



# 香山：用敏捷开发方法 实现开源高性能RISC-V处理器

---

香山团队

2021/11/4@海思OpenNetChip技术论坛

# 开源芯片包含三个层次

① 指令集 + ② 处理器微架构设计/实现 + ③ 设计流程/工具

处理器设计新方法、新流程、新平台

开源

微架构设计

工程开发

EDA工具

指令集手册

设计文档

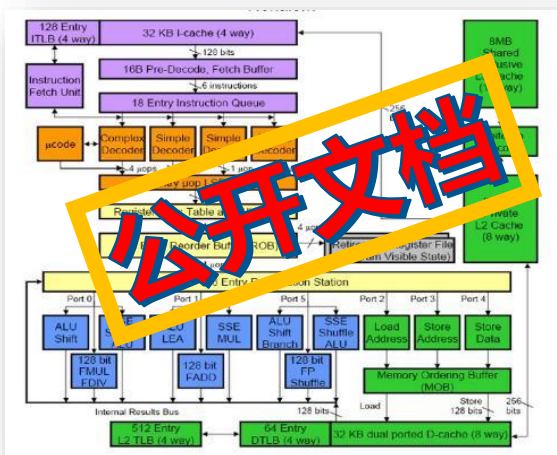
RTL代码

芯片版图

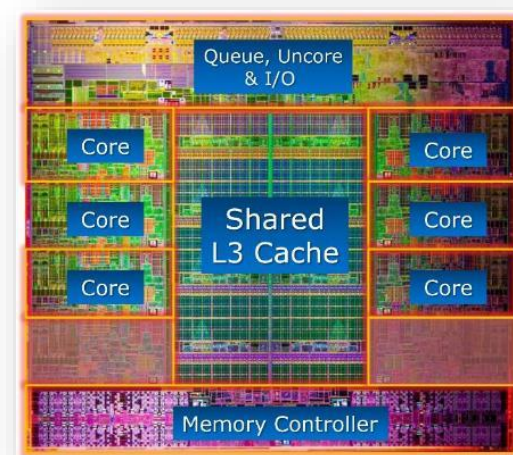
开放免费  
共同制定

公开文档

开放源码



```
component DebugCoreTop is
port (
  -- Trigger and Data
  cu_Clk      : in  std_logic_vector(31 downto 0) := (others => '0');
  cu0_Trig    : in  t_trig_0 := (others => '0');
  cu1_Trig    : in  t_trig_1 := (others => '0');
  cu2_Trig    : in  t_trig_2 := (others => '0');
  cu0_Data    : in  std_logic_vector(31 downto 0) := (others => '0');
  cu1_Data    : in  std_logic_vector(31 downto 0) := (others => '0');
  cu2_Data    : in  std_logic_vector(31 downto 0) := (others => '0');
  -- Downstream
  SCL         : in  std_logic := '0';
  SDA         : in  std_logic := '0';
  -- Upstream
  gt_RefClk_p : in  std_logic := '0';
  gt_RefClk_n : in  std_logic := '0';
  gt_RX_p     : in  std_logic_vector(2 downto 0) := (others => '0');
  gt_RX_n     : in  std_logic_vector(2 downto 0) := (others => '0');
  gt_TX_p     : out std_logic_vector(2 downto 0);
  gt_TX_n     : out std_logic_vector(2 downto 0);
);
end component;
```



# 为什么做开源高性能 RISC-V 核？

- RISC-V的一大优势——形成“**竞争前合作**”，实现各界**联合开发开源CPU架构**
- 国内外现有开源RISC-V项目**尚不满足业界对高性能处理器的需求**
- 建立像Linux的被工业界广泛应用的**体系结构创新开源平台——存活30年！**
- 用于研究和验证芯片**敏捷开发方法、流程与工具**

微架构设计 指令集	1 开放免费的设计	2 需授权的设计	3 封闭的设计	产品可选的设计 (对应各指令集)
开放免费的指令集 (RISC-V)	Berkeley的Rocket Chip/剑桥lowRISC/ 芯来科技蜂鸟E203	平头哥/SiFive/晶 心科技Andes的 RISC-V处理器核	Google和 NVIDIA的自研 RISC-V处理器	1 2 3
需授权的指令集 (ARM)		ARM的处理器设计, 如Cortex-A76等	基于ARM架构的 Apple处理器	2 3
封闭的指令集 (x86)			Intel和AMD的 处理器	3

# 香山：开源高性能 RISC-V 处理器

- **第一版：雁栖湖架构**

- 2020/6：代码仓库建立
- 2021/4：RTL完成
- 2021/6：在RISC-V中国峰会正式发布
- 频率：1.3GHz@28nm
- 性能：预估SPEC CPU2006 ~7分/GHz

- **第二版：南湖架构**

- 2021/5：开始RTL实现工作，同时持续进行设计讨论
- 2021/10：功能特性确定，进入时序收敛阶段
- 计划：11月完成RTL freeze，22年初投片，**目标SPEC CPU2006 20 分，2GHz@14nm**

- **开源情况**

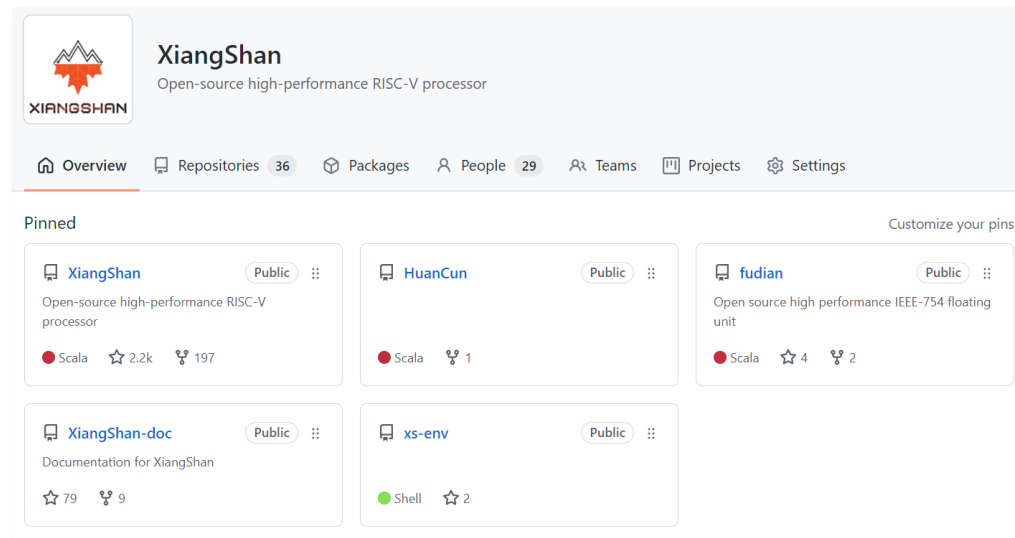
- 开源协议：**MulanPSLv2协议**（兼容Apache v2.0）
- 代码托管：GitHub (<https://github.com/OpenXiangShan/XiangShan>)；镜像：Gitee/Trustie/iHub
- 报告/文档/进展：微信公众号、知乎、微博 香山开源处理器
- 邮件列表：[xiangshan-all@ict.ac.cn](mailto:xiangshan-all@ict.ac.cn)



香山源码仓库

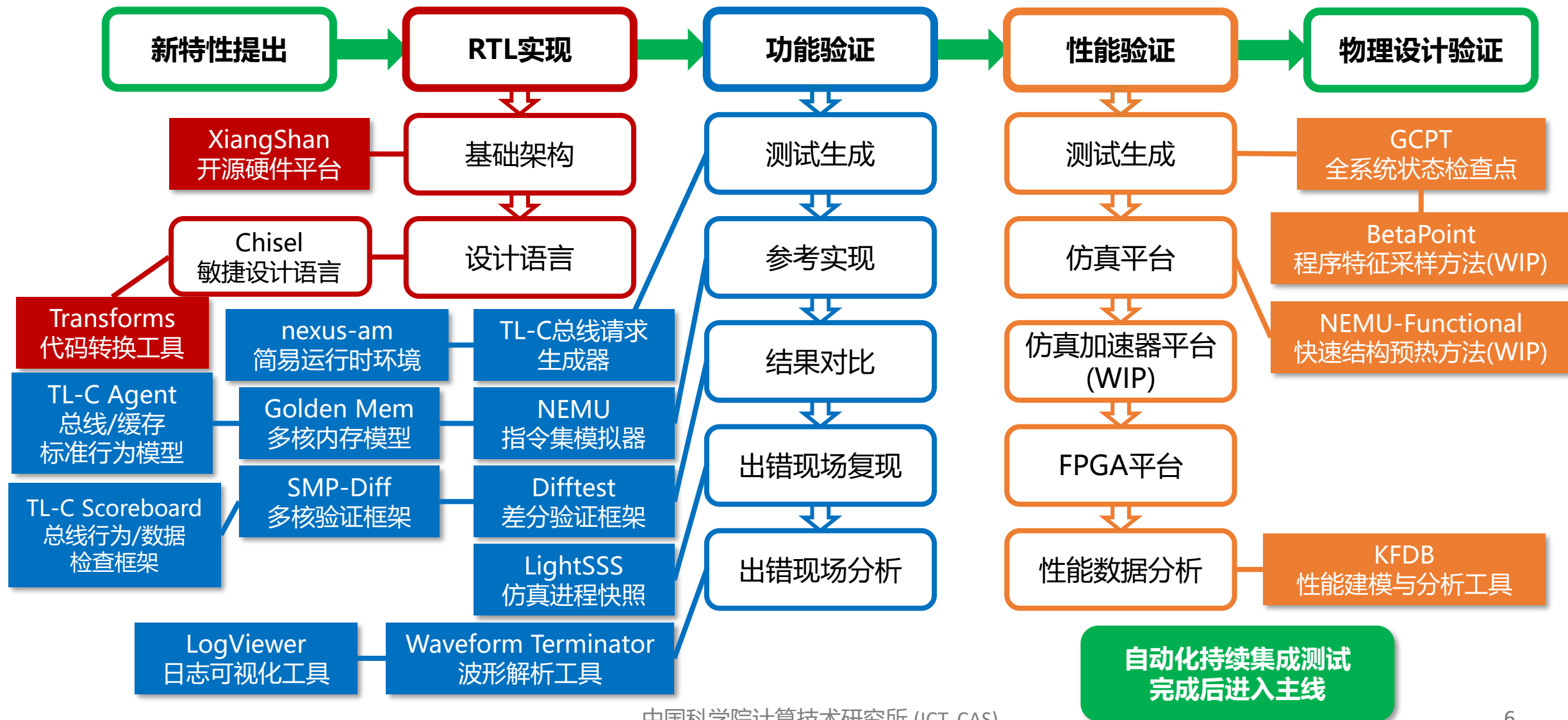
# 南湖：香山下一代微结构

- **改进方向①：取指与分支预测**
  - 采用解耦的前端取指与分支预测架构
  - 更高的吞吐率、更高的预测准确率
- **改进方向②：运算单元**
  - 支持 RISC-V 位操作 (B)、标量加密运算 (K) 扩展指令集
  - 开源高性能浮点运算部件 FuDian
- **改进方向③：功能支持**
  - 支持 RISC-V 物理内存保护机制 (PMP) 等安全特性
  - 支持自定义的可配置物理内存属性 (PMA)
- **改进方向④：Load Store Unit**
  - 提高 TLB、L1 Cache 等结构容量
  - 重构 L1 DCache 流水线设计，降低访问冲突，提高效率
- **改进方向⑤：L2/L3 Cache**
  - 支持 inclusive / non-inclusive 的开源高性能 L2/LLC 模块 HuanCun



**所有修改全部开源**

# 香山：高性能处理器敏捷开发实践







# 新型硬件设计语言带来编码效率的大幅提升

- 2018年：设计定量实验对比Chisel与Verilog，代码量降低80%

## Chisel vs. Verilog：芯片敏捷开发效率对比案例

- 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

- 敏捷开发的效率是传统 14 倍!
- 代码量约为传统开发的 1/5

## 开发质量对比：敏捷 vs 传统

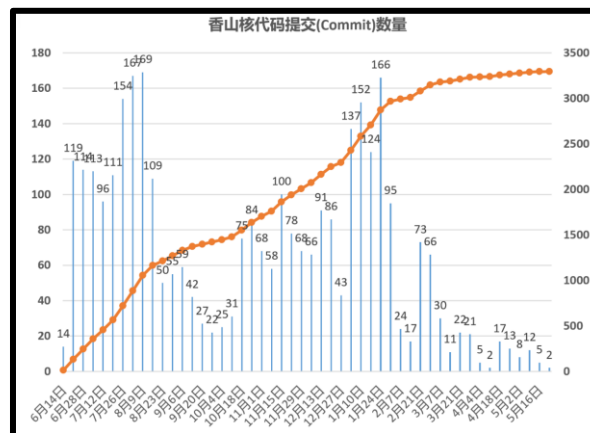
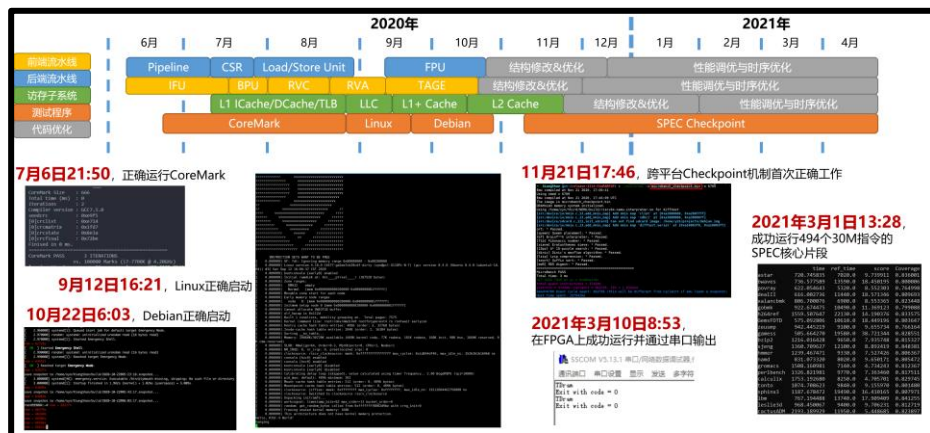
- 让另一名Chisel零基础的国科大本科生翻译工程师的核心模块并评估
  - Vivado 2017.01, FPGA 型号 xc7v2000tthg1716-1

	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

余子濠, 刘志刚, 李一苇, 黄博文, 王州, 孙凝晖, 包云岗. 芯片敏捷开发实践：标签化RISC-V. 计算机研究与发展, 2019

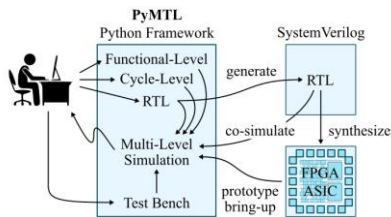
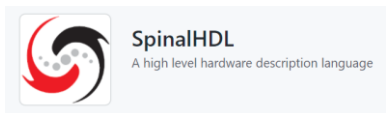
- 2020年：采用Chisel语言完成香山高性能处理器实现，三个月启动Linux



香山：开源高性能RISC-V处理器. RISC-V中国峰会, 2021

# 新型硬件设计语言驱动敏捷硬件设计范式变化

CHISEL



bluespec

cucapra/latte21

Languages, Tools, and Techniques for Accelerator Design

Scala

python

OCaml

Verilog

SystemVerilog

敏捷硬件设计语言

新型硬件设计范式

抽象层

高抽象层次的  
硬件描述语言

硬件描述语言

描述层

处理器体系结构

结构层

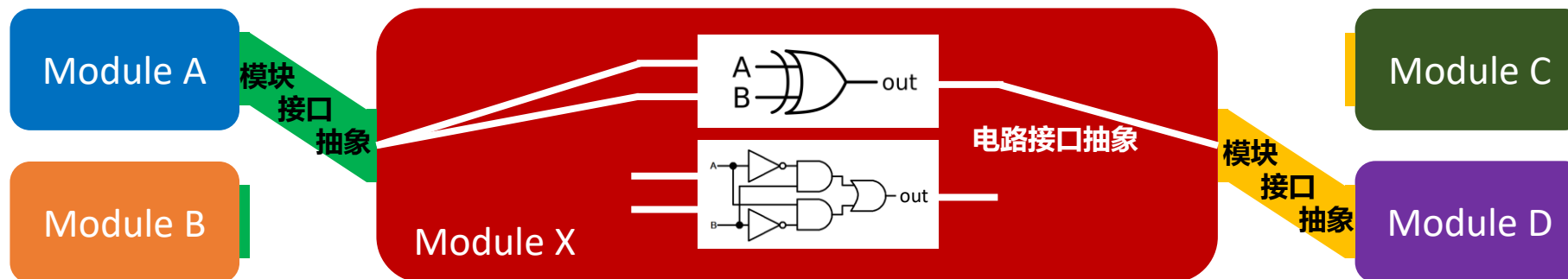
逻辑电路

电路层

敏捷硬件设计方法



# 抽象层次①：模块/电路接口



## • 示例：替换算法、数据校验算法

```
29 object ReplacementPolicy {
30   //for fully associative mapping
31   def fromString(s: Option[String], n_ways: Int): ReplacementPolicy = fromString(s.getOrElse("random"), n_ways)
32   def fromString(s: String, n_ways: Int): ReplacementPolicy = s.toLowerCase match {
33     case "random" => new RandomReplacement(n_ways)
34     case "lru"    => new TrueLRU(n_ways)
35     case "plru"   => new PseudoLRU(n_ways)
36     case t => throw new IllegalArgumentException(s"unknown Replacement Policy type $t")
37   }
38   //for set associative mapping
39   def fromString(s: Option[String], n_ways: Int, n_sets: Int): SetAssocReplacementPolicy = fromString(s.getOrElse("random"), n_ways, n_sets)
40   def fromString(s: String, n_ways: Int, n_sets: Int): SetAssocReplacementPolicy = s.toLowerCase match {
41     case "random" => new SetAssocRandom(n_sets, n_ways)
42     case "setlru"  => new SetAssocLRU(n_sets, n_ways, "lru")
43     case "setplru" => new SetAssocLRU(n_sets, n_ways, "plru")
44     case t => throw new IllegalArgumentException(s"unknown Replacement Policy type $t")
45   }
46 }
```

(1) 实现不同的替换算法

```
97 icacheParameters: ICacheParameters = ICacheParameters(
98   tagECC = Some("parity"),
99   dataECC = Some("parity"),
100   replacer = Some("setplru"),
101   nMissEntries = 2
102 ),
103 l1plusCacheParameters: L1plusCacheParameters = L1plusCacheParameters(
104   tagECC = Some("secced"),
105   dataECC = Some("secced"),
106   replacer = Some("setplru"),
107   nMissEntries = 8
108 ),
109 dcacheParameters: DCacheParameters = DCacheParameters(
110   tagECC = Some("secced"),
111   dataECC = Some("secced"),
112   replacer = Some("setplru"),
113   nMissEntries = 16,
114   nProbeEntries = 16,
115   nReleaseEntries = 16,
116   nStoreReplayEntries = 16
117 ),
```

(2) 基于Scala语言完成算法的实例化

# 抽象层次②：基础硬件模块/数据结构

Module X

Module X

Module X

Module X

- 示例：不同大小的硬件队列（包括队列指针、数据存储空间等）

```
class CircularQueuePtr[T <: CircularQueuePtr[T]](val entries: Int) extends Bundle {  
  def this(f: Parameters => Int)(implicit p: Parameters) = this(f(p))  
  
  val PTR_WIDTH = log2Up(entries)  
  val flag = Bool()  
  val value = UInt(PTR_WIDTH.W)  
}
```

- (1) 定义一个参数化的队列指针类型

```
11 class FtqPtr(implicit p: Parameters) extends CircularQueuePtr[FtqPtr]({  
12   p => p(XScoreParamsKey).FtqSize  
13 })  
14 {  
15   override def cloneType = (new FtqPtr).asInstanceOf[this.type]  
16 }  
17 object FtqPtr {  
18   def apply(f: Bool, v: UInt)(implicit p: Parameters): FtqPtr = {  
19     val ptr = Wire(new FtqPtr)  
20     ptr.flag := f  
21     ptr.value := v  
22     ptr  
23   }  
24 }
```

- (2) 基于队列大小参数继承得到对应队列的指针类型

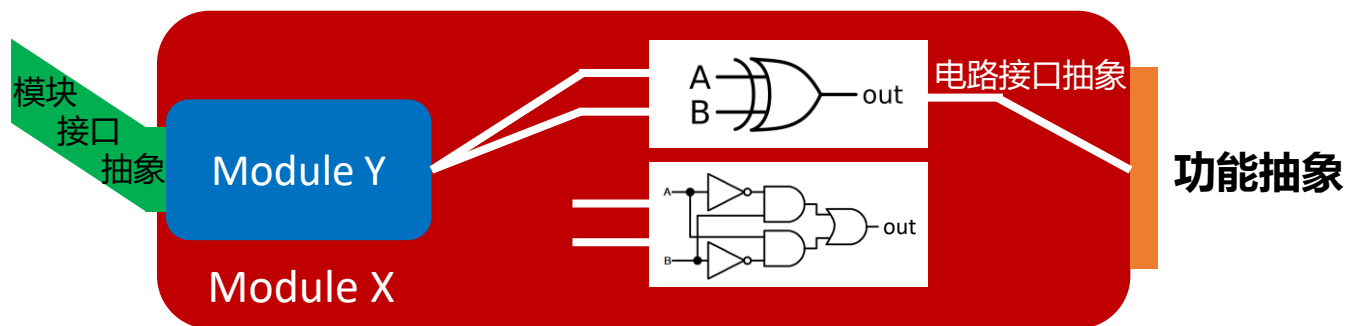
```
161 // multi-write  
162 val update_target = Reg(Vec(FtqSize, UInt(VAddrBits.W)))  
163 val cfiIndex_vec = Reg(Vec(FtqSize, ValidUndirectioned(UInt(log2Up(PredictWidth).W))))  
164 val cfiIsCall, cfiIsRet, cfiIsJalr, cfiIsRVC = Reg(Vec(FtqSize, Bool()))  
165 val mispredict_vec = Reg(Vec(FtqSize, Vec(PredictWidth, Bool())))  
166
```

- (3) 基于队列大小参数来例化数据存储

```
172 when(neal_fine) {  
173   val enqIdx = tailPtr.value  
174   commitStateQueue(enqIdx) := VecInit(io.enq.bits.valids.map(v => Mux(v, s_valid, s_invalid)))  
175   cfiIndex_vec(enqIdx) := io.enq.bits.cfiIndex  
176   cfiIsCall(enqIdx) := io.enq.bits.cfiIsCall  
177   cfiIsRet(enqIdx) := io.enq.bits.cfiIsRet  
178   cfiIsJalr(enqIdx) := io.enq.bits.cfiIsJalr  
179   cfiIsRVC(enqIdx) := io.enq.bits.cfiIsRVC  
180   mispredict_vec(enqIdx) := WireInit(VecInit(Seq.fill(PredictWidth)(false.B)))  
181   update_target(enqIdx) := io.enq.bits.target  
182 }
```

- (4) 使用参数化的队列指针完成对数据存储的操作

# 抽象层次③：具有功能抽象的参数化模块设计



## • 示例：一个参数化的发射队列

```
32 case class RSPParams
33 {
34   var numEntries: Int = 0,
35   var numEnq: Int = 0,
36   var numDeq: Int = 0,
37   var numSrc: Int = 0,
38   var dataBits: Int = 0,
39   var dataIdBits: Int = 0,
40   var numFastWakeup: Int = 0,
41   var numWakeup: Int = 0,
42   var hasFeedback: Boolean = false,
43   var delayedRf: Boolean = false,
44   var fixedLatency: Int = -1,
45   var checkWaitBit: Boolean = false,
46   var optBuf: Boolean = false,
47   // special cases
48   var isJump: Boolean = false,
49   var isAlu: Boolean = false,
50   var isStore: Boolean = false,
51   var isMul: Boolean = false,
52   var isLoad: Boolean = false,
53   var exuCfg: Option[ExuConfig] = None
54 }
55 def allWakeup: Int = numFastWakeup + numWakeup
56 def indexWidth: Int = log2Up(numEntries)
57 // oldestFirst: (Enable_or_not, Need_balance, Victim_index)
58 def oldestFirst: (Boolean, Boolean, Int) = (true, !isLoad, if (isLoad) 0
59 def hasMidState: Boolean = exuCfg.get == FmacExeUnitCfg
60 def needsScheduledBit: Boolean = hasFeedback || delayedRf || hasMidState
61 def needBalance: Boolean = exuCfg.get.needLoadBalance
```

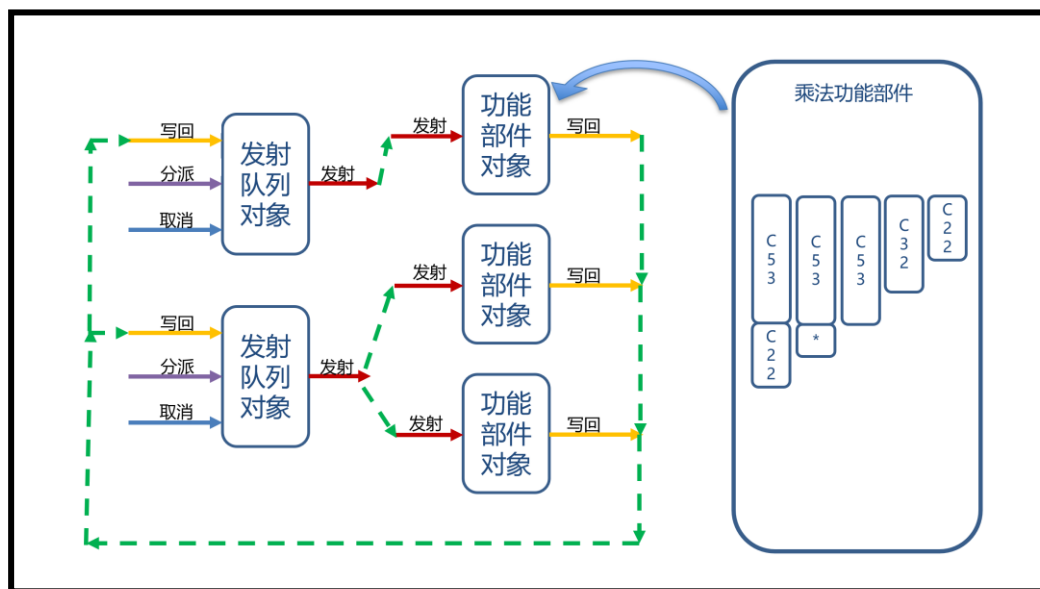
(1) 确定发射队列的参数需求  
并定义对应的参数类型

```
class StatusArray(params: RSPParams)(implicit p: Parameters) extends XModule
with HasCircularQueuePtrHelper {
  val io = IO(new Bundle {
    val redirect = Flipped(ValidIO(new Redirect))
    // current status
    val isValid = Output(UInt(params.numEntries.W))
    val canIssue = Output(UInt(params.numEntries.W))
    val flushed = Output(UInt(params.numEntries.W))
    // enqueue, dequeue, wakeup, flush
    val update = ec(params.numEnq, new StatusArrayUpdateIO(params))
    val wakeup = ec(params.allWakeup, Flipped(ValidIO(new MicroOp)))
    val wakeupMat = h = Vec(params.numEntries, Vec(params.numSrc, Output(UInt(params.allWakeup.W))))
    val issueGranted = Vec(params.numDeq, Flipped(ValidIO(UInt(params.numEntries.W))))
    val isFirstIssue = Vec(params.numDeq, Output(Bool()))
    val allSrcReady = Vec(params.numDeq, Output(Bool()))
    val updateMidstate = Input(UInt(params.numEntries.W))
    val deqRespWidth = if (params.hasFeedback) params.numDeq * 2 else params.numDeq
    val deqResp = Vec(deqRespWidth, Flipped(ValidIO(new Bundle {
      val rsMask = UInt(params.numEntries.W)
      val success = Bool()
      val respType = RSFeedbackType() // update credit if needs replay
      val dataInvalidSqrIdx = new SqrPtr
    })))
    val stIssuePtr = if (params.checkWaitBit) Input(new SqrPtr()) else null
    val memWaitUpdateReq = if (params.checkWaitBit) Flipped(new MemWaitUpdateReq) else null
  })
  val statusArray = Reg(Vec(params.numEntries, new StatusEntry(params)))
```

(2) 完成模块接口、基础硬件结构、  
内部逻辑的参数化实现

# 抽象层次④：模块对象化

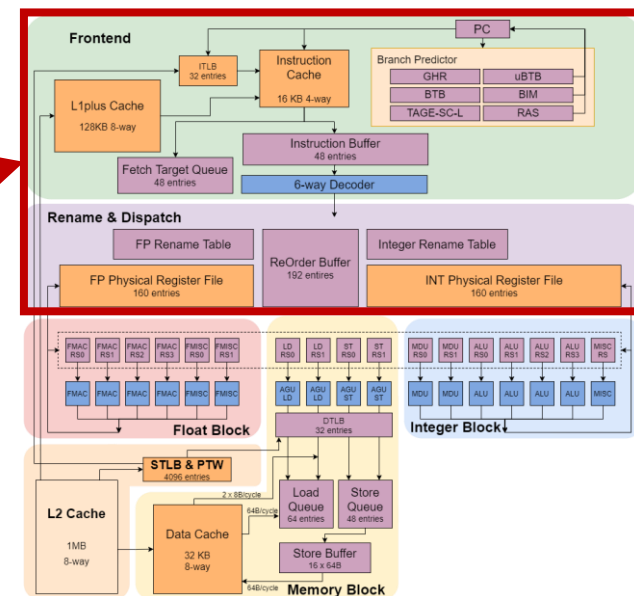
- 确定对象间交互方式之后，使用对象化的方式完成模块间组织



示例：发射队列、功能部件的对象化组织

```
23 // Synthesizable minimal XiangShan
24 // * It is still an out-of-order, super-scalaer arch
25 // * L1 cache included
26 // * L2 cache NOT included
27 // * L3 cache included
28 class MinimalConfig(n: Int = 1) extends Config(
29   new DefaultConfig(n).alter((site, here, up) => {
30     case SoCParamsKey => up(SoCParamsKey).copy(
31       cores = up(SoCParamsKey).cores.map(_copy(
32         DecodeWidth = 2,
33         RenameWidth = 2,
34         FetchWidth = 4,
35         IssQueSize = 8,
36         NRPhyRegs = 80,
37         LoadQueueSize = 16,
38         StoreQueueSize = 16,
39         RqSize = 32,
40         BrqSize = 8,
41         FtqSize = 16,
42         IBufSize = 16,
43         StoreBufferSize = 4,
44         StoreBufferThreshold = 3,
45         dpParams = DispatchParameters(
46           IntDqSize = 8,
47           FpDqSize = 8,
48           LsDqSize = 8,
```

All affected  
But it works



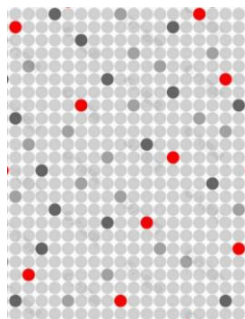
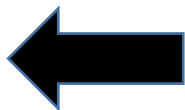
示例：将香山从六发射改为双发射

# 处理器敏捷设计范式



处理器

面向电路的  
处理器设计



像素  
(门电路)

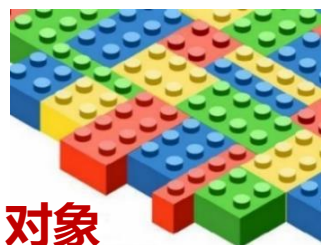
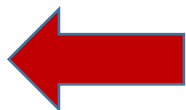
## 面向电路

- 以门电路作为基本设计单元
- 可模块化但接口混乱，不易复用



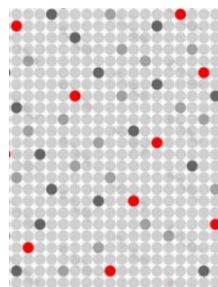
处理器

面向对象的  
处理器设计



对象  
(可复用的模块)

复杂度  
降低



## 面向对象

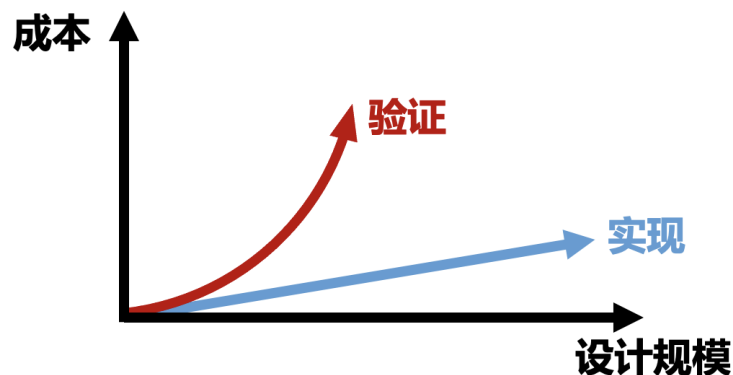
- 以对象作为基本设计单元
- 可模块化且接口标准，易复用

# 🏔️ 高性能处理器的验证方法

- 验证是处理器开发流程中耗时最长的一个阶段
  - 香山第一版9月启动Linux，次年4月基本完成验证工作，7月流片

- 随着设计效率的提高，验证将成为处理器开发的瓶颈

$$\text{Amdahl's Law: } S_{\text{latency}}(s) = \frac{1}{(1-p) + \frac{p}{s}} < \frac{1}{1-p}$$



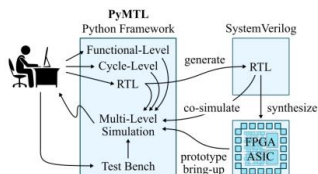
- 敏捷验证：使用更低的（人力、金钱、时间等）成本完成验证工作



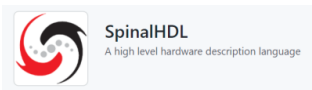
# 敏捷设计范式背景下的敏捷验证

敏捷工具支持下的  
敏捷开发范式

CHISEL



bluespec



cucapra/latte21

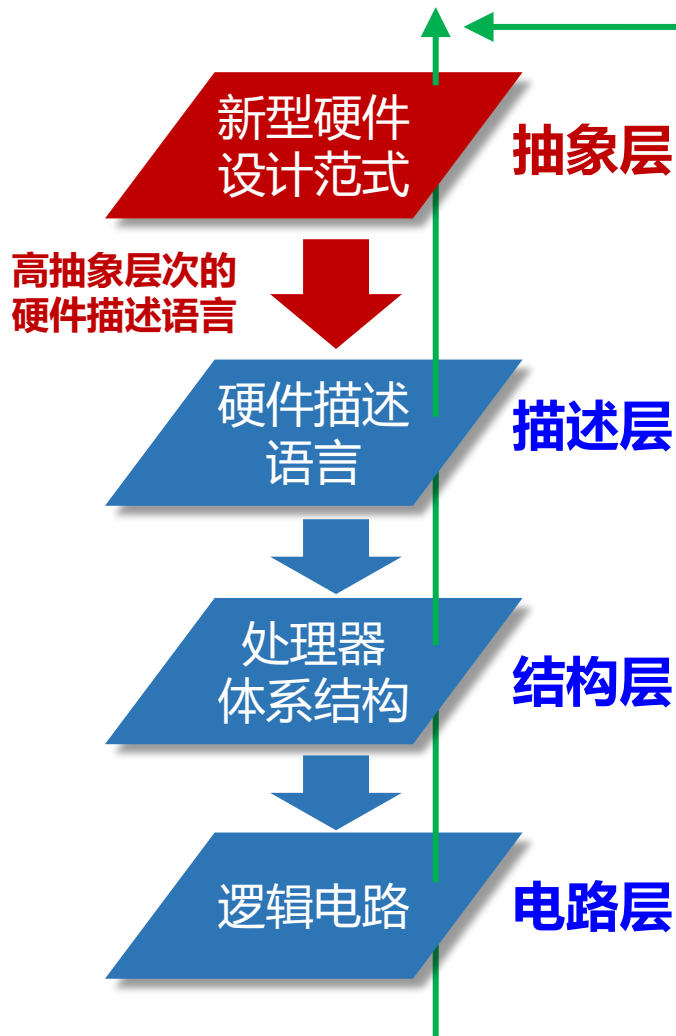
Languages, Tools, and Techniques for Accelerator Design



Verilog



敏捷硬件设计语言



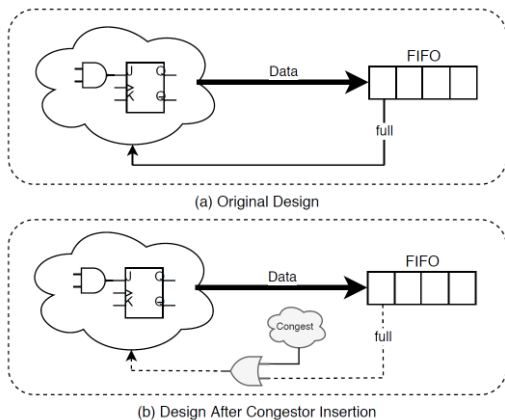
敏捷硬件设计方法



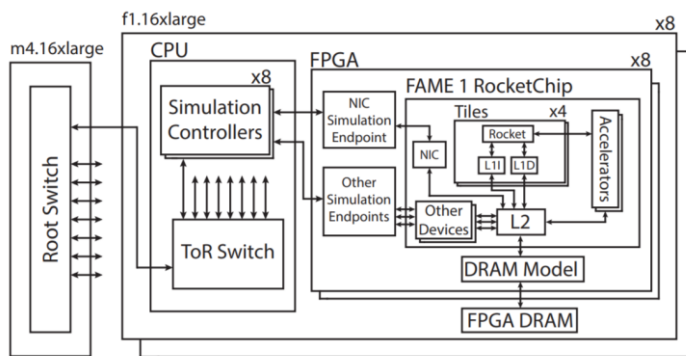
敏捷硬件验证方法

# 处理器敏捷验证方法

- 处理器的验证至少包括以下两个方面：
  - 功能验证：处理器的功能是否满足ISA定义的规范要求
  - 性能验证：处理器的性能是否达到了预期的性能指标
- 已有一些工作关注处理器敏捷验证方法

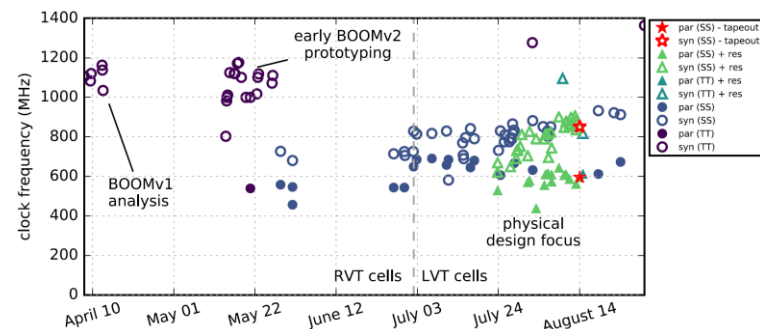


Dromajo[MICRO'21]: 通过Fuzzy Logic 设置随机事件以提高验证覆盖率



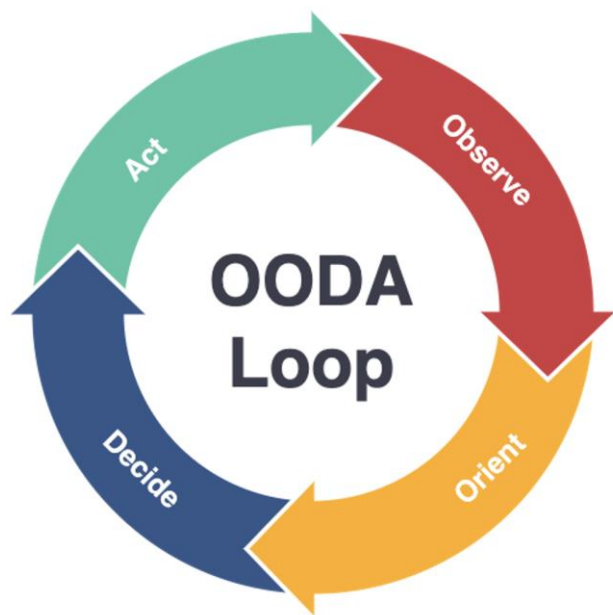
FireSim[ISCA'18]: 基于公有FPGA云 实现互联总线级大规模设计验证

## 目前香山前端团队 主要关注的内容



BROOM[HotChips'18]: 通过参数化 缩小结构参数完成关键路径评估

# 敏捷验证：快速完成验证迭代的循环



## 观察 Observe

判断目前的设计是否存在问题

## 定位 Orient

确定目前的设计存在什么样的问题

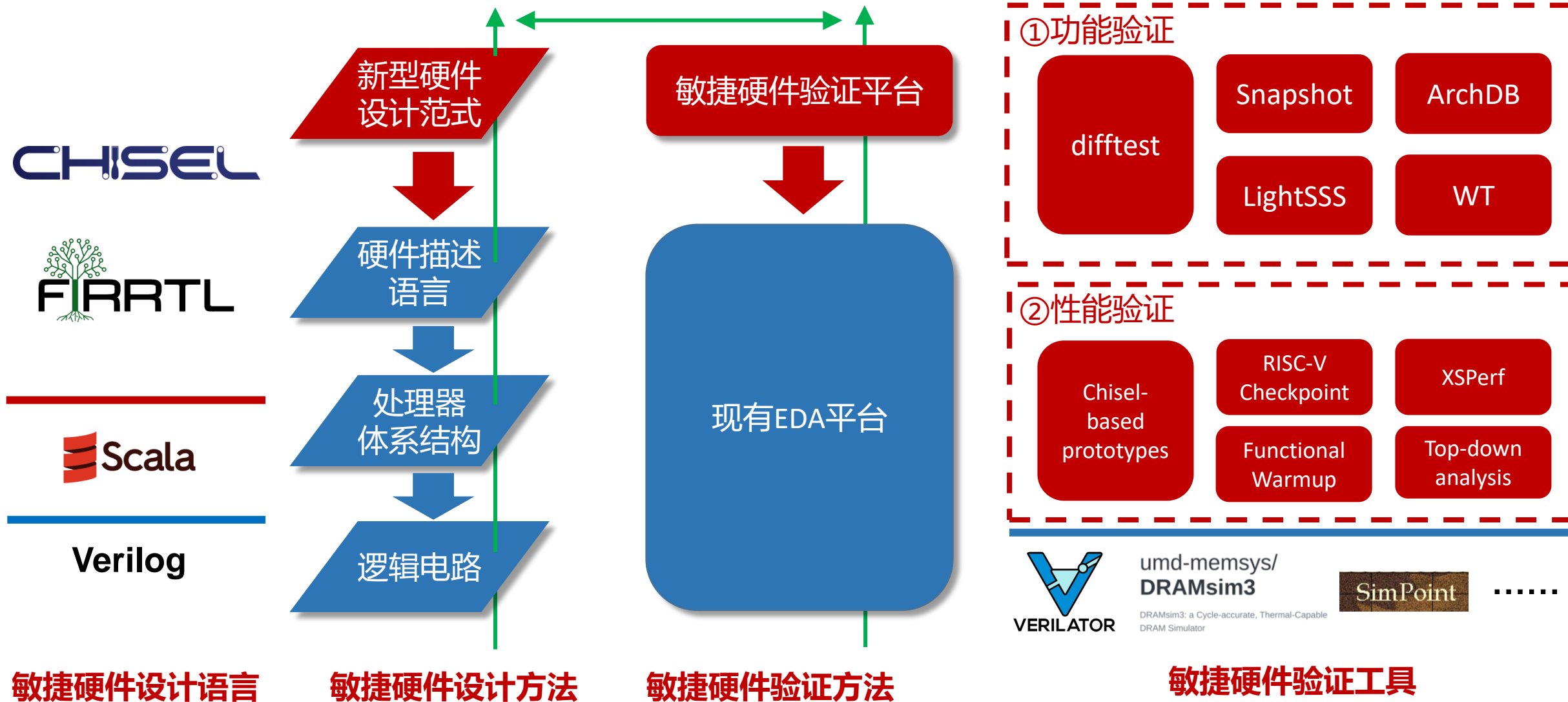
## 决策 Decide

决定如何解决目前的设计存在的问题

## 行动 Act

应用已知的解决方案至目前的设计中

# 建立匹配敏捷设计工具的敏捷验证平台



# ①敏捷功能验证方法



发现错误

保存出错现场

运行回归测试

定位并解决问题

Snapshot

LightSSS

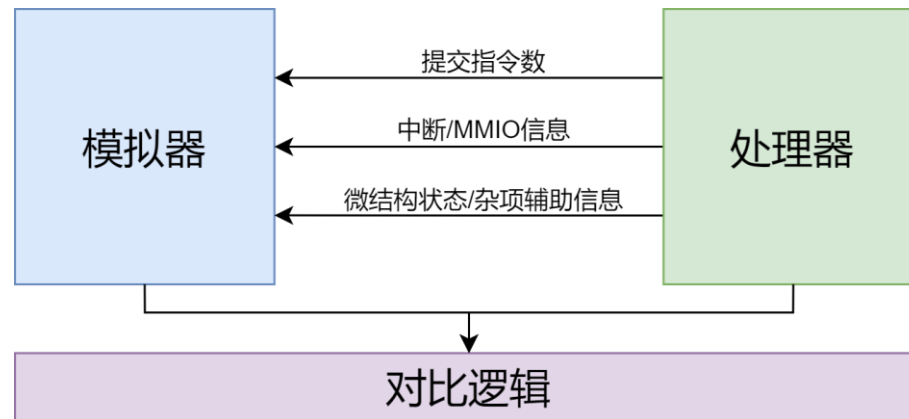
ArchDB

WT

# DiffTest: 指令级在线差分验证框架

## • 基本流程

- 处理器仿真产生指令提交/其他状态更新
- 模拟器执行相同的指令
- 比较两者状态
- 单步结果：报错或继续



基本验证框架

## • 提供Chisel和Verilog API

- 支持Verilator/VCS等RTL仿真器、NEMU模拟器
- SMP-DiffTest: 支持 SMP 结构的全系统仿真
  - 支持多线程程序、SMP Linux 内核等负载
  - 支持检测Cache一致性、内存一致性方面的软硬件问题

```
while (1) {  
    icnt = cpu_step();  
    nemu_step(icnt);  
    r1s = cpu_getregs();  
    r2s = nemu_getregs();  
    if (r1s != r2s) { abort(); }  
}
```

在线对比机制



# 如何更快地获得出错现场信息

方法1

带调试信息仿真

问题①：无法预知是否会出错——产生大量无用的调试信息

方法2

不带调试信息仿真



不带调试信息仿真

带调试  
信息仿真

问题②：无法预知什么时候出错——出错位置越晚，复现成本越高

软件容错技术：快照+恢复

不带调试信息仿真



带调试  
信息仿真

新方法

定期对仿真进行快照，并在出错时恢复相应快照进行调试

实现方案：应用级快照；系统级快照

# 快照方案1：应用级快照

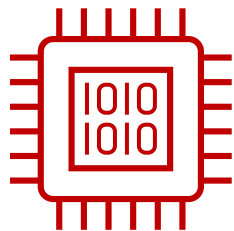
- **SSS**：由程序代码直接指定快照操作需要保存的变量和数据

RTL仿真电路状态

NEMU状态

非RTL外设状态

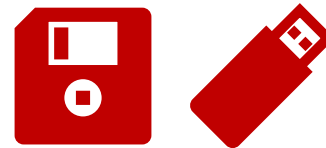
数据内容



所有电路信号值

Register	ABI Name	Description	Saver
x0	zero	Hard-wired zero	—
x1	ra	Return address	Caller
x2	sp	Stack pointer	Callee
x3	gp	Global pointer	—
x4	tp	Thread pointer	—
x5	t0	Temporary/alternate link register	Caller
x6-7	t1-2	Temporaries	Caller
x8	a0/fp	Saved register/frame pointer	Callee
x9	s1	Saved register	Callee
x10-11	a0-1	Function arguments/return values	Caller
x12-17	a2-7	Function arguments	Caller
x18-27	s2-11	Saved registers	Callee
x28-31	t3-6	Temporaries	Caller
f0-7	ft0-7	FP temporaries	Caller
f8-9	fs0-1	FP saved registers	Callee
f10-11	fs0-1	FP arguments/return values	Caller
f12-17	fs2-7	FP arguments	Caller
f18-27	fs2-11	FP saved registers	Callee
f28-31	ft8-11	FP temporaries	Caller

体系结构寄存器、  
内存等



状态、数据等

实现方式

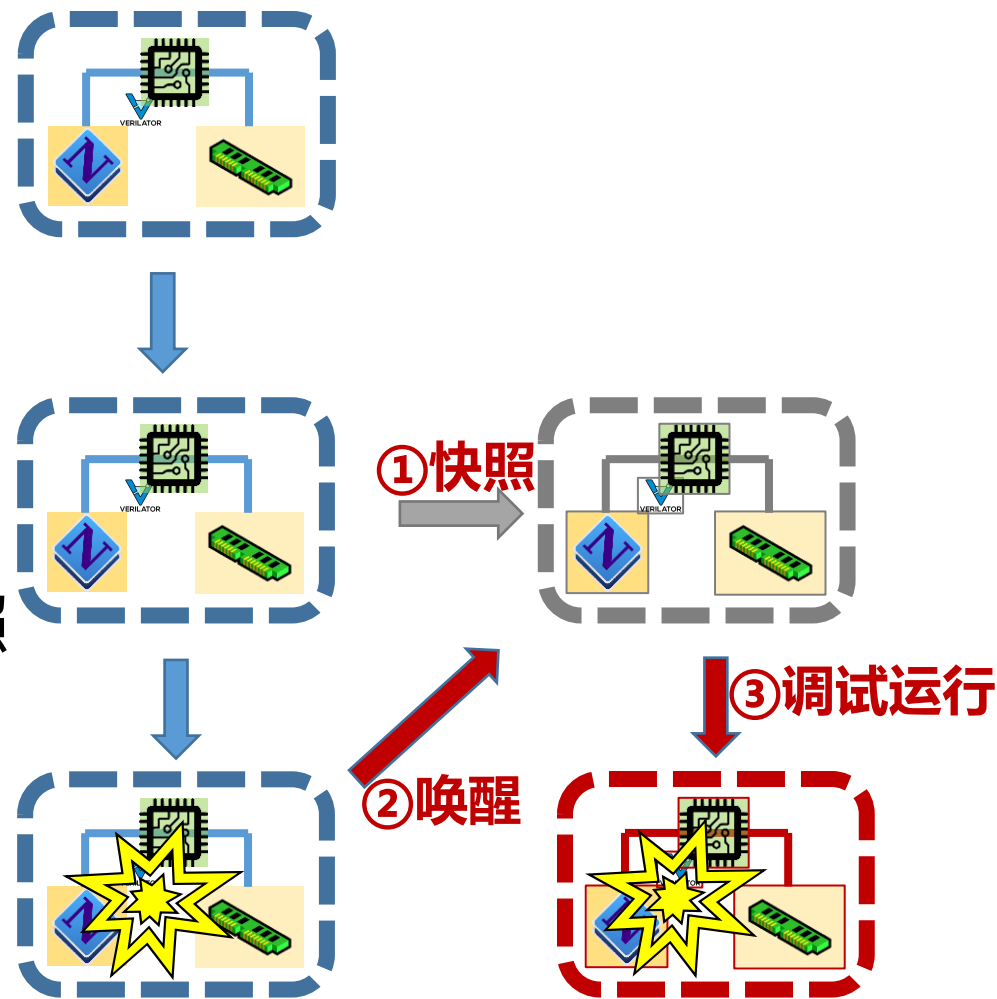


逐个保存

逐个保存

# 🚀 快照方案2：系统级快照

- **LightSSS: Light-weight Simulation SnapShot**
  - 一种轻量级仿真快照方法
- **关键技术：用fork对进程状态做快照**
  - 由OS提供写时复制技术（Copy On Write, COW）
- **优势1：可扩展性好，支持对任意外部模型的快照**
  - 无需理解外部模型的内部状态
- **优势2：快照效率高，LightSSS v.s. SSS**
  - Verilator编译：3.8X
  - g++编译：2.8X
  - 单次快照：6951.4X



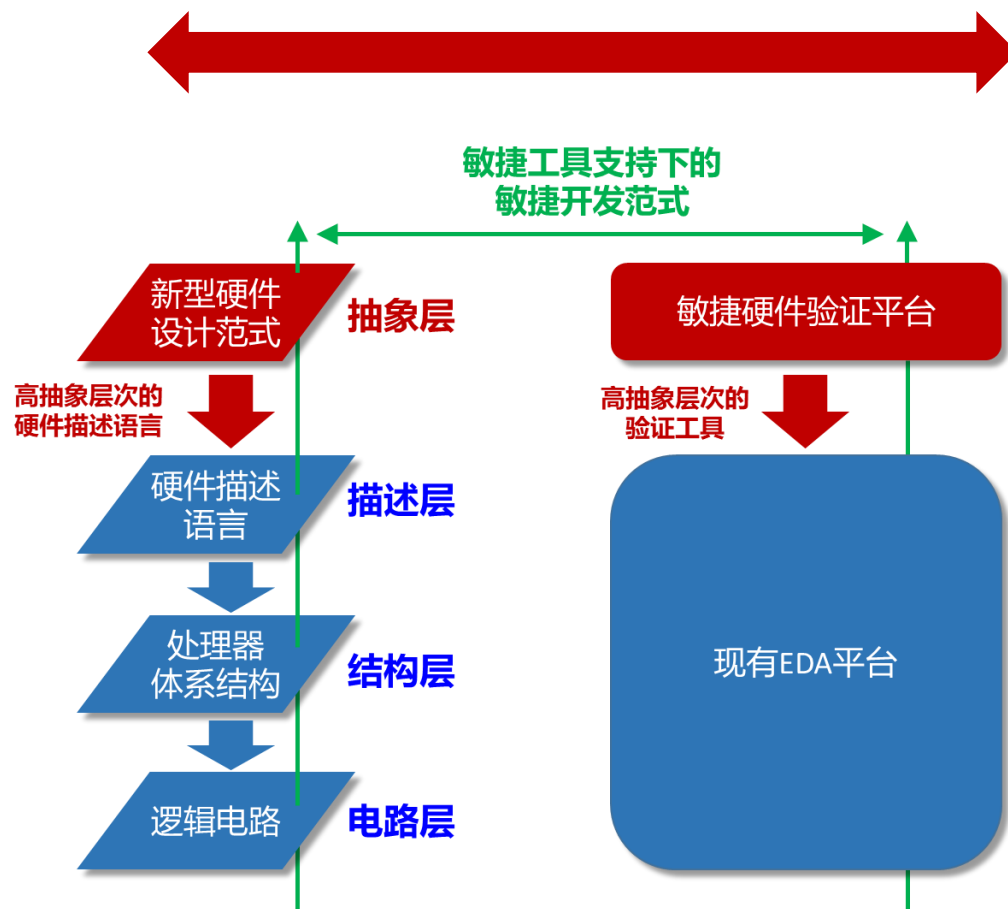
# 构建敏捷设计与验证工具的闭环

CHISEL

FIRRTL

Scala

Verilog



支持硬件敏捷设计范式的  
敏捷调试手段：  
Verilog与Chisel之间的桥梁

错误现场

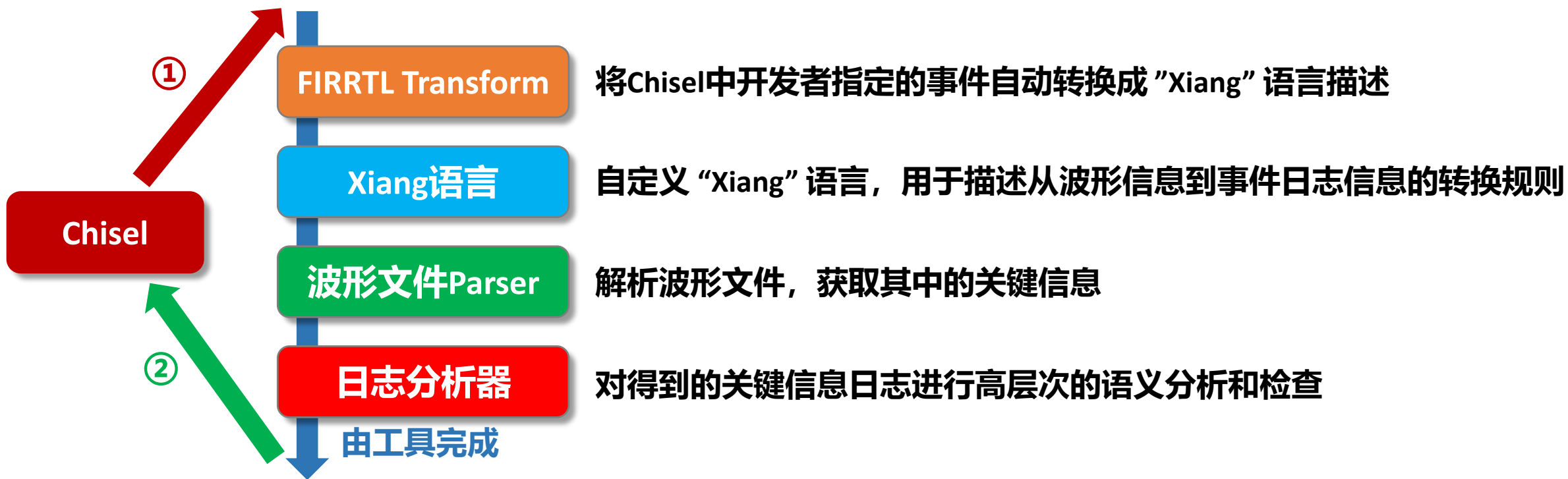
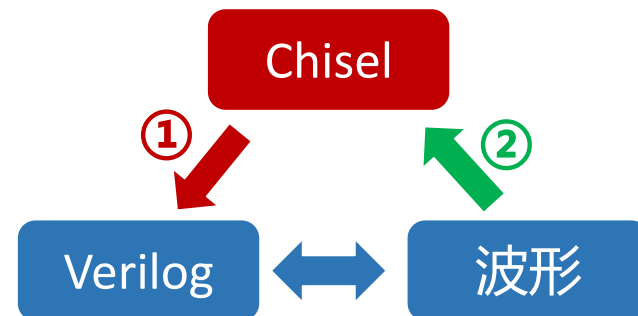
错误检查

VERILATOR

仿真波形文件

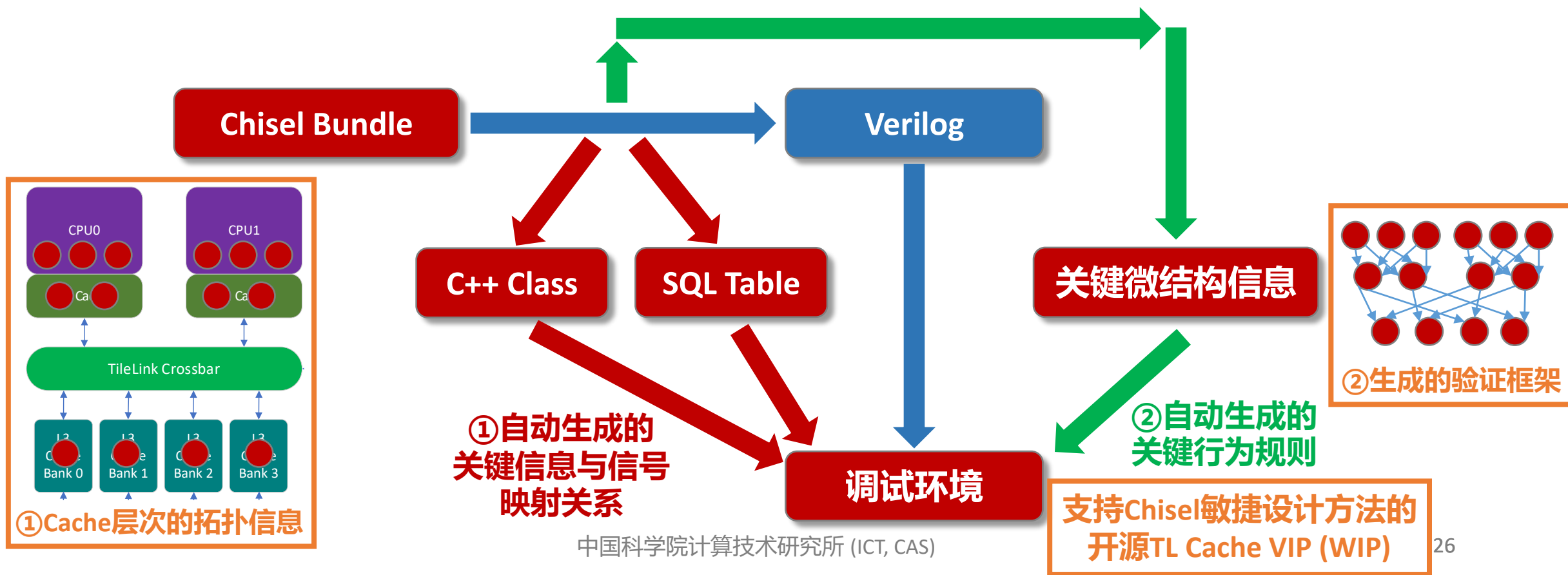
# 🏔️ Waveform Terminator: 敏捷硬件调试栈

- 核心目标：从波形中恢复被丢失的抽象语义信息
- 案例：基于高层次事件的调试方案



# ArchDB: 敏捷硬件调试工具

- 关键：使用软件工程的方法，提高①②的自动化程度
- 案例：针对关键微结构信息的调试方案（WIP）







## ②敏捷性能验证方法



# 基于Chisel的设计空间探索方法

- **解决了准确度的问题**：基于RTL代码的性能验证，与真实硬件表现完全一致
- **解决了重复工作的问题**：不再需要完成模拟器与硬件代码的对齐工作
- **带来了另一个问题：如何保证性能验证的效率？**
  - 敏捷设计方法保证了实现的效率，但仍然有大量的验证工作需要完成

	Verilog/SV	Cycle-approximate Simulator	Cycle-accurate Simulator	Chisel
准确程度	高	较高	低	高
实现速度	慢	较快	快	快
评估速度	慢	较快	快	快？慢？
分析难度	高	较低	低	高？低？
对齐成本	低	较高	高	低

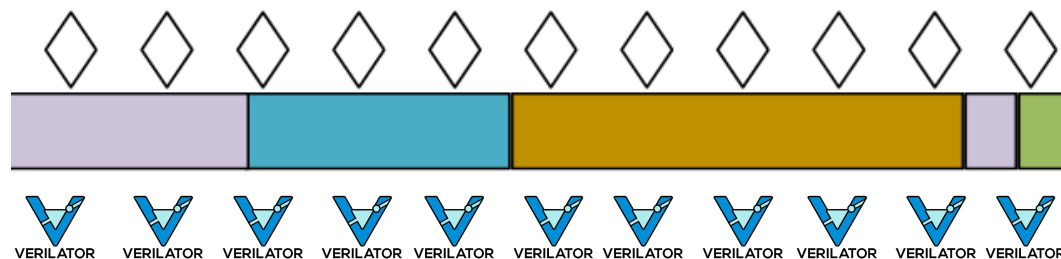
# 基于仿真的敏捷性能验证方法

- **敏捷验证**：使用**更低的（人力、金钱、时间等）成本**完成验证工作
- 然而，**传统的性能验证方式在高性能处理器场景下成本极高**

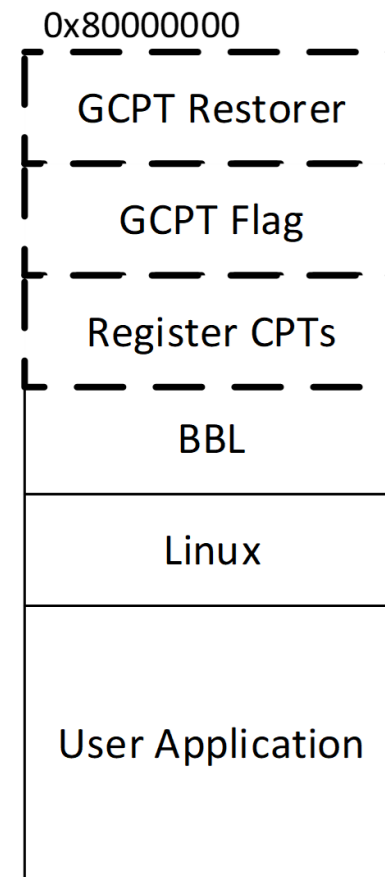


# 通过RISC-V Checkpoint提高仿真并行度

- 出发点：将程序切成小片段，片段间可以实现并行仿真



- 难点①：对程序在某一时刻的状态进行快照
  - 基于指令集模拟器完成，与Snapshot-NEMU类似
- 难点②：在香山仿真时完成快照的恢复
  - 关键问题：RTL并不确定，需要一种普适的方法完成状态恢复
  - 原理：通过特权指令实现各寄存器的初始化**
  - 已有ARM等指令集上的类似工作，需要针对RISC-V指令集进行设计



RISC-V Checkpoint  
基本格式

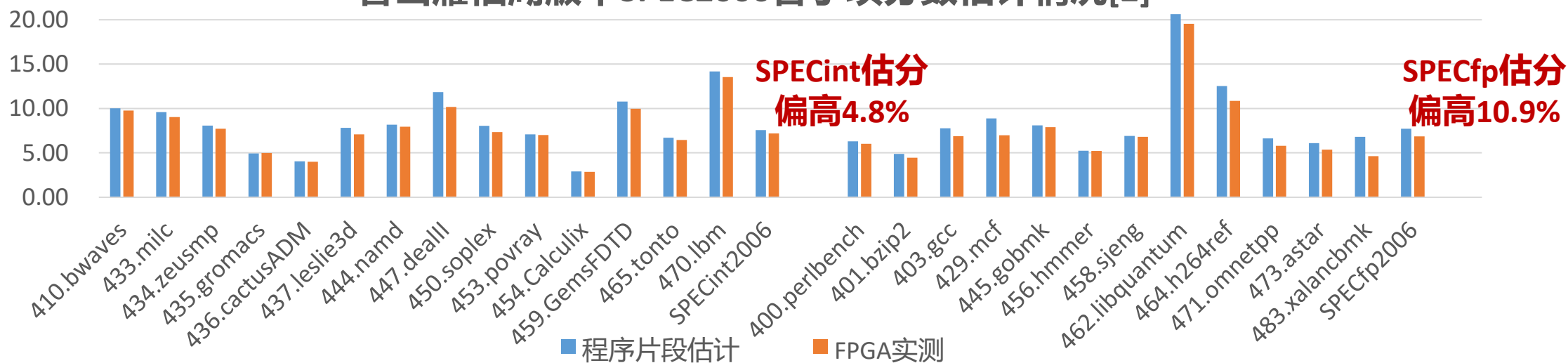
# 特征采样：进一步选出具有代表性的程序片段

- Simpoint[1]：针对程序基本块的聚类，从程序中选出具有代表性的片段



- 案例：结合片段IPC与聚类权重完成程序总体执行时间的估计

香山雁栖湖版本SPEC2006各子项分数估计情况[2]

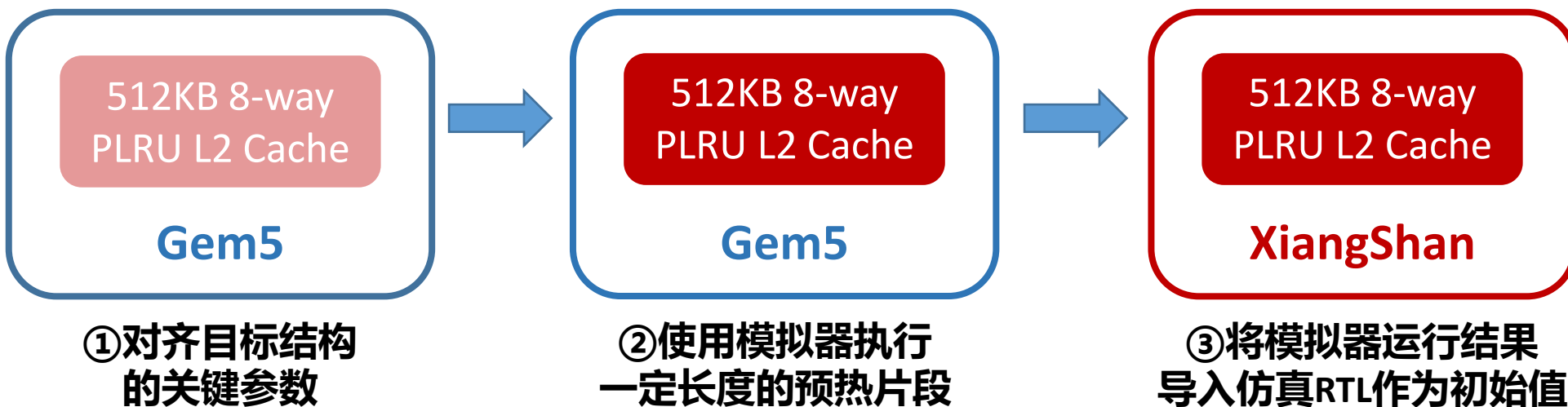


[1] Greg Hamerly, Erez Perelman, Jeremy Lau, and Brad Calder, SimPoint 3.0: Faster and More Flexible Program Analysis, Journal of Instruction Level Parallel, September 2005.

[2] 仿真和FPGA平台均使用90周期固定访存延迟, Gamess与wrf因当时存在的编译问题未包含在内。Simpoint设置Coverage 80%, maxK=30 中国科学院计算技术研究所 (ICT, CAS)

## 🔥 功能预热：进一步提高仿真效率

- 程序切片仿真存在需要**预热问题：分支预测、Cache等结构初始值是无效的**
- **效率低**：需要跑足够长的一个预热片段后，才能进入性能评估片段
  - 往往预热片段长度与性能评估片段长度在一个数量级
  - 导致**实际可达到的性能评估仿真速度只有仿真器理论速度的一半**
- **解决方案：基于模拟器完成结构的功能预热（WIP）**





# 支持敏捷设计语言的敏捷性能分析框架

- RTL仿真带来的优势：方便地获取**覆盖广泛的细节性能数据**
- **XSPerf：基于Chisel/Scala的性能计数器配置框架**
  - 支持参数化，匹配硬件敏捷设计范式需求
  - 自动化完成性能计数器的实例化、运行时开关等，提高代码复用程度
- **Top-down analysis：细粒度的性能分析工具**
  - 可视化展示处理器的性能瓶颈位置，指导设计空间探索

功能实现

并行仿真

针对性评估

性能分析

**我们的目标：在24小时内完成一个功能特性的性能验证**

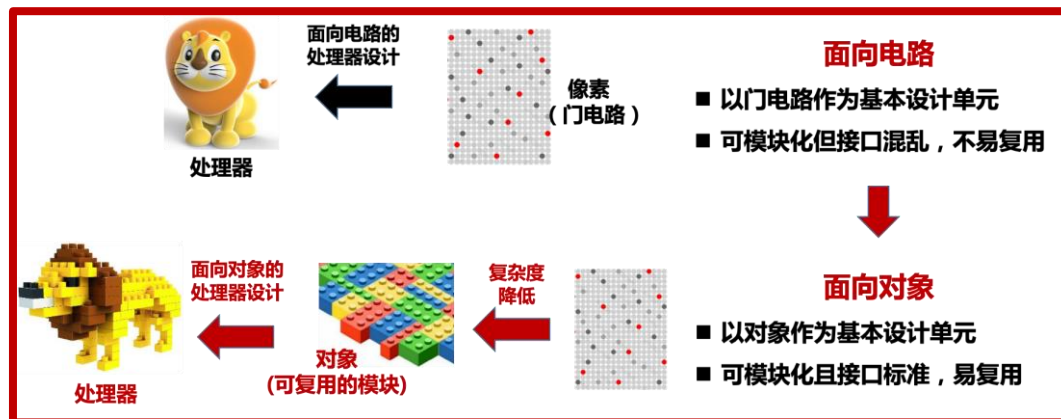
# 敏捷验证：进一步工作

- **问题①：**是否有可能通过EDA工具的优化进一步提高测试的运行速度？
  - 是否有可能采用测试之外的手段完成验证，如形式化方法？
- **问题②：**单次验证循环的效率得到了提升，如何进一步减少验证的循环次数？
  - 是否可以使用更少的测试集，达到相同的测试覆盖率？或达到相同的性能评估准确度？
- **问题③：**在敏捷设计的背景下，如何进一步优化敏捷物理验证流程？
  - 是否有可能进行敏捷频率、功耗、面积评估？

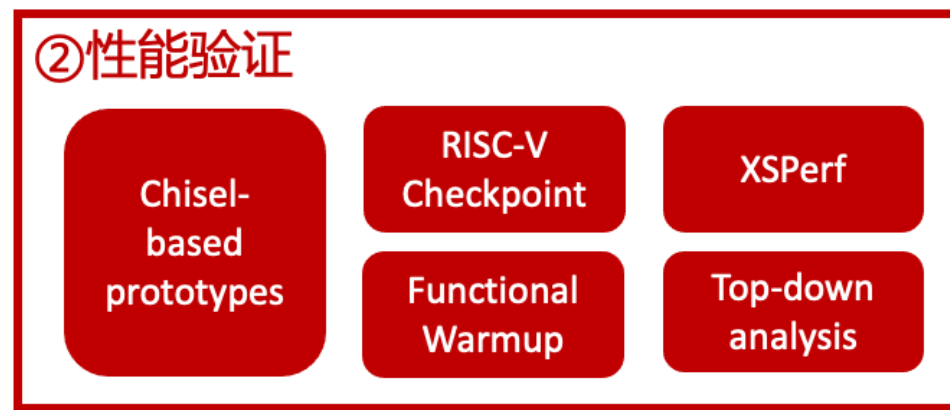
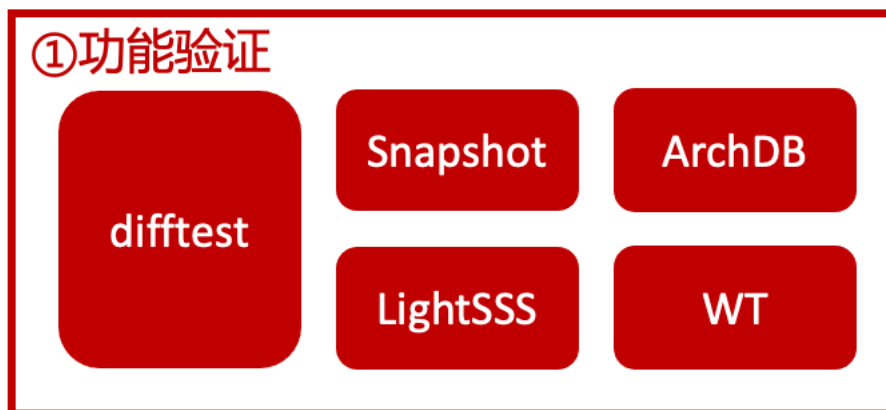


# 香山：面向高性能处理器的敏捷开发实践

- 敏捷设计：以Chisel为例实现对模块的功能抽象，探索面向对象的处理器设计



- 敏捷验证：针对敏捷设计范式搭建验证工具与平台，探索处理器敏捷验证方法



# 欢迎加入香山开源社区

- GitHub/Gitee/trustie/ihub 等开源代码托管平台
- 邮件列表: xiangshan-all@ict.ac.cn
- Gitter: OpenXiangShan/community
- 微信技术讨论群
- 微信公众号/知乎/微博技术研讨平台: **香山开源处理器**
  - ①技术类文章/点评: Hot Chips 等会议解读、香山技术分析
  - ②香山项目进展: 每双周香山开发进展同步
  - ③观点文章



微信搜一搜



香山开源处理器

**谢谢！**  
**请各位批评指教！**

