

光纤通道仲裁环网的容错技术研究

张 磊, 王卫东, 惠晓强

(中国航空计算技术研究所, 陕西 西安 710068)

摘 要: 为了满足机载分布式计算系统内部数据传输的高带宽和高可靠性需要, 文中提出了面向机载分布式计算系统应用的光纤通道仲裁环网的容错技术, 即硬件冗余和重构算法。重点讨论了光纤通道接口设计, 四余度光纤通道仲裁环网的互连结构, 以及四余度光纤通道仲裁环的故障分类、模式切换和重构算法。

关键词: 光纤通道; 仲裁环; 容错; 冗余; 重构

中图分类号: TP393

文献标识码: A

文章编号: 1000-7180(2011)11-0010-04

Fault-Tolerance Technology Research in Fibre Channel Arbitrate Loop

ZHANG Lei, WANG Wei-dong, HUI Xiao-qiang

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

Abstract: In order to satisfy the requirement of high throughput and high reliability of data transfer in airborne distributed computation system, this paper introduces fault-tolerance technologies based on fiber channel arbitrate loop (FC-AL) including hardware redundancy and reconstruction arithmetic. The emphases of discussion are design of Fiber Channel interface circuit, interconnection structure of four-redundancy FC-AL. Also this paper discusses fault classifications, mode switch and reconstruction arithmetic about FC-AL.

Key words: fibre channel; arbitrate loop; fault-tolerance; redundancy; reconstruction

1 引言

光纤通道(Fibre Channel, FC) 是美国国家标准委员会制定的集成了通道和网络优点的高速串行通信标准, 由于该协议具有高传输率、低误码率、支持多种拓扑结构、服务类型、上层协议、底层传输介质和流控制功能的特点, 使得它成为先进军用航电系统互连的首选^[1]。

FC 协议支持点到点、网络交换式结构和环网三种拓扑结构。其中点到点的拓扑结构不能实现多节点的互联通信, 故不作考虑。网络交换结构具有可靠性高、支持路由功能、扩展性好, 数据传输速率和链路利用率高等优点, 所以网络交换结构是航电系统通信最优的拓扑结构, 但是其结构复杂, 研制周期长, 成本较高。如果在节点数不多(几个到十几个)的小型系统中, 环形网可以较好的综合点到点拓扑和

网络交换结构的优缺点, 可以为系统互连提供一种高性能价格比的解决方案^[2]。光纤通道仲裁环网(Fiber Channel Arbitrate Loop, FC-AL)作为光纤通道中一种重要的拓扑结构, 提供了在共享式光纤网络中把多个节点连接在一起的方法, 它比点对点形式更灵活并且支持更多设备, 而且比交换式网络结构更经济, 在机载分布式系统中得到了广泛的应用^[3]。

2 FC-AL 初始化

FC-AL 初始化是在新节点加入环时为其分配仲裁环物理地址(AL_PA)的过程, 整个过程由环端口状态机(LPSM)完成状态转换^[4-6]。

(1) 环初始化

当端口全部处于激活状态时, 将不断发送“环路初始化原语”, LPSM 连续接收到三个 LIP 后自动

转入 OPEN-INIT 状态.

(2) 选择主机

端口发送环初始化主机选择帧(LISM)并监视接收到的 LISM 帧,当端口从其他端口接收到第一个 LISM 帧,就将该帧包含的端口名与本端口名比较,如果接收到的端口名具有较高优先权,它将转发收到的 LISM 帧;否则,它将再次发送本端口的 LISM 帧;如果接收的 LISM 端口名是自己,它就成为主机.

(3) 分配 AL-PA

主机依次启动环初始化光纤网地址帧(LIFA)、环初始化原始地址帧(LIPA)、环初始化硬地址帧(LIHA)和环初始化软地址帧(LISA),为环路上的节点分配 AL_PA.

(4) 位置映射

主机发送环初始化位置报告帧(LIRP),第一个收到 LIRP 帧的端口将自己的 AL_PA 按从前至后的顺序放入 LIRP 的净荷内,继续转发 LIRP 帧,以此类推,当主机收到该 LIRP 帧时,所有端口都填入了自身 AL_PA,同时获知环路中其它端口的 AL_PA,环路初始化结束.

3 FC-AL 容错技术

3.1 容错设计

光纤通道仲裁环上所有节点共享传输带宽,数据沿环单方向流动,各节点通过仲裁获得环的使用权^[7].不相邻的节点进行数据传输需要经过中转节点,一个节点出现故障,将造成整个环断路,因此仲裁环需要进行多余度设计.机载分布式计算系统由若干计算节点构成,每个节点配置一个网关模块,网关负责各节点之间互联通信,完成节点内部与外部通信协议之间的转换.网关的外部接口采用 FC 通道容错设计,实现节点之间的四余度 FC-AL 互连^[8].

3.1.1 容错设计方法

FC 接口采用时间容错和硬件容错相结合的设计方法.

(1) 时间容错

时间容错是以牺牲时间来换取计算机系统的高可靠性的一种手段,可以有两种方式:一种是有限度地降低机器的速度来增加系统的可靠性.另一种是以重复执行指令或加入等待时间来消除瞬时故障带来的影响,可以细分为指令复执和程序卷回.在机载实时容错分布式计算机系统原理样机的研制中为了

屏蔽 FC 接口的瞬时故障,网关采用了指令复执的容错方法.

(2) 硬件容错

容错系统的工作方式可分为静态余度、动态余度和混合余度.静态余度通过表决和比较,屏蔽系统中出现的错误,达到容错目的,该技术虽然容错速度较快,但需要大量的硬件资源;动态余度技术的主要方式是多重储备电路相继运行来维持系统正常工作,当检测到工作的电路出现故障时,即用一个备用电路顶替并重新运行,并将故障电路旁路掉,显然这里有一系列检测、切换和恢复的过程,该技术的硬件资源利用率很高,但在动态切换的过程中需要以时间为代价;混合余度是静态和动态两种余度技术的结合.为了提高 FC 接口的硬件资源利用率,网关采用了动态余度技术.

3.1.2 双余度 FC 接口设计

双余度 FC 接口电路是以 FC 协议处理器为核心,由两个硬件配置完全相同、互为备份的电路组成,支持与主机连接的 PCI 总线接口.由于光电收发器件的可靠性较低,在机载环境下会出现环路瞬间中断.因此,在电路设计时一个 FC 通道配置两个光电收发器、一个交叉开关 CSM 和一个旁路电路 PBC;两个 FC 通道之间交叉互连,以共享四个光电收发器;当系统检测到某一环路故障时,自动切换至其它环路,保证系统内各个节点的互连互通.这种资源共享、灵活配置的方法增强了 FC-AL 的健壮性.

如图 1 所示,FC 协议处理器与光电收发器的配对关系共 8 种.首先定义 FC 协议处理器 x 到光电收发器 y 的映射为 $FCCx \rightarrow PORTy$,其中 $x=1,2$, $y=1,2,3,4$.根据该定义,8 种映射关系为: $FCC1 \rightarrow PORT1$ 、 $FCC1 \rightarrow PORT2$ 、 $FCC1 \rightarrow PORT3$ 、 $FCC1 \rightarrow PORT4$ 、 $FCC2 \rightarrow PORT1$ 、 $FCC2 \rightarrow PORT2$ 、 $FCC2 \rightarrow PORT3$ 、 $FCC2 \rightarrow PORT4$.通过硬件逻辑检测故障后自动配置 PBC,实现光电收发器的输入信号经 PBC 环绕到发送端输出,并不经过 FC 协议处理器,进而将 FC 协议处理器 x 旁路;当两个 FC 协议处理器都被旁路时意味着当前节点从 FC-AL 上旁路.

3.1.3 通道控制逻辑

为了检测故障并控制 FC 容错电路,网关在实现时采用可编程控制逻辑设计了一系列状态、控制寄存器.网关上的主机通过访问这些寄存器就获得 FC 容错电路当前的工作状态,并对其进行控制.为保证 FC 接口电路状态的确定性,上电复位后网关的通道控制逻辑自动配置两个旁路器,使得四个光

一AL断开都将引发FC协议处理器向网关的主机发送中断请求,主机首先执行故障滤波,若该故障为永久性故障,选择可用资源,再进行模式切换,并将故障环的编号记录在资源健康表中。

3.3.4 模式切换

模式切换的实质是在预先获知网关硬件资源健康状况的基础上,将故障单元隔离,选择有效的工作

模式,建立可用的FC-AL。环网中任意一个网关进行模式切换都将引发其它网关的模式切换,从环网故障到模式切换,再到环网建立需要数百毫秒时间,包括控制命令的发出、模式切换、环网初始化、节点登录和系统识别。如图3所示为四余度FC-AL容错算法,体现了模式切换的全过程。该算法实现的前提条件是交叉开关和旁路器工作正常。

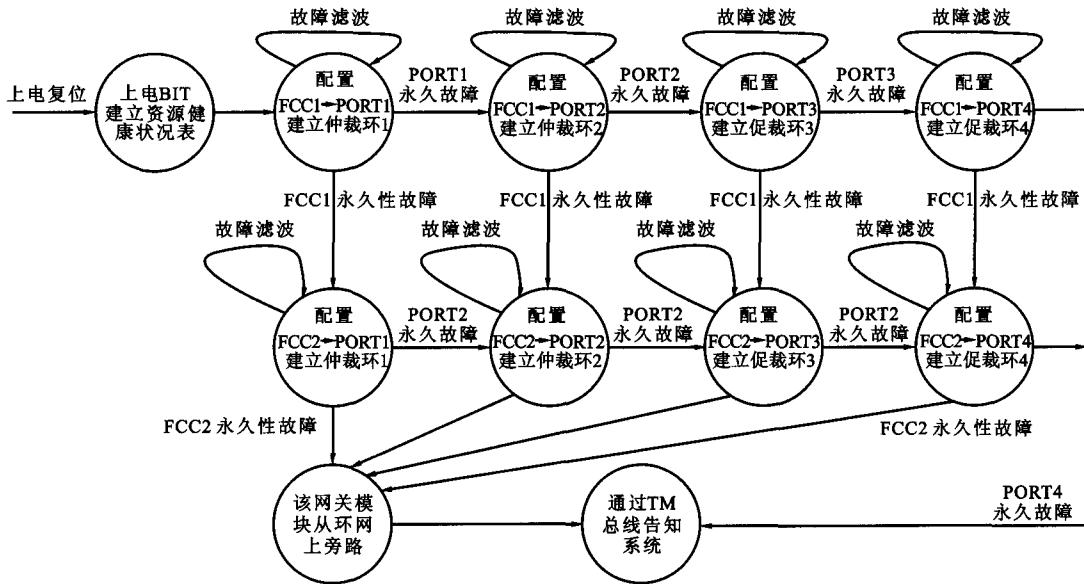


图3 四余度FC-AL容错算法

上电复位后网关进行BIT测试,获知其硬件资源的健康状况,并选择FCC1→PORT1建立连接,进行故障滤波,等待其它网关登录,组成FC-AL1;在正常工作中可能出现突发的环路断开,通过故障滤波将会把瞬时故障屏蔽掉,当确定是PORT1永久性故障引起的环路断开时,工作模式将切换到FCC1→PORT2,通过同样的方式建立FC-AL2。只有当四个光电收发器都确定为永久性故障时,FC-AL无法建立。然而发生这种事情的概率极小。如果在建立FCC1→PORT1连接后,发生环路断开的原因是FC协议处理器1永久性故障造成的,工作模式会切换到FCC2→PORT1,进行故障滤波,等待其它网关登录,组成FC-AL1;如果随后发生了FC协议处理器2永久性故障,网关将工作模式切换到节点旁路,通知系统进行降级重构。降级重构后的FC-AL有四个节点,可以继续采用该算法实现FC-AL的容错。该算法具有实现简单、效率高、突出实际应用的优点,适合小规模FC-AL应用。

4 结束语

机载分布式计算系统中的网关采用双余度FC

接口设计,实现了四余度光纤通道仲裁环网的容错控制,不仅保证该系统的实时性和可靠性,而且有效地降低了系统互连的成本,在军用航空电子系统领域具有较好的应用前景。

参考文献:

- [1] 林强,熊华钢,张其善. 光纤通道概述[J]. 计算机应用研究, 2006, 7(2): 9-13.
- [2] American National Standards Institute. Fiber channel arbitrated loop(FC-AL-2)Rev7.0[S]. 1995.
- [3] 高昆,凡叔军,徐海静,等. 光纤通道应用系统的设计方法[J]. 光电工程, 2004, 31(11): 31-34.
- [4] Alan F Benner. 存储区域网络光纤通路技术[M]. 胡先志,胡佳妮,译. 北京:人民邮电出版社, 2003.
- [5] 咸立民,裴如霞,沈绪榜. 光纤通道组帧控制器的研究与设计[J]. 微电子学与计算机, 2009, 26(5): 130-134.
- [6] 郑魁,杨志义,曾小芸. 基于Linux的光纤通道网卡驱动程序开发[J]. 微电子学与计算机, 2007, 24(1): 88-90.
- [7] 周敬利,韦丰,林芝,等. 基于光纤通道仲裁环的存储区域网的分析与实现[J]. 小型微型计算机系统, 2004, 25(1): 16-20.

(下转第17页)

压缩有很好的效果.

表 4 实验结果比较

	WinZIP	WinRAR	文中算法
Lena	6.572 3	5.476 8	4.905 6
Camera	5.944 6	5.292 0	5.216 0
House	5.796 9	4.553 6	4.626 4
Finger	6.880 0	6.584 8	4.945 6
平均值	6.298 5	5.476 8	4.923 4

3.2 算法改进

本算法还有可以改进的地方,数据的读入扫描还可以与分块扫描相结合,在对原始数据进行编码以后产生了大量的二值码表,在这些方面还有一些压缩空间有待进一步的挖掘.

4 结束语

文中针对图像无损压缩提出了新的算法,它利用图像数据的相关性和冗余性,把图像的原始灰度值数据按照差值的变化动态分段进行处理.算法中对扫描方法和差值的选择使得本算法有了一定的自适应性.与传统的压缩算法相比,文中提出了一种比较独特的思路.实验证明了本算法具有很好的压缩效果.

参考文献:

[1] Rafael C Gonzalez, Richard E Woods. 数字图像处理

[M]. 2 版. 阮秋琦,阮宇智,译. 北京:电子工业出版社, 2007:326—364.
[2] 张春田,苏育挺,张静. 数字图像压缩编码[M]. 北京:清华大学出版社,2006.
[3] David S Taubman, Michael W Marcellin. JPEG2000 图像压缩基础、标准和实践[M]. 魏江力,柏正尧,译. 北京:电子工业出版社,2004.
[4] 王学武,石跃祥. 对图像灰度级分段的压缩编码[J]. 计算机工程与设计,2006,27(2):222—223.
[5] Tian Jun. Reversible data embedding using a difference expansion[J]. IEEE Transaction on Circuits and Systems for Video Technology,2003,13(8):890—896.
[6] 孙全党,高美真. 基于颜色空间分布熵的图像检索[J]. 微电子学与计算机,2009,26(10):134—136.
[7] 谢颢,范九伦. 一种基于位运算的数字图像加密算法[J]. 微电子学与计算机,2009,26(12):117—120.
[8] 陈开英,胡永健,李建伟. 利用差值扩展进行可逆数据隐藏的新算法[J]. 计算机应用,2008,28(2):455—459.
[9] 孙鑫,余安萍. VC++深入详解[M]. 北京:电子工业出版社,2006.

作者简介:

高 健 男,(1961—),博士,副教授. 研究方向为图像处理、计算数学.
宋 奥 男,(1987—),硕士研究生. 研究方向为图像处理.

(上接第 13 页)

[8] 张磊,李鹏,冯晓东. 基于光纤通道的网关模块设计与实现[J]. 微电子学与计算机,2007,24(6):56—58.

作者简介:

张 磊 男,(1976—),硕士,工程师. 研究方向为民机航空

电子系统.
王卫东 男,(1963—),博士,研究员,硕士生导师. 研究方向为分布式计算系统.
惠晓强 男,(1978—),硕士,工程师. 研究方向为嵌入式计算机.