

王慕华,唐卫,丰德恩,等,2018.基于消息驱动的气象图形产品加工系统[J].气象,44(5):692-698.

## 基于消息驱动的气象图形产品加工系统\*

王慕华 唐卫 丰德恩 渠寒花 王阔音 鹿业涛

中国气象局公共气象服务中心,北京 100081

**提 要:** 高质量的气象图形产品是融媒体时代气象信息服务的重要载体。本文综合运用并行计算和大数据等技术,构建消息驱动的气象图形产品加工系统(WeatherStudio),分析了公共气象服务图形产品加工共性问题,探讨了基于消息驱动的产品加工系统的优势,设计了基于消息驱动和定时驱动相融合的气象图形产品加工系统的框架和功能,并以高速公路交通气象灾害风险预警业务为应用实例进行验证,显示出了其较好的应用前景。

**关键词:** 气象图形, 产品加工, 消息驱动

**中图分类号:** P409

**文献标志码:** A

**DOI:** 10.7519/j.issn.1000-0526.2018.05.010

### The Message-Driven Making System of Meteorological Graphic Products

WANG Muhua TANG Wei FENG De'en QU Hanhua WANG Kuoyin LU Yetao

CMA Public Meteorological Service Centre, Beijing 100081

**Abstract:** The high quality of meteorological graphic products is important for the public meteorological service (PMS) in the time of media convergence. Based on the techniques of parallel computing and big data, message-driven making system of meteorological graphic products was developed, namely WeatherStudio. After analyzing the common problems of making graphical products for PMS, this paper discussed the advantages of the message-driven product making system, and then focused on the overall structure and main functions of WeatherStudio based on the message and timing driven techniques. Finally, the application of WeatherStudio in the expressway meteorological disaster risk early warning service shows its potential prospect.

**Key words:** meteorological graphic, product making, message-driven

### 引 言

气象图形产品是融媒体时代气象信息服务的重要载体,国内外气象部门已建设了适用于不同应用场景的气象图形产品加工系统。国际上,美国马里兰大学开发的 GrADS、美国国家大气研究中心的 NCAR 绘图软件(赵连伟等,2000,刘旭林和赵文芳,2009),是美国大气科学研究和业务工作通用的计算机绘图工具,具有较强的专业性,但在三维图形

支持(张秀丽等,2002)和本地汉化方面(乔春贵等,2007)存在不足。美国 The Weather Company 公司开发的 WeatherCentral 软件针对影视图形图像产品制作,具有较强的产品渲染能力和较好的可视化效果,并针对部分气象业务场景实现了产品的多样化制作功能(洪萍等,2014)。国内而言,国家气象中心研发的气象信息综合处理系统(MICAPS)(李月安等,2010)和决策气象服务信息系统(MESIS)(吴焕萍,2010;吴林林等,2013;吕终亮等,2012;郑卫江等,2015),实现了多种气象标准格式数据的交互分

\* 公益性行业(气象)科研专项(GYHY201406029)资助

2016年11月9日收稿; 2017年8月15日收修定稿

第一作者:王慕华,主要从事GIS及遥感技术在公共气象服务领域的应用。Email:wangmuhua@cma.gov.cn

通信作者:唐卫,主要从事公共气象服务系统开发关键技术研究及气象大数据技术的研究与应用。Email:tangw@cma.gov.cn

析、可视化和预报制作,并且以气象预报业务逻辑为驱动,专业性较强、自动化程度较高;张振涛等(2014)研发的基于天气事件的公共气象服务产品制作系统,以天气热点事件为驱动,依托 GIS 组件、计算机图形学等技术开展产品的交互式加工。

上述系统在公共气象服务发展前期发挥了重要作用,但面对新的服务需求,在性能、质量和效率上又略显不足。公共气象服务对象多,包括公众、决策和专业用户;服务渠道多样,既包括影视传统媒体,又包括手机、网站、微博、微信等新媒体;服务领域广,包括交通(公路、铁路、航空)、森林草原火险、能源电力、海洋等;服务时效高,从分钟级到小时级的气象服务产品;图形质量要求高,融媒体及专业服务产品,对产品的表现形式及质量都有不同的要求。为了适应新时代对气象图形产品的自动化、个性化、高性能、高质量、多源数据融合的要求,急需建立新一代高性能气象服务图形产品加工系统。

## 1 基于消息和定时驱动融合的图形产品加工技术

### 1.1 基于 Kafka 的分布式消息

传统的图形产品加工多基于定时运行的工作模式,多采用FTP作为数据传输手段,以传统关系型数据库和文件系统作为存储介质,效率较低,在面对分

钟级的雷达数据及新一代静止卫星数据加工时,无法满足客户对数据和产品实时性的要求。同时,传统的单机、单线程系统架构在性能、吞吐量和容错性等方面都存在缺陷,无法适应海量气象服务产品的加工需求。而基于消息的分布式系统具有扩展性好、时效性高和异步通信能力等优点,被广泛应用于准实时性数据业务中。

Apache Kafka(Tharakan,2016)是一种分布式的,基于发布/订阅机制的消息系统,具有可水平扩展、高吞吐率和实时性好的特点而被广泛使用。一个简单的 Kafka 消息系统(图 1)包括三部分,消息的发布者称作生产者(Producer),消息的订阅者称作消费者(Consumer),中间的存储阵列称作代理者(Broker)。生产者将数据生产出来,推送(Push)给代理者进行存储,消费者需要消费数据时,就从代理者中拉取(Pull)数据,然后完成一系列对数据的处理。

基于 Kafka 的消息驱动气象图形产品加工技术框架如图 2 所示。多源数据集(Laps、OCF、MinCast 等)作为生产者生成消息日志文件;Zookeeper 作为分布式应用程序协调服务,通过设置 Zookeeper



图 1 简单的 Kafka 消息系统  
Fig.1 Simple structure of Kafka message system

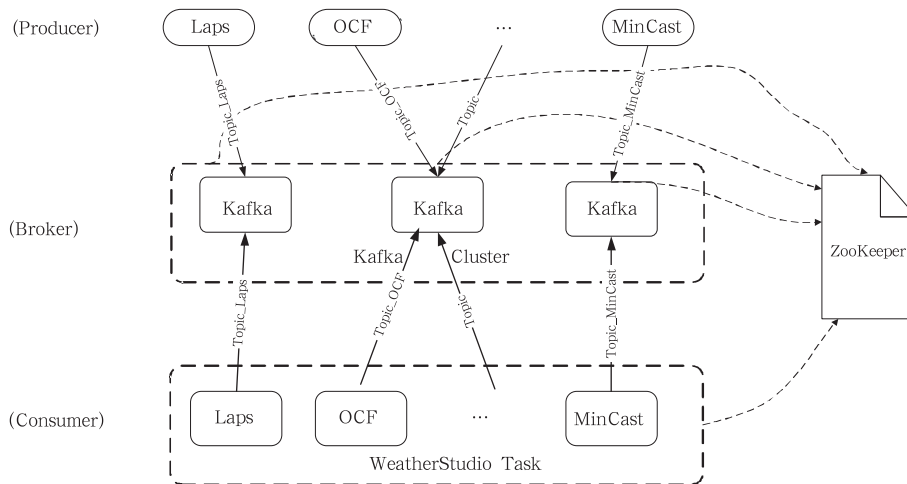


图 2 基于 Kafka 的消息驱动产品加工技术  
Fig.2 Message-driven product making technique based on Kafka

心跳,实时采集日志内容生成约定的 Kafka 集群的消息类别(Topic)并加入发送队列,消息由生产者主动推送给代理者。三台代理服务器组成的 Kafka 集群对不同消息进行接收和分布式存储管理。系统通过配置 Laps、OCF、MinCast 各产品加工任务相对应的消息类别,从代理拉取相应消息,进行加工任务参数解译和任务调度,从而完成消息驱动的产品加工任务。

## 1.2 定时驱动的产品加工技术

由于历史原因,很多已有的系统和数据源都没有基于消息的分布式架构实现,为了对接这些数据源和系统,采用定时驱动相补充的方式来设计图形产品加工系统。QUARTZ(2017)是 Java 编写的开源作业调度框架,具有支持对应时间点和精确定义时间点两种方式的定时任务触发机制,支持集群部署及多种错误处理形式(如任务执行错误后下次不执行、马上重新执行、下次继续执行等),支持多种任务漏触发处理(如关机漏触发、数据源未到漏触发等)的特点,在解决多源、多时效的气象服务产品定时加工任务时具有优势。

QUARTZ 主要包含调度器监听、作业及触发器监听三部分。通过重载监听类中处理函数实现具体的定时作业调度服务。调度器为框架的核心,为确保可伸缩性,QUARTZ 采用了基于多线程的架构。QUARTZ 依赖一个松耦合的线程池来管理部件及线程环境,每个对象根据业务需求都是可配置的。框架启动后,调度器初始化一套工作线程,该线程被用来执行已预定好的作业。基于 QUARTZ 的定时驱动产品加工技术,用户根据产品加工任务需求,在前台设置产品加工运行时间规律,经分析形成各类产品加工任务调度时间表达式持久化到数据库中。系统启动后,QUARTZ 根据服务器资源情况,将调度器总线程数采用 3:1 的分配原则分配给实时业务调度器和错误补做调度器。实时业务调度器读取数据库中的时间表达式,监听剩余时间,当调度程序监听到剩余时间小于某一临界范围时,QUARTZ 框架将启动具体产品加工调度类的作业函数传入相应的任务参数。如果作业失败,将该任务加入错误补做调度器线程池,根据预设的错误补做时间间隔重新作业。如果配置的作业在随后的运行中再次被调用,QUARTZ 框架将在预定的时间再次调用它。通过调用 QUARTZ 框架,便捷地实

现了定时任务的触发、数据源未到等原因导致的漏触发、错误补做等复杂的业务逻辑,实现系统的稳定运行。

## 2 系统架构和功能

### 2.1 系统架构

气象图形产品加工系统(WeatherStudio)首先要具有对海量气象产品高效率的加工能力,尤其是针对分钟级时间分辨率的产品,要求基于消息驱动,秒级完成一张图形产品的加工;其次,WeatherStudio 还需具备高稳定性系统架构,特别是面向实时直播类的产品加工任务,系统一旦发生故障,能够快速恢复,从而保障业务稳定运行;再次,面对融媒体个性化需求,如图形风格、标题、图例等,WeatherStudio 需具有强大的制图能力;最后,从公共气象服务产品加工业务一体化设计的角度,WeatherStudio 需具有先进的、开放式的技术架构,能够适应服务需求的不断变化,达到精准化、按需气象服务的目的。

利用大数据分布式处理架构,并分别从以下五个方面予以具体实现:(1)采用 Kafka 分布式消息系统,实现基于消息驱动的任务触发机制,达到每秒百万级别的数据处理能力,满足分钟级产品高性能加工的要求。(2)效率体现在对产品加工服务的分层设计上,包括下层的原子服务层和上层的组合服务层。原子服务层由最小单元的业务逻辑封装而成,组合服务层由原子服务通过标准接口,按照一定的业务流程编排而成,从而实现业务解耦,高效率响应业务需求的目的。(3)通过对数据库的读写分离,Kafka 集群、ArcGIS Server 服务集群(彭涛,2009)的构建和全业务流程的日志监控实现系统的高可用性。(4)通过消息和计划任务调度管理器实现消息与定时驱动的产品加工任务自动化运行。(5)通过地图制图技术,配置不同的产品加工模板,满足高质量和个性化的产品加工需求。

从技术架构上看,系统分为三层结构,如图 3 所示。

持久层采用关系型数据库 Oracle 与 HDFS 分布式文件系统两种方式实现原始数据与生成产品的存储和管理。关系型数据库用来存放离散点类气象数据(任芝花等,2015)。对于网格点类气象数据,由于具有时间、空间的多维特征,用传统的关系型数据

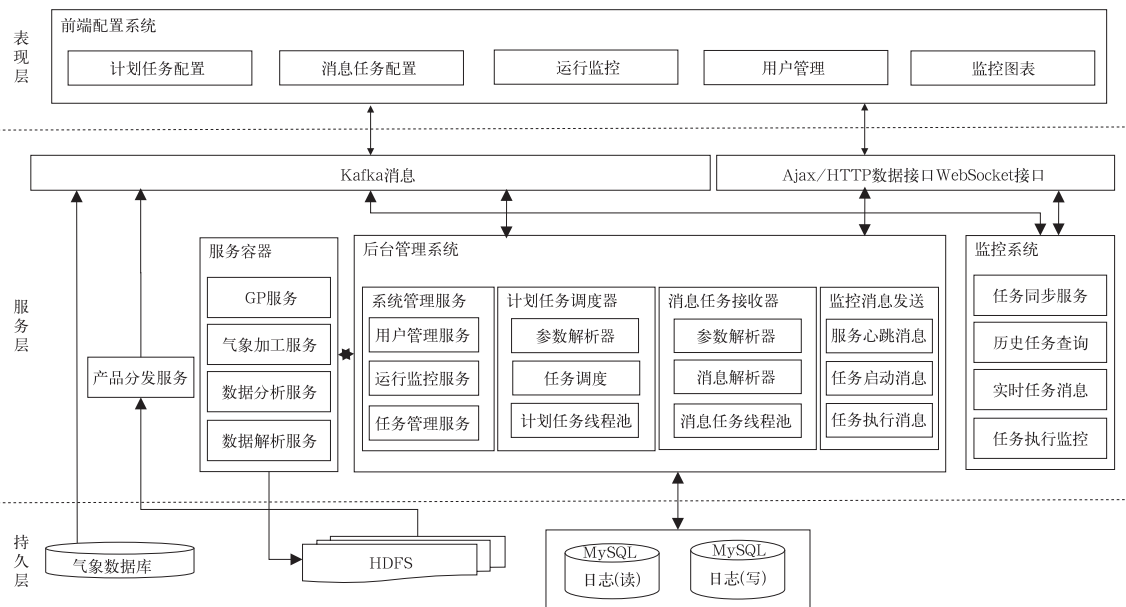


图 3 总体架构图

Fig. 3 Overall structure of WeatherStudio

库很难表征气象要素间的关系,且检索效率低下。HDFS 具有高容错、高可靠性、高扩展性、高吞吐率等特征为海量数据存储提供了可靠的持久化功能(徐海荣等,2015),为气象数值模式产品及雷达分钟级降水产品等超大数据集的应用处理带来了便利。

服务层通过标准服务接口实现与持久层和表现层的贯通。首先,通过构建统一的空间气象数据模型,对气象数据类型进行逻辑抽象,研发多种气象数据解析组件服务(即原子服务),实现对不同类型数据的解析处理,包括:MICIAPS 系统定义的多种数据格式,数据库(Oracle,MySQL,Access 等)格式,WebService 或 Restful 数据接口服务格式,NetCDF、GRIB、HDF 格点类数据集等,从而实现数据解析处理业务逻辑的服务化管理。根据不同的产品加工业务需求,还封装了插值分析服务、制图整饬服务、图形生成服务等。组合服务分为两类,一类是 ArcGIS GeoProcessing 服务(简称 GP 服务),即将多个原子服务通过 ArcGIS ModelBuilder 编排成流水线式气象加工模型形成 GP 组合服务;另一类则是自主研发的 Webservice 服务,遵从服务标准规范,跟开发语言无关。原子服务和组合服务统一由 ArcGIS Server 服务容器进行注册和管理,并针对不同的服务类型提供多种服务接口。

任务调度方面,利用 QUARTZ 技术,构建计划任务调度器,实现定时任务请求的参数解析和任务

调度。针对消息任务,采用 Kafka 消息发布订阅机制启动消息任务线程执行加工任务。

实时监控方面,综合利用 Zookeeper,Flume,Kafka 等技术完成。实时业务通过 Flume 收集系统中各个任务的运行日志,发送到 Kafka 日志总线中,并由系统中的监控模块分析处理 Kafka 中各个任务的日志信息,将每个任务的起止运行时间、任务运行状态(正在运行、运行成功、运行失败)、失败原因等信息进行可视化展示,从而实现系统监控的实时性与稳定性。

表现层与服务层间采用 Ajax/Http 方式和 WebSocket 方式进行通信。表现层实现了计划任务配置、消息任务配置、运行监控显示等功能,采用主流的 Html5 技术,并借鉴 JavaScript 开源框架,AngularJS 实现数据绑定功能,Bootcss 实现组件样式的定制,Highcharts 进行统计图表的展现,以及 JavaScript UI 实现前端界面的设计。

## 2.2 功能设计和实现

从系统逻辑结构上看,WeatherStudio 分为两部分:加工模型构建部分和任务配置调度部分。其中,加工模型构建部分由可扩充模型组件库、加工服务装配、服务注册发布三部分组成,针对每一类产品的加工需求,通过工作流引擎构建一套产品加工模型;任务配置调度部分由定时驱动加工任务、消息驱

动加工任务以及业务监控三部分组成,一方面通过用户界面实现加工任务参数的配置,另一方面由统一的任务调度器实现加工任务的并行有序开展。功能结构图如下所示(图 4)。

(1) 可扩充模型组件库,根据气象图形产品加工业务逻辑,将数据接入、数据分析、制图整饬、产品生成四个环节进行业务抽象,封装为模型组件,并根据模型组件标准,扩展新的业务模型服务。

(2) 加工服务装配,将基于业界主流技术标准的业务模型组件,通过定义良好的接口和契约联系起来,使业务人员通过简单拖拽等操作实现加工流程的设计与编排。

(3) 服务注册发布,将模型组件进行注册,实现对产品加工模型和数据源的发布和管理,对产品加

工制作服务连接池中的模型服务以 SOAP 服务和 REST 服务的形式提供调用。

(4) 定时驱动任务加工,将定常性的加工任务,通过参数化的配置,实现基于定时任务的触发加工模式。

(5) 消息驱动任务加工,基于 Apache Kafka 消息系统,对高时效性的加工任务,采用消息发布订阅机制进行任务调度,实现基于消息的高效产品加工模式。

(6) 全业务流程监控,包括对定时任务和消息任务的启动、暂停、执行、结束等环节的监控,以及数据库、数据文件、产品分发的监控,从而实现全业务流程的监控。

### 3 业务应用实例

利用 WeatherStudio 对高速公路大风、低能见度和强降雨灾害风险预警产品进行自动化加工(焦圣明等,2017),结果表明,系统的性能有了明显的提高。下面以强降雨灾害风险产品的加工为例,分别从原子组件服务、产品加工模型、任务执行时间三个方面进行具体说明。

表 1 从数据接入、数据处理、数据分析、制图整饬和产品生成五个环节,分析高速公路强降雨灾害风险预警产品主要应用的原子组件服务,并详细说明了组件的功能。

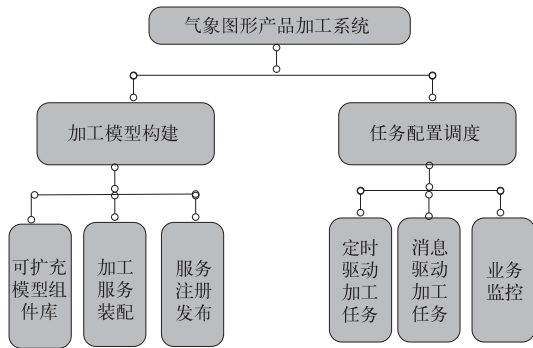


表 1 高速公路强降雨灾害风险预警产品原子组件服务

Table 1 Early warning product atomic component service on severe rainfall disaster risk on expressways

序号	流程环节	原子组件服务	说明
1	数据接入	MICAPS 数据解析	将 MICAPS 数据解析为矢量数据
		数据筛选	异常值排除
2	数据处理	栅矢转换	栅格数据转换为矢量数据
		属性删除	删除不必要的属性信息
		反距离权重插值	将气象离散点插值生成栅格曲面
		分类	根据产品分级标准对数据进行分类处理
3	数据分析	栅格计算器	利用栅格计算器编辑强降雨风险预警模型公式
		叠加分析	对数据源进行空间裁剪和空间叠加操作
		属性操作	对气象要素属性信息进行计算操作
4	制图整饬	更新地图模板	地图模板里定义了产品的背景图片、标题、图例、地图符号等样式,将气象产品作为专题图层叠加到地图模板上
5	产品生成	生成 JPEG	设置图形的分辨率、图幅大小

高速公路强降雨灾害风险预警产品加工模型(图 5)的构建思路如下:利用 MICAPS 数据解析组件读取强降雨预报数据,通过空间裁剪提取研究区

域的数据,并排除奇异点值,对离散点数据进行反距离权重插值生成栅格数据类型,按照产品分级标准对数值进行危险性分类,得到多值背景栅格数据,根

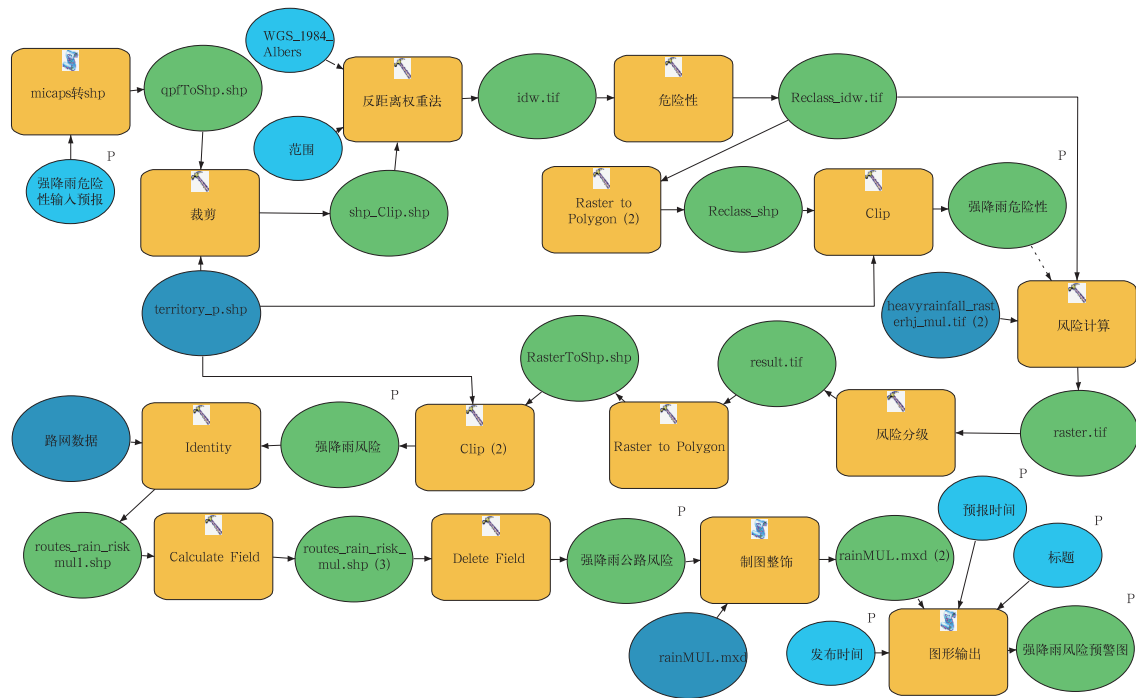


图 5 高速公路强降雨灾害风险预警产品加工模型

Fig. 5 Early warning product model for severe rainfall disaster risk on expressways

据历史 10 年的统计资料,建立强降雨灾害风险数学模型,通过栅格计算器组件实现数学模型的编辑,并利用属性操作组件删除不必要的属性信息。至此,数据接入、数据分析和数据的工作已经完成,后面针对产品的表现形式和输出利用更新地图模板和生成 JPEG 两个组件完成。

利用 WeatherStudio 对高速公路强降雨灾害风险预警产品进行自动化加工,图6给出了2016年9

月 10—20 日该产品的加工任务运行时间统计结果,最短耗时在 48 s,最长耗时在 54 s,与传统的单机加工模式对比,效率提高了 3 倍以上。除此之外,WeatherStudio 支持并行调度策略,在多个任务并发执行时性能更具有明显的优势。

### 4 结 论

基于消息驱动的气象图形产品加工系统,在充分吸纳 MESIS 软件产品优势的基础上,采用消息及分布式技术,满足了公共气象服务图形产品加工效率、质量、个性化等多方面的需求。在近一年的业务试运行过程中,表现出了良好的应用前景,但也存在的一些不足,主要表现在:(1)基于时间序列的动画产品加工功能尚不具备,在天气过程性气象服务表示时支撑不够;(2)大数据分析方法与集成较少,面向多源气象数据的融合分析彰显不足;(3)气象大数据可视化图表类产品还未涉猎。

### 参考文献

洪萍,姚立峰,俞卡莉,等,2014. WeatherCentral 系统在电视气象节

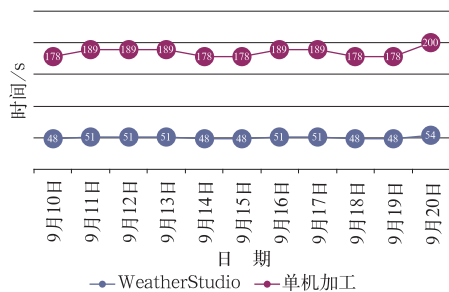


图 6 2016 年 9 月 10—20 日高速公路强降雨灾害风险预警产品任务执行时间统计对比

Fig. 6 Comparison of task execution time statistics of early warning product for severe rainfall disaster risk on expressways during 10—20 September 2016

- 目中的应用分析[J]. 浙江气象, 35(1):20-22, 44.
- 焦圣明, 郑媛媛, 王宏斌, 等, 2017. 灾害性天气个例库智能分析系统的设计与实现[J]. 气象, 43(3):354-364.
- 李月安, 曹莉, 高嵩, 等, 2010. MICAPS 预报业务平台现状与发展[J]. 气象, 36(7):50-55.
- 刘旭林, 赵文芳, 2009. 气象观测数据等值线自动绘制系统[J]. 气象, 35(4):102-107.
- 吕终亮, 罗兵, 吴焕萍, 等, 2012. MESIS 信息检索及可视化产品制作平台实现[J]. 应用气象学报, 23(5):631-637.
- 彭涛, 2009. 基于 ArcGIS Server 的城市公众地理信息服务系统研究[D]. 赣州:江西理工大学, 132-135.
- 乔春贵, 王国安, 杨建民, 2007. 两种气象绘图软件的对比及应用[J]. 气象与环境科学, 30(S1):159-160.
- 任芝花, 张志富, 孙超, 等, 2015. 全国自动气象站实时观测资料三级质量控制系统的研制[J]. 气象, 41(10):1268-1277.
- 吴焕萍, 2010. GIS 技术在气象领域中的应用[J]. 气象, 36(3):90-100.
- 吴林林, 刘黎平, 徐海军, 等, 2013. 基于 MICAPS 3 核心的人影业务平台设计与开发[J]. 气象, 39(3):383-388.
- 徐海荣, 陈闰叶, 张兴媛, 2015. 基于 Flume、Kafka、Storm、HDFS 的航空维修大数据系统[J]. 上海工程技术大学学报, 29(4):303-305, 311.
- 张秀丽, 徐桂荣, 俞康庆, 2002. 一种新型绘图软件在气象中的应用初探[J]. 气象科学, 22(2):223-229.
- 张振涛, 张正文, 陈宇, 等, 2014. 基于天气事件的公共气象服务产品制作系统[J]. 应用气象学报, 25(2):249-256.
- 赵连伟, 柳晓放, 赵远东, 2000. Windows 版本 NCAR 绘图软件包的开发与应用[J]. 辽宁气象, (4):27-28.
- 郑卫江, 吕终亮, 白新萍, 等, 2015. 区域决策气象服务产品自动化制作关键技术——以自动站小时累加降水量产品为例[J]. 气象, 41(9):1166-1169.
- QUARTZ, 2017. QUARTZ By software[EB/OL]. <http://www.quartz-scheduler.org/>.
- Tharakan R, 2016. Kafka: A high-throughput distributed messaging system[EB/OL]. (2010-12-03). <https://www.royans.net/wp/2010/12/03/kafka-a-high-throughput-distributed-messaging-system/>.