

全国高校计算机教育研究会

全国大学生计算机系统能力大赛组委会

2024 年全国大学生计算机系统能力大赛

数据库管理系统设计赛技术方案

全国大学生计算机系统能力大赛数据库系统设计赛面向全国高校大学生，旨在培养数据库系统领域的设计、分析、优化与应用人才，激发学生的想象力、创新力、工程实践能力和团队协作精神，以赛促学、以赛促教，为高质量专业人才搭建交流、展示、合作的开源社区平台，让参与数据库系统设计赛的学生、指导教师、学校和企业都能从中获得补益，助力我国高校与企业在数据库系统领域的产学研合作健康、快速发展。

本技术方案旨在说明数据库系统设计赛目标、数据库系统设计赛内容、评测方式等技术方面的规定，帮助各参赛队更好地参赛。

一、数据库系统设计赛目标

数据库系统设计赛目标如下：

- 大赛要求各参赛队综合运用各种知识（包括但不限于数据库管理系统原理与实现、编译原理、数据结构与算法、操作系统等），具备将上述知识用于设计、实现和优化数据库管理系统基本的存取管理、查询处理、事务处理三大核心功能的能力。
- 大赛鼓励各参赛队伍在充分了解数据库系统的基本原理和基本实现技术、现代 C++ 编程技巧和特点等基础上，尽可能提高数据库系统的运行效率。
- 为展示参赛队的设计和实现水平，增加竞赛的对抗性，进入决赛的参赛队还需要针对业务负载的变化，现场增加、调整或优化相关算法，按照实际系统运行的结果进行排名。

二、初赛内容与评分标准

2.1 初赛内容

参赛队伍需要基于大赛提供的原型系统框架，实现赛题指定的功能，并进行一定的性能优化，主要包括以下内容：

- 基于大赛提供的数据库原型系统框架，围绕存取管理、查询处理、事务处理三大核心功

能，补充和完善相应的代码（编程语言为 C++）。

2. 参赛队提交的源代码要求能够在比赛指定的环境下（具体参见第五节“参赛平台与测试程序”）运行，并能够通过指定用例的测试。
3. 在比赛结束前，参赛队不得将提交的参赛内容（包括代码和文档）发布在比赛指定之外的平台上。

2.2 功能与性能测试

初赛包括功能与性能两类题目，提交与测试方式如下：

1. 大赛提供知识点考察的样例代码，参赛队伍可以进一步增添其他知识点考察代码、扩展接口，但项目工程文件（如 CMakeLists.txt 等）不得更改，具体规定随原型系统框架代码一并发布。
2. 大赛提供数据库原型系统框架，围绕存取管理、查询处理、事务处理三大核心功能，设置一定数量的知识点考察和相应的测试用例。参赛队伍需要在指定的比赛托管平台代码库中拉取数据库原型系统框架，并将补充完善的代码提交到托管平台上。
3. 大赛对参赛队伍提交的代码进行编译运行，并对每个知识点考察，使用相应的测试用例进行自动测试。
4. 功能测试：每个功能题目对应若干测试点，参赛队伍只有通过测试点才能拿到该测试点对应的分数。
5. 性能测试：参赛队伍完成可运行的数据库原型系统后，测试平台会在参赛队伍提交的系统上运行相应的测试负载，以参赛队伍在测试负载下的性能为评价依据。

2.3 初赛评分标准

初赛总成绩为 100 分，分为功能与性能两类题目，其中功能分数为占比 90%，性能分数为占比 10%。

功能题目共包含四个模块，每个模块分别设置一定数量知识点的考察，参赛队伍提交的源代码只有通过了大赛提供的相应测试用例，才能拿到该知识点考察的分数，各模块成绩权重大致如下：

- 查询编译与执行、元数据管理模块的实现：30%
- 事务管理模块的设计与实现：30%
- 存储模块的设计与实现：20%
- 索引模块的设计与实现：20%

性能题目为端到端测试，参赛队伍在通过功能测试的基础上才可以通过性能测试。所有参赛队伍中性能测试最佳的队伍得 100 分，其余参赛队伍的得分，以其性能相较于所有参赛队伍中的最佳性能，映射至[0, 100]区间；未通过正确性测试的小组得 0 分。

各参赛队伍总成绩计算方法如下：

- 总成绩 = 功能测试得分*90% + 性能测试得分*10%

三、决赛内容与评分标准

3.1 决赛内容

参赛队伍继续完善初赛阶段提交的数据库内核，并按要求完成作品。参赛队所开发的数据库系统内核应能够运行于数据库系统设计赛提供的硬件平台之上，并以通过的功能与性能测试用例个数作为成绩的评定标准。参赛队在规定的截止时间内修改数据库系统源代码，并把数据库系统源码公开提交到评测系统进行测试。生成的数据库系统在指定的目标硬件平台上运行，并进行基准测试程序的性能评测。所有参赛队伍中性能测试最佳的队伍得满分，其余参赛队伍的得分，以其性能相较于所有参赛队伍中的最佳性能，映射至[0, 满分]区间（具体的测算方法将会在决赛前发布）；未通过正确性测试的小组的0分。

参赛队伍还需进行现场答辩，展示优化思路、实现重点等，评审专家对参赛队伍的答辩情况进行打分。

3.2 决赛成绩

全国赛总成绩为 100 分，各阶段成绩权重如下所示。

- 初赛成绩占比：20%
- 性能成绩占比：50%
- 答辩表现占比：30%

四、参赛项目提交

4.1 提交内容

各参赛队伍在初赛和决赛阶段需要分别提交对应的完整设计内容。

1. 初赛：提交到代码库的数据库内核的完整工程文件。必须包含全部数据库内核源代码（有代码注释，中英文均可，鼓励良好的代码风格）
2. 进入决赛的队伍需要补充设计与实现文档。文档内容包括但不限于：设计思路、系统框架、实现重点、代码注释(中英文)、开发过程中遇到的主要问题和解决方法。
3. 决赛：提交到代码库的数据库内核的完整工程文件和设计文档。工程文件必须包含全部数据库内核源代码（有代码注释，中英文均可，鼓励良好的代码风格）。

4.2 关于开源的说明

参赛队需独立构造数据库系统，允许部分代码参考遵循开源（GPL 等）协议的数据库系统或软件模块，但不能够直接使用第三方库或开源代码（包括但不限于往届比赛参赛作品）来完成系统功能，一经发现代码抄袭或技术抄袭等学术不端行为，组委会将立即告知参赛队，

参赛队可于一周内进行申诉。如参赛队放弃申诉或者组委会认定其申诉无效，参赛队参赛资格将被取消。对于被认定有学术不端行为的参赛队，组委会将通报其所在学校，并取消和追回其已获得的证书和奖金。

五、参赛平台与测试程序

5.1 托管平台和基准测试程序

- 数据库系统设计赛托管平台：支持各参赛队的群体协作与版本控制。
- 大赛评测系统：参赛队伍向数据库系统设计赛托管平台提交完整工程文件，平台生成基准测试程序，自动进行功能及性能测试。
- 基准测试程序：用于对参赛队实现的 DB 内核进行功能与性能评测。

5.2 开发与测试环境

大赛开发与测试环境要求如下：

- 操作系统：Ubuntu18.04(x86_64)
- CMake：3.16 及以上版本
- 编译器：编译器需要支持 C++17 等新标准，建议使用 GCC 7.1 及以上版本
- 代码框架的第三方依赖库：flex; bison; readline

5.3 性能评测服务器配置

- CPU：配置 ≥ 2 颗英特尔至强银牌 4314 2.4G, 16C/32T, 10.4GT/s, 24M 缓存，DDR4-2666
- 内存：配置 ≥ 256 GB 3200Mhz RDIMM 内存
- 硬盘：配置 ≥ 3 块 4TB 7.2K SAS 12Gb 热插拔 512n 硬盘。3 块硬盘配置 RAID5。
- 独立智能阵列卡，配置 Raid 0/1/10/5 8GB 缓存带闪存保护功能，掉电 cache 数据永不丢失。

六、数据库系统设计赛网站

数据库系统设计赛网站为 <https://db.educg.net/>，数据库系统设计赛网站将提供多种软件开发工具及设计资料。