

何林,吉庆,燕东渭,等.“秦云工程”气象数据对接

文章编号:1006-4354(2020)01-0047-04

“秦云工程”气象数据对接的设计与实现

何林^{1,2},吉庆³,燕东渭²,李亚丽²,王垒²

(1. 秦岭和黄土高原生态环境气象重点实验室,西安 710016;

2. 陕西省气象信息中心,西安 710014;

3. 渭南市气象局,陕西渭南 714000)

摘要:依托以 CIMISS 为核心的省级集约化气象数据环境,通过架构分析、接口定制、数据同步、存储和数据服务等关键技术环节的研究,实现“秦云工程”气象云的数据对接。为开展其他行业大数据接入“秦云工程”并进行交换和共享,具有一定的指导与参考意义。

关键词:秦云工程;大数据;气象数据;CIMISS;服务接口;交换共享

中图分类号:P409

文献标识码:B

在大数据技术及产业高速发展的背景下,以政府牵头、惠及民生为目标的行业数据交换共享与融合应用成为一个重要发展方向。“秦云工程”即陕西省大数据产业的“N+1”云工程,“N”是 22 朵行业云,“1”是大数据交换共享平台,是“秦云工程”建设的核心。大数据交换共享平台承载了行业云的数据交换、共享、开放、服务四大功能体系,在保证政府数据流通性和安全性的前提下,实现行业数据汇聚交换、互联互通和开放共享^[1-2]。气象云作为“秦云工程”22 朵“行业云”之一,通过大数据交换共享平台与环保、水利、交通、旅游其他行业云的数据进行融合分析,充分挖掘和发挥气象数据的应用效益。对此,研究数据对接技术实现气象数据高效地接入“秦云工程”是首要解决的问题。

1 设计原则

(1)保证气象数据使用的安全性。由于气象观测数据中涉及地理坐标等敏感信息,需严格遵循中国气象局第 4 号令《气象资料共享管理办法》,避免核心数据、加密数据、历史数据开放后,对国家安全造成危害^[3]。

(2)保证气象数据更新的及时性。按照数据

的观测频次,设计合理的同步策略,保证最新的观测数据在最小的时间损耗下接入气象云,为行业数据融合与应用提供高时效的数据服务。

(3)保证气象数据接入的可靠性。在“秦云工程”气象数据采集和交换的过程中,不对气象部门内现有各类常规业务的顺利开展、信息系统的稳定运行造成影响。

2 技术架构

“秦云工程”气象云的数据来源为全国综合气象信息共享平台(China integrated meteorological information sharing system,下简称“CIMISS”)。作为气象部门建立规范的国省两级气象数据环境,CIMISS 将各类气象数据纳入集约化管理,为气象业务和科研提供了权威、丰富的数据源,为气象大数据的应用与研究奠定了基础^[4]。由于存在网络隔离,需要在“秦云工程”和气象业务局域网之间搭建专线网络来保障数据交换通畅。实际的数据交换由部署在局域网内的一台前置机完成,包括气象数据的采集及推送代理。其中,数据采集代理通过调用 CIMISS 气象数据统一服务接口(meteorological unified service in-

收稿日期:2019-08-20

作者简介:何林(1987—),男,汉族,陕西武功人,工程师,硕士,从事气象信息软件研发、集约化气象数据环境建设。

基金项目:秦岭和黄土高原生态环境气象重点实验室开放研究青年基金课题(2019Y-6)

terface community, 下简称“MUSIC”), 据工程—气象云”所需的数据同步至前置机; 数据推送代理则采用主动共享方式, 将气象数据接入“秦

据交换共享平台实现 22 朵“行业云”之间的数据交互。数据对接的技术架构如图 1 所示。

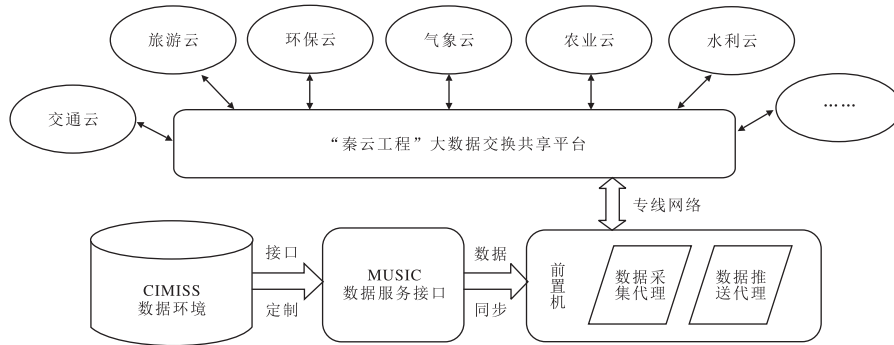


图 1 (秦云工程-气象云) 总体技术架构图

3 实现方案

3.1 接口定制

MUSIC 面向气象行业内部提供了开放、全量、标准的数据接入服务。而“秦云工程”面向社会大众, 为保证数据使用安全, 需要基于对 MUSIC 标准接口的定制, 实现对气象数据的适当裁

剪后才可共享。根据《基本气象资料和产品开放清单》向全社会开放共享的地面、高空、气象卫星、天气雷达、数值预报等 5 类 17 种基本气象资料和产品清单^[5], 结合“秦云工程”资料需求, 划定可共享的数据范围(表 1)。

表 1 “秦云工程 - 气象云”第一批核心资料共享范围

资料名称	CIMISS 资料编码	覆盖范围	时间范围
中国地面(国家站)逐小时观测资料	SURF_CHN_MUL_HOR_N	全省 88 站(不含延安等 11 站)	近 7 天
中国高空定时观测资料	UPAR_CHN_MUL_FTM	全省 3 站(不含安康)	近 3 天
多普勒雷达基数据资料	RADA_L2_UFMT	全省 7 部雷达	近 3 天

接口定制通过 MUSIC 后台管理平台进行操作, 其核心是更新接口元数据^[6], 实现对“秦云工程”API 账户的资料访问权限控制。设定该用户可访问的数据仅为表 1 所示的三类资料, 其他资料禁止访问。同时, 对三类资料时间、空间属性也须做裁剪, 通过修改元数据要素配置值, 划定明细的站号范围以及时间范围。以中国地面(国家站)逐小时观测资料的接口定制为例(表 2), 列出了主要的元数据配置项。

3.2 数据同步

通过调用定制的 MUSIC 接口, 研发数据采集代理程序, 并部署在前置机。前置机的最主要作用是屏蔽“秦云工程”数据对接对 CIMISS 业务库的影响, 同时避免了开通 CIMISS 公网端口访

表 2 MUSIC 接口定制的元数据配置示例

元数据要素名称	元数据要素配置值
MUSIC 资料编码	SURF_CHN_MUL_HOR_N
数据检索场景	按时间、地区检索地面要素数据
MUSIC 接口 ID	getSurfEleInRegionByTime
权限总控开关	USER_NAME='BEXA_QYGC' and ACCESS_CONTROL='Y'
时间控制类型	TIME_CONTROL='Y' and CONTROL_TYPE='R' and RECENT_DAYS='7'
空间控制类型	AREA_CONTROL='Y' and CONTROL_FIELD='V01301' and AREA_LIST='...'

问造成的安全隐患。此外,为保证气象数据更新,还需在前置机上根据不同资料的时效特点,配置不同的定时任务执行策略。以中国地面(国家站)逐小时观测资料为例,由于 CIMISS 中,当前时次整点数据的入库呈现类指数方式的增长,

01、02、03、05、07、10、30、59 分各定时执行一次。图 2 给出了通过调用 MUSIC 实现地面逐小时数据同步流程。

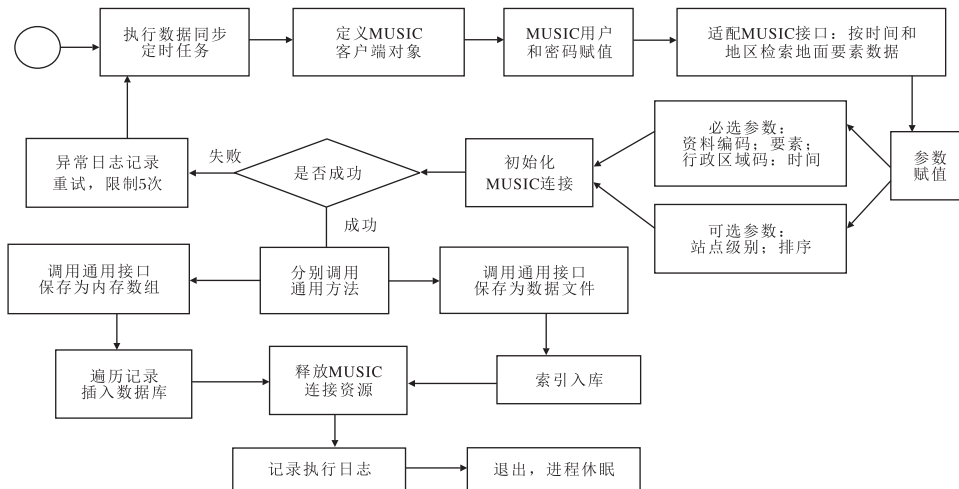


图 2 调用 MUSIC 实现数据同步流程

3.3 数据存储

数据同步到前置机进行存储。按照数据的组织结构特点,一般可分为结构化数据和非结构化数据两大类,其中非结构化数据含有自描述信息的半结构化数据^[7]。典型的结构化数据都以记录的方式存储在关系型数据库中,如 Oracle、MySQL 等。其他大部分数据都是以文档、图片等非结构化的形式存储,这些数据可直接保存于文件系统中^[8]。此外,为了提高访问效率和用户体验,对交互频次较高的数据,将其存储索引以记录的形式在 Redis 内存数据库中保留一份^[9]。前置机上的数据包括主动和被动两种共享方式。以前置机为中心,被动共享是指大数据共享交换平台可通过专线网络直接访问存储在前置机上的数据;主动共享是指通过前置机部署的数据推送代理,将数据推送至大数据交换共享平台以及其他对气象数据有需求“行业云”。两种方式皆可实现数据访问,本方案选取了主动共享方式。

3.4 数据服务

“秦云工程”大数据交换共享平台集结了已授权的行业云数据,提供了内容全面、功能丰富的数

据服务,满足各类用户多元化的数据服务需求。气象数据接入后,可通过大数据交换共享平台的开放接口获取相关数据服务,将气象数据进行可视化的展示。图 3 给出了一个大屏展示气象数据的截图。该应用为某政府部门通过“秦云工程”大数据交换共享平台调用气象数据进行服务的图例原型。

同时,通过大数据交换共享平台的代理服务,可将其他通过授权的行业数据,如环保、水文数据等推送至气象内网的前置机上,便于气象部门开展行业数据融合分析及相关应用研究。

4 结论与展望

通过采用本文设计的数据对接方案,目前,气象数据已成功接入“秦云工程”并形成示范。这不仅扩大了气象数据资源的共享开放力度,有助于打破政府部门间的“数据壁垒”,拆除“应用烟囱”^[10]。同时,与其他行业的数据交换、融合、分析,进一步推动了智慧气象的发展。未来,借助“秦云工程”,如何用好数据将是重点研究方向,只有将气象业务与大数据应用技术相结合,才能最大化发挥气象数据的效益。



图3 “秦云工程”陕西气象云大屏展示截图

参考文献:

- [1] 陕西省人民政府. 陕西省大数据与云计算产业示范工程实施方案;陕政办发〔2016〕50号[A]. 2016-06-17.
- [2] 王睿. 秦云工程数据交换共享平台上线[N]. 陕西日报, 2017-07-07(1).
- [3] 夏巧利,王小宁. 大数据时代陕西气象档案数据的管理与应用[J]. 陕西气象, 2017(2):46-48.
- [4] 何林,邓凤东,李亚丽,等. 陕西省气象数据共享系统[J]. 陕西气象, 2017(6):32-35.
- [5] 中国气象局. 基本气象资料和产品共享目录[Z]. 2015-09-29.
- [6] 吴彦平,董保华,张文龙,等. 河北省综合气象观测资料共享显示平台设计[J]. 现代电子技术, 2015(3):112-114.
- [7] 杨明,陈晔峰,陈晴,等. 气象数据云数据存储技术及应用[J]. 气象科技, 2017(6):1071-1021.
- [8] 涂俊英,李志敏. 云计算下非结构化大数据存储系统设计[J]. 现代电子技术, 2018(1):173-177.
- [9] 何林,薛春芳,邓凤东,等. 陕西省精细化气象格点预报数据环境的设计与实现[J]. 陕西气象, 2017(4):17-21.
- [10] 王垒,邸永强. “秦云工程”助推智慧气象发展探析[C]//陕西省通信学会 2016年学术年会论文集. 西安:陕西省通信学会, 2016:132-135.