

基于分布式和微服务架构的华信气象服务云平台

■ 崔宏 吴恩平 熊安元 王尊鹏 赵笔锋 龚伟伟 邓鑫

华信气象服务云平台应用数据库、通信、并行计算等传统技术结合新兴的云计算、大数据技术进行建设。基于Hadoop分布式处理框架与应用Spring Cloud框架搭建、Docker容器封装、RESTful API作为服务接口的微服务架构，平台实现了各服务模块的高内聚、低耦合、灵活可扩展，为气象服务产品开发和实时服务提供数据环境和基础资源支撑，以实现面向航空、电力、保险等不同行业提供定制化的气象信息服务。

近年来，随着我国经济的迅速发展，各行各业对气象服务的需求也日益凸显。《气象信息服务管理办法》于2015年6月1日起开始施行，促进了我国气象信息服务发展、气象信息服务市场的培育与气象信息服务活动的规范化，气象服务进入快速发展阶段。华信气象服务云平台即为开展B2B定位的气象服务业务而建设，基于云平台提供的数据库环境以及基础资源支撑，通过将气象行业数据与不同领域的需求结合进行信息再加工，提供实时更新的有针对性的气象服务，为相关行业生产经营规划与决策提供信息支撑，为防灾减灾服务。

作为气象服务云平台的核心内容，气象数据具有时空尺度跨度大、数据体量大（但可控）等特点，符合大数据体量大、数据种类繁多、数据需求处理速度快、数据价值高的基本特征，海量气象数据“密集型科学发现”以及“与其他相关行业或领域数据的深度融合以获得跨领域、跨学科的新知识”成为未来气象科学发展和进步的新途径和模式。因此，可以应用大数据相关技术解决海量气象数据的规范化存储管理、高时效应用等技术问题。由于气象服务具有业务不断丰富扩展与气象服务产品个性化定制需求不断增加的特点，因此需要建设一个基础资源与业务逻辑可灵活扩展的服务平台，分布式与微服务结合的架构可以满足气象服务平台可扩展性要求。其中，云平台的弹性计算、虚拟化、按需服务、异地灾备、分布式存储与处理等技术特性可以满足气象服务平台基础资源扩展的需求；专注于具体业务功能实现、模块边界明确、细分功能为独立服务的微服务架构具有技术选型灵活、独立按需扩展、可用性高等优点，能较好地满足快速迭代开发、业务灵活扩展的需求，将高度耦合的功能分解到各个离散的微服务中，可以实现对应用系

统的解耦，解决传统的单块式系统架构随着业务需求增长出现的交付周期长、灵活性降低、可伸缩性差、维护成本高等问题。

本文主要介绍基于分布式和微服务架构建设的华信气象服务云平台，其应用了传统的数据库、通信、并行计算等技术，并结合新兴的云计算、大数据技术，以实现面向不同行业的实时气象服务可定制、可扩展，为气象服务产品开发和实时服务提供数据环境和平台支撑。

1 平台简介

华信气象服务云平台的主要功能包括各种气象数据的收集、传输、存储管理，服务产品加工处理，以及数据服务与监控管理相关功能，在统一气象数据环境基础上实现面向不同行业应用的气象服务，其主要功能包括：

1) 数据收集、传输：负责定时或根据请求从数据源收集业务所需的各类数据，并传输到公有云平台上。数据收集方式支持基于数据库API接口、文件共享、FTP共享、Web Service等多种数据收集方式。

2) 数据存储管理：负责对收集到的各种数据进行入库存储，并提供数据备份、数据恢复、数据清除等功能。各种观测数据、产品数据采用关系型数据库和文件库相结合的形式进行存储。

3) 服务产品加工处理：根据不同领域气象服务用户的需求，对地面、高空、农气、雷达、卫星、数值预报产品、气象服务产品等各种气象基础数据结合行业信息进行加工处理，生成不同服务用户所需的定制气象服务产品。

4) 数据服务发布：以门户网站的形式提供服务数据发现导航、数据查询检索、数据下载功能，并提供数据和产品实时服务API接口以及数据FTP下载服务。

5) 平台监控管理: 实现平台业务和数据监控、用户管理等功能。

1.1 平台数据环境

平台通过建设统一气象数据环境支撑面向不同行业应用的气象产品定制服务, 主要业务数据包括地

面、高空、农气、雷达、卫星等基础观测资料, 多种数值预报模式产品、融合分析产品、天气预报产品, 以及平台加工制作的各种气象服务产品, 日增数据量约2.2 T。表1给出了平台数据环境中包含的数据类型、名称、更新频次、空间范围以及日增量。

表1 华信气象服务云平台数据环境说明

数据类型	数据名称	更新频次	空间范围	日增量
地面观测资料	全球(含中国)地面逐小时观测数据	1 h	全球	约8 G
	中国地面分钟自动观测数据	1~10 min	中国	
	中国地面观测日值数据	1 d	中国	
	中国地面观测旬值数据	1旬	中国	
	中国地面观测月值数据	1月	中国	
高空气象资料	中国气球探空资料	2次/d	中国	约13 G
	GNSS/MET高空水汽观测资料	1 h	中国	
	飞机AMDAR气象观测资料	不定时	中国	
	闪电定位观测资料	有闪电时10 min一次	中国	
农业气象资料	自动站土壤水分观测资料	1 h	中国	约2 M
	农作物生长发育观测资料	1旬	中国	
	农业气象灾害观测资料	1旬	中国	
雷达气象资料	新一代多普勒天气雷达基本产品及反演产品	1次/6 min	中国	约1.2 T
	全国以及分区多普勒天气雷达组网拼图产品	1次/6 min	全国以及8大区域	
卫星气象资料	FY-2E、FY-2F、FY-2G云图产品	2次/h(加密观测4次/h)	中国大陆、海区	约416 G
	FY-4A云图产品	5~15 min/次	中国大陆、海区	
	FY-3云及大气廓线反演产品	2次/d	全球	
	FY-2E、FY-2F、FY-2G反演降水、云、辐射等产品	2次/h	中国大陆	
	日本葵花-8(Himawari)卫星圆盘图	1次/10 min	东亚	
数值模式预报产品	T639东北半球中期数值预报模式产品	2次/d	东北半球	约208 G
	GRAPES_MESO中尺度数值预报产品	2次/d	中国范围	
	欧洲中心(ECWMF)数值预报模式产品	2次/d	全球	
	美国NCEP全球预报系统(GFS)产品	2次/d	全球	
	日本全球数值预报模式产品	2次/d	全球	
气象分析与预报产品	全国三源融合降水格点分析产品	1 h	中国	约285 G
	全国地面同化气象要素格点分析产品	1 h	中国	
	全国城镇精细化预报指导产品	2次/d	中国	
	灾害性天气预报产品	3次/d	中国	
	全国精细化逐小时格点预报产品	2次/d	中国	
	强对流天气短时临近预报产品	10 min	中国	
	台风路径监测与预报数据	1 h	西太平洋海域	
行业气象服务产品	机场跑道自动观测数据	1~5 min	中国	约128 G
	航空气象服务产品	1次/10 min~2次/d	中国	
	电力气象服务产品	1次/10 min~2次/d	中国	
	环保气象服务产品	1 h	中国	
	保险气象服务产品	1 h~1旬	中国	
	农业气象服务产品	1 h~1旬	中国	

1.2 公有云环境

随着气象服务业务的不断发展, 气象服务平台需要随时扩充服务器、存储等基础设施支撑资源, 而技术成熟的公有云具有按需服务、弹性计算、异地灾备等特性, 可以满足气象服务平台基

础资源扩展的需求。因此, 华信气象服务云平台部署在公有云上, 由公有云提供计算、存储与网络资源支撑。

公有云环境现有资源配置及用途说明见表2, 未来可能随着服务业务量增加补充公有云资源。

表2 平台公有云资源说明

分类	设备名称	数量	配置	用途
服务器资源 (虚拟机)	数据接收服务器	2台	8核CPU, 16 G内存, 340 G高效云盘, 带宽50 Mbps	接收传输上云的服务所需的各种气象数据和产品
	数据加工处理服务器	4台	2台8核CPU, 64 G内存, 4 T高效云盘, 带宽50Mbps; 2台32核CPU, 64 G内存, 140 G高效云盘, 带宽100 Mbps	部署数据加工处理系统, 实现数据解析、加工处理、服务产品加工等功能
	数据存储管理服务器	1台	8核CPU, 64 G内存, 380 G高效云盘, 带宽50 Mbps	部署数据存储管理系统, 实现数据入库存储、数据管理等
	Hadoop服务器集群	6台	2台管理节点8核CPU, 64 G内存, 680 G高效云盘, 带宽50 Mbps; 4台数据节点8核CPU, 64 G内存, 1 T高效云盘, 带宽20 Mbps	部署HBase集群、Redis内存数据库、Elastic Search分布式全文检索系统等, 实现数据分布式存储和检索服务
	MySQL集群服务器	6台	2台管理节点32核CPU, 64 G内存, 1700 G高效云盘, 带宽50 Mbps; 4台数据节点8核CPU, 16 G内存, 3700 G高效云盘, 带宽50 Mbps	部署MySQL集群, 包括2个管理节点和4个数据节点, 实现结构化数据分布式存储
	API接口服务器	2台	4核CPU, 16 G内存, 100 G高效云盘, 带宽50 Mbps	部署各种数据和产品的API接口
	Web服务器	2台	4核CPU, 16 G内存, 380 G高效云盘, 带宽50 Mbps	部署Web服务器中间件, 气象大数据共享服务平台软件
	平台监控管理服务器	1台	8核CPU, 16 G内存, 340 G高效云盘, 带宽20 Mbps	部署平台监控管理系统
存储资源	对象存储OSS资源及云盘	50 T	50 T云盘与对象存储OSS资源	存储非结构化气象数据文件
网络资源	网络及安全设备、负载均衡	按需		实现数据网络传输、安全保障以及网络负载均衡

2 平台架构设计

华信气象服务云平台基于公有云提供的IaaS资源

进行建设, 采用B/S架构进行设计, 其总体技术架构

见图1。

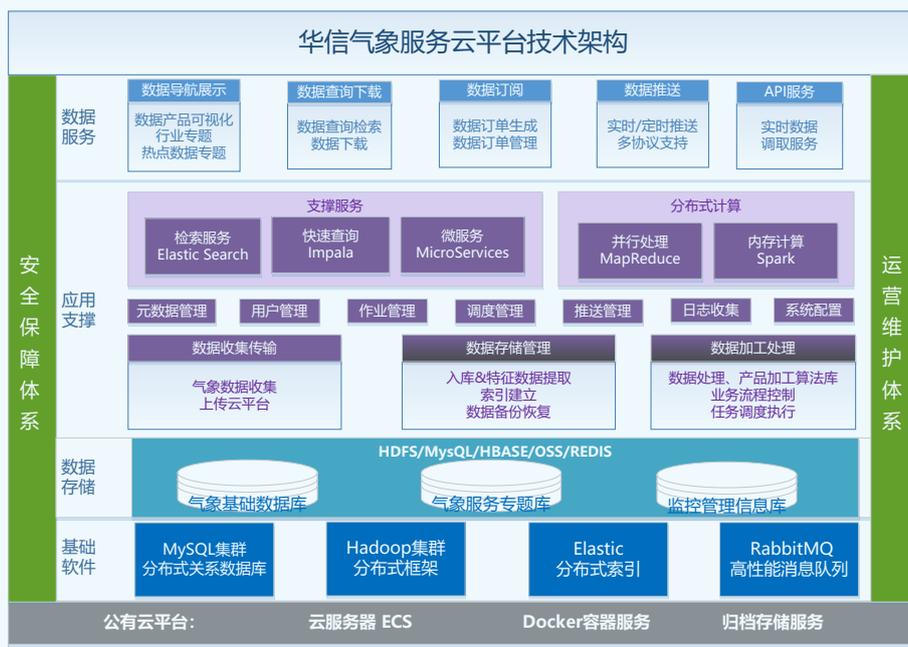


图1 华信气象服务云平台总体技术架构图

华信气象服务云平台采用的主要技术路线如下:

1) 基于公有云平台提供的虚拟化基础设施资源部署气象服务平台, 采用基于Spring Cloud的微服务架构, 结合Docker容器技术, 按服务数据类型与服务对象开发集成面向不同应用领域的气象服务功能, 实现服务业务的可伸缩、可灵活扩展。

2) 平台采用J2EE技术框架进行服务平台构建, 以层次化、组件化的方式进行平台总体架构设计, 集成元数据技术、中间件技术(包括Web中间件、消息

中间件等)等多种技术, 遵循统一的技术标准进行服务平台开发。

3) 采用Hadoop分布式处理框架进行云上的分布式平台构建, 采用分布式MySQL数据库集群、HBASE(Hadoop Database)分布式存储系统集群以及HDFS分布式存储文件系统结合分布式对象存储系统OSS存储管理各种结构化与非结构化气象服务数据与产品, 实现分布式的高容错数据存储; 采用Map Reduce并行处理技术以及Spark内存计算技术实现多种

并发数据的分布式处理。

4) 采用基于Lucene作为内部引擎的Elastic Search分布式全文检索系统实现大规模气象数据的快速检索, 采用Impala提供的基于HDFS和OSS的分布式快速查询引擎, 为整个系统提供基于分布式存储的快速数据访问服务。

5) 采用REDIS数据缓存技术, 将高频访问的实时服务数据以一定的数据结构保存在REDIS内存数据

库中, 提高实时更新数据的查询、页面响应、显示及应用效率, 满足实时数据服务的时效性要求。

6) 采用RabbitMQ高性能消息队列, 对多任务并发调度信息进行有序管理, 实现高可用的多并发任务高效调度管理。

与采用传统单体结构的普通服务平台相比, 采用分布式和微服务架构建设的服务平台具有的优势见表3。

表3 传统服务平台与分布式结合微服务架构平台性能对比

对比项	传统单体机构服务平台	分布式结合微服务架构平台
系统稳定性	发生故障影响范围大, 一个业务模块故障可能影响整个系统运行, 故障不易隔离	发生故障影响范围小, 一个业务模块故障不会影响其他业务模块以及系统整体运行
可扩展性	系统可扩展性差, 新增功能可能影响系统的整体框架, 需要重新部署; 硬件环境扩展性差, 一般业务运行过程中不能按需增加计算资源或存储资源	系统可扩展性好, 新增功能不会影响系统的整体框架, 微服务Docker容器封装新增功能模块可以在不影响系统及现有业务模块运行的情况下进行新增业务部署, 真正实现服务“热插拔”; 硬件环境扩展性好, 分布式架构可以根据业务需要按需增加计算资源或存储资源
时效性	数据量小、吞吐量小时系统操作响应快, 数据量大、吞吐量大、数据处理复杂时响应时间长	数据量小、吞吐量小时系统操作响应没有传统系统快, 数据量大、吞吐量大、数据处理复杂时时效性更好
系统开发	系统一次性开发时间长, 技术单一且封闭, 需要整体开发和部署	系统框架复杂, 技术多样且开放, 功能组件可以根据业务需要随时开发和部署
系统可维护性	系统不能经常更新, 部署复杂	系统可以经常更新, 单个服务容易开发和部署, 部署快速、简单

3 关键技术

为满足平台资源按需使用、气象服务业务不断扩展且新增业务不影响已有服务等灵活、可扩展性要求, 华信气象数据服务云平台采用分布式技术与微服务架构结合进行设计和开发, 以满足为不同服务对象提供可伸缩、可扩展、灵活定制的个性化气象数据服务的需求。

3.1 分布式技术

华信气象数据服务云平台分布式架构设计如图2。平台采用Hadoop分布式处理框架以可靠、高效、可伸缩的方式提供平台分布式支撑, 采用分布式存储技术管理气象服务数据与产品——地面、高空等结构化数据采用分布式MySQL数据库集群、HBASE集群进行存储管理, 雷达、卫星、数值预报模式产品等非结构

化数据采用HDFS分布式存储文件系统结合分布式对象存储系统OSS实现分布式的高容错数据存储, 基于Elastic Search的分布式索引系统提供对各类数据文件的索引和高效检索, 在此基础上采用Map Reduce并行处理技术以及Spark内存计算技术实现多种并发数据的分布式处理, 生成的服务产品在保存到专题产品库的同时通过API接口、FTP、推送服务等提供给不同的用户或对接用户的业务应用系统, 满足海量气象数据存储、检索服务、产品加工处理的高效、可靠、灵活等各种性能要求。

3.1.1 分布式存储

平台采用文件系统与关系型数据库相结合的分布式存储技术, 针对不同数据特点, 分别将不同类型的气象数据存储于文件系统或关系型数据库中。如图3所示, 平台将分布式对象存储规划为若干个存储空间, 按基础数据库、专题产品库、监控管理库及归档库对海量气象数据和产品进行分布式存储, 每个存储

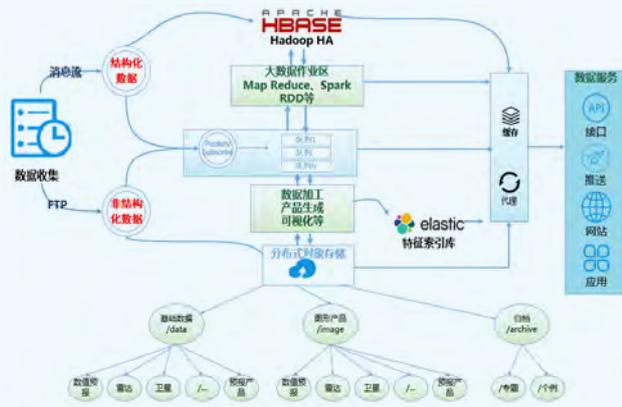


图2 华信气象服务云平台分布式架构图



图3 华信气象服务云平台分布式存储架构图

空间可以根据业务需要定义权限、生命周期等。其中，采用分布式存储系统HBASE与MySQL数据库集群结合存储各种关系型数据，采用分布式对象存储系统OSS存储各种非结构化数据，基于Elastic Search的分布式索引系统和Impala快速查询引擎提供对各类数据文件的索引和高效检索。

结构化数据分布式存储：为提高海量结构化气象数据实时服务的时效性，需要根据不同类型气象数据的特点和应用瓶颈，进行适当的数据分片与集群化部署。对于结构化数据中服务时效要求高的分钟级实况与预报预警数据（表4），平台采用高可靠性、高性能、面向列、可伸缩的分布式存储系统HBase进行存储，而采用MySQL数据库集群存储地面、高空、农气、台风路径监测预报以及精细化预报产品等结构化数据。MySQL集群分布式存储的结构化数据访问通过分布式集群中间件（图4）实现，中间件将SQL语句转换成对应的语法规析树，再结合预定义的分片及路由配置通过SQL路由/计划组件将其转换成执行计划，分发到相应的MySQL数据节点，MySQL节点执行后的结果通过中间件合并计算生成最后结果返回给应用。

表4 HBase分布式存储数据说明

数据类型	数据种类	数据频次	空间范围	日增量
地面观测资料	中国地面分钟自动观测数据	1~10 min	中国2万站左右	约10 G
	机场跑道自动观测数据	1~5 min	中国约200个机场	
气象预报产品	强对流天气短时临近预报产品	10 min	中国	约1.5 G
行业气象服务产品	全国机场危险天气短时临近预报产品	10 min	中国	约120 M

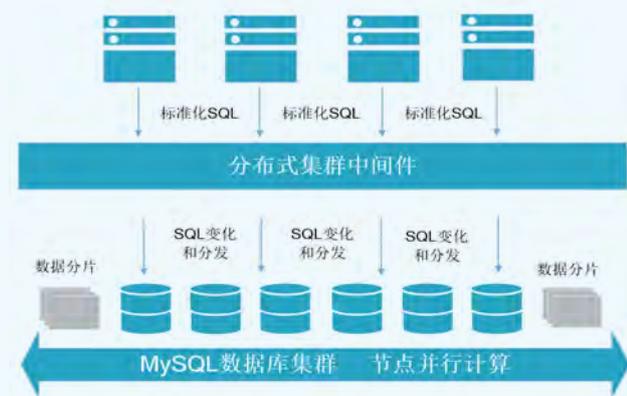


图4 结构化数据分布式存储MySQL集群逻辑拓扑图

图5出示了非结构化数据分布式存储及处理逻辑拓扑。平台中的雷达、卫星、数值模式预报产品、天气分析图等各种文件类、图形类、文档类非结构

化数据，采用基于Swift技术的分布式对象存储系统OSS进行存储管理，检索查询采用分布式全文检索系统Elastic Search，所有非结构化数据入库前提取的特征索引数据均存储于分布式全文检索系统。系统提供Impala快速查询引擎以实现对数据的分析和快速查询。用户的检索请求会被分解为多个并行检索任务，下发到多个节点并行执行，系统会将多个并行检索的结果进行合并、返回给用户，实现对各类数据文件的高效检索和服务。

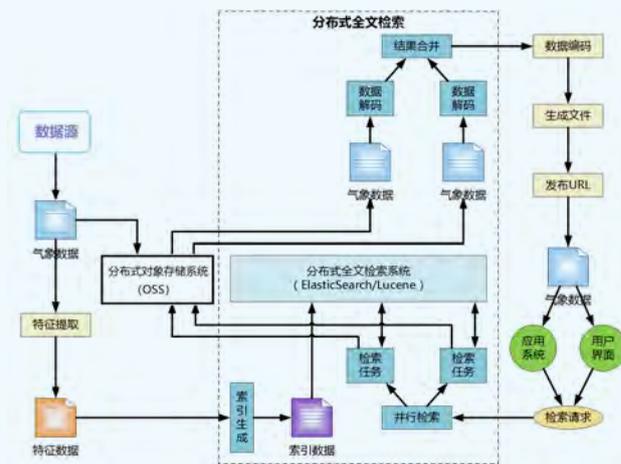


图5 非结构化数据分布式存储及处理逻辑拓扑图

3.1.2 分布式处理

由于气象数据具有数据类型多、时空尺度跨度大等特点，华信气象数据服务云平台常常需要同时加工处理雷达、地面、卫星、模式等不同类型的 meteorological data，将其加工制作成符合不同行业需求的服务产品，因此需要平台具有多任务并发处理能力，避免任务堆积。平台采用开源云平台Hadoop架构的分布式计算平台开发和运行处理海量气象数据的应用程序，采用分布式并行处理模型Map Reduce以及Spark内存计算技术实现海量气象数据并发的分布式处理，保证各种气象产品加工处理与服务的时效性。

Map Reduce作为面向大数据并行处理的计算模型、框架和平台，可以充分利用集群资源，自动完成计算任务的并行化处理，自动划分计算数据和计算任务，在集群节点上自动分配和执行任务以及收集计算结果，将数据分布存储、数据通信、容错处理等并行计算涉及的很多系统底层的复杂细节交由系统负责处理。Map Reduce提供了一种简便的并行程序设计方法，用Map和Reduce两个函数编程实现基本的并行处理任务，提供了抽象的操作和并行编程接口，以简单方便地完成大规模数据的编程和计算处理。

Spark是专为大规模数据处理而设计的快速通用

的计算引擎，其优点除了和Hadoop Map Reduce一样高效、高可靠、容错和可扩展之外，还可以通过启用弹性分布数据集RDD（Resilient Distributed Dataset）将中间输出结果保存在内存中，从而不再需要读写HDFS就可以实现交互式查询，优化迭代工作负载，提高计算效率。

下面以保存在HBase分布式存储系统中的地面分钟级观测降水数据分时段统计分析为例，说明采用Spark如何实现气象数据的分布式加工处理。Spark的运行架构由Driver（控制器）和Executor（执行端）组成，Driver负责把用户代码进行切分，划分为不同的任务阶段Stage，然后把每个Stage对应的任务调度提交到多个Executor进行计算，这样Executor就并行执

行同一个Stage的任务。如图6所示，地面分钟级观测降水数据分时段统计分析可以分为数据读取、处理与结果存储两个Stage，每个Executor负责根据Driver划分的阶段按流程执行相应的任务：首先从HBase分布式数据库中读取各个测站分钟级降水观测数据，进行质量控制、均一化处理，然后进行从1 min、5 min、10 min、15 min、20 min、30 min、45 min、1 h……到1 d的不同时间窗的降水数据处理，并分不同时段进行排序，最后生成每个站不同时段的最大降水量排序产品，返回给Driver进行结果组合并保存到HDFS分布式数据系统中。数据处理过程的中间数据通过分布式数据的逻辑抽象RDD进行缓存和流转，通过减少数据IO次数提高数据处理效率。

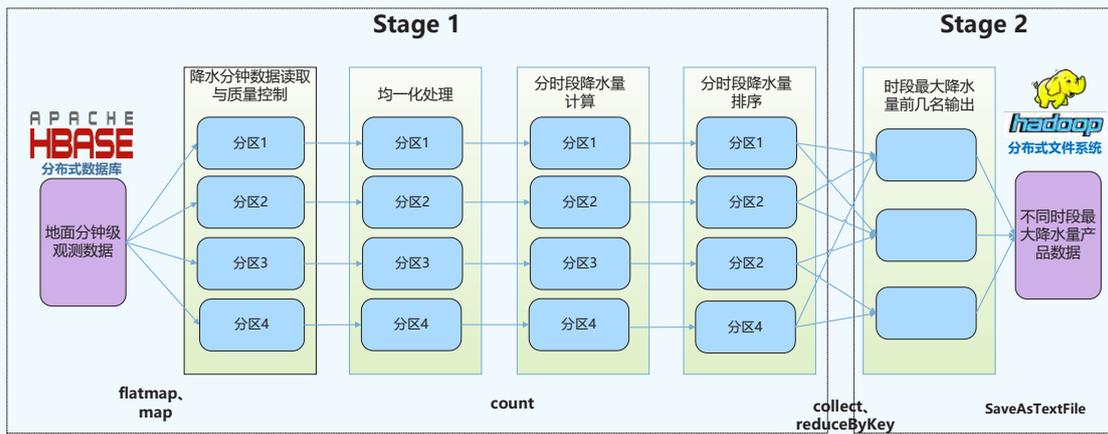


图6 基于Spark降水数据统计分析逻辑流程图

3.2 微服务架构

微服务架构是一种去中心化的系统架构，本质上也是一种分布式架构。微服务的核心思想是将应用功能尽可能细地划分成独立的服务，每个服务都专注于单一具体的业务功能实现，服务之间采用轻量级的通信机制（通常是基于HTTP的RESTful API）互相协调、互相配合，最终为用户提供符合不同需求的功能服务。微服务架构的目的是将大型的、复杂的、长期运行的系统构建为一组相互配合的服务，一个大型复杂软件每个应用都可以由一个或多个微服务组成。系统中的每个微服务都可被独立部署，各个微服务之间是松耦合的，每个微服务仅关注于完成一件任务并很好地完成该任务，在所有情况下，每个任务代表着一个小的业务能力。与传统的面向服务的架构（SOA）架构相比，微服务架构降低了系统的耦合性，提高了系统容错性、伸缩性、可扩展性，易于迭代开发、业务扩展与持续交付。

在实际应用中，微服务架构的服务拆分粒度在

尽量细化的同时还要与具体业务相结合，总的原则是服务内部高内聚，服务之间低耦合。如图7所示，华信气象服务云平台的微服务主要按照产品服务类型或服务对象进行划分，将基础气象数据地面、高空、辐射、海洋、农业气象、雷达、卫星、数值预报产品、气象服务产品等数据加工处理分别划分成微服务，或

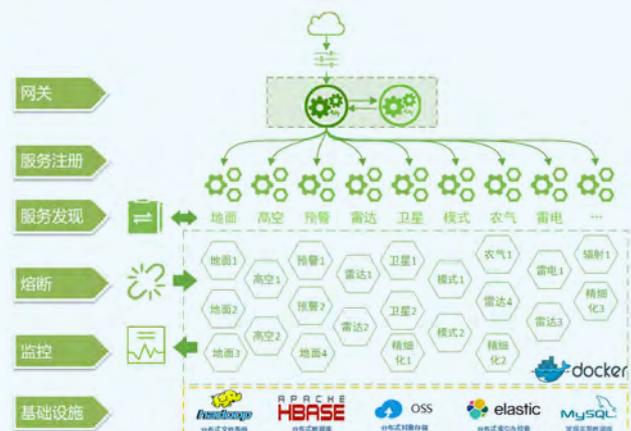


图7 华信气象服务云平台微服务架构图

根据气象数据服务对象的需求将不同类型的服务产品加工制作划分成相应的微服务，各服务之间彼此独立，也可以根据业务需求进行服务调用。

华信气象服务云平台的微服务由分布式MySQL数据库集群、HBase集群、HDFS分布式存储文件系统、分布式对象存储系统OSS结合Elastic Search的分布式索引系统提供基础数据支撑，将面向不同服务应用的气象产品微服务功能封装到可虚拟化部署的Docker容器中，通过Spring Cloud实现服务发现注册、配置管理、服务调用、服务网关、负载均衡、熔断、数据监控等功能，构建生产级的灵活扩展微服务应用，方便地进行气象服务的持续开发和集成。

3.2.1 Docker 容器

Docker是一种新兴的轻量级虚拟化容器技术，在复杂的集群系统搭建方面，具有可移植、易使用、跨平台等优势。所以，在复杂的分布式应用集群的部署中，Docker能够快速、准确、标准化封装应用程序并自动化部署整个运行环境。采用Docker容器在云平台上进行微服务的部署，可以实现系统从开发、部署到运维整个过程的微服务化，充分利用服务器的资源，构建高伸缩性的数据服务，实现扩展业务功能的快速部署、测试及上线业务运行。

平台采用Docker容器分别封装地面、高空、辐射、海洋、农业气象、雷达、卫星、数值预报产品等基础气象数据解析、转换、统计加工、绘图等服务功能，以及面向航空的积冰、颠簸、机场危险天气短临预报、对流指数预报等航空气象服务产品加工处理功能，面向环保的大气稳定度产品加工处理功能，以及面向电力的台风灾害监测预警服务功能等等微服务，这些服务功能应用C、Fortran、Java、python等不同编程语言开发，或应用到NCL等绘图工具包，业务部署时可以将应用及相关开发环境一起封装到可移植且相互独立的Docker容器中，实现在虚拟化公有云环境中快速简单的服务部署。

3.2.2 Spring Cloud

Spring Cloud 是从成熟的Spring framework上发展起来的，是基于Spring Boot实现的服务治理工具包，是一系列框架的有序集合，可以为微服务构建提供标准化的、全站式的技术方案。它利用Spring Boot的开发便利性巧妙地简化了分布式系统基础功能的开发，如服务发现注册、配置中心、消息总线、负载均衡、熔断器、数据监控等，都可以用Spring Boot的开发风格做到一键启动和部署。Spring将目前各家公司开发的比较成熟、经得起实际考验的服务框架组合起来，

通过Spring Boot风格进行再封装，屏蔽掉了复杂的配置和实现原理，给开发者留出了一套简单易懂、易部署和易维护的分布式系统开发工具包。

华信气象服务云平台的微服务构建主要应用到Spring Cloud的核心功能组件包括：分布式/版本化配置管理（Spring Cloud Config）、服务注册和发现（Eureka）、服务调用（Feign）、熔断器（Hystrix）、API网关（Zuul）、负载均衡（Keepalived）、数据监控（Actuator）。Spring Cloud各个组件可以插拔形式集成，各组件相互配合，合作支持了一套完整的微服务架构：Spring Cloud Config提供统一的配置中心服务，实现分布式系统的配置文件的统一管理；Eureka负责服务的注册与发现，避免服务之间的直接调用，方便服务后续的水平扩展、故障转移，将各服务连接起来并保持服务高可用；Feign负责服务之间通过RESTful API方式进行声明式的调用，网关使用Feign做数据验证；Hystrix负责监控服务之间的调用情况，连续多次失败进行熔断保护，并按一定间隔时间检查调用失败的服务，如果服务恢复将继续提供服务；轻量级网关Zuul负责服务转发，接收并转发所有内外部的客户端调用，实现相关的认证逻辑从而简化内部服务之间相互调用的复杂度；Keepalived负责在高并发访问时进行服务负载均衡，加强网络数据处理能力、提高网络的灵活性和可用性；Actuator负责监控服务间的调用和熔断相关指标，保证数据服务业务的连续性、可靠性。

3.2.3 RESTful API

RESTful API 是满足REST（REpresentational State Transfer）架构约束条件和原则的API，RESTful可以通过一套统一协议规范为不同形式的应用客户端提供统一的服务接口，不同的应用终端可以通过一个标准的Restful API 来访问数据库从而获取相同的数据。

平台中部分业务应用功能是通过多个微服务协调配合实现的，每个微服务都提供相应的轻量级RESTful API，服务之间通过调用API实现业务协同。例如，环保大气稳定度产品加工处理服务需要应用风云静止卫星的云量反演产品结合太阳高度角计算太阳辐射等级，再结合地面风观测数据进行大气稳定度判断。因此，大气稳定度产品加工处理微服务需要调用风云静止卫星的云量反演产品服务API获取云量数据、调用地面逐小时观测数据服务API获取地面风观测数据，再根据算法计算得到大气稳定度判断结果，最终产品也通过RESTful API为用户提供服务。

3.2.4 微服务封装与构建

华信气象服务云平台中，微服务封装主要根据具体的业务场景将服务流程细化拆分成一系列的微服务进行封装，各个微服务之间通过消息总线、负载均衡、配置中心等微服务框架协调计算资源进行处理并实现业务所需的服务。具体到技术层面，一个微服务的表现形式就是同一类型数据（产品）的一组 RESTful Web Service 的组合。因此，构建一个微服务时，首先基于划分好的服务类型或服务对象做API层面的定义，然后在业务层实现本地Repository和其他服务的聚合，数据访问层根据存储系统类型实现各自的Data Access Object，从而以Web Service形式提供数据或产品服务。其中，业务层聚合微服务可以借助Spring Cloud的Feign组件实现，如果调用服务不可用，则借助Hystrix组件结合业务要求实现熔断或服务降级，防止长时间等待导致故障蔓延、服务雪崩。此外，微服务需要开启服务发现客户端即Enable Discovery Client，并配置Eureka Server集群地址，实现自动化的服务注册与发现。

以环保大气稳定度产品加工处理服务为例，该服务由太阳辐射等级计算服务、大气稳定度判断服务以及大气稳定度接口服务三个微服务组成，服务构建与执行流程如下：

1) 太阳辐射等级计算微服务启动后通过Eureka注册服务，并从配置中心Spring Cloud Config获取配置，通过Feign组件由网关Zuul调用风云静止卫星的云量反演产品服务API获取云量数据，计算得到太阳辐射等级结果，以MQ消息方式发送到消息总线。

2) 大气稳定度判断微服务启动后通过Eureka注册服务，从配置中心获取配置后通过Feign组件聚合太阳辐射等级计算服务，从而不断获取辐射等级结果并通过网关Zuul调用地面逐小时观测数据服务API获取地面风观测数据，进行大气稳定度判断后将判断结果入库存储。

3) 大气稳定度判断结果通过启动大气稳定度接

口微服务，以API接口方式提供给用户。

其中，太阳辐射等级计算、大气稳定度判断微服务根据任务量可以动态启动多个服务实例，通过Hystrix组件对每个服务实例设置熔断阈值保证每一个服务实例服务正常，通过Keepalived负载均衡保证整个计算流程更加合理地利用计算资源，两者结合使得整个服务更加高可用。

对于大气稳定度接口微服务，可以根据产品访问量动态启动不同数量的接口服务实例，通过负载均衡使得接口服务请求均匀转发到每个服务，保证快速高效地为用户提供秒级产品服务。

4 结语

本文介绍了应用分布式技术基于微服务架构建设的华信气象服务云平台。平台采用Hadoop分布式处理框架构建云上的分布式平台，结合分布式存储、分布式处理技术实现海量气象数据的高效检索服务和产品加工处理；基于Spring Cloud的微服务架构提供灵活可扩展的数据服务，将面向不同服务应用的气象产品微服务功能封装到可虚拟化快速部署的Docker容器中，构建高伸缩性的数据服务，方便地进行气象服务的持续开发和集成，满足不断拓展的多种多样的气象数据服务需求。

深入阅读

- 李春阳, 刘迪, 崔蔚, 等, 2017. 基于微服务架构的统一应用开发平台. 计算机系统应用, 26(4): 43-48.
- 沈文海, 2014. 气象数据的“大数据应用”浅析——《大数据时代》思维变革的适用性探讨. 中国信息化, 30(11): 20-31.
- 张洁, 薛胜军, 2016. 云计算环境下气象大数据服务的应用研究. 安徽农业科学, 44(5): 298-301.
- 张晶, 王琰洁, 黄小锋, 2017. 一种微服务框架的实现, 计算机系统应用, 26(4): 82-86.
- 赵乐乐, 黄刚, 马越, 2016. 基于Docker的Hadoop平台架构研究. 计算机技术与发展, 26(9): 99-103.

(作者单位: 崔宏, 吴恩平, 王尊鹏, 赵笔锋, 龚伟伟, 华云信息技术工程有限公司; 熊安元, 邓鑫, 国家气象信息中心)