

一种用于信令寻径高速光纤交换网的信令降频方法

刘君瑞¹ 陈颖圉² 樊晓桠¹ 康继昌¹

(西北工业大学计算机学院 西安 710072)¹ (中航一集团第六三一研究所一室 西安 710068)²

摘要 针对高速光纤交换网络高质量的交换服务需求,提出了一种适用于信令寻径式高速光纤通道交换机的信令降频方法。该方法采用一套特殊的序列码对信令帧内容进行编码后传输,以达到信令帧信号频率成倍降低的目的,使交换机无需借助串并转换器对高频信令帧进行降频,就能够直接识别信令帧内容。试验结果显示,采用信令降频方法后,交换机的结构显著简化,可靠性得到很大提高,并且硬件成本大大降低。

关键词 光纤通道,信令寻径,信令降频码

中图分类号 TP393.02 文献标识码 A

Research of the Reducing Token's Frequency Method in Token-routing High-speed Fibre Channel Switch Network

LIU Jun-rui¹ CHEN Ying-tu² FAN Xiao-ya¹ KANG Ji-chang¹

(College of Computer Science, Northwestern Polytechnical University, Xi'an 710072, China)¹

(The 1st Department, Aeronautical Computing Technique Research Institute, Xi'an 710068, China)²

Abstract In order to acquire high speed switching in high-speed fibre channel switch network, a method applying to the token-routing fibre channel switch to reduce the token's frequency was put up. In this scheme, the token is coded with "the reducing-frequency code", that is a special sequence of '0' and '1', then it's frequency may be reduced multiply, and the switch can extract the information of token without the serial-to-parallel converter. The results show that the structure of switch is simplified while its reliability is greatly improved and its costs is significantly decreased.

Keywords Fibre channel, Token-routing, Reducing-frequency code

1 引言

目前,长距离高速通信的主要方式是采用串行方式进行数据传输,其实现机制一般是在传输过程中采用串行方式。由于发送或接收方电路无法分析太高频率的信号,因此需要经过串并转换之后才能进行接收或者解析后转发。以太网标准和 FC 标准中,每个数据帧携带自己的路径信息,传输过程由中间节点打开每个数据帧,提取路径信息,进行下一步路由控制。一直以来,用 FPGA 来接收高速数据始终是一个难题。信息源器件在发送信息时需要经过“并/串转换器件”把低频并行信号转换成串行信号,使信号频率成倍提高;而信息接收端器件需要“串/并转换器件”把串行信号转换成并行信号,使信号频率成倍降低才能够处理。因此在进行高速数据传输时,需要对每个数据帧进行串并转换操作来降低信号频率,使通信电路能够进行正确的信息获取,从而完成数据帧路由。本文通过分析高速以太网和千兆光纤网中数据传输的缺点,在原有项目成果“先锋信令寻径式交换网络”的基础之上提出了一种适用于信令寻径式高速光纤通道交换机的信令降频方法,使高速信号传输网络体系结构中无需串并转换器件,

并且在只解析最少量通信帧的情况下就能够完成数据的高速、正确传输。

2 先锋信令寻径式交换网络

2.1 先锋交换网与信令寻径技术

目前,国内外互连网络广泛采用基于消息头寻径的通信技术。由于消息头和消息体不分离,因此在每条信道中均需缓冲区来暂存消息体,控制逻辑复杂。为弥补其不足,西北工业大学计算机学院康继昌教授提出一种“先锋信令寻径式交换网络”(简称“先锋交换网”或 Pionet),该项研究得到国家航空科学基金、国家中小企业创新基金、国家 211 重点实验室和杭州奥利金并行计算技术有限公司的支持,并已获得国家发明专利(专利号:ZL 99 1 15908. X,授权证书号:164242)。

先锋交换网采用独特的“信令寻径(Token-Routing)”技术,它可使安装有先锋网卡的计算机之间通过基于交叉点开关的信令寻径式交换机直接相连,进行端到端通信。其特征在于:由相当于消息头的信令来寻径,申请信道。信道建立后,才发送消息体,消息体可经交换机直接到达目标机。交换机只需自动检测申请信道、撤销信道等少数几种信令,数据则

到稿日期:2009-01-22 返修日期:2009-04-07 本文受浙江湖州市 2008 年科技创新项目“万亿次高性能计算集群的研究与设计”资助。

刘君瑞(1978-),女,博士生,讲师,主要研究方向为高性能网络、集群计算,E-mail:liu.junrui@nwpu.edu.cn;陈颖圉(1978-),男,工程师,主要研究方向为智能信息处理技术、集群计算;樊晓桠(1964-),男,教授,博士生导师,主要研究方向为微处理器系统结构、集成电路设计;康继昌(1930-),男,教授,博士生导师,主要研究方向为高性能计算、并行计算。

可在接通的信道中自由通过,无需缓存,故交换机的结构可大大简化,且成本降低、可靠性提高。

在先锋交换网中,寻径信令与普通的消息头不同,它不是以目的地址来寻径,而是以交换机的端口号来进行寻径,因此其地址很短、控制简单,交换机的结构(如图1所示)也可进一步简化。设 Node1 和 Node2 为连接在信令寻径式交换机上的两个节点,它们所对应的端口号分别为 i 和 j 。若 Node1 要向 Node2 传输数据,则其通信过程如图1所示。首先,Node1 要向交换机发送一条寻径信令,申请与 Node2 接通。交换机收到该信令后,查询目标端口 j 当前的工作状态:若正在通信,则向端口 i 返回“忙”信息;否则,接通端口 i 与端口 j 之间的信道,并向端口 i 返回确认信息。收到“忙”信息后,建立信道失败,Node1 根据需要重试(Retry)或放弃当前请求(Terminate)。收到确认信息后,Node1 可发送消息,这些消息沿着已建立好的信道直接到达目标端口(目标节点),无需交换机的干预。通信结束后,Node1 发出撤销信道请求,交换机拆除它与 Node2 之间的信道,从而释放信道资源。如果有多个源/目标对需进行通信,则交换机依次接通多条信道,使得这些源/目标对可并行通信,这样大大提高了信道利用率和系统聚合带宽。

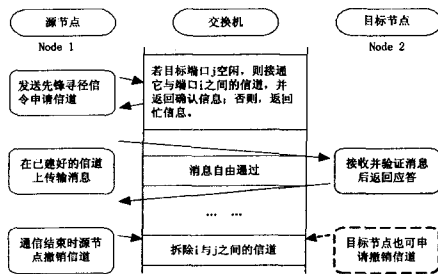


图1 先锋交换网通信过程

2.2 信令寻径式交换技术应用在高速信号传输中的不足

在信令寻径式光纤交换网络中,交换机依靠解析信令内容来控制信道的接通与断开,而数据信息则直接进入交换机信道进行转发,因此信令是交换机工作的重要依据。为了让交换机能够识别这些信令,现行的主流技术都是在交换机中增加串并转换器件,进行串并转换,以达到降频效果,使交换机电路能够识别解析信令。但这种方案带来两个问题:(1)增加串并转换器件,增加了许多高频信号布线,造成了交换机内部信号传输的不稳定。并且由于收发器的引出线很多,使得PCB板面积大幅度增加,交换机的硬件成本也随之增加。(2)通信过程中信令帧数量很少,为了少量的信令帧解析付出这些代价是值得商榷的。

为此,本文提出了一种适用于光纤通道交换机的信令降频编码方法。对信令帧进行降频后,交换机无需增加串并转换器件,即可提取信令,显著简化了交换机的结构,解决了PCB板设计难题,使得交换机的工作稳定性得到极大提高。

3 信令降频方法

3.1 信令降频码编码方法

3.1.1 编码规则

在高速信号通信中,发送设备多采用“8b/10b”编码方法对用户的8位数据进行10位编码后发送,以提高信号传输质量。鉴于此,本文采用的方法中信令降频码的编码遵循8b/

10b 编码规则。信令降频码有两种编码形式:“1111100000”和“0000011111”,分别表示信息‘0’和‘1’。从降频码的编码序列中可以看到有连续的5个‘1’和5个‘0’,降频码进行串并转换后频率就相当于降低为初始状态下信号频率的五分之一。因此,用于检测信号的时钟频率至少可以降低为原时钟频率的五分之一,从而能够正确识别降频码内容。

3.1.2 信令帧格式

信令帧中至少应包括3个内容:信令帧头 HD、信令类型 OP 以及端口信息 ID。

信令帧头是用于区分数据帧和信令帧的特殊标志,不能与正常数据序列冲突,因此本方法可使用降频码两种编码序列的交替重复来编码信令帧头,即信令帧头为“000001111111111100000”序列的重复。一般情况下,把这个序列重复两次就可以和正常数据区分开,所以信令帧头定义为“00000111111111110000000001111111111100000”。通过增加信令帧头中“000001111111111100000”序列的重复次数可以大大提高信令帧识别可靠性。

信令类型、端口信息与信令帧头的编码机制相同。如果实际应用中交换机有16种信令,那么在此方法中可以连续使用4个降频码的16种组合进行编码。扩展用于编码信令类型的编码长度,可以编码更多的信令种类。同样,如果交换机中共有16个端口,那么端口信息也可以连续使用4个降频码的16种组合进行编码。如果交换机的端口数量增多,可以扩展编码。

使用降频码编码后,信令帧格式如图2所示,与没有经过降频编码的信令帧长度相当。

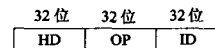


图2 降频码编码的信令帧格式

3.2 信令帧的解析方法

使用信令降频码编码后的信令帧,其解析分为3个步骤:识别信令帧头、解析信令类型以及获取端口信息。使用信令降频码进行编码后,信号频率降为原来信号频率的五分之一,因此在检测时使用的时钟频率比原有信号的检测时钟频率降低了很多。

由于没有办法固定检测时钟与信令的高频串行信号之间的相位差,为了保证不出现信令帧检测遗漏的现象,本方法中使用了两个频率都为原信号频率五分之一、反相的时钟,每个时钟都在其上升沿到来时检测信令帧,最少会有一个时钟在其上升沿检测到正确无遗漏的信令内容。因此,使用信令降频码进行编码后,信号的检测时钟频率至少降低为原信号检测时钟频率的十分之一。

下面对本方法的正确性进行证明。

证明:信令信号进行降频编码之后,频率降为原信号频率的五分之一。设降频后信号最高频率为 f ,从降频后的信令帧波形图进行分析可以得知,这是一个非周期性的连续信号。因此,降频后的信令信号最少会维持一个固定状态的时间长度 $T=1/2f$ 。依据信令降频码的编码规则可知,降频后的信令信号维持一个固定状态的时间只有两种可能性,分别为 $1/2f$, $1/f$,即为 T 的整数倍。只要保证在每个信号固定状态的维持时间段内都进行了信号采样,就能获得正确的信令内容。下面就只需要证明任何情况下上述两个时钟只要有一个能够保

证在每个信号维持状态的 T 时间段内进行了采样即可。

假设信令信号在时刻 t 开始,那么信令内容中所有状态的维持时间段可用下式表示:

$$T_s \in [t+nT, t+(n+1)T] \quad n=0,1,2,\dots \quad (1)$$

设 $clk0, clk1$ 是由同一个时钟源产生的满足上述条件的两个检测时钟,其中相位最早的时钟为 $clk0$,其次为 $clk1$ 。设时钟 $clk0, clk1$ 的频率为 f_c ,周期为 T_c ,则有 $f_c = f, T_c = 1/f_c = T$ 。两个时钟之间的相位相差为 $T_c/2$ 。如果 $clk0$ 在信令信号开始之后第一个上升沿的时刻为 $t+\Delta$,其中 Δ 是由时钟和串行信号之间相位差形成的,则 $clk0$ 时钟进行信号采样的时刻,即其上升沿时刻可用下式表示:

$$T_{clk0 \uparrow} = t + \Delta + gT_c = t + \Delta + gT \quad g=0,1,2,\dots \quad (2)$$

同样 $clk1$ 进行信号采样的时刻可用式(3)来描述:

$$T_{clk1 \uparrow} = t + \Delta + T_c/2 + kT_c = t + \Delta + T/2 + gT \quad k=0,1,2,\dots \quad (3)$$

由于时钟和串行信号之间的相位差有可能造成 $clk1$ 比 $clk0$ 的时钟上升沿先碰到信令信号,但是其可能早一个时钟周期碰到。

下面证明不论 Δ 值为多少,这两个时钟的采样时刻至少有一个位于式(1)描述的时间域中。

从系统分析得知, Δ 值域为 $[0, T_c)$,下面分别讨论 Δ 取值域中不同值时的时钟上升沿的情况:

当 $\Delta=0$ 时, $g=n, k=n$

$$T_{clk0 \uparrow} = t + gT_c = t + nT \in [t+nT, t+(n+1)T]$$

$$T_{clk1 \uparrow} = t + T_c/2 + kT_c = t + T/2 + nT \in [t+nT, t+(n+1)T]$$

从上述计算可以看到, $\Delta=0$ 时,时钟的信号采样时刻都位于式(1)描述的时间域中。但是因为信号稳定需要一定时间,所以这种情况下 $clk1$ 可作为采样时钟。

(1) 当 $\Delta \in (0, T_c/2]$ 时, $g=n, k=n$

$$T_{clk0 \uparrow} = t + \Delta + gT_c = t + \Delta + nT \in (t+nT, t+T/2+nT]$$

因此有 $T_{clk0 \uparrow} \in [t+nT, t+(n+1)T]$

$$T_{clk1 \uparrow} = t + \Delta + T_c/2 + kT_c = t + \Delta + T/2 + nT \in (t+nT + T/2, t+T+nT]$$

因此有 $T_{clk1 \uparrow} \in [t+nT, t+(n+1)T]$

从上述计算可以看到,此时两个时钟的信号采样时刻都位于式(1)描述的时间域中, $clk0, clk1$ 都可以作为采样时钟。

(2) 当 $\Delta \in (T_c/2, T_c)$ 时, $g=n, k=n-1$

$$T_{clk0 \uparrow} = t + \Delta + gT_c = t + \Delta + nT \in (t+T/2+nT, t+T+nT]$$

因此有 $T_{clk0 \uparrow} \in [t+nT, t+(n+1)T]$

$$T_{clk1 \uparrow} = t + \Delta + T_c/2 + kT_c = t + \Delta + T/2 + (n-1)T \in (t+nT, t+T/2+nT]$$

因此有 $T_{clk1 \uparrow} \in [t+nT, t+(n+1)T]$

从上述计算可以看到,此时两个时钟的信号采样时刻都位于式(1)描述的时间域中。 $clk0, clk1$ 都可以作为采样时钟。

从 Δ 的 3 种取值情况可以看出,不论 Δ 为值域中的任何值,至少有一个时钟的信号采样时刻位于式(1)描述的时间域中。因此,使用两个相位差相同的同频率时钟可以完全正确地检测出降频码的信息内容。证明完毕

信令检测过程中无论是识别信令帧头、解析信令类型以

及获取端口信息,都需要通过判断连续的几个时钟周期中采样的信息是否与定义好的编码序列相符才能完成,因此本方法中对于信令帧的每一步检测工作都使用有限状态自动机(FSM, Finite State Machine)技术来实现。用上述两个时钟分别控制两个相同机制的有限状态机来同时检测信令帧,一旦其中一个状态机检测到信令帧,就会关闭另一个有限状态机的运行,保证了交换机能正确检测出信令信息。若用数组 HEAD 存放降频信令帧的帧头信息序列, SHQ, NDQ, BC 存放降频编码信令申请信道、撤销信道以及广播 3 种命令的信息序列, COMMAND 存放实际接受到命令的信息序列, DID 存放实际接受到端口号的信息序列, i 暂存每个时钟上升沿采样的信息,则本方法中有限状态机的运行可以用下面的算法描述:

算法 1

- (1) 有限状态机等待在空闲状态。
- (2) 如果 i 为 '0', 则认为信令信号到来, 进入信令帧头解析状态。
- (3) 依次判断每个采样信息 i 与 HEAD 信息序列是否相符。如果两者不相符, 有限状态机回到空闲状态; 如果相符, 继续进行判断, 直到检测到整个信令帧的帧头, 进入信令类型解析状态, 同时关闭另一个有限状态机。
- (4) 依次把每个采样信息 i 存放在数组 COMMAND 中, 直至整个命令信息检测完毕为止。把 COMMAND 数组与数组 SHQ, NDQ, BC 进行比较, 判断用户发出了什么命令, 根据判断结果置相应的命令标识为有效状态, 或者发出非法命令警告。如果是申请信道命令, 有限状态机进入端口信息解析状态, 否则有限状态机回到空闲状态。
- (5) 依次把每个采样信息 i 存放在数组 DID 中, 直至整个端口信息检测完毕为止, 有限状态机回到空闲状态。

3.3 信令降频码特点

信令降频码有以下几个特点:

- (1) 这种信令降频码编码方法使信令信号频率成倍下降, 因此对于信令的识别无需串并转换器件进行信令信号的串并转换。
- (2) 降频编码后的信令帧与没有经过降频编码的信令帧长度相差不大。
- (3) 检测信令信息的时钟频率是原串行信号频率的五分之一。采用有限状态自动机 FSM 依次检测信令帧头、信令类型和目的端口号。采用多个不同相位的时钟控制多个相同机制的有限状态自动机 FSM 来解决交换机时钟与串行信号之间的相位随机现象带来的信令检测遗漏问题。

4 实验仿真

为了验证信令降频方法的正确性, 笔者给出了仿真实验结果。假设信令帧的帧头编码定义为“00000111110000011110000011111”, 有 3 种命令分别为申请信道 SHQ、撤销信道 NDQ 以及广播 BC, 它们经过降频码编码之后分别为“00000111110000011111”, “000001111111111100000”和“1111100000000011111”。由于只是为了验证本方法的正确性, 因此仿真实验中采用的信号频率较低。信号 XL_OUT 表示 50M 的串行信令数据, $clk0, clk1$ 是两个反相的 100M 时钟, 由外部输入的 125M 时钟 $clk125M$ 经过锁相环以及分频技术产生。 $clk0, clk1$ 用来控制两个相同机制的有限状态自动机 XLOStatus 和 XL1Status, SHQ, NDQ 和 BC 是测试中用到的

3种信令到来标志。仿真结果如图3所示。

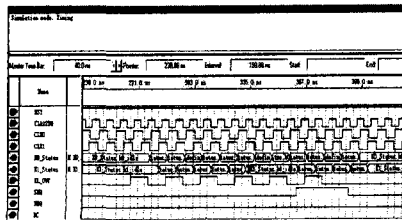


图3 使用信令降频码进行信令降频方法的仿真图

从仿真波形中可以看到,时钟 $clk0$ 上升沿首先检测到信令帧信息的到来并启动 0# 状态机,5ns 后时钟 $clk1$ 上升沿也检测到信令帧信息并启动 #1 状态机运行。6 个时钟上升沿过后,0# 状态机首先检测出完整的信令帧头,所以由 0# 状态机关闭 1# 状态机。信令帧头检测到之后的 4 个时钟周期中 0# 状态机根据检测到的信息识别出信令是 SHQ。仿真结果表明信令降频方法是正确、可行的。

5 分析与讨论

下面笔者以 16 端口交换机为例,分析使用信令降频方法给交换机带来的优势。

图 4 描述了在没有使用信令降频方法的 16 端口交换机的体系结构。由于没有使用信令降频编码方法,交换机中使用了 16 个电收发器,对串行信号进行并行转换,降低信号频率之后送给控制器。由于控制器接收分析的都是并行信号,传输这些并行信号占用了控制器和交换器件的大量引脚,使得一个交换器件根本无法满足 16 端口中 16 路并行信号的转发,因此 16 端口交换机中使用了多个交换器件级联来解决这个问题。另外,并行信号的传输给 PCB 板布线造成了巨大困难,并给交换机的稳定性带来了隐患。

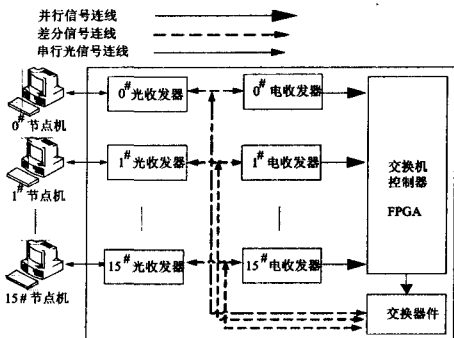


图4 使用降频编码方法前的 16 端口光纤通信网交换机

图 5 描述的是使用了信令降频编码方法之后的 16 端口交换机的结构。从图中可以看出,交换机省去了所有的电收发器器件,光收发器接收的串行信号直接送给交换机的控制器进行处理。由于只有信令帧经过了降频编码,因此交换机的控制器能够解析信令帧,以控制交换器件正常工作,对于普通数据帧则不作解析。与图 4 比较,图 5 中节省了 16 个电收发器,并且由于交换机控制器送出的是串行信号,所需引脚大大减少,从而使用的交换器件数量也得到减少。另外,使用降频编码方法之后,PCB 板上至少减少了 500 条高频信号线,其

面积至少缩小一倍。

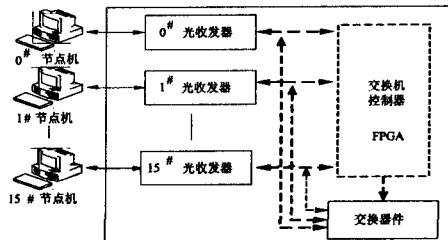


图5 使用降频编码后设计实现的 16 端口光纤通信网交换机

从上述分析中可以看出,使用降频编码方法之后,交换机的结构得到了很大简化,并且由于结构简化带来了硬件成本的大幅降低以及工作稳定性的大大提高。信令降频编码方法给交换机带来的优势在大规模交换机中会得到更显著的体现。

结束语 本文提出的信令降频编码方法在光纤通信中用降频码来描述信令信息,使得信令帧信号达到了频率成倍降低的效果,交换机无需串并转换器件就能够识别信令帧。交换机的硬件成本得到极大降低,PCB 板面积大大缩小,并且使得交换机的稳定性得到了极大提高。下一步的工作就是通过增加滤波等手段使信令帧的识别更准确,尽量缩短信令帧头长度,且不对交换机的工作稳定性产生影响。

参考文献

- [1] Yang Yang, Blum R S. Energy-efficient Routing for Signal Detection Under the Neyman-Pearson Criterion in Wireless Sensor Networks[C]//Proceedings of the 6th International Conference on Information Processing in Sensor Networks, 2007; 303-312
- [2] Lu Yufeng, Oruklu E, Saniie J. Fpga-based hardware/software co-design for chirplet signal decomposition[C]//Proceedings of the 18th ACM Great Lakes Symposium on VLSI, 2008; 347-350
- [3] Cortes F P, Bampi P S. A fully integrated CMOS RF front-end for a multi-band analog mixed-signal interface[C]//Proceedings of the 20th Annual Conference on Integrated Circuits and Systems Design, 2007; 271-275
- [4] 曹寒梅,杨银堂,蔡伟,等.应用于高速信号传输系统的多路选择器[J].半导体学报,2008(6):1040-1043
- [5] 雷艳静,曾晓荟,等.信令寻径式高速光纤传输交换机研究与设计[J].计算机工程与应用,2005,26:23-26
- [6] 刘树彬,郑伟,刘序宗,等. BES-III TOF 子触发系统中高速数据传输的实现[J].核技术,2008(3):223-228
- [7] 张曦林,刘海涛,李道本. 高速率的符号干扰传输及其编码的检测译码[J].北京邮电大学学报,2007(12):60-64
- [8] 冯德军,王雪松,肖顺平,等. 高速目标解调频处理的向各位特性与处理[J].电子与信息学报,2008(4):916-920
- [9] 王玮,冯永茂,丁铁夫. 基于 FPGA 的高速数据传输芯片控制器设计[J].电子器件,2008(6):845-848
- [10] 付志红,马静,谢品芳,等. 基于高速数字信号处理平台的实时仿真技术[J].系统仿真学报,2007(8):3680-3683
- [11] 刘能,徐晓,邓华秋,等. 一种新的周期信号的高速数据采集方法研究[J].仪器仪表学报,2007(4):765-768