陈增境,卓凤艳,韩格格,等. 基于混合云的气象大数据云平台改造及安全防护设计[J]. 陕西气象,2025(1):77-82.
文章编号:1006-4354(2025)01-0077-06

基于混合云的气象大数据云平台改造及安全防护设计

陈增境[1,2],卓凤艳[1,2],韩格格[1,2],高英育[1,2]
(1. 中国气象局旱区特色农业气象灾害监测预警与风险管理重点实验室,银川 750002;
2. 宁夏气象防灾减灾重点实验室,银川 750002)

摘 要:针对传统私有云受限于本地计算资源供给和部署空间不足、扩建成本和管理成本高的现状,提出将公有云与本地私有云有机融合,构建混合云运行模式,以气象大数据云平台中解码处理子系统作为改造对象,将业务流程向公有云延伸,并按照三级等保要求,从边界防护、流量管控和终端安全三方面设计构建网络安全防护体系。以宁夏为例,从数据完整性和传输时效性两方面开展对比分析,评估业务改造效果。结果表明,混合云运行模式下,总体时效性较单纯私有云环境稍有延迟,但数据解码入库完整,系统整体运行稳定可靠,能够满足实时业务需求。

关键词:气象;混合云;云安全
中图分类号:TP309 **文献标识码**:A

2020 年建成的气象大数据云平台[1](下简称“天擎”),即本地私有云,提供了“数算一体”平台化服务,支撑了“云+端”气象业务,很大程度上消除了“信息孤岛”,但随着信息化程度的提高,相应的运维压力和维护成本呈直线上升趋势。例如,自 2021 年天擎业务化运行以来,完成了 11 类 370 种数据的接入,加之气象核心业务逐步融入天擎,服务器性能已趋于饱和,常规计算资源建设速度和 UPS、空调等基础环境条件无法满足气象新增业务的发展,同时需要大量人力物力来运维管理,故利用云上资源,引进混合云[2-3]技术,选择扩展性强、安全性高的宁夏回族自治区电子政务公共云平台作为气象业务系统的部分承载平台,与本地私有云有机融合,实现业务系统在两朵云间协同工作。本文以天擎-数据交换与质控系统作为改造研究对象,搭建混合云运行模式,部署数据交换与质控系统中解码处理(Data Processing Center,简称 DPC)子系统,通过研究跨云业务协同调度运行、分工负载、应急备份切换等关键技术,建立气象数据异地解码业务流程,为实现天擎算力资源异地异构扩容提供实践经验。

1 本地私有云现状

天擎是为解决数据供应不足、数据交换传输不灵活、数据存储服务不够高效、业务系统烟囱林立等问题,基于云计算、大数据等新兴信息技术建设的私有云平台,主要由数据交换及质控、产品加工、挖掘分析、数据存储与服务、业务监控等 5 个系统构成[4],由基础设施资源和标准规范体系提供支撑保障。天擎因其处于业务内网,在数据传输、应用服务和网络安全方面具备较为可靠的保障,其结构如图 1 所示。

2 公有云简介

宁夏回族自治区电子政务公共云平台[5-6](下简称“公有云”)是按照智慧宁夏总体规划,在银川、中卫两市建成异地主备、互为冗余的双数据中心。公有云是运用云计算架构建设的具有基础

收稿日期:2024-03-18
作者简介:陈增境(1984—),男,汉族,福建福州人,硕士,高工,主要从事气象信息网络安全研究与应用。
通信作者:卓凤艳(1979—),女,汉族,黑龙江鸡西市人,本科,高工,主要从事气象信息系统运维管理。
基金项目:中国气象局旱区特色农业气象灾害监测预警与风险管理重点实验室开放研究项目(CAMF-202309)

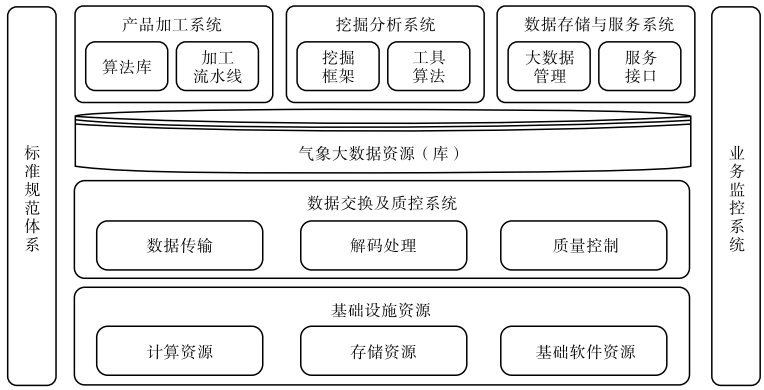

图 1 大数据云平台结构图

性、公共性的服务平台，采用多云云管平台，整合了电信、移动和联通公有云资源，形成统一的计算、存储、网络和安全资源池，为宁夏回族自治区内各级政府单位、企事业部门提供资源服务。同时公有云采用云网安全区域边界隔离，对接入用户进行安全管理，具备实名认证、业务域切换、审计溯源、防仿冒等安全管理功能。

3 混合云建设

3.1 架构设计

混合云作为公有云和私有云的有机结合[7-8]，兼具两者优点，私有云在边界防御、流量控制、安全准入等方面易于实现，数据传输及共享服务响应迅速，满足气象数据安全需求，而公有云可提供灵活多变的 IT 资源和服务，符合对基础资源实时扩容的需求，混合云是现阶段较为流行的一种云建设方式。根据业务实际需求，参照现有天擎架构模式，设计基于混合云的气象解码入库架构。混合云部署架构如图 2 所示，由公有云和私有云两部分组成，两朵云基础设施彼此独立运行，通过电子政务外网互连，将应用程序横向扩展部署在云上，实现应用的平滑上云。

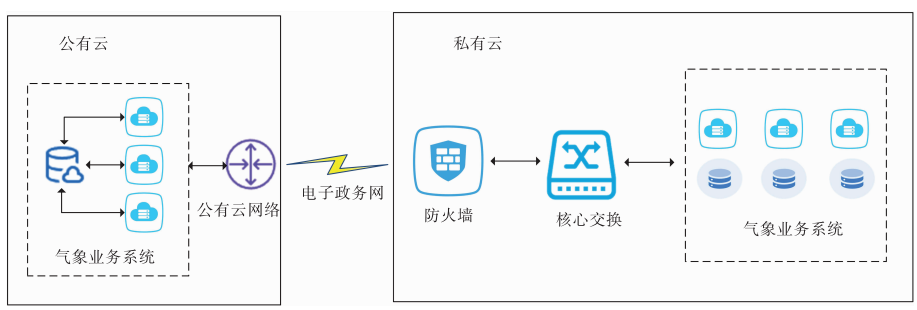

图 2 混合云部署架构图

3.2 业务流程设计

DPC 子系统目前在天擎内部署有 15 台服务器(dpc01～15)，用于处理结构化、半结构化、非结构化数据，主要功能为数据解析和数据入库，系统对接国内气象通信系统[9-11](China Telecommunication System，CTS)接收数据，经解码质控后，入库至数据存储与服务系统(service-oriented data storage system，SOD)。为保障业务运行稳定，高效利用公有云资源，参考天擎-DPC 系统部署模式，部署 8 台云服务器(yundpc01～08)，每台云主机部署多种解码程序，每种程序采用分布式部署模式，混合云解码部署程序部署情况如表 1 所示。

根据两朵云的特点及性质，在两朵云中部署相同的解码程序，同时对接上游 CTS 系统发来的气象数据。为兼顾业务流量负载均衡，将传输时效要求高的气象数据置于私有云处理，数据保存在本地 SOD，而时效要求低的气象产品数据则依托公有云解码，处理后回写至本地 SOD。当私有云或公有云一方出现故障时，另一方将承担所有

数据解码处理业务,并将数据推送至本地 SOD。　　业务流程如图 3 所示。

表 1　解码程序安装部署情况

序号	程序	云 DPC	天擎-DPC	作用
1	结构化流处理程序	yundpc01～03	dpc01～03	处理结构化消息数据
2	非结构化应用程序	yundpc03～04	dpc04～05	处理非结构化文件数据
3	数值模式预报应用程序	yundpc03～04	dpc06～07	处理半结构化文件数据
4	雷达流	yundpc05～06	dpc08～09	处理非结构化流式数据
5	消息转发	yundpc07～08	dpc10、dpc13	消息拆分与转发
6	RabbitMQ	yundpc07～08	dpc11、dpc12	消息中间件
7	结构化应用程序	yundpc05～06	dpc14～15	处理结构化文件数据
8	FTPserver	yundpc07～08	dpc01～10、dpc13～15	FTP 服务

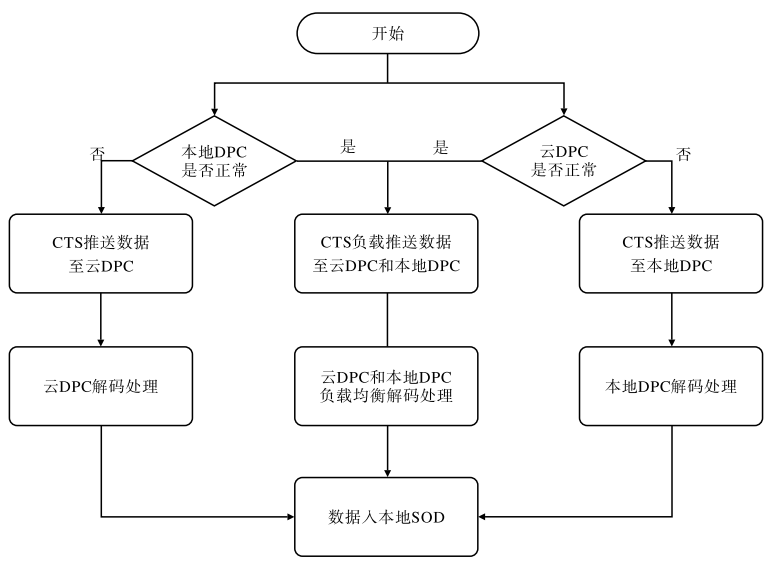

图 3　混合云数据解码业务流程图

3.3　网络安全设计

网络安全防护是业务系统建设中不可或缺的部分。混合云除传统通用威胁外,还有其特有的威胁形式,如云计算导致传统安全域划分不可用、东西向流量不可见、横向移动无法及时发现等安全问题。安全等级保护 2.0 提出在通用安全防护的同时,额外新增了针对云计算环境的威胁防护要求。根据等级保护基本要求,通信网络、区域边界、计算环境以及安全管理中心是等级保护 2.0 纵深防护理念的重要组成部分。基于等级保护视角构建混合云安全防护体系[12],主要从区域边界防护、数据流量管控和终端安全防护三方面开展设计,安全防护体系架构如图 4 所示。

3.3.1　区域边界防护　数据通过电子政务外网实现交互,两朵云间采用下一代防火墙进行安全隔离,连接公有云气象虚拟专网[13](virtual private cloud,简称 VPC)和本地业务内网,实现混合云南北向流量的安全防护。防火墙制定精准的访问控制策略,仅允许特定网段接入,放行业务使用的服务和运维端口。VPC 内通过安全组规则、访问控制白名单等方式,灵活控制访问专有网络内云资源的出入流量,同时每个 VPC 都有一个独立的隧道号,一个隧道号对应着一个虚拟化网络,与其他业务单位保持逻辑隔离,确保业务系统网络安全,有效防范云上安全威胁横向蔓延。

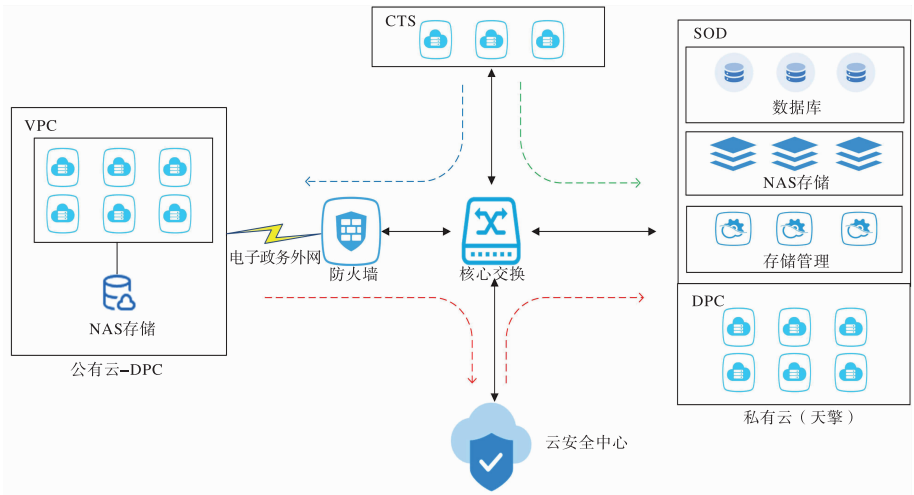

图 4 混合云网络安全防护体系架构

3.3.2 数据流量管控

(1)南北向流量防护 除采用物理防火墙进行内外网安全隔离外，将公有云出向数据引流至云安全中心进行流量检测[14]，通过多种机制的分析检测，对黑客入侵行为、SQL 注入、DDoS 攻击进行有效检测及防护阻断，合规安全的流量按原路由返回核心业务内网。

(2)东西向流量防护 在混合云内安装终端安全软件，采用微隔离技术，提升东西向流量的检测和控制能力，防止攻击者入侵内部业务网络后的东西向移动，提升对抗攻击的能力，同时按照 IP 地址、端口、流量类型以及流量方向配置防火墙规则。

3.3.3 终端安全防护 每台终端均安装安全防护软件，由云安全中心统一管控，定期更新病毒库，及时修复补丁。终端设置复杂口令，采用 SSH 协议远程管理服务器，防止鉴别信息在网络传输过程中被窃听。操作系统中禁用系统默认账户，删除多余、过期的账户，同时关闭多余的服务和端口。

4 数据处理分析

4.1 处理流程概述

为对比测试公有云上数据解码情况，分别采用双轨独立解码运行模式和云上云下数据解码分工负载的混合云模式。气象数据传输方式包括数据消息、文件消息和数据流 3 种，气象数据形态有结构化、半结构化和非结构化 3 种，两者不同的组合，产生 7 种数据解码处理框架，如表 2 所示。

表 2 气象数据传输与数据形态对应解码处理框架

传输方式	数据形态		
	结构化	半结构化	非结构化
数据消息	结构化消息数据流式处理 结构化消息数据多线程处理	无	无
文件消息	结构化文件流式处理 结构化文件多线程处理	半结构化文件多线程处理	非结构化文件多线程处理
数据流	无	无	非结构化数据流式处理

不同的处理框架，分别衔接上游的数据交换与下游的存储系统，完成从数据接收到入库的全部过程。每种处理框架可以处理类型相似的多种资料，根据数据落地形式将其存储至不同的数据库和 NAS，7 种数据处理框架如图 5 所示。

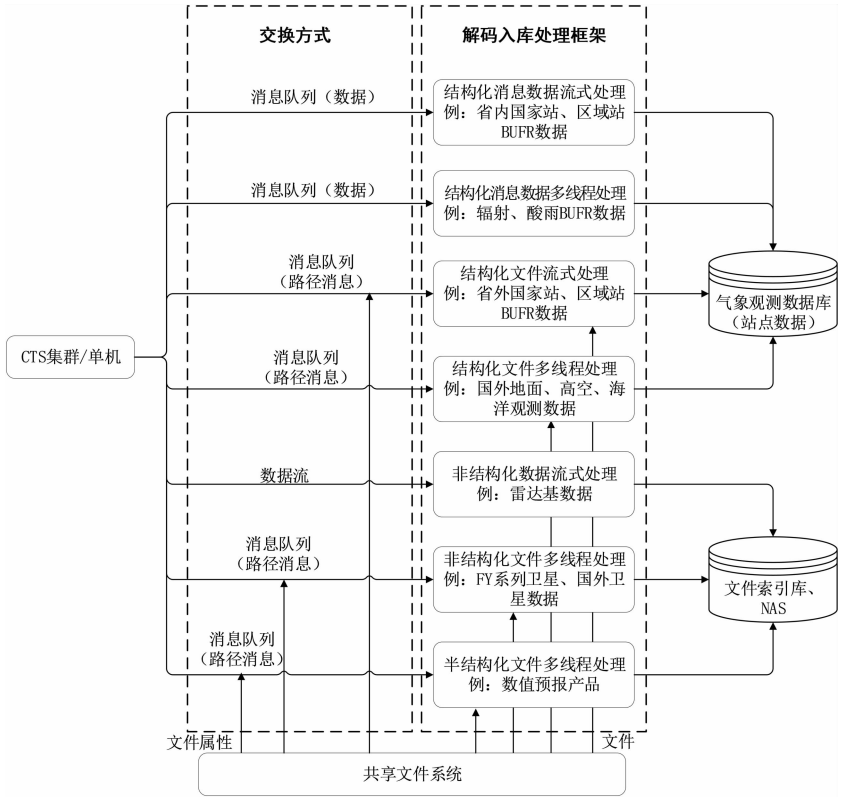

图 5　气象数据解码入库处理框架

4.2　处理对比分析

为确保业务切换的可靠性,从数据传输完整性、时效性方面,开展公有云和私有云解码处理对比分析。选择 5 种数据开展分析,分别为国家站地面小时数据(结构化消息数据流式)、辐射小时数据(结构化文件流式)、交通站小时数据(结构化文件多线程)、雷达基数据(非结构化数据流)和雷达基数据(非结构化文件多线程),时间为 2024 年 1 月 1 日至 2024 年 1 月 31 日。数据传输环境如表 3 所示。

表 3　宁夏气象局混合云 DPC 业务环境对比

序号	系统	带宽	线路	所在网络区域	是否经过防火墙
1	公有云-DPC	100 Mibit/s	电子政务专线	政务网-气象 VPC 区域	是
2	私有云-DPC	10 000 Mibit/s	局域网	核心业务网	否

4.2.1　完整性对比　测试时间段内,两朵云 DPC 接收数据正常,没有出现数据丢失和解码错误问题。存储过程中两朵云数据量相同,完整性均为 100%,其中解码入库国家站地面小时数据文件 19 940 个,辐射小时数据 1 440 个、交通站小时数据 5 040 个,雷达基数据(非结构化数据流)传输 14 440 个,雷达基数据(非结构化文件多线程)传输 17 224 个。

4.2.2　时效性对比　表 4 为一个月测试时段内,同等数据分别在私有云和公有云进行解码处理时效对比情况,可以看出:大部分数据在两种云环境下的处理时效基本一致,处理的数据量占比在 86.24%～99.35%之间;私有云处理时效较快的数据量占比区间为 0.1%～9.76%;公有云处理时效较快的数据量占比区间为 0%～6%。通过数值量对比,气象数据在公有云环境中进行异地处理时时效相对略低,尤其是单个文件较大的雷达基数据较为明显。由于从私有云到公有云进行数据交互,存在跨区域传输,加之用于测试的云服务器资源相对较少、网络链路节点多、带宽低,导

致公有云数据传输的效率较私有云有所下降，但总体而言两朵云 DPC 功能一致，时效差异不大，满足业务运行需求。后期可通过申请公有云上带宽、计算和存储资源的扩容，提高业务处理能力。

表 4 宁夏气象局公有云和私有云在测试时期内不同时效下数据量对比

不同处理时效	数据类型				
	国家站小时数据	辐射小时数据	交通站小时数据	雷达基数据（数据流）	雷达基数据（文件）
公有云比私有云快的数据量/个(占比/%)	1 167(6.00)	0	28(0.56)	146(1.00)	0
公有云和私有云一样快的数据量/个(占比/%)	16 766(86.24)	1 412(98.06)	5007(99.35)	13 069(90.51)	15 543(90.24)
私有云比公有云快的数据量/个(占比/%)	1 507(7.76)	28(1.94)	5(0.10)	1 225(8.49)	1 681(9.76)

5 小结

借助电子政务公有云资源，探索了业务“上云”的新途径，通过研究跨云业务协同调度运行、分工负载、应急备份切换等关键技术，实现业务在混合云中协同调度，有效缓解了单位信息化发展经费不足、管理成本高等问题，为后续其他核心业务上云提供理论基础和实践案例。从宁夏气象部门应用表明，基于行业垂直管理的业务体系和具备一定规模的私有云体量，通过引入公有云的信息资源优势，拓展了业务云化改造和行业云发展的思路，既能实现多元算力资源的精细管理和协同调度，又可确保网络安全和数据安全，混合云运行模式适应目前宁夏气象行业特点、业务发展趋势和用户对信息资源的需求。

参考文献：

[1] 刘媛媛，何文春，王研，等．气象大数据云平台归档系统设计及实现[J]．气象科技，2021，49(5)：688-705.

[2] 韩兵，李东明．基于混合云技术的农业信息化平台架构的研究[J]．吉林农业大学学报，2020，42(6)：693-698.

[3] 柳鹏，刘波，周娜琴，等．混合云工作流调度综述[J]．计算机科学，2021，49(5)：235-243.

[4] 刘洋，黄志，徐娟，等．气象大数据云平台监控告警系统[J]．计算机系统应用，2023，32(3)：86-94.

[5] 张旭辉．关于网络安全等级保护 2.0 在公有云中的应用研究[J]．数字通信世界，2020(6)：240-241.

[6] 殷涛．数字政府公有云管理体系建设思路[J]．广东通信技术，2023(11)：2-5.

[7] 杨武．混合云平台的设计及实现[J]．电脑知识与技术，2021，17(11)：77-78.

[8] 丁丽娜．“混合云＋”全国党校学术资源共建共享平台探索[J]．四川图书馆学报，2022(2)：71-75.

[9] 赵芳，何文春，张小缨，等．全国综合气象信息共享平台建设[J]．气象科技进展，2018，8(1)：171-181.

[10] 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.

[11] 余永城，翁秋华，段卿，等．RabbitMQ 在气象通信系统中的应用研究[J]．计算机技术与发展，2020，30(4)：216-220.

[12] 陈驰，于晶．云计算安全体系[M]．北京：科学出版社，2017：27-51.

[13] 牛红韦华，赵晖文，丁国强，等．云资源池异构 SDN 架构演进研究[J]．电信工程技术与标准化，2024，37(1)：18-25.

[14] 张瑞英，陈秀兰．混合云平台构建管理及安全性研究[J]．电脑知识与技术，2022，18(13)：40-42.