

光纤通道交换式网络的虚拟通道分配策略

王昊天, 丁 凡, 熊华钢

(北京航空航天大学电子信息工程学院, 北京 100191)

摘 要: 为了满足光纤通道(Fibre Channel, FC)交换式网络不断增长的带宽需求, 光纤通道协议提出了虚拟通道(Virtual Channel, VC)架构, 通过逻辑上的通信通道对整网流量进行合理分配. 本文将 VC 机制由“点到点”扩展为“端到端”, 并结合网络演算(Network Calculus)理论和信誉量漏桶机制, 在强实时条件约束下, 提出虚拟通道的分配策略, 以减小整网业务流量的端到端延迟时间为优化目标, 解决了 VC 架构的三项技术难题: 每种 VC 数目的确定、每个 VC 的信誉量(Credit)如何保证、以及每条 VC 的路径选择. 计算机仿真结果表明, 随着网络负载强度的不断加大, 扩展的 VC 分配策略具有适应能力, 增强了在 FC 交换式网络中应用 VC 架构的有效性.

关键词: 光纤通道; 交换式网络; 虚拟通道; 强实时性保证; 分配策略; 网络演算

中图分类号: TP393 **文献标识码:** A **文章编号:** 0372-2112 (2011) 10-2251-07

Allocation Scheme of Virtual Channel in Fibre Channel Switched Networks

WANG Hao-tian, DING Fan, XIONG Hua-gang

(School of Electronics and Information Engineering, Beijing University of Aeronautics and Astronautics, Beijing 100191, China)

Abstract: To satisfy the increasing bandwidth requirement of Fibre Channel (FC) switched networks, Virtual Channel (VC) architecture was proposed in new version FC protocols, in order to allocate traffic flows by logical communications channels. Under hard real-time constraint, VC has been expanded from “point-to-point” to “end-to-end”, and this paper proposes the allocation scheme of VC based on Network Calculus and token bucket to reduce the end-to-end delay of traffic flows in whole FC networks. Thus it has been solved the three hard works in FC protocols: how to determine the number of each kind of VC; how to guarantee each VC's credit; and how to choose the path of each VC. Computer simulations results show that VC allocation scheme has good flexibility with traffic increasing intensity, thus enhance the availability of VC architecture into FC switched networks.

Key words: fibre channel; packet switched networks; virtual channel; hard real-time guarantee; allocation scheme; network calculus

1 引言

光纤通道(Fibre Channel, FC)协议, 是由美国国家标准协会(American National Standards Institute, ANSI)提出的一种串行通信标准集, 因其具有很好的兼容性、高速率和适于远距离传输等特点被广泛应用到存储局域网^[1]和航空电子互连^[2]中. 光纤通道协议集中的交换式结构协议(Fibre Channel Switched Fabric, FC-SW)为交换式网络提供了一种互连方法, 协议不断更新, 至2009年末, 已经发展为FC-SW-5版本^[3], 与之相关的信号帧格式标准也更新至FC-FS-3版本^[4], FC网络体系结构有了较大的变化, 主要体现在: 针对FC网络提出了链路带宽由1Gbps提升为10Gbps的要求; 由单交换机互连结构演变成多个FC交换机互连的交换式网络结构; 引入虚拟

通道(Virtual Channel, VC)机制^[5,6]以管理交换式网络的业务流量. 其中, VC的使用作为FC网络体系结构的重要变化, 在FC协议中指出三点技术难题: VC数目(N_{VC})的确定; 每个VC的信誉量(Credit)保证; 以及VC的路径选择.

本文在早先的基于强实时约束的加权轮转(Weighted Round Robin, WRR)调度研究^[7]的基础上, 深入分析新版FC交换式网络体系结构和业务流量特性, 将VC由“点到点”扩展为“端到端”, 并且结合网络演算(Network Calculus, NC)和漏桶(token leaky bucket)技术, 提出一种以业务流量的整网端到端的延迟最小为目的的VC分配策略. 通过“活塞模型”对VC集进行二次优化, 确定出VC的数目; 在强实时保证下, 通过业务流量的信誉量漏桶模型得出每个VC的信誉量, 并进行处理加

工,得出更适合 VC 使用的相对信誉量值;通过一次遍历即可去除无用路径的 VC,获取每个 VC 的路径,最终优选出一组对于当前到达流量(每个消息最大允许延迟内)分配的 VC 集,解决了 FC 协议中 VC 机制的三项技术难题.

2 系统架构与问题描述

2.1 虚拟通道架构

FC 交换式网络由 FC 交换机(每个 FC 交换机的规格和性能相同)、FC 终端、内部交换链路(Internal-Switch Link, ISL)和普通链路组成,如图 1 所示,ISL 通过 E 端口将 FC 交换机互连,普通链路连接 FC 交换机的 F 端口和 FC 终端的 N 端口,所有的交换机与 ISL 组成多级交换结构(Multi-Switch Fabric),即图 1 中的“云型”部分.

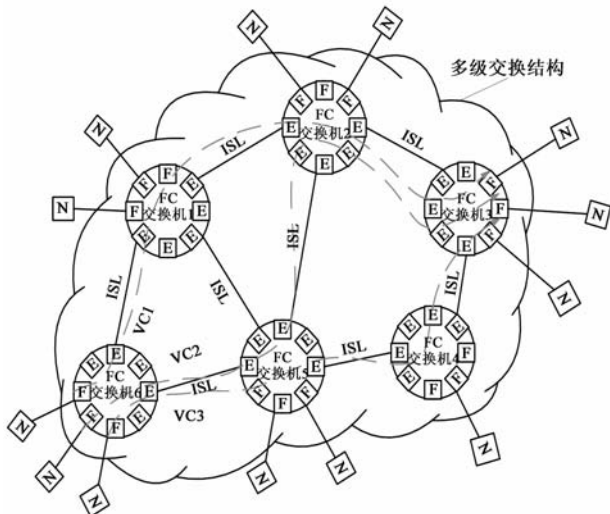


图1 扩展的VC机制在新版FC交换式网络中的系统结构图

VC存在于多级交换结构内部,FC协议中定义为“点到点”逻辑路径,但是本文经过分析,将“点到点”的概念扩展为“端到端”的概念,即VC表示在源交换机端口到目的交换机端口之间进行业务流量传输的逻辑路径.这样做一方面减少对不同VC的维护开销,另一方面可以将基于局部信息的分配策略转变为基于全局信息的分配策略.

例如在图1中,当前到达业务流量需要从FC交换机6发送至FC交换机3,可以将当前到达业务流量分为三条VC来发送数据(图1中的虚箭头所示),其中VC₁是FC交换机6-FC交换机1-FC交换机2-FC交换机3中ISL和FC交换机所组成的业务流量逻辑路径;VC₂是FC交换机6-FC交换机5-FC交换机2-FC交换机3中ISL和FC交换机所组成的业务流量逻辑路径;VC₃是FC交换机6-FC交换机5-FC交换机4-FC交换机3中ISL和FC交换机所组成的业务流量逻辑路径.并且VC₁、VC₂和VC₃是同时进行,互不干涉.在FC交换式网

络中,通过使用VC来分配业务流量,可以充分利用网络的带宽资源,将拥塞降低到最小,从而提高通信效率,减小端到端延迟时间.

2.2 问题的提出

在FC交换式网络中,如何能够合理地分配VC,使得整网通信数据的端到端延迟最小,成为在FC交换式网络中引入VC机制的难点.根据FC协议,针对当前到达的业务流量,VC分配的难点具体体现为三个方面:

- (1) VC数目(N_{VC})的确定;
- (2) 每个VC的信誉量(VC_Cr)保证;
- (3) 所选择的每条VC的路径信息.

本文通过漏桶模型,建立信誉量保证机制,保证在单个交换节点处传输的实时性;然后根据网络演算推导出分配策略,将到达流量分配到不同的VC中,同时得出VC的三个重要参数.

3 虚拟通道分配策略

3.1 网络流量模型

FC交换式网络的流量以VC进行划分,将相同源节点和目的节点的流量,做为输入流量,归属为一个VC集,即VCM,VCM中包括为了传输输入流量所需的所有VC:

$$VCM = \{ VC_1, VC_2, \dots, VC_{N_{VC}} \}$$

由于FC协议规定的通信基本单元是帧,参考实时通信中的周期性任务模型^[8],如图2(a)所示,本文定义划分到VC_{*i*}($1 \leq i \leq N_{VC}$)中的流量参数如下:

(1) VC_{*i*}所承载数据的间隔时间 P_i :表示对于周期性消息,表示划分在VC_{*i*}中数据帧的产生周期,而对于非周期性消息,则表示划分在VC_{*i*}中数据帧产生的最小时间间隔;

(2) VC_{*i*}所承载数据的最大长度 Fr_i^{\max} :由于FC交换式网络中的数据帧长度是可变的,用 Fr_i^{\max} 表示VC_{*i*}中数据帧长度的最大值,包括FC协议规定的分隔符、帧头、信息域和校验域等帧的全部内容;

(3) VC_{*i*}所承载数据的抖动 Jr_i 和 Jr_i^{\max} : Jr_i 表示由于竞争等因素使得VC_{*i*}中数据帧产生的时间偏移值, Jr_i^{\max} 为VC_{*i*}中数据帧的最大允许抖动值.

参考航空电子全双工交换式以太网(Avionics Full Duplex Switched Ethernet, AFDX)协议^[9],建立基于帧的信誉量漏桶模型,如图2(b)所示.

图2所示的信誉量漏桶模型规定:在FC交换机处,每到达一个数据帧,都要检查 VC_Cr_i 的值.如果 $VC_Cr_i \times 2112$ 的值超过 Fr_i^{\max} 的值,则该数据帧被接受,并且从 $VC_Cr_i \times 2112$ 的值中减掉一个 Fr_i^{\max} 的值;如果 $VC_Cr_i \times 2112$ 的值低于 Fr_i^{\max} 的值,则该数据帧直接被

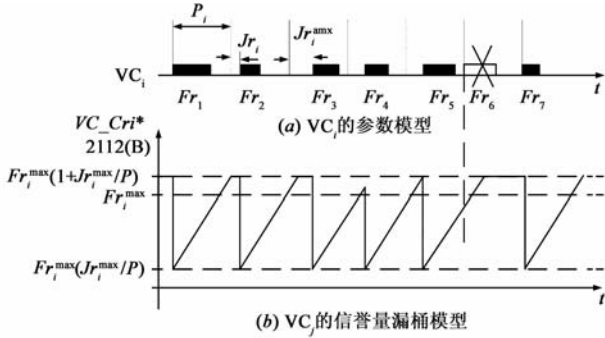


图2 FC交换式网络的流量模型

丢弃,且 $VC_Cr_i \times 2112$ 的值没有影响,如图 2(b) 中的 Fr_6 帧,可见同一 VC 上的两个连续数据帧的最小时间间隔应该为 $(P_i - Jr_i^{\max})$ 。

注:此处将 $VC_Cr_i \times 2112$,是根据 FC-FS-3 版本^[4] 中定义 FC 的数据帧的最大值为 2112(B),而 VC_Cr_i 的值为整数,在 FC 交换机中为 VC 所预留的缓存空间通常取 $VC_Cr_i \times 2112$ (B),而数据帧的单位为波特(B),故此处取 $VC_Cr_i \times 2112$ (B)。

信誉量漏桶模型的公式表示为:

$$VC_Cr_i \times 2112 = Fr_i^{\max} + \frac{Fr_i^{\max}}{P_i} Jr_i^{\max}, 1 \leq i \leq N_{VC} \quad (1)$$

每个 VC 在初始时,信誉量 VC_Cr_i 的值按式(1)设置,随着时间的推进, $VC_Cr_i \times 2112$ 的值按着 (Fr_i^{\max}/P_i) 的比率增长,但最高值不会超过 $Fr_i^{\max}(1 + Jr_i^{\max}/P_i)$,最小值也不会低于 $Fr_i^{\max}(Jr_i^{\max}/P_i)$ 。

3.2 基于网络演算的封装过程

在业务流量模型中,封装过程作为一种简单可行的绑定模型被广泛使用^[10]。根据网络演算理论^[11,12],设一个非负过程 $a(t), t > 0$,表示在 t 时刻的速率,令

$$A_i(t_1, t_2) = \int_{t_1}^{t_2} a_i(t) dt$$

作为在时间区间 $[t_1, t_2)$ 内累计到达的业务流量,则有 $A_i(t_1, t_2) \leq \bar{A}_i(t_2 - t_1), \forall t_1 < t_2$ 成立,其中 $\bar{A}_i(t_2 - t_1)$ 被称为 $A(t)$ 的封装过程,表示在时间区间 $[t_1, t_2)$ 内,对于任意 $t_1 < t_2$ 的所有的业务流量的上边界值。在网络演算中,还要求流量约束函数具有凹性 (concave) 和次可加性 (sub-additive)^[13],在简化计算中最常用的流量约束函数的形式为

$$A(t) = \sigma + \rho t \quad (2)$$

其中 σ 为突发度, ρ 为可持续流量,可简记为 $a(t) \sim (\sigma, \rho)$ 。

结合式(1),FC 交换式网络中的业务流量都可以用式(2)规整,在任何时间间隔 t 内,业务流量都被封装在 $A(t)$ 内,则有

$$\begin{cases} \sigma = Fr_i^{\max} \\ \rho = \frac{Fr_i^{\max}}{P_i} \end{cases} \quad (3)$$

如图 3 所示,本文将到达的网络流量分配到不同的 VC 中进行传输,将初始到达业务流量规整为 $\{\sigma, \rho\}$ 模型,每条 VC_i 分配的流量为 $\{\sigma_i, \rho_i\}, i = 1, 2, \dots, N_{vc}$ 。

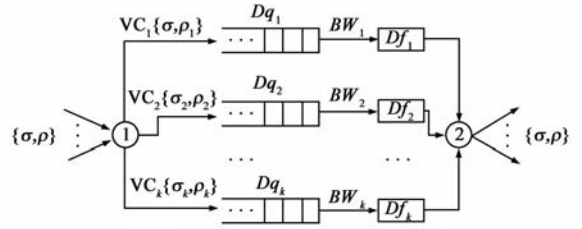


图3 业务流量的VC分配模型

考虑到在每条 VC 所经过的路径中,对延迟时间有影响的是 VC 经过路径上的瓶颈节点^[14],此处用瓶颈节点的队列延迟 Dq 和瓶颈链路带宽 BW 表示;而在高速网络中,固定延迟占据整端端到端延迟的主要部分,用瓶颈节点的队列延迟近似代替其它节点的队列延迟,而其它延迟(如链路延迟和传输延迟)用固定延迟近似,这种方法被证实是十分准确的^[15],所以此处每条路径的端到端延迟时间用瓶颈节点的队列延迟和每条路径的固定时延 Df 的值来代替。

VC 机制中所确定的端到端延迟,如图 3 中所示,是指输入的业务流量 $\{\sigma, \rho\}$ 被分配到 VC 中由分离点 1 到 VC 聚合点 2 的延迟时间,不包括源终端的排队时间和目的终端的排队时间。每条 VC 的延迟时间进一步划分为瓶颈节点的排队延迟 Dq 和每条路径上的固定延迟 Df 。由此可得某一 VC_i 的端到端延迟 D_{e2ei} 可以表示为

$$D_{e2ei} = Dq_i + Df_i, 1 \leq i \leq k \quad (4)$$

而对于一个连续工作的系统,排队延迟可以被限制在由 $\bar{A}(t)$ 的上限 Dq_i 内^[16],因而排队延迟 Dq 可以表示为:

$$Dq_i = \inf\{t \geq 0: \bar{A}_i(t) - BW_i \times t \leq 0\} \quad (5)$$

其中 BW_i 表示 VC_i 的端到端路径中的瓶颈链路带宽,即 VC_i 覆盖所有链路的带宽最小值; t 表示 Dq 的变化值。

强实时约束条件是指到达的业务流量必须满足在其最大允许延迟时间内被传输,根据在单交换机强实时调度的研究经验^[7],如果满足强实时条件,则有对于任意时间间隔 t ,排队的的数据被发送的最小时间量应大于等于该业务流量中数据的最大长度值。我们取 t 的最大值作为该 VC 的排队延迟限,结合式(5)得

$$Dq_i = \inf\{t \geq 0: \rho_i t + \sigma_i - BW_i \times t = 0\} = \frac{\sigma_i}{BW_i - \rho_i} \quad (6)$$

由于可持续流量 ρ 的最大值即为所对应的链路带宽,因此本文分配策略规定

$$\begin{cases} 0 \leq \rho_i < BW_i, \sigma_i > 0 \\ 0 \leq \rho_i = BW_i, \sigma_i = 0 \end{cases} \quad (7)$$

则有

$$\nabla Dq_i = \left[\frac{\partial Dq_i}{\partial \sigma_i}, \frac{\partial Dq_i}{\partial \rho_i} \right] = \left[\frac{1}{BW_i - \rho_i}, \frac{\sigma_i}{(BW - \rho_i)^2} \right] > 0$$

所以 Dq_i 在 $\{\sigma, \rho\}$ 组成的空间内单调递增, 其最小值是可以取得的, 且最小值一定发生在边界或最高点处^[17]. 而 Df_i 对于某一 VC_i 是固定值, 可以根据具体路径得出, 所以 VC_i 的端到端延迟 $D_{e2e,i}$ 的最小值是可以取得的, 根据式(4)和式(6)可得出最小值 $*D_{e2e,i}$ 的表示:

$$*D_{e2e,i} = \frac{\sigma_i}{BW_i - \rho_i} + Df_i, 1 \leq i \leq N_{VC} \quad (8)$$

而针对某次流量分配的所有 VC 组成一个集合 $VCM\{VC_1, VC_2, \dots, VC_N\}$. 根据 NC 理论可得:

初始到达业务流量的包封:

$$A(t) = \rho t + \sigma \quad (9)$$

每个 VC 的包封:

$$A_i(t) = \rho_i t + \sigma_i, 1 \leq i \leq N_{VC} \quad (10)$$

其中满足 $A_1(t) + A_2(t) + \dots + A_{N_{VC}}(t) = A(t)$, $\sigma_1 + \sigma_2 + \dots + \sigma_{N_{VC}} = \sigma$, $\rho_1 + \rho_2 + \dots + \rho_{N_{VC}} = \rho$, 因此, 使用 VC 来划分业务流量并没有给网络带来额外的业务流量开销.

而针对初始到达流量的 VCM 的端到端延迟 DM_{e2e} 应为所有 VC 端到端延迟的最大值, 即

$$DM_{e2e} = \max\{D_{e2e,1}, D_{e2e,2}, \dots, D_{e2e,N_{VC}}\} \quad (11)$$

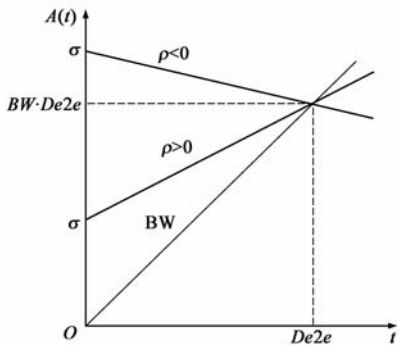


图4 端到端延迟(D_{e2e})与流量参数(σ, ρ, BW)的关系图

3.3 VCM 的优化

FC 协议中规定当某一 VC 的信誉量值 VC_Cr_i 为 0 时, 该 VC 不能传输数据. 因为 VC 的固定延迟 Df 是不变的, 是不可优化的, 所以信誉量模型主要是为了优化 VCM 中的队列延迟 Dq , 去除冗余的 VC, 得到一个最优的 VCM. 这里提出的最优定义为, 如果当前的 VCM 为最优, 则再加入任何一个 VC 都会增加 VCM 的整体端到端延迟限 $*DM_{e2e}$ 的值.

如图 5 所示, 取当前 VCM 中的瓶颈链路 VC_i, 将 VC 假想为一个“活塞模型”, 原有的固有延迟 Df_i 为图 5 中的阴影部分, 是不能注入流量的; 而其余部分是充分分配给 VC_i 的突发度 σ_i , 又因为 $*DM_{e2e}$ 的值要小于等于

VC_i 的总长度值, 从而限制了每个 VC 分配的突发度 σ_i 的大小.

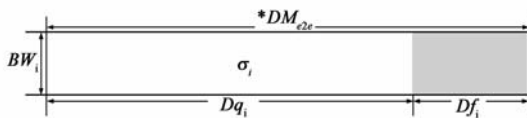


图5 VC_i的优化模型

当 VC_Cr_i 为 0 时, 即表示 VC 的端到端延迟为无穷大 ∞ , 数据帧不可能满足实时性要求, 不能分配到该 VC 上被传输, 符合协议要求.

基于 3.2 节的分析, 以 VC 分配业务流量并没有增加额外的流量值, 所以所有 VC 分得的业务流量应该等于初始到达的业务流量, 即

$$\sigma_1 + \sigma_2 + \dots + \sigma_{N_{VC}} = \sigma$$

即为

$$BW_1 \times (D_{e2e,1} - Df_1) + BW_2 \times (D_{e2e,2} - Df_2) + \dots + BW_{N_{VC}} \times (D_{e2e,N_{VC}} - Df_{N_{VC}}) = \sigma \quad (12)$$

取当前可用带宽的和为 BWM , 则有

$$BWM = \sum_{i=1}^{N_{VC}} BW_i \quad (13)$$

由式(12)和式(13)可得 VCM 的整体端到端延迟限 $*DM_{e2e}$ 的值和最优情况下分配的突发度 $*\sigma_i$ 的值:

$$*DM_{e2e} = \frac{1}{BWM} (\sigma + BW_1 \times Df_1 + BW_2 \times Df_2 + \dots + BW_{N_{VC}} \times Df_{N_{VC}}) \quad (14)$$

$$*\sigma_i = BW_i \times (*DM_{e2e} - Df_i) = \frac{BW_i}{BWM} (\sigma + \sum_{j=1}^{N_{VC}} BW_j \times Df_j - Df_i) \quad (15)$$

取 $\rho = \min\{\rho_1, \rho_2, \dots, \rho_{N_{VC}}\}$, 则可以得出每条 VC 的稳定可持续流量 $*\rho_i$ 的值:

$$*\rho_i = \frac{BW_i}{BWM} \times \rho \quad (16)$$

此时得到的 VCM 即为最优的 VCM, 最优 VCM 中的 VC 数即为输入流量分配的 VC 数 N_{VC} ; 每个 VC 满足应用式(1)信誉量 VC_Cr ; 剩下的 VC 所经过的路径即为 VCM 中所选择的 VC 的路径信息.

4 性能分析与验证

为了验证虚拟通道分配算法的正确性, 我们采用计算机仿真的形式, 使用 OPNET 14.5(教育版)仿真工具, 仿真环境如图 6 所示. 假设输入流量是图 1 中由 FC 交换机 6 发送至 FC 交换机 3, 为了表述清晰, 我们将图 1 中的相关拓扑提取, 得到测试拓扑结构, 得到如图 6 所示的测试拓扑结构. 为了与真实情况一致, 这里设定每个 FC 交换机的规格相同, 具有相同的转发性能. 设 FC 交换机的端口数为 N , 输入端口 i 至输出端口 j 的转发速率为 $\lambda_{i,j}$.

本文充分考虑到飞机在高空环境下飞行的严峻环境, 将输入业务流量设定为均匀流量(贝努力业务流)、非均匀流量(对角线业务流)和突发流量(对角线 ON-OFF 业务流)的混合流量, 三种不同类型的流量同时发送, 得到最苛刻情况下的仿真结果, 以确保仿真的可靠性。

业务流量的负荷 σ 采用正规化符合, $\sigma \in [0, 1]$. 为了建立较为准确的仿真环境, FC 交换机对于不同流量的处理有所区别, 每种流量的 FC 交换机处理参数见表 1 中的定义。

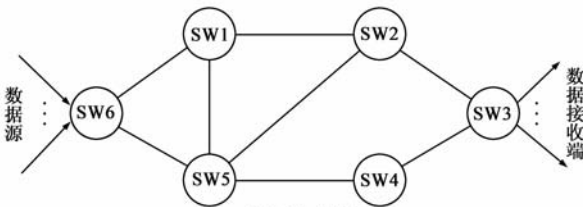


图6 测试拓扑结构

表 1 输入流量的参数

流量类型	流量参数
均匀流量	数据到达过程符合贝努力 i. i. d, $\lambda_{i,j} = \rho/N$
非均匀流量	数据到达过程符合贝努力 i. i. d, $\lambda_{i,i} = 2\rho/3, \lambda_{i,j} = \rho/(3 \times (N-1))$
突发流量	数据到达过程不再是贝努力 i. i. d, 而是符合两状态马尔科夫过程的 ON-OFF 模型, 在 ON 状态时, 均值为 32 的几何分布; 在 OFF 状态时, 到达的数据包为 0.

通过多次仿真, 应用 3.2 节中的式(8)计算出所有的 VC 的端到端延迟限, 然后应用 3.3 节中的式(14)进行优化, 得到了仿真环境下使用 VC 分配策略得到的 VCM 结果, 如表 2 所示, 包括了使用 VC 的数目, 每条 VC 的使用次数, 每条 VC 的最大抖动值 Jr_i^{max} , 每条 VC 的信誉量值, 和每条 VC 的路径信息. 分为 VC 瓶颈链路带宽与 VC 固有延迟成正比(正序)和 VC 瓶颈链路带宽与 VC 固有延迟成反比(逆序)两种模式下的结果. 由于 FC 网络流量是带有优先级的, 这里不妨设正序情况下, 为 VC_1 至 VC_7 的优先级逐次增高, 即 $VC_1 < VC_2 < \dots < VC_7$; 逆序情况下, $VC_1 > VC_2 > \dots > VC_7$.

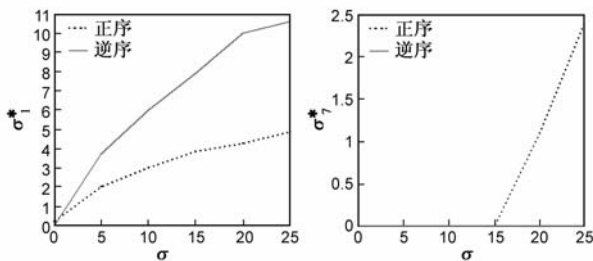


图7 输入流量总突发度与每条VC分配突发度的关系图

表 2 VCM 的分配结果

VCM	使用次数 (正序/逆序)	Jr_i^{max} (μs)	信誉 量值	路径(正序/逆序)
VC_1	2/248	50	5	正: SW6-SW1-SW2-SW3 逆: SW6-SW5-SW2-SW3
VC_2	14/216	48	7	正: SW6-SW1-SW2-SW5-W4-SW3 逆: SW6-SW1-SW5-SW2-SW3
VC_3	62/163	96	10	正: SW6-SW1-SW5-SW2-SW3 逆: SW6-SW5-SW1-SW2-SW3
VC_4	196/79	12	14	正: SW6-SW1-SW5-SW4-SW3 逆: SW6-SW5-SW4-SW3
VC_5	264/49	75	17	正: SW6-SW5-SW1-SW2-SW3 逆: SW6-SW1-SW2-SW3
VC_6	317/37	46	23	正: SW6-SW5-SW2-SW3 逆: SW6-SW1-SW2-SW5-W4-SW3
VC_7	415/5	87	24	正: SW6-SW5-SW4-SW3 逆: SW6-SW1-SW5-SW4-SW3

在仿真过程中, 由图 7 可以看出 VC_1 和 VC_7 被分配的突发度的值随输入流量的总突发度的值的增加而增加, 其中, VC_1 所分配的突发度无论是正序还是逆序, 都随着输入流量总突发度的增加而分段地线性增长(图 7(a)); VC_7 在正序情况下, 只有当输入流量的总突发度增长到某一个数值时才线性增长, 而在逆序情况下, 没有分配到突发度, 所分突发度的值近似为零(图 7(b)).

由图 8 可以看出, VC_1 的稳定可持续流量值随着总的输入可持续流量的增加而增加, 在正序情况和逆序情况下, 都会有波动, 这是因为每当有优先级高于 VC_1 的 VC 加入时, 致使 VC_1 的流量减小, 但随着总可持续流量的增加, VC_1 会稳定增加, 从而产生了波动(图 8(a)); 而对于 VC_7 , 随着总可持续流量的增加, VC_7 的稳定可持续流量一直在增加, 这是因为 VC_7 的优先级最高, 没有其它流与 VC_7 抢夺流量(图 8(b)).

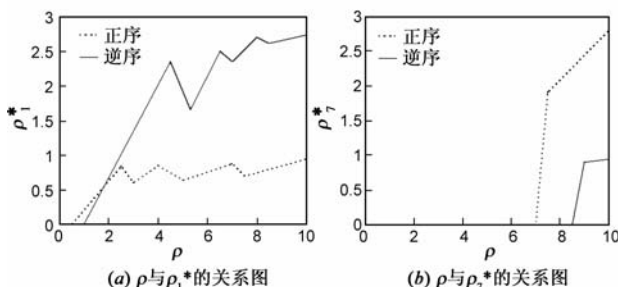


图8 输入流量总的可持续流量与每条VC的可持续流量的关系图

图 9 给出了每种 VC 所使用的数目, 可以看出正序情况下(图 9(a)), VC_1 的数目最少, VC_7 的数目最多; 逆序情况下(图 9(b)), VC_1 的数目最多, VC_7 的数目最少。

为了进一步验证虚拟路径配置策略对负载强度的适应性, 通过 OPNET 的负载流量注入机制, 用一个注入终端产生负载流量, 然后用过一个交换机与 FC 交换式网络中的所有交换机以“trunk”形式进行相连, 不断的

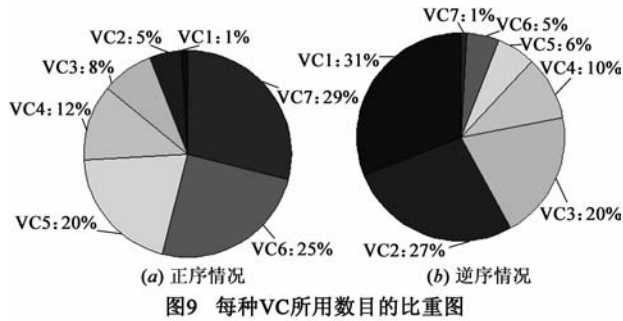


图9 每种VC所用数目的比重图

向 FC 交换式网络中发送广播包,使得负载强度不断增大,观察经典路由算法与虚拟路径配置策略下,到达流量的端到端延迟的变化情况。

将表 1 中的数据分为正序方式和逆序方式,在没有 VC 机制、原有 VC 机制和扩展的 VC 机制下进行传输,对比结果如图 10 所示。可以看出,在正序方式还是逆序方式下,轻度负载时(负载强度小于 40%),三者的端到端延迟十分接近;而随着负载强度的不断增加,特别是在重负载下(负载强度大于 70%),原有 VC 机制的性能明显恶化,在正序情况下甚至超过没有 VC 机制的端到端延迟时间,这是因为“点到点”VC 分配策略缺少整网信息,仅根据链路的部分信息分配流量,因而需要更多的开销,同时也使得交换机的固有转发延迟的累计结果较大;而扩展的 VC 机制(VC_{ex})分配策略,在负载重强度下,仍能保持端到端延迟性能,并没有使端到端延迟恶化。

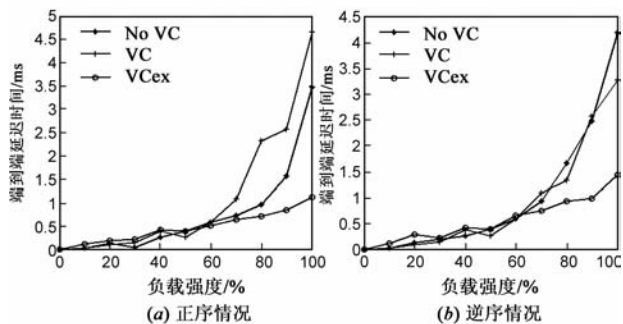


图10 不同算法的端到端延迟比较

5 结论

本文对 FC 协议中提出的虚拟通道架构进行研究,将“点到点”的 VC 架构扩展为“端到端”的 VC 架构,首次解决了 FC 交换式网络虚拟通道(VC)的分配策略的 VC 数目的确定、每个 VC 的信誉量(Credit)保证、以及 VC 的路径选择三项技术难点。

本文利用网络演算(NC)理论得出在强实时约束下,每个 VC 的端到端上限值,并以此值作为优化的参数,提高了算法针对不同流量特征的可用性。VC 分配策略并没有使用迭代的启发式算法来求解,而是使用“活塞模型”用过一次遍历而求得,大大降低了计算的复杂

度,适于在工程中实际应用。

通过将扩展的 VC 机制与没有 VC 机制、原有 VC 机制进行对比测试验证,当网络负载流量不断增大时,扩展的 VC 分配策略要比原有 VC 机制受到的影响小的多,特别在重负载(负载率超过 70%)下,扩展的 VC 分配策略的端到端延迟要远远小于没有 VC 机制的延迟。

今后准备展开将 WDM 波分复用技术为底层传输机制,上层使用 FC-SW 协议,通过 VC 分配策略实现光电双层航空网络的搭建。

参考文献

- [1] Shu J, Li B, Zheng W. Design and implementation of SAN system based on the fiber channel protocol[J]. IEEE Transactions on Computers, 2005, 54(4): 439 - 448.
- [2] 熊华钢, 王中华. 先进航空电子综合技术[M]. 北京: 国防工业出版社, 2009.
Xiong Huagang, Wang Zhonghua. Advanced Avionics Integration Techniques[M]. Beijing: National Defense Industry Press, 2009. (in Chinese)
- [3] International Committee for Information Technology Standards (INCITS). Fibre Channel-Switch Fabric-5 (FC-SW-5) [S]. 2009. 8.
- [4] International Committee for Information Technology Standards (INCITS). Fibre Channel-Framing and Signaling-3 (FC-FS-3) [S]. 2009. 6.
- [5] Bill Martin, Steve Wilson, Vince Guan. Virtual Channel Architecture [R]. The intelligent platform for networking storage, T11/03-369v0, 2003. 6.
- [6] Bill Martin, Steve Wilson, Vince Guan. Virtual Channel for Switched Fabric [S]. Virtual Channel for Switched Fabric 04-093v0, 2004. 1.
- [7] 林强, 熊华钢, 张其善. 强实时约束下光纤通道交换网络消息集优化[J]. 电子学报, 2006, 34(2): 343 - 347.
Lin Qiang, Xiong Huagang, Zhang Qishan. Optimization design of message set for fibre channel switched fabric under hard real time constraints[J]. Acta Electronica Sinica, 2006, 34(2): 343 - 347. (in Chinese)
- [8] Jane W S Liu. 实时系统[M]. 第一版. 北京: 高等教育出版社, 2003. 336 - 430.
Jane W S Liu. Real Time Systems [M]. First Edition. Beijing: Higher Education Press, 2003. 336 - 430. (in Chinese)
- [9] Airlines Electronic Engineering Committee (AEEC). ARINC 664 Aircraft Data Networks, Part 7: Avionics Full Duplex Switched Ethernet (AFDX) Network [S]. 2005. 6.
- [10] 邓钢, 纪阳, 张平. 一种基于分形布朗运动包封过程的 CAC 算法[J]. 电子学报, 2002, 30(12): 2102 - 2104.
Deng Gang, Ji Yang, Zhang Ping. A CAC algorithm based on the fraction brownian motion envelope process [J]. Acta Elec-

tronica Sinica, 2002, 30(12): 2102 – 2104. (in Chinese)

- [11] R Cruz. A calculus for network delay, part I[J]. IEEE Transactions on Information Theory, 1991, 37(1): 114 – 131.
- [12] R Cruz. A calculus for network delay, part II[J]. IEEE Transactions on Information Theory, 1991, 37(1): 132 – 141.
- [13] L B Jean-Yves, T Patrick. Network Calculus: A Theory of Deterministic Queuing Systems for the Internet[M]. Springer, LNCS 2050, 2004.
- [14] D Y Eun, N B Shroff. Simplification of network analysis in large-bandwidth systems[A]. Proceedings of IEEE/ACM INFOCOM[C]. USA: IEEE Press, 2003. 597 – 607.
- [15] C Centinkaya, E W Knightly. Opportunistic traffic scheduling over multiple network paths[A]. Proceedings of IEEE INFOCOM[C]. USA: IEEE Press, 2004. 1928 – 1937.
- [16] C S Chang. Performance Guarantees in Communication Networks[M]. Springer Verlag, 2000.
- [17] 黄平. 最优化理论与方法[M]. 北京: 清华大学出版社, 2009.
Huang Ping. Optimization Theories and Methods[M]. Beijing: Tsinghua University Press, 2009. (in Chinese)
- [18] Dijkstra E. A note on two problems in connection with graphs[J]. Numerische Mathematik, 1959, 1: 269 – 271.
- [19] Bellman R E. Dynamic Programming[M]. Princeton, NJ: Princeton University Press, 1957.

作者简介



王昊天 男, 1983 年 3 月出生, 吉林长春人. 2002 年毕业于吉林大学通信学院, 2006 年进入北京航空航天大学电子信息工程学院. 现为直博研究生, 从事航空电子系统、计算机网络、实时通信等方面的有关研究.

E-mail: qi1983xiang@ee.buaa.edu.cn



丁凡 男, 1980 年 1 月出生, 辽宁大连人. 2002 年和 2009 年在北京航空航天大学分别获工学学士和博士学位. 现为北京航空航天大学讲师, 主要从事航空电子综合化数字通信网络方面的研究工作.



熊华钢 男, 1961 年 12 月出生, 湖北武汉人. 教授、博士生导师、中国电子学会会士、IEEE 高级会员. 现为北京航空航天大学电子信息工程学院副院长, 主要从事航空电子、总线网络、实时性通信等方面的研究工作.