

文章编号: 2096-1618(2023)02-0160-06

# 基于 SATA3.0 的高速大容量存储系统设计

罗艺灵, 杜雨谔, 李一杰

(成都信息工程大学电子工程学院, 四川 成都 610225)

**摘要:**在通信技术快速发展的背景下,数据传输呈现速度快、容量大、带宽高的特点,为解决由此带来的数据缓存问题,设计一款基于 SATA3.0 协议的高速大容量存储系统。采用 Xilinx-Virtex7 系列 XC7VX690TFFG1158 作为控制核心,SRIO 协议缓存前级高带宽容量的数据,同 SSD 的数据通信采用 SATA3.0 协议,同上位机之间的数据通信采用 RGMII 协议。利用 FPGA 并行处理数据的优点,应用 RAID\_0 阵列实现并行控制 8 块 SSD。经实际测试,存储系统平均写入速度大于 20 Gb/s,整个 SSD 阵列平均写入速度可达 29.2 Gb/s,平均读取速度可达 35.7 Gb/s,在卫星通信等高速数字信号处理领域有广泛应用前景。同时为高速大容量存储系统的设计提供一定的参考意义。

**关键词:**FPGA;SATA3.0;SRIO;RAID\_0 阵列;GTH 收发器

**中图分类号:**TN914

**文献标志码:**A

**doi:**10.16836/j.cnki.jcuit.2023.02.006

## 0 引言

通信、嵌入式、现代雷达等领域需处理的数据量日渐增加,传输的数据呈现速度快、容量大、带宽高的特点,后级数字信号处理模块由于算力等原因无法实时匹配采集设备的数据率。因此设计一款能及时缓存高带宽大容量数据,为后级数字信号处理模块提供缓冲,从而解决高速度、大容量、高带宽数据的实时性处理问题的存储系统十分必要。近年来针对高速数据传输场景,涌现了多种高速传输技术<sup>[1]</sup>,如 Series Rapid IO<sup>[2]</sup>、PCI Express<sup>[3]</sup>、SATA(serial advanced technology attachment)和以太网等。白雪等<sup>[4]</sup>设计实现一款阵列控制器,存储容量为 4 TB,写入带宽为 26.4 Gb/s,读出带宽为 27.2 Gb/s。刘光祖等<sup>[5]</sup>提出一种基于 Rapid IO 协议的光纤通信方案,数据吞吐率为 1.25 Gb/s。徐玉杰等<sup>[6]</sup>提出一种基于 PCIE 接口的高速大容量存储系统,存储速度大于 12 Gb/s,但 PCIE 开发周期长、成本较高。张珮等<sup>[7]</sup>应用 RAID\_0 阵列于存储系统后,系统的读取速度为 3.72 Gb/s,约为单 SSD 的两倍,写入速度为 2.6 Gb/s,约为单 SSD 的 1.5 倍。本存储系统容量可达 8 TB,存储系统写入速度大于 20 Gb/s,整个 SSD 阵列平均写入速度可达 29.2 Gb/s,平均读取速度可达 35.7 Gb/s。

## 1 协议简介

SATA 3.0 理论传输速度为 1.5 ~ 6 Gb/s,使用 FP-

GA 实现的 SATA3.0 物理层传输速率近 6 Gb/s<sup>[8]</sup>。实际应用中采用 8b/10b 编/解码算法,其最高传输速度为 4.8 Gb/s。PCIE 单根信号线数据传输率达 2.5 Gb/s。通过增加信号线对,可以线性地扩展 PCI Express 的带宽,物理层可支持 X1、X2、X4、X8、X16 以及 X32 路带宽<sup>[9]</sup>。SATA3.0 协议是最成熟的 SSD 形式,同时考虑到开发周期和后期运行维护,最终采用 SATA3.0 协议。

SATA 结构如图 1 所示,SATA3.0 协议定义了一个高速串行 ATA 数据链路接口,包括 PHY 层、链路层、传输层和应用层<sup>[10]</sup>。应用层实现 ATA 命令的执行;传输层负责将应用层的数据封装发到链路层,将链路层的数据进行解析帧传到应用层;链路层对数据进行 8b/10b 编解码、加解扰、CRC 校验等;物理层负责与电气相关的特性,使数据通过差分线稳定的传输<sup>[11-12]</sup>。

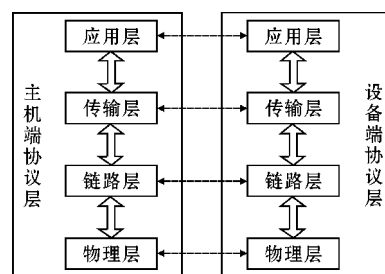


图 1 SATA 结构

以太网接口协议包括 XGMII、RGMII、GMII 及 SG-MII 协议,其中 XGMII 协议带宽为 10 Gb/s<sup>[13]</sup>,其信号线为 16 对。RGMII 协议信号线仅为 4 对,RGMII 协议的参考时钟为 125 M,采用双边沿有效传输数据,RG-

MII 协议带宽为  $4 \times 125 \times 2 = 1000 \text{ Mb/s}$ 。RGMII 协议的数据带宽能满足读取操作需求,且使用更少资源,因此采用 RGMII 协议。

SRIO 协议支持  $1 \times$ 、 $2 \times$  和  $4 \times$  三种链路通道,单通道支持  $1.25$ 、 $2.5$ 、 $3.125$ 、 $5.0$  和  $6.25 \text{ Gb/s}$ <sup>[14]</sup> 5 种传输线速率。写入存储系统速率需求为  $20 \text{ Gb/s}$ ,因此采用两路  $4 \times$  通道,单通道线速率为  $3.125 \text{ Gb/s}$  的 SRIO 协议可满足需求。

磁盘阵列(RAID<sup>[7]</sup>)技术将多个独立存在的硬盘重组成一个信息容量庞大的独立工作硬盘组,以扩大自身冗余率,提高信息储存效率<sup>[15]</sup>。

基于以上分析,设计一款支持 SRIO 协议、RGMII 协议、SATA3.0 协议,支持 RAID\_0 阵列并行控制 8 个 1TB 的 SSD<sup>[4]</sup> 的大容量高带宽数据存储系统。

2 系统总体方案

存储系统的设计框架如图 2 所示,应用 Xilinx Virtex-7XC7VX690T 作为存储系统的控制核心,数字接口 1、2 完成片间通信,数字接口 3 完成片内通信。数字接口 1、2、3 均由 FPGA 实现,其中数字接口 1 采用 SRIO 协议缓存高带宽大容量的数据,数字接口 2 采用 RGMII 协议与上位机通信,数字接口 3 采用 SATA3.0 协议与 SATA SSD 通信。

(1)实现 3 种高速数字接口,数字接口 1 采用两路 4 通道的 SRIO 协议,写入速度可达  $20 \text{ Gb/s}$ ;数字接口 2 采用 RGMII 协议写入速度可达  $1 \text{ Gb/s}$ ;数字接口 3 采用 SATA3.0 协议并且利用 RAID\_0 结构并行控制 8 块 SSD,写入速度可达  $29 \text{ Gb/s}$ 。

(2)电路结构简单、兼容性强,由于 FPGA 具有现场可编程以及并行处理的优势,所有协议皆用 FPGA 实现,因此各接口协议可根据实际应用需求做出调整。

3 硬件方案

存储系统的硬件架构如图 3 所示,主要包括控制模块、电源模块、时钟模块、SATA SSD、DDR 和 FLASH。

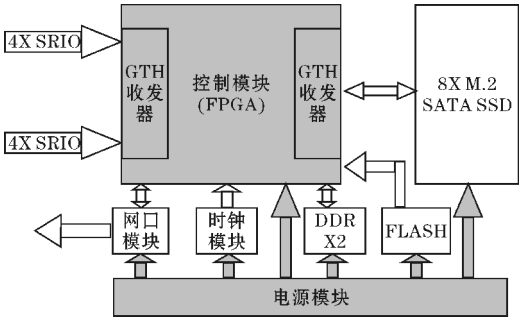


图 3 硬件架构

3.1 控制模块简介

控制模块主要实现数据处理和片内、片间传输,采用 Xilinx Virtex-7XC7VX690T 作为主控,利用其高速收发器 GTH、GTX 实现高速数据传输。Xilinx Virtex-V7 系列是业界最高性能的 FPGA,具有 1955k 逻辑单元、3600 个 DSP slices、X8 Gen3 的 PCIe 接口、1200 个 IO 接口、最高收发速率  $28.05 \text{ Gb/s}$ 。同 Xilinx 其他 7 系列性能对比如图 4 所示。

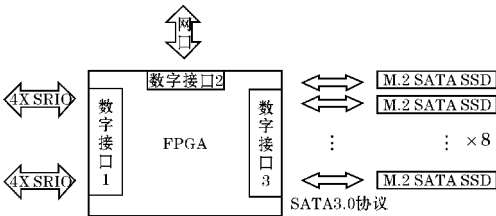


图 2 存储系统框图

该存储系统具有以下特点:

Max. Capability	Spartan-7	Artix-7	Kintex-7	Virtex-7
Logic Cells	102K	215K	478K	1,955K
Block RAM <sup>(1)</sup>	4.2 Mb	13 Mb	34 Mb	68 Mb
DSP Slices	160	740	1,920	3,600
DSP Performance <sup>(2)</sup>	176 GMAC/s	929 GMAC/s	2,845 GMAC/s	5,335 GMAC/s
MicroBlaze CPU <sup>(3)</sup>	260 DMIPs	303 DMIPs	438 DMIPs	441 DMIPs
Transceivers	–	16	32	96
Transceiver Speed	–	6.6 Gb/s	12.5 Gb/s	28.05 Gb/s
Serial Bandwidth	–	211 Gb/s	800 Gb/s	2,784 Gb/s
PCIe Interface	–	x4 Gen2	x8 Gen2	x8 Gen3
Memory Interface	800 Mb/s	1,066 Mb/s	1,866 Mb/s	1,866 Mb/s
I/O Pins	400	500	500	1,200
I/O Voltage	1.2V–3.3V	1.2V–3.3V	1.2V–3.3V	1.2V–3.3V
Package Options	Low-Cost, Wire-Bond	Low-Cost, Wire-Bond, Bare-Die Flip-Chip	Bare-Die Flip-Chip and High-Performance Flip-Chip	Highest Performance Flip-Chip

图 4 Xilinx 7 系列性能对比

Xilinx Virtex-7XC7VX69T 提供的高速收发器 GTH (图 5) 可实现数据的高速收发,单对数据收发线速率为 500 Mb/s ~ 13.1 Gb/s,支持SATA3.0的带外信号(out of

band,OOB)检测与生成、Serial RapidIO (SRIO)、PCI Express 等,因此能满足设计中的 SRIO 协议、RGMII 协议、SATA3.0协议的需求。

GTH Transceiver Performance						
Symbol	Description	Output Divider	Speed Grade			Units
			-3E/-2GE	-2(C/I)/-2LE	-1(C/I/M) <sup>(1)</sup>	
F <sub>GTHMAX</sub>	Maximum GTH transceiver data rate		13.1	11.3	8.5	Gb/s
F <sub>GTHMIN</sub>	Minimum GTH transceiver data rate		0.500	0.500	0.500	Gb/s

图5 Xilinx GTH 收发器

存储系统的 FPGA 硬件资源使用如表 1 所示。

表1 FPGA 资源利用表			
Resource	Utilization	Available	Utilization/%
LUT	49527	433200	11.43
LUTRAM	2825	174200	1.62
FF	52094	866400	6.01
BRAM	20550	1470	13.98
IO	165	350	47.14
GT	12	48	25.00
BUFG	17	32	53.13
MMCM	3	20	15.00
PLL	3	20	15.00

3.2 时钟资源

时钟模块可输出多路不同频率、不同电平标准的时钟信号,为系统各器件提供工作时钟。时钟资源如表 2 所示。

表2 时钟资源		
晶振	频率/M	用途
1	150	SATA 参考时钟
2	200	第一块 DDR3 的参考时钟
3	200	第二块 DDR3 的参考时钟
4	200	XC7VX690TFFG1158 的系统参考时钟
5	125	QSFP 参考时钟
6	125	SRIO 参考时钟
7	48	STM32F042K6T6
8	25	Gigabit PHY 芯片
9	100	LMK03806
10	106	LMK03806

3.3 电源管理

电源模块同时使用开关电源和线性电源,兼顾电源效率和稳定性,如图 6 所示。输入12 V 直流电源,通过开关电源 LTM4650A 和开关电源 LTM4644 得到其他器件所需的电压,然后由 TPS51200DRCT 得到 DDR 模块所需电压。电源模块 LTM4650A 的输入电

压4.5 ~ 12 V,输出电压0.6 ~ 5.5 V,支持单通道50 A,双通道25 A。LTM4650A 主要用于为 SSD 硬盘和 FPGA 的高速收发接口 GTH 供电。DC/DC 稳压器 LTM464,输入电压为4 ~ 14 V,输出电压0.6 ~ 5.5 V,支持 4 通道4 A的输出,单通道16 A。LTM464 主要为 STM32、板载晶振、以太网 PHY 芯片、时钟芯片 LMK03806、JTAG 及 QSFP 供电。以上两款开关模式降压型 DC/DC 稳压器,内置温度监视器、差分远端采样放大器,具有过流折返保护、输出过压保护的特点,常用于 FPGA、存储设备中。TPS51200 是德州仪器生产的 DDR 终端稳压器,满足 DDR、DDR2、DDR3、DDR4、DDR4VTT 总线终端的所有电源要求,LTM4644 产生的1.5 V和3.3 V经过 TPS51200 转换为 DDR\_VTT 和 DDR\_VREF。

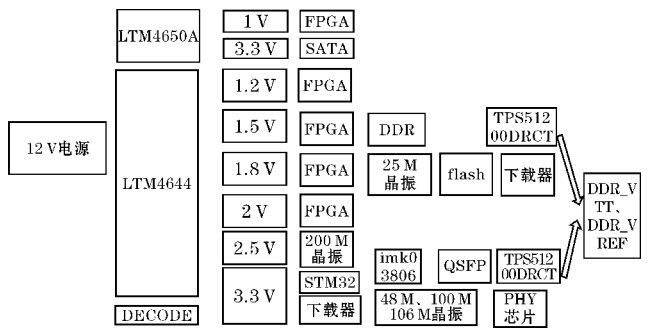


图6 电源管理

4 软件方案

首先通过 SRIO 协议缓存高带宽大容量的数据,再通过 SATA3.0 协议存入 SSD,最后通过 RGMII 协议完成 SSD 和上位机之间的数据传输。

软件框架如图 7 所示,从功能上区分,各协议对应的数字接口可分为完成数据传输的数据接口和完成指令传输的用户接口。通过数据接口,可向存储系统写入两种数据,采集数据和数据标题。为提高传输效率,两种数据分别应用数据流和扇区两种读写方式。数据标题和 RAM 缓存连接,采集数据和 FIFO 缓存连接。

通过用户接口,可对数据进行存储、查看和删除,可设置数据的读写方式以及完成对设备的自检、监测等。

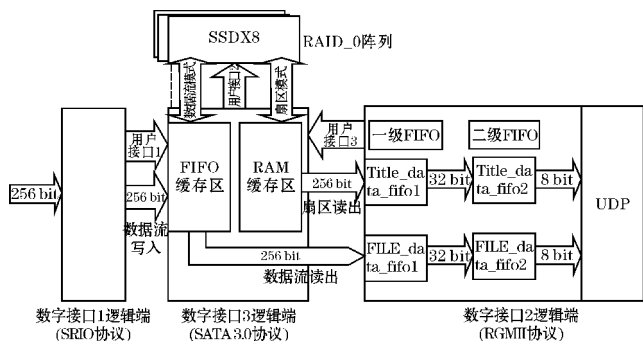


图 7 软件框架

## 4.1 数据传输方案

向存储系统写入数据的流程图如图 8 所示。

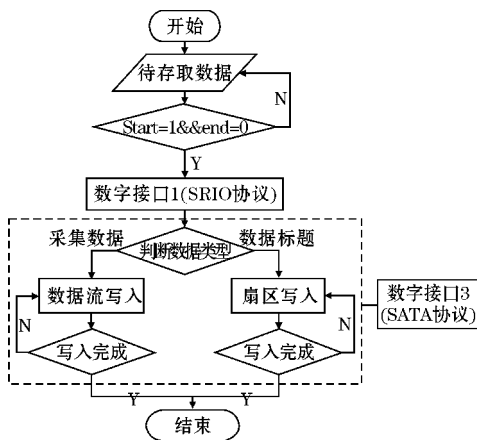


图 8 系统写数据流程图

读取存储系统内数据的流程如图 9 所示。

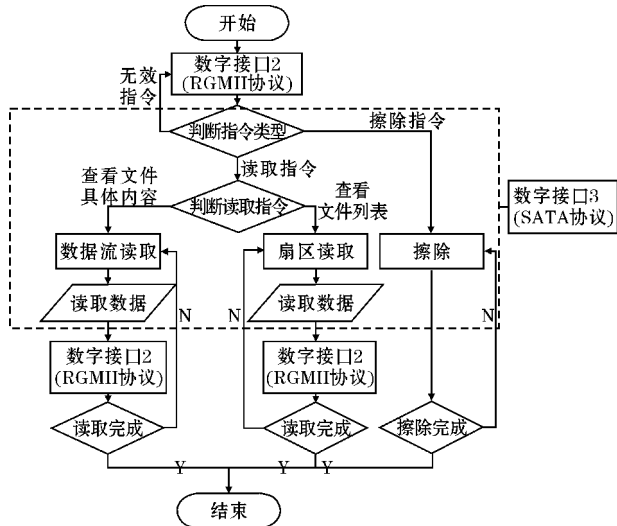


图9 系统读数据流程图

## 4.2 用户接口与通信协议

在数字接口 1、2、3 的逻辑端自定义用户接口 1、2、

3,各用户接口的信号描述如表3~5所示。可通过用户接口1向存储系统传达起/止信号、数据大小信号。

表3 SRIO 用户接口

指令	描述
start	数据开始传输
end	数据结束传输
File_size	本次数据大小
One_sector_over	传输完成 512byte 的提示信号

表4 SATA 控制器用户接口

指令	描述
0x00	停止
0x01	复位 SATA 设备
0x02	Identify SATA 设备
0x04	单扇区数据写入
0x08	单扇区数据读取
0x10	数据流写入
0x20	数据流读取
0x40	多扇区数据写入
0x80	多扇区数据读取

表5 千兆以太网用户接口

指令	描述
View_en	查看 SSD 硬盘中的文件列表
Trim_en	擦除 SSD 硬盘中的文件缓存
Read_en	查看 SSD 硬盘中的某个文件内容(文件最多 4096)

可通过用户接口 2 设置数据读写方式、获取 SATA 设备自检的情况和设备运行过程中的关键参数等。为提高传输效率,SATA Host Controller 提供扇区读写和数据流读写两种指令。通过用户接口 3 查看、擦除存储系统中的文件数据。

## 5 SRIO 协议、SATA3.0 协议、RGMII 协议测试

采用模拟数据验证系统的通信功能和 SRIO 的通信速率。设置本地 PC 的 IP 为 192.168.0.3,端口号为 8080,存储系统 FPGA 端的 IP 为 192.168.0.2,端口号为 8080。模拟设备产生固定帧头帧尾的数据以 20 Gb/s 的速度通过 SRIO 协议写入存储系统,然后通过 SATA3.0 协议写入 SSD,读出单帧数据并通过 RGMII 协议发送到上位机。SRIO 和 RGMII 协议测试如图 10 所示。上位机接收到帧头为 7e7e5a5a,帧尾为 a5a5e7e7 的单帧数据,该数据与生成的模拟数据完全吻合,表明已实现 SRIO 协议、SATA3.0 协议、RGMII 协议,SRIO 接口的速率大于 20 Gb/s。



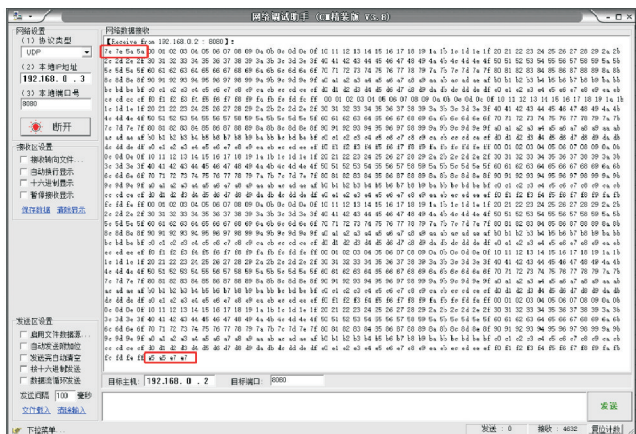


图10 SRIO和RGMII协议测试

对8块SATA SSD同时进行数据读写测试,取100组测试样本,主要步骤包括:

(1)设置写入起始扇区地址 0x200\_0000;(2)设置写入扇区长度 0x400\_0000;(3)设置写入DMA长度为1024 byte;(4)设置DMA写入/读取命令;(5)启动SATA写入/读取。读写速度:

$$\bar{V} = \text{Size} / \bar{t}$$

$$\text{Size} = \text{sector\_length} \times \text{sector\_size}$$

式中:sector\_length为用户设置的读写扇区长度,sector\_size为单个扇区大小,为定值512 byte,Size为传输的数据大小, $\bar{t}$ 为指令执行平均耗时, $\bar{V}$ 为系统平均读写速度。

写入测试数据如图11所示。 $\bar{t} = \frac{1}{N} \sum_{i=0}^{N-1} t_i (i=0,1, \dots, N-1)$ ,  $\bar{t} = 9390$  ms, sector\_length为67108864(0x400\_0000),代入式(2)、(3)可得 $\bar{V} = 29.2$  Gb/s。读取速度计算同理, $\bar{V} = 35.7$  Gb/s。实际读写速度低于理论速度,但已满足设计指标需求。SSD的实际读写速度主要受软件开销和硬件电路两个因素影响。存储系统的软件开销如图12所示,包括三部分:写数据初始化、信号同步、组合帧。写数据初始化,主要完成双向握手;信号同步,是因为当8块SSD的ready信号皆有

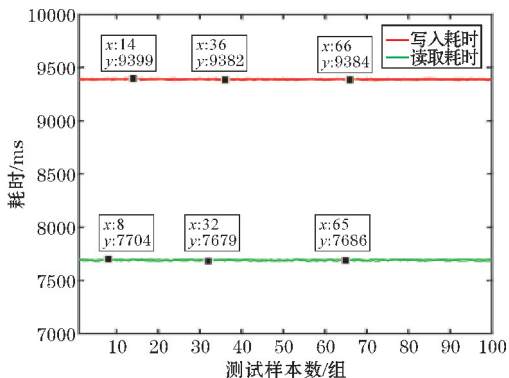


图11 8块SSD并行读写测试

效,才能进入下一状态;组合帧会产生延迟和损耗,实际传输中数据以数据帧的形式进行传输,帧结构由SOF、DATA、EOF以及CRC组成,其中DATA才为真实有效数据。综上所述,实际测得速度低于理论数值。

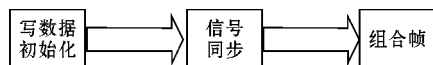


图12 软件开销

用Wireshark软件对实现RGMII协议的网口进行速率测试,如图13所示。存储系统(192.168.0.2)向PC端(192.168.0.3)连续发送帧长为1000 byte的数据包,Wireshark抓取10000个数据包总计花费0.081914 s,因此网口速率为 $10000 \times 1000 \times 8 / 0.081914 = 910$  Mb/s,考虑到数据包中的SOF、CRC、EOF所占字节以及软件开销,该速率符合理论标准。

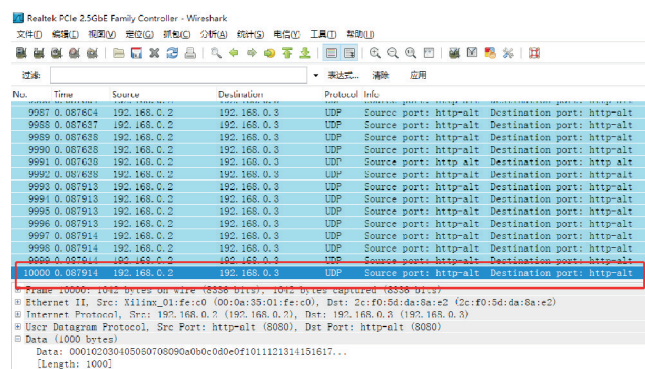


图13 Wireshark抓取

## 6 结束语

介绍一种基于SATA3.0协议的大容量数据存储系统的设计。测试基于FPGA的SATA3.0协议、SRIO协议和RGMII协议的高速稳定性、验证RAID\_0阵列并行控制SSD的正确应用,为现代高速数字信号通信中的大容量数据存储系统设计提供一定参考意义。

## 参考文献:

- [1] 张德智,戴跃飞,徐今,等.一种S波段T\_R组件的设计与制造[J].现代雷达,2008(2):76-78.
- [2] 杨磊,宋佳音. SRIO在高速信号处理系统中的应用分析[J].信息与电脑(理论版),2019(4):94-95.
- [3] Budruk R, Anderson D, Shanley T. PCI express system architecture [M]. Addison-Wesley Professional, 2004.
- [4] 白雪,王海鑫,郑桥,等. 基于SATA-3.0的高速

- 大容量嵌入式存储阵列控制器[J]. 电子设计工程, 2018, 26(10): 138-143.
- [5] 刘光祖, 张强, 仲雅莉. 基于 RapidIO 协议的光纤通信系统设计与实现[J]. 电讯技术, 2013, 53(7): 840-844.
- [6] 徐玉杰, 张磊. 高速大容量存储系统的设计与实现[D]. 北京: 中国科学院研究生院(空间科学与应用研究中心), 2007.
- [7] 张珮, 王晓曼. 固态硬盘 SSD 性能分析及 RAID\_0 方案设计[J]. 微型机与应用, 2016, 35(6): 26-28.
- [8] 杨亚涛, 台运娇, 江先阳. SATA-3.0 物理层设计与 FPGA 实现[J]. 信息技术, 2019, (10): 121-130.
- [9] 屈涛. CompactPCI Express 体系结构技术分析[J]. 工业控制计算机, 2006, 19(11): 10-12.
- [10] Serial ATA. International organization: Serial ata revision 3.0[J]. Gold Revision, June, 2009(2).
- [11] 罗桂娥, 康霞. 固态硬盘性能优化研究与实现[J]. 计算机工程与应用, 2015, 51(1): 43-48.
- [12] Woods L, Eguro K. Groundhog-a serial ata host bus-adapter(hba) for fpgas[C]. 2012 IEEE 20th International Symposium on Field-Programmable Custom Computing Machines. IEEE, 2012: 220-223.
- [13] Chang L, Ganga I. Dual speed interface between media access control unit and physical unit; U. S. Patent Application 10/851[P]. 2005-11-24.
- [14] 多卉枫, 任勇峰, 武慧军. 基于 SRIO 的数据传输设计[J]. 电子测量技术, 2021, 44(21): 7-11.
- [15] 孙桦栋. 磁盘阵列(RAID)在医院网络信息系统中存储解决方案探讨[J]. 数据, 2021(12): 95-96.

## Design of High-speed Mass Storage System based on SATA3.0 Interface

LUO Yiling, DU Yuming, LI Yijie

(College of Electronic Engineering, Chengdu University of Information Technology, Chengdu 610225, China)

**Abstract:** With the rapid development of communication technology, data transmission is characterized by high-speed, large-capacity and high-bandwidth, but data cache can be a problem. In order to solve this problem, a high-speed and large-capacity storage system is designed based on SATA3.0 protocol. Using Xilinx-Virtex7 series chip XC7VX690TFFG1158 as main controller, SRIO protocol is used to cache the data of Pre-device with high bandwidth and large capacity, SATA3.0 protocol is used to communicate with SSD, and RGMII protocol is used to communicate with the host computer. Taking advantage of the parallel processing characteristic of FPGA, RAID\_0 array is used to realize the parallel control of 8 SSDs. The test result shows that, for the whole storage system, the average data writing speed is greater than 20 Gb/s, the average write speed of the entire SSD array can reach 29.2 Gb/s, and reading speed can reach 35.7 Gb/s. This storage system has a wide application prospect in the field of high-speed digital signal processing such as satellite communication, and it provides some reference significance for the design of high-speed and large-capacity storage system.

**Keywords:** FPGA; SATA3.0; SRIO; RAID\_0 arrays; GTH transceiver