

多处理器数据交互体系架构研究与实现

边 庆¹ 田征戈¹ 张 炫¹

BIAN Qing TIAN Zhengge ZHANG Xuan

摘 要

随着应用场景对运算处理的要求愈发复杂,容错计算机在处理器性能层面持续迭代,为满足多样化的运算需求,不仅要配备多个处理器节点,各节点还需执行不同运算任务。于是,多处理器节点间高效的数据交互,便成为保障系统流畅运行的关键环节。基于此,文章提出了一种基于并行总线互连型架构、点到点全互连型架构的多处理器数据交互方法。主要介绍容错计算机系统总体设计,针对不同系统需求对两种体系架构分别进行研究、提出实现方法,实际测试数据表明,该系统性能稳定可靠。

关键词

多处理器;容错计算机;并行总线互连;点到点全互连;PCI总线

doi: 10.3969/j.issn.1672-9528.2025.04.002

0 引言

在高可靠容错计算机技术领域,容错计算机对处理器性能、可靠性要求较高,并且容错计算机任务功能越来越多。新型容错计算机往往具有多个处理器节点,执行多种不同的任务。传统容错计算机按任务功能分为多种计算机,每种计算机采用多余度机制保证可靠性,使得系统复杂、体积重量庞大、价格昂贵,更主要是缺乏灵活配置能力。随着时代的发展,容错计算机将多个功能融合在一个计算机中,必然使用多个处理器节点执行不同的功能。

多处理器节点间数据交互架构具有并行总线互连型架构、点到点全互连型架构,本文介绍容错计算机系统总体设计,针对系统需求对两种体系架构分别进行研究、提出实现方法。

1 系统总体设计

容错计算机系统主要由4个处理器模块组成,总体结构如图1所示。

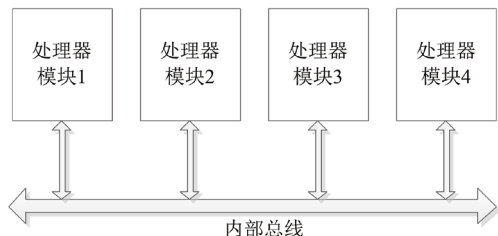


图1 容错计算机系统结构

容错计算机系统处理器模块之间通过内部总线进行数据

交互,内部总线可以是并行总线、串行总线,内部总线不同也决定了处理器之间的拓扑结构。

目前在嵌入式系统中,常用的设备间和板间互连高速串行总线主要有以太网、PCIe总线和RapidIO总线,3种总线采用的协议都较为复杂,存在传输延时、通信不确定的问题。此外,3种总线基于数据包传输,最大有效载荷长度均有限制,存在数据包开销。

在通信实时性、确定性要求较高的系统中,发送端应在确定时间范围内将数据传输到接收端;不仅支持数据包传输,也要支持控制指令的传输;对于数据包最大有效载荷长度也不应有限制,不应存在软件开销。本文采用Aurora协议作为传输协议,有效解决其他串行总线的传输效率低、实时性不高等问题。

因此,本文按照并行总线互连型架构、点到点全互连型架构两种进行详细阐述。

2 并行总线互连型架构

并行总线互连型架构硬件平台基于处理器和交换结构桥接器,桥接器具有CROSSBAR和PCI总线,通过CROSSBAR交换技术和高速PCI总线互连技术,多处理器节点进行DMA高速数据交互,不占用处理器资源,在系统要求任务周期内完成数据交互。

硬件平台采用交换结构的桥接器,桥接器利用内部的64位全双工数据通道来连接不同的单元,内部CROSSBAR体系结构决定了不同接口之间可进行并行数据交换。

硬件平台采用基于PCI总线方式实现多个处理器互连,如图2所示,处理器A作为PCI总线主设备,处理器B/C/D作为PCI总线从设备,4个处理器节点可通过PCI总线进

1. 中国航空工业集团公司西安航空计算技术研究所
陕西西安 710065

行高速数据交换，构建一个由多处理模块组成的计算机系统^[1]。

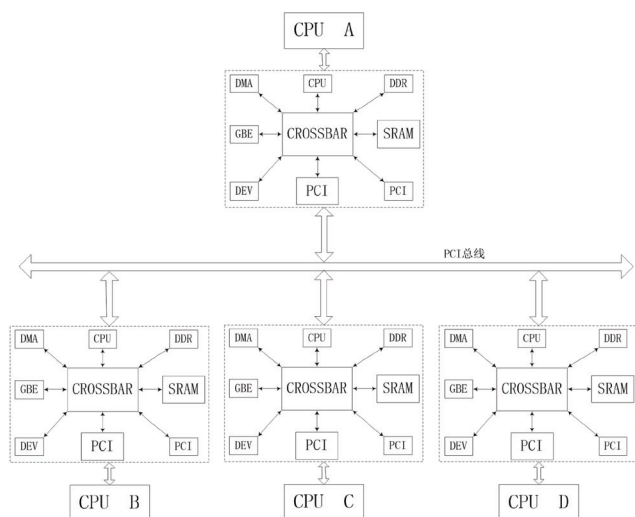


图2 并行总线互连型架构硬件平台结构图

桥接器具有2 MB片内SRAM存储器，作为多处理器的共享存储器，用于多处理器间的数据交互，如图3所示。处理器A负责PCI总线的初始化、管理等工作，通过PCI总线对处理器B/C/D片内SRAM进行访问。处理器B/C/D将PCI接口地址映射到片内SRAM，访问片内SRAM以本地内存的形式直接进行^[2]。

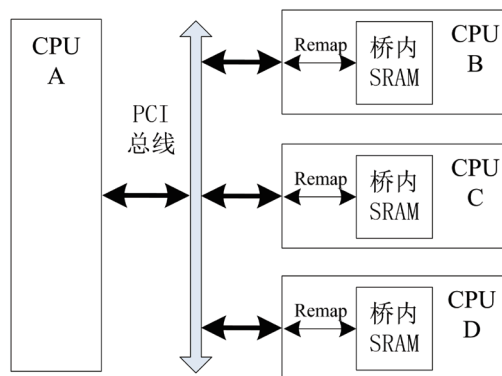


图3 多处理器节点PCI总线互连结构图

硬件平台4个处理器节点分别负责不同的运算控制功能，只有处理器A具有访问通信控制模块的能力，处理器A将处理器B/C/D数据向外转发。处理器B/C/D将本地片内SRAM存储器分为接收区和发送区，处理器A将外部数据写入处理器B/C/D的接收区，处理器B/C/D将发送数据写入发送区，可保证数据完整性，如图4所示。处理器A将任务周期分为3个时间区域，每个区域只能访问处理器B/C/D中一个，因此，既能在硬件上具有仲裁机制，又能在软件层次保证数据交互有效性^[3]。具体流程如下：

(1) 处理器A利用CROSSBAR交换结构将数据从DEV接口转发到DDR存储器，处理器A读取数据包标识将

数据再次转发，利用CROSSBAR交换结构从DDR存储器转发到PCI接口，此时外部数据发送到PCI总线，处理器B/C/D根据PCI总线协议，决定是否将数据进行读取。

(2) 处理器B/C/D向外部发送数据，只需将数据存储在本地桥内SRAM，处理器A将任务周期分为3个时间区域，每个时间区域处理器A只能通过CROSSBAR交换结构和PCI总线读取处理器B/C/D中一个，这样的话，既能在硬件上具有仲裁机制，又能在软件层次保证数据交互有效性。

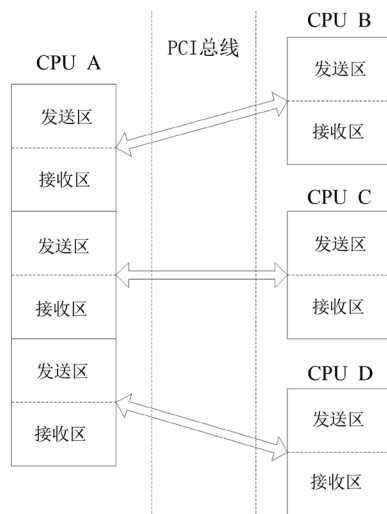


图4 多处理器节点软件分配空间

3 点到点全互连型架构

本文提出一种基于对等结构的多节点数据交互方法，采用多节点对等结构技术、高速串行总线技术和双端口存储技术，在有限任务周期内完成多节点间数据交互，多节点间采用点到点的传输链路，构建一个多处理器组成的容错计算机系统，如图5所示^[5]。

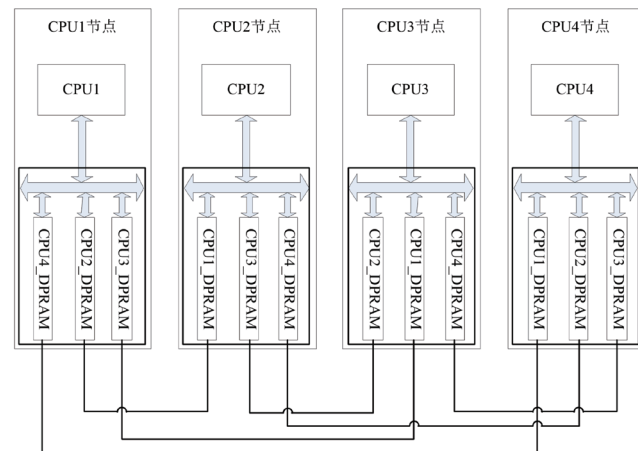


图5 基于对等结构的多节点数据交互结构图

硬件平台中处理器节点属于对等结构，节点间为点对点的方式传输，通过高速串行总线技术和双端口存储技术，多

节点进行 DMA 高速数据交互, 不占用处理器资源。

硬件平台包含 4 个处理器节点, 4 个处理器节点分别负责不同的运算控制功能, 处理器节点为点到点传输链路, 两者为对等结构, 不分主从, 均有能力发送和接收数据^[5]。

硬件平台每个节点具有不同数量的双端口存储器, 对应每个双端口存储器具有独立点到点的传输链路。每个处理器节点具有 3 个双端口存储器, 分别用于处理器间节点数据交互。另外多处理器间具有同步接口, 多处理器既能高速数据交互, 又能同步运行, 构建一个由多处理模块组成的高可靠容错计算机系统。

处理器 CPU1 节点双端口存储器分为接收区和发送区, 软件将发送数据

写入 DPRAM 发送区后, 软件可继续执行其他功能。硬件逻辑查询到 DPRAM 发送区数据已更新, 将该数据转发到串行总线逻辑部分, 串行总线逻辑包含数据链路层和物理层, 增加包头、包尾和校验码等信息后, 经过差分驱动器使用差分信号进行数据传输^[6]。

来自处理器 CPU1 节点的数据经过高速差分串行总线传输后, 硬件逻辑自动接收数据, 通过数据链路层和物理层, 去除包头、包尾和校验码等信息后, 保存在双端口存储器 CPU1_DPRAM 接收区^[7], 定时读取 CPU1_DPRAM 接收区数据, 再通过硬件逻辑发送到 CC 总线。高速串行总线发送和接收逻辑图如图 6 所示, 仅有数据链路层和物理层。

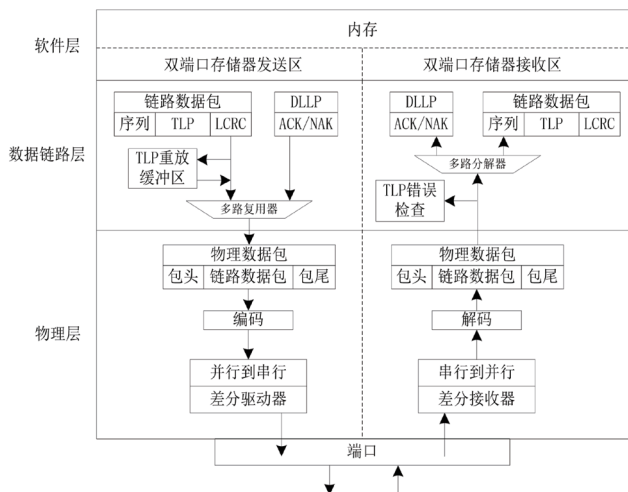


图 6 高速串行总线发送和接收逻辑图

另外, 其他节点与处理器 CPU1 节点的工作原理、应用方式均相同, 此处不再赘述。

每条点到点传输链路为高速差分串行总线, 使用 FPGA 芯片的高速 serdes 接口, 具备 8B/10B 编解码能力, 编码方式具有纠错、检错能力, 如图 7 所示。高速串行总线接口为全双工模式, 发送和接收可同时进行。软件仅需操作双端口存储器, 总线发送和接收数据操作都是使用硬件逻辑实现^[8]。

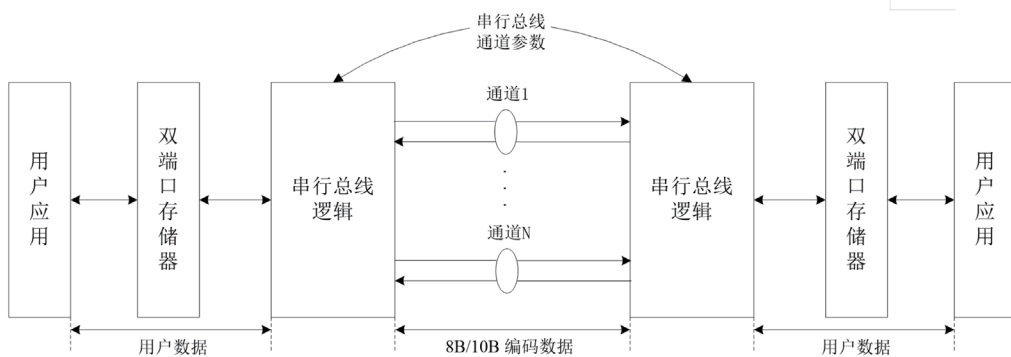


图 7 点到点传输链路结构示意图

4 并行总线互连型处理器模块设计

处理器模块采用 PC755CMGU400LE 作为其核心处理 CPU, 工作频率为 300 MHz; 桥接器采用交换结构桥接器, 桥接器利用内部的 64 位全双工数据通道来连接不同的单元, 内部体系结构决定了不同接口之间可进行并行数据交换, 内部数据交换可同时进行。桥接器内部 Crossbar 框图如图 8 所示。

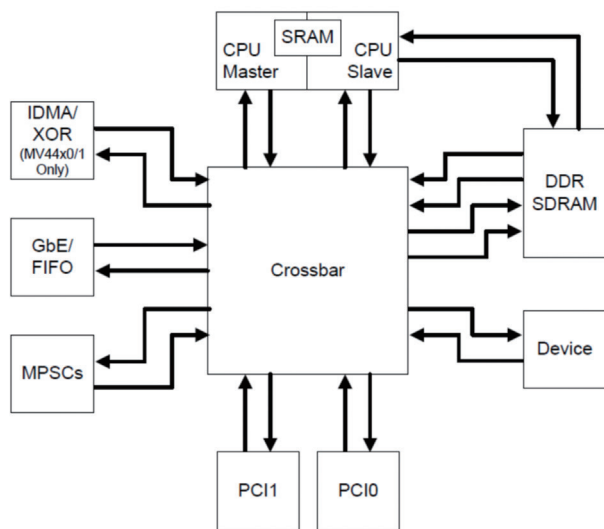


图 8 桥接器内部 Crossbar

桥接器内部 Crossbar 具有可编程仲裁机制, 避免多个接口访问同一资源的冲突, 多个请求不可能同时被响应。

硬件平台 4 个处理器节点通过 66 MHz/32 位 PCI 总线进行高速数据交换, 为了保证 PCI 总线信号完整性, 需采

用 PCI-PCI 桥设备拓展, 该设备支持透明模式、非透明模式和通用模式操作, 还支持一级端口和二级端口之间异步操作^[9]。

透明 PCI 桥是指 PCI 桥提供与系统的电气隔离, 允许总线上挂载更多的负载或设备, 还可以用来实现同一系统中运行不同的 PCI 总线宽度和速度。非透明 PCI 桥不仅提供电气隔离, 还提供地址隔离。桥两侧的设备都保留其独立的存储空间, 数据从桥的一侧传递到另一侧, 访问地址根据地址转换机制进行转换。

5 点到点全互连型处理器模块设计

处理器模块采用 SM755 作为核心处理器, 桥接器选用 SMPC107A 作为桥芯片, SM755 处理器 60X 总线和 SMPC107A 连接, 桥接器 SMPC107A 具有存储器总线和 PCI 总线, 通过存储器总线译码控制 SDRAM 和 Flash, 通过 PCI 总线转为局部总线控制串口、以太网接口、高速串行总线接口和 NVRAM。

本文高速串行总线采用 Aurora 协议, 该协议是 FPGA 芯片自带的针对高速传输开发的一种可裁剪的轻量级链路层协议。Aurora IP 核是 Xilinx 公司在 Aurora 协议和高速串行收发器 Rocket 基础上研发出来的硬核。Rocket I/O 支持多种通信协议标准, 支持 400 Mbit/s ~ 3.125 Gbit/s 的数据传输速率, 提供全双工或单工的数据传输模式。

一个完整的 Rocket I/O 收发器主要包括 PCS (物理编码子层) 和 PMA (物理媒介适配层) 两个子层, 其内部结构如图 9 所示。PMA 子层主要用于线路时钟恢复单元、高速时钟产生器、发送和接收缓冲区和串行器和解串器; PCS 子层主要包括 CRC 校验、8B/10B 编解码、时钟修正的弹性缓冲、Comma 检测对齐和通道绑定^[10]。

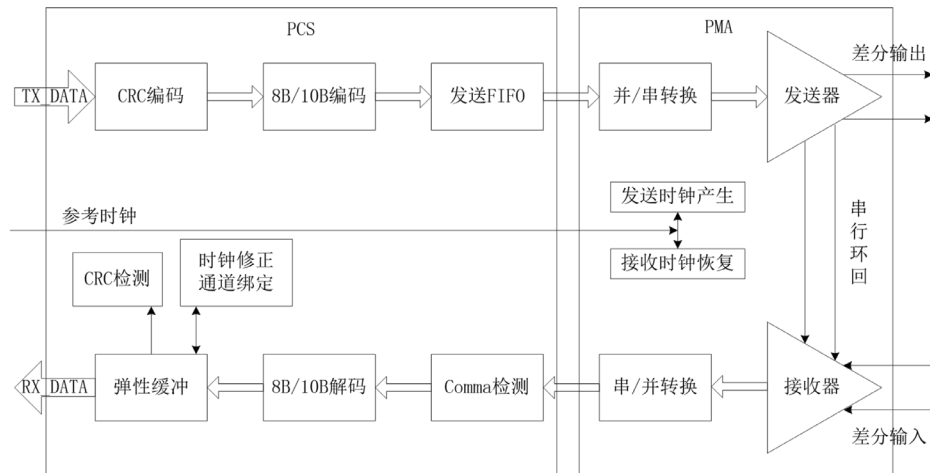


图 9 Rocket I/O 收发器内部结构

本系统采用自定义协议的高速串行通信系统, 包括: 发送部分、接收部分、数据解析模块、数据分发模块。

6 结论

本文基于并行总线互连型架构、点到点全互连型架构的多处理器数据硬件平台, 详细介绍了两种架构的容错计算机系统总体设计、处理器模块设计, 经仿真和硬件测试表明, 该方案针对系统需求对两种体系架构分别进行研究, 充分发挥了两种方案的技术优势, 已在机载计算机中成功应用, 对同类型嵌入式系统硬件设计具有较高的参考价值。

参考文献:

- [1] 张伟, 张祖良, 邹志强, 等. 基于共享内存的 PowerPC 系统 BSP 设计 [J]. 计算机工程, 2008,34(S1):11-12.
- [2] 李承伟, 杜鹃, 李波, 等. 多 PowerPC 7400/7410 处理器体系架构研究 [J]. 计算机工程, 2005,31(6):189-191.
- [3] 张旭. 基于 PCI 总线的多处理器协同机制研究 [J]. 科技风, 2017 (20): 68-71.
- [4] 张中华. 基于双 PowerPC7447A 处理器的嵌入式系统硬件设计 [J]. 现代电子技术, 2008,31 (24): 9-13.
- [5] 王齐. PCI Express 体系结构导读 [M]. 北京: 机械工业出版社, 2010.
- [6] 焦新泉, 杜凯旋, 杨志文, 等. Aurora 协议的高速光纤存储系统优化设计 [J]. 单片机与嵌入式系统应用, 2022, 22(4): 88-91.
- [7] 李维明, 陈建军, 陈星铸. 基于 Aurora 协议的高速通信技术的研究 [J]. 电子技术应用, 2013, 39(12): 37-40.
- [8] 岳振, 吕波, 张涌, 等. 基于 Aurora 协议的高速图像传输和通信平台设计 [J]. 电子技术应用, 2014,40(8):7-9.
- [9] 白曦, 张鹏. 基于 PowerPC755 的嵌入式数据采集系统

的设计与实现 [J]. 信息通信, 2018(5): 141-142.

- [10] Xilinx Corp. Aurora 8B/10B v11.0 LogiCORE IP Product Guide[Z]. 2016.

【作者简介】

边庆 (1986—), 男, 河北衡水人, 硕士, 高级工程师, 研究方向: 嵌入式计算机系统结构。

(收稿日期: 2025-01-10)