

气象大数据云平台归档系统设计及实现

刘媛媛 何文春* 王妍 李江涛 白金婷

(国家气象信息中心, 北京 100081)

摘要 本文针对“天擎”归档系统的建设需求、总体架构,提出系统实现的若干关键技术问题并进行深入分析,气象大数据云平台(以下简称“天擎”)归档系统设计为面向海量数据归档存储的支撑系统,保障数据的安全完整。系统基于在线存储、离线存储以及磁带库管理软件,通过层次性结构体系、消息队列和微服务技术框架,实现多协议数据归档、敏感数据归档及转储数据归档等功能,满足各种新增数据归档业务场景,同时具备丰富的数据回取服务、数据全流程的追溯和规范化的磁带管理保障等功能。2020 年 8 月“天擎”归档系统投入业务运行,每日归档的实时数据量约为 10 TB,后续最高可承载约 240 TB 日数据量的归档业务。

关键词 海量数据;归档;回取;磁带

中图分类号: P409 **DOI:** 10.19517/j.1671-6345.20210083 **文献标识码:** A

引言

随着对龙卷、冰雹和强对流大风等小尺度天气监测的需求提升,分钟级观测、秒级自动探测数据海量涌进,以及高分气象卫星发射和高时空分辨率数值模式的发展,近年来气象数据量增长明显加速,可以说真正进入了气象大数据时代,如何解决好海量气象数据存储的问题也成为信息化发展关注的焦点^[1-7]。海量数据规模通常会达到 PB 级,甚至已经发展到 ZB 级。虽然大数据技术发展得如火如荼,考虑到成本海量数据管理通常不会使用单一的存储技术或介质,会根据数据访问的情况决定数据的存储方式,这是因为从数据的服务特点来看,数据会呈现冷热不均的分布。数据总量中用户访问频次较高的数据称为“热数据”,用户访问频次一般的数据称为“温数据”,用户访问频度较低的数据称为“冷数据”。冷热数据分布通常符合二八原则。对于冷数据,大部分数据中心都会采用相对廉价的存储介质进行管理。由于磁带具有安全可靠、经济节能的优势,而且技术成熟,是存储冷数据较好的选择。在气象行业,用磁带库存储“冷数据”有很长的历史,欧

洲中心从 1985 年就开始在 MARS 系统使用^[8-10],目前仍在大规模应用。美国 NOAA 的 CLASS 系统、NOMDAS 系统与中国气象局的 9210 系统均在 2002—2004 年左右开始在数据归档业务场景中使用磁带库。

中国气象局国家气象信息中心承担世界数据中心中国气象科学数据分中心的职责,负责气象资料的归档业务。归档业务最初阶段为人工归档方式,通过光盘以及磁带介质载体对数据集产品以及少部分人工观测原始数据、数值预报模式产品等进行归档。自 2004 年起,信息中心启动了自动化归档业务,先后历经国家级气象资料存储检索系统归档子系统、国家级综合气象信息共享系统归档子系统(以下简称 CIMISS 归档系统)。国家卫星气象中心也先后建成气象卫星资料存档系统^[11]和风云气象卫星数据存档与服务系统^[12]。存储介质先后使用过 9940 B 磁带、LTO4 磁带。随着数据量的不断增加,自 2017 年以来,系统开始出现资源占用率高以及磁带读写性能低下、数据回取服务缓慢等问题。2018 年,气象大数据云平台(简称“天擎”)建设时,为满足大数据背景下的归档业务需求,“天擎”归档

<http://www.qxkj.net.cn> 气象科技

国家发展改革委员会项目“海洋气象综合保障一期工程”资助

作者简介:刘媛媛,女,1980 年生,硕士,高级工程师,主要从事海量气象数据存储管理、分级存储等研究工作,Email:liuyy@cma.gov.cn

收稿日期:2021 年 2 月 19 日;定稿日期:2021 年 6 月 18 日

* 通信作者,Email:hewc@cma.gov.cn

系统开始建设。

1 归档业务建设需求

“天擎”是以“云+端”为设计思路实现的“数算一体”的平台,改变了过去烟囱式的应用系统建设方式,各业务系统的输入输出数据均在“天擎”中进行统一集约化管理,算力靠近数据,数据不再冗余存储。“天擎”以数据为中心支撑各应用系统围绕数据结合算法和算力进行集约化运行,提升业务流程运行效率并节约系统资源。“天擎”管理的数据分为 3 部分:①核心气象数据,如气象观测、加工的气象数据和产品。②辅助气象数据,即辅助气象业务应用的其他数据,如装备信息、社会化观测、气象多媒体等数据。③行业社会交换数据,如农业、林业、交通、海洋、水利等外部门和社会交换的、与气象应用相关的数据。“天擎”管理的数据存储规模预计将达到 100 PB。

“天擎”一方面要采用合理的技术方案对海量数据进行集中存储管理,另一方面要满足不同应用场景的数据服务需求,因此需要对应用的特点及数据使用的特点进行详细的分析。“天擎”对于业务和科研应用均需进行良好的支撑,天气分析、决策服务、数值预报等实时性业务希望数据服务能够毫秒级响应,它们使用的数据范围通常为当前时刻数据或近期数据,科研工作需要进行的历史数据回算或分析,通常会使用 5 年前甚至更长时间前的数据,但数据服务响应要求相对较低,可以为小时级,还有一些业务应用需要使用近 5 年的数据,数据服务响应需求处于实时业务和科研工作之间。

“天擎”根据上述分析将所管理的数据按热度分为 3 类(表 1):①7~30 天数据为热数据,采用性能较高的如 SSD 等在线存储介质进行存储;②5 年内的数据为温数据,采用普通性能的在线存储介质;③5 年以前的数据为冷数据,采用磁带等近线、离线存储介质进行管理。

“天擎”归档系统承担的是对冷数据的管理和服务。由于“天擎”管理数据的范畴的扩大,归档数据范围和种类的也随之增长,2018 年建设时预计未来 5 年内,“天擎”归档业务日数据量会由 1 TB 上升到至少 10 TB,最高有可能达到 40 TB,同时归档系统还需支持数据分级分类归档等数据安全要求。基于以上对未来业务发展的预估,“天擎”归档系统

设计建设为一套支撑百 PB 级数据归档、具备高效自动化数据回取服务能力的标准、规范的气象资料归档管理系统:①实现归档业务功能的分布式集群化运行能力,建立可扩展的数据归档和回取系统,满足不断增长的实时历史数据的归档业务需求;②支持不同安全级别数据的归档及回取,磁带读写效率大于 150 MB/s;③综合考虑数据记带性能和数据回取服务需求,实现灵活的数据归档策略的制定;④实现 7000 盘存量 LTO4 磁带的自动化转储功能和支撑未来转储需求所需匹配的系统功能;⑤与“天擎”存储管理与服务系统进行协同,实现自动化数据回取服务;⑥在天镜中实现系统主要业务流程监控及数据全流程的监控;⑦实现对磁带安全规范的管理。

表 1 “天擎”归档系统数据存储策略

数据热度	数据种类	存储介质及规模	数据访问频次	服务响应时效
热	7~30 天站点数据和卫星、雷达、数值预报等数据	高性能 在线存储 (TB 级)	>1 次/h	ms 级
温	近 5 年的数据	普通性能 在线存储 (<20 PB)	>100 次/月	s 级至 min 级
冷	5 年前的数据	近线离线存储 (20~100 PB)	不定期	h 级

注:1 TB=1024 GB, 1 PB=1024 TB。

2 系统设计

2.1 系统架构

在“天擎”归档系统总体技术结构中,明确定义了系统建设的层次性结构体系,并将系统的各个层次加以详细划分,主要包括应用与接口交互层、归档业务处理与服务层、系统软硬件支撑层,如图 1 所示。

应用与接口交互层提供归档系统门户、专题及应用数据管理功能。在归档系统门户中,实现各类资料策略管理与配置功能、近线和离线数据实现回取功能、订单管理功能、资料归档统计信息功能等。应用与接口交互层同时与通信系统(CTS)、数据统一访问接口(MUSIC)、天镜等进行接口交互,提供资料数据归档、回取和归档系统各业务环节状态、告警信息上报等功能。

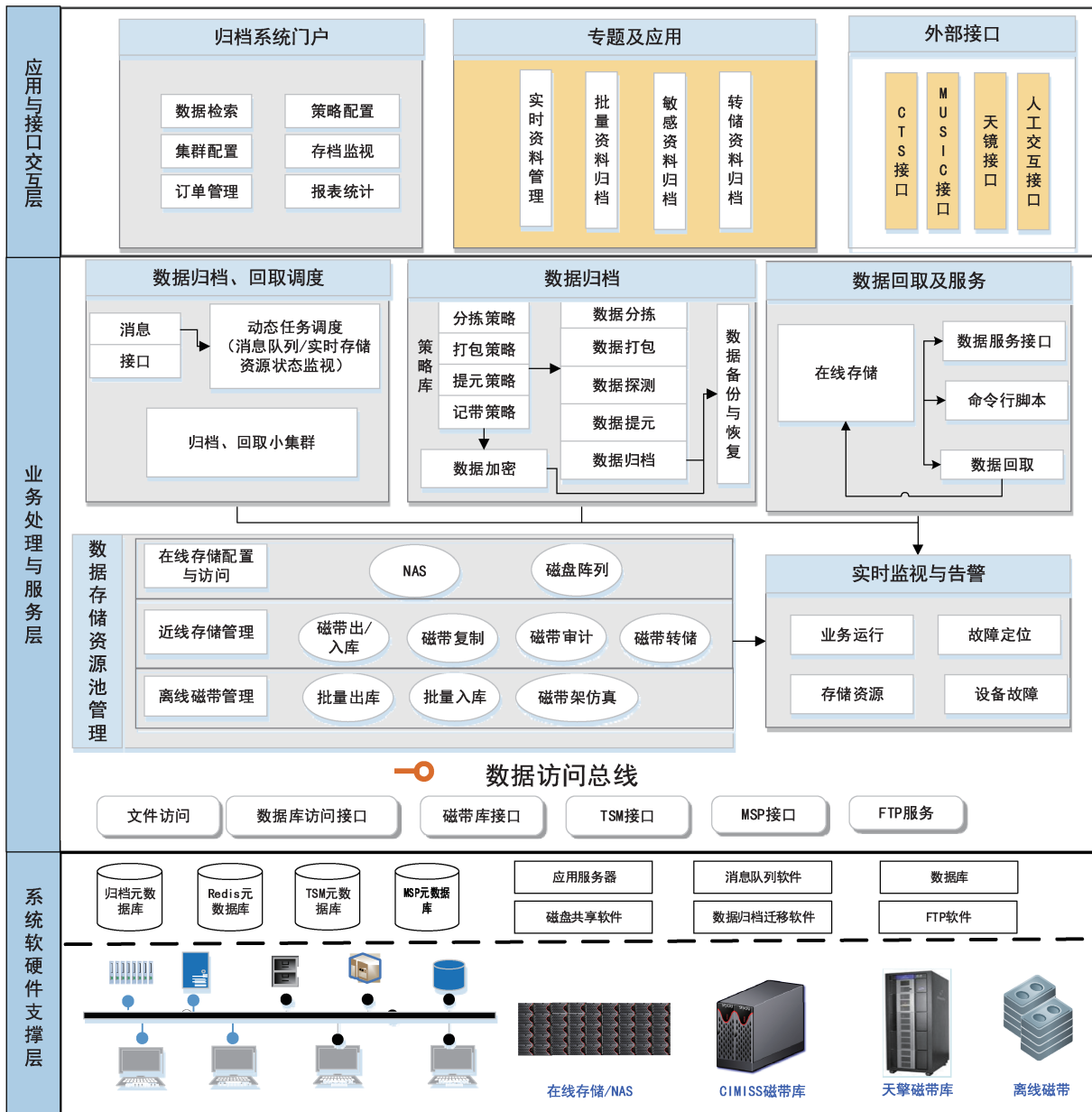


图 1 “天擎”归档系统架构

归档业务处理与服务层实现归档业务具体功能,以服务的形式分布式部署在归档服务器集群环境中,主要包括数据归档调度、数据归档、数据分发与服务、数据存储资源池管理和实时监视与告警。其中数据归档调度子系统以消息队列(RabbitMQ)为基础实现归档业务动态任务调度,同时支持在管理界面上灵活配置归档集群;数据归档子系统以微服务形式实现数据分拣、打包、探测、提元、归档、回取等具体业务逻辑,由数据归档调度子系统统一调度;数据回取与服务主要实现归档资料多种形式的

回取功能,包括订单式回取、脚本式回取、自动化回取和敏感资料回取等;数据存储资源池管理通过在线存储配置与访问、近线存储管理和离线磁带管理实现存储资源统一管理功能,在线存储配置与访问支持 NFS(网络文件系统)、S3(简单存储服务)、FTP(文件传输)等多种数据访问协议,近线存储和离线存储管理实现磁带出入库、磁带复制、磁带审计、磁带转储和离线磁带架管理等;数据访问总线实现归档元数据访问接口、磁带库访问接口、磁带库管理软件(MSP)访问接口、FTP 访问接口的统一数据

访问服务,支撑各服务业务逻辑的实现和归档门户各应用功能的集成。实时监视与告警将业务处理与服务层各环节运行问题及时向天镜上报。

系统软硬件支撑层由各类物理设备和基础软件构成。其中物理设备主要包括各软件部署所需要的服务器设备、在线存储设备、近线存储设备、网络设备等,基础软件包括操作系统、数据库管理软件、应用服务器软件、消息队列软件、磁盘共享软件、归档软件等组成。

2.2 硬件环境

“天擎”归档系统支撑硬件设备主要包括 X86 物理服务器、盘阵、磁带库以及网络设备。其中 X86 物理服务器用于部署归档系统软件,盘阵用于磁带库归档中间环节缓存,提供数据分拣、数据打包等的缓存空间,磁带库用于实现归档资料的近线存储和离线存储。这些硬件共同支撑归档系统承载 4 种类型数据的归档,即实时数据归档、批量数据归档、敏感数据归档和转储数据归档。实时数据指通信系统实时传输到归档系统的数据,如地面自动站分钟观测数据;批量数据指按批次提交归档系统的数据,如某次科考试验的数据集;敏感数据指仅能对部分用户开放的数据;转储数据是指已在归档系统中待更换存储介质的数据,如旧型号磁带中的数据为避免磁带失效需要改为新型号的磁带存储。

归档业务为 I/O 密集型应用,需对大量数据进行读写操作,为避免互相影响,对不同类型归档业务匹配对应的服务器进行支撑。实时数据归档每天要承担约 3 TB 数据的打包处理及归档,其中大量小、碎文件的打包处理对资源的消耗较高,目前配备了 6 台服务器处理实时数据。批量数据文件数量通常不高,但数据量会偏大,达到 10 TB 至 PB 级,目前分配了 3 台服务器;转储数据文件数量很高数据量也很大,但因为转储工作处于起步阶段,暂分配了 3 台服务器进行处理;敏感数据量较小,只是物理上需要与其他数据进行隔离,目前 1 台服务器即可满足需求。现在共计使用 13 台服务器,对于每类归档业务后续可根据数据量增长情况进行调整扩容。

系统采用华为 OceanStor 18500 V5 盘阵作为归档业务的临时缓存。

系统使用的磁带库为曙光 STL4500 高密度磁带库,磁带库包含 1 套磁带库系统库体、62 个 LTO7 磁带驱动器,可装载 10000 盘 LTO7 磁带。磁带库

目前所装载的 5000 盘磁带可用容量为 27 PB,满配后可用容量为 54 PB。磁带库驱动软件使用的是火星 MSP 归档存储中间件实现对 STL4500 磁带库的管理和磁带读写功能。

3 关键技术

“天擎”归档系统通过以下关键技术满足发生巨大变化的业务需求:①采用分布式集群运行技术实现系统的动态可扩展;②策略驱动设计实现系统的灵活性;③实现多协议数据归档、敏感数据归档、转储数据归档的功能满足多场景数据归档需求;④提供丰富的数据回取服务接口提升回取数据服务便捷性;⑤归档数据全流程追溯和归档业务各环节数据量统计保证数据归档的完整性和正确性;⑥规范化的磁带管理实现归档的磁带服务。

3.1 分布式小集群设计

为合理利用硬件资源并实现系统可扩展性、适应海量数据归档需求,“天擎”归档系统采用小集群(图 2)的方式对硬件资源进行规划使用:①服务器。根据开展的实际业务对服务器进行集群划分,如实时归档集群、历史归档集群、敏感资料归档集群、资料转储集群等,每个集群所需的服务器数量根据所承担的业务负载决定;②驱动器。驱动器的用途包括归档和回取,对于用于归档的驱动器再根据服务器集群进行匹配分组,包括实时归档驱动器组、历史归档驱动器组等,每组所需使用的驱动器个数可根据业务量动态调整;③磁带库。为方便后续数据服

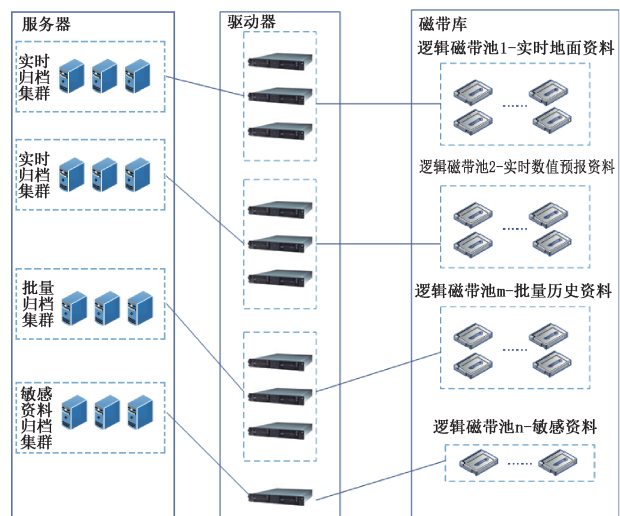


图 2 分布式小集群设计

务,满足同类型相近时间段的资料最好存储在同一批磁带的需求,归档系统对磁带池根据资料类型进行逻辑划分,共分配了 14 个主磁带池,实时数据使用 9 个主磁带池,批量数据和转储数据使用 3 个主磁带池,敏感资料使用 1 个主磁带池。

服务器、驱动器、磁带库均可形成小的逻辑单元,归档系统软件也需实现能够对这些小的逻辑单元进行调度使用,以满足系统整体的可扩展和动态灵活部署需求。系统建设过程中通过策略信息调度软件功能和硬件逻辑单元(图 3)来实现这一需求^[13]。策略信息即各类资料归档相关的软件功能对应服务器 IP 地址的规则,归档系统软件在元数据库中保存有归档整体业务的策略信息,调度管理服务在读取策略信息后形成一系列资料分拣、打包、探测、提元的任务到消息队列中,各个实际执行任务的进程主动到消息队列中提取相关的任务信息进行执行,当现有执行任务的进程无法满足系统要求时,则需增加新的策略记录,由新的服务器部署相关进程配合完成任务,实现系统的动态扩展。

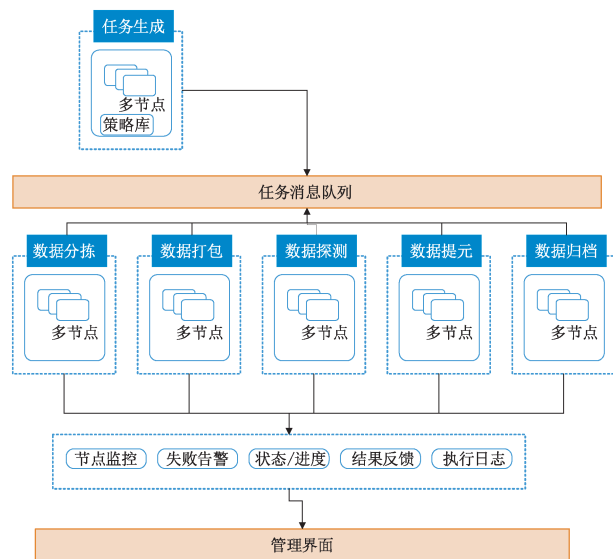


图 3 分布式调度管理流程

3.2 策略驱动数据归档

数据归档需要完成数据打包、数据提元、数据记带一系列操作。为适应各种资料以适当的方式进行归档,“天擎”归档系统将归档涉及的每一环节均以策略驱动运行,包括数据打包策略、数据提元策略、数据记带策略。

“天擎”归档系统每日接入的文件数量约为 410

万个,其中 94% 的文件为小于 1 MB 的文件,6% 的文件大于 1 MB,仅有约 100 个文件大于 1 GB。磁带库是机械设备,大量小文件的读写会造成频繁操作磁带而大大拉低磁带读写效率,因此对于归档系统一般建议归档文件大于 1 GB。“天擎”归档系统将大量文件打包以达到组织为大文件的目的。为了同时兼顾数据服务的方便性,对于不同的数据制定不同的打包策略,如地面资料按照省份名称将每日资料进行打包,雷达资料按照站号将一天的资料进行打包等。数据提元策略也可根据打包文件名情况进行灵活配置。

面对大量数据的归档需求,“天擎”归档系统通过减少磁带物理操作时间来提升磁带读写性能,磁带物理操作时间主要包括磁带的装载/卸载时间和磁带寻址时间。“天擎”归档系统通过调整归档系统数据记带策略,设定记带任务执行时间,以实现记带操作的集中并行运行,最大限度减少磁带装卸载次数和磁带寻址操作。通过比较,一天内执行 4 批次记带操作比 48 批次平均写带操作性能提升 50%,目前每日平均记带性能能够达到约 160 Mb/s。

3.3 多场景数据归档支撑

为了提高归档性能和归档服务的能力,满足大数据背景下的数据归档需求,“天擎”归档系统实现了多场景数据归档,包括多协议数据归档、敏感资料归档和磁带转储归档功能等。

3.3.1 多协议数据归档

“天擎”采用大数据技术支撑海量数据在线存储,采用了磁盘阵列存储、分布式文件系统和分布式对象存储等不同的存储技术实现数据存储,实现对不同应用的高效支撑。为支持海量数据异构存储的数据归档,归档系统增加主动获取资料数据归档功能,支持通过 NFS、FTP、S3 协议获取数据资料文件并进行归档(图 4),同时归档后的数据可在“天擎”归档系统门户中可以进行相应的数据检索、回取等功能。

3.3.2 敏感数据归档

“天擎”归档系统事先对不同安全级别数据的归档。为提升数据安全能力,“天擎”归档系统新增敏感性资料归档管理。如图 5 所示,归档系统对敏感资料划分单独的磁带池,通过配置密钥,将加密后的数据归档到磁带,具体功能为:①敏感资料集群划分。归档系统支持针对敏感性资料划分单独的文

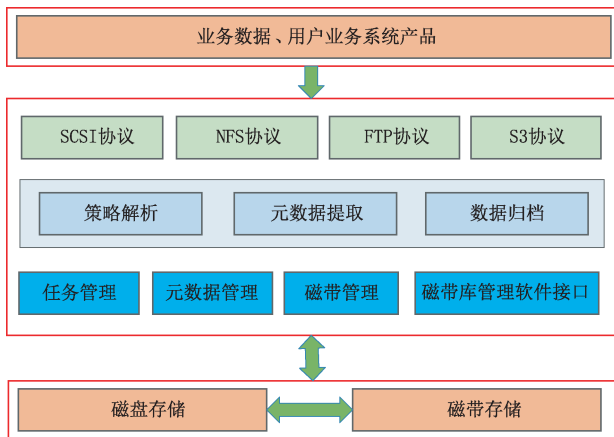


图 4 多协议数据归档架构

件系统目录,并指定单独的分拣、打包、探测、提元和归档服务器,同时划分单独的磁带池、指定单独的磁带进行归档。②敏感资料加密归档及回取。支持加密归档策略配置,可以设置敏感资料写磁带密钥,在归档资料写磁带环节,通过设置的密钥将资料文件加密,并将加密后的资料文件写到磁带上。

3.3.3 转储数据归档

磁带转储归档包括历史磁带转储和将来磁带转储。①历史磁带转储对 CIMISS 归档系统或其他系统通过 TSM 归档软件、IBM TS3584 磁带库写入的 LTO4 磁带采用 LTO7 磁带重新进行存储。这些 LTO4 磁带中有部分通过 TAR 命令直接写入磁带,导致磁带上记录的资料文件不规范,也缺乏归档资料元数据信息,造成转储这部分历史磁带的工作量和难度比较大。经过多次实验验证,通过导出 TSM 磁带元数据、分析磁带元数据、按资料批量回取、重新整理资料、记录元数据信息、数据重组、写入“天擎”归档系统磁带库等步骤能实现自动、快速、规范化的磁带转储功能(图 5)。②考虑到将来磁带技术的升级换代,如从 LTO7 升级到 LTO8、LTO9,归档系统通过磁带批量出入库操作,结合磁带复制技术、磁带元数据和归档元数据自动一致性更新、磁带审计等实现全自动化的磁带转储功能,满足未来磁带转储功能要求。

3.4 数据回取服务

“天擎”归档系统为满足不同用户在各自应用场景下回取数据的需求提供丰富的数据回取服务,包括订单式数据回取、脚本式数据回取、自动化数据回取以及敏感资料回取。

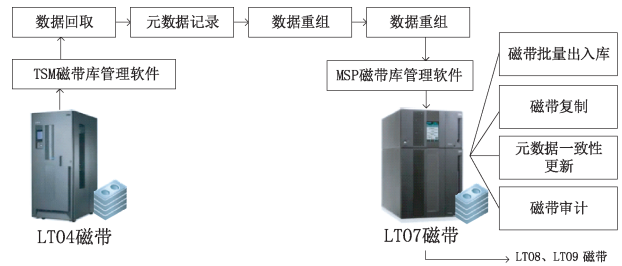


图 5 磁带转储流程

订单式数据回取是传统的数据回取方式,即用户通过“天擎”归档系统提供 Web 网站查询到所需回取的数据后提交数据回取订单,网站后台根据订单中数据文件请求清单逐一从磁带库中进行回取,然后将数据存放至用户指定的位置,订单完成后通知用户。

订单式回取主要满足的是用户在 WINDOWS PC 机进行数据回取的场景。现在大量业务运行在类 Linux 服务器中,用户对数据的使用可能会大量发生在服务器端。脚本式回取为“天擎”归档系统建设的新增功能,主要为满足用户在服务器端进行数据回取的需求。用户在类 Linux 的服务器上像使用 Linux 命令一样执行数据回取命令完成数据回取的操作。脚本式数据回取目前支持按照文件列表进行回取和按数据类型及时间范围进行回取两种方式。

由于“天擎”归档系统中存储的大部分非结构化数据与在线存储的数据形态一致,可与在线数据一同为用户服务,延长数据可服务的时长范围。为方便用户能够自动化使用数据回取服务,归档系统与数据服务接口(MUSIC 接口)和存储管理系统、解码入库系统共同配合完成归档近线、离线数据的自动化数据回取流程(图 6)。归档系统为 MUSIC 接口提供回取请求服务的 RESTFUL 接口,该接口能够接收 MUSIC 发起的回取请求,对请求进行解析后,向磁带库发起回取数据的流程,数据从磁带库回取后需要完成 2 个操作步骤:①将回取后的数据放在与解码入库系统共享的数据交换区;②向解码入库子系统发送数据已回取成功的消息。解码入库系统对回取后的数据进行入库操作后将回取状态对队列中的消息状态进行修改,改为数据可服务状态,MUSIC 接口发现消息状态为可服务状态后则反馈用户,用户即可获取数据。

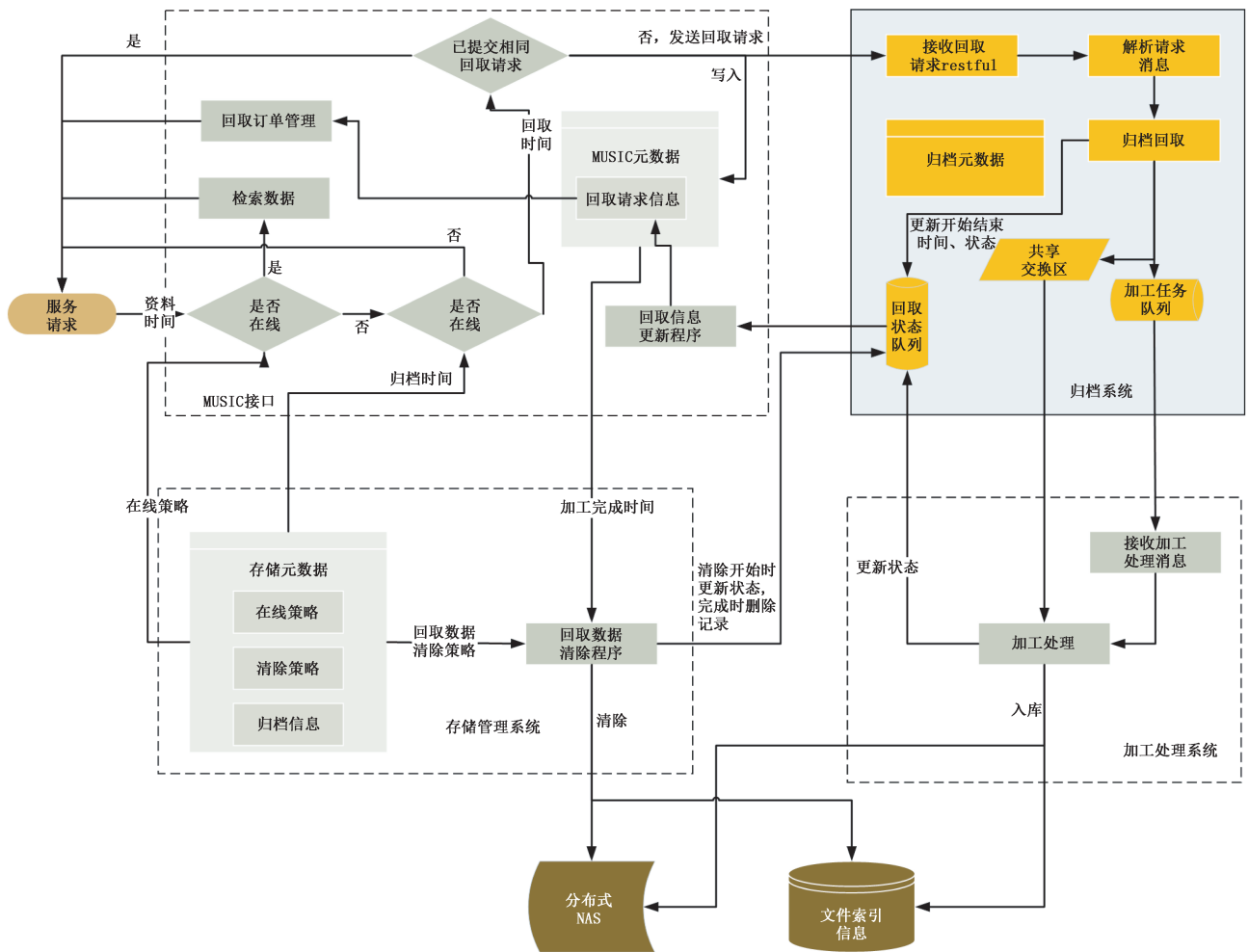


图 6 自动化数据回取服务流程

敏感数据是以加密方式进行归档的,因此敏感数据回取是通过密钥将磁带上的加密数据进行解密后回取。敏感性资料回取需要用户具备敏感资料访问权限才能进行,用户在通过认证和权限审核后,进行数据回取。回取时磁带驱动软件会对之前加密归档的文件进行解密,解密后将数据返回用户(图 7)。

3.5 归档数据全流程追溯

数据归档的目的是保障数据安全,所以数据归档的完整性、正确性对于完成归档业务和后续归档数据服务至关重要。“天擎”归档系统基于“天镜”(气象综合业务实时监控)实现对数据源和数据归档结果完整性对比的监控,在天镜的“天擎”归档系统的专题监视页面中显示归档完成的资料与数据源的资料数据量比较情况,以达到资料出现异常时,能够达到及时发现、及时告警和有效追溯的目的。监视页面如图 8 所示。

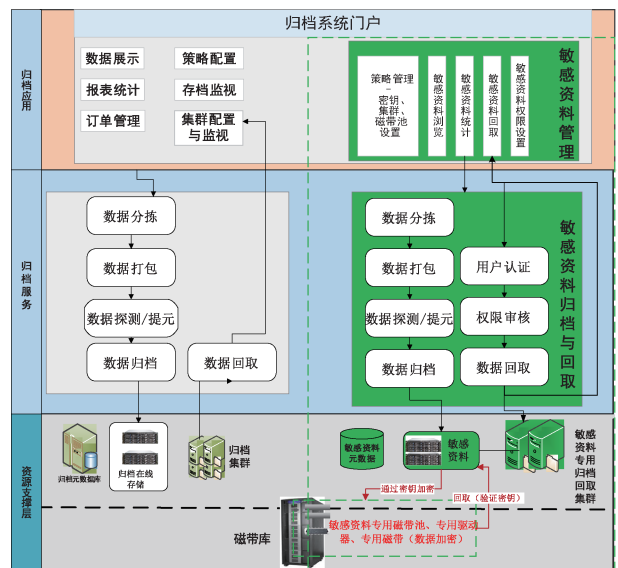


图 7 敏感资料归档与回取流程



图 8 “天擎”归档系统数据完整性监控页面

为方便技术人员熟悉了解资料的归档情况，“天擎”归档系统通过 H-Chart 图表插件设计了多种形式的报表对资料归档情况进行可视化展示，包括资料归档日报表、月报表和资料各环节归档监视展示，以及资料正确性和完整性验证功能展示(图 9)。

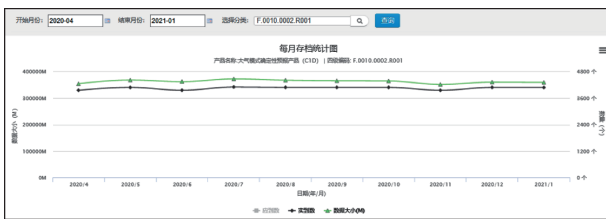


图 9 “天擎”归档系统数据统计报表页面

3.6 规范化的磁带管理

由于磁带对存储环境和保管方式都有特殊要求，需对磁带进行规范化的管理才能确保磁带中的数据可用。“天擎”归档系统中提供了全面的磁带管理功能，包括磁带池管理、磁带离线管理、磁带审记和磁带复制(图 10)。磁带池管理包括磁带池划分、磁带池监控。其中磁带池监控实现以磁带池为单位实时监控已用容量、剩余容量和空白磁带等，支持设定空白磁带告警阈值，提醒系统管理员及时向磁带库中添加磁带，以免影响归档业务运行；磁带离线管理中实现磁带的出入库、批量出入库管理。在批量出入库功能中，支持大于磁带库 I/E 口的上百盘磁带进行批量出库操作，大大提高了出入库操作的便利性和易用性。磁带审计根据磁带审计策略，自动对近线和离线磁带进行审计，以确认磁带介质可用性、磁带数据的一致性和完整性。磁带复制提供基于整盘的磁带介质克隆复制和资料文件级的复制功能，以应用于不同场景下磁带数据恢复。

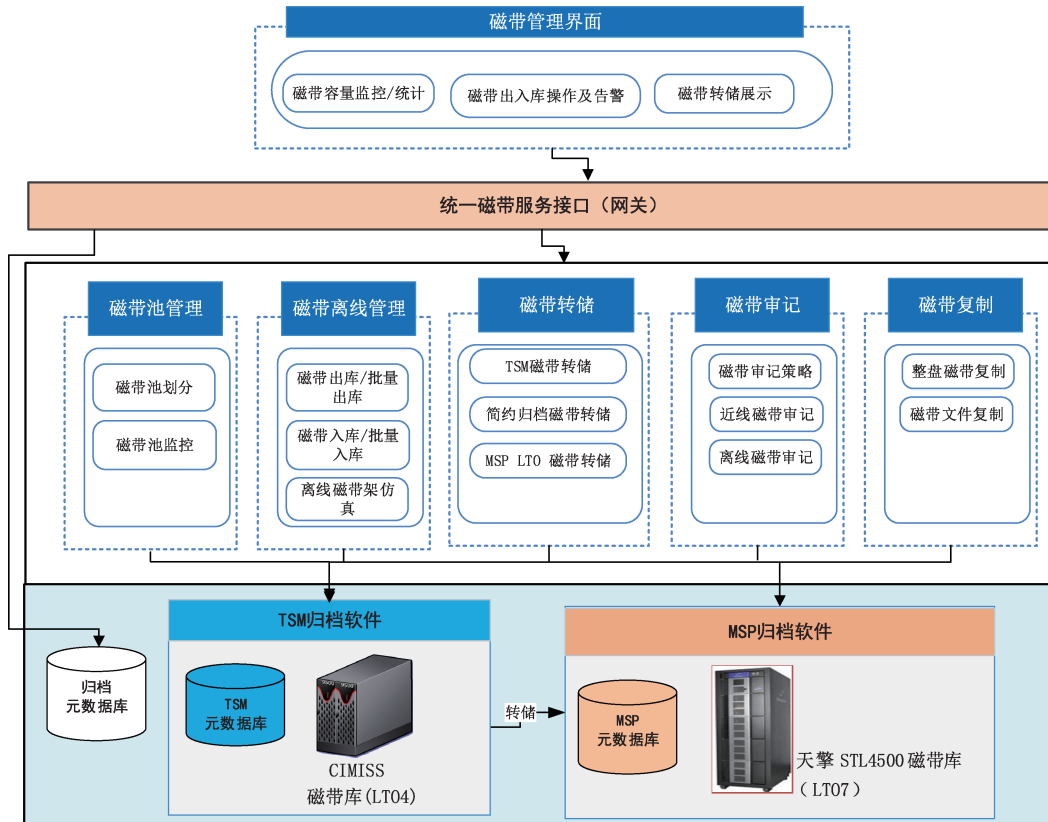


图 10 磁带管理功能

4 结论与讨论

“天擎”归档系统自2019年7月开始运行,2020年8月实现业务化,截止目前已完成12 PB数据的归档以及107 TB数据的回取,其中归档的数据包括实时数据9 PB以及2 PB高性能计算服务器上产生中国再分析资料和1 PB气候中心的CMIP6、CMOP模式历史数据。目前,每日实时数据的归档数据量为3.02~3.15 TB,接入的数据种类为703种,所有的业务数据均按照3副本的策略进行归档,每日实际实时数据的归档数据量达到近10 TB,磁带读写性能平均为160 MB/s。模式历史数据因文件均大于1 GB,归档时磁带写性能达到220 MB/s。从数据量来看,数值预报、卫星及雷达资料的数据量最大,三者数据量占总体数据量的96%;从文件数量来看,雷达和地面资料的文件数量最大,占总体文件量的85.7%以上。就目前的应用需求和发展趋势来看,数值预报资料将会是未来归档数据量的主要贡献者。

根据系统具备的硬件规模和系统目前服务器、驱动器的情况,如果将驱动器平均用于归档和回取,“天擎”归档系统最高可承载每日约240 TB数据的归档及回取检索,与欧洲中心MARS系统目前每日的250 TB数据增量相当。欧洲中心MARS系统和美国NOAA CLASS系统数据存储层采用High Performance Storage System(HPSS)来完成磁带库的管理操作,“天擎”归档系统使用国产归档存储中间件及自研数据处理和服务软件实现系统的整体管理。从可提供的日归档数据量级来看,“天擎”归档系统与发达国家气象归档业务管理系统相当,但由于刚投入业务使用,国内业务用户对“天擎”归档系统发起的实际业务请求数量与国外还有一定差距。在用户逐渐熟悉“天擎”归档系统的服务方式后,用户发起的业务请求会逐渐增多。

磁带存储能力持续提升和新技术的发展让磁带在大数据时代应用更具前景。2020年9月,LTO-9发布上市,LTO-9存储容量为LTO-7的3倍,达到了18 TB,同时LTO-9驱动器性能提高至400 MB/s。新的LTO路线图扩展到第12代,在磁带大小保持不变的情况下,未来LTO-12最高可达144 TB存储能力。考虑到磁带绿色节能和存储成本低的特点,在未来ZB级存储的时代,磁带存储可能发挥更大

的作用^[14-19]。磁带管理发展的新技术也改变了传统磁带的使用方式,如线性磁带文件系统(LTFS)技术使得磁带访问方式与磁盘一致。LTFS在磁带上创建数据区和索引区2个分区,数据区中存放了所有的数据内容与元数据,而索引区则存放数据区中数据的附属子集与附加元数据,这样用户搜索磁带上信息、搜索磁带的流程和磁盘存储一样。通过LTFS,用户能够很简易地和磁带进行互操作,可直接应用于备份和归档流程。LTFS还可与并行文件系统无缝衔接,通过策略配置可以自动将访问频率较低的数据通过LTFS迁移至磁带库,用户却可以做到无感知访问^[20]。这一技术将大幅提升用户体验,未来非常有必要在归档系统中探索使用。

参考文献

- [1] 沈文海. 试论智慧气象与数字化转型[J]. 中国信息化, 2019(3):7-9.
- [2] 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.
- [3] 赵芳,熊安元,张小纓,等. 全国综合气象信息共享平台架构设计技术特征[J]. 应用气象学报,2017,28(6):750-758.
- [4] 沈文海,赵芳,高华云,等. 国家级气象资料存储检索系统的建立[J]. 应用气象学报,2004,15(6):727-736.
- [5] 胡英楣,王甫棣,谭小华,等. 有状态消息队列在国内气象通信系统的应用[J]. 计算机系统应用,2020,29(3):121-126.
- [6] 林润生,孙周军,谭小华,等. 新一代国内气象通信系统设计与实现[J]. 气象,2011,37(3):356-362.
- [7] 张来恩,王鹏,韩鑫强. CTS2.0消息封装及交换控制策略设计及实践[J]. 气象科技进展,2018,8(1):271-273.
- [8] 侯明雨,李为冲,丁宁,等. 磁带新技术在数据长期归档中的应用[J]. 物探装备,2018,28(3):207-210.
- [9] 肖华东,孙婧,孙朝阳,等. 中国气象局S2S数据归档中心设计及关键技术[J]. 应用气象学报,2017,28(5):632-640.
- [10] 张玺,杨菲菲,董峰. TIGGE数据自动化归档技术研究及实现[J]. 气象科技,2015,43(4):622-626.
- [11] 钱建梅,郑旭东. 国家卫星气象中心气象卫星资料存档系统[J]. 应用气象学报,2003,14(6):756-762.
- [12] 钱建梅,孙安来,徐喆,等. 风云气象卫星数据存档与服务系统[J]. 应用气象学报,2012,2(3):369-376.
- [13] 王瑾,曹云鹏,王海峰. 面向大数据复杂应用的虚拟集群动态部署模型[J]. 计算机应用研究,2020,37(6):1760-1764.
- [14] Goodwin P. Tape and cloud: solving storage problems in the Zettabyte Era of Data[R]. IDC White Paper, 2019.
- [15] 胡燕平,胡冰. 大数据的长期归档存储研究[J]. 南京邮电大学学报(自然科学版),2014,34(4):33-38.
- [16] 蔡雨蒙,冷锴,单红伟,等. 大型医院影像分级存储架构下归档的策略及应用[J]. 中国数字医学,2017,12(1):48-51.

- [17] 刘颖,叶茂,王耀鲁,等.分级存储技术及其在水利大数据处理中的应用[J].水利信息化,2019(6):18-22.
- [18] Bradshaw P L, Brannon K W, Clark T, et al. Archive storage system design for long-term storage of massive amounts of data [J]. IBM J RES & DEV, 2008, 52(415): 379-388.
- [19] Lee D J, O'Sullivan M, Walker C G. Measurement for improving the design of commodity archival storage tiers [C] // Fourth IEEE International Conference on Utility and Cloud Computing, IEEE, 2012.
- [20] 谭乐娟,孙天.LTFS技术及其在“新华08”视频归档系统中的应用[J].中国传媒科技,2013(15):64-67.
- [19] Lee D J, O'Sullivan M, Walker C G. Measurement for impro-

Design and Implementation of Meteorological Big Data Cloud Platform Archive Storage System

LIU Yuanyuan HE Wenchun WANG Yan LI Jiangtao BAI Jinting
(National Meteorological Information Center, Beijing 100081)

Abstract: This article focuses on a detailed introduction of the construction requirements and overall architecture of the Meteorological Big Data Cloud Platform (“Tianqing”) archive storage system, and an in-depth analysis of several key technical issues of the platform implementation. “Tianqing” archive storage system is based on online storage, offline storage and tape drive management software. Through the hierarchical structure system, it establishes a dynamic and scalable archive cluster. The system implements multi-protocol data archiving and sensitive data archiving to meet various data archive scenarios. The system has data restoration, entire data process traceability, standardized tape management and other functions. In August 2020, the “Tianqing” archive storage system began to run operationally.

Keywords: big data; archive; retrieval; tape