

## 龙芯 2 号处理器设计和性能分析

胡伟武 张福新 李祖松

(中国科学院计算技术研究所计算机系统结构重点实验室 北京 100080)

([hww@ict.ac.cn](mailto:hww@ict.ac.cn))

## Design and Performance Analysis of the Godson-2 Processor

Hu Weiwu, Zhang Fuxin, and Li Zusong

(Key Laboratory of Computer System and Architecture, Institute of Computing Technology, Chinese Academy of Sciences, Beijing 100080)

**Abstract** In this paper, the design and the result of performance analysis of the Godson-2 processor are presented. The Godson-2 implements a 4-way superscalar pipelined architecture, contains two 64KB L1 caches for instruction and data, and supports up to 8MB off-chip L2 cache. To improve the pipeline efficiency, The Godson-2 utilizes out-of-order executing technologies such as advanced branch prediction unit, register renaming and dynamic scheduler, and dynamic memory access mechanism like non-blocking cache and load speculation. The Godson-2 is implemented on 0.18 $\mu$ m CMOS technology, with a maximum frequency of 500MHz under normal voltage and consumes 3-5 watts power under that frequency. The Godson-2 can perform one billion double-precision floating-point operations per second (two billion for single-precision), and the overall performance is comparable to Intel Pentium III with similar frequency. Presently a full Linux distribution (Debian) is running well on the Godson-2 prototype machines, including important desktop applications such as Mozilla web browsers, media players and OpenOffice.

**Key words** superscalar pipeline; out-of-order execution; branch prediction; register renaming; dynamical scheduling; non-blocking cache; load speculation; performance analysis

**摘要** 介绍龙芯 2 号处理器设计及其性能测试结果。龙芯 2 号采用四发射超标量流水结构,片内一级指令和数据高速缓存各 64KB,片外二级高速缓存最多可达 8MB。为了充分发挥流水线的效率,龙芯 2 号实现了先进的转移猜测、寄存器重命名、动态调度等乱序执行技术以及非阻塞的 Cache 访问和 load Speculation 等动态存储访问机制。龙芯 2 号处理器采用 0.18 $\mu$ m 的 CMOS 工艺实现,在正常电压下的最高工作频率为 500MHz,500MHz 时的实测功耗为 3~5W。龙芯 2 号单精度峰值浮点运算速度为 20 亿次/秒,双精度浮点运算速度为 10 亿次/秒,SPEC CPU2000 的实测性能是龙芯 1 号的 8~10 倍,综合性能已经达到 Pentium III 的水平。目前芯片样机能流畅运行完整的 64 位中文 Linux 操作系统,全功能的 Mozilla 浏览器、多媒体播放器和 OpenOffice 办公套件,可以满足绝大多数桌面应用的要求。

**关键词** 超标量流水线;乱序执行;转移猜测;寄存器重命名;动态调度;非阻塞的 Cache;load 指令猜测执行;性能分析

中图法分类号 TP302

收稿日期:2005-07-11;修回日期:2006-01-04

**基金项目:**国家自然科学基金杰出青年基金项目(60325205);国家“八六三”高技术研究发展计划重点基金项目(2002AA110010,2005AA110010,2005AA119020);国家“九七三”重点基础研究发展规划基金项目(2005CB321600);中国科学院计算技术研究所基础研究基金项目(20056020);中国科学院计算技术研究所知识创新课题(20056240)

## 1 引言

龙芯 2 号处理器是一款实现 64 位类 MIPS 指

令集<sup>[1]</sup>的通用 RISC 处理器. 龙芯 2 号处理器的微体系结构如图 1 所示.

龙芯 2 号采用乱序执行和激进的 Cache 设计<sup>[2]</sup>来提高流水线的效率.

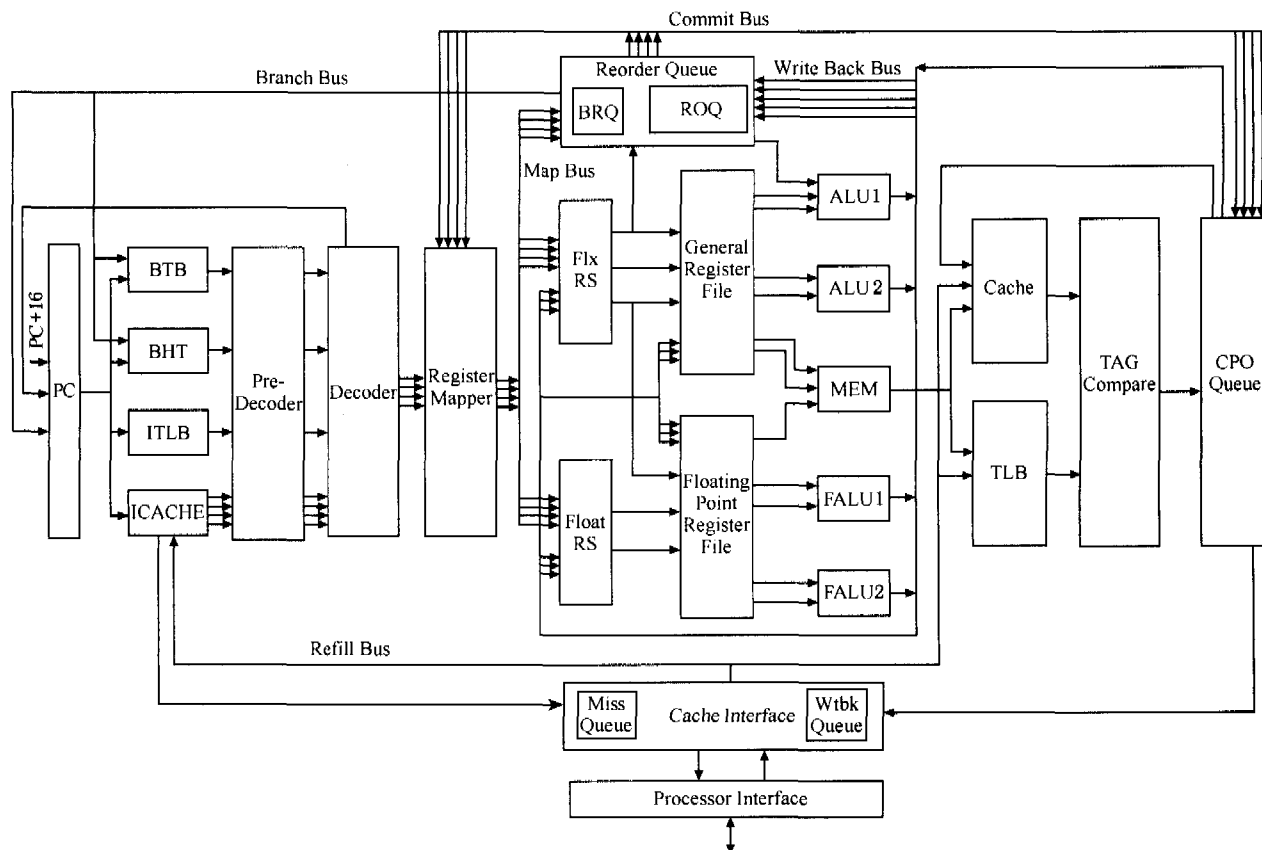


Fig. 1 Godson-2 architecture.

图 1 龙芯 2 号体系结构

乱序执行技术包括寄存器重命名、动态调度和转移预测技术. 现代的高性能处理器包括 Alpha 21264<sup>[3]</sup>, MIPS R10000<sup>[4]</sup>, Sun Ultra Sparc III<sup>[5]</sup>, HP PA-8000<sup>[6]</sup>, IBM Power4<sup>[7]</sup>, Intel Pentium IV<sup>[8]</sup> 和 Itanium<sup>[9]</sup>等, 都广泛采用了这些技术.

龙芯 2 号先进的 Cache 系统设计可以有效地提高流水线的效率. 它能动态地解决地址依赖, 实现了访存操作的乱序执行、非阻塞 Cache、取数指令猜测执行(load speculation)等.

龙芯 2 号处理器采用 0.18 $\mu\text{m}$  的 CMOS 工艺实现, 布线层为 6 层铝金属, 芯片晶体管数目为 1350 万个, 芯片面积为 6700 $\mu\text{m}$   $\times$  6200 $\mu\text{m}$ , 在正常电压下的最高工作频率为 500MHz, 500MHz 时的实测功耗为 3~5W. 龙芯 2 号单精度峰值浮点运算速度为 20 亿次/秒, 双精度浮点运算速度为 10 亿次/秒, Spec CPU2000 的实测性能是龙芯 1 号<sup>[10]</sup> 的 8~10 倍, 是威盛的 VIA-C3 500MHz Pro 处理器的 2~3

倍, 综合性能已经达到 Pentium III 的水平. 目前芯片样机能运行完整的 64 位中文 Linux 操作系统, 全功能的 Mozilla 浏览器、多媒体播放器和 OpenOffice 办公套件, 可以满足绝大多数桌面应用的要求, 第 1 版龙芯 2 号芯片已经成功应用于 Linux PC 的开发.

## 2 取指和转移预测

龙芯 2 号的流水线从取指流水级开始, 每次取 4 条指令, 但每次取指不能跨越 32B 的指令 Cache 行. 龙芯 2 号取指时同时访问指令 Cache 和指令 TLB(简称 ITLB). 为了降低延迟, Tag 比较在取指阶段进行, 但根据 Tag 比较结果进行指令选择在预译码阶段进行.

16 项的 ITLB 是主 TLB 的子集. 当 ITLB 不命中时, 龙芯 2 号处理器产生一个内部指令去查找主 TLB 并且填充 ITLB. 如果在主 TLB 中也不命中将

产生一个普通的 TLB 例外。

龙芯 2 号的第 2 级流水级是预译码流水级。这一级的主要工作是预测转移指令的跳转方向以及目标地址。不同的转移指令使用不同的方式进行预测:likely 类转移指令和直接跳转指令总是被预测为跳转,编译器可以通过编译出 likely 指令进行静态预测;条件转移指令通过 BHT 预测跳转方向;间接跳转指令则用转移目标表(BTB)或返回地址栈(RAS)预测目标地址。

BHT 包括一个 9 位的全局历史寄存器(GHR)和一个 4K 项的模式历史表(PHT)。PHT 的每项是一个两位的饱和计数器。

16 项的 BTB 用于预测寄存器跳转指令的目标地址。每项 BTB 保存转移指令的地址和目标地址,以及一个两位的饱和计数器。当发生替换时,计数器的值小于 2 的项会优先被替换。

龙芯 2 号的第 3 级流水级是译码流水级。在这一级,4 条指令被译成龙芯 2 号的内部指令格式送往寄存器重命名模块。由于定点乘法指令和定点除法指令要生成两个 64 位结果,所以被译为两条内部指令。为了简化转移指令的管理,龙芯 2 号每拍最多只进行一条转移指令的译码。

### 3 寄存器重命名

寄存器重命名是目前通用处理器普遍采用的方法,常见的实现包括 MIPS R10000 的基于 RAM 的方式和 Alpha 21264 的基于 CAM(content access memory)的方式。龙芯 2 号使用基于 CAM 的方法进行寄存器重命名,通过两个 64 项的物理寄存器映射表(PRMT)用来保存物理寄存器和结构寄存器间的关系。

在寄存器重命名流水级,每条指令通过查找 PRMT 表得到该指令的源寄存器对应的物理寄存器号,并分配目标寄存器。同一拍重命名的 4 条指令间的相关也要在同一拍检查。

## 4 动态调度

### 4.1 指令发射和保留站

寄存器重命名后的指令送到保留站调度执行。龙芯 2 号具有两个独立的分组保留站:定点和访存指令送到定点保留站;浮点指令送到浮点保留站。每一个保留站 16 项。

在寄存器重命名阶段,每条指令查找 PRMT 表确定相关操作数是否在寄存器堆中。如果查找重命名表时相关的操作数没有准备好,该指令在送入保留站的途中以及在保留站中要通过比较自己的源寄存器号和结果总线或 forward 总线的目标寄存器号以确定寄存器堆中需要的操作数何时准备好。

保留站每拍最多可以发射 5 个操作数准备好的指令分别到 5 个功能部件。如果在保留站中同一个功能部件有多个操作数准备好的指令,则选择最“老”的指令进行发射。在保留站中增加 age 域来记录每一条指令在保留站中的“年龄”。

### 4.2 读操作数和寄存器堆

龙芯 2 号有一个定点物理寄存器堆和一个浮点物理寄存器堆,大小都是  $64 \times 64b$ 。从保留站发射的指令到寄存器堆中读操作数送到功能部件执行。

定点寄存器堆有 3 个写端口、6 个读端口。两个定点部件和访存部件各使用一个写端口和两个读端口。浮点寄存器堆有 3 个写端口和 5 个读端口。两个浮点部件各使用一个写端口两个读端口;此外,浮点 load 指令使用一个写端口,浮点 store 指令使用一个读端口。定点和浮点寄存器堆间的 move 指令,如 MTC1,DMTC1,MFC1,DMFC1,CTC1 和 CFC1 使用访存数据通路传输数据,因此由访存部件执行。

特殊指令如 branch and link 指令的程序计数器或条件转移指令的 taken 位从转移队列中读出并且与寄存器堆中的操作数一起送到相关的功能部件。

### 4.3 指令提交和 Reorder 队列

在龙芯 2 号中,指令被顺序译码和重命名、乱序发射和执行,但有序结束。Reorder 队列负责指令的有序结束,它按照程序次序保存流水线中所有已经完成寄存器重命名但未提交的指令。指令执行写回后,reorder 队列按照程序顺序提交这些指令。Reorder 队列最多可以同时容纳 32 条指令。

每条完成寄存器重命名的指令在送入保留站的同时也送入 reorder 队列。reorder 队列一拍最多可以提交队列头上的 4 条已经确定写回状态的指令。当指令提交时,指令的重命名信息送到寄存器重命名模块修改重命名状态,它还通知存储队列相应的 store 指令可以开始修改存储器。

为了实现精确例外,在指令执行过程中发生例外时把例外原因记录在 reorder 队列相应的指令中。当例外指令成为 reorder 队列头时进行例外处理,把例外原因、例外指令的 PC 值等例外信息记录到有关的 CP0 寄存器中,并根据例外类型把例外处理程

序的入口地址送到程序计数器 PC 中。

#### 4.4 转移取消和转移队列

转移指令在重命名后进入 reorder 队列和保留站的同时进入转移队列。转移队列同时可以容纳多达 8 条转移指令。当转移指令发射执行时,转移队列提供该指令执行所需的信息,这些信息包括转移指令的 PC 值和条件转移指令的预测 taken 位等。

在转移指令执行后,转移指令的执行结果写回到转移队列。这些结果包括 JR 和 JALR 指令的目标地址、条件转移指令的转移方向和转移指令是否预测错误的标志位。转移指令的执行结果在提交前通过转移总线反馈到取指部分用来修正 BHT, BTB, RAS 和 GHR, 以进行接下来的转移预测。

预测错误的指令和它后面的指令都需要取消。龙芯 2 号为每条指令记录它所在的基本块号,用该信息来进行取消操作。

### 5 功能部件

龙芯 2 号包括两个定点部件 ALU1, ALU2 和两个浮点部件 FALU1, FALU2。

定点 ALU1 执行定点加减、逻辑运算、移位操作、比较操作、Trap 操作以及转移操作。所有 ALU1 执行的指令 1 拍完成执行并写回。

定点 ALU2 执行定点加减、逻辑运算、移位操作、比较操作以及乘除操作。定点乘法为全流水操作,延迟为 4 拍;定点除法采用 SRT 算法,非全流水操作,延迟根据操作数的不同从 4~37 拍不等;所有其他 ALU2 执行的指令 1 拍完成执行并写回。

浮点 FALU1 执行浮点加减、取绝对值、取反、格式转换、比较、转移等操作。所有操作为全流水操作。浮点取绝对值、取反、比较、转移操作延迟为 2 拍,浮点加减、格式转换操作延迟为 5 拍。

浮点 FALU2 执行浮点乘、除、开方操作。浮点乘采用两位 Booth 编码及 Wallace 树算法,为全流水操作,延迟为 5 拍;除法和开方使用 SRT 算法,非全流水结构,根据操作数的不同,单/双精度浮点除延时从 4~10/17 拍,单/双精度开方延时从 4~16/31 拍。

除了执行 MIPS III 浮点指令外,浮点功能部件还可以执行并行单精度浮点指令,即在 64 位数据通路上同时计算两个单精度操作(加、减和乘),32 位或 64 位定点指令(算术、逻辑、移位、比较、转移),并通过浮点指令的扩展格式域 fmt 执行 8 位或 16 位 SIMD 定点指令。

### 6 存储访问与存储管理

龙芯 2 号存储子系统对提高处理器的流水线效率起着重要作用。龙芯 2 号一级指令和数据 Cache 大小均为 64KB,采用 4 路组相联结构;TLB 共有 64 项,为全相联结构,每项映射一个奇数页和一个偶数页。龙芯 2 号通过一个 16 项的访存队列来动态解决存储相关,实现访存指令乱序执行、非阻塞 Cache, load speculation 和 store forwarding 等。

龙芯 2 号访存流水线分为 4 级。访存请求发射到地址运算部件后,第 1 拍地址运算部件计算虚地址并把访存请求送到 TLB 和 Cache;第 2 拍在 TLB 把虚地址转换为物理地址的同时访问 Cache;第 3 拍根据 TLB 和 Cache 的访问结果确定 Cache 是否命中并送到访存队列;第 4 拍把访问结果写回。

龙芯 2 号接口支持 split 读和 R5000 类的片外二级 Cache。片外二级 Cache 复用处理器的系统总线作为二级 Cache 总线,二级 Cache 的块大小为 32B,容量从 256KB~8MB 可调。

#### 6.1 TLB

龙芯 2 号采用 64 项全相联 TLB 进行虚实地址转换。该 TLB 包含一个 CAM 部分进行虚地址的全相联查找,一个 RAM 部分存储物理页号和页的保护位。TLB 的 CAM 比较部分与地址运算在同一拍以避免使用异步 RAM。为降低硬件代价,龙芯 2 号使用 40 位虚地址、36 位物理地址。

龙芯 2 号还通过在 TLB 的每一项增加一个执行保护位来实现可执行保护功能,它可以被用来防范基于缓冲区溢出的许多非法攻击。

#### 6.2 数据 Cache

龙芯 2 号的一级数据 Cache 有 64KB,为 4 路组相联结构。该 Cache 采用虚地址 Index 以及物理地址 Tag 以进行并行的 Cache 和 TLB 查找。替换算法为随机替换,但是通过硬件避免对同一个块进行两个连续的替换。

数据和 Tag 都采用单端口 RAM 以降低芯片的面积。龙芯 2 号允许对不同体的读和写同时进行以降低 Cache 访问冲突。在 Cache 失效时的回填操作 refill、地址运算后访问 Cache 以及存数操作提交后的写回 3 种操作访问 Cache 端口冲突时,refill 具有最高的优先级,存数操作的写回具有最低的优先级。

由于 Cache 的每一路包含 16KB(是最小虚页的 4 倍),虚地址 Index 的两位(13;12)可能与物理地址 Tag 的相应位不相等. 现在的龙芯 2 号不用硬件解决可能的一致性问题的,而是由操作系统解决. 对于虚地址 Index 引起的不一致问题,软件已经有了成熟的解决方案.

### 6.3 访存队列

访存队列是龙芯 2 号存储子系统的核心部件,它记录最多 16 个未执行完的 load 或 store 操作. 虽然 load 和 store 操作乱序进入队列,但在访存队列中按它们程序中出现的次序排列. 访存队列允许 Cache 失效访存操作后面的多个 Cache 失效或命中的访存操作继续进行.

龙芯 2 号在 Cache 失效或相关时不重新进行访存. 访存队列通过物理地址的全相联比较,动态解决访存操作间的相关. 取数操作进入队列时,通过地址比较按字节接收它前面的最近一个对同一地址的存数操作的值;存数操作进入队列时,通过地址比较按字节把所存的值传递给它后面对同一地址的取数操作,直到下一个对同一地址的存数操作.

访存队列有 4 个读端口:第 1 个读端口用来选择第 1 个结果已经准备好的操作写回;第 2 个读端口选择第 1 个写准备好的存数操作,把值写到 Cache,存数操作写准备好,指的是该存数操作 Cache 命中并且已经提交;第 3 个读端口把失效的请求发送到下一级存储器,直接访存操作和例外处理使用最后一个读端口.

## 7 物理实现及初步性能分析

### 7.1 物理设计

目前龙芯 2 号已经使用 SMIC(Semiconductor Manufacturing International Corporation)公司的 6 层铝金属,0.18 $\mu$ m 的 CMOS 工艺流片获得成功. 龙芯 2 号的物理设计在 ASIC 流程的基础上进行了手工布局和手工定制标准单元和宏单元的优化. 物理设计的 ASIC 流程使用 Design Compiler 进行逻辑综合,Physical Compiler 进行单元布局,并用 Astro 进行时钟树生成和布线.

为了降低延迟,龙芯 2 号物理设计在关键路径上进行了手工生成网表和手工布局的工作. 为了对关键路径进行手工生成网表,对部分 RTL 的结构进

行了修改,主要是把数据通路和控制通路分开. 在手工网表的基础上,对关键路径的单元按 bit-sliced 方式进行了手工布局,再与其他单元一起进行 PC 的布局和优化. 此外,还使用 clock skew 技术在关键路径的流水级借用相邻流水级的时间.

由于采用了很多的手工工作,龙芯 2 号在版图划分时把整个设计划分成几块并进行层次化的设计. 但龙芯 2 号的层次化设计与其他层次化设计不同. 一般的层次化设计是做完布线后再拼接起来,龙芯 2 号的层次化设计是在做完布局后就把各个分模块拼接起来,即层次化的布局,但布线是展平的. 这样做的好处是:①在模块划分时不用精确规定 pin 的位置,只要规定大致的方位即可;②在把分模块的布局结果拼接起来之后还可以在 PC 中进行全局的 inc 优化,调整逻辑结构和布局. 因为 Astro 的布局优化功能没有 PC 强.

为物理设计手工定制了高速的标准单元和宏单元是龙芯 2 号物理设计的另一个重要特点. 手工定制的单元包括一些基本单元如延时和功耗比 Artisan 库低的不同大小的触发器、NANDs, NORs, AOIs, Muxs, buffer 和反相器;一些双倍高度的单元,如 4, 6 或 8 位比较器,4 位寄存器和全加器;倒装焊 pad 单元. 手工定制的宏单元包括 3 个写端口、6 个读端口的 64 $\times$ 64b 的寄存器堆和为 TLB 定制的特殊 RAM 宏单元.

### 7.2 初步性能测试和分析

龙芯 2 号的初步性能测试包括 3 个方面:①基本性能参数的测试,包括不同电压下的主频、功耗、峰值速度等;②媒体播放速度测试;③ Spec CPU2000 性能测试和比较.

在基本性能参数的测试中,龙芯 2 号 1.8V 供电在拉偏  $\pm 10\%$  时,分别在 432MHz $\sim$ 500MHz 频率范围下工作正常. 芯片在 450MHz 主频工作时,内部功耗为 4.4W. 龙芯 2 号运行在 500MHz 时,实测得到的最高浮点运算速度单精度为每秒 19.9 亿次,双精度为每秒 9.93 亿次;定点加法为每秒 9.93 亿次;定点乘法为每秒 4.98 亿次.

在媒体播放速度的测试中,龙芯 2 号样机上的 mplayer 播放器能通过软解压流畅播放多种码率的 mpeg1, mpeg2 以及 mpeg4 视频流;在测试模式下,对于 mplayer 的标准测试视频流 matrix 的 mpeg1, mpeg2 以及 mpeg4 格式,450MHz 龙芯 2 号样机播放的帧速率分别 131fps, 29fps 以及 40.36fps. 具体测试结果见表 1.

Table 1 Media Play Speed Test

表 1 媒体播放速度测试

Video File	Format	Resolution	Bitrate (Kbps)	Normal Frame Rate (fps)	Total Frame	Achieved Frame Rate (fps)
matrix_vcd.mpg	mpeg1	352×288	1150.0	25	4675	130
matrix_mpg2.mpg	mpeg2	720×576	11421.6	25	4675	28.9
matrix_mpg4.avi	mpeg4	720×576	1655.7	25	4675	40.4
star_war.avi	mpeg4	720×576	1858.3	25	4750	31.2
hannibal.avi	mpeg4	480×260	153.7	24	3336	154
sample.avi	mpeg4	640×352	823.3	23	1518	54.2

Spec CPU2000 是一个国际通用处理器生产厂家公认的处理器性能测试标准程序包<sup>[11]</sup>, 共有 12 个定点程序和 14 个浮点程序. 龙芯 2 号的 Spec CPU2000 性能测试采用 gcc3.4.1 编译器, Spec CPU2000 测试程序的 4 个需要 F90 编译器的测试程序在龙芯 2 号样机上暂时还不能运行. 为了便于比较分析, 我们分别在 250MHz 的龙芯 1 号样机、450MHz 的龙芯 2 号样机、使用 500MHz 的 VIA-C3 处理器的样机以及一台采用 450MHz Pentium III 处

理器的微机上运行了 Spec CPU2000 测试程序. 图 2 给出了上述机器运行 Spec CPU2000 测试程序的分值比较. 从图 2 可以看出, 在使用 gcc3.4.1 编译器的情况下, 龙芯 2 号 Spec CPU2000 的定点分值是龙芯 1 号的 8~10 倍, 是威盛的 VIA-C3 500MHz Pro 处理器的 2 倍; 是浮点分值龙芯 1 号的 6~8 倍, 是威盛的 VIA-C3 500MHz Pro 处理器的 3 倍; 龙芯 2 号的总体性能达到了 Pentium III 处理器的水平.

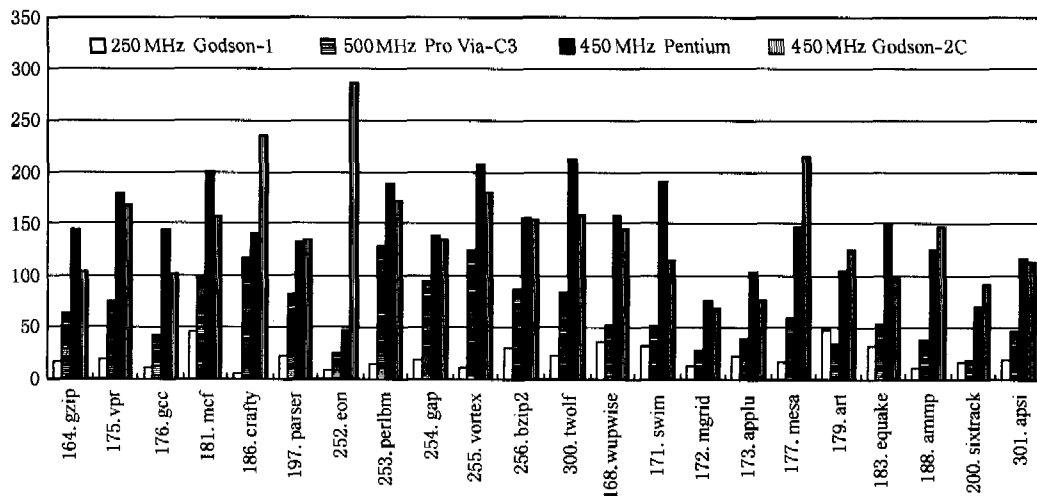


Fig. 2 Godson-2 SPEC CPU2000 marks.

图 2 龙芯 2 号 SPEC CPU2000 分值测试

图 3 给出了使用 CacheBench 测试程序测试龙芯 2 号样机和 Pentium III 参考机的访存带宽结果. 从图 3 可以看出:

① 龙芯 2 号的带宽在块大小为 64KB 时出现明显的下降, 而 Pentium III 参考机在块大小为 16KB 时就出现明显下降, 反映了龙芯 2 号 64KB 大小的数据 Cache 对 Pentium III 参考机 16KB 数据 Cache 的优势;

② 在访问块较小时, Pentium III 参考机带宽较低而龙芯 2 号的带宽比较稳定以及龙芯 2 号的读带

宽比 Pentium III 参考机高, 说明龙芯 2 号流水线效率较高;

③ Pentium III 参考机的二级 Cache 对带宽的效果明显, 而龙芯 2 号的二级 Cache 效果不明显, 主要是由于 Pentium III 有专门的二级 Cache 接口而龙芯 2 号使用系统总线作为二级 Cache 接口导致龙芯 2 号二级 Cache 的块大小较小而且访问延迟长效率低;

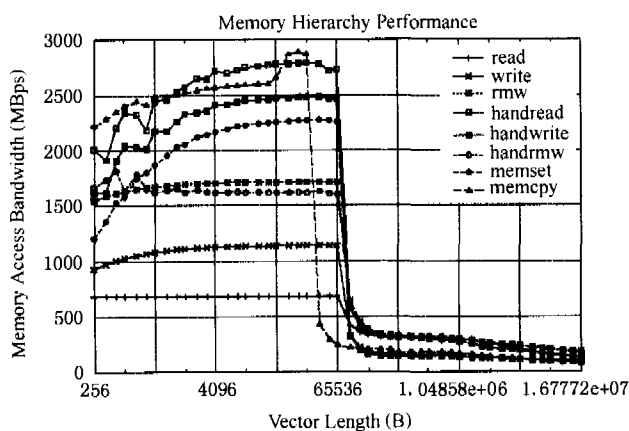
④ 在访问的块很大导致一级和二级 Cache 都发挥不了作用时, Pentium III 的读带宽达到 200MBps 以上, 而龙芯 2 号的读带宽还不到 100MBps. 从

Spec CPU2000 的测试结果也可以看出,对于访存带宽要求较高的程序如 164.zip, 175.gcc, 171.swim 以及 173.aplu 等,龙芯 2 号样机的性能明显低于相同主频的 Pentium III 参考机。

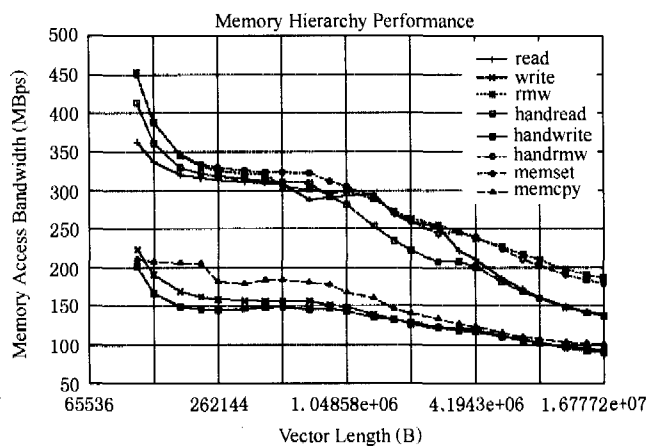
上述带宽测试结果表明,龙芯 2 号的性能瓶颈很大程度上在于访存带宽,而访存带宽和处理器套片的性能直接相关。使用正在研制的龙芯 2 号套片

可以比目前使用的 Marvell 套片访存带宽提高一倍左右。

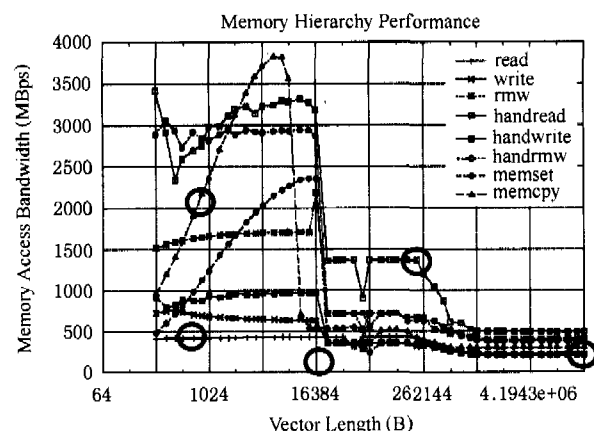
此外,Spec CPU2000 的性能和编译器也直接相关,我们在使用 MIPS R10000 处理器的 SGI 工作站上的测试表明,使用 MIPS Pro 编译器比使用 gcc 编译器 SPEC CPU2000 的定点分值高 75%。



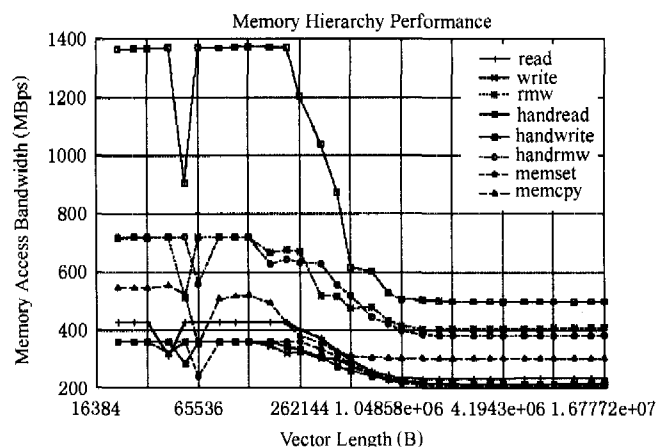
(a)



(b)



(c)



(d)

Fig. 3 Bandwidth comparison of Godson-2 and Pentium III.

图 3 龙芯 2 号和 Pentium III 带宽测试与比较

## 8 总结和未来的工作

本文介绍了龙芯 2 号微处理器体系结构。龙芯 2 号是一款 64 位、4 发射、乱序执行的 RISC 处理器,实现类 MIPS 指令集。采用先进的乱序执行技术(如寄存器重命名、转移预测和动态调度)和 Cache 技术(如非阻塞 Cache, load 猜测和动态内存相关),使龙芯 2 号在不是很高的主频下就获得很高的性能。

我们目前正在进行的工作包括进一步改进处理器核和存储系统。处理器核的改进包括支持条件转

移指令和浮点乘加指令以及增加有关队列大小;存储系统的改进包括片内二级 Cache 的实现和片内 DDR 存储接口。并在上述部分或全部改进的基础上,通过更多的定制设计和更先进的工艺实现下一版主频达到 1GHz 的龙芯 2 号。此外,与龙芯 2 号匹配的高性能套片正在设计中。该套片将通过包括与 DDR 控制器结合在一起的预取以及页管理策略猜测在内的各种技术来提高带宽。

在未来的龙芯 3 号中将实现单芯片多处理器来开发多线程并行,并努力争取使用多核来加速单线程的执行。

## 参 考 文 献

- 1 MIPS IV instruction set. <http://www.mips.com>, 1995
- 2 David Patterson, John Hennessy. Computer Architecture: A Quantitative Approach. San Francisco: Morgan Kaufmann, 1996
- 3 R. Kessler. The Alpha 21264 microprocessor. IEEE Micro, 1999, 19(2): 24~36
- 4 Kenneth Yeager. The MIPS R10000 superscalar microprocessor. IEEE Micro, 1996, 16(3): 28~41
- 5 Tim Horel, Gary Lauterbach. UltraSparc-III: Designing third-generation 64-bit performance. IEEE Micro, 1999, 19(3): 73~85
- 6 Ashok Kumar. The HP PA-8000 RISC CPU. IEEE Micro, 1997, 17(2): 27~32
- 7 J. Tendler, S. Dodson, S. Fields, *et al.* Power 4 system microarchitecture. New York: IBM Corporation, 2001
- 8 G. Hinton, D. Sager, M. Upton, *et al.* The microarchitecture of the Pentium IV processor. Intel Technology Journal, 2001. <http://www.intel.com/technology/itj/q12001/pdf/art.2.pdf>
- 9 J. Huck, *et al.* Introducing the IA-64 architecture. IEEE Micro, 2000, 20(5): 12~23
- 10 Hu Weiwu, Tang Zhimin. Architecture of the Godson-1 processor. Chinese Journal of Computers, 2003, 26(4): 385~396 (in Chinese)  
(胡伟武, 唐志敏. 龙芯 1 号处理器结构设计. 计算机学报, 2003, 26(4): 385~396)
- 11 J. Henning. SPEC CPU2000: Measuring CPU performance in the new millennium. Computer, 2000, 33(7): 28~35



**Hu Weiwu**, born in 1968. Received his B. S. degree in computer science from the University of Science and Technology of China in 1991 and his Ph. D. degree in computer science from the Institute of

Computing Technology, the Chinese Academy of Sciences in 1996, respectively. He is currently a professor and Ph. D. supervisor in the Institute of Computing Technology. His research interests include high performance computer architecture, parallel processing, and VLSI design.

胡伟武, 1968 年生, 博士, 研究员, 博士生导师, 主要研究方向为高性能计算机体系结构、并行处理和 VLSI 设计等。



**Zhang Fuxin**, born in 1976. Received his B. S. degree in computer science from the University of Science and Technology of China in 2000 and his Ph. D. degree in computer science from the Institute of Computing Technology, the Chinese Academy of Sciences in 2005, respectively. He is currently associate professor in the Institute of Computing Technology. His research interests include high performance computer architecture, microprocessor architecture, and performance analysis.

张福新, 1976 年生, 博士, 副研究员, 主要研究方向为高性能计算机体系结构、微处理器体系结构和性能分析等。



**Li Zusong**, born in 1977. Received his B. S. degree from Tsinghua University in 2000. He is currently a doctorate candidate in the Institute of Computing Technology, the CAS. His main research interests include high performance computer architecture, on-chip multi-threading technology, and VLSI design.

李祖松, 1977 年生, 博士研究生, 主要研究方向为高性能计算机体系结构、片内多线程技术和 VLSI 设计等。

## Research Background

In this paper we introduce the design and preliminary performance analysis of the godson-2 processor, which is China's effort to make its own high performance microprocessor. The detailed micro-architecture is presented, along with the physical design methods and performance results. Future work includes making more architectural improvements on the core, transferring to more advanced CMOS technology, researching on fault tolerance techniques and multi-core designs. Our work is supported by the National Science Foundation of China(60325205), Hi-Tech Research and Development Program(863) of China(2002AA110010, 2005AA110010, 2005AA119020) and the National Basic Research Program (973) of China(2005CB321600).