

长时间序列格点数据管理平台的设计与实践

贾晓振 胡江凯* 王大鹏 梁晨

(中国气象局地球系统数值预报中心, 北京 100080)

摘要 随着数值模式时空分辨率的提升,数据量急剧增加,长序列数据很难直接通过文件拷贝或者网络传输方式为用户提供数据服务。为此,笔者设计实现了一种分布式管理平台,该平台根据用户定制的数据需求,运用预报要素、空间范围、时间尺度等约束条件,抽取或根据区域参数裁剪指定气象要素,生成精简数据进行用户服务。该平台集成了搜索引擎、格点数据解码、内存数据库技术以及分布式框架,实现跨操作系统的统一接口调用和数据快速获取,有效解决用户访问长时间序列历史资料的难题。实验测试显示,该平台在格点数据管理规模和访问效率方面均表现出色。特别是在北京 2022 年冬季奥运会和冬残奥会气象保障服务中,该平台发挥了重要作用,展现了其实际应用的价值和潜力。

关键词 分布式架构;搜索引擎;数值预报业务;内存数据库;数据检索;跨平台

中图分类号: P409 **DOI**: 10.19517/j.1671-6345.20230430 **文献标识码**: A

引言

近些年,随着数值预报业务的发展^[1],气象模式分辨率、预报时效的增长,模式输出的预报数据总量呈爆发式增多。中国气象局地球系统数值预报中心,作为国家级气象模式研发及业务中心,截至 2023 年,其主要业务模式的格点数据输出量,达到每日 0.76 TB,年输出量 155.12 TB。2016—2020 年的历史模式版本的格点数据量达到 134.56 TB,2021—2023 年的历史数据量达到 418.75 TB。此外,模式逐步从单一的气象、气候或海洋模式向地球系统模式耦合发展^[2-3]。届时,模式所生产出的格点数据的规模更加庞大。Wagemann 等^[4]指出,科研用户在未来更期望从云平台或其他种类的接口来获取数据,而非只通过下载这一渠道。Ryan 等^[5]也曾指出,60%的科研时间花费在数据的组织和清理上。而针对不同用户的个性化数据集需求,数据管理员需使用 wgrib2^[6]工具反复从二进制编码的气象格点数据中提取某些要素。此过程需要不断重复对压缩二进制编码进行解码,在不同数据服务场景

中提取同一条数据时重复解码将会降低访问性能。

目前,国内外已有一些研究诸如 ECMWF 的数据管理系统 MARS 等技术用于解决这个问题。例如 MARS 等技术,是将数据切片存储至分布式存储以加速提取数据^[7-8],采用此种方法管理数据,往往需要付出更高规模的冗余存储来存放切片数据。为解决数据服务过程中遇到的长时间序列历史格点数据提取难的问题,同时考虑实际的存储规模,笔者提出一种轻量级^[9]分布式并行处理^[10]格点数据管理框架,针对格点数据原始文件,通过获取数据文件存放路径,不增加原始数据存储需求,对预报格点数据文件只进行一次解码,实现多次检索和提取的加速管理,并且支持跨平台获取数据,提高检索提取数据效率。

1 总体框架

结合搜索引擎^[11]和格点数据解码技术^[12-13],设计了图 1 所示架构。通过搜索引擎快速检索元数据,而解码技术则用于元数据获取。为提高效率,采用可扩展的消息队列框架^[14]进行分布式处理,加速

http://www.qxkj.net.cn 气象科技

作者简介:贾晓振,男,1992,学士,工程师,主要从事数值预报业务运行维护及数据管理方面的工作,Email:jiaxzh_cma@163.com

收稿日期:2023 年 12 月 12 日;定稿日期:2024 年 10 月 9 日

* 通信作者,Email:hujk@cma.gov.cn

©《气象科技》编辑部 (CC BY-NC-ND 4.0)

数据提取。同时,数据编目模块通过构件式开发^[15]以支持 NetCDF^[16-18]等多种文件格式的文件,可灵活升级和扩展。

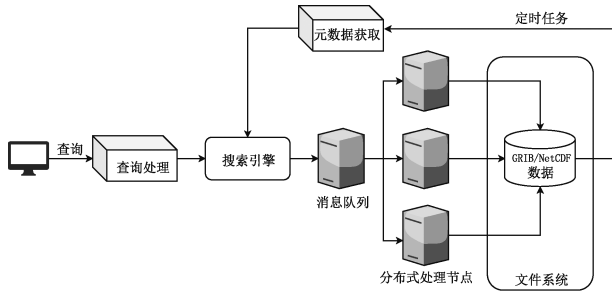


图 1 本设计的分布式架构

Fig.1 Distributed architecture of this design

图 2 展示了设计的三层架构体系,数据层存储数据、算法层实现功能、业务层对接用户。①数据层存搜索与共享功能所需数据,未来还将增加用户行为数据,以优化整体设计。②算法层涵盖搜索与编目核心算法。编目模块建索引,数据抽取模块检索并提取数据,网络交互模块调度算法。③业务层提供数据编目与抽取接口,解析用户指令,与网络交互模块通讯,并接收数据。

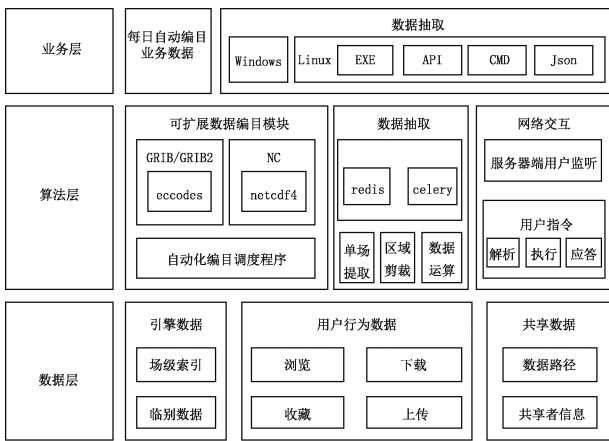


图 2 本设计的原型架构

Fig.2 Prototype architecture of this design

2 主要功能

2.1 模式格点数据编目

用户需编目原始数据文件时,可调用编目接口并输入表 1 所示参数。系统解析文件关键信息,如单要素场特征值、偏移量等,生成元数据索引,完成编目。

表 1 模式格点数据逐文件编目模块介绍

Table 1 Introduction to the module for cataloging the grid point data in each file

Table with 2 columns: 名称 (Name) and 介绍 (Introduction). It lists parameters like 原始数据 (Original Data), 索引名称 (Index Name), and 索引类型 (Index Type), along with their descriptions and output information.

2.2 模式格点数据抽取

调用检索接口,给定检索参数获取指定数据。图 3 展示了在没有输入规范参数,或者使用-h 参数调用检索接口时命令行调用结果。

```
usage:
getdata.py -h 帮助
            -d [date]预报日期YYYYMMDD
            -t [HH]起报时次HH, 多时次选择可用冒号(:)隔开
            -l [levels]要素层次, 多层次选择可用冒号(:)隔开
            -s [stepRange]预报时效, 多时效选择可用冒号(:)隔开
            -n [shortName]要素场/变量名称,多要素选择可用冒号(:)隔开
            -f [function]数据操作:
               区域截取 latlon:lon1:lon2:lat1:lat2
            -i <index>索引类型
               GRAPES_GFS模式格点数据gfs
               GRAPES_MESO_3KM模式格点数据meso_3
               GRAPES_MESO_10KM模式格点数据meso_10
            -o <outputdir>下载数据目录
            -w [filename]指定下载数据名称
```

图 3 实现的用户命令行接口

Fig.3 Implementation of the user command line interface

输入规范参数调用检索接口,会返回如图 4 展示的检索信息,包括索引名称、检索条件、数据检索命中条目、抽取结果和任务启动与完成时间。抽取数据为 2023 年 1 月 00:00 起报的 2 m 相对湿度 (RH),全月共抽取到不同层次、预报时效 RH 共计 2263 个,总文件大小 8.4 GB。

2.3 模式格点数据裁剪

配置数据检索接口的-f 参数裁剪指定经纬度的数据。图 5 展示了调用裁剪功能获取范围经度 113°~120°E,纬度 44°~37°N 的数据的命令行参数及获取

```

Fri Nov 24 11:21:52 GMT 2023
Starting the Searching ...
eccode_meso_3_2023_grb_index
['-d','202301','-t','00','-l','2','-n','RH','-i','meso_3_2023','-o','./TEST_CASES_RESULT',
'-w','TESTDATA.2023.01.RH.2m.00.GRIB2']
connect redis ok
unlock_num:
2263
cutList_num:
2263
Finish file extract!
Fri Nov 24 11:23:02 GMT 2023
8.4G Nov 24 11:22 TESTDATA.2023.01.RH.2m.00.GRIB2.grib2

```

图 4 数据检索接口检索数据

Fig. 4 Data retrieval interface for data retrieval

的结果数据, latlon 参数用于指明经纬度范围, 参数格式为 latlon:起始纬度:截止纬度;起始经度:截止经度, latlon 与各经纬度参数之间使用半角冒号“:”相连接。原始数据为 2023 年 1 月 00:00 起报的 2 m 相对湿度。裁剪前数据量为 8.4 GB, 裁剪后数据量为 9.8 M, 裁剪总耗时在 1 min 以内。

```

Sun Nov 26 15:20:16 GMT 2023
Starting the Searching ...
eccode_meso_3_2023_grb_index
['-d','202301','-t','00','-l','2','-n','RH','-f','latlon:44.37:113:120',
'-i','meso_3_2023','-o','./TEST_CASES_RESULT',
'-w','TESTDATA.2023.01.RH.2m.00.GRIB2']
connect redis ok
unlock_num:
2263
cutList_num:
2263
Finish file extract!
Sun Nov 26 15:20:59 GMT 2023

9.8M Nov 26 15:20 latlon.TESTDATA.2023.01.RH.2m.00.GRIB2.grib2
8.4G Nov 24 11:22 TESTDATA.2023.01.RH.2m.00.GRIB2.grib2

1:0:vt=2023013015:2m above ground:63 hour fcst:RH Relative Humidity [%]:
ndata=2501:undef=0;mean=34.3956:min=9.26:max=85.9163
grid_template=0:winds(N/S):
lat-lon grid:(2501x1)units 1e-06 input WE:NS output WE:SN res 48
lat 60.099998 to 60.099998 by 0.000000
lon 70.000000 to 145.000000 by 0.030000 #points=2501

```

图 5 数据检索接口裁剪数据

Fig. 5 Data retrieval interface cropping data

3 关键技术

3.1 格点数据访问技术

以 CMA-GFS^[19] 模式输出的 GRIB2 数据文件为例。其二进制段结构如图 6 所示, 由多个独立的要素场串联而成。但由于 GRIB2 文件的自描述性, 传统方式访问 GRIB2 文件, 需要将整个 GRIB2 文件加载到内存才能进行解码, 进而选取所需气象要素场^[20]。因此, 重复抽取大量单要素场时, 需要反复解码整个二进制文件, 带来较高的时间开销。

为减少解码时间开销, 利用二进制存储特

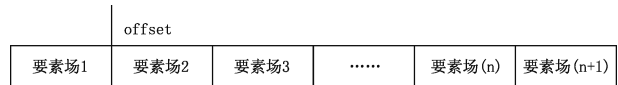


图 6 GRIB2 资料格式示意

Fig. 6 Schematic of GRIB2 data format

性^[21], 根据要素场在格点数据中的偏移量和段长度创建索引。这样, 复用数据时无需重复解码。偏移量和段长度是抽取要素场的关键参数, 而约束条件则用于快速检索这些参数。以 CMA-GFS 的 GRIB2 数据为例, 约束条件包括气象要素、起报信息、预报时效等, 与 GRIB2 数据中的关键字有明确的对应关系(表 2)。

表 2 GRIB2 数据参数对照

Table 2 GRIB2 data parameter comparison

特征值	GRIB2 编码中的关键字
气象要素名称	shortName
起报日期	date
起报时次	time
预报时效	stepRange
层次	level

图 7 为某一气象单要素场的 GRIB2 头信息, 其中包含建立索引模型所需的所有约束条件, 包括 ShortName、date、time、stepRange、level。

```

1 #===== MESSAGE854(length=1083) =====
2 GRIB{
3 #Meteorological products(grib2/tables/4/0.0.table)
4 discipline = 0;
...
11 dataDate = 20210601;
12 dataTime = 1200;
...
42 #Therm odynamic Stability indices(grib2/tables/4/4.1.0.table)
43 parameterCategory = 7;
44 #Convective available potential energy (Jkg-1) (gtib2/tables/4/4.2.0.7.table)
45 parameterNumber = 6;
...
56 stepRange = 9;
...
69 level = 0;
...
72 shortName = unknown;
73 }

```

图 7 GRIB2 数据头信息中的约束条件

Fig. 7 Constraints in GRIB2 data header information

结合要素场约束条件、偏移量、段长度实现单要素场检索并抽取。图 8 展示了约束条件检索偏移

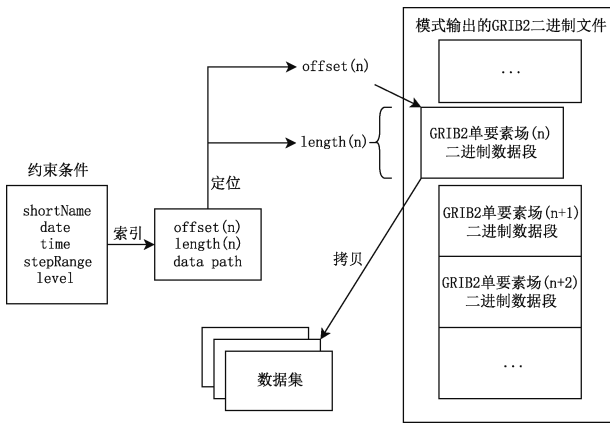


图 8 通过索引获取段数据

Fig. 8 Obtain segment data through indexing

量、段长度和文件位置,利用偏移量和段长度提取单要素场二进制数据段的过程。

3.2 基于搜索引擎的索引设计

设计采用倒排索引 value-id 模型^[22],其优势是通过属性值快速检索数据存储位置,但建立索引时时间开销较大^[23]。相比之下,传统工具 wgrib2 在大规模提取单要素场时,需多次解码二进制文件,效率较低(表 3)。

表 3 wgrib2 提取要素场耗时

Table 3 Time consuming of extracting feature fields with wgrib2

提取要素场个数	数据在 GRIB 文件中的位置	耗时/s	
		平均单要素	提取总
30	头部	0.58	17.34
30	中部	0.62	18.53
30	尾部	0.58	17.51

GRIB2 数据具有自解释性,提取任意位置的单要素场都需对整个数据进行解码。通过建立索引,利用约束条件获取偏移量和段长度,可避免重复解码的时间开销。虽然建立倒排索引的时间开销较大,但对用户来说是透明的。

表 4 为设计的第一代索引模型,该索引模型在检索气象要素时存在问题,因为某些要素的 shortName 可能为 unknown^[24],导致无法精确定位。

表 4 索引模型-I

Table 4 Index Model-I

Id	约束条件
DataPath # Offset # Length	shortName, date, time, stepRange, level

为解决这一问题,设计引入使用图 7 中 GRIB2 三位编码^[25]; discipline、parameterCategory 和 parameterCategory 作为新的约束条件,如表 5 所示三位编码能唯一指定一个变量类型,从而替代 shortName 的功能,实现更准确的检索。

表 5 GRIB2 数据三位编码示例

Table 5 Example of three-digit encoding for GRIB2 data

三位编码	三位编码含义	取值示例	取值示例含义
discipline	产品类型编码	0	气象产品
parameterCategory	参数类别	7	热力学稳定性指标
parameterNumber	参数编码	6	对流有效位能

结合表 4 中 date、time、stepRange、level 约束条件和单要素场偏移量、段长度信息,得到满足索引设计需求的索引模型-II 如图 9 所示:

```

1  {
2      "_id": "DataPath&offset&length",
3      "_value": {
4          "data": "dataDate",
5          "stepRange": "stepRange",
6          "level": "level",
7          "shortName": "shortName",
8          "time": "dataTime",
9          "discipline": "discipline",
10         "parameterCategory": "parameterCategory",
11         "parameterNumber": "parameterNumber",
12     }
13 }

```

图 9 索引模型-II

Fig. 9 Index Model-II

3.3 分布式集群框架

为提升数据处理性能,设计采用如图 10 所示的分布式节点构建 GlusterFS 共享存储^[26],并使用 Celery 分布式消息队列^[27]管理各节点 CPU。数据编目模块、数据检索模块作为应用部署到集群中。

在分布式架构下,各节点将抽取的单场数据存储于本地,数据集成为主要问题^[28]。为解决此问题,设计采用 Redis 内存数据库集群^[29]。内存数据库部署在主节点,各分节点通过网络访问。检索完成后,分节点发送抽取到的单要素场至内存数据库中。主节点读取内存数据库内容完成数据汇聚如图 11。

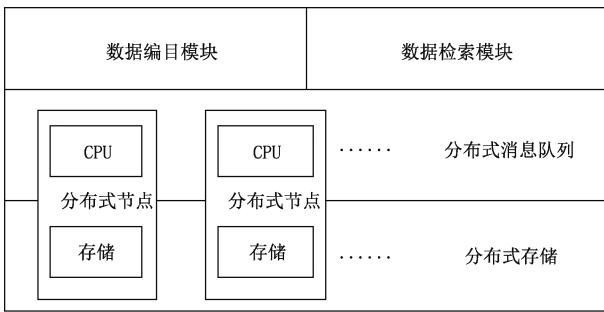


图 10 分布式集群架构

Fig. 10 Distributed cluster architecture

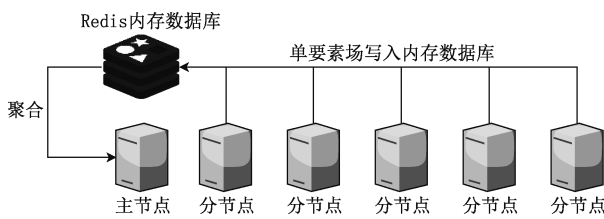


图 11 内存数据库部署及数据流

Fig. 11 Memory database deployment and data flow

在数据处理流程中,由于使用了异步触发^[30],主节点需要判断是否所有分节点数据已进入内存数据库。设计采用的判别方式为:判断检索命中条目数是否等于内存数据库中的数据条目数。条目数相等即表示所有节点数据已进入内存数据库,这种方法简单有效,能够确保数据的完整性和一致性。

3.4 网络交互

网络交互模块中,设计采用 gRPC 技术实现 C/S 通信^[31-33]。客户端程序将收集到的用户指令信息进行转码,转换为适合数据检索模块指令的格式,加入消息头 Action,生成如图 12 所示消息体结构,最后将转码数据发送至 gRPC 中间件,gRPC 通过网络向服务端发送指令。

HEAD	BODY
Action	Parameters
Index	FILE_PATH, MODEL_NAME TRAP SLEEP
Extract	INDEX, JSON, NOUN

图 12 C/S 传输的消息体示意

Fig. 12 Schematic of the message body transmitted by C/S

消息体由 Action 和 BODY 两部分组成。Action 部分决定 Server 端执行的调度行为,BODY 中保存了执行调度行为必须的参数。参数说明见表 6。

表 6 C/S 端传输消息体详表

Table 6 Detailed table of C/S transmission message body

参数名	含义	说明
Action	调度行为	指定 Server 调度编目 Index 或数据抽取 Extract
FILE_PATH	数据路径	指定需要编目的数据文件路径
MODEL_NAME	模式名称	创建索引名称时使用模式名做区分
TRAP	重做次数	异步创建索引失败自动重做次数
SLEEP	挂起时间	自动重做的时间间隔
INDEX	抽取数据所在索引	eccode_gfs_2023_grb_index
JSON	抽取数据约束条件	根据用户指令转码的约束条件 JSON 文件
NOUN	抽取数据执行其他操作	根据用户指令转码的其他计算操作,如裁剪参数

服务端的守护进程持续监听 gRPC 通讯端口。在接收到客户端消息后,首先解析消息头,判断指令类型。当指令类型为编目 Index 时,服务端按照编目消息格式解析消息体,获取编目关键字数据目录、模式名等信息。之后服务端会调度编目程序对数据进行编目。当指令类型为抽取时,服务端会从消息体中提取数据目录、偏移量等信息,并调度抽取程序,抽取完成后,数据汇聚至主节点。并最终通过 HTTPS 服务发送回客户端如图 13。

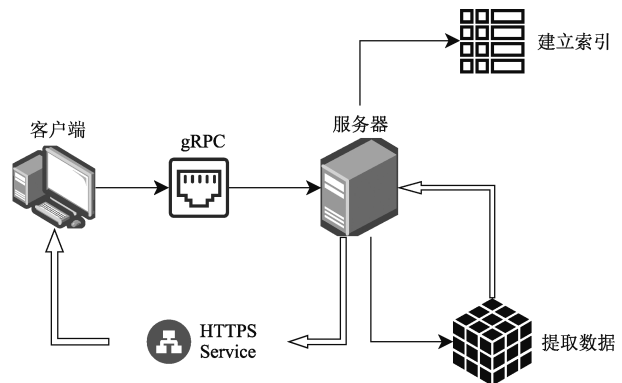


图 13 C/S 调度流程

Fig. 13 C/S Scheduling Process

4 测试与讨论

设计并实现的分布式系统已完成框架搭建及关键模块开发,涵盖索引编码、数据检索处理与 GRPC 远程通信,从而初步构建了数据管理与服务能力。在格点数据管理、访问效率及多用户支持方面成果显著。特别是在北京冬奥会气象保障中,本系统成功处理了 PB 级的格点数据,为“冬奥会气象条件预测保障关键技术”项目提供了 2016—2020 年的历史数据处理服务。此举不仅打破了传统数据处理的限制,更将数据量精简至原来的 1/500,为冬奥预报服务提供了坚实支撑。表 7 详细展示了本系统处理 CMA-GFS 模式输出的冬奥气象保障数据集及其原始数据规模。

表 7 北京 2022 年冬奥会和冬残奥会气象保障服务数据集
Table 7 Beijing 2022 Winter Olympic and Paralympic Winter Games meteorological support service dataset

格点数据来源	提取后大小	要素场个数	原始数据大小
	GB		TB
CMA-GFS(2016)	1.6	938432	7.75
CMA-GFS(2017)	2.6	1632205	13.60
CMA-GFS(2018)	4.0	2556373	20.66
CMA-GFS(2019)	3.6	2224602	48.10
CMA-GFS(2020)	3.6	2287021	44.45

4.1 格点数据管理规模

目前,已对中国气象局的主要数值预报业务和准业务模式输出的格点数据进行了编目管理。这涵盖了 CMA-GFS 全球模式从 2020—2023 年输出的原始格点数据,以及 CMA-MESO-3 km 区域 3 km 预报模式在相同时间段内的原始格点数据。然而,由于存储空间的限制,部分年份中的某些数据已被清理。表 8 详细列出了各模式已管理的数据条目数。

各模式输出的 GRIB2 数据信息见表 9,包括每类数据的单时效、全层次、全要素文件大小以及单时效文件中所有预报场数量,以及单数据索引(全年)原始数据的数据大小。

4.2 格点数据访问效率

为测试本平台原型访问格点数据效率,设计了 5 个测试用例(表 10)。实验从已管理的数据中抽取指定数据,抽取得到的结果数据大小超过 1 GB,共

进行 5 次实验,每次实验检索不同类型、不同时间的数据具体检索内容。

表 8 各模式已管理数据条目

Table 8 Managed data entries in each mode

格点数据来源	索引名称	索引条目
CMA-GFS(2020)	eccode_gfs_2020_grb_index	96763059
CMA-GFS(2021)	eccode_gfs_2021_grb_index	62323020
CMA-GFS(2022)	eccode_gfs_2022_grb_index	32663925
CMA-GFS(2023)	eccode_gfs_2023_grb_index	50146285
CMA-MESO-3 km(2020)	eccode_meso_3_2020_grb_index	41364308
CMA-MESO-3 km(2021)	eccode_meso_3_2021_grb_index	61695504
CMA-MESO-3 km(2022)	eccode_meso_3_2022_grb_index	14875429
CMA-MESO-3 km(2023)	eccode_meso_3_2023_grb_index	57766800

表 9 各模式单要素场数据大小

Table 9 Size of single-element field data in each mode

格点数据来源	平均单时效		全年数据量
	文件大小	单时效要素场个数	
	MB		TB
CMA-GFS(2020)	420	837	44.45
CMA-GFS(2021)	420	837	29.79
CMA-GFS(2022)	420	858	30.13
CMA-GFS(2023)	420	859	67.86
CMA-MESO-3 km(2020)	970	508	77.35
CMA-MESO-3 km(2021)	970	508	116.13
CMA-MESO-3 km(2022)	996	557	147.50
CMA-MESO-3 km(2023)	1203	575	136.48
总计			649.69

表 10 GB 级结果分钟级抽取实验编号及内容

Table 10 GB-level results extraction experiment number and content

实验编号	测试内容
TEST1	2023 年 2 月露点温度(DPT)
TEST2	2023 年 1 月露点温度(DPT)
TEST3	2023 年 1 月 2 m 温度(TMP)
TEST4	2023 年 1 月 2 m 相对湿度(RH)
TEST5	2023 年 6 月 2 m 相对湿度(RH)

每次实验共重复 10 组,通过统计各组试验的总耗时的最大值、最小值、中位数、均值,生成表 11。总耗时包括检索数据耗时,检索后对数据进行抽取

的耗时,以及将抽取到的单预报场文件合并的数据合并耗时三个部分。

表 11 各用例在 10 组实验中的耗时

Table 11 Time consumption in 10 sets of experiments for each use case

实验 编号	数据大小 GB	文件 数量	耗时/s			
			最小值	最大值	中位数	均值
TEST1	12	22	14	16	3.0	1036
TEST2	13	23	17	17	3.2	1147
TEST3	12	20	16	16	3.0	1147
TEST4	17	24	20	20	4.3	1147
TEST5	16	22	18	18	3.9	1110

通过 10 组重复实验得到的耗时统计结果达到抽取 GB 级结果数据可在分钟级完成的目标。

各实验所涉及任务数量级各节点执行情况。如表 12 所示,每个实验的查询抽取任务,由分布式调度系统,根据各节点健康状态,分配至各分节点中执行。以 GROUP.5 为例,包含主节点在内的 6 个分布式集群节点,分别执行 5 个用例的任务个数。

表 12 抽取实验各分布式节点任务执行数量

Table 12 Number of tasks executed by each distributed node in the extraction experiment

实验编号	主节点	节点-1	节点-2	节点-3	节点-4	节点-5	合计
TEST.1	211	192	170	153	154	156	1036
TEST.2	200	198	188	188	186	187	1147
TEST.3	265	213	159	170	168	172	1147
TEST.4	223	207	184	179	175	179	1147
TEST.5	252	195	166	166	165	166	1110

表 13 为 5 个用例的平均抽取总耗时与串行抽取总耗时对比。

表 13 抽取实验各分布式节点执行情况

Table 13 Execution status of each distributed node in the extraction experiment

实验编号	TEST1	TEST2	TEST3	TEST4	TEST5
预报场个数	1036	1147	1147	1147	1110
10 组实验	16	17	16	20	18
平均总耗/s					
串行抽取数据	953	1114	1126	1137	1021
总耗时/s					

通过对比可看出,抽取 1000 余个预报场时,使用数据平台的总耗时均小于 20 s,而串行方式耗时均大于 900 s。不对预报场进行抽取,直接对全要素

进行拷贝,以 2023 年 1 月整月的 CMA_MESO 数据量为例,全要素拷贝数据量约为 12326 GB。

5 结论与讨论

该平台设计结合了分布式架构、数值预报业务、数据管理、数据检索和跨平台等技术,实现了对大规模格点数据的高效管理和服务。随着实验测试的验证和北京 2022 年冬季奥运会和冬残奥会气象保障服务的成功应用,证明该长时间序列历史格点数据管理平台的设计与实践在定制 GRIB/GRIB2 格式数据的长时间序列数据集工作上是有用的,特别是从已编目数据的数据文件中,反复提取不同要素方面的工作,其效率比反复使用无索引工具提取数据,有显著提升。但是,对于从尚未进行编目管理的数据文件中提取要素场的情况,本方法需要进行额外的处理。例如开发对应数据文件格式的编目组件,或者编目录入数据。为此带来的时间开销,也应当一并计入数据集制作的时间开销中。随着气象模式的不断发展和数据量的持续增加,该平台将为气象预报和气候预测提供更加可靠和高效的数据支持,并为气象服务领域的发展做出贡献。

参考文献

- [1] 沈学顺,王建捷,李泽椿,等. 中国数值天气预报的自主创新发展[J]. 气象学报, 2020, 78(3): 451-476. Shen X S, Wang J J, Li Z C, et al. China's independent and innovative development of numerical weather prediction [J]. Acta Meteorologica Sinica (in Chinese), 2020, 78(3): 451-476.
- [2] 赵宗慈,罗勇,黄建斌. 从检验 CMIP5 气候模式看 CMIP6 地球系统模式的发展[J]. 气候变化研究进展, 2018, 14(6): 643-648. Zhao Z C, Luo Y, Huang J B. The detection of the CMIP5 climate model to see the development of CMIP6 earth system models [J]. Advances in Climate Change Research (in Chinese), 2018, 14(6): 643-648.
- [3] 袁林旺,闫国年,谢志仁,等. 地球系统科学数据共享网的气候-海面变化数据集建设[J]. 地球信息科学, 2008, 10(1): 1-6. Yuan L W, Lu G N, Xie Z R, et al. Discussion on the construction of climate & sea level change dataset in the earth system science data sharing network [J]. Journal of Geo-information Science (in Chinese), 2008, 10(1): 1-6.
- [4] Wagemann J, Clements D O, Figuera R M, et al. Geospatial web services pave new ways for server-based on-demand access and processing of Big Earth Data [J]. International Journal of Digital Earth, 2018, 11(1/3): 7-25.
- [5] Abernathey R, Augspurger T, Banihirwe A, et al. Cloud-

- native repositories for big scientific data [J]. *Computing in Science & Engineering*, 2021, 23(2): 26-35.
- [6] 王兵, 李杰. 基于通用模型的 GRIB 格式数据读取技术[J]. *航空计算技术*, 2018, 48(6): 96-101. Wang B, Li J. Data reading method of GRIB format based on general model [J]. *Aeronautical Computing Technique (in Chinese)*, 2018, 48(6): 96-101.
- [7] 肖华东, 孙婧, 张玺, 等. MARS 软件在数值预报模式产品数据管理中的应用[J]. *应用气象学报*, 2015(2): 247-256. Xiao H D, Sun J, Zhang X, et al. Application of the MARS to data management of NWP productions [J]. *J Appl Meteor Sci (in Chinese)*, 2015, 26(2): 247-256.
- [8] 李永生, 李高洁, 陈逸智, 等. 高频精细化气象格点数据实时处理系统设计与实现[J]. *气象科技*, 2022, 50(6): 776-782. Li Y S, Li G J, Chen Y Z, et al. Design and implementation of a high-frequency fine meteorological grid data real-time processing system [J]. *Meteorological Science and Technology (in Chinese)*, 2022, 50(6): 776-782.
- [9] 吴秀芸, 王海江, 梁寒冬. 轻量级空间数据引擎的应用研究[J]. *地理空间信息*, 2017, 15(12): 48-50. Wu X Y, Wang H J, Liang H D. Application research of lightweight spatial data engine [J]. *Geospatial Information (in Chinese)*, 2017, 15(12): 48-50.
- [10] 何险峰, 马力, 罗永康. 分布式气象内容管理系统设计[J]. *气象科技*, 2013, 41(6): 1036-1042. He X F, Ma L, Luo Y K. Design of a distributed meteorological content management system [J]. *Meteorological Science and Technology (in Chinese)*, 2013, 41(6): 1036-1042.
- [11] 王甫棣, 王帅, 赵希鹏. 基于 Elasticsearch 的异构数据迁移方法[J]. *气象科技*, 2022, 50(1): 45-50. Wang F D, Wang S, Zhao X P. A method for heterogeneous data migration based on Elasticsearch [J]. *Meteorological Science and Technology (in Chinese)*, 2022, 50(1): 45-50.
- [12] 刘媛媛, 应显勋, 赵芳. GRIB2 介绍及解码初探[J]. *气象科技*, 2006, 34(z1): 61-64. Liu Y Y, Ying X X, Zhao F. Introduction of GRIB2 and GIB2 decoding [J]. *Meteorological Science and Technology (in Chinese)*, 2006, 34(z1): 61-64.
- [13] 郑波, 李湘, 何文春, 等. 基于 CIMISS 全国精细化格点预报业务数据环境系统设计与实现 [J]. *气象科技*, 2018, 46(4): 670-677. Zheng B, Li X, He W C, et al. Design and implementation of data environment for CIMISS-Based nationwide refined forecast system [J]. *Meteorological Science and Technology (in Chinese)*, 2018, 46(4): 670-677.
- [14] 霍达, 宋利. 基于 Celery 的分布式视频计算处理框架[J]. *电视技术*, 2016, 40(4): 12-17. Huo D, Song L. Distributed video processing system based on Celery [J]. *Video Engineering (in Chinese)*, 2016, 40(4): 12-17.
- [15] 夏洁武, 钱乐秋. 基于构件/构架的软件开发方法[J]. *计算机应用与软件*, 2001, 18(10): 24-27. Xia J W, Qian L Q. Component/architecture-based software development [J]. *Computer Applications and Software (in Chinese)*, 2001, 18(10): 24-27.
- [16] 刘俊, 黄兴友, 周红根, 等. 基于 NetCDF 的国产新型气象雷达基数据通用编码试验[J]. *气象科技*, 2016, 44(2): 171-176. Liu J, Huang X Y, Zhou G H, et al. Unified coding for domestic new weather radar data based on NetCDF [J]. *Meteorological Science and Technology (in Chinese)*, 2016, 44(2): 171-176.
- [17] 查石祥. NetCDF 数据的程式访问[J]. *气象科技*, 2004, 32(z1): 40-43. Zha S X. Programming access to NetCDF Files [J]. *Meteorological Science and Technology (in Chinese)*, 2004, 32(z1): 40-43.
- [18] 刘培宁, 韩笑, 杨福兴. 基于 R 语言的 NetCDF 文件分析和可视化应用[J]. *气象科技*, 2014, 42(4): 629-634. Liu P N, Han X, Yang F X. Visualized application and analysis of NetCDF files based on R language [J]. *Meteorological Science and Technology (in Chinese)*, 2014, 42(4): 629-634.
- [19] 张进, 孙健, 沈学顺, 等. CMA-GFS V4.0 模式关键技术研发和业务化[J]. *应用气象学报*, 2023, 34(5): 513-526. Zhang J, Sun J, Shen X S, et al. Key model technologies of CMA-GFS V4.0 and application to operational forecast [J]. *J Appl Meteor Sci (in Chinese)*, 2023, 34(5): 513-526.
- [20] 冯建设, 王建源, 崔玉东, 等. 气象要素场客观分析软件[J]. *气象科技*, 2005, 33(4): 382-384. Feng J S, Wang J Y, Cui Y D, et al. An objective analysis software for meteorological elements [J]. *Meteorological Science and Technology (in Chinese)*, 2005, 33(4): 382-384.
- [21] 李忠武, 陈丽清. 存储器的管理机制—分段技术与寻址方案[J]. *数字技术与应用*, 2016(1): 70-70. Li Z W, Chen L Q. Memory management mechanism - segmentation technology and addressing scheme [J]. *Digital Technology and Application (in Chinese)*, 2016(1): 70-70.
- [22] 白福均, 高建瓴, 李宛蓉, 等. 一种倒排索引压缩方法[J]. *计算机应用研究*, 2019, 36(1): 106-109. Bai F J, Gao J L, Li W R, et al. Method of inverted index compression [J]. *Application Research of Computers (in Chinese)*, 2019, 36(1): 106-109.
- [23] 刘小珠, 彭智勇. 全文索引技术时空效率分析[J]. *软件学报*, 2009, 20(7): 1768-1784. Liu X Z, Peng Z Y. Time and space efficiencies analysis of full-text index techniques [J]. *Journal of Software (in Chinese)*, 2009, 20(7): 1768-1784.
- [24] 佟华, 胡江林, 张玉涛. GRAPES 模式后处理技术改进应用研究[J]. *气象科技*, 2020, 48(4): 511-517. Tong H, Hu J L, Zhang Y T. Improvement and application of post-processing technology in GRAPES Model [J]. *Meteorological Science and Technology (in Chinese)*, 2020, 48(4): 511-517.
- [25] NCEP Central Operations. NCEP WMO GRIB2 Documentation [EB/OL]. [2023-06]. https://www.nco.ncep.noaa.gov/pmb/docs/grib2/grib2_doc/
- [26] 杨勇. 基于 GlusterFS 的分布式冗余存储[J]. *西安文理学院*

- 学报(自然科学版), 2010, 13(4): 67-70. Yang Y. Methods of vector field analysis of special solution at some riccati differential equations [J]. Journal of Xi'an University (Natural Science Edition) (in Chinese), 2010, 13(4): 67-70.
- [27] 肖红玉, 贺辉, 陈红顺. 在线评测教学辅助系统设计[J]. 计算机技术与发展, 2017, 27(11): 141-145. Xiao H Y, He H, Chen H S. Design of online judge teaching assistant system [J]. Computer Technology and Development (in Chinese), 2017, 27(11): 141-145.
- [28] 岳昆, 王晓玲, 周傲英. Web 服务核心支撑技术: 研究综述 [J]. 软件学报, 2004, 15(3): 428-442. Yue K, Wang X L, Zhou A Y. Underlying techniques for Web services: a survey [J]. Journal of Software (in Chinese), 2004, 15(3): 428-442.
- [29] 曾超宇, 李金香. Redis 在高速缓存系统中的应用[J]. 微型机与应用, 2013, 32(12): 11-13. Zeng C Y, Li J X. Redis application in cache system [J]. Cyber Security and Data Governance (in Chinese), 2013, 32(12): 11-13.
- [30] 李玉帅, 李天义, 高炜, 等. 基于异步动态事件触发通信策略的综合能源系统分布式协同优化运行方法[J]. 自动化学报, 2020, 46(9): 1831-1843. Li Y S, Li T Y, G W, et al. Distributed collaborative optimization operation approach for integrated energy system based on asynchronous and dynamic event-triggering communication strategy [J]. Acta Automatica Sinica (in Chinese), 2020, 46(9): 1831-1843.
- [31] 钟建峰, 孟宇坤, 王石生, 等. 基于 gRPC 分布式数据通信的地铁线网指挥中心平台设计与实现[J]. 都市轨道交通, 2023, 36(2): 190-197. Zhong J F, Meng Y K, Wang S S, et al. Design and implementation of a metro network control center based on gRPC distributed data communications [J]. Urban Rapid Rail Transit (in Chinese), 2023, 36(2): 190-197.
- [32] 李兴宝, 侯方. 基于 C/S 模式的气象资料自动下载系统[J]. 气象科技, 2010, 38(5): 609-612. Li X B, Hou F. Automatic C/S-based meteorological data downloading system [J]. Meteorological Science and Technology (in Chinese), 2010, 38(5): 609-612.
- [33] 孙嫣. 基于 C/S-B/S 联合模式的气象计量管理系统[J]. 气象科技, 2008, 36(4): 502-505. Sun Y. Management system of meteorological metrology based on client/server and browser/server [J]. Meteorological Science and Technology (in Chinese), 2008, 36(4): 502-505.

Design and Practice of Long-Term Sequential Grid Data Management Platform

JIA Xiaozhen HU Jiangkai WANG Dapeng LIANG Chen

(CMA Earth System Modeling And Prediction Center, Beijing 100080)

Abstract: With the rapid development of numerical weather prediction services, the resolution and forecasting lead time of meteorological models have significantly improved, leading to an exponential growth in the volume of forecast data output. As a national meteorological model research and operational centre, CMA Earth System Modeling and Prediction Center (CEMC) currently produces daily gridded data outputs of 0.76 TB, with an annual output reaching 155.12 TB. Given the enormous data volumes, researchers' preferences for data access are evolving. Wagemann predicts that future scientific users increasingly prefer cloud platforms or other interfaces for data access rather than solely relying on downloads. To address these issues, this paper proposes a lightweight distributed parallel processing framework for gridded data management, aiming to streamline data management processes and enhance data access speed. The core design philosophy revolves around leveraging search engine technology for rapid metadata retrieval and gridded data decoding techniques for efficient data acquisition. To mitigate performance penalties from repetitive decoding, the framework decodes gridded data files once and supports multiple retrievals and extractions, significantly accelerating data access. Additionally, it supports cross-platform data access, facilitating easier data acquisition for researchers. The framework adopts a three-tier architecture: the data layer stores data, the algorithm layer implements core search and cataloguing algorithms, and the business layer interfaces directly with user needs. The framework implements crucial functions such as gridded data cataloguing, extraction, and clipping. During cataloguing, users invoke the cataloguing interface and input parameters (e.g., original data file paths, index names, index types), and the system automatically parses file metadata and generates indexes. For data extraction, users call the retrieval interface with specific parameters to obtain designated data. Moreover, the framework supports precise extraction of specified latitudinal and longitudinal data segments by configuring cropping parameters. It reduces decoding time by creating indexes based on binary storage characteristics, utilises an inverted index value-id model for rapid data location retrieval, enhances processing performance through GlusterFS shared storage and Celery distributed message queues, and ensures efficient and stable data transmission using gRPC technology for C/S communication. Practical tests and applications demonstrate the framework's exceptional performance in handling massive meteorological data. Notably, it successfully processes petabyte-scale gridded data during the Beijing Winter Olympics meteorological support services, significantly improving data access efficiency. Additionally, the framework supports flexible processing and scalable upgrades for various file formats to meet diverse user needs. By integrating advanced search engine technology, gridded data decoding methods, and a distributed cluster framework, the platform not only enables rapid data retrieval and efficient access but also satisfies researchers' urgent demand for cross-platform data access. As meteorological data continues to grow, this platform holds significant potential to play a pivotal role in various fields, offering more robust data support for weather forecasting, scientific research, and operational applications.

Keywords: distributed architecture; search engine; numerical prediction business; in-memory database; data retrieval; cross-platform

Received: 2023-12-12 Accepted: 2024-10-9

© Editorial Office of *Meteorological Science and Technology* (CC BY-NC-ND 4.0)