

# 基于 SATA3.0 的存储系统优化设计

李晋涛<sup>1</sup>,任勇峰<sup>1,2</sup>,杨志文<sup>1</sup>,李辉景<sup>1</sup>

(1.中北大学 电子测试技术国家重点实验室,山西 太原 030051;

2.中北大学 仪器科学与动态测试教育部重点实验室,山西 太原 030051)

**摘要:** 针对遥测试验中,获取的高分辨率、高帧频的大容量图像数据的实时存储以及传输稳定性问题,对存储系统进行了优化设计。该系统以 Kintex 7 系列 FPGA 为控制核心,将接收到的大容量图像通过 SATA3.0 传输总线,存储到固态硬盘中,并且优化了链路状态机,通过实时标志监测解决了传输中断的问题。存储完成后,通过千兆以太网将固态硬盘中的数据读取到上位机进行回放和分析。在常温以及-40 ℃和+60 ℃条件下完成大量的读写测试,结果表明,优化后的存储系统可以完成高速率、长时间的稳定存储,在大容量存储方面具有很好的参考价值。

**关键词:** SATA3.0; 固态硬盘; OOB; 大容量存储

中图分类号: TN919

文献标识码: A

DOI:10.16157/j.issn.0258-7998.200612

中文引用格式: 李晋涛,任勇峰,杨志文,等. 基于 SATA3.0 的存储系统优化设计[J]. 电子技术应用, 2021, 47(1): 86–90.

英文引用格式: Li Jintao, Ren Yongfeng, Yang Zhiwen, et al. An optimized design of a storage system based on SATA3.0[J]. Application of Electronic Technique, 2021, 47(1): 86–90.

## An optimized design of a storage system based on SATA3.0

Li Jintao<sup>1</sup>, Ren Yongfeng<sup>1,2</sup>, Yang Zhiwen<sup>1</sup>, Li Huijing<sup>1</sup>

(1.State Key Laboratory of Electronic Technology, North University of China, Taiyuan 030051, China;

2.Key Laboratory of Instrument Science and Dynamic Testing(Ministry of Education)

North University of China, Taiyuan 030051, China)

**Abstract:** Aiming at the problem of real-time storage of high resolution, high frame frequency and large capacity image data obtained in telemetry test, as well as the problem of transmission stability, we optimized a storage system. This system takes 7 series FPGA as the control core, and stores the received large-capacity image into the solid state hard disk through SATA3.0 transmission bus. The state machine is optimized in the stored procedure to solve the problem of transmission interruption through the real-time marker monitoring. After the storage was completed, the data in the SSD will be read through the gigabit Ethernet to the computer for playback and analysis. Lots of reading and writing data were completed at room temperature and -40 ℃ and +60 ℃. The results show that the optimized storage system can achieve high speed and long time stable storage and has a good reference value in large capacity storage.

**Key words:** SATA3.0; solid state disk; OOB; bulk storage

## 0 引言

SATA3.0(SerialATA Revision3.0)采用差分信号传输系统,该系统对共模噪声有很强的抵抗力,因此 SATA 可以采用更低的电压去抑制噪声。在数据传输这一方面,SATA3.0 的传输速度在 SATA2.0 的基础上翻倍,理论上可以达到 6 Gb/s。另一方面,SATA 总线使用了嵌入式时钟频率信号,具备了比以往更强的纠错能力。在传输过程中能对相应的指令进行检查,即使出现错误的指令,也可以进行及时的修正,因此广泛地应用于实际工程中。但是,基于测试系统产生的数据愈加地趋向于大体积、高速率,多数存储系统采用存储阵列的方式进行存储。然而,随着存储量的增加以及传输环境的复杂,SATA 总

线会出现个别通道传输中断的现象,虽然只是偶尔现象,但却能造成不可挽回的损失。

针对以上问题,该设计基于 FPGA 控制的 SATA3.0 总线,从硬件和代码方面优化了存储系统并进行验证。实测存储速率不低于 800 MB/s,并且数据传输稳定,在大容量存储应用领域具有很好的参考和使用价值。

## 1 系统总体设计

基于 SATA3.0 的存储电路能够实现对大量数据的快速、稳定的存储,该系统的总体设计方案如图 1 所示。系统由四部分组成,即数据接收模块、FPGA 控制模块、数据存储模块和数据分析模块。数据接收模块采用了 4 路 SRIO 数据接口,接收外部数据,单路传输速率为

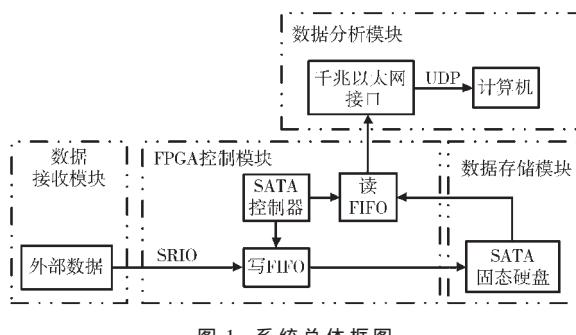


图 1 系统总体框图

3.125 Gb/s, 总体理论传输速率为 12.5 Gb/s。控制模块采用 FPGA, 主要负责逻辑控制, 即通过控制 SRIO 协议、SATA 协议以及 UDP 协议来完成数据的接收、存储和调用。数据存储模块采用 4 片 500 GB 的固态硬盘, 以单次记录的方式, 先擦除再写入, 断电停止记录。数据分析模块采用 UDP 协议, 通过千兆以太网接口进行交互<sup>[1]</sup>。对试验数据进行分析时, 只需要将硬盘内存储的数据读取到计算机上即可。

## 2 硬件电路设计

此模块采用 Xilinx 的 Kintex-7 系列 FPGA 作为主控

芯片, 该芯片采用 RocketIO 技术进行串行数据传输, 即串行收发器 GTX 接口, 单路最高传输速率可达到 12.5 Gb/s<sup>[2]</sup>, 串行收发器配备有可配置的阻抗、摆幅以及耦合等级。SATA 传输采用全双工的方式, 2、3 引脚作为发送端将指令或者数据发送至硬盘, 5、6 引脚作为接收端接收硬盘返回的信息。相比于以往的并行传输方式, 串行传输采用差分走线, 具有很强的抗噪声能力, 不需要考虑多排传输线之间的干扰, 只需要提高时钟频率便可以提高传输速率, 串行传输一次传输一个比特(bit)包含有源同步时钟。而时钟是通过 PMA PLL 将一个高质量的参考时钟倍频出高速串行时钟, 用来驱动 GTX 收发器工作, 本设计采用的时钟芯片为 CDCM61002<sup>[3]</sup>, 该芯片内置板载锁相环, 可以提供稳定的差分时钟作为参考时钟, 并且具有功耗低、高性能、低相位噪声等优点。在时钟配置方面, 该芯片可配置出专有的 SATA 时钟、PCI Express 以及 GigE 等。根据芯片手册, 此处选用 25 MHz 的晶振作为时钟输入, 同时将 PR0、PR1、OD2 配置为低电平, OD1 和 OD0 配置为高电平, 即实现设计需要的 SATA 参考时钟 150 MHz。同时, 将 OS1 配置为低电平、OS0 配置为高电平, 从而控制时钟输出方式为 LVDS。硬件电路图如图 2 所示。

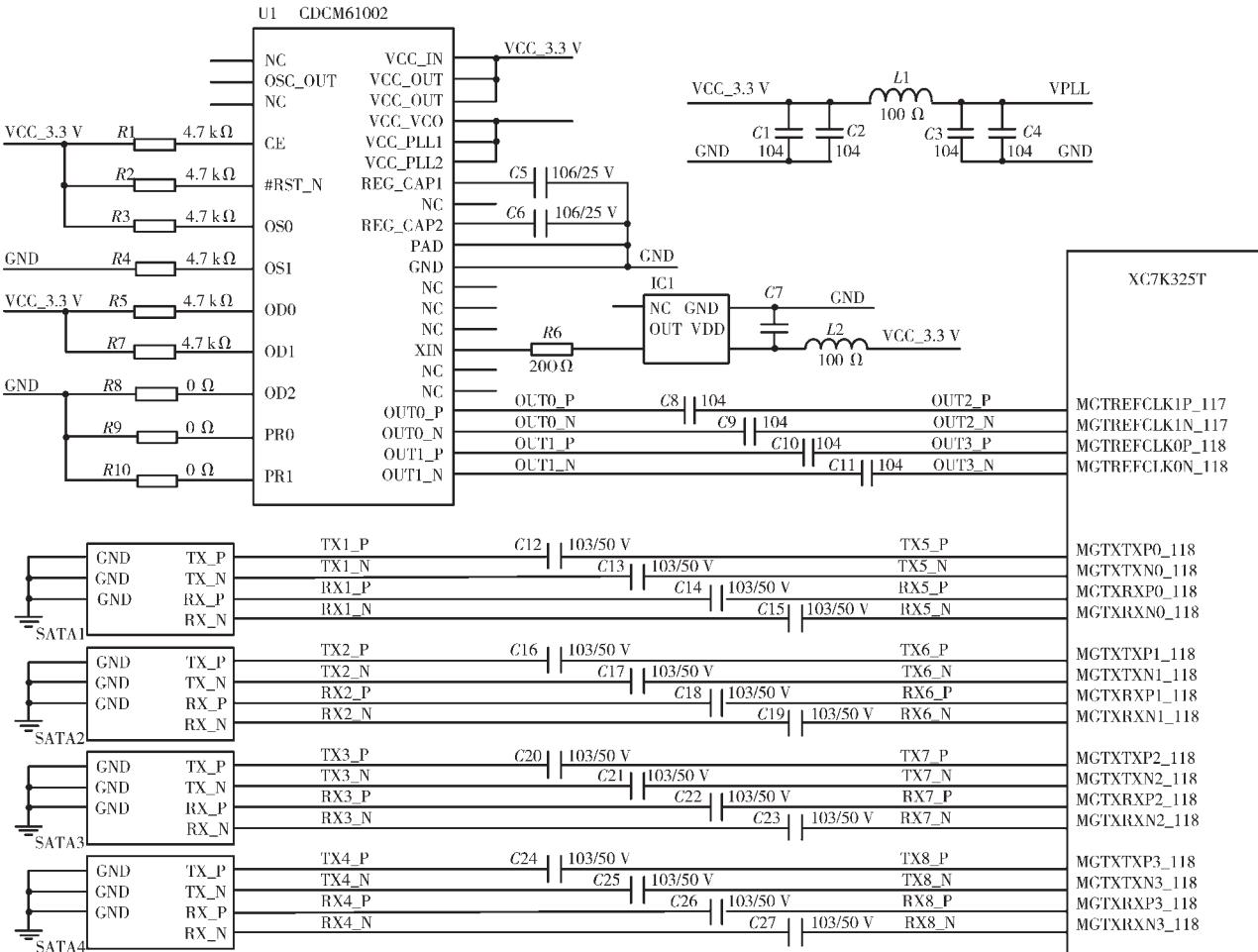


图 2 SATA 硬件原理图

### 3 系统软件设计

#### 3.1 SATA3.0 结构分析

SATA3.0 接口主要包括四个层次：应用层、传输层、链路层以及物理层，总体框架如图 3 所示。

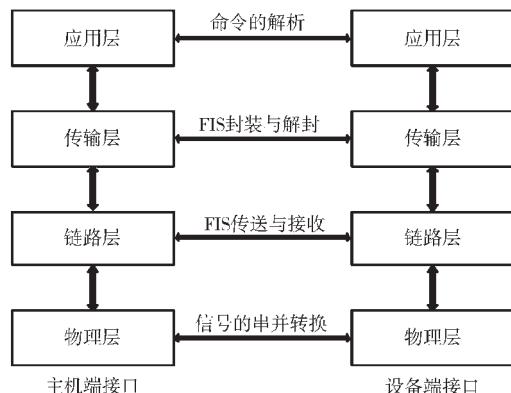


图 3 SATA 协议结构图

(1) 物理层主要是由差分对 TX 和 RX 来实现串行全双工传输，主要功能是主机端与设备端的链路初始化<sup>[4]</sup>。为了提高 SATA 接口的主机端和设备快速建立通信的能力，SATA 协议在物理层引入了 OOB 信号，通过原语交互来完成链路初始化和传输速度的协商。

(2) 链路层主要负责的是 FIS 帧收发，通过 CRC 校验和加扰码解扰码的方式来提高数据传输的准确性<sup>[5]</sup>。当链路层发送数据时，SATA 协议将 0x52325032 作为初始值<sup>[6]</sup>，根据 CRC32 校验方式，将最多 2 048 个 DWORDS 有效数据依次校验，生成一个最终的校验和，然后将校验和加在所发送的有效数据后，并进行加扰，最后进行 8B/10B 编码后发送到物理层，在接收到设备端的数据时，先对数据进行 8B/10B 解码以及解扰，然后将帧头原语 SOF、CRC 校验以及帧尾原语 EOF 去除，再次计算接收到的有效数据的 CRC 校验和<sup>[7]</sup>，并与接收到的 CRC 校验和做对比，然后向传输层报告当前传输状态。CRC32 的生成多项式如式(1)所示：

$$G(X) = X^{32} + X^{26} + X^{23} + X^{22} + X^{16} + X^{12} + X^{11} + X^{10} + X^8 + X^7 + X^5 + X^4 + X^2 + X + 1 \quad (1)$$

(3) 传输层主要负责将相关的数据或命令构造成相应的 FIS 帧。其中，帧类型主要有数据帧、寄存器帧以及

控制帧这三大类。当传输层接收到来自链路层的帧时，会将该帧解析出帧类型以及帧内容，并将其中的数据写入寄存器中。当检测到应用层发出的数据传输操作请求后，传输层作为响应，会从寄存器中读取数据，并按照相应的帧格式组成帧发送至链路层<sup>[8]</sup>。

(4) 应用层主要负责解析各个寄存器以及执行寄存器中的命令。根据命令的要求，应用层会接收来自主机端传输的数据，并将数据写入 SATA 硬盘中，或者将硬盘中的数据读取出来传送给主机端。存储或者读取的方式可以是 PIO 或 DMA，为了提高数据交互速率，本文采用 DMA 方式<sup>[9]</sup>。在数据进行缓冲时，分别采用一个读 FIFO 和一个写 FIFO 来完成。写 FIFO 负责缓存传输层解析出来的数据帧，而读 FIFO 负责把来自总线的数据保存下来<sup>[10]</sup>。在这一系列动作完成之后，应用层将构造数据帧的指令发给传输层。

#### 3.2 OOB 状态机优化

SATA 物理链路的建立主要是通过 OOB(Out of Band, OOB)信号来实现的<sup>[11]</sup>。FPGA 和硬盘上电以后，FPGA 端首先发出重置信号 COMRESET，硬盘在检测到 FPGA 发出的 COMRESET 后，会发出初始化信号 COMINIT 作为回应，并且硬盘可以在任意时间通过发送 COMINIT 来发起会话。收到硬盘回应以后，FPGA 会进行校准并发出 COMWAKE 信号。硬盘在 RX 差分对上接收到 COMWAKE 信号后会校准发送器，并且会将 COMWAKE 信号连续传输六次，接着连续发送 ALIGN 原语与主机端进行传送速度的协商。FPGA 接收到 COMWAKE 信号，锁存硬盘发来的原语然后在 54.6  $\mu$ s 内回传 ALIGN，并且在 COMWAKE 信号收到之后连续地发送 D10.2 原语。返回 ALIGN 信号给硬盘的速率应与接收时保持一致，如在 880  $\mu$ s 内未检测到硬盘返回 ALIGN，则进入复位状态重新连接。硬盘在检测到 ALIGN 原语后，会以 SYNC 同步信号作为回应，表示准备开始正常的传输操作。当 FPGA 接收到三个来自硬盘的非 ALIGN 原语后，linkup 信号由低电平变为高电平，链接建立完成，进入正常操作。

如图 4 所示，以往的数据传输过程中，可能会遇到物理链路连接断开的现象<sup>[12]</sup>，此时 linkup 信号由高变低，数据传输中断，OOB 状态机停留在当前 linkready 状态。从硬盘回放的数据可以看出，递增数已经不再连续，

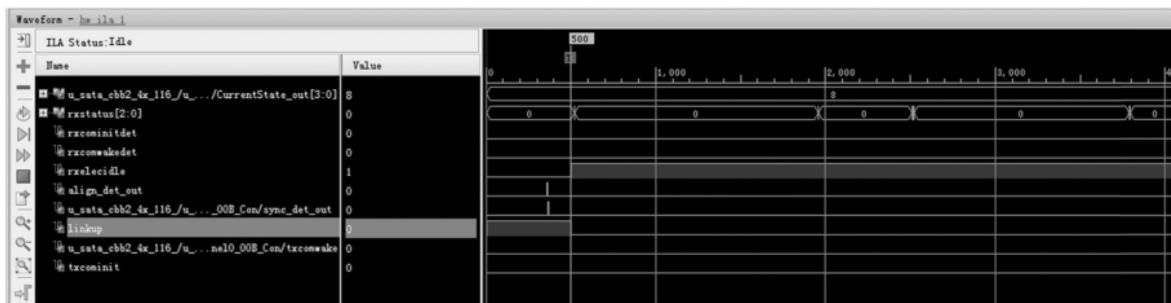


图 4 链路优化前

虽然这种问题出现的概率比较小,但却能造成无法挽回的后果。为了提高系统链路的稳定性以及抗干扰能力,使得系统建立起来以后数据更加稳定地传输,避免在系统工作中出现中断,需要对 OOB 状态机增加一个通路标志循环监测状态。当出现链路断开时,在物理层立即重新建立物理链路,及时地恢复数据传输。图 5 是改进后的 OOB 状态机。

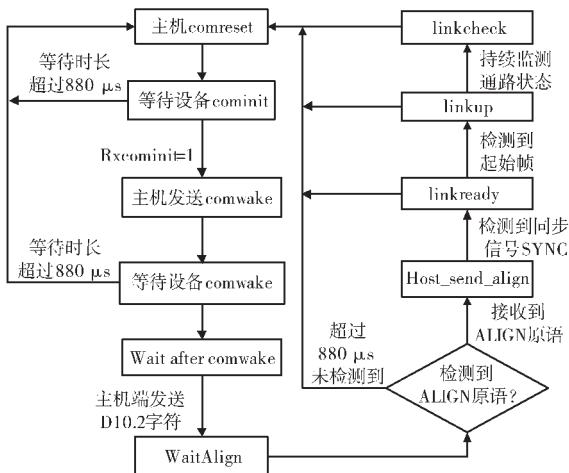


图 5 优化后的 OOB 状态机

linkup 状态之前的状态与原来相同，是物理通路的建立过程。linkup 状态下认为发送的已经是正常的数据帧，在此状态下可以检测帧头 SOF 原语和 EOF 原语判断是否有正常数据帧的交互。当检测到帧头 SOF 原语后进入下一状态。linkcheck 是链路检测状态，也是物理层正常工作时需要保持的状态。在正常工作状态下，链

路需要保持稳定,这就需要实时检测当前链路的工作状态,此状态下需要检测的标志位有 rxidle、linkup 等。可以看到链路断开以后,linkup 信号由高变低,此时 OOB 状态机立即进入 0 状态,即复位状态;txcominit 拉高,说明链路开始初始化;经过短时间的链路重建 linkup 信号拉高,说明链路重新恢复连接,rxidle 信号持续拉低,说明数据已经开始继续传输。图 6 是 ILA 抓取的一条通路重新建立的采样图。

## 4 系统整体测试

首先,使用外部发图机以 800 MB/s 的速率循环发送递增数,递增数以四个字节的帧计数,再加上 EB 90 作为帧结尾<sup>[13]</sup>,每帧包含 256 个字节。接收数据时,FPGA 以 156.25 MHz 的时钟接收数据,然后写入 64 位的 FIFO。由于 SRIO 采用 8B/10B 编码,4 路 SRIO 接口理论接收速率可达到  $156.25 \text{ M} \times 64 \div 10 = 1000 \text{ MB/s}$ 。紧接着,例化一个 64 位转 128 位的 FIFO,读时钟为 150 MHz,此时 FIFO 的读速率为  $150 \text{ M} \times 128 \div 10 = 1920 \text{ MB/s}$  完全大于接收速率。之后 FPGA 将打包后的 128 位的数据拆分成四路 32 位的数据,通过 SATA 控制器,将 32 位的数据转为串行的差分数据,通过 SATA 串口同时存储到四块硬盘中。图 7 为上位机读取到的递增数,图 8 为递增数校验结果。

在高温 60 °C 和低温 -40 °C 条件下, 以 800 MB/s 的写入速率每次写入硬盘 600 GB 的数据进行测试, 经过 100 次测试及数据分析, 发现数据传输稳定, 递增数连续且没有出现传输中断的情况。

5 结论

本文优化了一种基于 FPGA 控制的存储系统, 系统

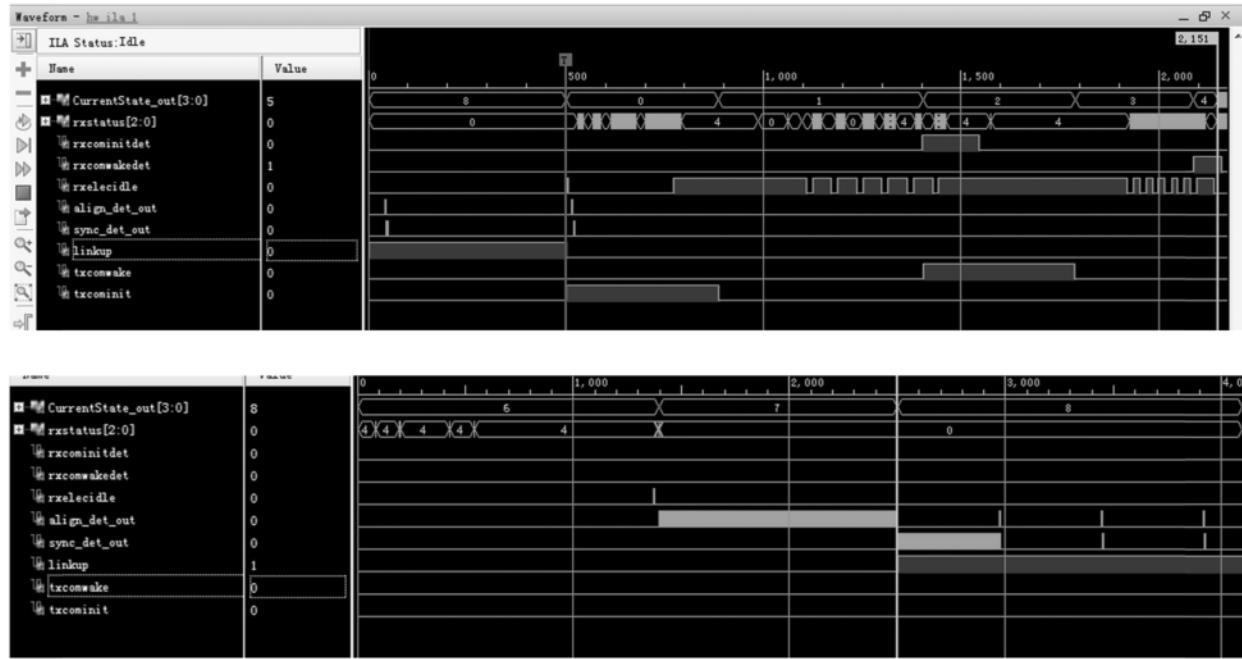


图 6 链路优化后

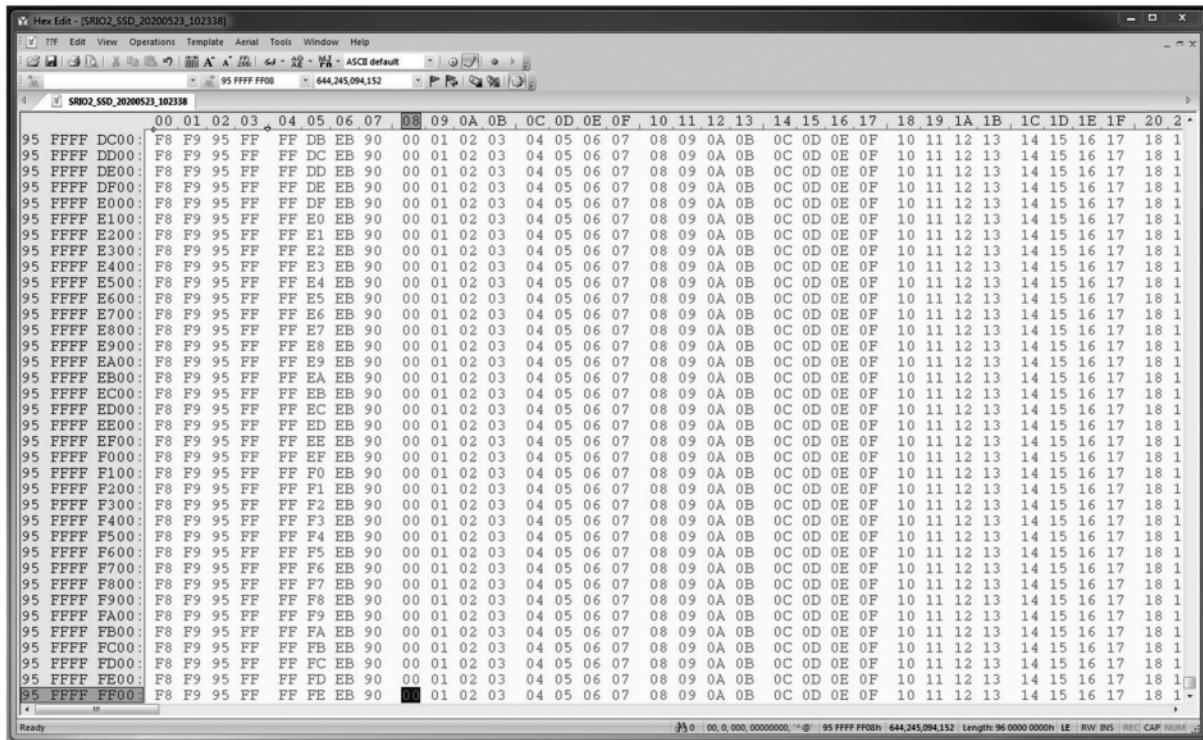


图 7 接收数据回放



图 8 递增数校验

以 SATA3.0 作为传输总线, 固态硬盘阵列作为存储介质, 对存储逻辑进行了设计, 并对硬件电路和 OOB 状态机进行了优化。最终实现了一种高速率, 高稳定性的存储系统, 解决了大容量存储系统的传输稳定性问题。经过大量的读写测试, 表明该系统可以对高速数据进行可靠的存储, 满足多数遥测试验的需求, 具有很好的应用价值。

#### 参考文献

- [1] 程晓航. 基于 FPGA 的 SATA 盘大容量存储系统的设计与实现[D]. 西安: 西安电子科技大学, 2019.
- [2] 黄刚. 大容量高速回放系统设计与实现[J]. 物联网技术, 2019, 9(6): 68–70.
- [3] 张虎威. 基于 SATA 接口的高速图像存储器的研究与设计[D]. 太原: 中北大学, 2017.

- [4] 台运娇, 江先阳. SATA3.0 物理层设计与实现[J]. 信息技术, 2019(10): 121–125, 130.
- [5] 张毅. 基于 FPGA 的 SATA3.0 硬盘阵列控制器设计[D]. 保定: 河北大学, 2017.
- [6] 刘赞. 基于 FPGA 的 SATA3.0 主机端控制器的设计[D]. 保定: 河北大学, 2016.
- [7] 崔梦强. 基于 FPGA 的 SATA 控制器的设计及实现[D]. 西安: 西安电子科技大学, 2017.
- [8] 陆晓玲. 基于 FPGA 和 SATA3.0 接口的高速大容量存储系统的设计与实现[D]. 南京: 南京邮电大学, 2017.
- [9] 刘森帆. 基于固态盘的数据存储与回放技术研究[D]. 西安: 西安电子科技大学, 2018.
- [10] 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.
- [11] 杨亚涛, 张松涛, 李子臣, 等. SATA 3.0 物理层设计与 FPGA 实现[J]. 计算机工程与应用, 2017, 53(20): 38–42.
- [12] 周玉山. 基于 FPGA 的多通道 SATA 接口设计与实现[D]. 西安: 西安电子科技大学, 2017.
- [13] 杨玉华, 秦菲, 单彦虎, 等. 基于 FPGA 的高速图像存储系统的设计[J]. 电子器件, 2018, 41(4): 970–975.

(收稿日期: 2020-07-03)

#### 作者简介:

李晋涛(1994-), 男, 硕士研究生, 主要研究方向: 测试计量技术及仪器。

任勇峰(1968-), 男, 教授, 主要研究方向: 测试计量技术及仪器、电路与系统。

## 版权声明

经作者授权，本论文版权和信息网络传播权归属于《电子技术应用》杂志，凡未经本刊书面同意任何机构、组织和个人不得擅自复印、汇编、翻译和进行信息网络传播。未经本刊书面同意，禁止一切互联网论文资源平台非法上传、收录本论文。

截至目前，本论文已经授权被中国期刊全文数据库（CNKI）、万方数据知识服务平台、中文科技期刊数据库（维普网）、DOAJ、美国《乌利希期刊指南》、JST 日本科技技术振兴机构数据库等数据库全文收录。

对于违反上述禁止行为并违法使用本论文的机构、组织和个人，本刊将采取一切必要法律行动来维护正当权益。

特此声明！

《电子技术应用》编辑部

中国电子信息产业集团有限公司第六研究所