

基于FPGA的PCIe总线DMA传输的设计与实现

邹晨, 高云

(中航工业西安航空计算技术研究所, 西安 710068)

摘要: 串行的PCIe接口是第3代I/O互连标准,具有高速率和高带宽等特点,克服了传统PCI总线在系统带宽、传输速度等方面的固有缺陷,具有很好的应用前景。使用FPGA来设计基于PCIe总线的数据传输,可以降低硬件的设计成本,提高硬件集成度的同时还能利用FPGA的可编程特性提高设计的灵活性与适应性。在对PCIe总线、FPGA内嵌PCIe硬核结构以及PCIe传输报文进行简要概述的基础上,提出了一种在FPGA内嵌PCIe硬核的基础上实现DMA传输的解决方案,较为详细地阐述了设计方案,对设计进行了评估与分析,并给出了传输带宽的测试结果。

关键词: 数据传输; FPGA; PCIe总线; DMA; IP硬核

中图分类号: V243

文献标志码: A

文章编号: 1671-637X(2015)07-0084-05

Design and Implementation of DMA Transmission with PCIe Bus Based on FPGA

ZOU Chen, GAO Yun

(Xi'an Aeronautics Computing Technique Research Institute, AVIC, Xi'an 710068, China)

Abstract: The serial PCIe interface is the third-generation I/O interconnection standard. Characterized by high speed and high bandwidth, it overcomes the inherent shortcomings of traditional PCI bus in system bandwidth and transmission speed, and has a good application prospect. Use of FPGA to design PCIe bus-based data transmission can reduce the design cost of hardware and improve its integration level, as well as improving the design flexibility and adaptability by making use of the programmable feature of FPGA. Based on a brief overview of PCIe bus, PCIe hard core structure inserted into FPGA and PCIe transmission message, a solution to realize DMA transmission based on PCIe hard core inserted into FPGA is proposed, with a detailed description of the design proposal. An evaluation and analysis of the design is made, and the test results of transmission bandwidth are given.

Key words: data transmission; FPGA; PCIe bus; DMA; IP hard core

0 引言

随着嵌入式技术的飞速发展,嵌入式系统在对大数据量交互方面提出更高需求的同时还对总线传输速率以及数据完整性等方面提出了越来越高的要求。虽然过去的十几年中PCI总线在嵌入式领域得到了广泛的应用^[1],但由于PCI总线的并行特性,整体设计难度大,故主流的技术已经成为系统整体性能提升的瓶颈。因此,串行PCI Express (PCIe)总线标准凭借其高速串行传送特性以及能够支持更高的传输频率、无需共

享总线带宽等优势,一经推出就得到了迅速的发展,并开始逐步取代PCI总线^[2]。

PCIe总线串行传输的特性是通过报文的形式进行传输,每个数据报文在PCIe的事务层被封装成一个或者多个TLP数据包,PCIe设备之间则通过这些数据包进行数据通信^[3]。由于TLP的数据包中包含TLP前缀、TLP头以及TLP摘要等信息,因此,当设备在进行单次数据传输(每个报文数据负载长度为1)时PCIe总线的性能优势并不明显,其传输速度甚至还不如PCI总线。为了得到更高的传输效率,在使用PCIe总线进行数据传输时往往需要使用DMA的传输方式。

1 PCIe总线的概述

PCIe总线技术是取代PCI的第三代I/O技术,也称

收稿日期:2014-07-21

修回日期:2015-05-08

基金项目:装备预研共用技术基金(9140A16010311HK6101)

作者简介:邹晨(1987—),男,江苏南通人,硕士,研究方向为计算机系统结构。

为 3GIO。PCIe 总线是为将来的计算机和通讯平台定义的一种高性能、通用 I/O 互连总线^[4]，在其物理实现上使用了高速差分的方式来实现数据传输，与此同时，其端到端的连接方式使得每条 PCIe 链路中只能连接两个设备，因此相比于 PCI 所有设备共享总线带宽来说，PCIe 总线具有独享传输通道数据带宽的特性^[5]。与 PCI 总线相比，PCIe 总线主要有下面的技术优势：1) 是串行总线，进行点对点传输，每个传输通道独享带宽；2) 支持双向传输模式和数据分通道传输模式，支持 x1, x4, x8, x16 等模式，x1 单向传输带宽可达到 250 MByte/s，双向传输带宽更能够达到 500 MByte/s；3) 充分利用先进的点到点互连、基于交换的技术和基于包的协议来实现新的总线性能和特征；4) 对 PCI 总线具有良好的继承性，可以保持软件的继承和可靠性；5) 充分利用先进的点到点互连，降低了系统硬件平台设计的复杂性和难度，从而大大降低了系统的开发制造设计成本，极大地提高系统的性价比和鲁棒性^[4]。

2 FPGA 内部的 PCIe 接口

目前，PCIe 接口的设计通常可通过专用接口芯片或 FPGA 这两种方法来实现。使用 PCIe 专用接口芯片实现方式的优点是设计开发简单，但缺乏可配置性和灵活性^[6]。为了更好地发挥 PCIe 总线的高带宽优势，一般采用 FPGA 来实现 PCIe 接口。

2.1 PCIe 核的层次结构

目前在很多 FPGA 内部都嵌入了可选的 PCIe 硬核，其逻辑结构如图 1 所示。PCIe 硬核实现了 PCIe 总线的物理层、数据链路层、配置管理层的协议，上层提供一个事务层接口，用户通过该接口完成应用逻辑，最终实现对 PCIe 设备的控制与访问^[4]。

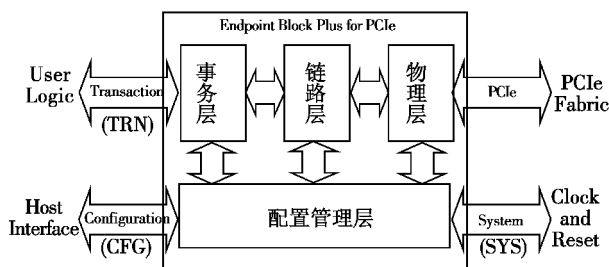


图 1 PCIe 核的层次结构

Fig. 1 Functional structure and interface of PCIe IP core

2.2 事务层接口逻辑结构

事务层接口逻辑主要由 4 部分逻辑电路构成，如图 2 所示，分别是核配置与辅助控制模块、数据接收控制器、BAR 地址空间管理模块和数据发送控制器。其中，PCIe 硬核对外负责与其他 PCIe 设备协议通信，对内负责与数据接收模块以及数据发送模块进行 64 位并行与

高速串行的转换；核配置与辅助控制模块负责对 PCIe 核进行配置与辅助控制；数据接收控制器负责对 PCIe 设备进行单字写以及 DMA 方式下的写操作；数据发送控制器负责对 PCIe 设备的单字读以及 DMA 方式下的读操作；BAR 地址空间管理模块为 PCIe 空间映射几个 BAR 地址空间，最终通过 BAR 地址空间与 PCIe 设备控制逻辑进行交互，实现 PCIe 主设备与 PCIe 从设备之间的数据传输。

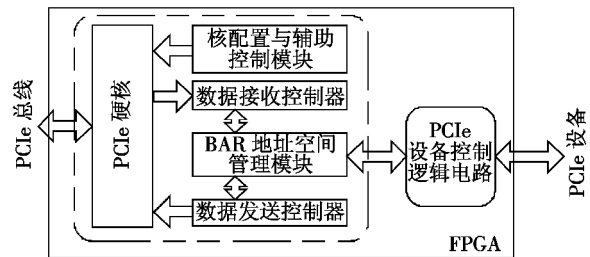


图 2 事务层接口逻辑结构

Fig. 2 Logic structure of transaction layer interface

2.3 事务层数据组织方式

当处理器或者其他 PCIe 设备访问 PCIe 设备时，所传送的数据报文首先通过事务层被封装为一个或者多个 TLP 数据包，之后才能通过 PCIe 总线的各个层次发送出去。TLP 的基本格式如图 3 所示。一个完整的 TLP 数据包由 1 个或者多个 TLP 前缀、TLP 头、数据有效负载和摘要组成^[7]。PCIe 设备每发一次数据都要包含 TLP 数据包格式中的所有信息，但只有 TLP 头和数据有效负载需要在 FPGA 的代码中实现，其他信息都是 PCIe 核自动生成并封装至 TLP 数据包中的。

TLP 前缀 (Optional)	TLP 前缀 (Optional)	TLP 头	数据有效负载	TLP 摘要 (Optional)
-------------------	-------------------	-------	--------	-------------------

图 3 TLP 数据包的格式

Fig. 3 Specification of TLP packet

事务的传输是在 PCIe 请求者和完成者之间进行的操作，一般来说，事务的类型包括存储器事务、I/O 事务、配置事务和消息事务^[4,8]。与此同时，TLP 数据包包头是 TLP 数据包的最重要的标志，其中包含了当前 TLP 的总线事务类型、路由信息等一系列信息。本文讨论的 DMA 读写操作是针对存储器空间的操作，在此本文仅对 PCIe 的存储器操作的 TLP 头的格式进行一些简要介绍。对于一个 32 位地址空间的存储器读写请求来说，其 TLP 头格式如图 4 所示。

	+0				+1				+2				+3											
	7	6	5	4	3	2	1	0	7	6	5	4	3	2	1	0	7	6	5	4	3	2	1	0
Byte 0 >	R		Fmt		Type		R		TC		Reserved		TID		Attr		R		Length					
Byte 4 >	Requester ID								Tag				Last DW BE		1st DW BE		R							
Byte 8 >	Address[31:2]																							

图 4 存储器写请求 TLP 头格式

Fig. 4 Specification of memory write TLP head

请求者通过填写 TLP 头内的正确信息以及地址并且将数据放在 TLP 头的后面发送给接收者,接收者便会解析包内的信息,并将正确的数据放到自己的地址空间中。

在具体的 FPGA 设计中,对于存储器写请求,事务层接口中的数据接收控制器接收到 TLP 头和数据之后,只需要通过 BAR 地址空间管理模块将数据存入寄存器或者双口缓冲区中即可。存储器读请求则稍微复杂一些,请求者向接收者发送存储器读请求 TLP,在这个数据包中没有有效数据负载,其 TLP 头的格式与存储器写请求的 TLP 头相同,接收者根据 TLP 头中的“Type”字段、“Fmt”字段区别操作类型。另一方面,当接收者在收到这个存储器读操作后,将请求的数据以完成报文的形式发送给请求者,请求者通过解析完成报文可读到有效的数据。图 5 所示为完成报文的 TLP 头格式。

	+0				+1				+2				+3																			
Byte 0 >	7	6	5	4	3	2	1	0	7	6	5	4	3	2	1	0	7	6	5	4	3	2	1	0	7	6	5	4	3	2	1	0
Byte 4 >	Completer ID				Reserved				Reserved				Reserved				Comp	Status														
Byte 8 >	Requester ID				Tag				Reserved				R																			

图 5 完成报文的 TLP 头格式

Fig. 5 Specification of completion packet TLP head

3 PCIe 的 DMA 传输设计

PCIe 的 DMA 传输设计分成两个部分,即 DMA 写操作(请求者到接收者)和 DMA 读操作(接收者到请求者)。由于 FPGA 内嵌的 PCIe 硬核的 PIO 模块本身已经支持 32 位地址空间的单字存储器读写请求、64 位地址空间的单字存储器读写请求和 IO 读写请求。本设计主要是在 PIO 模块的基础上完成对 32 位地址空间 DMA 传输方式的存储器读写请求的支持。同时,考虑到实际应用中 DMA 控制器一般只支持数据 8 字节对齐的情况,所以本设计也只支持数据 8 字节对齐的 DMA 传输方式的读写访问。

3.1 DMA 传输方式的存储器写操作

对于存储器写请求来说,实现的主要方法是数据接收控制器将接收的 TLP 头中各个字段正确解析,并区分是单字写还是 DMA 传输写,确定写操作的数据长度,如果是单字写(数据长度字段值为 1),就将接下来的 64 位数据存入寄存器或者双口缓冲区中;如果是 DMA 方式写(数据长度字段值大于 1),就将数据按照 64 位写入双口缓冲区中,每个时钟周期,双口缓冲区的地址自动增加,直到数据全部写入双口缓冲区中。DMA 传输方式存储器写状态机如图 6 所示。

图 6 中 RX_MEM_WR64_DW1DW2 状态、RX_MEM_

WR32_DW1DW2 状态为 PCIe 的 32/64 位单字存储器写操作,RX_IO_WR_DW1DW2 状态为 PCIe 的 IO 写操作。在这些状态下,写完 TLP 数据包中有效数据负载后即可跳转 RX_RST_STATE 状态。在 DMA 传输方式下,RX_MEM_WR_BST_QW1 状态处理第一个 32 位数据状态(TLP 头包含 3 个 32 位大小空间,需要分两次 64 位进行解析),RX_MEM_WR_BST_QWN 状态处理剩余有效数据负载,并将数据写入双口存储器中。另一方面,RX_IO_MEM_RD32_DW1DW2 状态以及 RX_MEM_RD64_DW1DW2 状态为请求者发送 32/64 位读请求命令,在此类 TLP 数据包中就不包含有效数据负载,在解析 TLP 头后,将读操作的地址空间即操作类型等关键信息传至数据发送控制。

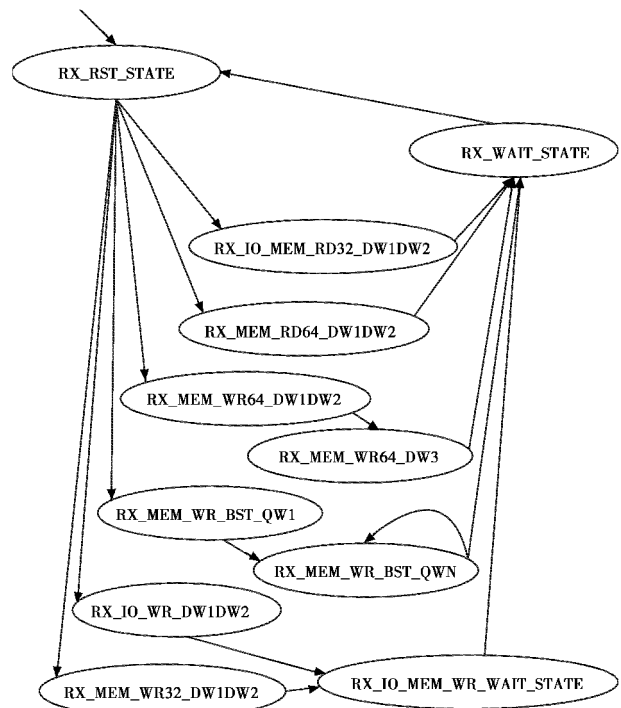


图 6 DMA 传输方式存储器写状态机

Fig. 6 Memory write state machine of DMA

3.2 DMA 传输方式的存储器读操作

对于存储器读请求,数据发送控制器将根据数据接收控制器发送来的读地址空间信息从双口存储器中读取数据,然后数据发送控制器将读到的数据封装成若干个读完成数据报文,将数据发送至读操作请求者。DMA 传输方式存储器读状态机如图 7 所示。其中,TX_CPL_QW1 状态为发送不带数据的完成报文,TX_CPLD_QW1 状态发送只带一个有效数据负载的完成报文。在 DMA 传输方式下,TX_CPLD_BST_HEADER 状态发送完成报文头,TX_CPLD_BST_QW1 状态发送完成报文头的最后一个双字和第一个有效负载数据,TX_CPLD_BST_QWN 状态发送有效数据负载。

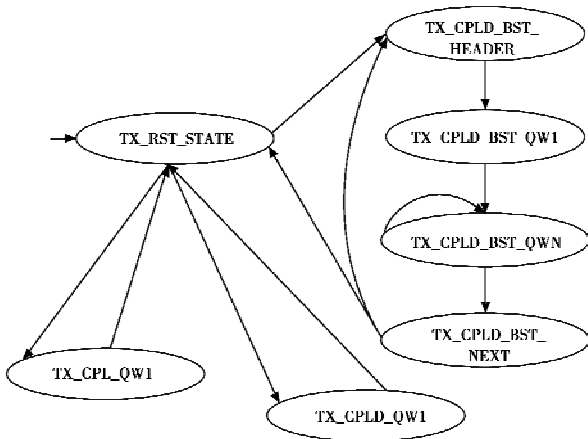


图 7 DMA 传输方式存储器读状态机
Fig. 7 Memory read state machine of DMA

3.3 DMA 传输的周期拆分

在 PCIe 总线中,请求者在发出一个存储器读请求之后,可能会收到目标设备回的多个完成报文,因为在 PCIe 总线中,一个存储器读请求最多可以请求 4 kB 大小的数据报文,当读请求跨越了 RCB(Read Completion Boundary)边界时,目标设备则需要分多个完成报文才能将数据传递完毕。在 PCIe 设备的配置空间中,Link Control 寄存器的 RCB 位决定了读完成报文的边界,本文设计为 64 字节。

因此,当请求者向本文所设计的 PCIe 接口发起读请求后,数据发送控制器则会按照 RCB 边界进行完成周期拆分,具体的算法流程如图 8 所示。

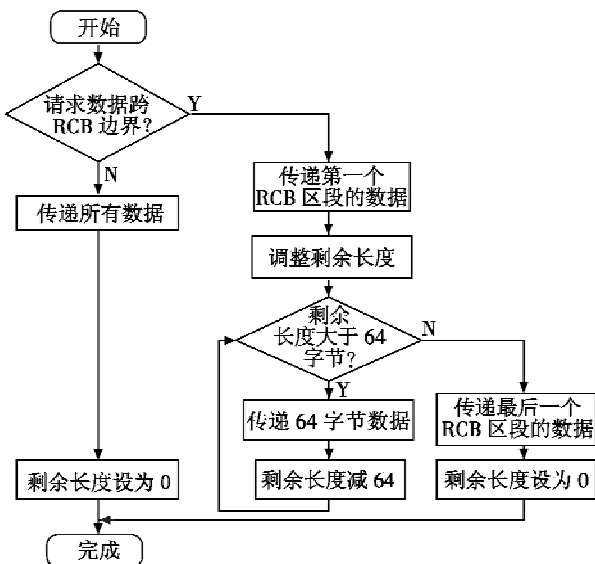


图 8 DMA 传输的周期拆分
Fig. 8 Cycle split of DMA

对于数据请求长度在 RCB 边界之内的读请求来说,数据发送控制器在第一个 DMA 传输周期内完成所有数据的传输;对于数据跨 RCB 边界的读请求来说,先处理第一个周期的数据,使得剩余数据的起始地址

按 64 字节对齐,接下来每个 DMA 周期传递 64 字节,直到数据传递结束。

4 设计实现与性能分析

由于本文设计的 PCIe 总线的 DMA 传输的实现方式是硬件描述语言(VHDL),因此设计仿真是整个系统设计中十分关键的环节^[9]。设计完成后,本文通过仿真软件(modelsim)对其进行功能仿真,同时实现 TestBench 作为激励加载单元,对接收数据控制器和发送数据控制器进行功能仿真验证。当功能仿真正确后,在实践中进一步使用 ChipScrop 抓图,主要分析通过本文所设计的数据接收控制器和数据发送控制器所产生的 PCIe 总线传输信号的时序是否满足要求。

PCIe 总线 DMA 传输中的写过程相对比较简单,其时序如图 9 所示,从捕获的波形可以看出,在 DMA 传输过程中数据接收控制器正确解析 DMA 写数据报文,并成功写入 PCIe 的数据缓冲区。

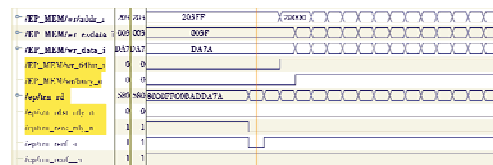


图 9 DMA 写时序

Fig. 9 Write sequence of DMA

读操作的过程相对复杂,数据接收控制器接收到读请求后启动数据发送控制器,数据发送控制器将请求者发起的一个存储器读请求分成多个存储器读完成报文进行数据传输,其时序如图 10 所示。

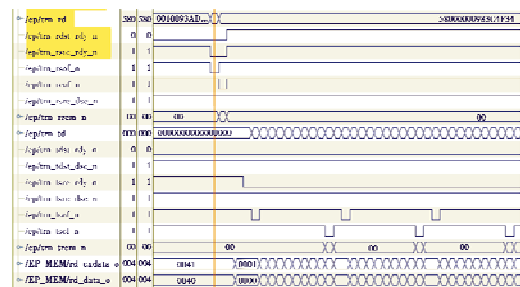


图 10 DMA 读时序

Fig. 10 Read sequence of DMA

经实验测试,DMA 访问与单字访问相比,写速度提升了近 30 倍,读速度提升了约 80 倍,具体读写性能如表 1 所示,可见,采用 DMA 传输方式的数据传输能够有效提高整个系统的传输带宽。

表 1 读写带宽测试

Table 1 Bandwidth test of read and write operation
MByte/s

测试项	测试带宽	测试项	测试带宽
单字写	6.13	DMA 写	167
单字读	4.45	DMA 读	370

5 结束语

使用 FPGA 来设计基于 PCIe 总线的数据传输,可以省去专用的 PCIe 接口芯片,降低硬件的设计成本,提高硬件集成度的同时还能利用 FPGA 的可编程特性提高设计的灵活性与适应性^[10]。与此同时,目前很多 FPGA 内都嵌入了 PCIe 的硬核,此硬核支持 PCIe 传输的单字读写操作以及 DMA 读写操作,由于 PCIe 具有一定的协议开销,其单字传输效率不甚理想,因此,本文在 PCIe 单字传输的基础上提出了一种基于 PCIe 接口的 DMA 传输的设计方案,此方案可以稳定快速地实现 PCIe 总线的 DMA 传输。经测试,DMA 传输方案在传输带宽方面满足设计要求。

参考文献

- [1] 马鸣锦,朱剑冰. PCI、PCI-X 和 PCI Express 的原理及体系结构[M]. 北京:清华大学出版社,2007. (MA M J, ZHU J B. Architecture and principle of PCI, PCI-X and PCI Express [M]. Beijing: Tsinghua University Press, 2007.)
 - [2] BUDRUK R. PCI Express 系统体系结构标准教材[M]. 田玉敏,王崧,张波,译. 北京:电子工业出版社,2005. (BUDRUK R. PCI Express system architecture [M]. Translated by TIAN Y M, WANG S, ZHANG B. Beijing: Electronic Industry Press, 2005.)
 - [3] 王齐. PCI Express 体系结构导读[M]. 北京:机械工业出版社,2011. (WANG Q. Introduction of PCI Express architecture[M]. Beijing: China Machine Press, 2011.)
 - [4] 汪精华,胡善清,龙腾. 基于 FPGA 实现的 PCIe 协议的 DMA 读写模块[J]. 微计算信息,2010,26(10):7-9. (WANG J H, HU S Q, LONG T. The implementation of DMA reading and writing module on FPGA based on PCI Express protocol [J]. Microcomputer Information, 2010, 26(10):7-9.)
 - [5] PCI Special Interest Group. PCI Express base specification revision 1.0a[EB/OL]. [2014-07-21]. <http://netyi.net/Book>.
 - [6] Xilinx. LogiCore™ IP endpoint for PCI Express V3.7 user guide[EB/OL]. [2014-07-21]. <http://www.xilinx.com>.
 - [7] 王嘉良,赵曙光. 用 FPGA 实现 PCI-E 接口和 DMA 控制器设计[J]. 计算机技术与发展,2011,21(6):181-184. (WANG J L, ZHAO S G. PCI-E interface implementation and DMA engine design with FPGA[J]. Computer Technology and Development, 2011, 21(6):181-184.)
 - [8] 石峰,吴建飞,刘凯. 基于 Xilinx FPGA 的 PCIE 接口实现[J]. 微处理机,2008(6):19-21. (SHI F, WU J F, LIU K. PCIE port develop based on Xilinx FPGA[J]. Microprocessors, 2008(6):19-21.)
 - [9] STILIADIS D, VARMA A. Selective victim caching: a method to improve the performance of direct-mapped caches [J]. IEEE Transactions on Computers, 2007, 46(5):603-610.
 - [10] 闫振兴. 基于 FPGA 的 PCI Express 接口的设计与实现[D]. 北京:北京理工大学,2009. (YAN Z X. The design and implementation of PCI Express interface based on FPGA [D]. Beijing: Beijing Institute of Technology, 2009.)
-
- (上接第 83 页)
- frared system constellation [C]// Proceedings of IEEE Aerospace Conference, 2001:385-393.
 - [3] 刘仁,梁维泰,赵敏. 一体化空天防御系统[M]. 北京:国防工业出版社,2011. (LIU R, LIANG W T, ZHAO M. Integrated aerospace defense system [M]. Beijing: National Defense Industry Press, 2011.)
 - [4] ANDREAS N S. Space-Based Infrared System (SBIRS) system of systems [C]//Proceedings of IEEE Aerospace Conference, 1997:429-438.
 - [5] SMITH M S. Military space programs: issues concerning DOD's SBIRS and STSS programs [R]. USA: Library of Congress Washington DC Congressional Research Service, 2005.
 - [6] BEAULIEU M R. Launch detection satellite system engineering error analysis [D]. Monterey: Naval Postgraduate School, 1996.
 - [7] DANIS N J. Space-based tactical ballistic missile launch parameter estimation [J]. IEEE Transactions on Aerospace and Electronic Systems, 1993, 29(2):412-424.
 - [8] 刘刚,吴诗其. 区域性星座设计[J]. 系统工程与电子技术,2002,24(7):15-18. (LIU G, WU S Q. Designing satellite constellations for partial coverage [J]. Systems Engineering and Electronics, 2002, 24(7):15-18.)
 - [9] 张育林,范丽,张艳,等. 卫星星座理论与设计[M]. 北京:科学出版社,2008. (ZHANG Y L, FAN L, ZHANG Y, et al. Theory and design of satellite constellations [M]. Beijing: Science Press, 2008.)
 - [10] 张毅,杨辉耀,李俊莉. 弹道导弹弹道学[M]. 长沙:国防科技大学出版社,1999. (ZHANG Y, YANG H Y, LI J L. Missile ballistics [M]. Changsha: National University of Defense Technology Press, 1999.)
 - [11] TORRIERI D J. Statistical theory of passive location systems [J]. IEEE Transactions on Aerospace and Electronic Systems, 1984, 20(2):183-198.
 - [12] HAIM Z B, ELDAR Y C. On the constrained Cramer-Rao bound with a singular fisher information matrix [J]. IEEE Signal Processing Letters, 2009, 16(6):453-456.
 - [13] BISHOP A N, FIDAN B, ANDERSON B D O, et al. Optimality analysis of sensor-target geometries in passive localization: part 1-bearing-only localization [C]//Proceedings of 3rd International Conference on Intelligent Sensors, Sensor Networks and Information (ISSNIP), 2007:7-12.