

龙芯处理器服务器芯片组的适配与实现

郑臣明¹⁾, 姚宣霞^{1)✉}, 周 芳¹⁾, 郑雪峰¹⁾, 杨晓君²⁾, 戴 荣³⁾

1) 北京科技大学计算机与通信工程学院, 北京 100083 2) 海光信息技术股份有限公司, 北京 100193 3) 中科曙光信息产业成都有限公司, 成都 610213

✉通信作者, E-mail: yaoxuanxia@ustb.edu.cn

摘 要 针对龙芯中央处理器(CPU)无对应高性能服务器芯片组的现状,设计开发了一种为龙芯 CPU 筛选芯片组的架构,并实现了一种龙芯 CPU 和芯片组适配的方法。提出了采用现场可编程门阵列(FPGA)串联在龙芯 CPU 和即将适配的多组芯片组之间的架构。借助于此架构,设计实现了在 CPU 和芯片组之间待处理物理信号线的连接方法,设计了两组之间上下电时序配合的调试方法,设计实现了规避两者信号协议差异的方法。借助该架构和这些方法能够实现同时筛选多款芯片组的目的,避免了以前需要设计多款主板进行适配的情况,节省了重复研发主板的成本;找到了可以适配龙芯 CPU 的高性能服务器芯片组;其芯片组规格参数和性能高于目前龙芯 CPU 所用的芯片组,开拓了其在服务器领域的应用。

关键词 龙芯; 芯片组; 适配; 服务器; 现场可编程门阵列

分类号 TP302.1

Adaption and implementation of server chipsets for the Loongson CPU

ZHENG Chen-ming¹⁾, YAO Xuan-xia^{1)✉}, ZHOU Fang¹⁾, ZHENG Xue-feng¹⁾, YANG Xiao-jun²⁾, DAI Rong³⁾

1) School of Computer and Communication Engineering, University of Science and Technology Beijing, Beijing 100083, China

2) Haiguang Information Technology Co., Ltd., Beijing 100193, China

3) Dawning Information Industry Co., Ltd., Chengdu 610213, China

✉ Corresponding author, E-mail: yaoxuanxia@ustb.edu.cn

ABSTRACT The CPU is the core part of all integrated circuits. Although some homemade CPUs of proprietary intellectual property rights are rapidly developed, few high-performance chipsets are available, especially in server domains, to match them. Thus, the total systems designed using these CPUs and low-performance chipsets do not have proper performance. The Loongson CPU faces the same problem. To seek better chipsets for it, certain architecture and some methods are designed and implemented to adapt different types of chipsets. In this architecture, a field-programmable gate array (FPGA) is linked between a CPU and these chipsets. An FPGA is divided into three domains: an HT (hyper transport) bus domain, a processing domain for important but temporarily indeterminate signals, and a CPLD (complex programmable logic device) function domain. In these adaption processes, HT bus signals, the temporarily indeterminate signals, and power signals in CPUs and chipsets are respectively linked into three domains in an FPGA and treated by a programming FPGA to perform all types of possible signal combinations. The power sequence between the CPU and chipsets is coordinated to the right order using an FPGA. The signal integrity difference between them is avoided and trimmed to the right state by amending their signals in the FPGA. In this system, the experimental results show that this architecture and these methods simultaneously make more chipsets work together to be adapted than before in a single motherboard. This combination avoids researching and developing many different motherboards for every type of possible chipset and greatly reduces costs. High-performance server chipsets can be found to properly match the Loongson CPU and have better specifications and higher performance than those

收稿日期: 2021-10-08

基金项目: 国家重大科技专项“核心电子器件、高端通用芯片及基础软件产品”资助项目(2017ZX01028-102)

currently used for the Loongson CPU. A prototype system composed of the Loongson CPU and five types of chipsets is designed and implemented. Using the above architecture and methods, a type of optimal server chipsets SR5690 + SP5100 has been found, and the matching principles or correct settings for the signal connection and power sequence have been concluded. The Loongson 3B4000 two-way SMP motherboard with SR5690 + SP5100 chipsets is also produced. On this motherboard, the results of evaluation experiments on computing performance tests by the SPEC CPU 2006 program, storage performance tests by the IO zone program, and network performance tests by the Netperf program are performed. Compared with the current Loongson 3B4000 server with a 7A1000 chipset, the test results show the performance on three items is improved by approximately 10%. The combination of the Loongson CPU and this type of server chipset provides wider applications in the server market and promotes the development of the Loongson CPU in its ecosystem.

KEY WORDS Loongson; chipsets; adaption; server; field-programmable gate array

集成电路是国家信息产业发展的核心领域,而 CPU(中央处理器)则是集成电路中的核心元器件。目前国家对信息安全有着迫切的要求,如果没有 CPU 的国产化与自主化,国家安全将无从谈起^[1-2]。龙芯 CPU、飞腾 CPU、申威 CPU 是目前被国家认可的、应用领域比较广的三款自主可控 CPU^[3-5]。“龙芯”处理器是龙芯中科技术有限公司自主研发的通用 CPU。“飞腾”处理器是飞腾信息技术有限公司开发的飞腾系列高性能通用 CPU。“申威”处理器由国家高性能集成电路(上海)设计中心自主研发,具体由江南计算所研制实施。上述处理器都得到了国家“863”、国家科技重大专项“核高基”(核心电子器件、高端通用芯片及基础软件产品)等项目^[6]的支持,在党政、关键基础设施、高性能计算等领域已得到了大量的应用,从应用的结果来看,飞腾^[7-8]、申威^[9-10]、龙芯^[11]性能表现不错,可以胜任目前绝大部分应用^[12]。

这三款 CPU 性能不错,如以申威 CPU 研发的神威·太湖之光超级计算机在 2021 年中国高性能计算机(HPC)TOP100 排行榜中位列第 2 名^[13],但存在一个较大的短板,即缺乏与其配套的高性能芯片组,特别是在服务器领域,虽有强大的大脑,但无强壮的躯干和四肢,严重地限制了 CPU 高性能的发挥、应用领域的拓展和产业化进程。在服务器产品上,飞腾和申威 CPU 采用 PLX 公司的 PCIE(Peripheral component interconnect express)交换芯片作为北桥,例如飞腾的腾云 S2500 CPU 推荐的参考设计采用型号为 PEX8764 的 PCIE 交换芯片,申威的 1621 CPU 推荐的参考设计采用型号为 PEX8748 的 PCIE 交换芯片,利用这些 PCIE 交换芯片转换出更多的 PCIE 总线,再外接具有 PCIE 接口的 SAS(Serial attached SCSI,其中 SCSI:Small computer system interface)/SATA(Serial advanced technology attachment)控制器(例如 Marvell

公司的 88SE9215)生成硬盘接口;再外接具有 PCIE 接口的显卡芯片产生显示器接口等。此种芯片组的方案是迫不得已的拼凑方案,缺点是为了得到所要的 IO(Input output)接口而需要通过多级串联不同种类芯片的方式来实现,稳定性和兼容性很差,信号延迟增大,性能变差。

与龙芯 CPU 相配套的芯片组,最初有 PCI(Peripheral component interconnect)接口桥片方案^[14],但 PCI 总线速度太低,已逐渐被淘汰,然后有龙芯 2H 方案^[15],但龙芯 2H 桥片性能太弱,只有 4 对 PCIE 信号线和 2 个 SATA 接口,而且龙芯 2H 其实是一款嵌入式 CPU,在龙芯 CPU 无桥片可用的情况下把它作为桥片是不得已而为之。在 2017 年底,龙芯中科技术有限公司又推出了龙芯 7A1000 桥片^[16]作为龙芯处理器的配套芯片组。龙芯 7A1000 桥片通过 HT(Hyper transport)总线接口与龙芯 CPU 相连,接口包括 32 个 PCIE2.0 通道、3 个 SATA2.0 通道、6 个 USB(Universal serial bus)2.0 接口,可以满足桌面领域对 IO 接口的应用需求。但龙芯 7A1000 桥片在 IO 数量、RAS(Reliability 可靠性、Availability 可用性、Serviceability 可维护性)和 IOMMU(Input/output memory management unit,输入输出内存管理单元)等方面的不足难以满足服务器芯片组的苛刻要求。

本文针对龙芯 CPU 的特点,基于目前市场上现有的芯片组,筛选适配出一套高性能的芯片组来满足龙芯 CPU 在服务器级别上应用。由于所要筛选的芯片组不是专为龙芯 CPU 设计的,所以在龙芯 CPU 和芯片组适配嫁接过程中,需要解决在物理信号定义、上下电时序、信号协议、BIOS(Basic input output system)层等方面的差异。本文为此设计了一套筛选芯片组的架构,提供 CPU 和芯片组两者之间在物理信号连接、上下电时序、规避信号协议差异方面的设计方法,增加调试的

灵活性和信号裕度,能够简单、反复地调整参数和设计,求证芯片组的可用性和兼容性,为龙芯 CPU 很便捷地找到服务器级别的芯片组。

本文的主要内容如下:

(1)基于目前市场上现有的芯片组,提出了一种为龙芯 CPU 筛选适配芯片组的架构,此架构同样适用于其他缺少配套芯片组的国产 CPU。

(2)提出了一种 CPU 和芯片组两者之间在物理信号连接、规避信号协议差异方面的设计方法,为在适配过程中暂时不知如何处理的物理信号找到正确的连接方法或者处理方式。

(3)提供了一种灵活调试主板上下电时序的方法,在不修改硬件的条件下,可以随意调整 CPU、芯片组、电源之间的时序,直到找到正确的上下电时序为止,节省了大量研发时间和成本。

(4)为龙芯 CPU 找到了一组高性能的服务器级别的芯片组,改变了一直采用低性能、嵌入式/桌面级芯片组充当服务器芯片组的局面,更好地发挥龙芯 CPU 性能,拓宽龙芯 CPU 的应用领域,进一步促进龙芯 CPU 生态环境的发展。

1 龙芯 CPU 和芯片组适配的架构设计

CPU 和芯片组适配的常规方法^[17-18]是做一块主板把两者直接连接起来进行调试,但这种调试方法存在明显的缺点。在适配不成功时很难判断是由于芯片组和 CPU 两者之间的不兼容引起的,或者本来两者兼容只是没有正确连接处理引起的,还是由于 CPU 或芯片组的缺陷引起的,原因不容易定位。即使原因定位了,由于主板不能修改,特别是由 CPU 或芯片组的设计缺陷造成的适配成功,在不可更改硬件的情况下不得不重新设计主板、修改 CPU 或者芯片组的设计,浪费了大量时间和成本。

在芯片研发设计中,为了降低研发风险,缩短研发周期,避免多次流片产生的高额成本,通常都会在芯片流片加工之前进行充分的仿真模拟。目前常用的仿真模拟方法有三种,基于软件仿真模拟验证的传统方法^[19],基于硬件仿真加速器的仿真模拟验证方法^[20],以及基于现场可编程门阵列(FPGA)的原型验证方法^[21]。基于软件仿真模拟验证的传统方法,灵活性强,使用简单,但随着逻辑单元规模的增大模拟验证速度下降明显,一般只能达到几十赫兹(Hz),与现在达到上吉赫兹(GHz)系统相差甚远,并且软件模拟仿真与真实的物理环境有差距,存在许多难于发现的问题。基于硬件

仿真加速器的仿真模拟验证方法,运行速度能达到几兆赫兹(MHz),其运行速度比基于软件仿真模拟验证的传统方法提升了上千倍,但仍然达不到上吉赫兹的要求,并且此方法专用性强,维护成本高^[22-24]。基于 FPGA 的原型验证方法相对于其他两种方法,FPGA 的运行速度可以达到百兆级别,而且逻辑资源、IO 资源比较丰富,能够更加真实地模拟真实芯片的行为,一旦系统模拟成功,可以采用硬件拷贝(Hardcopy)方式直接快速地固化为 ASIC(Application specific integrated circuit)芯片。FPGA 具有非常灵活的在线编程特点^[25-26],在不改变硬件的情况下能够反复地修改代码改变硬件行为^[27-28],能模拟出目前常见的 IO 接口和复杂的总线^[29-31]、能灵活地变换不同时钟频率、能调节改变信号的电平,并内含在线逻辑分析仪功能^[32]。本文利用 FPGA 这些优势,在龙芯处理器和芯片组之间串联一颗 FPGA 芯片,每组芯片组通过 FPGA 芯片后再和龙芯处理器相连,如图 1 所示,这样可以在一块主板上实现多组芯片组和龙芯处理器的适配,既增加了适配成功的几率又节省了成本。

基于 FPGA 的原型验证方法,传统的使用方法是只模拟验证单一架构/单一功能,如果模拟多架构/多功能需要使用多块 FPGA 互联来实现。本文对其进行了改进,在 FPGA 内部模拟划分了三个不同的功能模块,利用一块 FPGA 芯片来实现,并且通过设计带宽匹配模块(Bandwidth matching module)实现不同频率、不同带宽总线的自协商、自适应,内嵌逻辑分析仪(Logic analyzer)可以做到实时调试分析,加快研发调试进度。

FPGA 内部分为三个功能域:HT 总线^[33-34]处理域、重要暂不确定控制信号线处理域和 CPLD (Complex programmable logic device)功能域,如图 2 所示。

HT 总线处理域用于处理龙芯 CPU 和芯片组之间 HT 总线的适配。通过采用第三方或者按照 HT 总线协议自行编写 HT 总线的 IP 核,在 FPGA 上实例化 HT 总线接口,用来分别连接龙芯 CPU 和芯片组的 HT 总线。由于 FPGA 的运行速度一般在几百兆赫兹,而 HT 总线频率可达上吉赫兹,所以在 FPGA 内设计协议转化模块(Protocol converting module)把 HT 总线的高速串行信号线转化为较为低速的并行内部总线,然后连接到交叉开关总线(Crossbar bus)上。交叉开关总线可以通过软件控制选择切换任意一组芯片组和龙芯 CPU 建立 HT 总线互联,实现在一块主板上能同时调试适

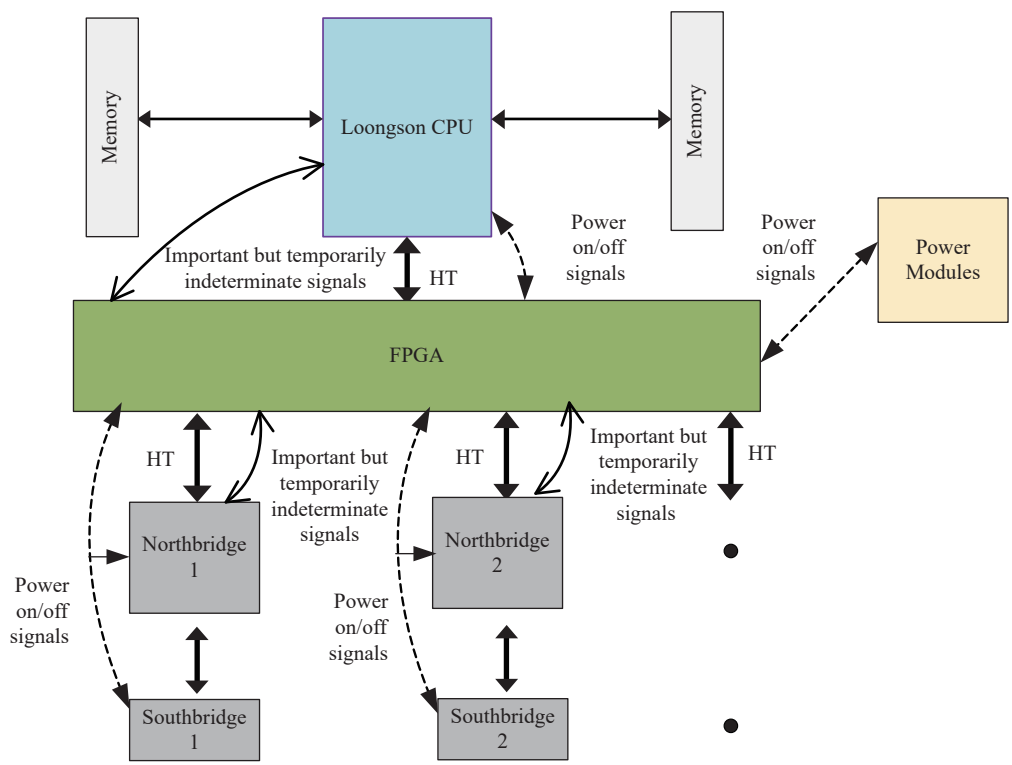


图 1 龙芯 CPU 和芯片组之间的适配架构

Fig.1 Adaptation architecture between Loongson CPU and chipsets

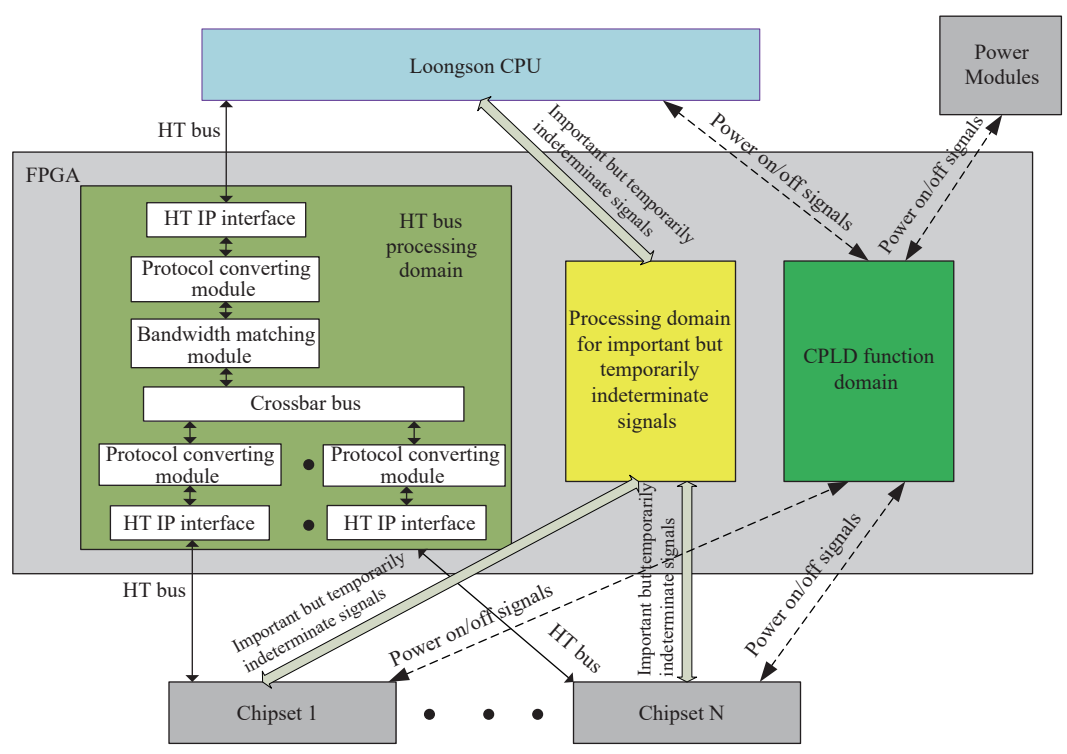


图 2 FPGA 内部总体架构图

Fig.2 Overall architecture of the FPGA

配多款芯片组，十分灵活。由于芯片组和龙芯 CPU 有可能存在 HT 总线差异，例如有的是 HT 2.0 总线协议（最高工作频率可达 1.4 GHz），或者

HT 3.0 总线协议（最高工作频率可达 2.6 GHz），需要在 FPGA 内部设计一个带宽匹配模块，自动协商解决两者之间带宽差异问题。

龙芯 CPU 有一些独属于自己的信号线, 在即将要适配的芯片组里也有一些特殊的信号线, 由于芯片组不是专门为龙芯 CPU 配套而设计的, 在未适配成功前, 有些信号是不知道如何连接或者处理的. 本文把这些信号线命名为“重要暂不确定控制信号线”, 将其全部引入 FPGA 内部设置的“重要暂不确定控制信号线处理域”, 借助于 FPGA 灵活编程的特性, 可以在 FPGA 内部把两者的信号连接起来, 或者单独地做上拉、下拉、悬空等处理. 通过反复地实验和调试, 总能找出“重要暂不确定控制信号线”的处理方式, 给出到底能否兼容的结论.

龙芯 CPU 和芯片组都有自己的上下电控制时序, 彼此之间怎样配合也是未知的. 上下电时序信号主要包括三部分: CPU 上下电控制信号, 芯片组上下电控制信号、电源模块上下电控制信号. 这三部分信号都需要引入 FPGA 内部设置 CPLD 功能域. 在主板设计中, 常采用 CPLD 芯片来控制整个主板的上下电时序, 本文采用在 FPGA 内部分离出一部分逻辑资源实现一个 CPLD 功能域, 依靠 CPLD 来调整三部分的上下电时序来达到 CPU 和芯片组协调工作的目的. 通过反复地实验和调试, 找出相互配合工作的最佳时序或者给出无法兼容的结论.

在龙芯 CPU 和芯片组适配的过程中, 通过本架构可以实现如下功能或优点:

(1) 一块主板可以适配多款芯片组(同一时间只能实现一款芯片组和 CPU 互联), 不用再为每种芯片组单独研发一种主板, 节省了研发不同主板的时间和成本.

(2) 通过 FPGA 的隔离, 能有效地屏蔽 CPU 和芯片组在 HT 总线方面的差异, 例如协议版本的不同、信号线差异, 甚至 HT 设计的缺陷. 通过 FPGA 补偿龙芯 CPU 和芯片组的 HT 总线差异或缺陷, 使两者无缝地连接起来, 而不会因为 HT 总线的差异使适配过程半途而废, 所以此架构具有很大的灵活性, 节省了时间和成本.

(3) 无论龙芯 CPU 还是芯片组在未适配成功前, 那些无法确定怎样连接或者处理的信号线通过 FPGA 可以实现任意连接或处理, 增加了调试的手段, 当找出确定的处理方式后, 可指导下一版龙芯 CPU、芯片组或者主板的改进.

(4) CPU 和芯片组两者的上下电时序配合通过 FPGA 编程来灵活的调节时序和时间间隔, 可以实现任意的组合尝试, 增强了调试的灵活性.

2 龙芯 CPU 和芯片组适配的实现过程

2.1 芯片组的选择

根据龙芯 CPU 的特点, 在理论上推断目前市场上有如下的芯片组可与龙芯 CPU 进行适配实验:

(1) 矽统科技 (SiS) 公司的芯片组, 北桥芯片型号 SiS761GX, 南桥芯片型号 SiS968;

(2) 英伟达 (NVIDIA) 公司的芯片组, 型号 MCP68;

(3) AMD 公司的芯片组, 北桥芯片型号 M690T, 南桥芯片型号 SB600;

(4) AMD 公司的芯片组, 北桥芯片型号 RS780E, 南桥芯片型号 SB710;

(5) AMD 公司的芯片组, 北桥芯片型号 SR5690, 南桥芯片型号 SP5100.

龙芯 CPU 和即将要适配的芯片组按照图 1 的架构设计一块主板试验品, 龙芯 CPU 采用子板的方式与芯片组主板互联, 试验品实物图如图 3 所示. 主板上包含了如上五种芯片组和 FPGA, 并把 PCIE、USB、GPIO (General purpose input/output)、控制线等信号线引入到相应的连接器上.

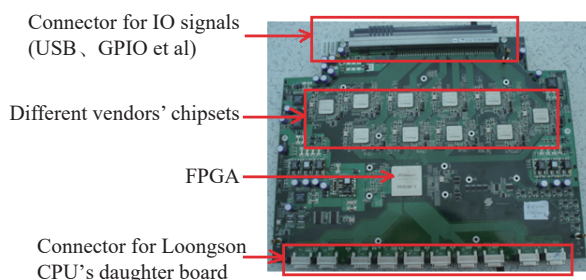


图 3 包含 FPGA 和各种芯片组的主板

Fig.3 Motherboard sample containing the FPGA various chipsets

FPGA 采用 Intel Stratix 10 GX 2800 系列 (简称 GX 2800), 具体型号为 1SG280HU2F50E2LG. FPGA 选型主要从“可编程 IO 单元”和“可编程逻辑单元”两个方面考虑. 在可编程 IO 单元方面, 龙芯 CPU 和五种芯片组共有 6 组 HT 总线, 其中芯片组 1 和芯片组 2 的 HT 总线速率最高为 1 GHz, 龙芯 CPU 和芯片组 3、芯片组 4、芯片组 5 的 HT 总线速率最高为 2.6 GHz, 每个 HT 总线有 20 对差分信号线, 所以 FPGA 需要提供 40 对可达 1 GHz 高速串行信号线和 80 对可达 2.6 GHz 高速串行信号线. 因为 GX 2800 具有 576 对可达 1.6 GHz 高速串行总线和 96 对可达 17.4 GHz 高速串行总线, 所以满足本设计要求. 另外 GX 2800 可提供 1160 个其他类型的 IO 针脚, 足以满足本设计的要求. 在可编程逻辑单元方面, 每种芯片组的 HT 总线和协议

转换模块大约需要 9 万个可编程逻辑单元,龙芯 CPU 和五种芯片组大约共需 54 万个;交叉开关总线和带宽匹配模块大约需要 12 万个可编程逻辑单元;重要暂不确定控制信号线处理域大约需要 4000 个可编程逻辑单元;CPLD 功能域大约需要 3000 个可编程逻辑单元,所以共需大约 67 万个可编程逻辑单元. GX 2800 具有 93.3 万个可编程逻辑单元,本设计资源利用率大约在 72%,所以完全满足本设计的资源需求.

2.2 信号线的设计处理

2.2.1 HT 总线的设计处理

标准的 HT 总线协议包括三类信号线:连接信号线(数据、地址、控制和时钟)(表 1)、复位和初始化信号线(表 2)和电源管理信号线(表 3).

表 1 HT 总线的连接信号线

Table 1 Hyper transport bus link signals

Signal	Width	Description
CAD	2, 4, 8, or 16	Command, addresses, and data (CAD). Carries HyperTransport™ requests, responses, addresses, and data. CAD width can be different in each direction.
CTL	1, 2, or 4	Differentiates control and data. Each byte of CAD has a control(CTL) signal in the Gen3 protocol. One CTL signal is used for an entire link in the Gen1 protocol.
CLK	1, 2, or 4	Clocks(CLK)for the CAD and CTL signals. Each byte of CAD and its respective CTL signal has a separate clock signal.

表 2 HT 总线的复位/初始化信号线

Table 2 Reset/Initialization signals of the HT bus

Signal	Width	Description
PWROK	1	Power and clocks are stable
RESET#	1	Reset the HyperTransport™ chain

表 3 HT 总线的电源管理信号线

Table 3 Power management signals

Signal	Width	Description
LDTSTOP#	1	Enables and disables links during system state transitions
LDTREQ#	1	Indicates link is active or requested by a device

对于连接信号线,龙芯 CPU 和上面所提到的芯片组(或北桥)都有 16 对接收差分信号线,16 对发送差分信号线;2 对接收信号时钟差分对,2 对发送信号时钟差分对;2 对接收控制信号差分对,2 对发送控制信号差分对. 两者的 HT 总线连接信号线都直接引入 FPGA 内部的 HT 总线域进行处理,需要注意的是不同型号的龙芯 CPU 和所选择

的芯片组遵循 HT 协议版本不同,有的是 HT2.0 协议,有的是 HT3.0 协议,即使都支持 HT3.0 协议,也存在 1.8、2.0、2.4、2.6 GHz 四种频率等级,所以在适配的过程中需要采用带宽匹配模块进行调整.

在实际的适配中,利用 FPGA 内部设置的交叉开关只需连接龙芯 CPU 和所要适配芯片组的 HT 总线,而断开其他芯片组连接. HT 总线调试采取频率、通道数从低到高逐步增加的步进调节方法. HT 可支持 2、4、8 和 16 位(bit)四种通道模式,HT 协议的最低工作频率可低至 200 MHz,HT2.0 协议支持最高频率为 1.4 GHz,HT3.0 协议支持的最高频率为 2.6 GHz.

HT 调试流程如图 4 所示,在 HT 调试过程中,先固定工作频率,从最低的 2 位通道数进行调试,如果适配成功,就依次增加通道数继续调试,即按照 2 位、4 位、8 位、16 位顺序调试. 如果适配不成功,就利用 FPGA 内部的逻辑分析仪采样信号进行分析,修正补偿相应的信号线后,再重新进行适配. HT 总线频率调试顺序依次为 200、400 和 800 MHz,一直到芯片工作最高频率(F_{max}).

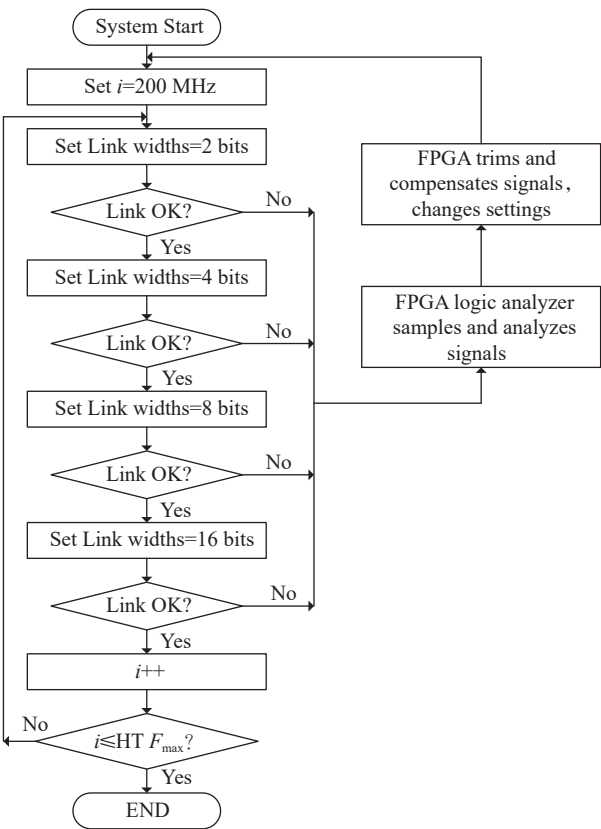


图 4 HT 总线调试流程

Fig.4 Flow of the HT bus debug

通过如上 HT 总线的适配,可以确定信号的正确连接方式或者处理方法,能够定量地找出 CPU

和芯片组的 HT 总线在主板 PCB(Printed circuit board)设计中信号线的最佳分组方式、长度限制、线宽限制、相邻信号线的间距限制,能够找出两者之间 HT 总线的最佳工作频率甚至下一版芯片设计中需要修改的错误或者改进之处。

利用如上的 HT 调试流程分别对如上五种芯片组进行调试,结果表明,第一种和第三种芯片组可以调试成功的最高频率是 800 MHz,第二种芯片组在达到 HT 总线 400 MHz 时出现不稳定的现象,第四种和第五种可以达到 2.0 GHz HT 总线频率。第四种芯片组,即 RS780E + SB710 组合,PCIE 总线有 22 对信号线,不支持服务器领域要求的 RAS 特性,但成本低,所以非常适合作为龙芯的桌面芯片组。第五种芯片组 SR5690 + SP5100,拥有 42 对 PCIE 信号线、14 个 USB 接口,具有强大的 IO 能力,并支持 RAS 特性,所以可以得出在服务器领域与龙芯 CPU 适配成功的最高端、最适合的芯片组为 SR5690 + SP5100。在本文后面的其他章节中,在阐述芯片组适配与实现的通用原理和方法后,为了文章的简洁,主要以 SR5690 + SP5100 芯片组进行阐述。

2.2.2 重要暂不确定控制信号线的设计处理

龙芯 CPU 的 HT 总线由于其独特的设计比标准的 HT 总线协议多出了一些额外的、特殊的信号,所以在硬件层和协议层需要正确地处理这些特殊信号。既不能影响龙芯 CPU 自身正常工作,又能和芯片组无缝的兼容,这是一个比较关键的技术点。在标准的 HT 总线协议中,信号 PWROK、RESET#、LDTSTOP#、LDTREQ#(表 2 和表 3)只有一组,而龙芯 HT 总线接口信号有两组,一组是 HT_HI_POWEROK、HT_HI_RSTn、HT_HI_LDT_STOPn、HT_HI_LDT_REQn,另一组是 HT_LO_POWEROK、HT_LO_RSTn、HT_LO_LDT_STOPn、HT_LO_LDT_REQn。龙芯 CPU 和芯片组还有其他各自特殊的信号线,无法直接确定两者之间的对应关系。以上这些重要暂不确定控制信号线都引入 FPGA,如图 1 和图 2 所示。

这些信号如何处理,怎样和北桥、南桥互连在未适配成功前是不清楚的,但可以依靠 FPGA 在其内部灵活地处理,例如连接、断开或者采用某种终端方式来反复地实验调试,直到找到某种确定的处理方式或者结论为止。经过本文的实验,对于适配成功的芯片组 SR5690 + SP5100 和龙芯 CPU 在 HT 总线上可以找到如下正确的连接方式,如图 5 所示。

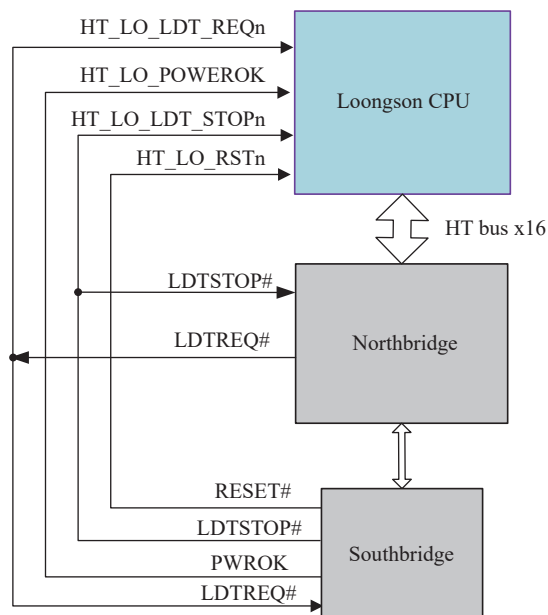


图 5 HT 重要暂不确定控制信号线经适配证明后找到的正确连接方式

Fig.5 Appropriate connection of the important but temporarily indeterminate HT signals after effective adaptation

2.2.3 上下电控制信号线的设计处理

在无法得知龙芯 CPU 和芯片组的上下电时序如何配合的情况下,需要把 CPU 上下电控制信号、芯片组的上下电控制信号、电源模块的控制信号线都引入 FPGA。这些信号线主要包括:(1)龙芯 CPU 的系统复位信号线 SYSRESETN,PCI 总线的复位信号线 PCI_RESETN;(2)南桥的上电完成信号线 PWR_GOOD,北桥的上电完成信号线 NB_PWRGD;(3)各个供电模块的上电使能信号线和电源完成信号线 PWRGOOD;(4)龙芯 CPU 的 GPIO 配置信号线,内存时钟频率的设置信号线,HT 总线时钟频率的设置信号线等。

用 FPGA 使能或者关闭电源模块的使能信号、龙芯 CPU 和芯片组的电源完成信号、复位信号线,并调整各个电源控制信号的先后顺序和时间间隔进行反复实验,一直尝试到得到确定的答案为止,或者能找到最佳的电源时序适配方案,或者证明两者无法兼容。

芯片组 SR5690 + SP5100 是本文适配成功的芯片组,图 6 是为其找到的已经验证成功的上下电时序,包含的信号有正 5 V standby 电源+5VSB,正 3.3 V standby 电源+3.3V_DUAL, FPGA 的开机信号线 FPGA_SB_RSMRST#,南桥的上电信号线 SB_PWRON,系统唤醒信号线 SLP_S3#_R,12 V 供电信号线 12V_SLP_S3#, ATX(Advanced technology extended)电源上电正常指示信号线 ATX_PWRGD,北桥 1.8 V 电源 NB_1V8,CPU 的 1.2 V 电源 CPU_NB_SB_1.2V,北桥 1.1 V 电源 NB_1V1,CPU 的 1.1 V

电源 CPU0/1_VCCINT_1V1, CPU 内存的 1.5 V 电源 CPU0/1_1.5V_LS3_DDR3, 系统上电正常指示信号线 SYS_PWRGD, 北桥上电正常指示信号线 NB_PWRGD, FPGA 的复位信号线 FPGA_KBRST_L, 龙芯 HT 总线上电正常指示信号线 LS3_HT1_LO_

POWEROK, 南桥复位信号线 SB_A-RST_L, 南桥 PCI 总线复位信号线 SB_PCI_RST_SR_L, 系统复位信号线 SYSTEM_RST, 以及 CPU HT 的复位信号线 CPU_LDT_RST. 为了进一步验证上下电时序的正确性和可靠性, 采用两种方法进行验证.

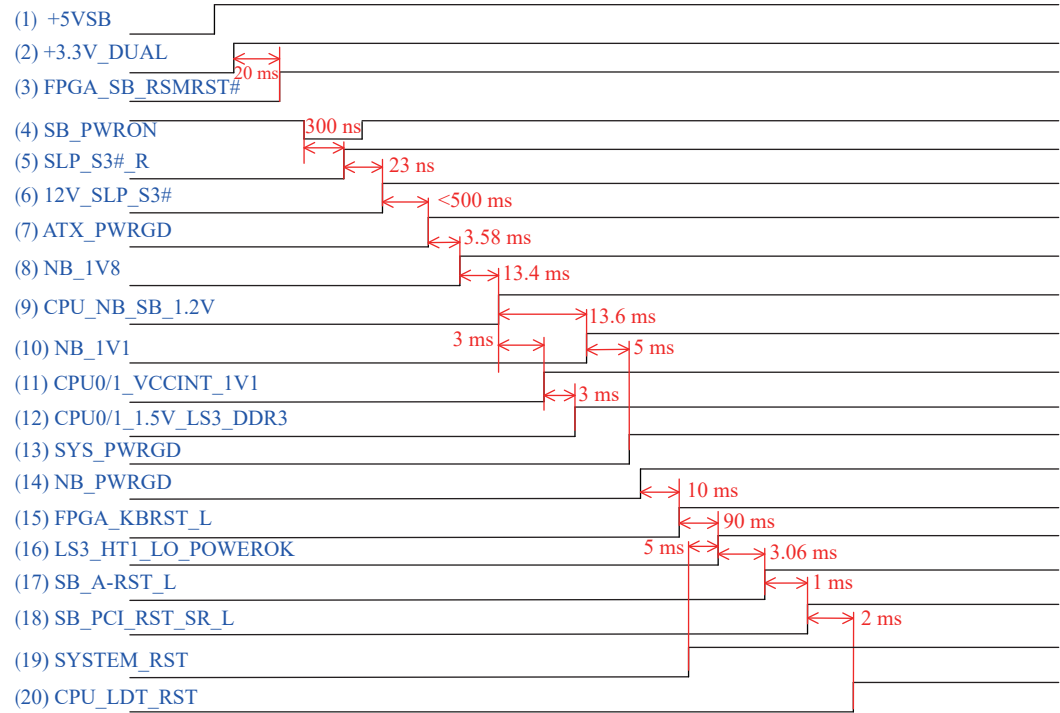


图 6 经适配证明后找到的正确电源时序

Fig.6 Correct power sequence after effective adaptation

在 ACPI(Advanced configuration and power interface)规范中规定了计算机系统的 7 种状态: G3(Mechanical off, 即彻底切断所有电源, 包括外置电源), S0(工作状态, 即所有设备都运行), S1(CPU 关闭工作状态, 其他的部件仍然正常工作), S2(CPU 停止工作状态, 其他的部件仍然正常工作), S3(除了内存外的部件都停止工作), S4(内存数据写入硬盘后所有部件都停止工作状态)和 S5(关机状态, 只有 standby 电源存在). 对于服务器系统来说, 其中必然经历三种状态, 即 G3、S5、S0. 鉴于此, 龙芯 CPU 和芯片组 SR5690 + SP5100 组成的服务器系统从启动到运行过程中, 通过调试工具在 G3、S5 和 S0 状态读取龙芯 CPU 和芯片组的寄存器值, 和产品文档中标准值进行对比从而验证上下电时序的正确性. 测试表明, 图 6 设计的上下电时序是正确的.

另一种方法是采用服务器产品化过程中常用的电源循环压力测试方法: AC Power on/off cycle test 和 DC Power on/off cycle test. AC Power on/off cycle test, 是交流电循环上下电测试, 从 G3 到 S0

再到 G3 再到 S0 不间断循环 1000 次, 如顺利完成则表明系统设计的正确性和可靠性. DC Power on/off cycle test, 是直流电循环上下电测试, 从 S0 到 S5 再到 S0 再到 S5 不间断循环 1000 次, 如顺利完成则表明系统设计的正确性和可靠性. 压力测试表明, 图 6 所示的上下电时序顺利完成了 1000 次循环压力测试, 达到了产品化的目标.

2.3 龙芯服务器产品的设计

当找到能够正确适配龙芯 CPU 的芯片组后, 在实际的龙芯服务器产品设计中, 去掉中间 FPGA, 按照适配好的规则直接把龙芯 CPU 和适配成功的芯片组连接起来.

在适配过程中已经总结出物理信号线的正确连接或者处理方式、正确的上下电时序、最佳的 HT 工作频率、确定的 PCB 布线规则(如线长、线宽、线距)和信号完整性的正确处理方式等设计要点. 在龙芯服务器产品设计中只需遵循已总结出的这些设计要点按照常规主板设计的方法进行设计就可以了.

图 7 是根据以上总结得出的设计规则去掉

FPGA 后设计的龙芯服务器主板实物图. 此龙芯服务器主板采用龙芯双路 SMP(Symmetrical multi-

processing)架构设计, 芯片组是 SR5690 + SP5100, 已经实现了产品化, 稳定性好, 性能高.

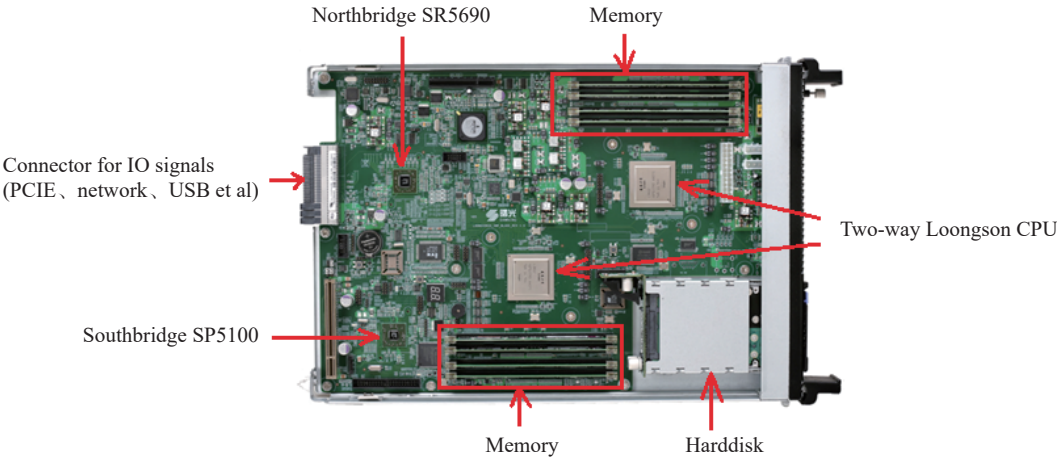


图 7 SR5690+SP5100 龙芯双路服务器产品主板

Fig.7 Loongson two-way SMP motherboard product using SR5690 + SP5100 chipsets

3 讨论

本文设计的筛选芯片组的架构和方法具有很强的通用性, 不但从如上五种芯片组中筛选出服务器级别的 SR5690 + SP5100 芯片组, 能够完全适配龙芯 CPU, 而且利用此架构和方法在其他的研究项目中也成功筛选出多种适用于龙芯 CPU 的桌面级别芯片组和服务器级别芯片组, 对于桌面级别芯片组, 例如北桥芯片型号有 RX781、RS785G、RD790, 南桥芯片型号有 SB700; 对于服务器级别芯片组, 例如北桥芯片型号有 SR5650、SR5670, 南桥芯片型号有 SB750. 对于这些已经成功筛选出的芯片组, 用户可以根据其功能、性能、成本等综合因素选择高性价比的芯片组来使用, 极大地拓宽了龙芯 CPU 的使用领域, 促进了龙芯的产业化.

SR5690 + SP5100 芯片组功能丰富、性能强大, 完全可以作为龙芯 CPU 服务器级别的芯片组使用, 其和目前龙芯中科技术有限公司推出的龙芯 7A1000 桥片在规格参数上的对比如表 4 所示.

表 4 芯片组规格对比

Table 4 Comparison of different chipset specifications

Item	Features of 7A1000	Features of SR5690 + SP5100
HT bus	HT3.0 × 16	HT3.0 × 16
PCIE	32 lanes	42 lanes
SATA	3 × SATA2.0	6 × SATA2.0
USB Ports	6 × USB2.0	14 × USB2.0
RAS	No	Yes
IOMMU	No	Yes

高 IO 带宽、支持 IO 虚拟化的 IOMMU 功能和支持 RAS 功能是目前服务器所应具备的基本要求. 从这三种芯片组规格的对比来看, 芯片组 SR5690 + SP5100 的 IO 带宽特别是 PCIE 数量远大于龙芯 7A1000 桥片, SATA 和 USB 数量也远大于龙芯 7A1000 桥片, 而且还支持 IOMMU 和 RAS 功能. 所以, 芯片组 SR5690 + SP5100 在规格参数上完全可以作为龙芯 CPU 的服务器芯片组.

在性能方面, 本文设计的 SR5690 + SP5100 双路龙芯 3B4000 服务器与市场上已有的双路龙芯 3B4000 + 7A1000 服务器进行了实际测试对比. 两种服务器都采用相同的硬件配置, 即两路龙芯 3B4000 CPU(主频 1.8 GHz, 4 核), 4 条紫光国芯 16 GB DDR4 内存(型号: SCC16GP02H1F1C-26V), 4 块东芝 2 TB 的 SATA 硬盘(型号 MG04ACA200N), 网卡采用 Intel I350 千兆网卡. 采用国际上公认的 benchmark 程序进行测试, 用 SPEC CPU 2006 测试了 CPU 计算性能, 单核整型性能(int_speed_base)、多核整型性能(int_rate_base)、单核浮点性能(fp_speed_base)和多核浮点性能(fp_rate_base), 如表 5 所示; 用 IOzone V4.3.0 测试了 SATA 硬盘性能, 512 Byte 读写性能和 1 MB 读写性能, 如表 6 所示; 用 Netperf V2.7.0 测试了网络性能, TCP(Transmission control protocol) 吞吐率、TCP 传输速度、UDP(User datagram protocol) 吞吐率和 UDP 传输速度, 如表 7 所示. 从性能对比的数据来看, 采用 SR5690 + SP5100 作为龙芯 3B4000 配套的芯片组可以较大幅度地提高龙芯 CPU 的性能, 特别是 IO 方面性能, 可以提高到 10% 以上.

表5 SPEC CPU2006 性能对比

Table 5 Analysis of SPEC CPU2006 performance				
Server	int_speed_ base	int_rate_ base	fp_speed_ base	fp_rate_ base
7A1000 server	12.30	78.07	12.02	74.90
SR5690+ SP5100 server	13.02	83.60	12.80	82.60
Performance improvement/%	6	7	6	10

表6 IOZone 性能对比

Table 6 Analysis of IOzone performance				
Server	512 Byte read speed/ (MB·s ⁻¹) (Average of three results)	1 MB read speed/ (MB·s ⁻¹) (Average of three results)	512 Byte write speed/ (MB·s ⁻¹) (Average of three results)	1 MB write speed/ (MB·s ⁻¹) (Average of three results)
7A1000 server	38.56	696.31	1.25	306.76
SR5690+ SP5100 server	43.19	800.76	1.53	383.45
Performance improvement/%	12	15	22	25

表7 Netperf 性能对比

Table 7 Analysis of Netperf performance				
Server	TCP Throughput/ (MB·s ⁻¹) (Average of three results)	TCP transfer rate/ (Times·s ⁻¹) (Average of three results)	UDP Throughput/ (MB·s ⁻¹) (Average of three results)	UDP transfer rate/ (Times·s ⁻¹) (Average of three results)
7A1000 server	850.51	8738.91	852.64	8999.10
SR5690+ SP5100 server	935.56	9787.58	946.43	9989.00
Performance improvement/%	10	12	11	11

4 结论

本文为龙芯 CPU 设计了一种筛选芯片组的架构,给出了龙芯 CPU 和芯片组之间适配的方法,并设计了筛选出的容纳多套芯片组的主板,对筛选出的芯片组进行了产品化,测试了真实的运行性能。

本文筛选出的芯片组 SR5690 + SP5100 无论在 PCIE、SATA、USB、IOMMU 和 RAS 等参数方面,还是在实际运行性能上都高于目前龙芯 CPU 所用的芯片组,所以非常适合作为龙芯 CPU 在服务器领域的芯片组,有力地开拓了龙芯 CPU 更多的适用领域。

参 考 文 献

[1] Hu W W. Developing our own CPU should take the road of marketing driven technology. *J Inf Secur Res*, 2019, 5(5): 450
(胡伟武. 发展自主CPU应该走市场带技术的道路. 信息安全研究, 2019, 5(5): 450)

[2] Ni G N. Adhere to the self-reliance and self-improvement of IT innovation system technology, build a powerful network country and digital China. *J Inf Secur Res*, 2021, 7(1): 2
(倪光南. 坚持信创科技自立自强建设网络强国和数字中国. 信息安全研究, 2021, 7(1): 2)

[3] Ma W, Yao J B, Chang Y S, et al. Current situation and prospect of CPU development in China. *Appl IC*, 2019, 36(4): 5
(马威, 姚静波, 常永胜, 等. 国产CPU发展的现状与展望. 集成电路应用, 2019, 36(4): 5)

[4] Xiong J, Xia Z P, Lin J, et al. Study of performance test scheme of information system based on domestic CPU and OS. *Comput Eng*, 2015, 41(12): 82
(熊婧, 夏仲平, 林军, 等. 基于国产CPU/OS的信息系统性能测试方案研究. 计算机工程, 2015, 41(12): 82)

[5] Zhang Z G, Zheng N B, Zhou Z F, et al. The research and design of office information system based on homemade software and hardware. *Comput Inf Technol*, 2012, 20(5): 8
(张忠革, 郑年斌, 周泽峰, 等. 基于国产CPU/OS的办公信息系统研究与设计. 电脑与信息技术, 2012, 20(5): 8)

[6] Hu X D, Yang J X, Zhu Y. Shenwei-1600: a high-performance multi-core microprocessor. *Sci Sin Information*, 2015, 45(4): 513
(胡向东, 杨剑新, 朱英. 高性能多核处理器申威1600. 中国科学:信息科学, 2015, 45(4): 513)

[7] Shen J, Long B, Jiang H, et al. Implementation and optimization of vector trigonometric functions on phytium processors. *J Comput Res Dev*, 2020, 57(12): 2610
(沈洁, 龙标, 姜浩, 等. 飞腾处理器上向量三角函数的设计实现与优化. 计算机研究与发展, 2020, 57(12): 2610)

[8] Fang J B, Du Q, Tang T, et al. Performance comparison between FT-1500A and Intel XEON. *Comput Eng Sci*, 2019, 41(1): 1
(方建滨, 杜琦, 唐滔, 等. 飞腾处理器与商用处理器性能比较. 计算机工程与科学, 2019, 41(1): 1)

[9] Hu X D, Ke X M, Yin F, et al. Shenwei-26010: A high-performance many-core processor. *J Comput Res Dev*, 2021, 58(6): 1155
(胡向东, 柯希明, 尹飞, 等. 高性能众核处理器申威26010. 计算机研究与发展, 2021, 58(6): 1155)

[10] Hong W J, Li K L, Quan Z, et al. PETSc's heterogeneous parallel algorithm design and performance optimization on the Sunway TaihuLight system. *Chin J Comput*, 2017, 40(9): 2057
(洪文杰, 李肯立, 全哲, 等. 面向神威·太湖之光的PETSc可扩展异构并行算法及其性能优化. 计算机学报, 2017, 40(9): 2057)

[11] Meng X F, Gao X, Cong M, et al. System performance optimization and analysis of Godson-3A multiprocessor. *J Comput Res Dev*, 2012, 49(Suppl 1): 137

- (孟小甫, 高翔, 从明, 等. 龙芯3A多核处理器系统级性能优化与分析. 计算机研究与发展, 2012, 49(增刊1): 137)
- [12] Zhao H, Wan J W, Bao Z G, et al. Application of independent and controllable technology in test missions. *J Spacecr TT& C Technol*, 2015, 34(2): 109
(赵辉, 万俊伟, 鲍忠贵, 等. 自主可控技术在试验任务领域的应用研究. 飞行器测控学报, 2015, 34(2): 109)
- [13] Yuan G X, Zhang Y Q, Yuan L. State of the art analysis of China HPC 2021. *Comput Eng Sci*, 2021, 43(12): 2091
(袁国兴, 张云泉, 袁良. 2021年中国高性能计算机发展现状分析. 计算机工程与科学, 2021, 43(12): 2091)
- [14] Cai F, Shen H H, Gao X. The design and implementation of north-bridge used in Godson-2 prototype system. *Chin High Technol Lett*, 2010, 20(1): 61
(蔡飞, 沈海华, 高翔. 龙芯2号原型系统北桥的设计与实现. 高技术通讯, 2010, 20(1): 61)
- [15] Liu D, Li X, Xu S Y, et al. Design and implementation of homemade information processing platform. *J Telem Track Command*, 2018, 39(6): 7
(刘达, 李鑫, 徐松艳, 等. 国产化信息处理平台设计与实现. 遥测遥控, 2018, 39(6): 7)
- [16] Zhu S S, Lu Y K, Liu L, et al. Design of AIO security computer based on Loongson CPU. *Ind Control Comput*, 2020, 33(11): 16
(朱书杉, 路永轲, 刘磊, 等. 基于龙芯处理器的一体式安全计算机设计. 工业控制计算机, 2020, 33(11): 16)
- [17] Zhao B, Yang M H, Liu W, et al. Research on security & trust computer based on Loongson CPU. *Comput Technol Dev*, 2015, 25(3): 126
(赵斌, 杨明华, 柳伟, 等. 基于龙芯处理器的自主可信计算机研究. 计算机技术与发展, 2015, 25(3): 126)
- [18] Wu J. *The Design of North-bridge Used in Godson System* [Dissertation]. Hefei: University of Science and Technology of China, 2003
(武杰. 龙芯系统中的北桥设计[学位论文]. 合肥: 中国科技大学, 2003)
- [19] Evans A, Silburt A, Vreckovnik G, et al. Functional verification of large ASICs // *Proceedings of the 35th annual Design Automation Conference*. New York, 1998: 650
- [20] Ganapathy G, Narayan R, Jorden G, et al. Hardware emulation for functional verification of K5 // *Proceedings of the 33rd Design Automation Conference*. Las Vegas, 1996: 315
- [21] Ray J, Hoe J C. High-level modeling and FPGA prototyping of microprocessors // *Proceedings of the 2003 ACM/SIGDA Eleventh International Symposium on Field Programmable Gate Arrays*. Monterey, 2003: 100
- [22] Li X B, Tang Z M, Li W. FPGA verification for heterogeneous multi-core processor. *J Comput Res Dev*, 2021, 58(12): 2684
(李小波, 唐志敏, 李文. 面向异构多核处理器的FPGA验证. 计算机研究与发展, 2021, 58(12): 2684)
- [23] Liu Y C, Wang J, Chen Y J, et al. Survey on computer system simulator. *J Comput Res Dev*, 2015, 52(1): 3
(刘雨辰, 王佳, 陈云霁, 等. 计算机系统模拟器研究综述. 计算机研究与发展, 2015, 52(1): 3)
- [24] Gateley J, Blatt M, Chen D, et al. UltraSPARC-I emulation // *Proceedings of the 32nd ACM/IEEE Conference on Design Automation Conference*. San Francisco, 1995: 13
- [25] Zhou S J, Prasanna V K. Accelerating graph analytics on CPU-FPGA heterogeneous platform // *2017 29th International Symposium on Computer Architecture and High Performance Computing*. Campinas, 2017: 137
- [26] Zhou H W, Xu S, Wang Z Y, et al. FPGA verification for memory link interface of many-core processor. *J Natl Univ Def Technol*, 2018, 40(3): 176
(周宏伟, 徐实, 王忠奕, 等. 众核处理器访存链路接口的FPGA验证. 国防科技大学学报, 2018, 40(3): 176)
- [27] Pang K, Shi Z F, Zhou J H, et al. Network topology exploration of coarse-grained reconfigurable architecture based on FPGA. *J Tianjin Univ Sci Technol*, 2018, 51(5): 507
(庞科, 史再峰, 周佳慧, 等. 基于FPGA的粗粒度可重构系统拓扑网络结构开发. 天津大学学报(自然科学与工程技术版), 2018, 51(5): 507)
- [28] Liu Y F, Liu P, Jiang Y T, et al. Building a multi-FPGA-based emulation framework to support networks-on-chip design and verification. *Int J Electron*, 2010, 97(10): 1241
- [29] Hu W W, Wang J, Gao X, et al. Godson-3: A scalable multicore RISC processor with x86 emulation. *IEEE Micro*, 2009, 29(2): 17
- [30] Kalla R, Sinharoy B, Tendler J M. IBM Power5 chip: A dual-core multithreaded processor. *IEEE Micro*, 2004, 24(2): 40
- [31] Kongetira P, Aingaran K, Olukotun K. Niagara: a 32-way multithreaded Sparc processor. *IEEE Micro*, 2005, 25(2): 21
- [32] Chen X M, Jha N K. A 3-D CPU-FPGA-DRAM hybrid architecture for low-power computation. *IEEE Trans Very Large Scale Integr (VLSI) Syst*, 2016, 24(5): 1649
- [33] Wang H D, Gao X, Chen Y J, et al. Interconnection of Godson-3 multi-core processor. *J Comput Res Dev*, 2008, 45(12): 2001
(王焕东, 高翔, 陈云霁, 等. 龙芯3号互联系统的设计与实现. 计算机研究与发展, 2008, 45(12): 2001)
- [34] Feng K K, Jia F, Du X J, et al. Design and realization of HT interconnection and memory fault diagnosis method for Loongson-3 mainboard. *Comput Meas Control*, 2020, 28(6): 1
(冯珂珂, 贾凡, 杜晓杰, 等. 龙芯3号板卡HT互联及内存故障诊断方法的设计与实现. 计算机测量与控制, 2020, 28(6): 1)