开源生态与计算机系统教学

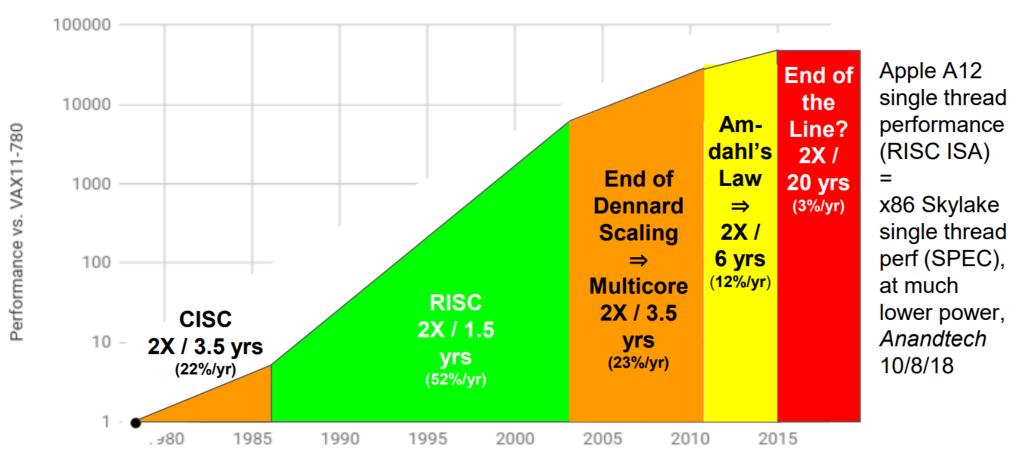
唐丹 中科院计算所 2019.10.17



计算机系统方向人才 从来没有如此紧缺!

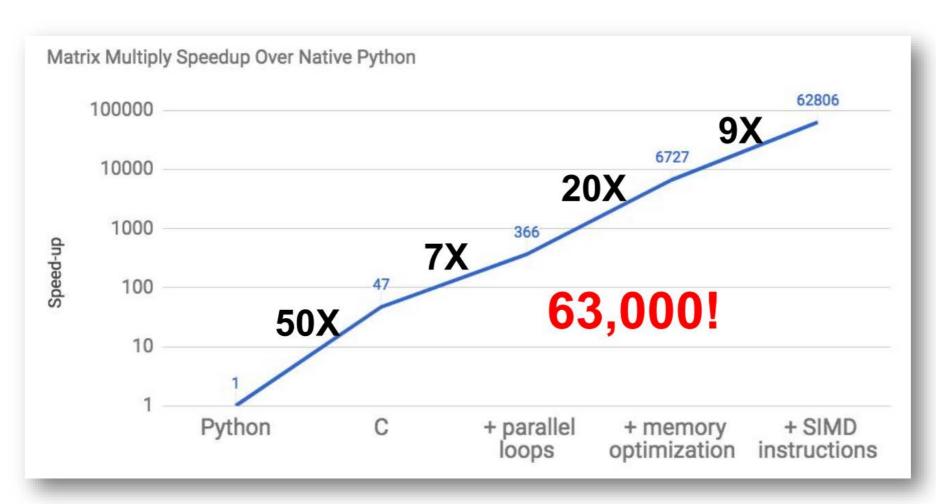
① 摩尔定律逐步放缓

40 years of Processor Performance



David Patterson, A New Golden Age for Computer Architecture: History, Challenges, and Opportunities, 2018

领域专用体系结构兴起



- 如何弥补软硬件 性能鸿沟?
- 硬件加速器
 - DomainSpecificArchitecture

David Patterson, A New Golden Age for Computer Architecture: History, Challenges, and Opportunities, 2018

AloT市场面临碎片化新挑战

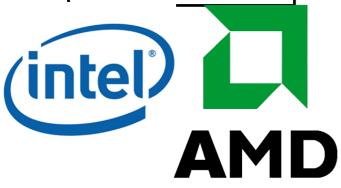
Intel/Linux 与Wintel生态 只有2-3种主流 处理器芯片方 案



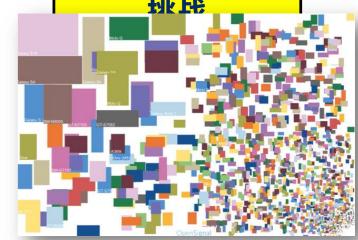
ARM/Andr iod生态有几十 种主流处理器 芯片方案



AloT市场将会出现成于上万种定制芯片需求,面临碎片化业战







②贸易摩擦:处理器芯片卡脖子问题凸显

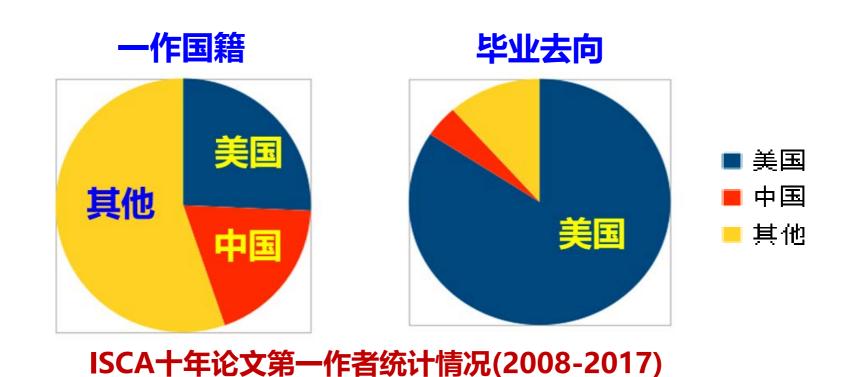
- 美国对华为、曙光等中国企业断供,企业出现以下燃眉之急:
 - (1) 工艺和生产渠道要迁移
 - (2) 原来借助的美国人才无法继续使用
- <u>处理器与基础软件生态被美国控制是</u> 最大的心腹之患

虽然企业被卡的是处理器技术,但原因也在于计算机系统方向人才严重不足!



现状: 优秀人才储备严重不足

• 85% vs. 4%: 2008-2017十年间体系结构国际顶级会议ISCA论文的第一作者85%选择在美国就业,仅有4%在中国就业,差距巨大

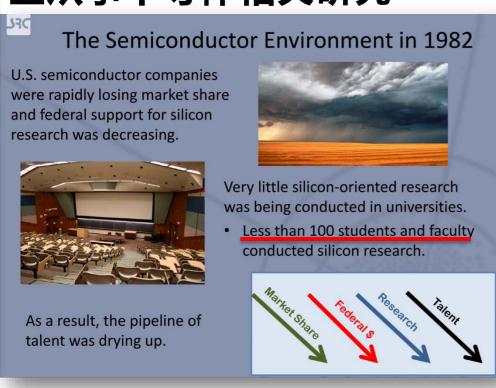


降低芯片设计门槛 是人才危机一种破解之道

国外的经验

1970年代末至1980年代中期,美国也曾遭遇人才储备不足问题

1982年,全美上干所大学中只有不到100位教授和学生从事半导体相关研究





降低芯片设计门槛 让学生也能流片

1981年,美国启动MOSIS项目,为大学提供流片服务。提出MPW模式,大幅降低芯片设计门槛。三十余年来为大学和研究机构流了60000多款芯片,培养了数万名学生

降低芯片设计门槛,推动产业变革

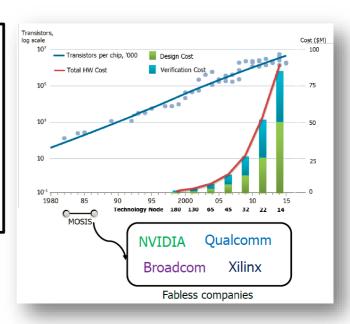
1981年, DARPA启动
 MOSIS项目, 提出了
 MPW (Multiple Project Wafer)模式,数量级降低芯片设计成本



三十余年来为大学和研究机构流了60000多款芯片,培养了数万名学生



- 催生半导体产业 新商业模式
 - 无晶圆企业
 - 代工企业





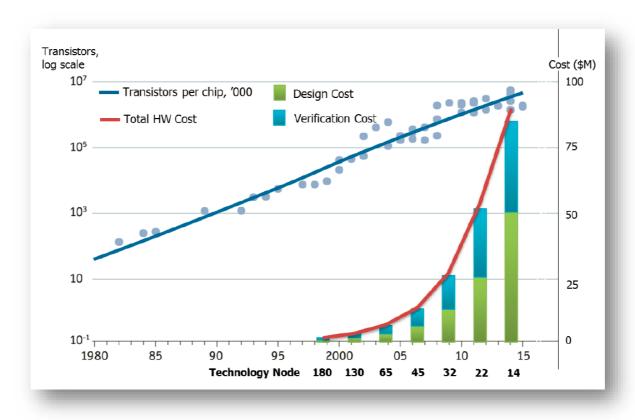
Source: Andreas Olofsson, Intelligent Design of Electronic Assets (IDEA),2017

芯片设计门槛极高

• 时间长: 英伟达Xavier SoC设计用了8000人年

• 成本高:终端芯片14nm工艺为例,上亿元研发经费

 只有少数企业能承受中高端芯片研发成本,大学无能力开展芯片研究,制约芯片 领域的创新,计算机系统/芯片人才培养能力严重不足



Source: Andreas Olofsson, Intelligent Design of Electronic Assets (IDEA),2017

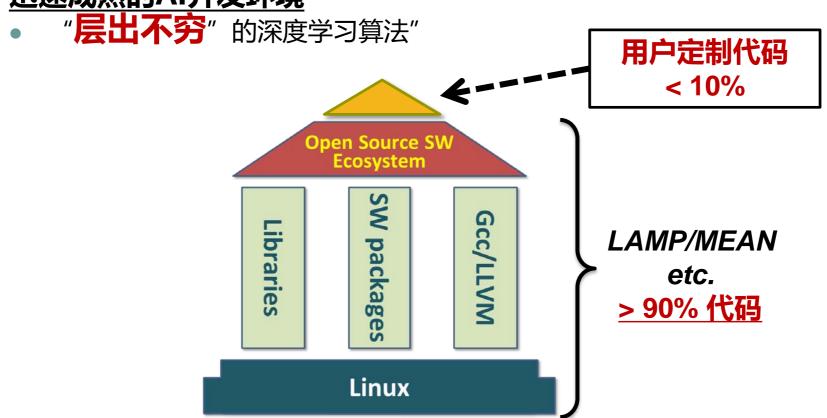
能否降低芯片设计门槛,实现......

- 让学生不再害怕做芯片
- 让本科生可以带着自己设计的芯片毕业
- 让3-5人的团队可以创办芯片创业公司
- 让做芯片像写APP那么简单
- 让天下没有难做的芯片

•

开源软件的成功经验

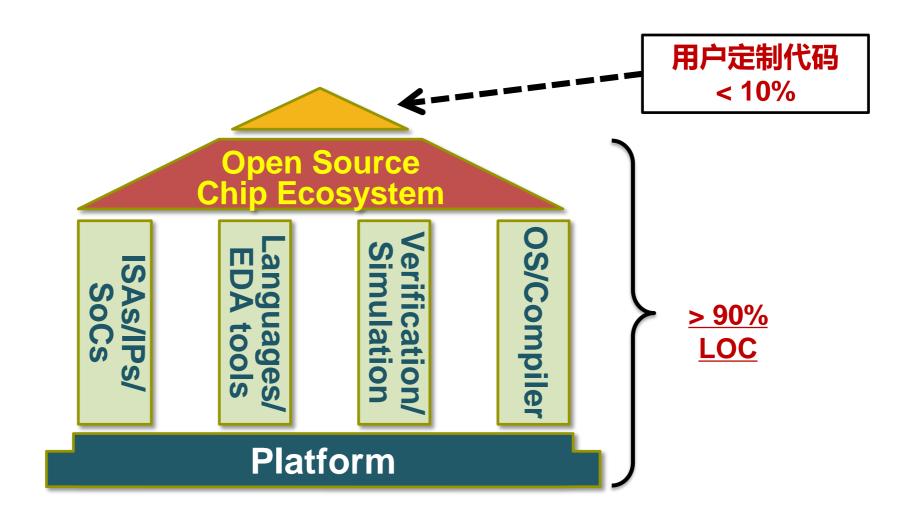
- 降低互联网创新的门槛
 - **3-5位**开发人员用**几个月**即可快速开发创新业务,滴滴、摩拜等
- 提高互联网企业自主能力
 - 互联网公司敢于 "去IOE"
- 迅速成熟的AI开发环境



降低系统芯片开发/教学门槛的 四个要素

开源芯片生态四要素

• 开源芯片+敏捷开发:降低芯片设计的人力、EDA、IP成本



① RISC-V: 降低IP成本的新模式

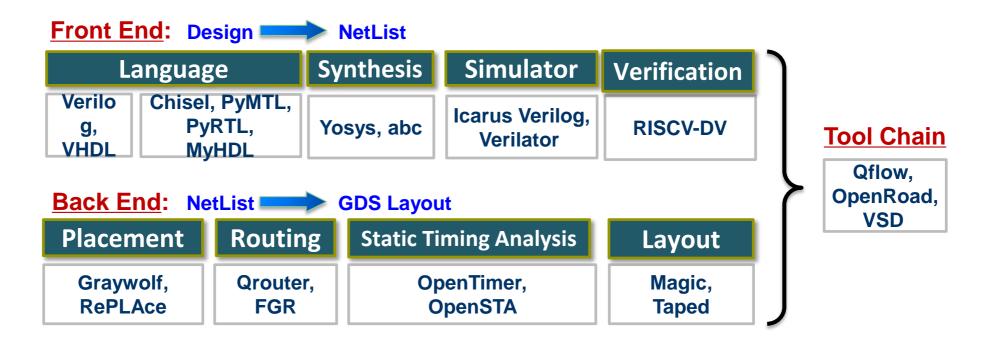


Field	Standard	Free, Open Impl.	Proprietary Impl.
Networking	Ethernet, TCP/IP	Many	Many
os	Posix	Linux, FreeBSD	M/S Windows
Compilers	С	gcc, LLVM	Intel icc, ARMcc
Databases	SQL	MySQL, PostgresSQL	Oracle 12C, M/S DB2
Graphics	OpenGL	Mesa3D	M/S DirectX
ISA	??????		x86, ARM, IBM360

开放指令集:无需担心ISA知识产权风险,教学/开源无缝衔接

②硬件开发语言与开源EDA工具链

- 语言: 更高抽象的硬件开发新语言可提高开发效率数倍
 - Chisel, PyMTL, ...
- EDA工具链:已存在一系列开源 EDA组件

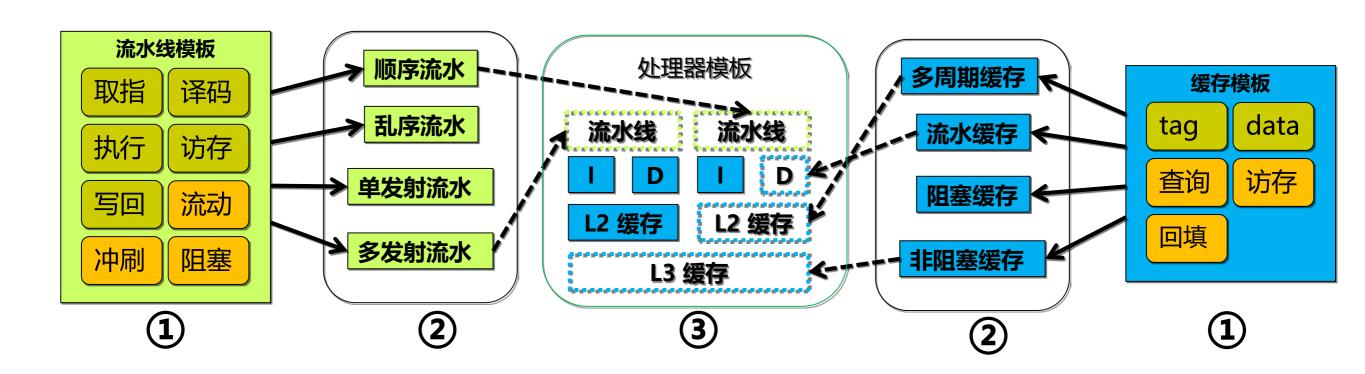


不需要昂贵的商用工具工具链,进一步降低芯片开发/教学门槛

面向对象的处理器结构设计

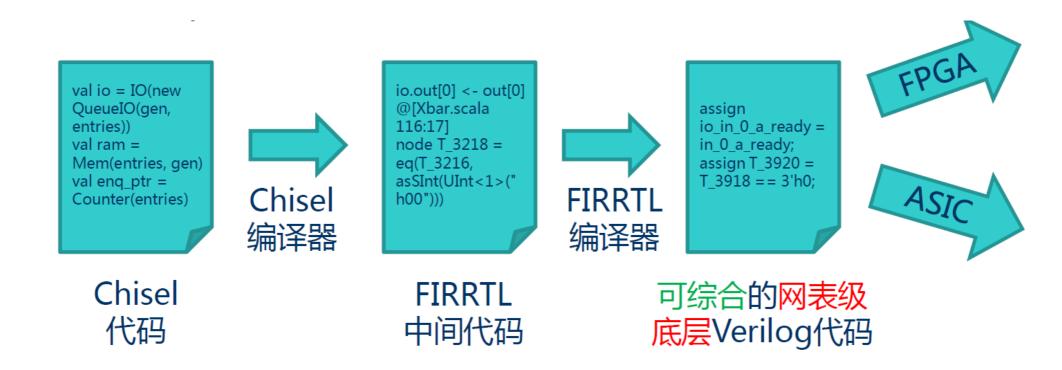
利用面向对象三个基本特征,实现处理器设计的可分解、可扩展、可组合

- ① 通过封装对硬件对象及其操作进行分解
- ② 通过继承从封装好的模板扩展出不同设计
- ③ 通过多态实现不同设计的组合切换



芯片敏捷开发

- 效率更高的硬件开发语言Chisel
 - Constructing Hardware in an Scala Embedded Language



开发效率对比

• 相同任务: 快速实现L2 Cache, 接入RISC-V

	一位工程师	一位本科生	
项目经验	熟悉OpenSparc T1, 修改过Xilinx Cache	做过CPU课程设计, 有9个月Chisel开发经验	
开发模式	传统开发	敏捷开发	
开发语言	Verilog	Chisel	
是否复用已有代码/测试环境	否, 独立开发/构建 测试环境(花费约3周)	是,使用Chisel库和Labeled RISC-V项目的测试环境	
周期	6周	3天	
有效代码/行	~1700	~350	
效果	目前仍无法启动Linux	可启动多核Linux, 支持DMA模式的以太网	

• 开发的效率高!

代码量约为仅为1/5

开发质量对比

- 让另一名Chisel零基础的国科大本科生翻译工程师的核心模块并评估
 - Vivado 2017.01, FPGA 型号 xc7v2000tfhg1716-1

	Verilog	Chisel (直接翻译)	Chisel-opt (高级特性与库)
Frequency/MHz	135.814	136.388 (+0.42%)	154.107 (+13.47%)
Power/W	0.770	0.749 (-2.73%)	0.749 (-2.73%)
LUT Logic	5676	6422 (+13.14%)	2594 (-54.30%)
LUT Storage	1796	1264 (-29.62%)	1492 (-16.93%)
FF	4266	3638 (-14.72%)	747 (-82.49%)
LOCS	618	470 (-23.95%)	155 (-74.92%)

• 敏捷开发方法可达到传统开发质量

《计算机研究与发展》60周年特刊

• 2019年1月发表

计算机研究与发展

Journal of Computer Research and Development

DOI:10.7544/issn1000-1239.2019.20180771

56(1): - , 2019

芯片敏捷开发实践:标签化 RISC-V

余子濠1,2 刘志刚1,2 李一苇1,2 黄博文1 王 卅1,2 孙凝晖1,2 包云岗1,2

1(计算机体系结构国家重点实验室(中国科学院计算技术研究所) 北京 100190)

2(中国科学院大学 北京 100049)

(yuzihao@ict. ac. cn)

Practice of Chip Agile Development: Labeled RISC-V

Yu Zihao^{1,2}, Liu Zhigang^{1,2}, Li Yiwei^{1,2}, Huang Bowen¹, Wang Sa^{1,2}, Sun Ninghui^{1,2}, and Bao Yungang^{1,2}

下载地址

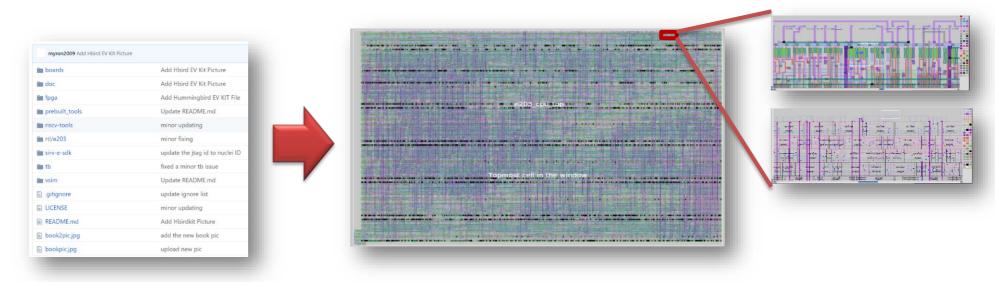


¹(State Key Laboratory of Computer Architecture (Institute of Computing Technology, Chinese Academy of Sciences), Beijing 100190)

² (University of Chinese Academy of Sciences, Beijing 100049)

开源EDA工具链

- 开源EDA工具链可实现从Verilog到GDS版图的整个流程
- 但应需要稳定、快速、高质量、维护好的开源EDA工具链发行版

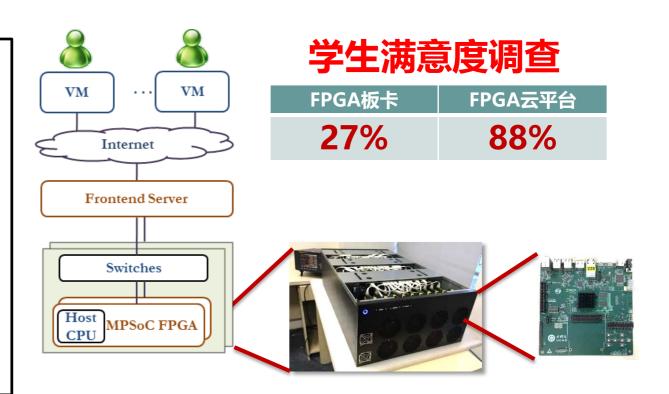


RTL of an open-source RISC-V core (E203)

GDSII w/ SIMC 180nm

③ 仿真/模拟

- 基于云的仿真与模拟: "即付即用"云模式可降低购买昂贵设备、服务器的成本
 - Berkeley's FireSim, 阿里平头哥, ...
 - **国科大教学实践**:设计与实现 FPGA云平台
 - 支撑2018/2019年超过**200位国科** 大本科生的计算机组成原理实验
 - 超过35%的学生由"不喜欢"转变为"喜欢"芯片与系统方向

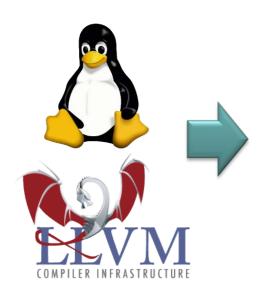


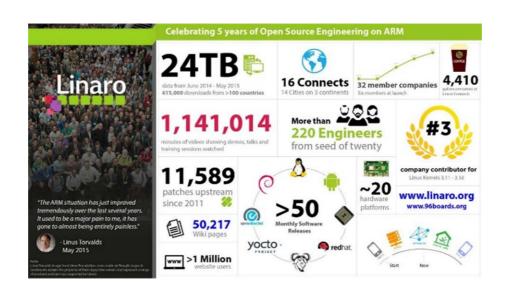
发表于国际计算机教学领域旗舰会议ACM SIGCSE 2019,国科大首篇、大陆第四篇长文

④ 软件工具链: OS, Compiler, ...

• 软件支持的三个层次

- 面向芯片 (Linux, LLVM): extended ISA, sensor drivers etc.
- 功能软件 (Linaro): languages, libraries, JVM, DBMS etc.
- 面向用户/程序员 (Android): APIs, UI etc.

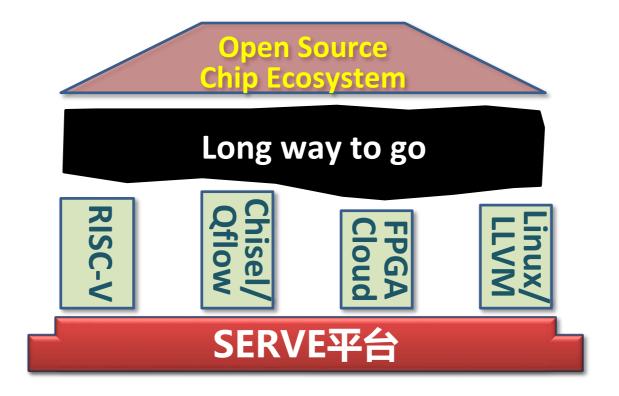






基于开源生态的计算机系统教学平台

集成开源芯片四要素的教学平台



思沃(SERVE): RISC-V系统级原型验证服务平台

System Emulation and Prototyping for RISC-V Environment



SERVE.r

基于Xilinx PYNQ-Z2低成本 精简普及版

SERVE.i

基于FIDUS Sidewinder-100 高性能增强版

SERVE.s

多节点集群版 (精简/高性能集群)

SERVE.c

云服务版 (精简/高性能云)

SERVE.v

标签化冯·诺伊曼 体系结构版本 (LvNA)

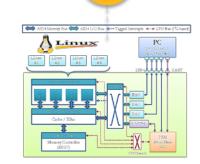












思沃(SERVE): RISC-V系统级原型验证服务平台

System Emulation and Prototyping for RISC-V Environment



SERVE.r

基于Xilinx PYNQ-Z2低成本 精简普及版

SERVE.i

基于FIDUS Sidewinder-100 高性能增强版

SERVE.s

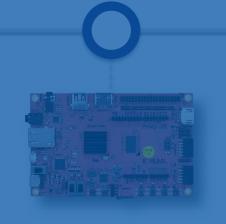
多节点集群版 (精简/高性能集群)

SERVE.c

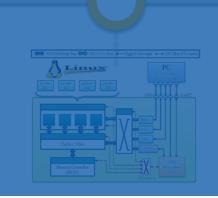
云服务版 (精简/高性能云)

SERVE.v

标签化冯·诺伊曼 体系结构版本 (LvNA)



首个面向RISC-V生态的 FPGA系列化系统级原型服务平台





SERVE.r: 低成本RISC-V全系统原型验证平台



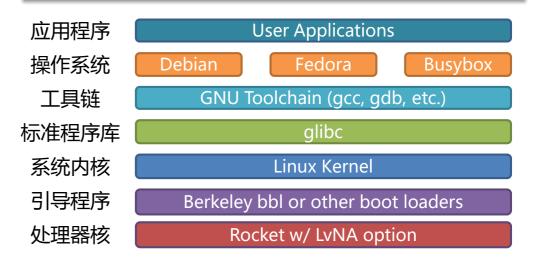
```
The programs included with the Debian GNU/Linux system are free software; the exact distribution terms for each program are described in the individual files in /usr/share/doc/*/copyright.

Debian GNU/Linux comes with ABSOLUTELY NO WARRANTY, to the extent permitted by applicable law. root@rv-hfu-01:~# uname -a Linux rv-hfu-01 4.19.0-00056-g3735af2 #1 SMP Wed Jul 24 22:08:39 CST 2019 riscv64 GNU/Linux
```

干兆以太网通路测试

平台特性配置

- 低门槛、低成本、低功耗、快速自启动
- 可配置Rocket单/双核@50-100MHz
- **外设硬核控制器**(串口/以太网/SD/多模USB等)
- 可扩展FPGA加速逻辑IP与HDMI等接口
- Linux v4.19内核+Debian OS社区生态
- 快速开展操作系统及应用层**软件开发**



开源开放的全系统软硬件栈环境



SERVE.r: 低成本RISC-V全系统原型验证平台

文件系统



The programs included with the Debian GNU/Linux system are free software; the exact distribution terms for each program are described in the individual files in /usr/share/doc/*/copyright. 挂载Debian

Debian GNU/Linux comes with ABSOLUTELY NO WARRANTY, to the extent permitted by applicable law.

root@rv-hfu-01:~# uname -a

inux rv-hfu-01 4.19.0-00056-g3735af2 #1 SMP Wed Jul 24 22:08:39 CST 2019 riscv64 GNU/Linux

oot@rv-hfu-01:-# ip addr show eth0
: eth0: <BROADCAST,MULTICAST,UP,LOWER_UP> mtu 1500 qdisc pfifo_fast sta
 link/ether a6:b2:91:23:3b:e5 brd ff:ff:ff:ff:ff
 inet 192.168.100.10/24 scope global eth0
 valid ff from 192.168.100.20: icmp_seq=1 ttl=128 time=0.672 ms

 inet 192.168.100.20: icmp_seq=2 ttl=128 time=0.672 ms

千兆以太网通路测试

平台特性配置

- 低门槛、低成本、低功耗、快速自启动
- 可配置Rocket单/双核@50-100MHz
- **外设硬核控制器**(串口/以太网/SD/多模USB等)
- 可扩展FPGA加速逻辑IP与HDMI等接口
- Linux v4.19内核+Debian OS社区生态
- 快速开展操作系统及应用层软件开发

应用程序 User Applications
操作系统 Debian Fedora Busybox
工具链 GNU Toolchain (gcc, gdb, etc.)
标准程序库 glibc
系统内核 Linux Kernel

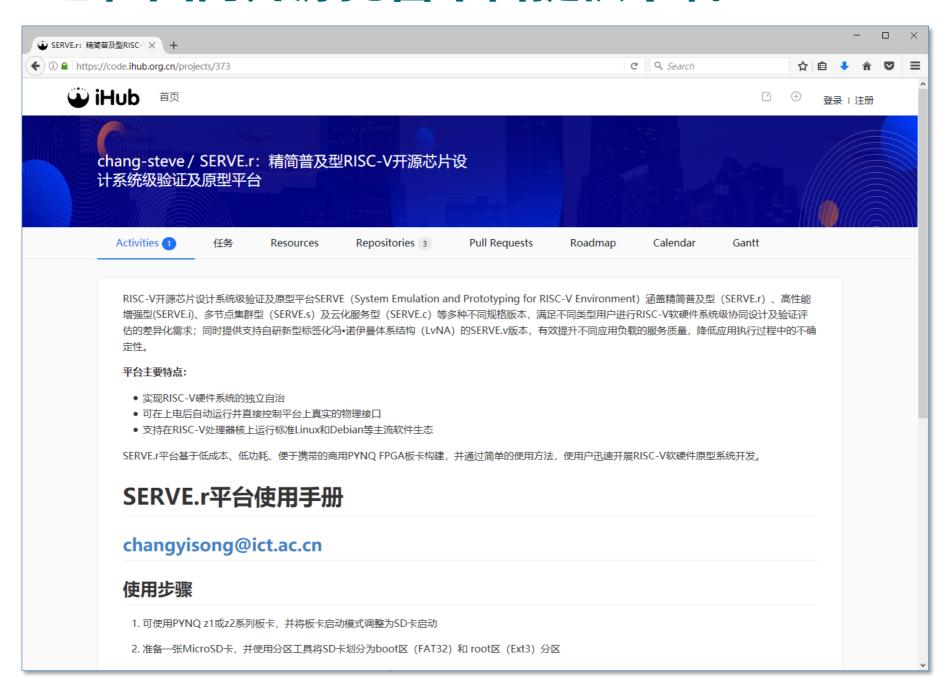
致捷验证提供可能

Rocket w/ LvNA option

开源开放的全系统软硬件栈环境



SERVE.r已在国内开源托管平台提供下载





SERVE.i: 高性能RISC-V全系统平台



PCle 扫描 axi-pcie: No bus range found for /amba_plee/axi-pcie: No bus range found for /amba_plee/axi-pcie: MEM 0x70000000..0x/fffffff -> 0x/0 axi-pcie: PCI host bridge to bus 0000:00 bus resource [bus 00-ff] bus resource [mem 0x70000000-0x73fffffff] 0: assigned [mem 0x70000000-0x740ffffff] 0: assigned [mem 0x74000000-0x7400fffff] 0: assigned [mem 0x74000000-0x74003fff 64bit] bridge to [bus 01] ridge window [mem 0x74000000-0x740fffff] ion 0000:01:00.0

Debian 登录

cpuinfo 输出

: θ : rv64imac

arch : sifive,rocket0
coot@label-riscv [10:15:27 ~] \$



SERVE.s: 集群版本RISC-V全系统平台



- 集成多个通过标准干兆以太网络 互连的SERVE.r或SERVE.s节点
- 便于多人同时同地开展RISC-V 系统级验证
- 基于Debian生态,可构建并行 编程环境

Ke Zhang, Yisong Chang, Mingyu Chen, Yungang Bao, Zhiwei Xu. **ZyCube: An In-House Mini-Cluster for Agilely Developing and Conducting Computer System Course Projects** (Poster), in the Proceedings of ACM Global Computing Education Conference 2019 *(CompEd2019)*, May 2019, Chengdu, China



SERVE.c: 云平台版本RISC-V全系统平台

基于网络 7 x 24小时 不间断服务

Scale-Out 验证评估 提供更多FPGA 逻辑资源及 异构加速能力 提供更加真实的 数据中心与云计算 应用负载评估环境





32路基于 Zynq FPGA的 RISC-V 全系统节点

动态弹性分配硬件板卡资源

高密度 低功耗 支持更大规模 的并发验证 和远程调试

RV-Prototypeas-a-Service

Ke Zhang, Yisong Chang, MingyuChen, YungangBao, and ZhiweiXu. 2019. **Computer Organization and Design Course with FPGA Cloud.** In Proceedings of the 50th ACM Technical Symposium on Computer Science Education *(SIGCSE2019)*. 927-933.



SERVE.v: 标签化冯诺依曼体系结构LvNA

Labeled RISC-V FPGA原型系统配置

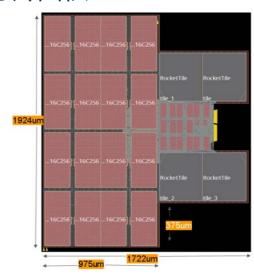
- · rocket核心 * 4
- 16KB L1 I\$, 16KB L1 D\$, 2MB L2\$
- 千兆以太网
- Linux v4.18内核
- 可运行Redis、Memcached等应用
- · 支持标签化特色的性能调控技术



基于 FIDUS 板卡部署

GF 14nm ASIC流片评估结果

- Slow Corner, 0.72V, 125C
- 核心频率1.25GHz
- L2 Cache频率625MHz
- 芯片面积1.9mm x 1.7mm



硬件应用案例: 我们的开源微结构实现改进计划



均在 SERVE.i 或 SERVE.v 上开 展

▶ Bug修复

- FIRRTL编译性能bug
- RAS性能bug
- BOOM的LSQ功能bug

- 性能改进

- 神经网络分支预测器
- DRRIP末级缓存替换算法
- Non-blocking的流水化末级缓存
- L1/LLC预取
- Load-to-use延迟优化

- 创新尝试

- 资源低开销的OoO调度器设计
- 自动化末级缓存容量划分
- 多bank的寄存器堆
- LSQ内存模型

- 测试验证和设计方法

- 基于Lock-step的差分测试
- BOOM的断点支持
- 自动流水化设计范式
- •



软件应用案例:帮助RedHat移植Fedora

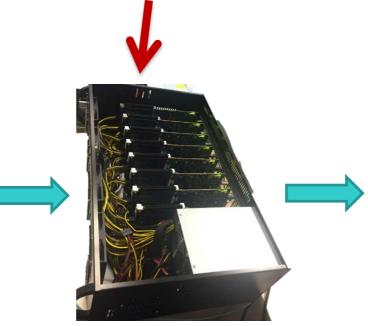


- RedHat工程师远程登录使用火苗进行Fedora的移植

火苗 = SERVE.c + SERVE.v







| Started Permit User Sessions. | Starting Terminate Plymouth Boot Screen... | Starting Hold until boot process finishes up... | Started Terminate Plymouth Boot Screen. | Started Hold until boot process finishes up. Welcome to the Fedora/RISC-V disk image https://fedoraproject.org/wiki/Architectures/RISC-V Build date: Wed Jul 3 20:19:49 UTC 2019 Kernel 4.18.0-ga57318a4-dirty on an riscv64 (hvc0) The root password is ..riscv... To install new packages use 'dnf install ...' To upgrade disk image use 'dnf upgrade --best' If DNS isn..t working, try editing ../etc/yum.repos.d/fedora-riscv.repo... For updates and latest information read: https://fedorapeople.org/groups/risc-v/disk-images/readme.txt Fedora/RISC-V http://fedora-riscv.tranquillity.se/koji/ http://fedora-riscv.tranquillity.se:3000/
Distribution rep.: http://fedora-riscv.tranquillity.se/repos-dist/
Koji internal rep.: http://fedora-riscv.tranquillity.se/repos/
[230.410000] tx_irq = 4
[230.410000] rx_irq = 5 fedora-riscv login: root [root@fedora-riscv ~]# passwd

在火苗上成功登录Fedora

全自主开发的教学SoC芯片年底流片

- ▶ RV64G指令集,单核,5级静态流水,32K/32K L1 Cache
- ► Chisel语言开发,面向对象的处理器结构设计
- 流水线行为抽象,利于学生理解
- 指令随机验证平台
- 教学所需外设 (SDRAM/SPI Flash/MAC/SDHC/GPIO)
- ► SMIC 0.11um工艺,商用EDA/开源EDA双后端设计流程
- ▶ 启动Linux,运行完整Debian/Fedora
- ▶ 开源所有RTL源代码、验证平台、物理设计脚本

希望与大家一起推动 计算机系统方向人才培养

謝謝! 请批评指正