

计算机体系结构 (期末复习)

胡伟武、汪文祥

计算机系统结构基础

- 计算机的基本概念
 - 什么是计算机
 - 计算机的基本组成
 - 衡量计算机的因素
- 影响计算机设计的主要因素及其发展趋势
 - 摩尔定律和工艺的发展
 - 计算机市场发展趋势
 - 多核结构的发展趋势
- 计算机重要指标
 - 性能、成本、功耗
- 我国计算机发展历史

二进制与逻辑电路

- 计算机中数的表示
 - 补码和原码
- CMOS门电路及工艺
 - 电路原理与延迟模型
- CMOS组合逻辑与时序逻辑
 - 逻辑表达式、真值表、卡诺图、逻辑图
 - 常见逻辑电路：译码器、选择器、ALU电路
 - 触发器原理：三个主要的时序指标
- CMOS电路延迟
- 从Verilog到GDSII
 - Verilog代码的编写
 - 电路与版图
- 其它“0”和“1”表示方法

指令系统结构

- 指令系统结构的设计原则
 - 硬件和软件的界面
 - 设计要求：兼容性、通用性、高效性、安全性
- 影响指令系统结构的因素
 - 工艺、系统结构、操作系统、编译、应用等
- 指令系统的分类
 - 堆栈型指令、累加器型指令、寄存器型指令
- 指令系统的组成部分
 - 操作、操作数、编码
- RISC指令系统结构 *简化指令本身, 简化指令间互*
 - 操作码：常用的十种操作占指令执行的96%
 - 操作数：load-store结构，简单寻址方式用得最多
 - 编码：定长
- RISC系统结构的简史
 - CDC6600=>801, MIPS, RISC-2项目
- 不同RISC结构的比较
 - MIPS, Solaris, HP-PA, PowerPC
 - MIPS指令系统与操作系统的关系

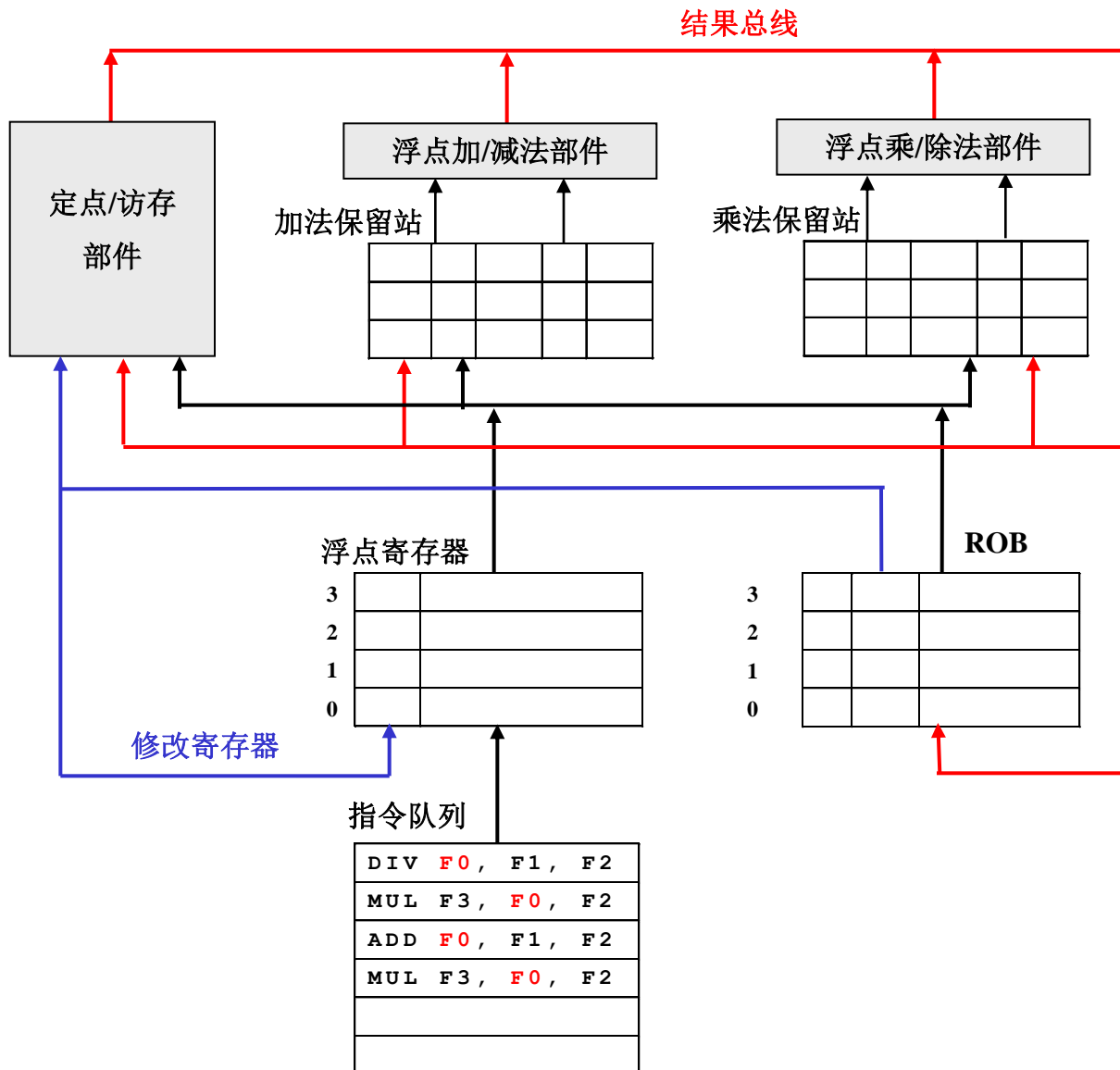
静态流水线

- 一个简单的**RISC CPU**
 - 指令系统、数据通路、控制逻辑、时序、时序优化==>流水线
- 指令流水线
 - 标准五级流水线: **IF, ID, EX, MEM, WB**
- 指令相关与流水冲突
 - 数据相关: **RAW、WAW、WAR**
 - 控制相关: 与PC有关的相关, 通过阻塞、**delay slot**、猜测解决相关
 - 结构相关: 资源冲突, 通过增加资源解决结构相关
- 流水线的前递技术
 - 缓解**RAW**相关引起流水线冲突
- 例外与流水线
 - 把例外信息保存下来到**WB**阶段再处理
- 多功能部件及多拍操作
 - **WAW**相关、**RAW**相关更突出, 没有**WAR**相关

动态调度流水线

- 影响流水线效率的因素
 - **Pipeline CPI=Ideal pipeline CPI + Structural stalls + RAW stalls + WAR stalls + WAW stalls + Control stalls**
- 指令调度技术
 - 软件循环展开与重命名的例子
- 动态调度技术：把译码分成发射和读操作数
 - 保留站的作用：有序进入、乱序发射
 - 名字依赖性与寄存器重命名：消除WAR和WAR
- Tomasulo算法
 - IBM360/91
- 动态流水线的精确例外处理
 - 增加提交流水级，修改机器状态
 - **ROB**的作用：有序退出

Tomasulo算法



多发射数据通路

- 乱序执行的关键技术
 - 保留栈、重命名寄存器、重排序缓存
 - 有序进入、乱序执行、有序退出
- 动态调度流水线数据通路
 - 保留站的组织
 - 寄存器与保留站的关系
 - 寄存器重命名方法
- 常见处理器的数据通路
- 多发射结构数据通路
 - 不仅要前后看，而且要左右看
- 龙芯2号多发射结构简介



转移预测

- 转移指令
 - 条件/无条件、直接/间接、绝对/相对
- 程序的转移行为
 - 很频繁、局部性、可预测性（重复性、相关性）
- 软件方法解决控制相关
 - 循环展开、软流水
- 硬件动态转移预测
 - 转移的重复性和偏向性：BHT、BTB
 - 转移指令的相关性问题：两层转移预测，Yeh和Patt分类
 - 分支别名干扰问题：Gshare等
 - 混合预测器：不同的分支预测只能对某类的分支行为有效
- 常见处理器的转移猜测
 - ALPHA处理器的转移猜测原理

功能部件

- 定点加法器
 - 先行进位加法器
- 定点ALU设计
 - 加减、逻辑、移位运算
- 定点补码乘法器的设计
 - Booth一位乘与二位乘算法
 - Wallace树的结构

高速缓存（Cache）

- 存储层次的基本概念
 - 程序访问的局部性：时间局部性和空间局部性
 - 越小越简单的硬件越快，越快的硬件越昂贵
- Cache结构 
 - Cache的内容是主存储器内容的一个子集
 - 同时存储数据和地址
 - 需要考虑所需要的数据不在Cache中的情况
- Cache性能优化 
 - 降低失效率（MissRate）
 - 降低失效延迟（MissPenalty）
 - 降低命中延迟（HitTime）
 - 提高Cache访问并行性
- 常见处理器的存储层次
 - 运用之妙、存乎一心

存储管理

- 虚拟存储的基本原理
 - TLB的原理和逻辑
- MIPS处理器对虚存系统的支持
 - TLB、控制寄存器、专用指令、专用例外
- LINUX操作系统的存储管理
 - TLB、页表和内存分配过程
 - 例：

```
array=(int*)malloc(0x1000);  
for (i=0;i<1024;i++) array[i] = 0;
```
- TLB的性能分析和优化
 - 提高TLB覆盖空间的大小（多页支持）
 - 提高TLB异常处理的效率（软TLB等缓存技术）

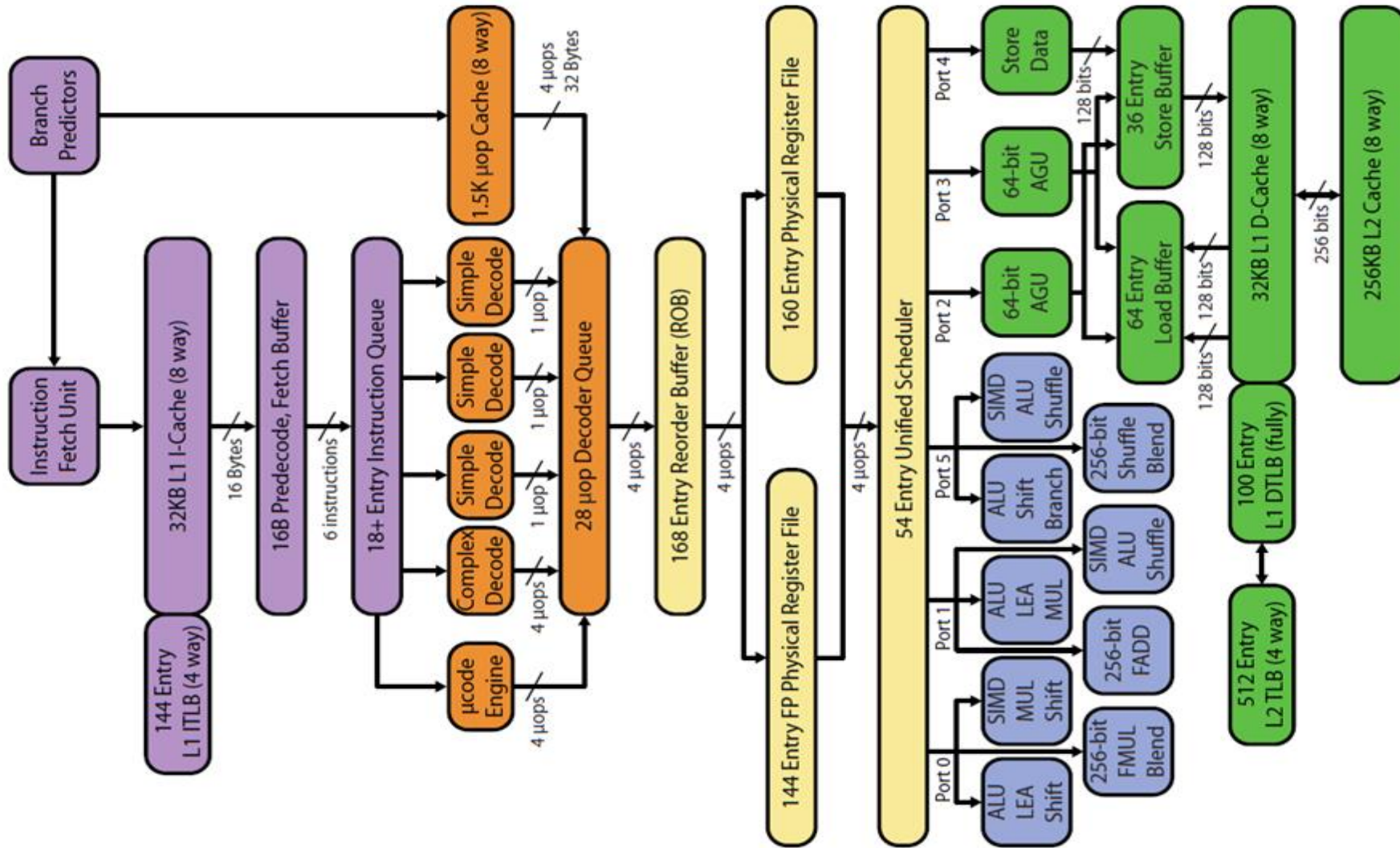
多处理器

- 消息传递与共享存储
 - 可伸缩性和通用性
 - 常见的共享存储系统：SMP、CC-NUMA、NCC-NUMA、SVM
- 共享存储系统的指令相关
 - 程序PRG的执行 $E(\text{PRG})$ 正确的充要条件是 $E(\text{PRG}) \cup \text{PO}(\text{PRG})$ 无圈
- 共享存储系统的访存事件次序
 - 任一访存操作u被分割成N个子操作，N为处理器个数
- 存储一致性模型
 - 结构设计者与应用程序员之间的一种约定，程序的正确性标准
 - 弱一致性模型：同步操作在共享存储系统中的作用
- CACHE一致性协议
 - 一种把新写的值传播到其他处理机的机制

什么是CPU

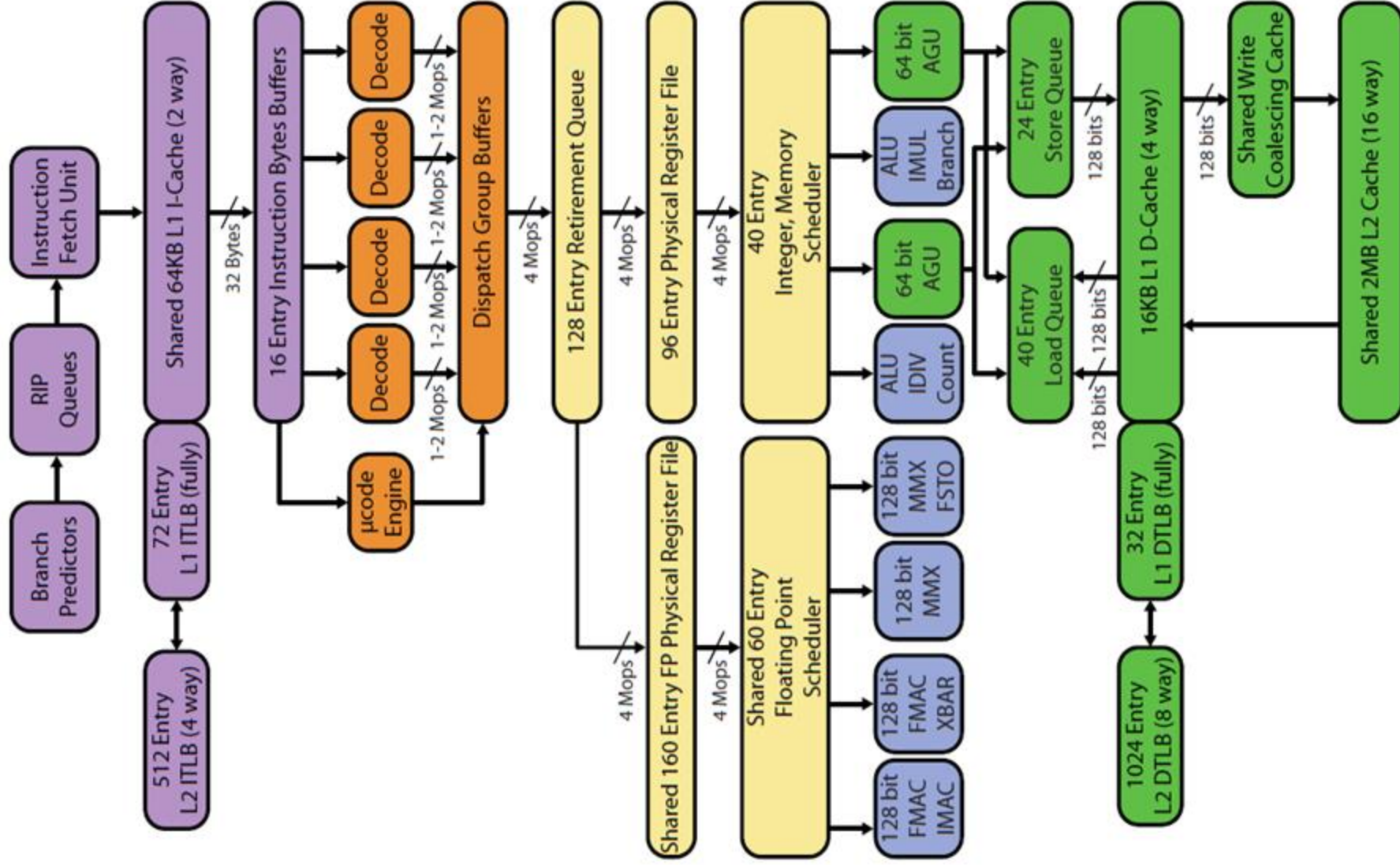
- 一个6岁小孩的答案
 - CPU就是在一张纸上画一些方块，
 - 然后用线和箭头把这些方块连起来，
 - 再写上几个字，涂上点颜色，
 - 最后一烧，烧出一个亮晶晶的小方块
- 我们比她多知道些什么？

Sandy Bridge

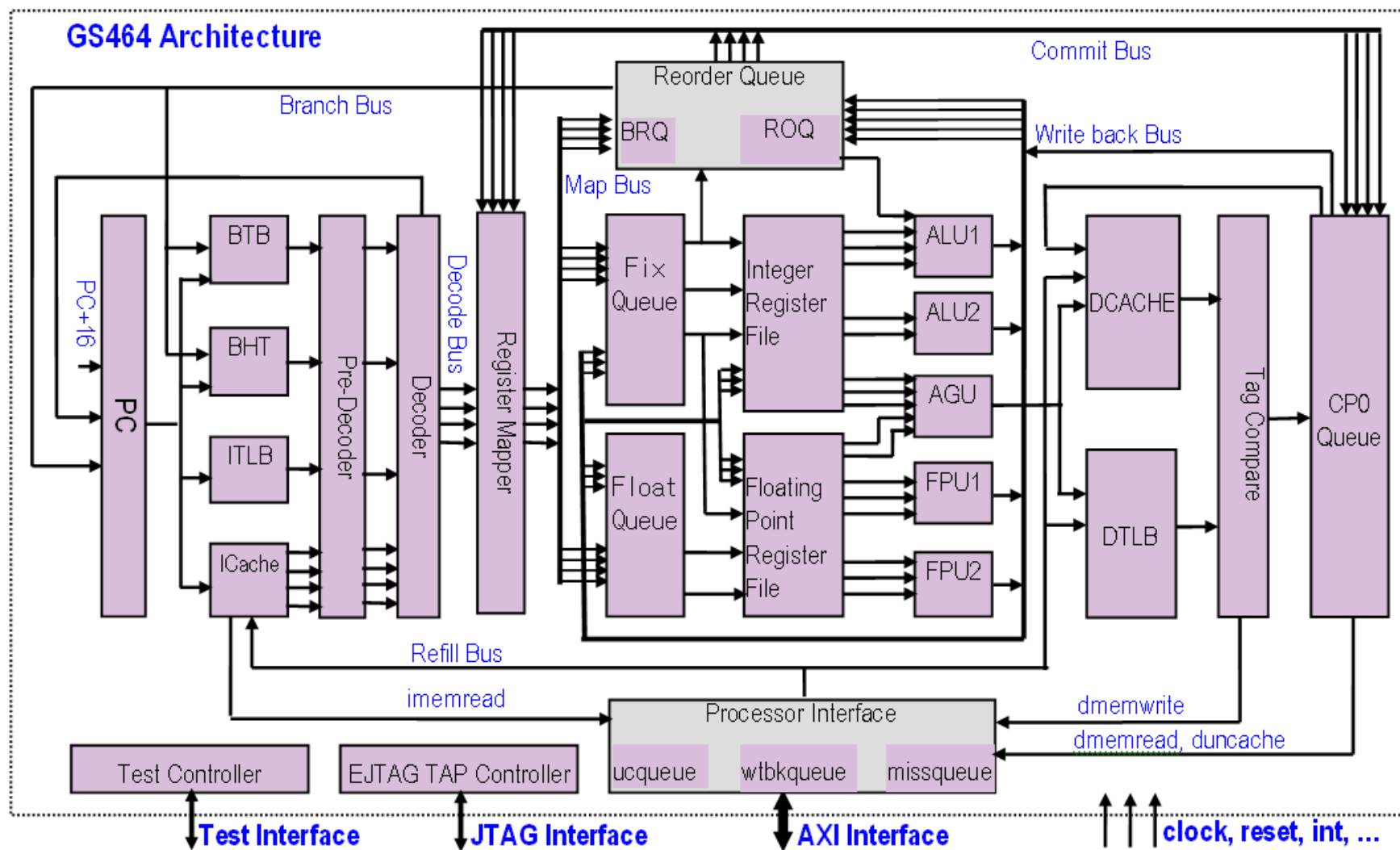


AMD的Bulldozer

Bulldozer



龙芯的GS464



什么是计算机体系结构

- 计算机体系结构（**ComputerArchitecture**）是描述计算机各组成部分及其相互关系的一组规则、方法和机制，是程序员所看到的计算机的属性，即概念性结构与功能特性。
 - 计算机体系结构主要研究内容包括指令系统结构（**Instruction Set Architecture**）和计算机组织，微体系结构（**Microarchitecture**）是处理器级别的计算机组织。
 - 高性能、低功耗、低成本是计算机体系结构的主要设计目标。
 - 冯诺依曼结构的存储程序和指令顺序执行是现代计算机体系结构的基础。

计算机体系结构的表达方式

- 计算机体系结构可以有不同层次和形式的表现方式。
 - 计算机体系结构通常用结构框图来表示，结构框图中的方块表示计算机的组成部分（如运算器、寄存器、存储器、输入/输出设备等），线条和箭头表示指令和数据在功能模块中的流动，结构框图可以不断分解一直到门级或晶体管级。
 - 计算机体系结构也可以用高级语言如C语言来表示，形成体系结构模拟器，用于对体系结构进行性能评估和分析。
 - 以Verilog为代表的硬件描述语言描述的体系结构可以通过EDA仿真工具进行功能验证和性能分析，并可以通过EDA布局布线工具转换成门级以及晶体管级网表并最终转换成版图，用于芯片制造。

关于期末考试

- 主要是考能力、考认识，不考知识
- 考试时间：2019年1月14日晚
- 考试方式：课堂闭卷
 - 11道解答题，2道5分，9道10分
 - 70%以上题目来自作业
- 期末成绩占总成绩的67%

谢 谢