

RISC-V 高性能开源计算库 OpenMPL

张先轶

PerfXLab 澎峰 (北京) 科技有限公司

xianyi@perfxlab.com

2023.8

关于澎峰科技

- 2016年, 澎峰科技 (PerfXLab)成立, 推出智能边缘计算全栈解决方案, 服务于国产芯片和行业边缘智能
- 2018年,加入RISC-V基金会,再次发力RISC-V基础软件能力和生态建设
- 公司一直致力于研发算力基础软件(高性能计算库、异构计算软件栈以及软硬融合解决方案等),为国内外知名企业和院所提供了加速计算解决方案(中科院,华为,燧原,平头哥=,华大九天,中船等)
 - 2016年,中国计算机学会科技讲步二等奖
 - 2017年,中国科学院杰出科技成就奖
 - 2018年,北京雏鹰人才计划,国家高新企业
 - 2021年,数字中国·集成电路赛道特等奖
 - 2021年, 创芯中国·决赛一等奖
 - 2021年, CRVA联盟, 软件工作组副组长单位
 - 2022年, OpenCAX SIG10组长单位
 - 2022年11月15日,ChinaSC中国超级算力大会荣获"算力软件基建领军企业"和"中国智能计算卓越贡献奖"双项荣誉
 - ・ 2023年,入选北京市"专精特新"中小企业
 - 2023年,入选中国互联网协会算网云协同系统工作委员会成员单位







公司核心技术与产品

• 公司为算力行业构建了从数学计算库、异构计算框架、直到新的科学计算工具全栈解决方案。

③ PerfXPy面向科学家的高性能Python计算平台

PerfXAPI 异构计算,并行优化, AI and HPC

PerfMPL 高性能数学库和各类算子库

for 高性能异构计算平台

x86/ARM/RISC-V CPU, GPU, NPU/DSA, FPGA

赋能科研新范式。建设软硬件 融合算力的全栈国产可控生态, 支持源生创新。

赋能跨架构的计算平台。为异构计算编程平台提供系统解决方案。实现算法跨平台敏捷迁移,实现加速计算和自主可控。

赋能新一代计算芯片。为各类 处理器架构提供高性能的、稳定 可靠的计算库软件。

RISC-V 软硬件融合类产品

澎峰科技与中科计算技术西部研究院共同研发!

衡山系列(高密度RISC-V计算型服务器)

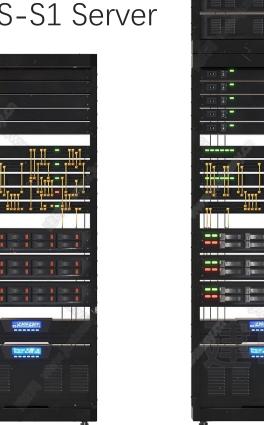
- 高密度计算型服务器
- 处理器: 国产RISC-V4X<u>64核@2.0GHz</u>
- 支持RDIMM/NV-DIMM 3200MHz的DDR4 ECC
- PCle4.0 x8 *3 (可扩展100G通信)
- 1.5U 标准机箱,支持GPU/NPU异构
- 软件集成: Linux + PerfMPL + PerfXAPI + PerfXPy

1024_{cores} 16T_{flops} 4 X HS-S1 Server

 Stronger performance

2048 cores 32T_{flops}

8 X HS-S1 Server





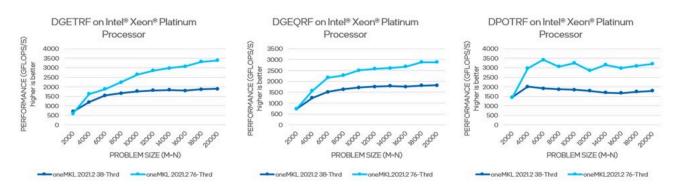
成熟架构计算领域软件生态



- Intel公司
 - MKL库 (数学库、几千个函数, 闭源)
 - 稠密矩阵计算BLAS/LAPACK
 - 快速傅里叶变换FFT
 - 稀疏矩阵计算 SparseBLAS
 - 稀疏解法器 (直接、迭代)
 - 随机数发生器
 - 向量数学库VML
 - 统计加速...
 - DAAL库(数据分析)
 - IPP库(图像、信号处理)
 - MKL-DNN库 (深度学习)
 - 持续增长、持续优化

oneMKL Linear Algebra Package (LAPACK) Scales Best at 76 Threads

Intel® oneAPI Math Kernel Library (oneMKL) 2021.2 LAPACK on Intel® Xeon® Platinum Processor



Testing Date: Performance results are based on testing by Intel as of April 26, 2021 and may not reflect all publicly available security updates.

Configuration: Frode, 2x Intel® Xex.PF Platinum 8368 processor on Coyote Pass platform with 256 GB (36 slots/16GB/3200) total DDR4 memory, ucode 0xd000270, HT off, Turbo off, Red Hait 8.3, 418.0-240 el8x86, 64, brite® PSD960GB OS Drives, the State of t

Performance results are based on testing as of dates shown in configurations and may not reflect all publicly available updates. See configuration disclosure for details. No product or component can be absolutely secure

Performance varies by use, configuration, and other factors. Learn more at www.Intel.com/PerformanceIndex. Your costs and results may vary

开源数学计算库:分散



常用稠密线性代数方程组求解器

名称	简介	主导方	类型	备注
OpenBLAS	高性能BLAS和LAPACK库	中国, 澎峰科技	开源	澎峰科技提供商业 版本PerfMPL
BLIS	BLAS和扩展接口库	美国, UT Austin	开源	
ATLAS	BLAS库	美国	开源	
LAPACK	稠密线性和特征值解法 器	美国	开源	

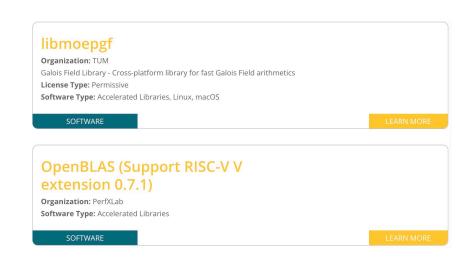
常用稀疏特征值求解器

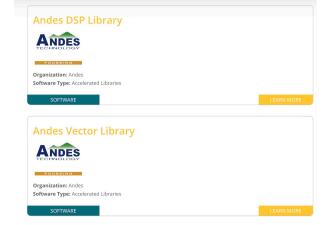
名称	简介	主导方	类型	备注
ARPack	稀疏特征值求解器	美国	开源	已停滞
ARPack-NG	基于ARPack开发	法国	开源	
Spectra	基于C++模版的特征值 求解	中国,上海财经大学	开源	基于Eigen

RISC-V计算领域现状 (2023)



- RISC-V Vector扩展
 - RVV 0.7.1 量产 (阿里平头哥 C906/C910, 全志D1)
 - RVV 1.0 发布
- 多核/众核
 - 单核
 - 2核、4核
- 软件生态支持
 - 编译工具链对Vector支持
- 加速计算库优化
 - 新增 Libmoepgf
 - 可惜没有优化

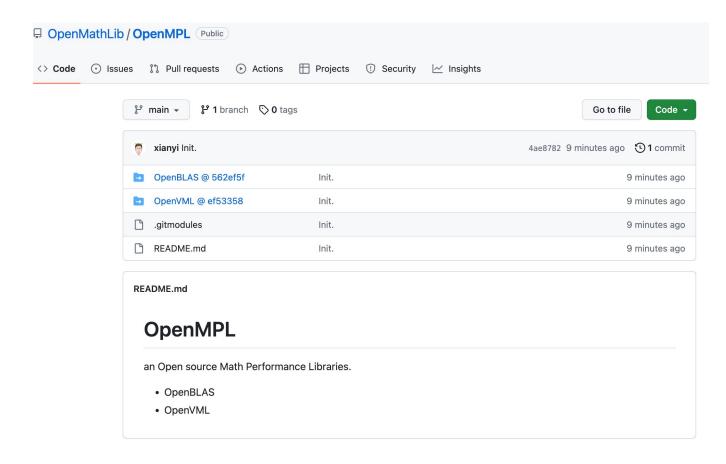




OpenMPL开源数学计算库

PerfXLab

- 统—API和库
 - 参考Intel MKL
- 持续针对各个架构处理器性能优化
 - X86
 - ARM
 - RISC-V



OpenBLAS



- 开源矩阵计算库,12年历史
 - https://github.com/xianyi/openblas
- 矩阵和向量基本计算
 - 当前版本: 0.3.21
- 支持主流处理器
 - X86/ARM/Power/龙芯
- 支持多种操作系统
 - Linux/Windows/Mac OSX/Android
- 目前支持RISC-V架构
 - 已经支持C910 RVV指令优化 (RVV 0.7.1, 1.0支持)
 - RVV 1.0进行中,已merge sifive x280 PR























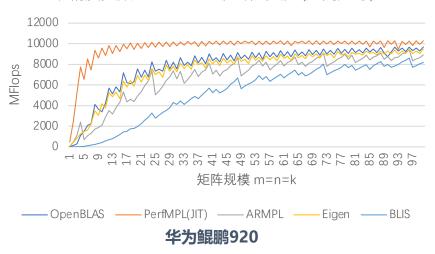




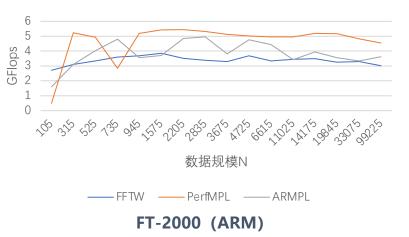
OpenMPL on ARM

优势: 1) 性能对标国际; 2) 一致精度; 3) 安全、可靠、稳定。

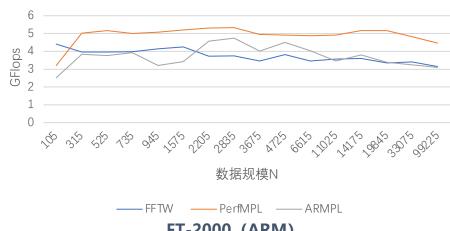
双精度实数 GEMM NN单线程性能 (越高越好)



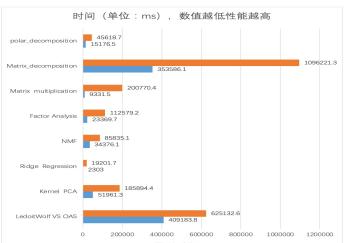
1D FFT C2C InPlace单线程性能(越高越好)







FT-2000 (ARM)

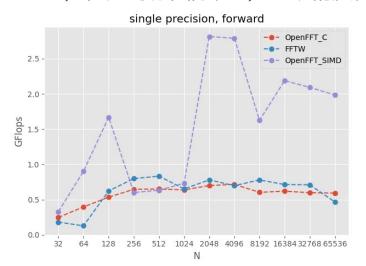


FT-2000/4

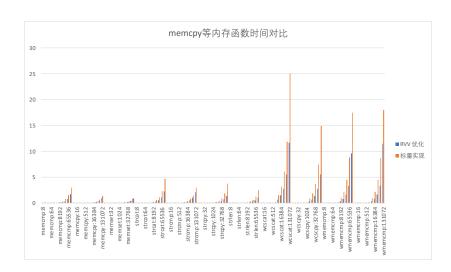
OpenMPL应用在 PerfXPy科学计算软件上, 与原生Python比较,带 来性能大幅提升。

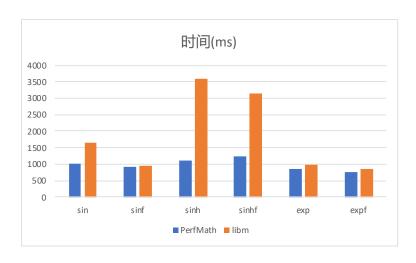
OpenMPL on RISC-V

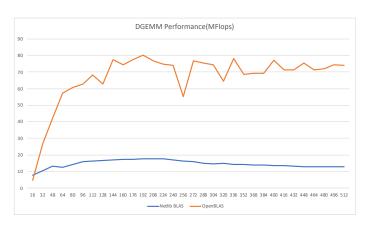
优势: 1) 性能对标国际; 2) 一致精度; 3) 安全、可靠、稳定



RISC-V C920







总结



- OpenMPL
 - 统一的开源数学计算库,降低迁移成本
 - OpenBLAS
 - OpenVML
- 未来工作
 - 开源FFT模块: fftw, openfft
 - 其他的稀疏解法器库



感谢各位

期待与各位交流,为RISC-V计算生态发展贡献力量

欢迎全职/兼职/实习生





