

NutShell-本科生设计的可运行Linux的 RISC-V芯片

中国科学院大学

王华强 张紫飞 金越 张林隽 王凯帆

2020.7.18

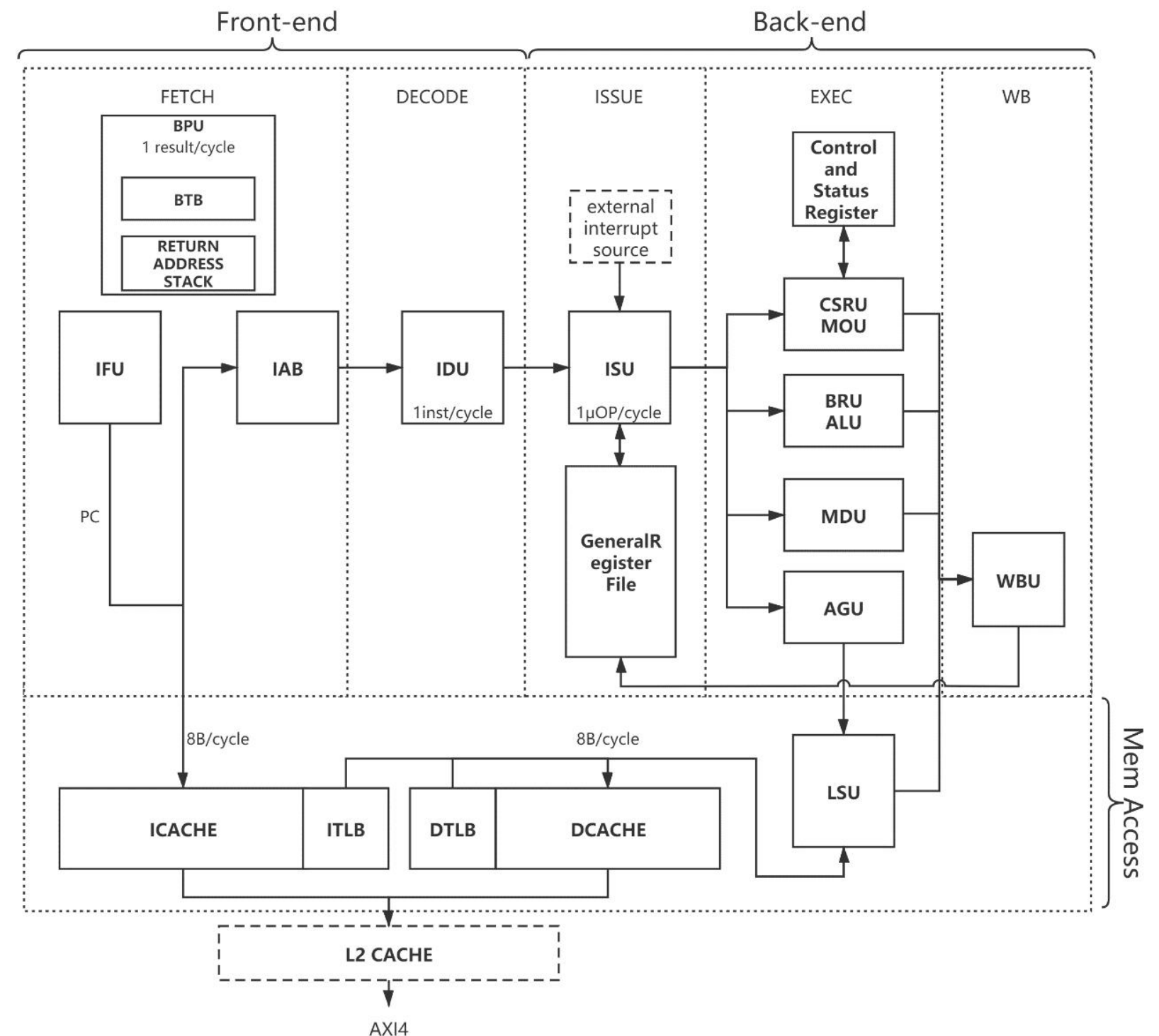
目录

- NutShell简介
- 经验分享
 - 使用Chisel进行硬件开发
 - 差分测试框架

NutShell简介

顺序RISC-V核: NutCore

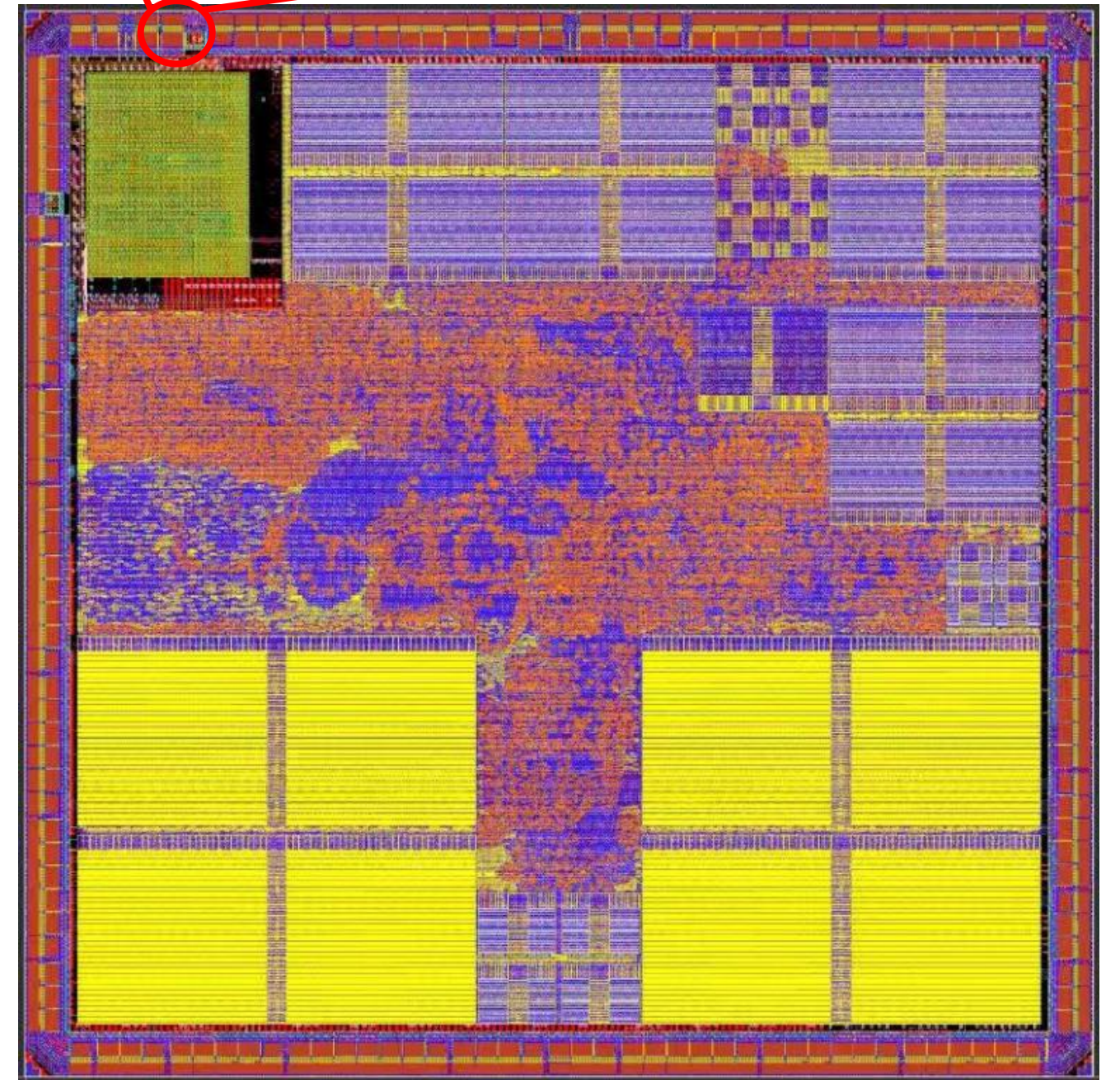
- **单发射顺序核**
- RV64IMAC指令集，支持M/S/U特权级
- 两位饱和计数器分支预测
- 512项BTB，16项RAS
- **支持Sv39分页机制, 支持硬件TLB填充**
- 32K指令/数据L1 cache
- 支持L1 指令/数据cache读一致性
- 128K L2 cache，支持next line预取
- **使用Chisel语言开发**



SoC芯片: NutShell (果壳)

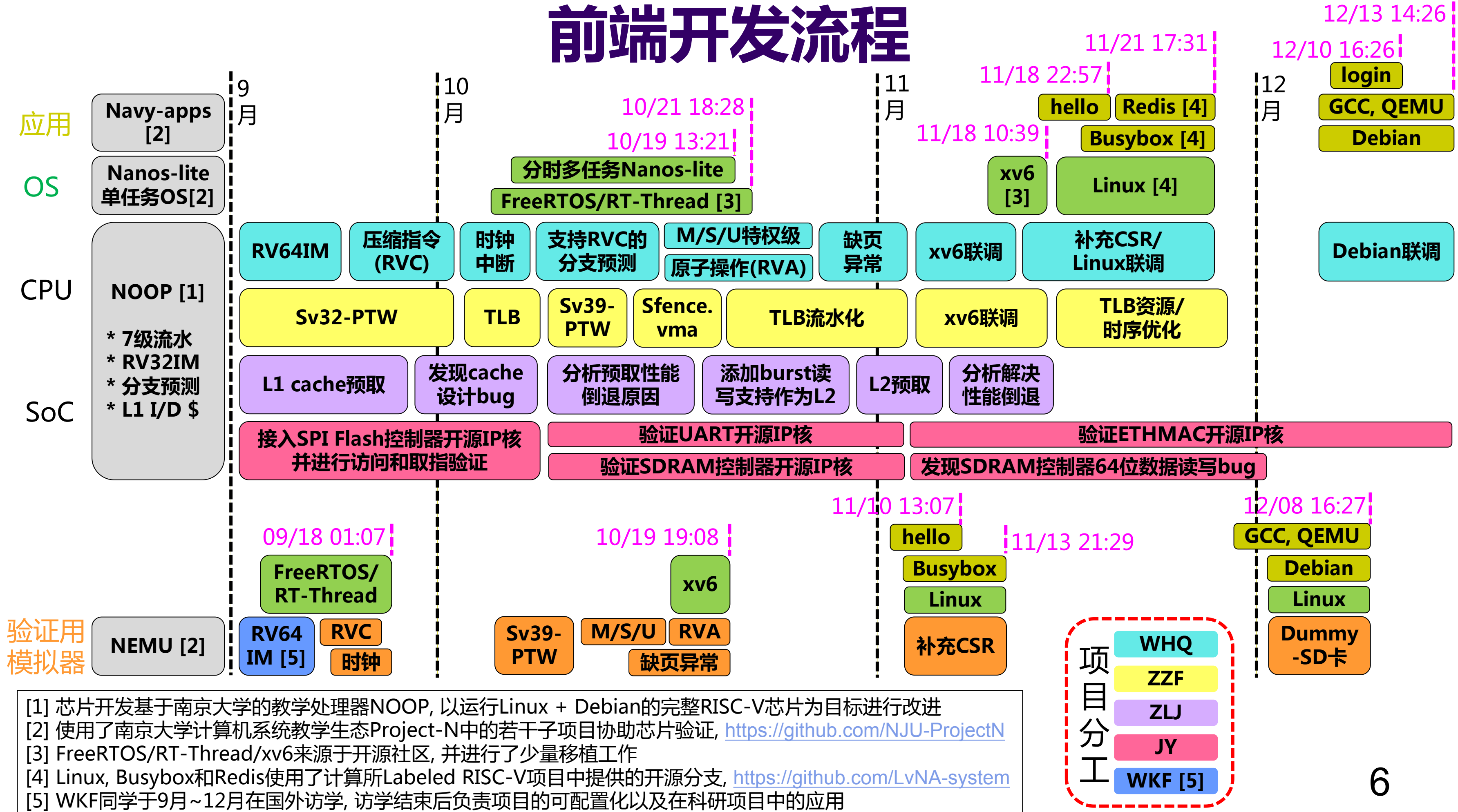
COOSCA (内部代号)

- 使用Chisel语言开发[1]
- 支持SDRAM、SPI flash、UART等外设
- 支持启动Linux 4.18.0内核
- 支持运行Busybox套件
- 在仿真/FPGA环境下支持启动Debian 11
- SMIC 110nm工艺
- 10mm²
- 200mw@350MHz Typical
- TQFP100封装



[1] 外设部分仍使用Verilog进行开发

前端开发流程



[1] 芯片开发基于南京大学的教学处理器NOOP, 以运行Linux + Debian的完整RISC-V芯片为目标进行改进

[2] 使用了南京大学计算机系统教学生态Project-N中的若干子项目协助芯片验证, <https://github.com/NJU-ProjectN>

[3] FreeRTOS/RT-Thread/xv6来源于开源社区, 并进行了少量移植工作

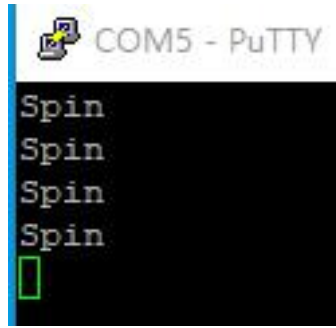
[4] Linux, Busybox和Redis使用了计算所Labeled RISC-V项目中提供的开源分支, <https://github.com/LvNA-system>

[5] WKF同学于9月~12月在国外访学, 访学结束后负责项目的可配置化以及在科研项目中的应用

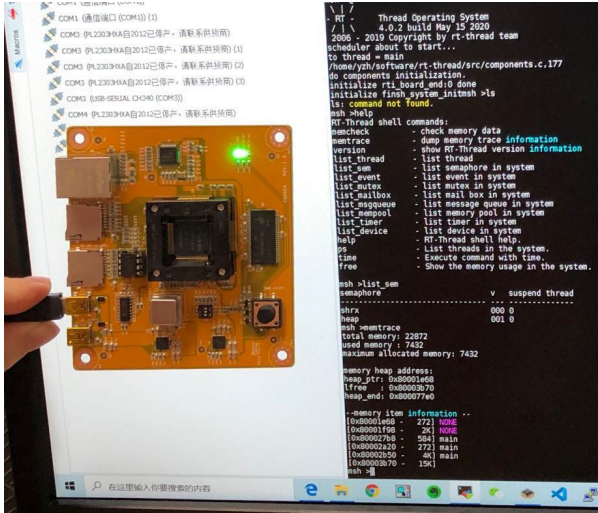
NutShell测试流程和功能展示



封装后的芯片



串口首次输出



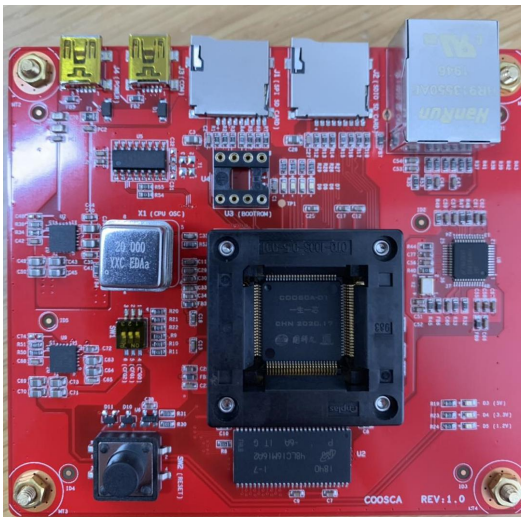
成功运行
RT-Thread

本科
毕业
答辩

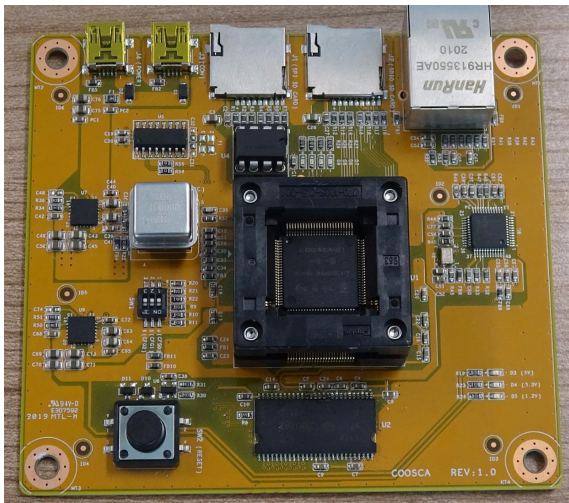


在本科毕业远程答辩中，代表一生一芯团队向评委展示芯片运行的过程

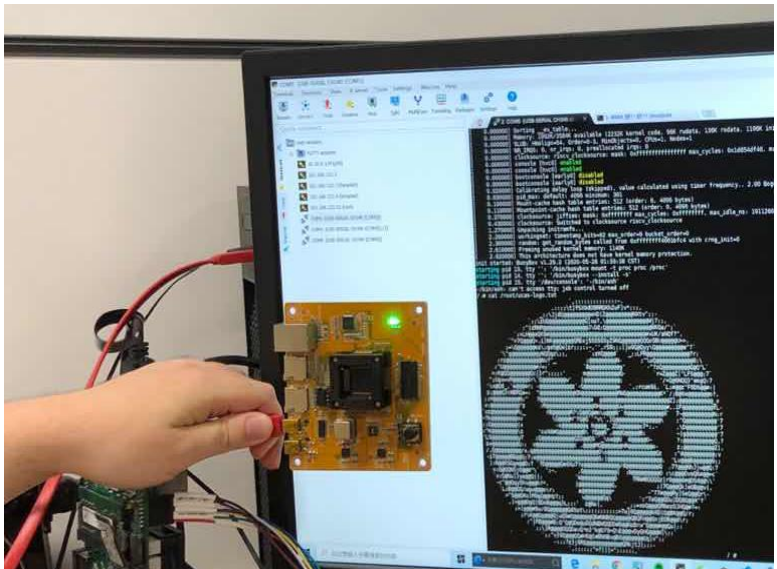
由于SD控制器的缺陷，无法在芯片上启动Debian



第一版
测试板



第二版
测试板



成功运行Linux
并展示中科院logo

分工

- YZH
- CY
- LT

学生成员因疫情无法参与芯片测试流程

NutShell性能展示

由于功能缺陷，只能配备50MHz晶振来使用SDRAM

芯片工作频率

配置开关	倍率	50MHz晶振	100MHz晶振
000	1	50MHz	100MHz
001	1.5	75MHz	150MHz
010	2	100MHz	200MHz
011	2.5	125MHz	250MHz
100	2.75	137.5MHz	275MHz
101	3	150MHz	300MHz
110	3.5	175MHz	350MHz
111	4	200MHz	400MHz

使用SDRAM（如运行Linux）最高频率

无输出

不使用SDRAM（如运行RT-thread）最高频率，与后端时序报告一致

```
vs. 100000 Marks (i7-7700K @ 4.20GHz)
Exit with code = 0
freq = 200MHz
Hello world! Build time: May 29 2020 11:57:58
Filling RAS with legal address...
sp = 0x80300000
Hello world! Build time: May 29 2020 17:04:04
zeroing memory region [0x800056a0, 0x80005eb0)
jump to 0x80000000...
@@@
@@@
@@@
Running CoreMark for 1000 iterations
2K performance run parameters for coremark.
CoreMark Size      : 666
Total time (ms)    : 3344
Iterations         : 1000
Compiler version   : GCC8.3.0
seedcrc            : 0xe9f5
[0]crclist         : 0xe714
[0]crcmatrix       : 0x1fd7
[0]crcstate        : 0x8e3a
[0]crcfinal        : 0xd340
Finised in 3344 ms.
CoreMark 1.0 : (29904 / 100) / GCC8.3.0

=====
CoreMark PASS      873 Marks
vs. 100000 Marks (i7-7700K @ 4.20GHz)
Exit with code = 0
```

CoreMark跑分为1.49/MHz

尝试芯片设计验证的前沿技术

使用Chisel语言进行敏捷开发

```
val ifu = Module(new IFU)
val idu1 = Module(new IDU1)
val idu2 = Module(new IDU2)
val isu = Module(new ISU)
val exu = Module(new EXU)
val wbu = Module(new WBU)

def PipelineConnect2[T <: Data](left: DecoupledIO[T], right: DecoupledIO[T],
  isFlush: Bool, entries: Int = 4, pipe: Boolean = false) = {
  right <-> FlushableQueue(left, isFlush, entries = entries, pipe = pipe)
}

PipelineConnect2(ifu.io.out, idu1.io.in, ifu.io.flushVec(0))
PipelineConnect(idu1.io.out, idu2.io.in, idu2.io.out.fire(), ifu.io.flushVec(1))
PipelineConnect(idu2.io.out, isu.io.in, isu.io.out.fire(), ifu.io.flushVec(1))
PipelineConnect(isu.io.out, exu.io.in, exu.io.out.fire(), ifu.io.flushVec(2))
PipelineConnect(exu.io.out, wbu.io.in, true.B, ifu.io.flushVec(3))
idu1.io.flush := ifu.io.flushVec(1)
idu2.io.flush := ifu.io.flushVec(1)
isu.io.flush := ifu.io.flushVec(2)
exu.io.flush := ifu.io.flushVec(3)
```

国内首个使用Chisel设计的支持Linux+Debian的开源RISC-V核

借鉴软件工程领域的差分测试技术进行处理器的仿真验证

与模拟器进行指令级别在线差分测试 [1] [2]
捕捉了99%的处理器功能bug
5天成功启动Linux + 运行Busybox
4天成功启动Debian + 运行GCC/QEMU

仿真4小时, 经过117亿个周期, 33亿条指令, 仍可瞬间捕捉出错现场, 无需通过波形回溯

2天修复6个启动Debian过程中的复杂bug
皆通过一次定位成功修复
1分钟定位RVC与缺页异常交互的某极端bug

[1] Yu, EasyDiff: An Effective and Efficient Framework for Processor Verification, CRVF 2019, <https://crvf2019.github.io/pdf/14.pdf>
[2] 2018年龙芯杯南京大学二队决赛答辩 报告
<http://www.nscscc.org/a/wangjie/NSCSCC2018/2018/1010/46.html>

使用Chisel进行硬件开发

Chisel 不是 HLS

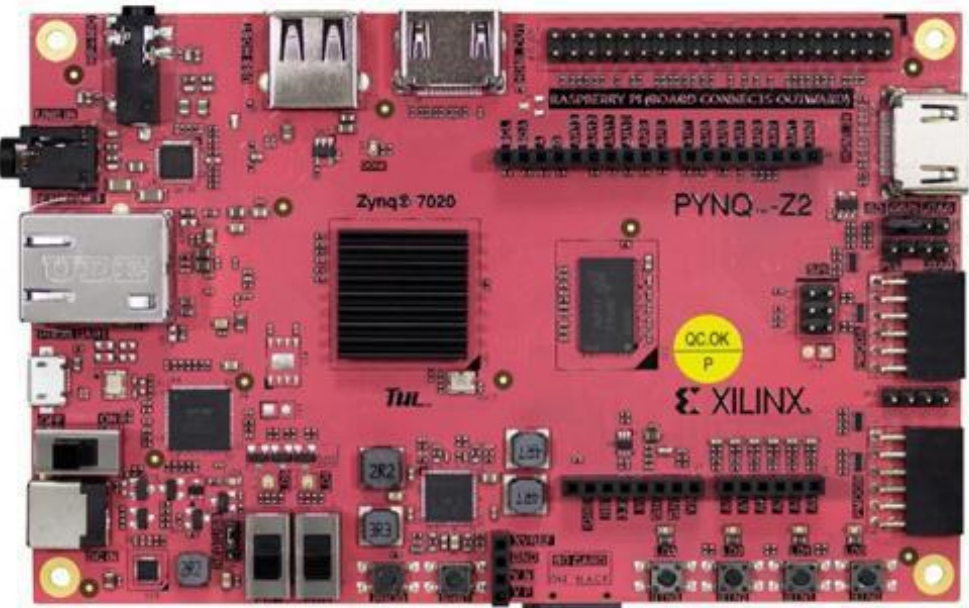
- Chisel是硬件描述语言
- 可以通过Scala中的函数式编程、面向对象特性等来**方便地描述电路**
- 仍能做到对电路的**精确控制**
 - HLS描述算法, **失去了**对电路的精确控制

```
val hitVec = VecInit(metaWay.map(  
  m => m.valid && (m.tag == addr.tag) && io.in.valid  
)).asUInt
```

```
assign _T_24 = metaWay_0_tag == addr_tag; // @[Cache.scala 173:59]  
assign _T_25 = metaWay_0_valid & _T_24; // @[Cache.scala 173:49]  
assign _T_26 = _T_25 & io_in_valid; // @[Cache.scala 173:73]  
assign _T_27 = metaWay_1_tag == addr_tag; // @[Cache.scala 173:59]  
assign _T_28 = metaWay_1_valid & _T_27; // @[Cache.scala 173:49]  
assign _T_29 = _T_28 & io_in_valid; // @[Cache.scala 173:73]  
assign _T_30 = metaWay_2_tag == addr_tag; // @[Cache.scala 173:59]  
assign _T_31 = metaWay_2_valid & _T_30; // @[Cache.scala 173:49]  
assign _T_32 = _T_31 & io_in_valid; // @[Cache.scala 173:73]  
assign _T_33 = metaWay_3_tag == addr_tag; // @[Cache.scala 173:59]  
assign _T_34 = metaWay_3_valid & _T_33; // @[Cache.scala 173:49]  
assign _T_35 = _T_34 & io_in_valid; // @[Cache.scala 173:73]  
assign hitVec = {_T_35, _T_32, _T_29, _T_26}; // @[Cache.scala 173:90]
```


使用Chisel开发的处理器频率符合预期

- 仅针对TLB部分和压缩指令译码部分做少量优化
 - xc7z020-1-clg400 60MHz
 - xczu3cg-sfvc784-1-e 200MHz
 - SMIC 110nm 350MHz



Power

Methodology

DRC

Timing

×

Q

—

H

◇

U

●

Intra-Clock Paths - coreclk_system_top_clk_wiz_0_0 - Setup

Name	Slack	Levels	Routes	High Fanout	From	To	Total Delay	Logic Delay	Net Delay	R
Path 1	0.525	14	14	217	system_top_i..._0/CLKARDCLK	system_top_i...RBWRADDR[8]	15.203	4.945	10.258	
Path 10	0.687	14	14	217	system_top_i..._0/CLKARDCLK	system_top_i...RARDADDR[6]	15.165	4.945	10.220	
Path 11	0.687	14	14	217	system_top_i..._0/CLKARDCLK	system_top_i...RARDADDR[6]	15.165	4.945	10.220	
Path 9	0.686	14	14	217	system_top_i..._0/CLKARDCLK	system_top_i...RARDADDR[6]	15.162	4.945	10.217	

Vivado针对Pynq开发板生成的关键路径报告

优化关键路径时可追溯到Chisel源码

- 对于名如_T_464的信号, 参照Verilog代码中的注释仍能追溯到其来源
 - 这种信号来源于开发过程中生成的匿名信号(如使用map生成的信号)
 - 已有命名的信号名称会保持原样

			Site: SLICE_X47Y71	noop_i/NOOPSoC_0/inst/noop/itlb/tlbExec/_T_8_req_addr[31]_i_6/1
LUT5 (Prop_F6LUT_SLICEL_I1_0)	(f) 0.039	3.663	Site: SLICE_X47Y71	noop_i/NOOPSoC_0/inst/noop/itlb/tlbExec/_T_8_req_addr[31]_i_6/0
net (fo=12, routed)	0.190	3.853		noop_i/NOOPSoC_0/inst/noop/itlb/tlbExec/satp_reg[52]
			Site: SLICE_X46Y71	noop_i/NOOPSoC_0/inst/noop/itlb/tlbExec/_T_464[0]_i_2/12
LUT6 (Prop_H6LUT_SLICEL_I2_0)	(f) 0.149	4.002	Site: SLICE_X46Y71	noop_i/NOOPSoC_0/inst/noop/itlb/tlbExec/_T_464[0]_i_2/0
net (fo=76, routed)	0.194	4.196		noop_i/NOOPSoC_0/inst/noop/itlb/tlbExec/_T_464[0]_i_1_n_0
			Site: SLICE_X46Y70	noop_i/NOOPSoC_0/inst/noop/itlb/tlbExec/ram_icachePF_reg_0_3_0_0_i_4/12

```
8028 wire _T_431 = _T_425 & _T_430; // @[EmbeddedTLB.scala 311:87]
8029 wire _T_433 = io_pf_status_mxr & _T_386[3]; // @[EmbeddedTLB.scala 313:68]
8030 wire _T_434 = _T_386[1] | _T_433; // @[EmbeddedTLB.scala 313:51]
8031 wire _T_435 = _T_431 & _T_434; // @[EmbeddedTLB.scala 313:36]
8032 wire _T_458 = _T_435; // @[EmbeddedTLB.scala 327:19]
8033 wire _T_464 = _T_458 & _T_333; // @[EmbeddedTLB.scala 327:29]
8034 wire _T_436 = _T_431 & _T_386[2]; // @[EmbeddedTLB.scala 314:37]
8035 wire _T_465 = _T_436; // @[EmbeddedTLB.scala 327:50]
```

```
326 if (tlbname == "dtlb") {
327   when ((!permLoad && req.isRead()) || (!permStore && req.isWrite())) {
328     state := s_miss_slpf
329     loadPF := req.isRead() && !isAMO
330     storePF := req.isWrite() || isAMO
331   }.otherwise {
332     state := Mux(updateAD, s_write_pte, s_wait_resp)
333     missMetaRefill := true.B
334   }
335 }
```


处理器在线差分测试仿真验证[1][2]

处理器在线差分测试仿真验证

- 传统的测试:
 - 生成正确的Trace, 运行处理器并对比Trace, dump波形用于观察

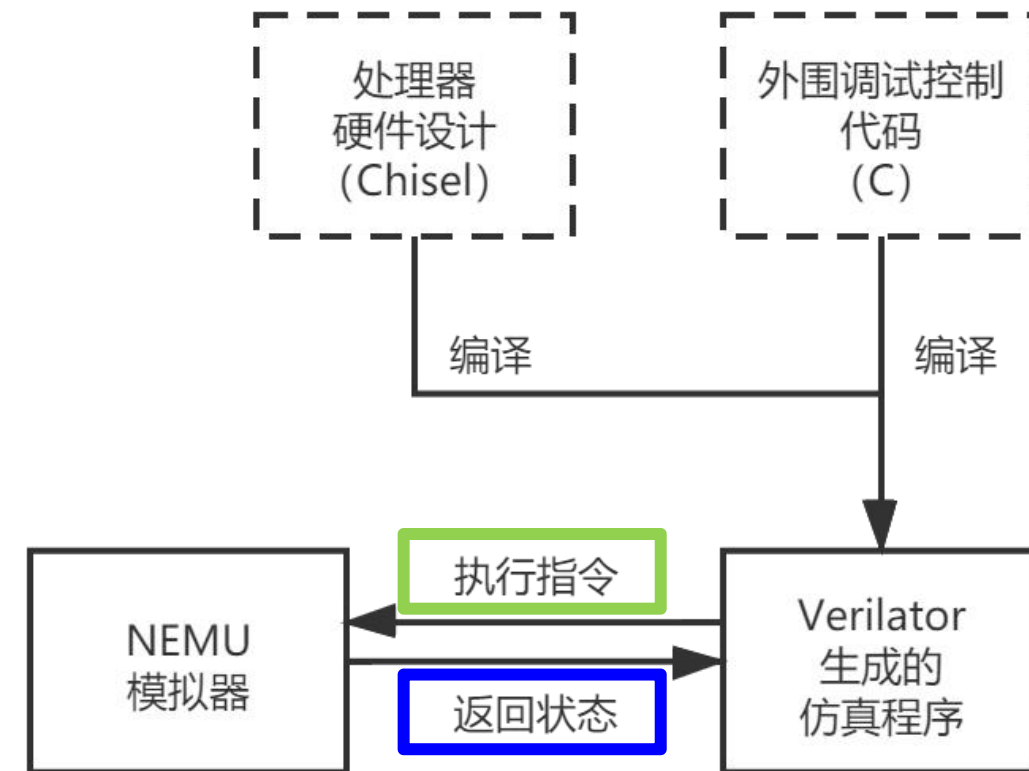
	传统测试方法
存储空间占用	高: 波形与Trace需要占用大量空间
验证效率	较低: 需要dump波形, 产生大量IO开销, 拖慢验证速度
灵活性	低: 无法应对变化的外部输入 (如外部设备状态的变化)
调试难度	较高: 波形可读性差

处理器在线差分测试仿真验证

	传统测试方法	我们使用的测试方法
存储空间占用	高: 波形与Trace需要占用大量空间	低
验证效率	较低: 需要dump波形, 产生大量IO开销, 拖慢验证速度	较高
灵活性	低: 无法应对变化的外部输入 (如外部设备状态的变化)	高
调试难度	较高: 波形可读性差	较低

处理器在线差分测试仿真验证-原理

- 基于指令级在线比对的验证框架
 - 模拟器作为动态链接库连接到Verilator生成的仿真程序中
 - 处理器在指令完成时发出比对请求
 - 在比对结果不一致时报错并输出当前处理器状态



测试框架结构示意图

```
while (1) {  
    verilator_step();  
    nemu_step(1);  
    verilator_getregs(&r1);  
    nemu_getregs(&r2);  
    if (r1 != r2) { abort(); }  
}
```

在线比对机制

处理器在线差分测试仿真验证-原理

- 特殊情况的处理
 - 外部中断: 处理器传出中断信息, 模拟器进入相同的处理流程
 - 其他例外: 模拟器支持处理例外, 处理器和模拟器的行为应当一致
 - Memory mapped IO: 处理器标识出这样的访存指令, 在这样的指令提交时从处理器将访存结果复制到模拟器中
 - 例如, 读取(内存地址映射的)时钟计数器
- 这样的处理使得我们的比对机制拥有很高的灵活性

```
if (isMMIO) {  
    // MMIO accessing should not be a branch or jump, just +2/+4 to get  
    int pc_skip = 0;  
    pc_skip += isRVC ? 2 : 4;  
    pc_skip += isMultiCommit ? (isRVC2 ? 2 : 4) : 0;  
    reg_scala[DIFFTEST_THIS_PC] += pc_skip;  
    nemu_this_pc += pc_skip;  
    // to skip the checking of an instruction, just copy the reg state  
    ref_difftest_setregs(reg_scala);  
    pc_retire_pointer = (pc_retire_pointer+1) % DEBUG_RETIRE_TRACE_SIZE;  
    pc_retire_queue[pc_retire_pointer] = this_pc;  
    inst_retire_queue[pc_retire_pointer] = this_inst;  
    need_copy_pc = 1;  
    return 0;  
}  
  
if (intrNO) {  
    ref_difftest_raise_intr(intrNO);  
} else {  
    ref_difftest_exec(1);  
}
```

测试框架中处理中断/MMIO的部分代码

处理器在线差分测试仿真验证-优势

- 模拟器在仿真过程中同步生成golden trace , 提高验证效率
 - 不需要提前生成trace
 - 可以支持外部中断
 - 模拟器速度快, 不会对仿真速度造成较大负担
 - 无需dump过多波形, 减少IO开销和磁盘空间占用
- 易于自行修改模拟器实现来辅助验证
 - 例如在访存相关测试时, 针对性地给出特定访存的golden trace
 - 模拟器NEMU是实现了所需功能的最小子集
 - 代码精简, 容易实现
- 直观
 - 直接使用print进行调试, 可读性高
 - 几乎无需观察波形

处理器在线差分测试仿真验证

- 测试使用的指令负载包括:

代码片段	性能测试基准程序	操作系统
ProjectN提供的CPU功能测试与性能测试	CoreMark	教学操作系统 XV6
从riscv-test引入的部分测试	Dhrystone	实时操作系统 RT-Thread
riscv-torture 随机指令测试		实时操作系统 FreeRTOS
		Linux Kernel
		Debian

- 依托这一套框架, 在仿真中启动Debian进行验证

处理器在线差分测试仿真验证-示例

- 使用测试框架定位复杂BUG示例
 - 示例: 跨页指令相关的BUG
 - 在Debian上运行gcc编译复杂程序时发现错误
 - 测试框架在仿真**4小时, 117亿**个周期, **33亿**条指令后捕捉到了这一错误
 - 暴露出某些很难被检查到的边界问题

```
===== Reg Diff =====
$0: 0x0000000000000000 ra: 0x000000000000005fb0 sp: 0xffffffe00acefee0 gp: 0x00000000000d01348
tp: 0xffffffe00abba6c0 t0: 0x0000000000000046000 t1: 0x0000000000000000 t2: 0x000000000000003ff
s0: 0x00000003ffcd32e0 s1: 0x800000000000006020 a0: 0x000000000000000d a1: 0x0000000000000000
a2: 0x00000000000000068 a3: 0x0000000000000030000 a4: 0x0000000000000000 a5: 0x000000000000000d
a6: 0x00000000000d6a022 a7: 0x00000000000000007 s2: 0x0000000000000008 s3: 0x00000000000d02000
s4: 0x00000000000000007 s5: 0x000000000000004000 s6: 0x0000000200067d708 s7: 0x0000000000000000
s8: 0x0000000000000000 s9: 0x0000000000000000 s10: 0x00000000000e04808 s11: 0x0000000000f2d470
t3: 0x0000000200032dcf8 t4: 0x000000000000000f t5: 0x00000000000000012f t6: 0x0000000000000001
pc: 0xffffffe00001dd38 mstatus: 0x00000000000000a0 mcause: 0x0000000000000002 mepc: 0x000000000020756c
sstatus: 0x0000000000000020 scause: 0x000000000000000c sepc: 0x00000000000005ffe

===== Csr Diff =====
privilegeMode: 0x0000000000000001
mstatus: 0x00000000000000180 mcause: 0x0000000000000002 mepc: 0x000000000020756c
sstatus: 0x00000000000000100 scause: 0x000000000000000c sepc: 0x00000000000005ffe
x 9 different at pc = 0xffffffe00001dd34, right= 0x800000000000006020, wrong = 0x800000000000006100
ABORT at pc = 0x412316982584
total guest instructions = 3317556127
instrCnt = 3317556127, cycleCnt = 11736118051, IPC = 0.282679
Guest cycle spent: 11736118052
Host time spent: 18251437ms
make[1]: *** [Makefile:90: emu] Error 1
make[1]: Leaving directory '/home/yzh/projectn/noop'
make: *** [Makefile:110: noop] Error 2
[7:21:22 ~/projectn/riscv-pk]%

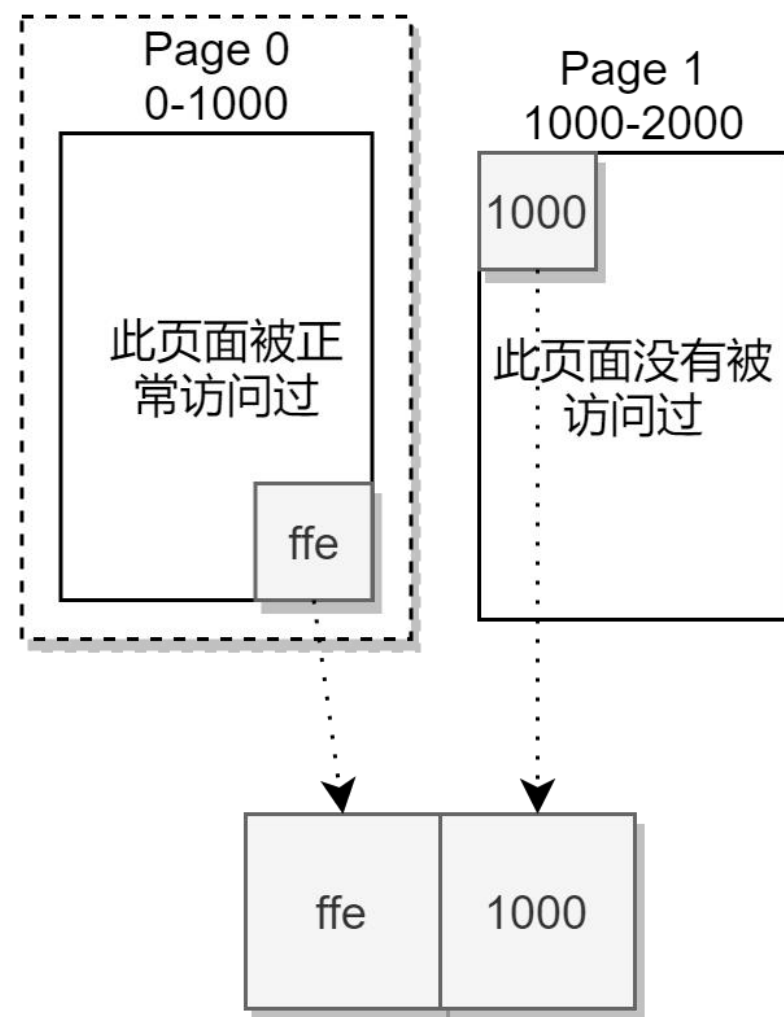
30155 ffffffe00001dd28: f9fa sd t5,240(sp)
30156 ffffffe00001dd2a: fdfe sd t6,248(sp)
30157 ffffffe00001dd2c: 000462b7 lui t0,0x46
30158 ffffffe00001dd30: 02023403 ld s0,32(tp) # 20 < __vdso_rt_sigreturn-0x7e0>
30159 ffffffe00001dd34: 1002b4f3 csrrc s1,sstatus,t0
30160 ffffffe00001dd38: 14102973 csrr s2,sepc
30161 ffffffe00001dd3c: 143029f3 csrr s3,stval
30162 ffffffe00001dd40: 14202a73 csrr s4,scause
```

测试框架在此处报错

处理器在线差分测试仿真验证-示例

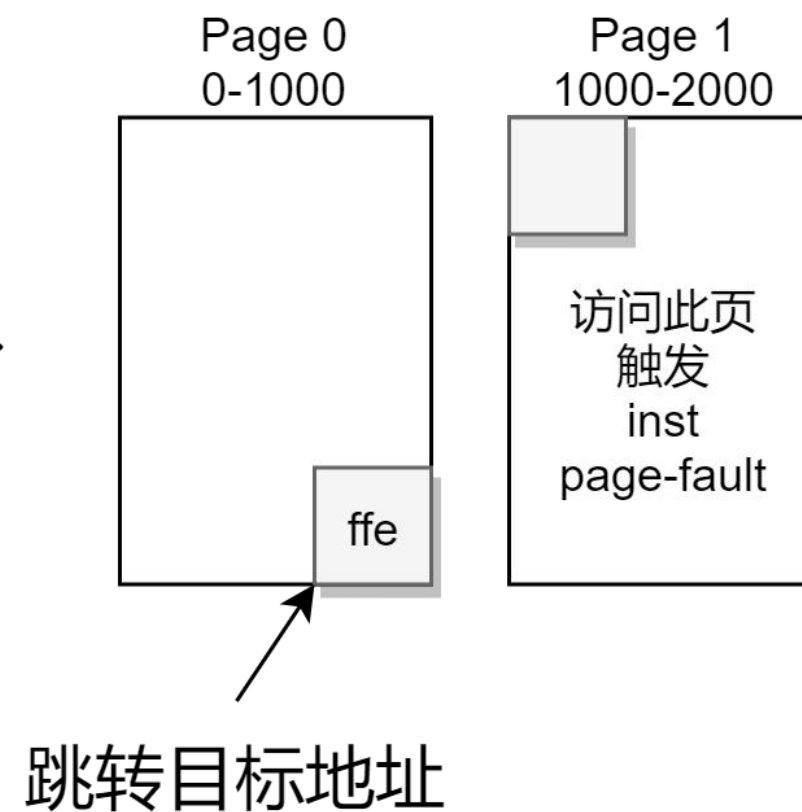
- 示例: 跨页指令相关的BUG

一个地址为0x0-0x1000的页面, 其末尾0xffe处为一条跨页的指令



阴影部分是一条32位指令的前后两个部分

处理器跳转到地址为ffe的指令且恰好触发指令page-fault



处理器在线差分测试仿真验证-示例

- 使用测试框架定位复杂BUG示例
 - 示例: 跨页指令相关的BUG

Volume II: RISC-V Privileged Architectures V1.12-draft

41

For misaligned loads and stores that cause access-fault or page-fault exceptions, `mtval` will contain the virtual address of the portion of the access that caused the fault. For instruction access-fault or page-fault exceptions on systems with variable-length instructions, `mtval` will contain the virtual address of the portion of the instruction that caused the fault while `mepc` will point to the beginning of the instruction.

指令集手册对这一情况的描述

开发期间针对这一问题的讨论

- 指令集手册要求在这种指令发生IPF时, 向`mepc`与`mtval`寄存器写入不同的值

直接跳转到 ffe, 而且这里恰好还是个 IPF



嗯, 也有可能是 1000 的页面之前已经被访问过了

所以就算 ffe 放了 4 字节指令, 也不会 IPF

对, 只要不是直接跳转到 ffe 就没有四字节指令的问题

额, 等下, 还是有的

毕竟 linux 的 lazy paging 只会在第一次的时候触发

如果第一次访问一个页面不是以这种跨页面的 4 字节形式访问的, 这个页面就不会暴露出这种问题

处理器在线差分测试仿真验证-示例

- 很多bug被光速修复, 仅仅留下了git log
 - 基于差分测试框架, 在两天内处理了12个启动Linux时的Bug

Commits on Nov 16, 2019

Merge branch 'dev-linux-priv' into 'dev-linux' ...

AugustusWillisWang committed on 16 Nov 2019

add(riscv64,CSR): set mtval when triggerring page fault in M-mode ...

AugustusWillisWang committed on 16 Nov 2019

add(riscv64,decode): implement inst wfi as nop

AugustusWillisWang committed on 16 Nov 2019

add(riscv64,exc,CSR): read/write unimplemented csr now triggers an il... ...

AugustusWillisWang committed on 16 Nov 2019

add(riscv64,intr,CSR): set mtval and stval to 0 except for page fault

AugustusWillisWang committed on 16 Nov 2019

Merge branch 'dev-linux-priv' into 'dev-linux' ...

AugustusWillisWang committed on 16 Nov 2019

fix(LSU): fix AMO inst decode error

AugustusWillisWang committed on 16 Nov 2019

add(riscv64,CSR): add mstatus.sd bit

AugustusWillisWang committed on 16 Nov 2019

chore(debug): add noop csr display for difftest

AugustusWillisWang committed on 16 Nov 2019

Commits on Nov 17, 2019

Merge branch 'dev-linux' into dev-linux-tlb

Lemover committed on 17 Nov 2019

Merge branch 'dev-linux-fix-tlbpriv' into 'dev-linux' ...

AugustusWillisWang committed on 17 Nov 2019

fix bug: disable vmEnable at ModeM && add isAMO: loadPF -> storePF wh... ...

Lemover committed on 17 Nov 2019

noop,fu,CSR: fix wrong mtval/stval for Instruction PF ...

sashimi-yzh committed on 17 Nov 2019

Merge branch 'dev-linux-priv' into 'dev-linux' ...

AugustusWillisWang committed on 17 Nov 2019

fix(IDU): ECALL will not influence exceptionVec in IDU

AugustusWillisWang committed on 17 Nov 2019

add(CSR): add exeception priority

AugustusWillisWang committed on 17 Nov 2019

mod(CSR): write satp will now flush pipeline ...

AugustusWillisWang committed on 17 Nov 2019

fix(decode): fix decode for invalid inst 0

AugustusWillisWang committed on 17 Nov 2019

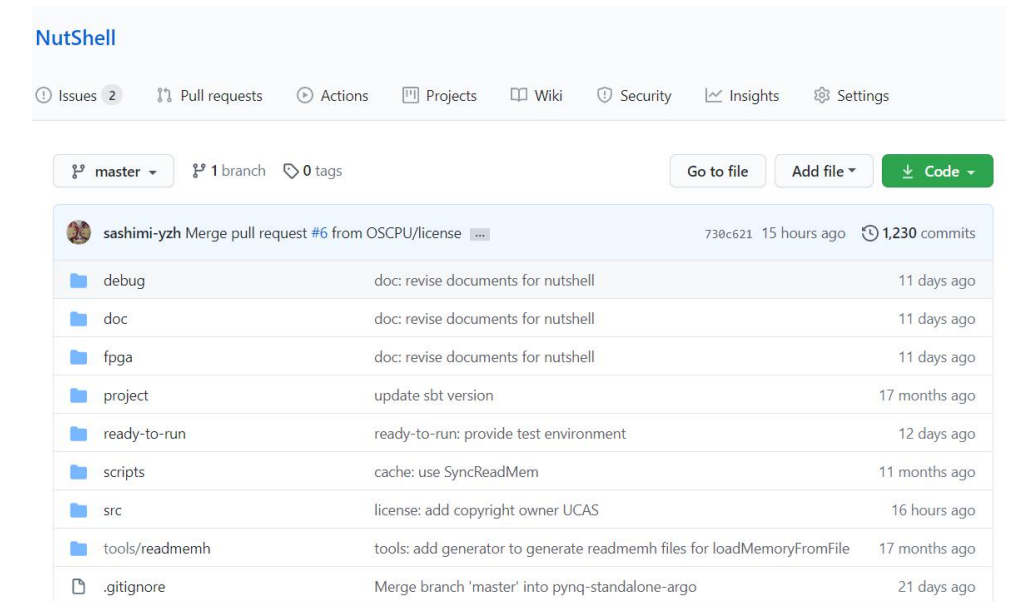
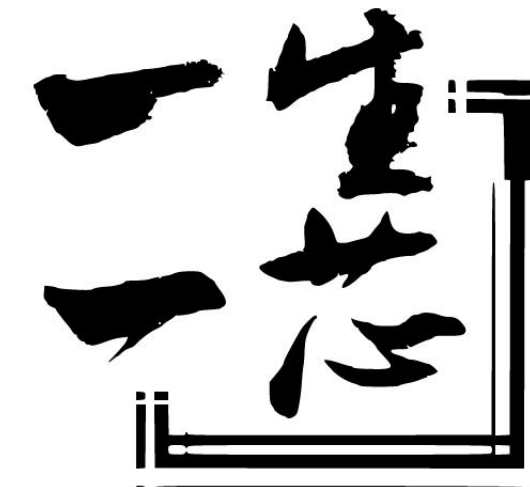
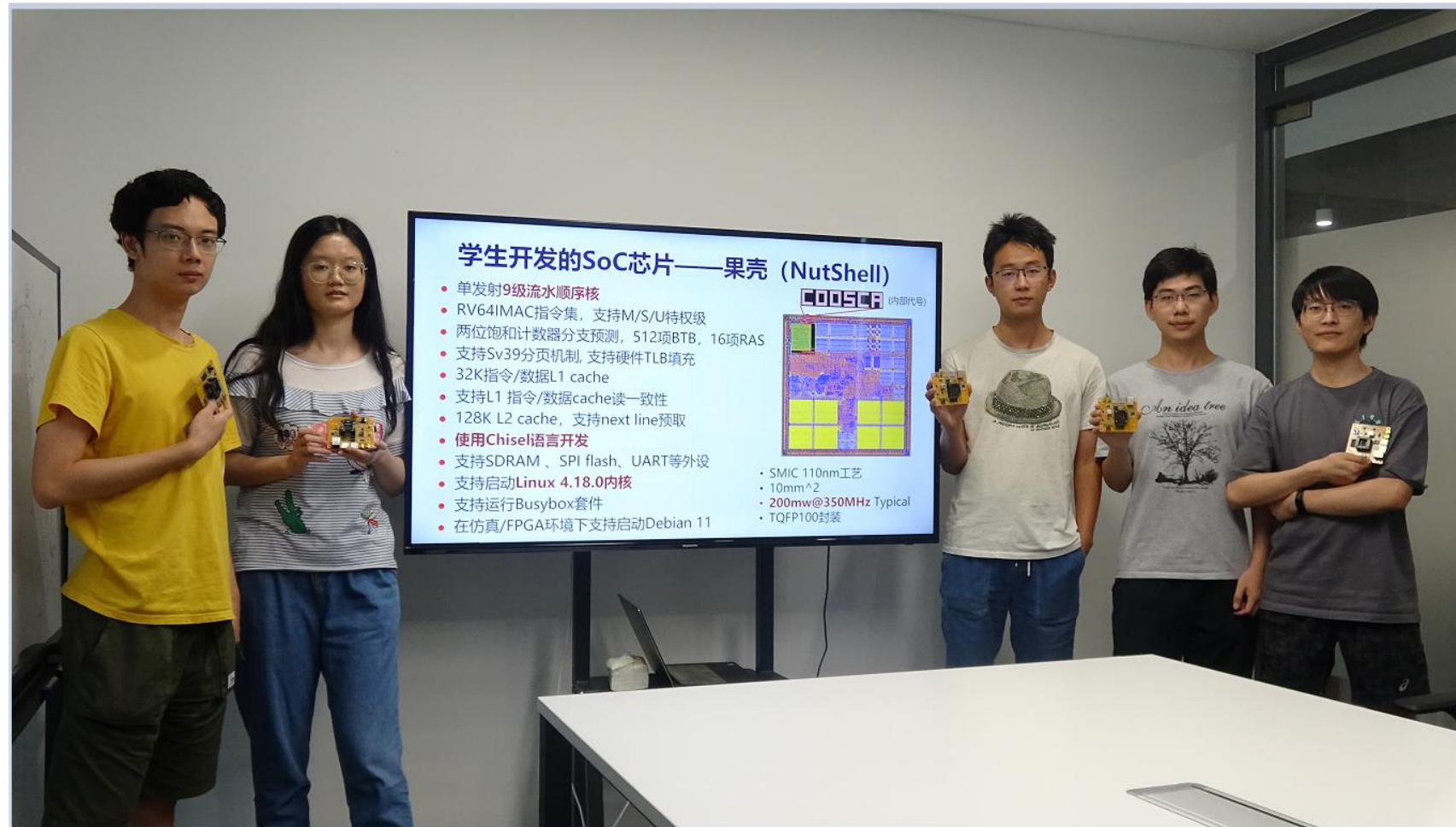
chore(debug): add linux to Makefile

AugustusWillisWang committed on 17 Nov 2019

fix(CSR): add inst set 'u' to misa

AugustusWillisWang committed on 17 Nov 2019

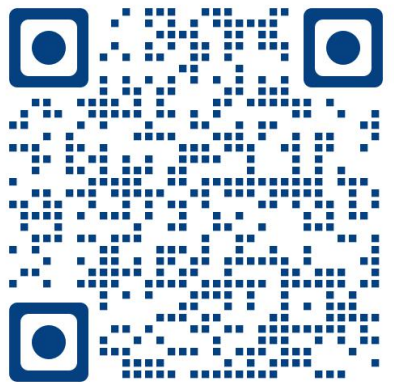
开源与致谢



Open Source Chip Project by University (OSCPU)

Let students design their own chips!

<https://github.com/OSCPU/NutShell>



感谢参与学生指导和为本项目提供帮助的以下人员（姓氏拼音排序）：

蔡晔老师(深圳大学)、李峰工程师(鹏城实验室)、刘彤工程师、唐丹高级工程师、王海喆博士生、徐易难博士生和余子濠博士生

感谢中国科学院计算技术研究所开源芯片工作组对本项目的指导

感谢南京大学计算机系统教学生态Project-N对本项目的开发和验证提供的帮助

NutShell-本科生设计的可运行Linux的RISC-V芯片

Q & A

中国科学院大学

王华强

2020.7.18