类型。其中,单次触发适用于单文件调度,统计触发适用于基于到报百分率的调度,到报百分率的阈值可在页面上配置。针对单次触发类型,数据感知调度支持从DI消息中截取文件名、文件路径、站号、数据时间等信息作为参数传递给算法,简化算法处理逻辑。如算法获得传递文件名和文件路径后,算法可直接定位至需处理的数据文件,无需进行目录轮询。此外,通过配置多条数据源触发规则,实现多数据源的耦合触发,即当所配置的数据均到报后才触发算法运行。

1.4.5 任务调度可靠性设计

任务调度的可靠性对气象业务运行至关重要。从技术层面采用4项措施提高任务调度的可靠性:①调度服务端采用主备双节点设计,实时监听进程运行状态和调度服务可用性,发生异常时双节点间自动切换。②采用缓存技术,将每次任务调度所需的算法信息和任务信息等存储在缓存中,减少对数据库等外部系统的依赖,屏蔽外部系统异常导致的任务调度失败。③计算节点基于容器技术,当某计算节点出现故障时,自动将算法调度至其他可用的节点。④数据感知调度一般辅以定时调度,通过设定任务的最迟启动时间,确保当数据源完整性不足时,算法也能定时运行生成产品。

2 关键技术

2.1 数据状态感知技术

数据状态感知技术利用天镜的DI消息实现数据到报状态的实时统计,当数据到报状态满足算法启动条件时,快速启动算法。DI消息内容包括四级编码、站号、数据时间、数据到报时间、文件名等。针对不同的场景需求数据状态感知包括两种方式:一是基于到报百分比的数据感知调度,按照数据四级

编码和数据时间分类进行到报状态统计,即统计数据在最新时次的到报百分率(到报站数与应到报站数之比),该方式适用于站点类的结构化数据,如地面自动气象站小时数据等;二是基于单文件到报的数据感知调度,每到报1个数据文件即进行1次任务调度,该方式适用于雷达、卫星、数值模式等文件类数据。

2.2 任务优先调度技术

当出现计算资源不足等导致的任务调度积压时,数据感知调度支持任务优先级调度,保障高优先级任务的优先调度。任务调度优先级技术基于等待队列实现,在等待队列中按照任务优先级策略对调度消息进行动态排序,排序靠前的优先调度。任务优先级策略支持按时间先后和自定义指定优先级等。按时间的优先级调度,可设置按时间正序或倒序调度,正序调度按照业务时序先到先调度,倒序调度实现最新调度时次的优先调度,以保障最新时次产品的优先生成。时间倒序调度适用于雷达拼图、实况等对实时产品要求较高的业务。自定义优先级可设置不同任务间的调度优先级,实现重要任务的优先调度。

2.3 性能保障缓存技术

数据感知调度应用缓存技术提升调度性能和调度可靠性^[23-24]。在任务调度指令生成环节,每次任务调度均需读取算法信息(包括算法标识、算法执行命令、算法运行参数、算法运行所需资源等)和任务信息(包括调度策略、运行节点、任务重做配置、任务告警配置等),将以上信息由天擎支撑库动态加载至缓存中,减少任务调度与数据库的交互次数,既可显著提升任务调度时效(由2s提升至约200ms),又可以屏蔽因数据库异常导致的调度失败,提升任务调度的可靠性。缓存中间件部署采用集群模式,三主三从共6个节点,具备对集群内单节点故障的绝对容错能力,及对双节点故障的部分容错能力。

2.4 容器技术

天擎加工流水线的算法均以容器方式运行,其中容器管理功能基于K8s实现^[25],由管理节点和计算节点组成(图3)。管理节点接收加工流水线调度服务端的任务调度指令,基于接口管理组件、容器控制组件、容器调度组件、数据存储组件等实现集群的资源管理、容器启停与扩缩容、系统监控等。计算节点基于容器执行、容器网络管理等组件,完成容器创建、算法执行和容器状态监视。管理节点和计算节

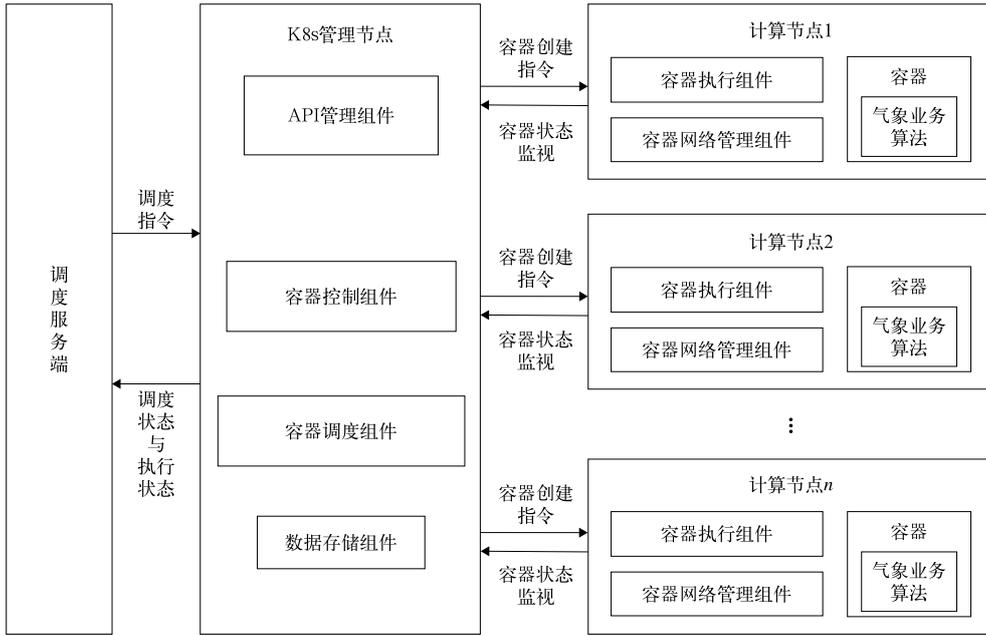


图 3 基于 K8s 的容器管理

Fig. 3 Container management based on K8s

点通过 http 通信^[20]。

3 调度性能测试

以国家级天擎业务环境实时运行的雷达单站产品生成算法为例,在资源量与性能均相同的 3 个计算节点上,同时以数据感知调度、逐 1 s 定时扫描目录、逐 10 s 定时扫描目录 3 种调度方式运行,并以调度延迟为指标表征任务调度的性能。调度延迟定义为雷达基数据到报时间与算法开始处理该基数据文件的时间的差值。该算法的数据感知调度基于单文件到报方式。测试期间(2024 年 2 月 5 日 13:00—

14:00(北京时,下同))3 种任务调度方式的调度延迟如图 4 所示,平均调度延迟分别为 11 ms、502 ms 和 3784 ms(表 1)。同时,测试期间 3 个计算节点的资源使用如图 5 所示,CPU 平均使用率分别为 36.7%、51.1%和 49.1%(表 1)。可见,数据感知调度在计算资源占用率最低的情况下,能够达到最高的调度性能,较定时调度体现出较大优势。

同时以逐小时区域极值统计算法为例,测试基于到报百分比的数据感知调度的性能(测试时间为 2024 年 3 月 20 日—4 月 19 日),该算法的数据源是中国地面逐小时数据。测试结果与单文件到报的数据感知调度相似,数据感知调度的平均调度延迟为

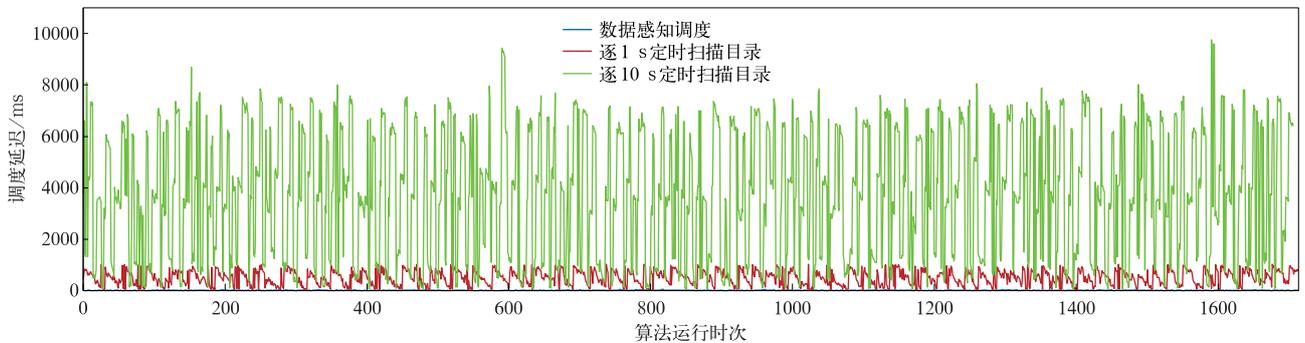


图 4 2024 年 2 月 5 日 13:00—14:00 雷达单站产品生成算法数据感知调度与定时调度性能对比测试

Fig. 4 Comparison of performance between data-arrival scheduling and timer-triggered scheduling for algorithm of single radar product from 1300 BT to 1400 BT on 5 Feb 2024

表 1 2024 年 2 月 5 日 13:00—14:00 雷达单站产品生成算法数据感知调度与定时调度平均性能对比

Table 1 Comparison of average performance between data-arrival scheduling and timer-triggered scheduling for algorithm of single radar product from 1300 BT to 1400 BT on 5 Feb 2024

任务调度方式	平均调度延迟/ms	平均计算资源使用率/%	
		CPU	内存
数据感知调度	11	36.7	12.1
逐 1 s 定时扫描目录	502	51.1	11.2
逐 10 s 定时扫描目录	3784	49.1	10.5

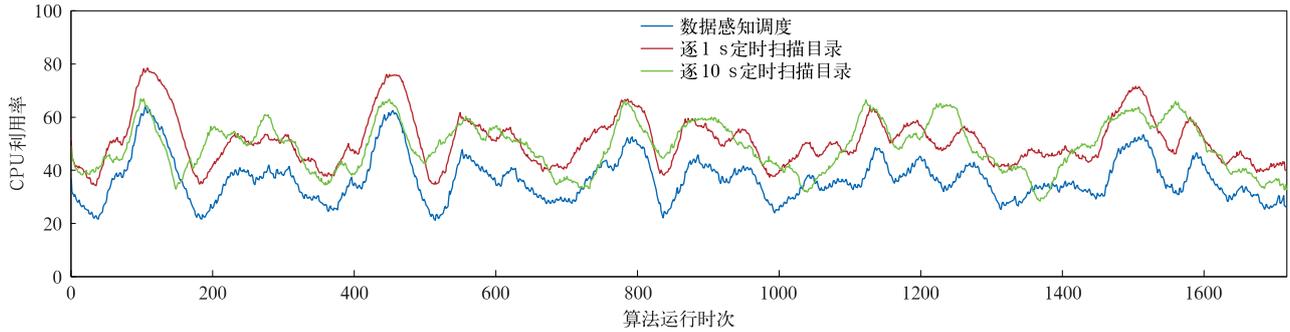


图 5 2024 年 2 月 5 日 13:00—14:00 雷达单站产品生成算法数据感知调度与定时调度对比测试期间计算资源使用

Fig. 5 Comparison of computing resource usage during between data-arrival scheduling and timer-triggered scheduling for algorithm of single radar product from 1300 BT to 1400 BT on 5 Feb 2024

4 ms, CPU 平均使用率为 15.1%; 逐 10 s 定时检索数据库的平均调度延迟为 4432 ms, CPU 平均使用率为 19.7%(图略)。

4 业务应用效益

4.1 提升气象算法运行及时性

数据感知调度一般用于高时效算法的任务调度, 截至 2024 年 5 月, 国家级天擎业务环境中共有 19 种算法(表 2)使用该调度策略, 日调度次数约为

6.67×10^5 , 平均调度延迟为 31 ms, 调度性能较原定时调度机制最高提升百余倍。其中, 基于天气雷达逐仰角基数据的单站雷达流数据处理算法的调度频次最高, 日调度次数超过 5.0×10^5 , 调度延迟为 110 ms, 有力支撑了实况业务产品的快速加工。天气雷达单站质量控制产品生成算法实现对天气雷达全体扫描标准格式基数据的快速处理, 日调度次数约为 4.6×10^4 , 调度延迟为 98 ms, 实时支撑天气雷达拼图系统 V3.0 的业务运行。

数据感知调度在各省级业务中均已部署, 并在

表 2 国家级天擎业务环境中基于数据感知调度算法清单

Table 2 Algorithms based on data-arrival scheduling in national operations of CMA Big Data and Cloud Platform

序号	算法名称	数据源	所属业务系统	平均日调度次数	日平均调度延迟/ms
1	时段内最高气温算法	中国地面逐小时数据	气象服务信息系统	2	25
2	雷达单站产品生成算法	天气雷达全体扫描标准格式基数据	面向实况应用的雷达实况产品系统	46375	15
3	单站雷达流数据处理算法	天气雷达逐仰角标准格式基数据	面向实况应用的雷达实况产品系统	507385	110
4	基于位置的实况服务-天气现象反演算法	全国逐小时总云量 融合实况分析产品(0.05×0.05)/ 中国区域多源融合实况分析 1 km 逐小时温湿风产品/中国区域多源融合实况分析 1 km 逐 10 min 降水产品	基于位置的实况服务系统	288	14
5	微波辐射计数据产品文件质量控制算法	微波辐射计基数据	基础数据产品生产	64918	101

续表 2

序号	算法名称	数据源	所属业务系统	平均日调度次数	日平均调度延迟/ms
6	逐小时区域极值统计算法	中国地面逐小时数据	基础统计产品加工系统	24	6
7	基于 GRAPES 数值预报业务系统的全球模式预报算法	GRAPES_GFS 0.25° 原始分辨率产品	数值预报业务系统产品制作平台	244	57
8	天气雷达单站质量控制产品生成算法	天气雷达全体扫描标准格式基数据	天气雷达拼图系统 V3.0	46075	98
9	天气雷达组网拼图基本反射率因子图像产品生成算法	天气雷达拼图系统 V3.0 组网基本反射率因子产品	天气雷达拼图系统 V3.0	240	5
10	天气雷达组网拼图组合反射率因子图像产品生成算法	天气雷达拼图系统 V3.0 组网组合反射率因子产品	天气雷达拼图系统 V3.0	240	31
11	天气雷达组网拼图回波顶高图像产品生成算法	天气雷达拼图系统 V3.0 组网回波顶高产品	天气雷达拼图系统 V3.0	240	5
12	天气雷达组网拼图垂直积分液态水含量图像产品生成算法	天气雷达拼图系统 V3.0 组网垂直积分液态水含量产品	天气雷达拼图系统 V3.0	240	5
13	天气雷达组网拼图 1 h 降水估测图像产品生成算法	天气雷达拼图系统 V3.0 组网 1 h 降水估测产品	天气雷达拼图系统 V3.0	240	5
14	天气雷达组网拼图雨强图像产品生成算法	天气雷达拼图系统 V3.0 组网雨强产品	天气雷达拼图系统 V3.0	240	4
15	天气雷达组网拼图 3 h 降水估测图像产品生成算法	天气雷达拼图系统 V3.0 组网 3 h 降水估测产品	天气雷达拼图系统 V3.0	24	4
16	天气雷达组网拼图 24 h 降水估测图像产品生成算法	天气雷达拼图系统 V3.0 组网 24 h 降水估测产品	天气雷达拼图系统 V3.0	24	4
17	天气雷达组网拼图未经质量控制组合反射率因子图像产品生成算法	天气雷达拼图系统 V3.0 组网未质控组合反射率因子产品	天气雷达拼图系统 V3.0	240	5
18	GFS 原始数据绘图算法	GRAPES_GFS 0.25° 原始分辨率产品	西北区域人影指挥系统	244	57
19	MESO 原始数据绘图算法	中国气象局区域模式 CMA-MESO 数值预报产品(0.03×0.03)	西北区域人影指挥系统	368	56

黑龙江、江西、山东等省级业务应用。截至 2024 年 5 月,3 省共有 14 种算法应用该调度策略(表 3),其中黑龙江省日调度次数为 1728,平均调度延迟为 44 ms;江西省日调度次数约为 5.9×10^4 ,平均调度

延迟为 415 ms;山东省日调度次数约为 1.9×10^4 ,平均调度延迟为 8 ms。江西省的数据感知调度延迟较国家级和其他省级业务偏高,经排查是因为计算资源紧张致部分时次调度积压所致。

表 3 省级天擎业务环境中基于数据感知调度的算法清单

Table 3 Algorithms based on data-arrival scheduling in provincial operations of CMA Big Data and Cloud Platform

序号	算法名称	数据源	所属业务系统	所属单位	平均日调度次数	日平均调度延迟/ms
1	黑龙江天气雷达拼图 V3.0 组网回波顶高图像产品加工算法	黑龙江天气雷达拼图 V3.0 组网回波顶高产品	雷达拼图系统	黑龙江省气象局	240	47
2	黑龙江天气雷达拼图 V3.0 组网 3 h 降水估测图像产品加工算法	黑龙江天气雷达拼图 V3.0 组网 3 h 降水估测产品(1 h 分辨率)	雷达拼图系统	黑龙江省气象局	24	47
3	黑龙江天气雷达拼图 V3.0 组网雨强图像产品加工算法	黑龙江天气雷达拼图 V3.0 组网雨强产品	雷达拼图系统	黑龙江省气象局	240	51
4	黑龙江天气雷达拼图 V3.0 组网 1 h 降水估测图像产品加工算法	黑龙江天气雷达拼图 V3.0 组网 1 h 降水估测产品(6 min 分辨率)	雷达拼图系统	黑龙江省气象局	240	44
5	黑龙江天气雷达拼图 V3.0 组网 24 h 降水估测图像产品加工算法	黑龙江天气雷达拼图 V3.0 组网 24 h 降水估测产品(1 h 分辨率)	雷达拼图系统	黑龙江省气象局	24	8
6	黑龙江天气雷达拼图 V3.0 未经质量控制组网组合反射率因子图像产品加工算法	黑龙江天气雷达拼图 V3.0 组网未经质量控制组合反射率因子产品	雷达拼图系统	黑龙江省气象局	240	45

续表3

序号	算法名称	数据源	所属业务系统	所属单位	平均日调度次数	日平均调度延迟/ms
7	黑龙江天气雷达拼图 V3.0 组网组合反射率因子图像产品加工算法	黑龙江天气雷达拼图 V3.0 组网组合反射率因子产品	雷达拼图系统	黑龙江省气象局	240	55
8	黑龙江天气雷达拼图 V3.0 组网液态垂直积分液态水含量图像产品加工算法	黑龙江天气雷达拼图 V3.0 组网垂直积分液态水含量产品	雷达拼图系统	黑龙江省气象局	240	49
9	单站逐仰角雷达基数据解码算法	天气雷达逐仰角标准格式基数据	实况融合业务系统	江西省气象局	23173	461
10	单站全体扫雷基数据解码算法	天气雷达全体扫雷标准格式基数据	实况融合业务系统	江西省气象局	1300	580
11	单站雷达降水粒子相态识别算法	天气雷达逐仰角标准格式基数据	实况融合业务系统	江西省气象局	23173	91
12	单站逐仰角 X 波段雷达基数据解码算法	X 波段天气雷达逐仰角基数据	实况融合业务系统	江西省气象局	11674	531
13	雷达基数据解析算法	天气雷达逐仰角标准格式基数据	威海市气象信息服务综合业务平台	山东省气象局	19320	8

4.2 提升气象业务间的协同性

数据感知调度在提升任务调度时效的同时,可实现上下游业务的无缝串联,在上游业务产品生成的第一时间启动下游相关业务算法,有效解决上下游业务脱节问题。中国气象局气象探测中心天气雷达拼图系统 V3.0 的图片产品生成算法应用数据感知调度,实现了上游雷达拼图产品生成后的极速启动运行,调度延迟不超过 5 ms。目前相关产品在国家级天擎中实时生成,并下发至全国应用。

数据感知调度支持多数据源的耦合触发调度。国家气象信息中心基于位置的实况服务系统^[26]中天气现象反演算法,同时使用中国逐小时总云量融合实况分析产品、中国区域多源融合实况分析 1 km 逐小时温湿风产品和中国区域多源融合实况分析 1 km 逐 10 min 降水产品 3 种数据作为输入源,为保证天气现象反演的准确性,要求 3 种数据均到报后启动运行,数据感知调度可分别判断 3 种数据源的到报情况,并在数据源均到报后的快速启动算法。

数据感知调度可实现分省/区域/流域的精细化调度。国家气象中心决策气象服务信息系统(Meteorological Service Information System, MESIS)及国家气象信息中心基础统计产品加工系统应用该调度功能,结合台站元数据等信息,解决了分省/区域/流域产品的快速加工需求。如针对东北、华北、华中等区域的逐小时最低气温的统计加工,数据感知调度通过实时判断各区域内中国地面逐小时数据的到报状态,当某区域到报率达到阈值要求时(如到报站数达 95%),立即启动对应算法计算该区域的

最低气温,同时,其他区域达到阈值后可并行计算。分区域的精细化调度同样适用于雷达拼图、实况融合产品加工等业务。

5 小结

本文详细介绍了适用于气象数据产品加工流程的数据感知调度的设计与应用,经对比测试与业务运行效果分析表明:

1) 数据感知调度在实时分析气象数据到报状态基础上,实现数据到报即算法启动运行的高效任务调度,调度延迟较定时调度显著降低,由 3784 ms 降至 11 ms。

2) 数据感知调度在国家级业务中已支撑 19 种核心业务算法的高效调度,日调度次数约为 6.67×10^5 ,平均调度延迟仅为 31 ms;在省级业务中支撑 14 种算法的高效调度,日调度次数约为 8×10^4 ,平均调度延迟为 156 ms。

3) 数据感知调度在具备高效调度性能的基础上,同时实现了气象业务上下游算法的无缝串联,可为解决气象业务间的脱节问题、提升气象业务间的协同性提供解决方案。

参考文献

- [1] 田倬璟,黄震春,张益农. 云计算环境任务调度方法研究综述. 计算机工程与应用, 2021, 57(2): 1-11.
Tian Z J, Huang Z C, Zhang Y N. Review of task scheduling methods in cloud computing environment. *Comput Eng Appl*, 2021, 57(2): 1-11.

- [2] 陈红华,崔脩龙,王耀杰.基于多种云环境的任务调度算法综述.计算机应用研究,2023,40(10):2889-2895.
Chen H H,Cui X L,Wang Y J. Summary of task scheduling algorithms based on multiple cloud environments. *Appl Res Comput*,2023,40(10):2889-2895.
- [3] 杨戈,赵鑫,黄静.面向云计算的任务调度算法综述.计算机系统应用,2020,29(3):11-19.
Yang G,Zhao X,Huang J. Survey on task scheduling algorithms for cloud computing. *Comput Syst Appl*,2020,29(3):11-19.
- [4] 曹禹,李晓辉,刘忠麟,等.云环境大数据工作流编排管理系统研究综述.计算机与现代化,2022(1):41-53.
Cao Y,Li X H,Liu Z L,et al. Review of big data workflow orchestration and management system in cloud environment. *Comput Mod*,2022(1):41-53.
- [5] 赵立成,沈文海,肖华东,等.高性能计算技术在气象领域的应用.应用气象学报,2016,27(5):550-558.
Zhao L C,Shen W H,Xiao H D,et al. The application of high performance computing technology in meteorological field. *J Appl Meteor Sci*,2016,27(5):550-558.
- [6] 赵春燕,孙婧,肖华东,等.基于工作流的气象超算作业调度应用研究.计算机技术与发展,2022,32(9):193-199.
Zhao C Y,Sun J,Xiao H D,et al. Application research of workflow-based meteorological job scheduling on super-computing. *Comput Technol Dev*,2022,32(9):193-199.
- [7] 张明,杨效,邹丽,等.作业调度软件 LoadLeveler 在气象模式中的应用.气象科技,2021,49(1):63-68.
Zhang M,Yang X,Zou L,et al. Application of job scheduling software LoadLeveler in meteorological mode. *Meteor Sci Technol*,2021,49(1):63-68.
- [8] 朱智,师春香,刘瑞霞,等.我国三维云融合分析业务系统(3DCloudA-V1.0)研制与开发.气象科技,2021,49(1):55-62.
Zhu Z,Shi C X,Liu R X,et al. Research and development of three-dimensional cloud merging analysis operation system (3DCloudA-V1.0) in China. *Meteor Sci Technol*,2021,49(1):55-62.
- [9] 马强,颜京辉,魏敏,等.北京气候中心 CMIP6 数据共享平台及应用.应用气象学报,2022,33(5):617-627.
Ma Q,Yan J H,Wei M,et al. Implementation and application of BCC CMIP6 experimental data sharing platform. *J Appl Meteor Sci*,2022,33(5):617-627.
- [10] 张进,孙健,沈学顺,等.CMA-GFS V4.0 模式关键技术研发和业务化.应用气象学报,2023,34(5):513-526.
Zhang J,Sun J,Shen X S,et al. Key model technologies of CMA-GFS V4.0 and application to operational forecast. *J Appl Meteor Sci*,2023,34(5):513-526.
- [11] 何文春,高峰,许艳,等.气候监测业务专题数据服务系统的设计与实现.应用气象学报,2012,23(5):624-630.
He W C,Gao F,Xu Y,et al. Design and implementation of special data service system for climate monitoring. *J Appl Meteor Sci*,2012,23(5):624-630.
- [12] 熊安元,赵芳,王颖,等.全国综合气象信息共享系统的设计与实现.应用气象学报,2015,26(4):500-512.
Xiong A Y,Zhao F,Wang Y,et al. Design and implementation of China Integrated Meteorological Information Sharing System(CIMISS). *J Appl Meteor Sci*,2015,26(4):500-512.
- [13] 孙超,霍庆,任芝花,等.地面气象资料统计处理系统设计与实现.应用气象学报,2018,29(5):630-640.
Sun C,Huo Q,Ren Z H,et al. Design and implementation of surface meteorological data statistical processing system. *J Appl Meteor Sci*,2018,29(5):630-640.
- [14] 吕终亮,白新萍,薛峰.基于 WebGIS 的气象服务产品制作系统及关键技术.应用气象学报,2018,29(1):120-128.
Lü Z L,Bai X P,Xue F. WebGIS-based meteorological service system and its key technology. *J Appl Meteor Sci*,2018,29(1):120-128.
- [15] 丰德恩,唐卫,王慕华,等.基于 WebGIS 的气象服务产品自动加工关键技术.气象与环境科学,2020,43(1):130-136.
Feng D E,Tang W,Wang M H,et al. Key technologies of automatic processing of meteorological service products based on WebGIS. *Meteor Environ Sci*,2020,43(1):130-136.
- [16] 顾文静,王鹏,戴晴,等.基于流程控制的国际气象通信数据收集管理系统的设计与实现.气象科技,2023,51(1):31-39.
Gu W J,Wang P,Dai Q,et al. Design and implementation of international meteorological communication data collection and management system. *Meteor Sci Technol*,2023,51(1):31-39.
- [17] 李莹,王国复.气象灾害风险管理系统设计与应用.应用气象学报,2022,33(5):628-640.
Li Y,Wang G F. Design and implementation of Meteorological Disaster Risk Management System. *J Appl Meteor Sci*,2022,33(5):628-640.
- [18] 杨和平,张强,罗兵,等.气象综合指挥平台建设与应用.应用气象学报,2023,34(1):117-128.
Yang H P,Zhang Q,Luo B,et al. Construction and application of Meteorological Integrated Command Platform. *J Appl Meteor Sci*,2023,34(1):117-128.
- [19] 梁丽,马舒庆,滕玉鹏,等.天气雷达空中生态监测系统建设和应用.应用气象学报,2023,34(5):630-640.
Liang L,Ma S Q,Teng Y P,et al. Construction and application of Weather Radar Aerial Ecological Monitoring System. *J Appl Meteor Sci*,2023,34(5):630-640.
- [20] 霍庆,何文春,何林,等.气象大数据云平台算法集约化环境设计与应用.应用气象学报,2024,35(1):80-89.
Huo Q,He W C,He L,et al. Design and application of algorithm intensive environment for CMA big data and cloud platform. *J Appl Meteor Sci*,2024,35(1):80-89.
- [21] 赵芳,熊安元,张小纓,等.全国综合气象信息共享平台架构设计技术特征.应用气象学报,2017,28(6):750-758.
Zhao F,Xiong A Y,Zhang X Y,et al. Technical characteristics of the architecture design of China Integrated Meteorological Information Sharing System. *J Appl Meteor Sci*,2017,28(6):750-758.

- logical Information Sharing System. *J Appl Meteor Sci*, 2017, 28(6): 750-758.
- [22] 孙超,肖文名,陈永涛,等.气象综合业务实时监控系统的设
计.气象科技进展,2018,8(1):153-157.
Sun C, Xiao W M, Chen Y T, et al. Design of real-time moni-
toring system of integrated meteorological operation. *Adv
Meteor Sci Technol*, 2018, 8(1): 153-157.
- [23] 曾超宇,李金香. Redis 在高速缓存系统中的应用. 微型机与应用,
2013, 32(12): 11-13.
Zeng C Y, Li J X. Redis application in cache system. *Micro-
comput Appl*, 2013, 32(12): 11-13.
- [24] 李焱,顾乃杰,黄增士,等. Redis 集群可靠性的研究与优化.
计算机工程,2018,44(5):40-46.
Li Y, Gu N J, Huang Z S, et al. Research and optimization of
redis cluster reliability. *Comput Eng*, 2018, 44(5): 40-46.
- [25] Jaramillo D, Nguyen D V, Smart R. Leveraging Microservices
Architecture by Using Docker Technology//Southeast Con.
IEEE, 2016: 1-5.
- [26] 徐拥军,倪学磊,郑波,等. 基于位置的天气实况数据服务接口
设计与应用. 计算机系统应用, 2023, 32(5): 77-86.
Xu Y J, Ni X L, Zheng B, et al. Design and application of live
weather data service interface based on user's location. *Com-
put Syst Appl*, 2023, 32(5): 77-86.

Design and Application of Meteorological Algorithm Scheduling Framework Based on Data Perception Technology

Huo Qing He Wenchun Gao Feng Chen Shiwang Xu Yongjun Wang Qi

(National Meteorological Information Center, Beijing 100081)

Abstract

The generation efficiency of data products depends on both the computational efficiency and the startup efficiency of algorithms. In meteorological operations, algorithms are typically data-driven, meaning they are initiated immediately upon data arrival to accelerate the generation time of data products. Therefore, data-driven meteorological services urgently require an efficient task scheduling framework to achieve the goal of starting and running algorithms as soon as data arrive, and to improve the generation efficiency of meteorological data products. CMA Big Data and Cloud Platform, referred to as Tianqing and led by National Meteorological Information Center, began nationwide business operations in December 2021. The data processing line (DPL), as the core function of Tianqing, enables the unified management and centralized scheduling of meteorological algorithms. DPL has established various task scheduling capabilities, including timer-triggered scheduling, sequential scheduling, data-arrival scheduling, and manual scheduling. Among these methods, data-arrival scheduling based on data reporting status enables the algorithm to start immediately after data reporting, greatly improving the startup time of meteorological algorithms and the generation time of meteorological data products.

Core functions of data-arrival scheduling include data state awareness components, task scheduling execution components, task scheduling post-processing components, and configuration management. Among these components, the data state perception component realizes real-time analysis of the reporting status of various meteorological data in the sky engine, and sends scheduling messages to the task scheduling execution component when the data reporting rate meets scheduling requirements. The task scheduling execution component combines the necessary resource information for the algorithm and computing nodes optimization, generates task scheduling instructions, and implements algorithm startup execution. The post-processing phase of task scheduling includes gathering and updating algorithm execution status, calculation node status, and sending of alarm information for algorithm execution abnormalities. Configuration man-

agement supports configuring data-aware scheduling parameters on the front-end page.

Real-time analysis of data status and scheduling enables efficient task scheduling that starts and runs the algorithm as soon as data are reported. The scheduling delay is significantly reduced, compared to the original schedule, from 3784 ms to 11 ms. Data-arrival scheduling, as the core capability of CMA Big Data and Cloud Platform, is deployed and operated in real-time in provinces and regions. Currently, the system supports the efficient scheduling of 19 core business algorithms at the national level, with a total of approximately 6.67×10^5 daily scheduling tasks and an average scheduling delay of 31 ms. Efficient scheduling is achieved with 14 algorithms supported at the provincial level, encompassing a total of approximately 8×10^4 daily scheduling times and an average scheduling delay of 156 ms. In addition, data aware scheduling achieves seamless integration of upstream and downstream algorithms in meteorological services, providing a solution to eliminate the problem of disconnection between meteorological services and improve collaboration between meteorological services.

Key words: data perception technology; task scheduling; meteorological algorithm; CMA Big Data and Cloud Platform; data processing line