

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.

文章编号:1006-4354(2017)04-0017-05

陕西省精细化气象格点预报数据环境的设计与实现

何林¹,薛春芳²,邓凤东¹,贾晨刚¹,王建鹏³

(1. 陕西省气象信息中心,西安 710014;2. 陕西省气象局,西安 710014;

3. 陕西省气象台,西安 710014)

摘要:为了有效支撑陕西省精细化气象格点预报业务运行,从底层架构和数据流分析入手,提出了基于 CIMISS 标准框架的精细化格点预报数据环境的设计方案。包括硬件平台、数据清单、数据存储、数据共享、数据管理等关键环节,并遵循“冷热分离、读写分离、集中化管理、多样化共享”的原则,给出了优化实现方法。

关键词:精细化格点预报;CIMISS;数据环境;共享管理

中图分类号:P409

文献标识码:B

随着社会经济的发展及信息化水平的不断提高,传统的城镇、乡镇站点预报已经不能完全适应防灾减灾工作的需求,以时空分辨率更高、要素定量化和多样化、预报精准、服务贴心、无缝隙、智能化为特征的精细化格点(智能网格)预报技术将是未来天气业务发展的主要方向。为了推进和支撑精细化格点预报技术研究及业务运行,需要搭建规范、合理的基础数据环境。该数据环境除了满

足预报业务的需求外,还需要符合全国综合气象信息共享平台(China integrated meteorological information sharing system,简称“CIMISS”)标准框架^[1],为天气业务数据资源纳入省级集约化数据环境奠定基础。

1 数据环境介绍

陕西省精细化气象格点预报数据环境旨在建立以基础数据接收、加工处理、存储管理、共享服

收稿日期:2017-03-15

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:陕西省气象局青年科研基金项目(2016Y-11);中国气象局气象预报业务关键技术发展专项 YBGJXM(2017)03-13;陕西省气象局研究型业务重点科研项目(2015Z-6)

酸雨,5次强酸雨主要发生在2007—2010年间。

(3) 汉台区酸雨日数和酸雨发生频率呈波动递增趋势。9—11月是酸雨发生频率最高的时段,其中9月最高,为24.8%;5月酸雨发生频率最小,为4.2%。

(4) 汉台区酸雨发生频率随着降水量级的增大而减少。酸雨日的地面主导风向是偏东风。

参考文献:

[1] 张新民,柴发合,王淑兰,等. 中国酸雨研究现状[J]. 环境科学研究,2010(5):527-532.

[2] 赵艳霞,侯青. 1993—2006年中国区域酸雨变化特征及成因分析[J]. 气象学报,2008,66(6):1032-1042.

[3] 中国气象局. 酸雨观测业务规范[M]. 北京:气象出版社,2005:23-24.

[4] 中国气象局. 地面气象观测规范[M]. 北京:气象出版社,2003:104-111.

[5] 高龙. 汉台区2006—2009年酸雨特征及策略建议[J]. 陕西气象,2013(3):25-27.

[6] 郑凤琴,孙崇智,于文杰,等. 城市酸雨预报方法及业务系统[J]. 气象科技,2006,34(6):684-687.

务功能于一体的标准数据流程及支撑平台。该流程和平台具备结构化和非结构化资料相结合、存取高效、管理便捷、交互性强、可快速扩展的特点；在保证数据质量和稳定性的前提下，不断提升访问效率。主要包括构建原始数据库、降尺度格点

数据库、格点背景场数据库、省市格点预报编辑库、预报产品库、解析产品库等六大模块。同时，在 CIMISS 的标准框架内实现数据接入和管理，并梳理整个精细化格点预报业务的数据清单及优化数据流程(图 1)。

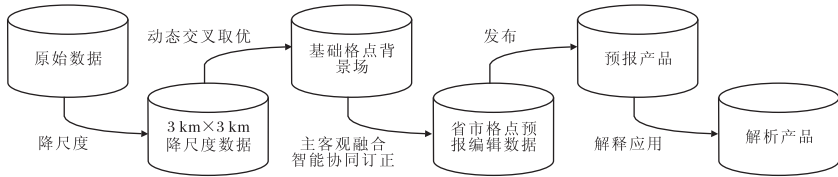


图 1 陕西精细化格点预报数据流程

2 基础平台架构

精细化格点预报数据环境基础平台部署在陕西省气象局自建的统一虚拟化资源池之上。此外，为满足精细化格点预报系统所存在的大量文件交互、高并发读写的个性化功能需求，在虚拟资源池中设立了逻辑专区，引入 VMware 公司发布的分布式存储技术。虚拟化网络存储空间(virtual store area network, 简称 VSAN), 是 VMware 全新的软件定义的存储层, 可扩展 vSphere 虚拟化管理程序以将计算和直连存储池化。通过建立服务器直连硬盘和固态硬盘集群, 创建专门针对虚拟机设计和优化的分布式共享数据存储^[2]。

通过资源池底层的改造, 构建了分布式的数据存储环境, 将精细化格点预报对数据的加工处理及存储全部迁移到基于 VSAN 的分布式平台上, 从而有效解决了数据环境的性能瓶颈。同时, 遵循“冷热分离”原则, 存储架构采用了混合模式^[3], 分为以实时处理、高频访问为特征的热数据区和以备份、回算、检验为目的的冷数据区。其中, 冷数据区建立在 SATA (serial advanced technology attachment) 硬盘构成 NAS 集中式存储之上, 热数据建立在以 SATA + SAS (serial attached SCSI) 硬盘构成的 VSAN 分布式存储之上。图 2 为基础平台架构图。

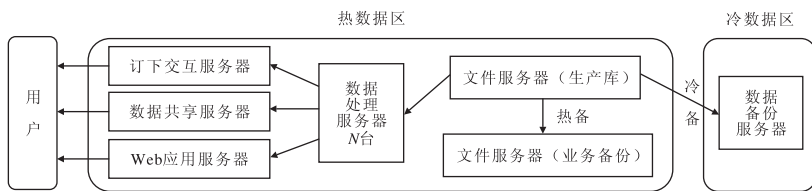


图 2 陕西精细化格点预报基础平台架构图

3 数据接入与数据清单

原始数据无论从种类还是数量, 都占据了整个数据环境的主体部分, 成为支撑精细化预报的基础环节。资料范围涵盖了精细化格点预报所需的地面观测站点实况资料, 国内外的数值预报模式或集合预报产品资料, 格点实况和预报产品, 雷达、卫星、闪电定位等资料。原始数据来源主要分为 CIMISS、CMACAST 卫星广播系统、国家气象信息中心实时数据共享平台、国家气象中心数据共享服务器、区域中心推送区域模式产品、国家数

值预报云等。接入方式包括 CIMISS 接口、FTP 下载、卫星广播等途径。通过对不同数据来源的分析对比, 选取最优的数据接入方式, 如表 1。

中间数据主要包括处理后不同格式的站点实况、对各类模式产品进行解码和降尺度处理后生成的数据。加工产品主要包括多模式动态交叉取优得到的格点背景场客观产品, 为地市共享的 MICAPS 格式的格点、站点预报产品, 省市预报员基于客观产品进行主观订正生成的预报产品, 最终确认预报产品, 预报解析产品(基于规则库一

表 1 原始数据清单及接入方式

资料大类	资料小类	频次	接入方式
地面 观测 资料	中国地面逐小时资料	每小时 1 次	CIMISS
	中国地面日值	每日 1 次	CIMISS
	中国地面月值	每月 1 次	CIMISS
	中国地面年值	每年 1 次	CIMISS
	中国地面分钟资料	每 10 分钟 1 次	CIMISS
模式 产品 资料	EC 细网格/粗网格/EFI 指数/卫星模拟产品	每日 2 次	CMACAST
	NCEP 细网格	每日 2 次	CMACAST
	RJTD 细网格/粗网格	每日 4 次	CMACAST
	T639 细网格	每日 4 次	CMACAST
	精细化气象要素客观预报系统 MEOFIS 下发产品	每日 2 次	外网 FTP
	GRAPES_MESO 资料	每日 2 次	外网 FTP
	GRAPES_RAFS 资料	每日 8 次	外网 FTP
	集合预报(EC/NECP/RJTD)	每日 2 次	内网 FTP
	CAUCE 环境气象预报模式产品	每日 2 次	内网 FTP
	气象中心下发 5 km 格点预报/实况	每小时 1 次	内网 FTP
	气象中心强对流天气概率预报产品	每小时 1 次	CMACAST
	西北区域中心中尺度数值模式产品	每日 2 次	内网 FTP
	WARMS 和 WRF 本地模式产品	每日 2 次	内网 FTP
雷达 资料	多普勒雷达基数据	每 6 分钟 1 次	CIMISS
	多普勒雷达 PUP 产品	每 6 分钟 1 次	CIMISS
	雷达 QPF 定量降水预报	每 6 分钟 1 次	内网 FTP
卫星 资料	日本葵花 8 圆盘图	每 10 分钟 1 次	内网 FTP
	FY-2G/2E 资料	每小时 1 次	CIMISS
其它 资料	闪电定位	不定时	CIMISS
	地面站点信息	不定时更新	CIMISS

键生成并发布的专题预报服务产品及材料), 预报检验数据等。

4 数据组织结构

遵循对数据“集中化管理”的原则, 将所有原始数据、中间数据、加工产品都存放在统一规划的服务器上, 只是数据组织结构和存储形式不同。

将精细化格点预报数据环境中的所有数据按组织结构分为两大类: 结构化数据和非结构化数据。典型的结构化数据如地面观测资料、站点信息、预报检验数据以及一些配置信息都以记录的方式存储在数据库中, 本数据环境选用了 Oracle

数据库存储结构化数据。除此之外的绝大部分资料都是以文件、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^[4]。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。

对于原始的结构化观测数据, 为方便预报检验算法的使用, 需要将其转换为 MICAPS 格式的文本文件。这种由结构化向非结构化的数据组织形式的转换, 通过使用数据统一服务接口对原始观测数据进行二次加工得以实现。气象数据统一

服务接口 (meteorological unified service interface community, 简称“MUSIC”), 基于国家级—省级统一的 CIMISS 数据环境, 面向气象业务和

科研, 提供了统一、标准、丰富、具有多种调用形式的数据访问服务和应用编程接口, 为各级应用系统提供直接支撑的数据接入服务^[5] (图 3)。

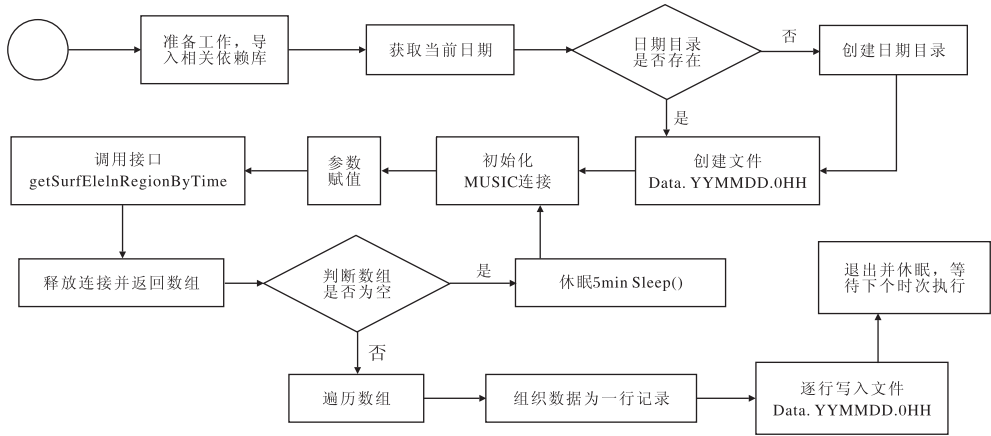


图 3 调用 MUSIC 实现数据组织形式转换范例流程

5 数据管理

5.1 数据存储策略

由于存储空间并不是无限提供的, 因此为了保持数据环境存储空间的动态平衡, 需设置合理的数据存储策略, 包括备份和清除两种情况。数据备份是指将生产环境中的热数据保存若干日后向备份区迁移, 成为冷数据。数据清除是指将生产环境的热数据保存若干日, 超过使用时效后直接删除, 以释放存储空间。适配不同存储策略的基本原则根据资料的重要性以及单日的数据量评级确定, 除个别试验性资料需特殊配置外, 基本策略如表 2 所示。

备份策略可编写脚本程序或借助第三方备份软件实现。数据清除策略, 记录级的数据使用 SQL 语言中的删除语句实现, 文件系统中的数据则根据操作系统的不同, 可使用如下参考命令并配置定时任务实现。参考命令中的 Path 表示数据存储目录, NDay 表示数据保存的日数。

Windows 平台: `forfiles /p "Path" /s /m * . * /d -NDay /c "cmd /c del @path"`

Linux 平台: `find . -type f -mtime +NDay -exec rm -f {} \;`

5.2 访问控制策略

为了保证数据安全, 对文件服务器上存储的数据进行了分级的用户权限管理, 以实现数据

表 2 陕西省精细化格点预报数据环境的数据存储策略

资料类别	重要性 评级	数据量 评级	存储策略
原始地面观测数据	中	小	长期存储, 备份
原始模式产品数据	中	大	保存 3 日, 清除
原始雷达、卫星数据	低	大	随用随取, 清除
系统元数据、站点信息	高	小	长期存储, 备份
模式降尺度数据	高	中	保存 10 日, 备份
格点实况和预报产品	中	中	保存 30 日, 备份
动态交叉取优 DCOEF 数据	高	小	长期存储, 备份
格点编辑数据	低	小	按需保存, 清除
预报算法研究产生的 中间数据	低	大	保存 10 日, 清除
地市共享 MICAPS 格式预报产品	中	中	保存 30 日, 清除
订正后的预报产品	中	小	保存 30 日, 备份
预报解析产品	中	中	保存 30 日, 清除
预报检验数据	高	小	长期保存, 备份

的访问控制。遵循“读写分离”原则, 将访问权限分为三级。首先, 使用超级管理员账户创建了多个的操作系统账户, 并将其设置为管理员属组, 分配给研究人员使用。研究人员使用各自用户生成的数据文件具备可读写操作权限, 同时, 生成的数

刘璐,屈振江,张勇,等. 陕西猕猴桃果品气候品质认证模型构建[J]. 陕西气象,2017(4):21-25.

文章编号:1006-4354(2017)04-0021-05

陕西猕猴桃果品气候品质认证模型构建

刘璐,屈振江,张勇,李艳莉

(陕西省经济作物气象服务台,西安 710014)

摘要:气象条件是猕猴桃品质形成的关键环境因子。基于猕猴桃品质调查数据和加密气象观测资料,综合考虑不同物候期气象适宜度、气象灾害损失度及生产管理条件,利用模糊函数和专家打分法,建立了分物候期气象适宜度、主要气象灾害灾损度及生产管理措施影响度的气候品质综合评价模型。利用认证模型对2013年陕西10个猕猴桃(海沃德)种植园果品品质进行业务检验,对猕猴桃品质评价的准确率达80%,具有业务应用价值。

关键词:猕猴桃;海沃德;气象适宜度;气候品质;认证模型

中图分类号:S162.5

文献标识码:A

陕西是全国最大的猕猴桃生产省,种植面积和产量分别达6.2 hm²和124.4万t,其中秦岭北麓和汉江流域是我国优质猕猴桃的最佳适宜区^[1]。随着猕猴桃产量的增加和市场对果品品质

要求的不断提高,猕猴桃产业已从产量效益型向质量效益型转变,而气象条件是影响猕猴桃品质形成的关键因素^[2]。建立猕猴桃果品气候品质认证模型,对影响品质的气候条件优劣等级进行评

收稿日期:2016-12-23

作者简介:刘璐(1981—),女,陕西西安人,汉族,硕士,高级工程师,从事果业气象服务技术研究。

基金项目:中国气象局气象关键技术集成与应用(重点/面上)项目(CMAGJ2015M60)

据对管理员属组的各个成员皆可见。对有交互的成员可读可写,对无交互成员只读。其次,使用超级管理员创建一个用户,设置为普通用户属组,用于对最终发布产品的共享。该用户具备只读权限,可以查看和下载数据。最后,超级管理员对文件服务器上的所有数据具有完全控制权限,可以做读写、删除、修改、执行等操作,以便于实施备份和清除等数据存储策略。

5.3 共享服务策略

对于保存在文件服务器上最终发布的预报产品、解析产品以及其他需要提供给用户使用的数据,遵循“多样化共享”原则,提供FTP、磁盘映射、标准API接口三种访问方式^[6]。为缓解服务端压力,目前不提供主动的订阅式推送服务。共享用户权限参考访问控制策略中设置的三级用户,赋予分级的数据共享级别。

参考文献:

- [1] 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[J]. 应用气象学报,2015,26(4): 500-512.
- [2] 于耳. 一种基于VSAN的分布式存储系统构建和应用[J]. 中国教育信息化,2014(16):72-75.
- [3] 朱海亚. 数据存储中分级存储管理研究[J]. 数字化用户,2013(4):55-59.
- [4] 何婉文,孙周军,肖文名. 基于CIMISS的广东省现代农业气象业务数据环境建设[J]. 成都信息工程学院学报,2012(4):210-217.
- [5] 张正阳,朱倩雨. 基于CIMISS环境的气象数据统一访问接口简介[J]. 沙漠与绿洲气象,2015,9(1): 134-136.
- [6] 何林,吉庆. 气象-环保数据交换共享方案的设计与实现[J]. 陕西气象,2016(4):36-39.