

2020 全国大学生计算机系统能力大赛

编译系统设计赛技术方案

详情访问大赛官网: <https://compiler.educg.net>

一、评价方式的基本说明

第1条 大赛要求各参赛队综合运用各种知识（包括但不限于编译技术、操作系统、计算机体系结构等），构思并实现一个综合性的编译系统，以展示面向特定目标平台的编译器构造与编译优化的能力。

第2条 大赛鼓励各参赛队充分了解目标语言及目标硬件平台（CPU 指令集、Cache、各类并行加速能力等）特点，使编译出的目标码能够尽可能利用目标硬件平台能力以提高目标码的运行效率。

第3条 为展示参赛队的设计水平，增加竞赛的对抗性，进入决赛的参赛队还需要针对目标语言或目标平台的变化，现场调整编程系统。

第4条 除本技术方案特别要求、规定和禁止事项外，各参赛队可自行决定编译器体系结构、前端与后端设计、代码优化等细节。

二、预赛评分标准

第5条 比赛内容。开发支持特定语言、面向 ARM 硬件平台的综合性编译系统。

1. 基于 C、C++语言开发，能够在 Ubuntu18.04(64 位)操作系统的 x86 评测服务器上编译。
2. 能够将符合自定义程序设计语言 SysY2020 的测试程序编译为 ARM 汇编语言程序（**32 位，ARMv7**），并通过汇编链接后在 Raspbian GNU/Linux 10 操作系统的 Raspberry 4B 设备上运行。

第6条 功能测试。参赛队所开发的编译器应能够编译大赛提供的 SysY2020 语言编写的基准测试程序。

1. 具有词法分析、语法分析、语义分析、目标代码生成与优化等能力，并支持编译错误的准确识别、定位与错误处理功能。
2. 对于正确编译通过的 SysY2020 基准测试程序，应生成符合要求的 ARM 汇编文件（**32 位，ARMv7**）。

功能测试通过将每个基准测试程序，基于大赛给出的 ARM 汇编器、链接器等工具，生成每个基准测试程序对应的二进制可执行文件，并在安

装有 Linux 操作系统的指定 ARM 硬件平台上加载并运行，根据给定的输入数据，比对输出结果，计算得分。每个基准测试程序未能正确编译出可执行二进制文件，或所有测试点都未通过计 0 分，所有测试点都通过计 100 分，部分测试点通过的，按所通过测试点的比例计算功能测试得分。参赛队的最终功能测试成绩为每个基准测试程序功能测试成绩的平均值。

第7条 性能测试。在通过功能测试的前提下，记录每个基准测试在目标硬件平台上的执行时间作为评价依据。每个基准测试程序按照执行时间最小者的性能测试分值被定义为 100 分，其余各参赛队依据在该基准测试程序上的运行时间与最小运行时间的比值除 100 计算参赛队在该基准测试程序上的性能得分（性能得分=100 / （运行时间/最短运行时间））。性能测试分值越大越好。参赛队的最终性能测试成绩为每个基准测试程序的性能成绩的平均值。

第8条 初赛总成绩 100 分，各分项成绩权重如下：

1. 编译器的功能测试成绩：50%
2. 基准测试程序的性能测试成绩：50%

三、决赛评分标准

第9条 决赛阶段的任务在预赛提交的最终版本编译系统上完成。

第10条 比赛内容。大赛组委会现场公布一项编译优化任务，包括但不限于对目标语言 SysY2020 或 ARM 架构目标执行平台的扩展或限制，并公布新的基准测试程序。参赛队根据变化，在限定时间内自行修改编译系统源代码，提交给竞赛系统。生成的编译系统需要以新给出的基准测试程序集作为输入，编译输出对应的 ARM 汇编程序，并在指定的目标硬件平台上完成性能评测。

第11条 功能测试与性能测试方法与初赛一致。参赛队团队协作及现场答辩的评分标准在决赛阶段另行公布。

第12条 决赛总成绩 100 分，各分项成绩权重如下：

1. 编译器的功能测试成绩：20%
2. 基准测试程序的性能测试成绩：70%
3. 参赛队团队协作及现场答辩：10%

四、参赛作品提交

第13条 各参赛选手预赛阶段需要在大赛的竞赛平台提交完整的设计内容：

1. 综合编译系统设计的完整工程文件（必须包含全部 C/C++源代码、`Make` 文件等），并在竞赛平台中至少有一次完整通过性能测试的记录和有效成绩。
2. 综合编译系统统设计文档。

第14条 如果需要使用第三方 IP 或者借鉴他人的部分源码，必须在设计文档和源代码的头部予以明确说明。

第15条 参赛队需从头构造编译系统，允许基于 `Lex`、`Yacc`、`Bison` 等通用词法、语法解析器生成工具帮助生成部分代码，但不得直接使用 `GCC`、`LLVM` 等现有、开源编译器及框架的源代码及裁剪。

第16条 参赛队必须严守学术诚信。一经发现代码抄袭或技术抄袭等学术不端行为，取消参赛队的参赛资格。

五、竞赛平台与测试程序

第17条 大赛提供的竞赛平台和测试程序包括：

1. 代码托管平台，支持各参赛队的群体协作与版本控制。
2. 竞赛评测系统，根据参赛队的申请从代码托管平台获取指定版本，生成编译系统，并加载基准测试程序，自动进行功能及性能测试。
3. 基于 `SysY2020` 语言的基准测试程序（包括 `SysY2020` 源码及评测点数据），用于在 `ARM` 平台上对参赛队编译器生成的可执行文件进行性能评测。

六、软硬件系统规范

第18条 `SysY2020` 语言是用于本次竞赛的高级程序设计语言，是 C 语言基本语法的子集，支持全局变量声明、自定义函数、`int` 类型和元素为 `int` 类型的多维数组类型，支持赋值、表达式、语句块、常见控制流结构，表达式支持基本算术运算、关系运算和逻辑运算，优先级与结合性与 C 语言规范保持一致。

第19条 复赛阶段的语言语法、目标硬件平台特性、基准测试集调整，由大赛组委会在决赛阶段发布。

第20条 大赛指定的编译环境用于编译参赛队提交的编译系统源码，参数如下：

1. CPU: Intel x86, 24 物理核心、开启超线程。
2. 内存: 128GB LPDDR4-3200 SDRAM。
3. Docker 容器操作系统: Ubuntu 18.04。
4. 编译器: LLVM/Clang-10.0.0, 编译选项:

C 语言: `clang --std=c11 -O2 -lm`

C++语言: `clang++ --std=c++17 -O2 -lm`

第21条 大赛指定的目标程序性能基准测试实验设备为 Raspberry Pi 4 Model B Rev 1.1, 主要参数如下:

1. CPU: Broadcom BCM2711, Quad core Cortex-A72 (ARM v8) 64 位 SoC, 主频 1.5GHz。
2. 内存: 2GB LPDDR4-3200 SDRAM。
3. GPU: 500Mhz VideoCore VI。
4. 操作系统: **32 位**, Raspbian GNU/Linux 10 (buster), 基于 Debian。
5. 汇编和链接器: gcc version 7.5.0 (Ubuntu/Linaro 7.5.0-3ubuntu1~18.04),
编译与链接命令: `gcc -march=armv7`

(用于汇编和链接参赛队编译系统输出的汇编代码)

七、大赛网站

第22条 大赛网站网址为: <http://compiler.educg.net>

第23条 大赛网站提供多种软件开发工具及设计资料, 包括但不限于下列内容:

1. SysY2020 程序设计语言规范、文法及说明。
2. Raspberry 4B 硬件技术规范。
3. 竞赛平台(代码托管平台、竞赛测试系统)使用文档。
4. 性能基准测试程序及其文档。
5. 参赛队远程调试及本地调试指南。