

强实时约束下光纤通道交换网络消息集优化

林 强, 熊华钢, 张其善

(北京航空航天大学电子信息工程学院, 北京 100083)

摘 要: 实时系统应当具有可预期的定时行为, 强实时约束要求系统必须满足定时约束和时限要求. 以周期性任务模型作为网络负载可以准确地描述强实时约束下的系统性能. 本文以光纤通道交换网络强实时约束下的性能研究为背景, 根据调度实时消息传输的加权轮转服务规程, 提出了特定消息集的优化设计算法, 推导并证明了采用该优化算法后系统最坏情况下的可达负载率比不采用该算法时可提高至少一倍.

关键词: 强实时约束; 光纤通道; 消息集; 分组调度; 负载率

中图分类号: TP393 **文献标识码:** A **文章编号:** 0372-2112 (2006) 02-0344-04

Optimization Design of Message Set for Fibre Channel Switch Fabric under Hard Real-time Constraints

LIN Qiang XIONG Hua-Gang ZHANG Qi-Shan

(School of Electrical Engineering, Beijing University of Aeronautics and Astronautics, Beijing 100083 China)

Abstract A real-time system is required to complete its work and deliver its services on a timely basis. We lays much emphasis on hard real-time characteristics of the system, because hard real-time performance is one important performance requirement of special system such as avionics network. The periodic task model is a well known deterministic workload. With its various extensions, the model characterizes accurately many traditional hard real-time applications. Switched fabric is one of basic fibre channel topology. Fixed-length packet scheme is mainly used in fibre channel switches. Taking the study on fibre channel switch fabric under hard real-time constraints as background, an optimization design of typical message set is proposed according to the weighted round-robin algorithm. It is derived and formally proved that the achievable utilization using the optimization design can be doubled at least.

Key words hard real-time constraints; fibre channel; message set; packet scheduling; utilization

1 引言

光纤通道 (Fibre Channel) 协议是由 ANSI (American National Standards Institute) 提出的新型串行通信标准. 光纤通道交换型结构 (Fibre Channel Switched Fabric, 以下简称 FC-SW) 提供了在交换式网络中将多个节点连接在一起的方法, 是光纤通道中一种重要的拓扑结构. 由于光纤通道帧格式数据域较大 (2kbyte), 所以目前光纤通道交换机中大多采用定长分组, 将消息帧在交换节点中重新打包成信元, 通过时隙在入线与出线间进行数据交换, 参考文献 [1, 2] 给出了交换内核采用 ATM 定长分组设计方案.

商用领域中, FC-SW 的应用目前多集中于存储局域网 (Storage Area Network), 在研究方法上侧重于平均消息延时、平均队列长度和吞吐量等网络统计特性方面的研究^[3-7], 但是在航空电子环境等实时通信领域中, 要求网络能够提供强实时约束下的消息传输. 强实时约束下的系

统服务质量问题 (QoS) 在研究方法上与传统的基于统计平均的计算机网络性能评价方法有较大不同, 目前的思想多集中于根据实时通信模型从系统互连的角度对链路带宽分配方法进行优化. 本文根据通用的实时通信模型和目前已经在工程中得到部分形式实现的用于调度实时消息传输的加权轮转服务规程, 对网络性能评价参数进行逆向求解, 提出了一种强实时通信模型中逻辑消息集的优化设计方案并对优化前后的网络可达负载率进行了分析比较. 为解决目前在 FC-SW 实时性研究中出现的“瓶颈”问题提供了一条新的途径.

2 FC-SW 交换节点模型

2.1 消息模型

采用实时通信中的周期性任务模型 (periodic task model)^[8], 交换用交换节点表示. 假设交换节点有 n 条入线和 n 条出线相连接, 每个入线均有一个实时消息流 S_i ,

需要通过交换节点传送, 消息流优先级相同, 组成一个消息集合 M , 即: $M = \{S_1, S_2, \dots, S_n\}$ (1)

对于消息流 S_i 和交换节点, 有如下假设: (1) 消息流产生周期 P_i : 表示消息流 S_i 的消息产生周期. 对于非周期性消息, 则表示消息产生最小时间间隔; (2) 消息流最大长度 C_i : 表示第 i 个消息流的传输时间, 包括网络协议规定的分隔符、帧头、信息域和校验域等帧的全部内容; (3) 基本时间单位: 交换节点中传输一个定长信息包所需要的时间, 包括处理信息包头需要的时间以及对信息包所做的用于传输操作所需要的时间; (4) 输出链路的轮长 RL : 交换节点在分组调度中每次轮转所允许的最大权值, 分组调度轮询周期的上限.

$$\text{消息由一个二维数组表示: } S_i = (C_i, P_i) \quad (2)$$

$$\text{消息流负载率 } U_i: \text{ 定义为: } U_i = C_i / P_i \quad (3)$$

$$\text{网络总的负载率为: } U = \sum_{i=1}^n U_i \quad (4)$$

2.2 队列调度模型

如图 1 所示, 进入交换节点的消息帧首先按时隙进行重新打包, 设每个入线队列 Q_i 对应消息集 M 中一个消息 S_i , 考虑最差情形, 调度消息包特指使用同一个输出链路的所有连接的信息包调度.

3 分组调度算法

3.1 实时限制条件

在一个输出链路上, 用 n 来表示该链路上已建立连接的数目, 调度程序为每个连接分配一个权值 w_{t_i} . 在强实时条件下, 分组调度算法为每个连接所分配的权值应同时保证消息传输的轮转权值约束条件和时限约束条件.

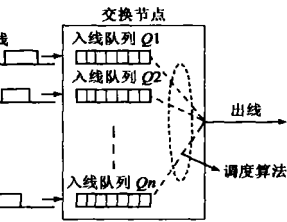


图 1 分组调度模型

(1) 轮转权值约束条件 考虑到交换节点中调度表的建立与查询开销, 在每次轮转中, 交换节点端口用于发送消息的总权值 (轮转长度) 应满足:

$$\sum_{i=1}^n w_{t_i} \leq RL - \delta \quad (5)$$

其中 w_{t_i} 表示连接 i 所分配的权值, δ 表示轮转过程中调度表的建立与维护开销.

(2) 时限约束条件 对于任意时间间隔 t 用 $X_i(t)$ 表示输入队列 Q_i 发送消息的最小时间量. 消息集合 M 中每个消息在其最大允许延迟时间内, 应有足够发送该消息的时间, 因此对于任意消息流 S_i 应有:

$$X_i(P_i) \geq C_i, \quad i = 1, 2, \dots, n \quad (6)$$

上式称为消息的时限约束条件.

对于特定的消息流, 如果所采用的分组调度算法既能满足轮转权值约束条件又能满足时限约束条件, 那么在该分组

调度算法下对特定消息流可实现实时传输.

3.2 分组调度方法

根据加权轮转服务规程, 对于轮转中的任意时间间隔 t , 输入队列 Q_i 发送消息的最小时间量, 可以表示为:

$$X_i(t) = [t / RL] w_{t_i} \quad (7)$$

从式 (7) 可知, $[P_i / RL] w_{t_i}$ 为任意时间间隔 ($t_0, t_0 + P_i$) 内可用于发送消息的时间, C_i 为消息流传输时间, 将二者相比较, 从负载匹配的角度让输出连接 i 的权值最小, 则可得 w_{t_i} 的确定方法:

$$0 \leq w_{t_i} - \frac{C_i}{[P_i / RL]} < 1 \quad (8)$$

3.3 强实时的充要条件

定理 1 对于任意给定的消息集 M , 如果输入连接权值 w_{t_i} 满足 $0 \leq w_{t_i} - \frac{C_i}{[P_i / RL]} < 1$, 则当且仅当 $\sum_{i=1}^n w_{t_i} \leq RL$ 时, M 中所有消息满足最大允许延迟时间要求.

证明 由轮转权值约束条件可得定理的必要性; 由下式可得定理的充分性.

$$X_i(P_i) = [P_i / RL] w_{t_i} \geq [P_i / RL] \frac{C_i}{[P_i / RL]} = C_i$$

4 负载率分析与消息集优化

在分组调度算法作用下, 如果网络负载率小于 U_A 时, 网络中所有消息的实时性都能得到保证, 那么称 U_A 为网络在该分组调度算法下的可达负载率.

引理 1 对于消息集 $\{S_n\}$, 如果在分组调度中以下条件成立: $0 \leq w_{t_i} - \frac{C_i}{[P_i / RL]} < 1$, 并且消息集负载率满足 $U < \frac{[P_{min} / RL]}{[P_{min} / RL] + 1} \left(1 - \frac{n}{RL}\right)$; 那么 $\sum_{i=1}^n w_{t_i} \leq RL$.

证明 由已知条件可得:

$$\begin{aligned} \sum_{i=1}^n w_{t_i} &< \sum_{i=1}^n \frac{C_i}{[P_i / RL]} + n = \sum_{i=1}^n \frac{C_i}{P_i} \frac{P_i}{[P_i / RL]} RL \\ &+ n \leq \sum_{i=1}^n U_i \frac{[P_i / RL] + 1}{[P_i / RL]} RL + n \end{aligned}$$

因为上式右边为 P_i / RL 的减函数, 所以可以推出:

$$\begin{aligned} \sum_{i=1}^n w_{t_i} &< \sum_{i=1}^n U_i \frac{[P_{min} / RL] + 1}{[P_{min} / RL]} RL + n \leq \frac{[P_{min} / RL] + 1}{[P_{min} / RL]} RL \\ \sum_{i=1}^n U_i + n &= \frac{[P_{min} / RL] + 1}{[P_{min} / RL]} RL + n < \frac{[P_{min} / RL] + 1}{[P_{min} / RL]} RL \\ &\times \frac{[P_{min} / RL]}{[P_{min} / RL] + 1} RL \left(1 - \frac{n}{RL}\right) + n = RL \quad \text{证毕.} \end{aligned}$$

定理 2 对于消息集 $\{S_n\}$, 如果分组调度中输入连接权值 w_{t_i} 满足 $0 \leq w_{t_i} - \frac{C_i}{[P_i / RL]} < 1$, 则网络最坏情况下的可达负载率满足:

$$U_A^* < \frac{[P_{min} / RL]}{[P_{min} / RL] + 1} \left(1 - \frac{n}{RL}\right) \quad (9)$$

证明: (1) 由定理 1 和引理 1 知式 (9) 成立意味着消息实

时性得到保证.

(2)下面证明对于任意给定实数 $\xi, 0 < \xi < 1$ 至少存在一个消息集 $\{S_i\}$ 的网络负载率 $U \leq \frac{[P_{min}RL]}{[P_{min}RL] + 1}$

$\left(1 - \frac{n}{RL}\right) + \xi$ 使得 $\sum_{i=1}^n w_{t_i} > RL$, 即不满足协议限制条件.

对于给定的 RL, ξ 假设 $P_{min} \geq RL, \alpha = [P_{min}RL], \xi' = \min(\alpha + 1 - P_{min}RL, \xi)$, 则: $\alpha \geq 2, 0 < \xi' < 1$.

构造一个消息集, 具有下列参数:

$$P_1 = P_{min}; C_1 = \xi' P_{min}; P_2 = (\alpha + 1 - \xi')RL;$$

$$C_2 = P_2 \frac{\alpha}{\alpha + 1} \left(1 - \frac{2}{RL}\right);$$

$$\text{则: } U_1 = \xi', U_2 = \frac{\alpha}{\alpha + 1} \left(1 - \frac{2}{RL}\right).$$

$$U = U_1 + U_2 = \xi' + \frac{\alpha}{\alpha + 1} \left(1 - \frac{2}{RL}\right)$$

$$\leq \frac{\alpha}{\alpha + 1} \left(1 - \frac{2}{RL}\right) + \xi$$

$$= \frac{[P_{min}RL]}{[P_{min}RL] + 1} \left(1 - \frac{2}{RL}\right) + \xi$$

$$w_{t_1} = \left[\frac{C_1}{[P_1RL]}\right] + 1 \geq \left[\frac{U_1 P_1}{[P_1RL]}\right] + 1 = \left[\frac{\xi' P_{min}}{\alpha}\right] + 1$$

$$w_{t_2} = \left[\frac{C_2}{[P_2RL]}\right] + 1 \geq \left[\frac{U_2 P_2}{[P_2RL]}\right] + 1$$

$$= RL - 2 - \frac{\xi'}{\alpha + 1} (RL - 2) + 1$$

$$\sum_{i=1}^2 w_{t_i} \geq RL + \left(\left[\frac{\xi' P_{min}}{\alpha}\right] - \frac{\xi' RL}{\alpha + 1}\right) + \frac{2\xi'}{\alpha + 1} \quad (10)$$

因为:

$$\begin{aligned} &\left[\frac{\xi' P_{min}}{\alpha}\right] - \frac{\xi' RL}{\alpha + 1} > \left[\frac{\xi' P_{min}}{[P_{min}RL]}\right] - \frac{\xi' RL}{P_{min}RL} \\ &\geq \left[\frac{\xi' P_{min}RL}{[P_{min}RL]}\right] - \frac{\xi' RL}{P_{min}} \geq \left[\xi' RL\right] - \frac{\xi' RL}{P_{min}} \\ &= \frac{[\xi' RL]}{\xi' RL} \xi' RL - \frac{RL}{P_{min}} \xi' RL \geq \left(\frac{[\xi' RL]}{\xi' RL} - \frac{1}{\alpha}\right) \xi' RL \quad (11) \end{aligned}$$

因为 $1 \leq \frac{\xi' RL}{[\xi' RL]} \leq 2$ 所以 $\frac{1}{2} \leq \frac{[\xi' RL]}{\xi' RL} \leq 1$; 又由已知条件得: $\frac{1}{\alpha} \leq \frac{1}{2}$; 将这两个结果代入式(11)得:

$$\left[\frac{\xi' P_{min}}{\alpha}\right] - \frac{\xi' RL}{\alpha + 1} > 0 \quad (12)$$

将式(12)代入式(10)得:

$$\sum_{i=1}^2 w_{t_i} \geq RL + \left(\left[\frac{\xi' P_{min}}{\alpha}\right] - \frac{\xi' RL}{\alpha + 1}\right) + \frac{2\xi'}{\alpha + 1} > RL.$$

证毕.

引理 2 对于消息集 $\{S_n\}$, 如果在分组调度中以下条件成立: $0 \leq w_{t_i} - \frac{C_i}{[P_iRL]} < 1, U \leq \left(1 - \frac{n}{RL}\right), RL | P_i$; 那么

$$\sum_{i=1}^n w_{t_i} < RL.$$

证明: 由已知条件得:

$$\begin{aligned} \sum_{i=1}^n w_{t_i} &< \sum_{i=1}^n \frac{C_i}{[P_iRL]} + n = \sum_{i=1}^n \frac{C_i}{P_i} \frac{P_i}{[P_iRL]} + n \\ &= \sum_{i=1}^n \frac{C_i}{P_i} \frac{P_i RL}{[P_iRL]} + n = \sum_{i=1}^n U_i \frac{P_i RL}{[P_iRL]} + n \quad (13) \end{aligned}$$

因为 $RL | P_i$, 所以 $\frac{P_i}{RL} = \left[\frac{P_i}{RL}\right]$; 式(13)可简化为:

$$\sum_{i=1}^n w_{t_i} < \sum_{i=1}^n U_i RL + n = RL \sum_{i=1}^n U_i + n = RL U + n$$

$$\leq RL \left(1 - \frac{n}{RL}\right) + n = RL$$

证毕.

定理 3 对于消息集 $\{S_n\}$, 如果分组调度中输入连接权值 w_{t_i} 满足 $0 \leq w_{t_i} - \frac{C_i}{[P_iRL]} < 1, RL | P_i$, 则网络最坏情况下的可达负载率为: $U_A^* = 1 - nRL$ (14)

证明 (1)由定理 1和引理 2知式(14)成立意味着消息实时性得到保证.

(2)下面证明对于任意给定实数 $\xi, 0 < \xi < 1$, 至少存在一个消息集 $\{S_i\}$ 的网络负载率 $U \leq 1 - nRL + \xi$ 使得 $\sum_{i=1}^n w_{t_i} > RL$, 即不满足协议限制条件.

构造一个消息集, 具有下列参数:

$$P_1 = P, C_1 = \xi P, P_2 = (1 - \xi)RL, C_2 = P_2(1 - 2RL);$$

$$\text{则: } u_1 = \xi, u_2 = 1 - 2RL, U = u_1 + u_2 = \xi + (1 - 2RL);$$

$$w_{t_1} = \left[\frac{C_1}{[P_1RL]}\right] + 1 \geq \left[\frac{C_1}{[P_1RL]}\right] + 1$$

$$= [U_1RL] + 1 = [\xi RL] + 1$$

$$w_{t_2} = \left[\frac{C_2}{[P_2RL]}\right] + 1 \geq \left[\frac{C_2}{[P_2RL]}\right] + 1 = [U_2RL] + 1$$

$$= \left[\left(1 - \frac{2}{RL}\right)RL\right] + 1 = RL - 1$$

$$\sum_{i=1}^2 w_{t_i} \geq [\xi RL] + 1 + RL - 1 > RL. \text{证毕.}$$

由于 $RL < P_{min}$, 所以消息集优化前网络最坏情况下的可达负载率不超过 50%, 由定理 2和定理 3可以看出, 消息集优化后网络最坏情况下的可达负载率比优化前提高了至少一倍.

5 仿真结果与分析

将消息延迟时间率作为衡量网络强实时性的指标, 消息延迟时间率定义为消息的实际传输延迟时间与消息的最大允许延迟时间的比率. 最大延迟时间率、最小延迟时间率和平均延迟时间率分别表示消息传输的最大延迟时间、最小延迟时间和平均延迟时间与消息的最大允许延迟时间的比率. 通过 C++ 建模, 基于事件的调度方式对多个消息集进行了离散事件仿真实验, 均得到相近的结果, 现列举其中两例进行说明:

表 1 优化前节点消息流特征

消息名	消息长度(字)	发送周期(m s)
A	2000	0.15
B	2000	0.22
C	2000	0.34
D	2000	0.4
E	2000	0.05
F	1000	2.0
G	1000	3.5
H	1500	3.5
I	200	8.0
J	100	8.0

表 2 优化后节点消息流特征

消息名	消息长度(字)	发送周期(m s)
A	2000	0.1
B	2000	0.2
C	2000	0.3
D	2000	0.4
E	2000	0.05
F	1000	1.5
G	1000	3.0
H	1500	3.0
I	200	3.0
J	100	6.0

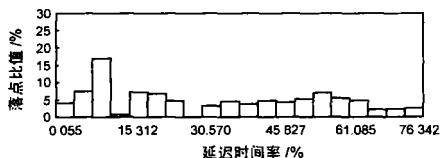


图 2 消息集优化前的延迟落点仿真统计

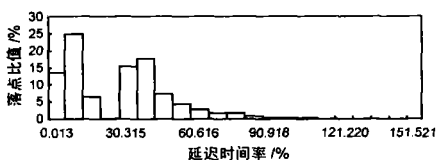


图 3 消息集优化后的延迟落点仿真统计

表 3 消息集优化前后仿真结果比较

比较参数	优化前	优化后
仿真时间	1000m s	1000m s
传输消息总数	37975	42666
网络通信负载率	59.78%	66.75%
最小延迟时间率	0.013%	0.015%
最大延迟时间率	151.521%	92.612%
平均延迟时间率	30.594%	32.049%
超过最大允许延迟周期消息数	298	0

通过表 1和表 2可以看出,消息集优化后比优化前在消息长度保持不变的前提下,增强了消息的周期时限。通过图 2图 3和表 3的统计数据对比可以看出,优化后的消息负载率为 66.75%,最大延迟时间率 92.612%,超过最大允许延迟的周期消息数为 0。优化前的消息负载率为 59.78%,最大延迟时间率为 151.521%,超过最大允许延迟周期消息数为 298。这表明,优化后的消息集提高了网络通信负载率并保证了消息传输的强实时要求,而优化前的消息集不仅负载率较低而且不符合消息的强实时要求。仿真实验结果与理论分析所得到的结论相吻合,验证了理论分析结果的正确性。

6 结论

强实时条件下,合理设定特定消息集对提高系统实时性能具有重要作用。特定消息集的编制应以体现和优化网

络各节点性能与精度为根据。本文根据光纤通道交换型网络协议,通过可达负载率,对系统实时性能进行了理论分析,提出了特定消息集的优化设计,并对消息集优化前后系统在最坏情况下的可达负载率进行了分析比较,为强实时条件下光纤通道交换网络消息集优化设计提供了理论分析基础。值得一提的是,本文的研究成果是以光纤通道交换网络为背景进行的,其设计思想和结论不失一般性,也适用于采用其他协议的交换网络消息集在强实时条件下的优化设计。

参考文献:

- [1] Anzaloni A, Agnitelli N, Avaltoni F. Fiber channel FCS/ATM interworking design and performance study [A]. Global Telecommunication Conference [C]. Houston Texas USA, IEEE, 1994 3: 1801-1807.
- [2] Anzabni A, De Sanctis M, Avaltoni F. Fiber channel (FCS) /ATM interworking a design solution [A]. Global Telecommunication Conference [C]. Houston, Texas USA, 1993 2: 1127-1133.
- [3] Cherkasova L, Kotov V, Rokicki T. Designing fibre channel fabrics [A]. Computer Design VLSI in Computers and Processors [C]. Austin, Texas USA, 1995: 346-351.
- [4] Cherkasova L, Kotov V, Rokicki T. Fibre channel fabrics evaluation and design [A]. System Sciences 1996 Proceedings of the Twenty-Ninth Hawaii International Conference [C]. Maui Hawaii USA, 1996: 53-62.
- [5] Mokoro X, Silla E, Santonja V, et al. On the switch architecture for fibre channel storage area network [A]. Parallel and Distributed Systems [C]. Kyongju City, Korea 2001: 484-491.
- [6] Vama A, Sahai V, Bryant R. Performance evaluation of a high-speed switching system based on the fibre channel standard [A]. High Performance Distributed Computing [C]. Spokane Washington USA, 1993: 144-151.
- [7] Vama A, Murthy S, Bryant R. Using cam-pion to improve the performance of a Fibre Channel switch [A]. Local Computer Networks [C]. Minneapolis Minnesota USA, 1993: 247-255.
- [8] Jane W S Liu. Real Time Systems [M]. Beijing Higher Education Press 2002.

作者简介:

林 强 男,山东烟台人,北京航空航天大学博士生,主要研究方向为数字通信网络的性能优化与仿真评价。E-mail: BUAA_linqiang@163.com.

熊华钢 男,湖北武汉人,北京航空航天大学教授,博士生导师,主要研究领域为嵌入式计算机网络、电子综合化和数字通信。

张其善 男,浙江浦江人,北京航空航天大学教授,博士生导师,主要研究领域为信息处理、卫星通信和智能交通系统。