

46-49,37

计算机科学 1999 Vol. 26 No. 1

混洗交换网络

路由算法

性能分析

性能分析

混洗交换网络旁路路由算法的性能分析及稳定性研究

Study on the Performance and Stability of Shuffle-exchange Networks with Deflection Routing

雷开友

TP393

(西南师范大学计算机科学系 重庆 400715)

Abstract Shuffle-exchange network is a kind of interconnection networks, and widely used in multiprocessor system and ATM switching system. In this paper the performance of shuffle-exchange network is analyzed and the instability of system with random deflection routing is depicted. Therefore a deflection routing algorithm of shortest-distance priority is presented. The system with this algorithm is stable and there is a marked improvement in the system performance.

Keywords Shuffle-exchange network, Random deflection routing algorithm, Shortest-distance priority, Performance, Stability

混洗交换网络是一种互连网络,用其组成的 delta 网络常用于多机系统的互连^[1]。此外,由于混洗交换网络具有路由算法简单,便于硬件实现等优点,也有人将其用于 ATM 交换网络中^[2]。因此,研究混洗交换网络的性能和稳定性具有重要意义。

在互连网络中,当两个或多个信息分组到达同一结点而又需从同一条输出线输出时,必然引起竞争,解决竞争有两种方法:一是胜者输出,败者缓存,待胜者输出完后,败者再输出;二是旁路路由法,将败者信息分组导入到“错误”的输出线,试图让该分组将来经过不同的路径到达目的地。显然,旁路路由法具有控制简单便于实现的优点,而且旁路路由法还能使信息分组自动从一阻塞链扩散,可以有效防止网络阻塞,所以这种方法得到了广泛应用。

对于混洗交换网络,已有人通过分析其时延来研究网络的性能,得出了网络的“时延-负荷(delay-offered load)”曲线^[3]。虽然使用分析网络时延的方法是一种常规做法,但“时延-负荷”曲线却不能揭示网络潜在的不稳定性。注入网络相同的负荷,网络可以处于两个不同的工作点,其中之一是不稳定的,网络吞吐率不能维持在给定负荷的水平上,系统很快进入一个稳定且饱和的工作点,这将导致网络性能的崩溃,因此,我们不希望网络处于一个不稳定的工作区,而应通过控制网络的整体负荷使其工作在一个稳定的区域。

本文分析混洗交换网络使用随机旁路路由算法

及最短距离优先旁路路由算法时的性能和稳定性问题。通过分析及模拟实验可以得出随机旁路路由算法使网络具有不稳定性,而最短距离优先旁路路由算法能较高地提高网络性能,使网络处于稳定的工作状态。

1 网络模型及路由算法

1.1 网络模型

本文研究单级混洗交换网络,将输出链直接连回到输入链,其网络模型可定义如下:

定义 1 (网络结点) 网络结点为 2×2 开关,两条输入线分别为 (I_0, I_1) , 两条输出线分别为 (O_0, O_1) , 可以组成 4 条不同的路由: (I_0, O_0) ; (I_0, O_1) ; (I_1, O_0) ; (I_1, O_1) 。

显然,4 条路由可能存在以下冲突: (I_0, O_0) 和 (I_1, O_0) 竞争输出线 O_0 ; (I_0, O_1) 和 (I_1, O_1) 竞争输出线 O_1 。

定义 2 (单级混洗交换网络) 设 $n = \log_2 N$, 那么 N 个结点的单级混洗交换网络可以按如下方式构成:

- 使用二进制数 $x_n x_{n-1} \cdots x_1$ 来标识网络结点,其中 $x_i = 0$ 或 $1, i = 1, \cdots, n$;
- 不失一般性,设每个网络结点上面的输出为 O_0 , 下面的输出为 O_1 ;
- 结点 $x_n \cdots x_1$ 经上面的输出 O_0 连到结点 $x_{n-1} \cdots x_1 0$; 经下面的输出 O_1 连到结点 $x_{n-1} \cdots x_1 1$

通过以上方式即可构成 N 个结点的单级混洗交换网络,在这种网络中,一个信息分组将要访问的下一个结点由当前结点 $x_n \cdots x_i$ 通过路由信息而获得。根据定义,由 4 个结点可构成如图 1 所示的混洗交换网络。

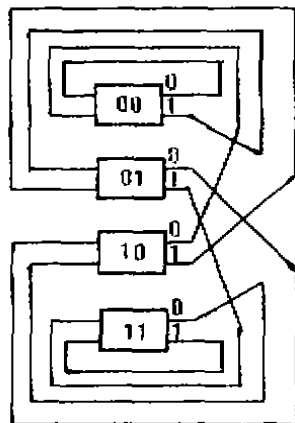


图 1 四结点混洗交换网络

1.2 路由算法

现在考虑一个分组从源结点 $s_n \cdots s_1$ 到目标结点 $d_n \cdots d_1$ 的路由问题。在此网络中,我们将目标结点地址 $d_n \cdots d_1$ 作为分组头并用于路由选择,算法如下:

- 在源结点,使用目标地址的 d_n 做路由选择: If $d_n = 0$ Then 分组送到结点 $s_{n-1} \cdots s_1 0$ Else 分组送到结点 $s_{n-1} \cdots s_1 1$;

- 设分组已传到第 i 步,此时已到达结点 $s_{n-1} \cdots s_1 d_n d_{n-1} \cdots d_{i+1}$,那么使用目标地址的 d_{n-1} 做路由选择: If $d_{n-1} = 0$ Then 分组送到结点 $s_{n-1} \cdots s_1 d_n d_{n-1} \cdots d_{i+1} 0$ Else 分组送到结点 $s_{n-1} \cdots s_1 d_n d_{n-1} \cdots d_{i+1} 1$;

- 当 i 等于 n 时,分组到达目标结点 $d_n \cdots d_1$ 。

显然,如果没有其他分组相冲突,那么一个分组从源结点经过 n 步即可到达目标结点,它所经过的结点序列如下:

$$s_n \cdots s_1 \rightarrow s_{n-1} \cdots s_1 d_n \rightarrow s_{n-2} \cdots s_1 d_n d_{n-1} \rightarrow \cdots \rightarrow d_n \cdots d_1$$

本文假设固定长度分组,时隙同步网络操作,每一时隙对应的时间为一个分组从一个结点传到另一个结点的时间。所以,一个分组需要 n 个时隙(n 步)才能到达其目的地。如果一个分组在某个结点与另一个分组竞争同一输出端,那么它可能被旁路到“错误”的输出端,从而需要更多的时隙才能到达其目的

地。对于此情况,随机旁路算法为:

- 当两个分组竞争同一输出端时,以相同概率随机产生胜者;
- 胜者按正常路由送到输出端,败者被旁路到另一个空闲的输出端;
- 败者被旁路后按源结点方式从头开始选择路由。

最短距离优先旁路路由算法:

- 当两个分组竞争同一输出端时,离目标结点距离最近者获胜;
- 胜者按正常路由送到输出端,败者被旁路到另一个空闲的输出端;
- 败者被旁路后按源结点方式从头开始选择路由。

显然,一个分组一旦被旁路,它离目标结点的距离被重新置为 n 。可以肯定这种算法不是一种最优算法,因为在混洗交换网络中两个结点之间的最短路径并不总是 n 步,然而,按最短路径的最优路由策略实现非常复杂;在每个结点查路由表或进行实时计算。最短路径的路由策略的分析更加困难,它的稳定性问题是将来进一步研究的很好课题。

1.3 网络访问

分组一旦进入网络,到达其目的地只是个时间问题,不存在分组丢失,然而,如果注入网络的负荷超过最大的吞吐率,分组可以从网中卸出来。

图 2 为一个结点的内部结构,对每一个结点存在一种机制让分组进入网络或从网络中输出。想进入网络的分组在输入队列中排队,队列中最前面的一个分组等到输入链的空时隙出现时即可进入网络。如果分组进入输入队列的速度高于空时隙到达的速度(即想注入网络的负荷大于网络吞吐率)时,输入队列的缓冲区会溢出,一些分组被迫撤除。

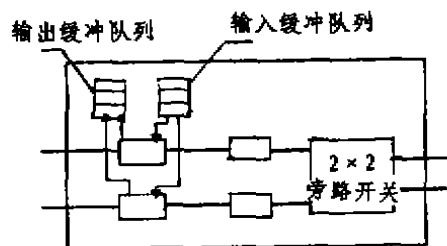


图 2 混洗网络结点的内部结构

对网络的访问有许多种不同的策略,最简单的是贪婪策略。这种策略是:一旦发现输入线之一有空

的负荷: $\Lambda = \Lambda_0$; 当 $\Lambda_{sat} \leq \Lambda_0 \leq \Lambda_{max}$, 系统处于稳定或不稳定区域。在稳定区域, 吞吐率等于提供的负荷: $\Lambda = \Lambda_0$; 在不稳定区域, 当 $\rho > \rho_{max}$ 时, 系统很快终止于 $\rho = 1$ 和 $\Lambda = \Lambda_{sat}$, 进入饱和状态, 当 $\rho' \leq \rho \leq \rho_{max}$ 时, 系统处于非饱和状态, 且 $\Lambda = \Lambda_0$ 。这种不稳定性与 Aloha 多路存取网络中的不稳定性属同类性质^[4]。

对任意对称网络(不仅仅是混洗交换网络), 我们可导出更一般的网络稳定条件: 设系统中有 M 条链, 根据李特定律(Little's Law)有 $\Lambda = M\rho/T$, 其中 T 为平均时延。 Λ 与 ρ 的微分

$$d\Lambda/d\rho = [TM - M\rho(dT/d\rho)]/T^2$$

因此, $d\Lambda/d\rho > 0$ 当且仅当 $dT/d\rho \leq T/\rho$

则稳定性的一般条件为 $dT/d\rho \leq T/\rho (0 \leq \rho \leq 1)$ 。

图 5(a) 为每结点的吞吐率, Λ/N 与 ρ 的关系曲线。四条曲线分别对应 $n=5, 10, 15, 20$ ($n=5$ 为最上面的曲线, $n=20$ 为最下面的曲线), 实线为分析结果, 叉线为模拟结果。

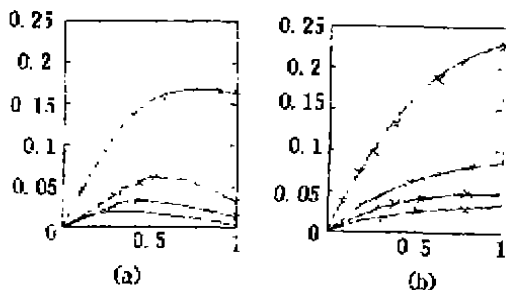


图 5 每结点与链负荷的关系曲线
($n=5, 10, 15, 20$)

从曲线分析可知, 每条曲线有一个不稳定区, 我们的模拟实验进一步证实了这一结论, 在不稳定区, 因不存在系统平衡点而无法获得模拟数据。

当网络变大(即 n 变大)每结点的吞吐率很快下降, 由不稳定性所带来的问题就更加严重, ρ' 和 ρ_{max} 变小, 以至于系统很容易波动到不稳定区; Λ_{sat} 变为 Λ_{max} 的更小的部分。

表 1 为 $n=10$, 采用贪婪访问方法的一个模拟结果。可以看出, 最大吞吐率可达到 0.42, 当注入输入队列的负荷大于网络可支撑的负载时, 网络进入不稳定区, 链负荷变饱和。不稳定发生时, 吞吐率降到一个很低的点上。表 1 还告诉我们, 尽管在分析中可以导出这一区域, 但在这一区域获得“吞吐率-链负荷”数据是不可能的。

表 1 混洗交换网络随机旁路由算法的模拟结果($n=10$)

提供的负荷	链负荷	吞吐率	平均时延
0.010000	0.054383	0.010005	10.871390
0.020000	0.120500	0.020010	12.041059
0.030000	0.207470	0.030016	13.823852
0.040000	0.355659	0.040036	17.767030
0.041000	0.382201	0.041031	18.529957
0.042000	0.417855	0.042020	19.888816
0.043500	1.000000	0.0282167	70.881493
0.045000	1.000000	0.028201	70.919990

2.2 最短距离优先旁路由算法的网络性能与稳定性

从前面的分析可以看出, 随机竞争策略虽然实现简单, 但系统具有不稳定性, 严重影响了系统性能。为此, 我们寻求一种新的竞争解决策略, 以消除系统的不稳定性。在此提出最短距离优先竞争策略, 经分析它可消除系统的不稳定性。

当一个分组在状态 i 时被旁路回状态 n , 那么前面已走过的 $n-i+1$ 步浪费了, 使用最短距离优先策略实际上就是要减少这种浪费。

从前面介绍的最大距离优先路由算法可知, 旁路概率与状态相关, 我们使用 q_i 代表在状态 i ($1 \leq i \leq n$) 时的旁路概率。为了分析方便, 我们假设网中存在 $2N$ 个虚拟令牌, 网络中每条链对应一个令牌。在输入队列中的分组必须获得一个自由令牌才能进入网络。当分组到达目标结点时, 令牌释放而成为自由令牌, 它可以被另一个分组所获得。

在状态 i ($1 \leq i \leq n$), 如果一个分组使用了令牌, 则称该令牌是活动的, 如果一个令牌未被使用, 或处于状态 0, 则该令牌是自由的。就令牌的状态变化而言, 它没有停止状态, 因为任何一个自由令牌可被一个分组所获得而离开状态 0。

我们追踪一个特定令牌的状态变化过程。当一个时隙开始时, 设 $P(i)$ 为某令牌在状态 i 的概率, 有:

$$P(i-1) = P(i)(1-q_i) \quad 2 \leq i \leq n \quad (5)$$

定义 $\pi(i) = P(1) + \dots + P(i) \quad 1 \leq i \leq n$ (6)
为令牌在状态 $1, 2, \dots, i$ 时的概率。

根据定义, 令牌为活动的概率为 $\pi(n)$, 一个活动的令牌意味着占有该令牌的链路也是活动的, 因为每条链路总有一个令牌, 活动令牌的概率即为活动链路的概率:

$$\pi(n) = \rho \quad (7)$$

一个分组在状态 i 被旁路的概率等于在状态 i 被一个分组旁路的概率以及所有小于 i 的状态被

(下转第 37 页)

息,而 MPOASS 上记录了子网中,或局域网中所有 MPC 的不可信任访问的记录。所以 MPOASS—MPC 模型更具实用性。

3. 可维护性 MPOASS—MPC 认证模型为系统管理人员提供一个统一的界面,来配置网络安全的所有设置,对某台 MPC 的特殊配置也可以在 MPOASS 上完成,整个安全系统维护性很好。在 MPCLFW 中,需要对每个 MPC 进行单独设置,维护性差。

4. 灵活性 MPOASS—MPC 通过对 ATM 主机按信任级别分类的方法,在对不同的主机提供不同的安全策略方面,较 MPCLFW 模型有明显的灵活性,特别是我们在原型系统中,圆满地处理了用户先登录到网络防火墙上,利用 ATM C 类主机的 CheckIn, CheckOut 方式,使该主机由 C 类变为 B 类主机,再登录到子网内部的其它主机上,使该主机

充分使用子网提供的优质服务的情况。在 MPCLFW 模型下是难以实现的。

参考文献

- 1 ATM Forum Multi-Protocol Over ATM Version 1.0, 1997. 5
- 2 Anixter, 1996, ATM Update 1
- 3 吴定一,等,异步传递模式的理论与应用,1997
- 4 ATM Forum, LAN Emulation Over ATM Version 2.0, 1997. 7
- 5 Paul Wernick, British Telecom URI Security: Project Outline, 1995
- 6 Anthony Alles, ATM Internetworking, 1995
- 7 Routing over Large Clouds Working Group, NBMA Next Hop Resolution Protocol (NHRP)
- 8 Cisco, Anixter, ATM Forum, UB Networks 等公司,组织的主页

(上接第 49 页)

一个分组旁路的概率之和:

$$q_i = \frac{P(i)}{4} + \frac{P(1) + \dots + P(i-1)}{2} = \frac{P(i)}{4} + \frac{\pi(i-1)}{2} \quad 1 \leq i \leq n \quad (8)$$

将(8)代入(5),经推导可得:

$$\pi(i-1) = [\pi(i) - \pi(1)] \left\{ 1 - \frac{\pi(1)}{2} - \left[\frac{\pi(i) - \pi(1)}{4} \right] \right\} \quad 2 \leq i \leq n \quad (9)$$

一个分组在状态 1 不被旁路的概率为 $1 - \pi(1)/4$, 不被旁路意味着此令牌携带的分组到达其目标结点。因此网络的吞吐率为:

$$\Lambda = 2N\pi(1) \left[1 - \frac{\pi(1)}{4} \right] \quad (10)$$

虽然我们无法从(7)和(9)解出 $\pi(1)$,但可以证明系统不存在不稳定区域。

$$\text{证明: } \frac{d\Lambda}{d\pi(1)} = 2N \left[1 - \frac{\pi(1)}{2} \right] > 0 \quad (11)$$

$$\text{据(9) } \frac{d\pi(i)}{d\pi(i-1)} = \frac{\left\{ 1 - \left[1 - \frac{\pi(1)}{2} \right] \frac{d\pi(1)}{d\pi(i-1)} \right\}}{1 - \frac{\pi(1)}{2}} \quad 2 \leq i \leq n \quad (12)$$

$$\text{令 } i=2 \text{ 可得 } \frac{d\pi(2)}{d\pi(1)} > 0$$

$$\text{据(12)可推出 } \frac{d\pi(i)}{d\pi(1)} > 0, 1 \leq i \leq n$$

$$\text{因此 } \frac{d\Lambda}{d\varphi} = \frac{d\Lambda}{d\pi(n)} = \frac{d\Lambda}{d\pi(1)} \times \frac{d\pi(1)}{d\pi(n)} > 0 \quad (13)$$

图 5(b)示出了每结点对应于 $n=5, 10, 15, 20$ 的吞吐率—链负荷曲线,在可达到的最大吞吐率方面与随机竞争策略相比有明显的提高,而且使系统工作在稳定区域。

结语 通过上面的分析可知,在混洗交换网络中,如果使用随机旁路算法解决竞争冲突,当网络规模较大($n \geq 5$)时,系统存在不稳定性,性能得不到保障;而使用最短距离优先旁路算法时,较好地解决了系统的稳定性问题,提高了系统的整体性能,但最短距离优先策略较随机旁路策略复杂,在选取获胜者时需比较两个分组已走过的步数,增加了实现的难度。因此,在实际的系统中,应具体情况具体分析,根据需要适当选择冲突解决算法。

参考文献

- 1 童颖,程代杰,多处理机及智能多机系统,重庆大学出版社,1988
- 2 杨宗凯,ATM 理论及应用,西安电子科技大学出版社,1996
- 3 Krishna A, et al. Performance of shuffle-like switching networks with deflection. In: Proc. IEEE INFOCOM'90, 1990. 473~480
- 4 Bertsekas D, et al. Data Networks, 2nd ed. Englewood Cliffs, NJ: Prentice-Hall, 1992