

题目编号：CQ-16

面向地球系统科学的海量网格类数据的 分布式文件系统设计比赛方案

一、发榜单位

国家气象信息中心

二、题目名称

面向地球系统科学的海量网格类数据的分布式文件系统设计

三、题目介绍

在地球系统科学领域，随着地面、雷达、卫星等观测技术快速发展，高性能计算能力持续增强，全球大气、海洋、陆地、冰冻、生物等多圈层的监测和预报的分辨率不断提升，网格类数据的空间分辨率达到公里级，时间分辨率达到小时级，单个文件达到 10GB，单数据中心在线存储量已超 100PB，且持续不断增长。

目前，对于地球系统科学领域的网格类数据，在存储技术上，一般采用分布式文件系统，基于文件系统内的元数据，定位到文件物理位置。在数据格式上，一般采用国际通用的 GRIB、HDF 等格式，这几类格式均为自描述格式，将多要素、多层次的数十个到上百个网格场压缩在一个文件中，而用户一般一次只使用其中的部分场，甚至一个场的一部分，因此，这几类数

据格式配套提供了相应的数据处理开发包，基于其自描述信息，实现并提供对局部数据块的网格场抽取、空间裁剪等方法，支持用户访问、解析和使用数据。

随着文件数量以及单个文件体量的快速增长，对地球系统科学领域网格类数据的使用效率也逐步下降，这已成为地球系统科学研究和应用领域“卡脖子”的瓶颈问题。因此，有必要开展将网格类数据格式的自描述信息融入到分布式文件系统的元数据体系，加速对文件内局部数据块的物理寻址，提升访问效率的研究。

本选题针对地球系统科学海量网格类数据的高效访问使用需求，选择 **GRIB** 或 **HDF** 一类格式，设计定制化的分布式文件系统。基于通用分布式文件系统，扩展设计其元数据体系，融入网格类数据格式的自描述信息，包括时空坐标、物理量纲、层级结构等，支持对网格内局部数据块的快速寻址定位；在分布式文件系统中，基于其扩展的元数据，设计网格类数据常用的读写函数库，代替数据格式自带的数据处理开发包，函数包括网格场写入、网格场读取、空间裁剪、空间插值等。

题目相关解释：

1. 关于通用分布式文件系统：参赛者可自行选择一类，推荐使用开源文件系统。

2. 关于数据格式说明及其自带的开发包：**GRIB**、**HDF** 均为公开通用的数据格式，参赛者可以在互联网上学习和了解。

参赛者自行下载试用 GRIB 或 HDF 自带的数据处理开发包，如有需要，可联系出题单位提供。

3. 关于测试数据文件：参赛者自行准备，文件数 ≥ 1 千，单文件数据量 30MB-10GB，如有需要，可联系出题单位提供数据示例。

4. 关于设计的试验环境：分布式文件系统的部署规模 ≥ 3 个节点。

5. 出题单位赛事支持网址：<https://data.cma.cn/jbgsjts>，网站发布内容包括上文提到的 GRIB 与 HDF 数据格式说明文档、GRIB 与 HDF 数据处理开发包、测试数据文件以及其他赛事支持相关材料。网址将在 5 月 30 日前开放，敬请关注。

四、参赛对象

本题目设学生赛道和青年科技人才赛道。

1. 学生赛道

参赛对象为 2025 年 6 月 1 日以前正式注册的全日制非成人教育的各类高等院校在校专科生、本科生、硕士研究生、博士研究生（不含在职研究生）。参赛人员年龄在 40 周岁以下，即 1985 年 6 月 1 日（含）以后出生。

同一作品不得同时参加第十九届“挑战杯”全国大学生课外学术科技作品竞赛（以下简称第十九届“挑战杯”竞赛）其他赛道的评比。

2. 青年科技人才赛道

参赛人员年龄在 40 周岁以下,即 1985 年 6 月 1 日(含)以后出生,在高等院校、科研院所、企业等各类创新主体中具有较高科研热情和较强科研能力的青年科技工作者。

高校青年教师在指导学生参赛的同时不得以参赛人员身份参加同一选题比赛。发榜单位及同发榜单位有相关隶属关系单位的青年不得参加本单位选题比赛。

各赛道参赛对象可以团队或个人形式参赛,每个团队不超过 10 人,每件作品可由不超过 3 名指导教师进行指导。可以跨专业、跨学校、跨单位、跨地域组队,但同一团队所有成员均应符合本赛道相关年龄、身份要求。每件作品只可由 1 所高等院校、科研院所或企业等作为参赛主体提交申报。

五、答题要求

提交作品为设计报告,主要体现对通用分布式文件系统的元数据体系的扩展设计,网格类数据常用读写函数库及其实现的设计。

六、作品评选标准

1. 文件系统设计的完整性与先进性(70 分)

优秀(56-70):文件系统的元数据体系扩展设计完整,覆盖数据格式全部的自描述信息,支持对数据块的快速寻址定位。已开发原型系统并对设计进行验证的作品另加 1-5 分,该项总分不超过 70 分。

良好（41-55）：文件系统的元数据体系扩展设计基本完整，覆盖数据格式主要的自描述信息，支持对数据块的寻址定位。

一般（0-40）：文件系统的元数据体系扩展设计不完整，覆盖数据格式少量的自描述信息，支持部分场景的数据块寻址定位。

2. 函数库设计的完整性与先进性（30分）

优秀（21-30）：函数库设计完整，包含网格场写入、网格场读取、空间裁剪、空间插值等4种，函数的逻辑和流程设计合理、高效。已开发原型软件并对设计进行验证的作品另加1-5分，该项总分不超过30分。

良好（11-20）：函数库设计基本完整，包含网格场写入、网格场读取、空间裁剪、空间插值等4种中的2-3种，函数的逻辑和流程设计基本合理。

一般（0-10）：函数库设计不完整，包含网格场写入、网格场读取、空间裁剪、空间插值等4种中的1-2种，函数的逻辑和流程设计不太合理。

七、作品提交时间

2025年5月-8月，各高校、企业、科研机构等组织协调机构组织学生和青年科技工作者参赛，安排专业人员给予指导，为参赛团队提供支持保障。

2025年8月15日前，各参赛团队通过大赛申报系统提交作品，具体要求详见作品提交方式。

2025 年 8 月底前，由大赛组委会会同发榜单位共同完成初审，确定入围终审擂台赛的晋级作品和团队。

2025 年 9 月，发榜单位安排专门团队提供帮助和指导，各晋级团队完善作品，冲刺攻关参加终审擂台赛，角逐“擂主”。

八、参赛报名及作品提交方式

（一）报名方式

1. 参赛选手登录“挑战杯”官网 2025.tiaozhanbei.net，在“揭榜挂帅”擂台赛报名入口注册账号，登录大赛申报系统在线填写报名信息。报名信息提交后，下载打印系统生成的报名表。

2. 申报人在报名表对应位置加盖所在学校或所在单位公章。

3. 将盖章版报名表扫描件上传至报名系统，等待系统审核。请参赛选手注意查看审核状态，如审核不通过，需重新提交。

4. 系统开放报名时间为 2025 年 5 月 30 日—6 月 30 日，逾期后系统将自动关闭报名功能。

（二）作品提交方式

同时提交申报作品（即设计报告）的 Word 版本和 PDF 版本文档，如有附件需一并提交。申报作品统一打包压缩提交至大赛申报系统，压缩包命名方式为：申报人所在单位-申报人姓名-作品名称-联系电话（例如：XX 大学-张 XX-XX 方案-手机号）。

九、赛事保障

5月-6月，面向参赛者组织集中的应用场景参观调研，统一提供选题讲解；按需提供公开的数据格式相关处理开发包、数据示例等。

5月-9月，配备专门的联系和指导人员，提供线上咨询和服务。

十、设奖情况及奖励措施

（一）设奖情况

本选题设置“擂主”1个，特等奖6个（含“擂主”），一等奖6个，二等奖8个，三等奖10个。其中，特等奖、一等奖、二等奖、三等奖分学生、青年科技人才两个赛道，独立评审、单独设奖；“擂主”在两个赛道的所有特等奖获得者中择优选出；最终奖项数量根据申报数量和质量的实际情况，按需调整。两个赛道的奖项设置如下：

1. 学生赛道

特等奖：3个；

一等奖：3个；

二等奖：4个；

三等奖：5个。

2. 青年科技人才赛道

特等奖：3个；

一等奖：3个；

二等奖：4 个；

三等奖：5 个。

2025 年“揭榜挂帅”擂台赛学生赛道获奖情况将按照一定分值计入第十九届“挑战杯”竞赛学校团体总分，具体分值以第十九届“挑战杯”竞赛章程为准。青年科技人才赛道获奖情况不纳入学校团体总分计分范围。

（二）奖励措施

1. 学生赛道奖励措施

（1）奖金额度：“擂主”（如果从本赛道产生）12 万元/个，特等奖（不含“擂主”）1 万元/个，一等奖 6 千元/个，二等奖 3 千元/个，三等奖 2 千元/个。

（2）成果转化：本单位择优与“擂主”（如果从本赛道产生）、特等奖获得团队合作，针对本选题的作品进行深入研究、成果转化和应用。

（3）实习机会：本单位对获奖团队成员提供不少于半年的实习锻炼机会。

（4）就业应聘：获奖团队成员在应届毕业参加本单位招聘时，同等条件下优先考虑。

2. 青年科技人才赛道奖励措施

（1）奖金额度：“擂主”（如果从本赛道产生）12 万元/个，特等奖（不含“擂主”）1 万元/个，一等奖 6 千元/个，二等奖 3 千元/个，三等奖 2 千元/个。

(2)成果转化：本单位择优与“擂主”(如果从本赛道产生)、特等奖获得团队合作，针对本选题的作品进行深入研究、成果转化和应用。

(3)交流机会：本单位对获奖团队成员提供不少于半年的实习锻炼机会。

(4)技术合作：中国气象局气象软件重点开放实验室为获奖团队提供进一步联合研究和技术攻关的合作机会。

(三) 奖金发放方式

比赛结束后1周内，本选题比赛专班工作人员与获奖团队联系，获奖团队填写奖金领取单，包括银行卡等信息。领取单反馈后1个月内，奖金一次性发放至获奖团队银行卡中。

十一、比赛专班联系方式

1. 专家指导团队

顾问专家：倪老师，联系电话：010-58994324（座机）、13581933632（手机）

顾问专家：傅老师，联系电话：010-58993098（座机）、18511489789（手机）

负责比赛期间技术指导保障。

2. 赛事服务团队

联络专员：刘老师，联系电话：010-58993284（座机）、13121787088（手机）

联络专员：冒老师，联系电话：010-68407846（座机）、

18161060207（手机）

负责比赛期间组织服务及后期相关赛务协调联络。

3. 联系时间

比赛期间工作日（9:00-17:00）。

附：发榜单位简介

国家气象信息中心（National Meteorological Information Centre）是中国气象局直属的司局级法人事业单位。中心承担着全球观测基础数据和气象产品的收集分发、气象数据加工处理与归档管理、气象数据产品研发与服务、高性能计算资源调度与并行计算技术支持、气象电子政务技术支持、信息系统基础设施资源管理与服务、信息网络安全防护及业务运行保障等任务职责。