

中国海洋大学本科生研究发展计划

项目申报书

项目编号			
项目名称	Nim 语言编译器面向 LoongArch 迁移研究		
项目负责人	蔺春名	联系电话	17513728619
所在学院	信息科学与工程学部		
学号	23020007072	专业班级	23 计算机科学与技术 2 班
指导教师	曲海鹏		
E-mail	lcm@stu.ouc.edu.cn		
申请日期	2024/11/13		
项目期限	一年期		
中国海洋大学教务处			

填写说明

1. 本申请书所列各项内容均须实事求是，认真填写，表达明确严谨，简明扼要。
2. 申请人可以是个人，也可为创新团队，首页只填负责人。“项目编号”一栏不填。
3. 本申请书为大 16 开本（A4），左侧装订成册。可网上下载、自行复印或加页，但格式、内容、大小均须与原件一致。

一、基本情况

项目名称	Ni m语言编译器面向 LoongArch 迁移研究						
所属学科	学科一级门：工学 学科二级类： 计算机类						
项目来源	<input checked="" type="checkbox"/> A、学生自主选题，来源于自己对课题的长期积累与兴趣 <input type="checkbox"/> B、学生来源于教师科研项目选题 <input type="checkbox"/> C、学生承担社会、企业委托项目选题 <input type="checkbox"/> D、拔尖专项 <input type="checkbox"/> E、竞赛专项 <input type="checkbox"/> F、研修专项						
申请金额	500元	项目期限	一年期	拟申报项目级别		校级	
负责人	蔺春名	性别	男	民族	汉族	出生年月	2005年3月
学号	23020007072	联系电话	宅： 手机： 17513728619				
指导教师	曲海鹏	联系电话	宅： 手机： 15905320001				
项目简介		迁移 Ni m 语言编译器到国产 CPU龙芯					
负责人曾经参与科研的情况		无					
指导教师承担科研课题情况		主持和参与多项国家级课题，主持国家重点研发子课题，为百度、奇安信等公司提供技术服务					
指导教师对本项目的支持情况		提供软硬件开发环境和技术指导					

项目组 主要成员	姓名	学号	学院	专业班级	联系电话	项目分工
	蔺春名	23020007072	信息科学与工程学部	23 级计算机科学与技术 2 班	17513728619	项目计划，环境搭建与进度监督
	於佳杰	23020007151	信息科学与工程学部	23 级五八班	18385638183	项目推进，分析错误模块
	李佳潼	22020007043	信息科学与工程学部	22 级计算机科学与技术 1 班	15564196227	资料查阅与材料收集
	赵禹惟	23020036096	信息科学与工程学部	23 级计算机科学与技术(中外合作办学) 2 班	13578774880	材料撰写，文案整理
	徐启睿	23020007133	信息科学与工程学部	23 级数据科学与大数据技术班	19100410509	测试与 bug 修复
指导教师	姓名	工号	学院/单位	职称	联系电话	电子邮件
	曲海鹏	2006018	信息科学与工程学部	副教授	15905320001	quhai peng@uc.edu.cn

二、立项依据（可加页）

（1）研究目的

Nim语言(nim-lang.org)作为静态类型的编译型系统级语言，因其简洁优雅而富含现代化特性的语法而拥有可超越 Python 的开发效率，语法简洁而灵活，于此同时，却拥有着接近 C 的运行效率，充分发挥了零成本抽象(zero-overhead)的特色。它整合了 Pascal，Python，Ada，Modula 等众多语言的优秀特性，也给许多现代化语言的设计带来灵感，如 Rust、Cangjie(仓颉)。

Nim代码在可以作为 **NimScript** 解释执行的同时，以编译执行作为主要运行方式，且在相当程度上做到了一套代码，多后端执行，即可编译为 C、C++、JavaScript 代码，拥有极高的跨语言性、语言互操作性和灵活的 FFI，极大提高了开发的便捷度。

不同于 C 的手动内存管理，也不同与 C++/Rust 的 RAI，**Nim**使用 ORC 作为内存管理，

即 ARC+小型循环检测器，保证了内存安全的同时，也通过 sink 参数、copy on write 等技术极大程度避免了对性能的影响。同时，也摆脱了 Rust 风格的强制所有权检查带来的各种不便，极大提高了开发效率与开发负担。

Loongarch(龙芯) CPU 作为国产自主研发 CPU，吸收近年来指令集发展的先进技术成果，具有完全自主、技术先进、兼容生态三方面特点，拥有代码效率高，支持高效二进制翻译等优点，不仅在硬件方面更易于高性能低功耗设计，而且在软件方面更易于编译优化和操作系统、虚拟机的开发。同时，在 ARM x86 都受到美国对华政策制约的大背景下，Loongarch(龙芯)作为国产自研 CPU，对突破对华技术封锁、增强综合国力、保障国家安全具有重要意义。

完成 Nim 到 Loongarch(龙芯) 的迁移，不仅是对龙芯生态的一次促进，更为龙芯开发提供了一个更安全，高效，便捷的选项，有助于推动软件国产化进程，同时在全球呼吁使用内存安全语言开发软件的大背景下，为软件安全化进程提供了一个比 Python 运行高效进百倍、比 Rust 更易于上手，开发与维护的方案。

Nim 作为开源语言，开源于 [github\(github.com/nim-lang/Nim\)](https://github.com/nim-lang/Nim)，众多维护者中不乏国人。拥有丰富健全的标准库、第三方库，且使用 nimble(github.com/nim-lang/nimble) 作为官方包管理器，极大方便了第三方库的安装、使用与维护。由于其语法优雅而表现力强，库源码具有简洁、易于阅读、易于维护的特点，对于相关领域的研究具有较高的参考学习价值。

希望能通过完成 Nim 到 Loongarch(龙芯) 的迁移：

- 完善国产龙芯 CPU 生态
- 为软件国产化提供更优方案进行探索
- 完成 Nim 语言编译器在国产龙芯 CPU 下的迁移，提高 Nim 跨平台性
- 推进 Nim 在软件开发中的应用

(2) 研究内容

- 完成 Nim 语言编译器伪交叉编译（获得龙芯 CPU 下可运行的 Nim 编译器）
- 凭借上述编译器完成龙芯本机下的 Nim 语言编译器编译
- 通过 Nim 语言集成测试，根据错误原因更改相关源码

(3) 国、内外研究现状和发展动态

已有的龙芯相关 Pull Request 链接:

- [Add support for LoongArch](#)
- 对生成编译脚本的 compiler/installer.ini 的更新 [Add linux/loongarch64 support in 'compiler/installer.ini'](#)
- 相关文档修改: [doc\(format\): system nim doc of host CPU for loongarch64](#)

总的来说,已经有对龙芯支持的尝试,但由于疏忽等原因实际未能完成整个流程(没有完整提交到源代码仓库),并且由于对龙芯不是一级支持(first class support),所以相关的测试失败也尚未被关注。

(4) 创新点与项目特色

- 利用 Loongarch(龙芯)与 MIPS 的高相似度、Nim 的基于 C 语言转译的编译原理,使用 Nim 源码面向 MIPS 架构的生成的 C 源码进行本机编译,获得龙芯 CPU 下可运行的 Nim 编译器
- 利用 Nim 语言(nim-lang.org) 自举(编译器由 Nim 语言实现而非 C 语言)的优势,避免多次交叉编译。
- 为龙芯架构提供更现代化、特性丰富的编译系统
- 为多个代码仓库提交龙芯相关贡献,助力国产处理器架构得到更广泛支持与认同

(5) 技术路线、拟解决的问题及预期成果

- 技术路线

配置 cpu 为 mips64le 编译 Nim 编译器源码为 C 代码,将 C 代码在 Loongarch64(龙芯)本机编译获得 Nim 编译器,在龙芯架构下使用 Nim 编译器编译 Nim 编译器源码获得 cpu 信息为 loongarch64 的 Nim 编译器,运行集成测试,根据测试报错部分修改相应源码直至通过,最终重新编译获得通过测试的 Nim 编译器

- 拟解决问题
 - N m语言的 csour ce_v2 无 loongarch(龙芯) 的 C 代码(即无法直接通过编译得到龙芯 N m 编译器)
 - 龙芯下 N m 的集成测试存在失败部分

预期成果

1. 编译获得龙芯本机下的 N m 编译器
2. 为 ni m-l ang/ csour ce_v2 仓库提交代码贡献, 以支持 N m 的内建编译器构建工具 koch 在龙芯下直接本机编译获得 N m 编译器
3. 在龙芯架构下通过所有 N m 语言的集成测试
4. 为官方源代码仓库 gi t hub. com/ ni m-l ang/ ni m 提交相关 Pul l Request , 以使官方支持龙芯架构
5. 为龙芯架构操作系统的包管理器的官方包索引配置 N m 作为可安装包 (参考: [ht t ps: // gi t hub. com/ ni m-l ang/ csour ces_v1/ pul l / 4#i ssuecomment -1868855583](https://gi t hub. com/ ni m-l ang/ csour ces_v1/ pul l / 4#i ssuecomment -1868855583))

(6) 项目研究进度安排

阶段一：准备与环境配置(2024.11 – 2024.12)

- 目标：

完成项目环境的搭建，包括 C 语言编译环境与 N m 语言编译环境。

确认龙芯环境下的基本工具链支持。
- 任务：
 1. 配置龙芯 Loongarch64 环境，并安装基本开发工具链（GCC 或 Clang 等）。
 2. 获取并准备 N m 的源码（包括 N m 主仓库和 csour ces_v2 仓库）。
 3. 阅读 N m 编译器源码，特别是与架构相关的部分，了解现有 ni ps64l e 的支持情况。
 4. 初步测试从 ni ps64l e 配置编译的 N m 生成的 C 代码在 Loongarch64 下的运行情况，确保能够执行基本命令（ni m —ver si on 等）。

阶段二：伪交叉编译与初步移植(2025.01– 2025.02)

- **目标：**

实现 N m 编译器的伪交叉编译，生成 Loongarch64 版本的 N m 编译器。

- **任务**

1. 配置 N m 编译器进行伪交叉编译，使用 mips64le 作为初始配置，编译为 C 代码。
2. 在 Loongarch64 本地编译生成的 C 代码，获得 N m 编译器。
3. 针对初步移植过程中出现的编译错误，记录并分析错误原因，进行简单修正。
4. 验证生成的编译器是否能在 Loongarch64 本地执行简单的 N m 命令。

阶段三：本地编译与深入源码修改(2025.02– 2025.04)

- **目标：**

使用 Loongarch64 环境下生成的 N m 编译器编译 N m 源码，并解决出现的错误。

- **任务：**

1. 使用 Loongarch64 环境下的 N m 编译器重新编译 N m 源码，生成本地架构版本的 N m 编译器。
2. 详细分析编译过程中涉及 Loongarch 架构的报错，定位并修改相关 C 代码与 N m 源码。
 3. 提交初步的源码修改，记录每次修改后的测试结果。
4. 与 N m 社区保持沟通，了解其他架构移植过程中可能遇到的类似问题。

阶段四：集成测试与优化(2025.04– 2025.09)

- **目标：**

通过 Nim 语言的所有集成测试，并进一步优化源码以兼容 Loongarch64。

- **任务：**

1. 使用 Loongarch64 版本的 Nim 编译器运行 Nim 的集成测试（`nim testament all`）。
2. 收集并分析测试失败的用例，确定问题是否与 Loongarch 架构相关。
3. 根据测试结果，修改对应的源码部分（可能涉及 Nim 编译器、C 代码和 Nim 标准库）。
4. 在 Nim 社区中发布进度更新，并请求社区成员协助测试和提供反馈。

阶段五：代码贡献与 Pull Request 提交(2025.9– 2025.10)

- **目标：**

为 Nim 官方仓库提交代码贡献，并推动 Nim 对 Loongarch 架构的官方支持。

- **任务：**

1. 整理所有修改过的源码文件和补丁，准备 Pull Request 提交给 Nim 的官方仓库。
2. 在 GitHub 上提交 Pull Request，详细说明每个修改的原因和测试结果。
3. 与 Nim 的维护者沟通，回答可能的代码审查问题，并根据反馈进行进一步调整。
4. 在龙芯架构操作系统的包管理器中配置 Nim 语言包，并测试其安装和运行效果。
5. 撰写项目总结报告，记录从交叉编译到本地编译、源码修改和测试的全过程。

(7) 已有基础

1. 与本项目有关的研究积累和已取得的成绩

1. 研究积累

考察了 Nim 仓库(github.com/nim-lang/Nim)已有的关于 Loongarch 支持的

Pull Request:

<https://github.com/nim-lang/nim/pulls?q=is%3Apr+loongarch>

2. 已取得的成绩

初步测试了 cpu 配置为 mips64le 生成的 C 代码可在 Loongarch64（龙芯）本机编译，获得的 Nim 编译器可执行 `nim --version` 等简单指令（详情：

<https://github.com/nim-lang/Nim/issues/24118>）

2. 已具备的条件，尚缺少的条件及解决方法

已具备条件：Nim 源码(github.com/nim-lang/Nim)，Nim csource 源码(github.com/nim-lang/csources_v2)

尚缺少条件及解决方法：

- a. loongarch(龙芯) 架构的主机较少。解决方法：使用 qumu 虚拟化，ssh，git 等方式构建多人协助环境
- b. Nim 中文资料，Nim 对 C 的 FFI 等相关的中文资料。解决方法：在中文论坛等交流群中搜寻自发翻译的资料，阅读源文档(英文) 等。

三、经费预算

开支科目	预算经费（元）	主要用途	阶段下达经费计划（元）	
			前半阶段	后半阶段
预算经费总额				

1. 业务费	500		200	3000
(1) 计算、分析、测试费	500	租用云服务器	200	300
(2) 能源动力费				
(3) 会议、差旅费				
(4) 文献检索费				
(5) 论文出版费				
2. 仪器设备购置费				
3. 实验装置试制费				
4. 材料费				
学校拨款				
财政拨款				

四、项目组成员签名

蔺春名 徐启睿 赵禹惟 李钰潼 於建杰

五、指导教师意见

导师（签章）：
年 月 日

年 月 日 盖章:

年 月 日 盖章: