

doi: 10.7690/bgzd.2025.12.010

国产高性能无阻塞 RapidIO 交换机设计与实现

阮翔, 牛文举, 蒋辰浩, 刘诗畅, 谢龙兵

(中国电子科技集团公司第五十二研究所, 杭州 311100)

摘要: 为解决并满足国产化设计需求, 设计一款 20 路 SRIO x4 高性能无阻塞全互连交换机。基于龙芯处理器 LS2K0500、复旦微 FPGA 和国产 NRS1800 交换芯片, 搭配设计的交换机管理软件实现枚举组网与路由均衡分配, 并对交换机无阻塞进行测试。结果表明: 该交换机读写速率可达到 1 600 MB/s, 满足 SRIO 交换的国产化需求。

关键词: non-blocking; 网络交换; NRS1800; SRIO

中图分类号: TP312 **文献标志码:** A

Design and Implementation of Domestic High Performance Non-blocking RapidIO Switch

Ruan Xiang, Niu Wenju, Jiang Chenhao, Liu Shichang, Xie Longbing

(No. 52 Research Institute of China Electronics Technology Group Corporation, Hangzhou 311100, China)

Abstract: In order to solve and meet the requirements of domestic design, a 20-way SRIO x4 high-performance non-blocking fully interconnected switch was designed. Based on the Loongson processor LS2K0500, Fudan Microelectronics FPGA, and the domestic NRS1800 switching chip, the designed switch management software was used to achieve enumeration networking and balanced routing distribution, and the non-blocking performance of the switch was tested. The results showed that the read and write speed of the switch can reach 1 600 MB/s, meeting the localized requirements for SRIO switching.

Keywords: non-blocking; network switching; NRS1800; SRIO

0 引言

近年来, 随着国产芯片研发实力的崛起以及中国军事力量逐渐壮大, 国外开始对中国技术水平进行限制。由于芯片限制, 致使国产设备在某些高端领域出现无芯可用。为突破国外芯片“卡脖子”及打破国外技术封锁, 国内大力研发国产芯片, 攻破关键技术。然而在关键技术研发中, 高速数据传输是所有技术发展的关键技术基础。

目前, 高速数据传输技术主要包括 PCIe、SRIO 和万兆网络等, 其中 SRIO 是一种高可靠性、高性能的嵌入式数据传输系统, 是基于包交换的新一代高速互连技术^[1-2]。SRIO 技术是由 Motorola 等公司率先倡导的一种高带宽性能、传输延时小、低引脚数、基于数据包交换的互连体系结构^[3-5]。其高效的传输和易扩展的互连结构可应用于芯片到芯片、板间和背板互连, 满足多种场景下的通信需求^[6]。SRIO 交换技术可通过 SRIO 路由配置构建一个星型或网状的拓扑结构, 实现系统内所有节点任意、高速的数据传输^[7]。

前期 SRIO 的研究大部分是基于国外进口 SRIO 交换芯片 CPS1848 的应用研究^[8]。国内于 2018 年

成功研制出首款具有自主知识产权的 SRIO 交换芯片 NRS1800^[9]。SRIO 交换机是整个 SRIO 交换网络的核心单元, 关系到整个网络的带宽性能、时延和数据传输质量等^[10]。笔者基于目前国产芯片研究状态, 设计出一款国产高性能无阻塞 RapidIO 交换机, 并对此板卡进行了相关测试, 性能满足目前对高速数据传输以及全国产化的需求。

1 总体组成及工作原理

SRIO 交换机系统结构如图 1 所示, 交换机硬件主要由管理单元、交换网络和时钟电源电路组成。管理单元由 CPU+FPGA 构成, 其中 CPU 采用的是 LS2K0500 处理器, FPGA 选用的复旦微 JFM7K325T8, 处理器通过 PCIe 与 FPGA 相连, 通过 FPGA 逻辑设计扩展出 SRIO 接口与交换网络通信, 管理单元承担交换网络的管理功能。管理单元同时还承担 BMC 的功能, 监测模块内的电流、电压、温度, 对外提供 IPMI 管理接口。

交换网络中交换芯片采用的是国产 NRS1800 SRIO 交换芯片, 通过 5 个 NRS1800 芯片互连组成星型网络, 内部提供 20 路 SRIO x4 互联链路, 对

收稿日期: 2024-10-22; 修回日期: 2024-11-22

第一作者: 阮翔(1980—), 男, 安徽人, 硕士。

外提供 20 路 SRIO x4 接口。

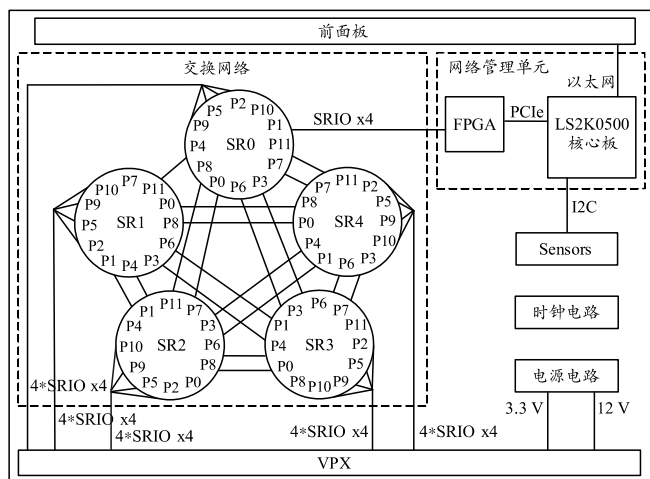


图 1 SRIO 交换机系统结构

交换机配套有管理软件，该软件负责 SRIO 交换网络的组网与管理，主要功能包括枚举组网、路由均衡分配、网络监控和速率自适应等。软件采用 C/S 架构，其中客户端运行在上位计算机中，提供 UI 人机界面。服务端运行在 LS2K0500 处理器中，通过以太网与客户端通信。

2 硬件设计

2.1 网络管理单元

2.1.1 CPU 电路设计

交换机的网络管理及维护是通过 LS2K0500 芯片完成的。LS2K0500 基于自主核心技术，是一款高集成度处理器芯片，面向工控互联网和 BMC 应用等领域。该处理器内部集成了 32 位 DDR2/3 控制器、2 路 PCIe2.0、2 路 SATA2.0、2 路 GMAC 及其他常用接口。LS2K0500 处理器模块是一个处理器最小核心系统，该核心系统包含处理器 LS2K0500 芯片所能工作的最小单元。该处理器模块中主要包含 LS2K0500 芯片、DDR 颗粒、Flash、PHY 芯片和电源等。LS2K0500 处理器模块通过连接器扣在交换机载板中，其应用如图 2 所示。

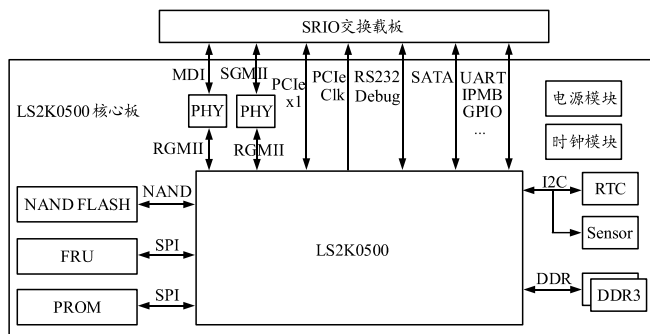


图 2 LS2K0500 处理器模块应用

2.1.2 FPGA 电路设计

FPGA 是一种可编程器件，通过编程可以在芯片上实现 PCIe 总线接口及 SRIO 总线接口与外界通信。该设计中 FPGA 起到 PCIe 与 RapidIO 之间数据交互桥梁的作用；因此 LS2K0500 可通过 FPGA 上实现 PCIE 转 SRIO 的功能，并与 SRIO 网络中的其他节点通信，为控制管理交换芯片建立物理通路基础。FPGA 电路设计如图 3 所示。

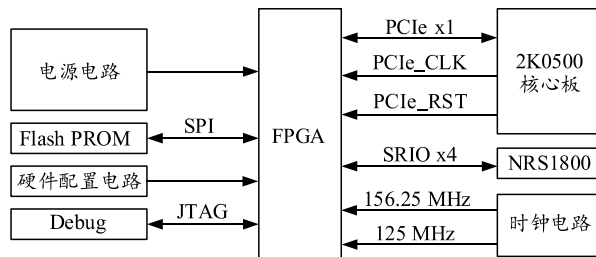


图 3 FPGA 电路设计

2.2 交换网络

SRIO 交换单元采用的是国产自主可控的 NRS1800 芯片。芯片具有 48 个高速 SerDes 通道，芯片支持 4x, 2x, 1x 模式的端口宽度配置。设计中 5 颗 NRS1800 交换芯片外围电路基本相似，本节以 NRS1800_0(SR0)为例进行阐述。

如图 4 所示，NRS1800 外围电路主要由复位电路、时钟电路、电源电路和配置电路等组成。

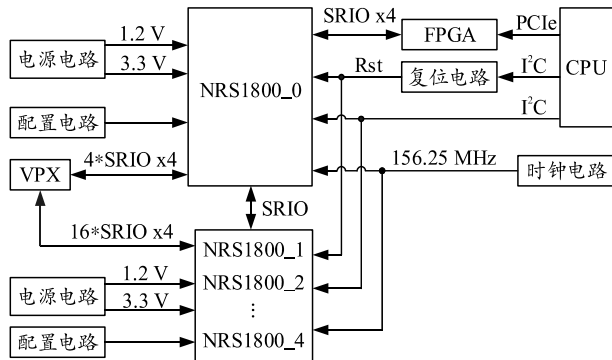


图 4 NRS1800 芯片外围电路设计

NRS1800 的复位信号分为硬复位和软复位，硬复位信号连接至板内复位电路，由 CPU 进行统一上电管理复位；软复位由芯片的 I²C 接口完成，同时，I²C 接口也可进行芯片初始配置和状态监控等功能。NRS1800 的参考时钟由时钟缓冲器提供时钟源。电源电路主要为 NRS1800 交换芯片提供 1.2 和 3.3 V 高精度电压，以满足芯片稳定工作需求。芯片外围配置电路，主要是系统内部时钟配置、端口速率配置、端口宽度配置以及 I²C 接口主从模式、地址、地址宽度等配置。

设计中 5 颗 NRS1800 芯片互连组成星型网络, 每颗芯片配置为 12 路全交换无阻塞 SRIO 5G x4 接口。如图 1 的 SR1~SR4, 每颗交换芯片通过 2 路 SRIO x4 进行互连后, 将其剩余 4 路 SRIO x4 引至连接器对外端口。对于 SR0 交换芯片, 由于需要引出一路 SRIO x4 与 FPGA 连接, 在与 SR2~SR4 互连之后, 将剩余的 5 路 SRIO x4 中, 其中 4 路引至连接器对外端口, 1 路与 SR1 进行互连。5 颗 NRS1800 芯片由此组成的交换网络可对外提供 20 路 4×SRIO 数据交换业务。

2.3 供电设计

对板卡所用芯片工作时所需的电压及最大工作电流进行梳理与分析, 表 1 列举了板卡主要元器件的电源需求。

表 1 主要元器件电源需求情况

芯片型号	数量	电压网络	电流/A
GD32F407	1	3.3V_STBY	0.1
1800	5	3.3V_VDD	2.6
		1.2V_VDD	63.8
		3.3V_VDD	0.5
FPGA	1	1.8V_VDD	0.8
		1.2V_VDD	1.0
		1.0V_VDD	10.0
		3.3V_STBY	1.5
2K500 核心板	1	3.3V_STBY	1.5
其他	-	3.3V_VDD	1.0

由于供电电压为 12 V 供电, 在电源设计中, 需要将输入电压进行降压以满足芯片元器件供电需求, 因此电源芯片选用的是长工微电子的 IS6608A 和 IS6607A 2 颗 DC-DC 降压芯片。设计中选用 3 颗 IS6608A 并联输出 1 路 1.2V_VDD, 选用 3 颗 IS6607A 电源模块分别输出 3.3V_STB、1.8V_VDD 和 3.3V_VDD, 以满足板卡所需电压和电流需求。

对于电源上电时序, 可通过电源芯片的 PG 信号输入到电源芯片的 EN 控制引脚 RUN, 从而控制整板的上电时序, 并在电源芯片的 RUN 使能端设计 RC 电路调整上电间隔时间。

2.4 时钟电路设计

设计中需要 5 片 NRS1800 芯片同步使用, 因此需要 5 路 156.25 MHz 同源的差分时钟, 芯片通过该时钟为芯片内部的 PLL 提供参考时钟, 用于倍频生成波特率时钟^[5]。由于 LS2K0500 芯片的 PCIe 接口作为 RC 模式, 因此 FPGA 的 PCIe 参考时钟可由 LS2K0500 提供。为了满足 FPGA 的 SRIO 接口适应不同的 SRIO 速率需求, FPGA 还需接入 125 和 156.25 MHz 差分时钟; 因此, 156.25 MHz 差分时钟总共需要 6 路, 设计中采用差分晶振产生时钟源,

并通过时钟驱动器输出 6 路高精度时钟。FPGA 的 1 路 125 MHz 时钟, 采用一颗振荡器为其提供时钟。其余芯片的单端时钟均由振荡器提供。

3 软件设计

3.1 交换机管理 UI 设计

为了更好地维护管理 SRIO 网络, 笔者研究了一款交换机管理软件客户端。该客户端软件基于 QT 设计实现, 提供一键组网功能, 组网完成后通过动态实时绘制的图元展示 SRIO 交换网络的拓扑连接关系, 如图 5 所示。UI 展示的方法相比用打印信息展示网络拓扑的传统方式, 对交换机的使用者更加友好直观, 大幅度降低了学习成本。

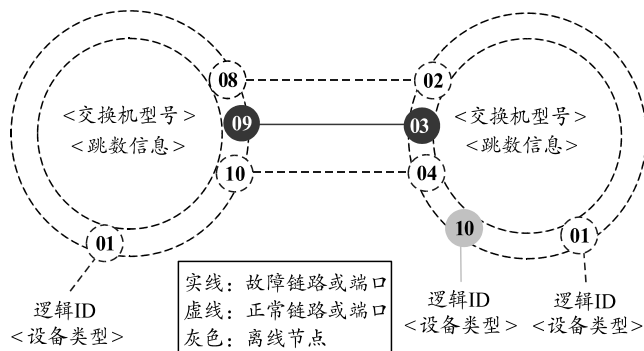


图 5 交换网络 UI 设计

3.2 管理软件核心功能设计

交换机管理软件核心功能由运行在 LS2K0500 处理器中的服务端程序实现。交换机上电后服务端程序会自主执行枚举组网。组网软件会使用本文中深度优先递归搜索算法对交换网络进行遍历, 得到网络的完整拓扑连接关系并存储在数据结构中。软件在枚举过程中会为所有节点分配目的 ID, 此时网络中分配的路由是最小路由, 该路由仅沿网络的最小生成树路径分配, 树根节点为 LS2K0500 管理节点。最小路由仅能保证节点间互相可通信。

为进一步实现无阻塞交换网络, 软件在枚举扫描完成后会基于本地存储的网络拓扑数据结构, 使用本文中算法^[11]对路由进行均衡规划。该路由分配算法将网络中的各节点目的 ID 均匀分布在各级交换机间的并行链路中。算法在保证均衡度的同时使用路径搜索算法使节点间的平均路由转发跳数最小(路径最短), 同时可为特定节点间的大数据流规划独立通信专线, 专线独占路线上的传输带宽。SRIO 路由分配流程如图 6 所示。路由算法运行结束后软件将新的路由配置到网络中的各交换机中, 至此无阻塞交换网络构建完成。

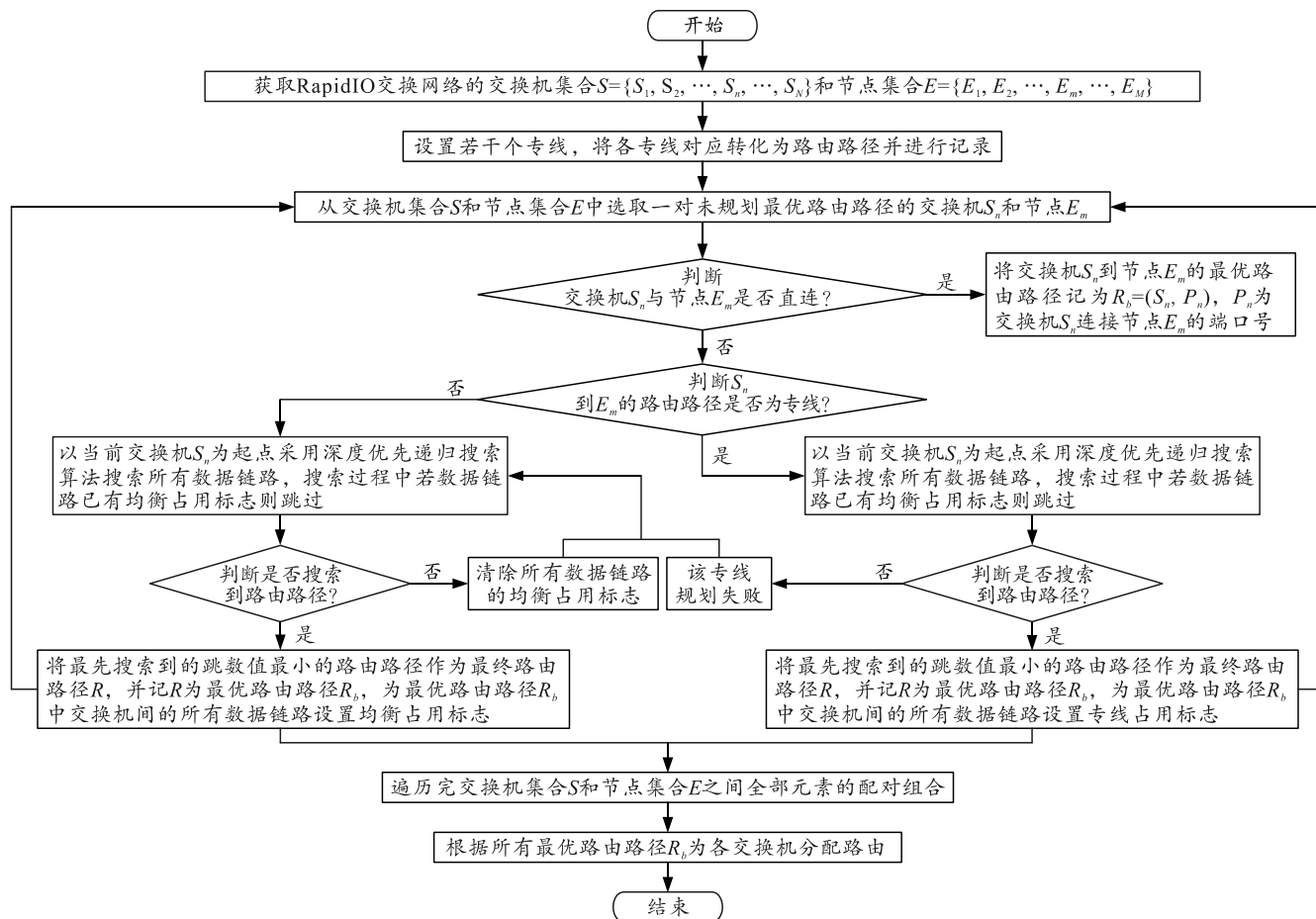


图 6 SRIO 路由均衡算法流程

4 功能以及性能验证

为验证 SRIO 交换机无阻塞数据传输性能，采用 4 个计算刀片和一台 PC 机搭建一个测试平台，如图 7 所示。每块计算刀片中包含 4 个 RapidIO 节点分别连接到交换机的 4 个端口。测试中使用计算刀片 1、2 中的 8 个设备节点作为数据发送端，使用计算刀片 3、4 中的 8 个设备节点作为数据接收端测试 NWRITE、NREAD 事务的数据传输速率。

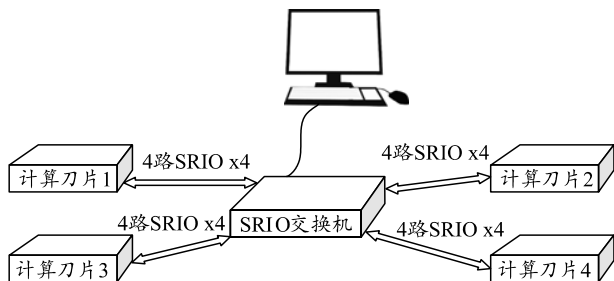


图 7 SRIO 交换测试平台

测试步骤：

1) 将 SRIO 交换机的网络接口连接到 PC 机，通过 PC 机上的交换机管理软件对交换网络进行组网，组网结果如图 8 所示。

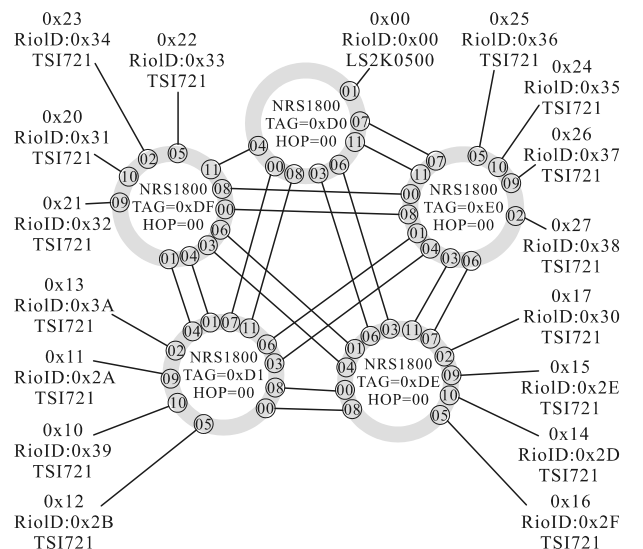


图 8 组网扫描结果

2) 在管理软件中根据测试场景设置 8 对通信专线并执行路由均衡规划，规划完成后各发送端与接收端之间的路由转发路径变为独占路径，均拥有 20 Gbaud 带宽的物理链路。

3) 通过速率测试程序在 8 个数据发送端同时向 8 个数据接收端发送测试数据，单次数据发送量

为 64 MB，测试时长 10 min。

在此 8 进 8 出无阻塞测试环境下，与采用进口 CPS1848 芯片的交换机进了对比测试，测试数据如表 2 所示。在 8 组同时测试下，设计的国产交换机

每个节点读写速率均在 1 600 MB/s 以上，且均能够正确并长时间稳定工作。进口交换机(CPS1848)读写速率在 1 500 MB/s 以上，设计的国产交换机性能优于进口交换机。

表 2 SRIO 交换机性能测试数据

操作方式	方向	端口宽度	节点 ID	NRS1800 实际速率/(MB/s)	CPS1848 实际速率(MB/s)
WRITE	计算刀片 1→计算刀片 3	4x	0x20-0x24	1 603	1 578
			0x21-0x25	1 642	1 588
			0x22-0x26	1 622	1 596
			0x23-0x27	1 638	1 532
NWRITE	计算刀片 2→计算刀片 4	4x	0x10-0x14	1 602	1 542
			0x11-0x15	1 611	1 533
			0x12-0x16	1 652	1 591
			0x13-0x17	1 637	1 566
NREAD	计算刀片 1←计算刀片 3	4x	0x20-0x24	1 623	1 531
			0x21-0x25	1 639	1 523
			0x22-0x26	1 607	1 589
			0x23-0x27	1 612	1 590
NREAD	计算刀片 2←计算刀片 4	4x	0x10-0x14	1 608	1 542
			0x11-0x15	1 625	1 536
			0x12-0x16	1 644	1 522
			0x13-0x17	1 618	1 568

设计中，NRS1800 的端口链路速率为 5.0 Gbaud，端口宽度 4x，理论速率为 2.5 GB/s，考虑物理编码开销 80%，RIO 协议开销 90%，实际理论速率 $2.5\text{ GB/s}\times 80\%\times 90\%=1.8\text{ GB/s}$ 。在测试中，还有数据收发端门铃交互的时间开销，交换机之间的数据交换开销(错误重传/流控)等，1.6 GB/s 的测试速率已接近物理带宽所能达到的理论上限。

5 结束语

笔者以国内自主研发的 SRIO 交换芯片 NRS1800 为核心，设计一款实现全国产化 20 路 SRIO x4 高性能无阻塞交换机，阐述了交换机的组成以及工作原理，对交换机的硬、软件设计进行了详细介绍。制作出实物样机，完成对样机硬件调试与软件开发，并对实物样机进行了 NREAD、NWRITE 无阻塞读写测试，功能性能达到了设计的预期目的，满足 SRIO 交换的国产化需求，对国产化的高速数据传输技术设计具有一定的参考意义。

参考文献：

[1] 吕鹏, 赵润卓, 王春燕, 等. 基于中标麒麟操作系统的 SRIO 路由设计与实现[J]. 无线电工程, 2021, 51(6):

492-497.
[2] 胡孔阳, 韩琮磊, 顾大晔. 一种 SRIO 交换器内部网络设计[J]. 微电子学与计算机, 2018, 35(9): 4.
[3] 雷志勇, 陈莹莹, 李兴明, 等. 一种基于 RapidIO 交换网络的多播路由方法及电子设备: CN111131018A[P]. 2020-05-08.
[4] 胡孔阳, 韩琮磊, 顾大晔. 一种 SRIO 交换器内部网络设计[J]. 微电子学与计算机, 2018, 35(9): 4.
[5] 陈刚, 康林, 陈航, 等. 基于 FPGA 的 SRIO 端点设计与实现[J]. 兵工自动化, 2021, 40(2): 49-52.
[6] 任勇峰, 多卉枫, 武慧军. 基于 FPGA 的多通路 SRIO 数据传输设计[J]. 电子测量技术, 2022(14): 45.
[7] 于东英, 陈俊, 康令州. 基于国产 FPGA 的高速 SRIO 接口设计与实现[J]. 通信技术, 2019, 52(1): 4.
[8] 李荣乐, 罗长洲, 李龙华, 等. 基于 FPGA 的交换机芯片配置器设计[J]. 计算机仿真, 2020, 37(2): 6.
[9] 王浩宇, 李雨航, 李龙杰, 等. 基于 NRS1800 的 SRIO 互连技术的研究与设计[J]. 机电产品开发与创新, 2021, 51(6): 28.
[10] 顾双豪, 韩崇伟, 王天石, 等. 基于 ARM 的国产化以太网交换机设计[J]. 兵工自动化, 2021, 40(11): 7.
[11] 蒋辰浩, 叶茂, 周细平, 等. 一种 RapidIO 交换网络的路由自动规划方法: CN115208820A[P]. 2022-10-18.