

基于布尔集线器的线速多播自路由交换结构

尘福兴¹, 李挥¹, 崔凯¹, 张博²

(1. 北京大学 深圳研究生院 深圳市云计算关键技术和应用重点实验室, 广东 深圳 518055;

2. 国家数字交换系统工程技术研究中心, 河南 郑州 450002)

摘要: 现有的多播交换结构均是从单播交换结构发展而来, 而目前的各种单播交换结构都存在各自的技术瓶颈, 当被应用于多播交换时, 既不能突破原交换结构的技术局限, 又限制了多播交换的性能, 因此现有的多播交换结构基本都不具备大规模可扩展的能力。提出基于分配格理论的线速多播交换结构, 它不仅能满足大规模可扩展和线速多播要求, 而且具有低时延、无抖动, 不需要排队缓存等性能优势。分析并证明了该多播交换结构既能实现最优化的多播, 又支持具有优先级的多播, 从而在时延上保证了多播 QoS(quality of service)。

关键词: 线速多播交换结构; 多播集线器; 分配格; 布尔集线器

中图分类号: TN919

文献标识码: A

文章编号: 1000-436X(2014)07-0129-11

Novel wire-speed multicast self-routing switch fabric based on Boolean concentrator

CHEN Fu-xing¹, LI Hui¹, CUI Kai¹, ZHANG Bo²

(1. Shenzhen Key Lab of Cloud Computing Technology and Application, Shenzhen Graduate School, Peking University, Shenzhen 518055, China;

2. National Digital Switching System Engineering Technological R&D Center, Zhengzhou 450002, China)

Abstract: The traditional multicast switch fabrics (MSF) are mainly developed from the unicast switch fabrics (USF). Owing to the bottleneck of the current USF, almost all the existing MSF are not able to achieve high efficiency or excellent large-scale scalability. A wire-speed MSF based on lattice theory is presented. It cannot only meet the requirements of superior scalability and wire-speed multicast, and has low delay, jitter free, and no queuing cache performance advantages. It is proved that the MSF can both achieve optimal multicast performance and support priority-based multicast which can guarantee the multicast QoS (quality of service) in delay.

Key words: multicast switching fabric; multicast concentrator; distribution lattice; Boolean concentrator

1 引言

1984年, Huang和Knauer首次对ATM网络中多播数据的交换问题予以关注, 并设计出一种新型交换结构来支持多播, 提出了第一个多播交换结构——Starlite^[1]。自此, 业界对高效的多播结构进行了持续的研究。其中, Saito提出了基于共享存储式的交换结构来实现多播的方法^[2]。但是由于存储器

带宽的限制使该结构不能够满足大规模扩展需求。Yeh提出了一种基于Knockout理论多播交换结构, 每个输入端口采用完全连接的方式连接到输出接口模块^[3]。Chao提出了基于该结构的多播设计方案, 但是由于文献[3,4]结构的级间采用总线式完全连接方式, 所以该结构存在线路器件驱动能力的极限问题, 使该类方案在大规模扩展性能上存在瓶颈。

收稿日期: 2013-03-26; 修回日期: 2013-05-20

基金项目: 国家重点基础研究发展计划(“973”计划)基金资助项目(2012CB315904); 国家自然科学基金资助项目(NSFC61179028); 深圳基础研究基金资助项目(JCYJ20130331144502026, JC201104210120A, JC201005260234A); 广东省自然科学基金资助项目(S201101000923, S2013020012822)

Foundation Items: The National Basic Research Program of China (973 Program) (2012CB315904); The National Natural Science Foundation of China (NSFC61179028); The Basic Research of Shenzhen (JCYJ20130331144502026, JC201104210120A, JC201005260234A); The Natural Science Foundation of Guangdong Province (S201101000923, S2013020012822)

基于 banyan 的多播交换结构大多采用多播复制的方法, Lee 提出了一种内部无阻塞的复制网络设计方案^[5], 由于该方案累加和网络及虚拟地址编码器不适合大规模扩展, 会发生满溢(overflow)的现象; 多播地址转换所需缓存开销庞大, 结构过于复杂, 难以满足大规模扩展的需求。文献[6]提出了一种能够为复制网络提供公平机制的方案, 然而它进一步增加了结构复制的复杂性, 使其难以应用于大规模的交换结构设计中。

文献[7]提出了一种基于 Clos 交换结构多播调度算法, 引入了一种路径分配向量, 但整个路径决策时间复杂度较高, 无法大规模扩展。文献[8]证明了对 Clos 网络进行多播路径匹配是 NP 完全问题。

基于 Crossbar 多播交换结构的研究则大都是关于多播调度算法。Hui 和 Ali 分析了在随机情况下扇出和非扇出交替类的多播调度算法的性能, 这些研究都证明了由于信元头阻塞 (HOL blocking) 会导致吞吐率大大降低^[9,10]。Marsan 在理论上证明了对于采用虚拟输出队列 (VOQ) 的大规模多播交换结构, 不存在一个有限的提速因子能使许可的多播流量达到 100% 的吞吐率^[11]。Pan 提出了一种将信元载荷与目的地址信息分开存储的方案, 但其要求的 $(N+1)N$ (N 为输入/输出的端口数) 倍于外部链路带宽的加速比, 本身就是大规模扩展的瓶颈^[12]。

从以上对多播交换结构的研究中可以看出, 现有的多播交换结构都无法满足大规模可扩展、线速多播的要求, 本文在分配格理论基础上, 通过将“统计线组”的技术应用到分治网络中, 用合适的集线器来代替该网络中节点, 并从理论上对该结构性能证明, 从而得到具有低时延、无抖动和无需排队缓存优势以及满足大规模可扩展以及线速多播需求的线速多播交换结构。该结构在一定条件下能很好地解决多播交换结构大规模可扩展以及线速多播问题的同时, 又能保证较好的服务质量。

2 多级互联完全分布式自路由模型和布尔多播交换单元

2.1 多级互联分布式自路由模型

多级互连完全分布式自路由模型^[13]以分治网络(divide & conquer)^[14]为基础, 而构建的具有高度模块化, 最低的器件复杂度 ($N \lg N$) 等优点的多播交换结构模型。以 $N=64$ [id:(6 5):(6 5 4):(6 3)(5 2)(4 1):(6 5):(6 5 4):id] 的分治网络为例, 如图 1

所示。其中, “:” 表示一级 2×2 单元, 而两级间连线方式用一组位置换群元素表示, 如(6 3)(5 2)(4 1)。

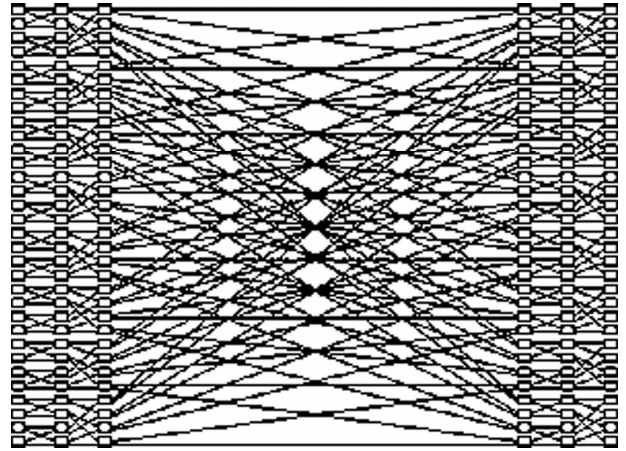


图 1 64×64 的分治网络

该类网络的自路由特性由其“迹 trace”和“导 guide”^[15]序列共同决定, 给出了网络的“迹 trace (TK)”和“导 guide(RK)”序列及其自路由过程。对于一个 $N \times N$ 交换结构, 多级互连网络 (MIN, multi-stage interconnection network) 由于其分布式自路由特性, 以及其理想的单元复杂度 $\Theta(N \lg N)$, 因而适用于构造大规模交换结构。交换结构的艺术就是由最基本的 2×2 元件通过各种拓扑结构以最低的代价构造出符合性能指标的大型交换系统。理想的结构是其规模可以任意递归扩展, 不存在任何瓶颈。因此基于该分布式自路由模型构建多播交换结构模型的理由和动机是充分的、明显的。构建多播交换结构的第一步应该分析分布式自路由模型的基本交换单元及其原理。

图 1 所示基本单元是由图 2(a)~图 2(c)状态的自路由排序单元构成, 进入该 2×2 基本自路由排序单元的数据分组排序交换过程可以使用带内信令来实现, 例如, 在数据分组的前面加上 2 位带内信令。第一位 A_1 表示当前时隙是否有数据分组, 1 表示存在有效数据分组, 0 表示不存在有效数据分组即为空分组; 当 A_1 为 1 时第二位 D_1 表示分组的输出目标端口地址才有意义, 否则 D_1 为无效位。故 $A_1 D_1$ 等于 ‘10’、‘11’、‘00’、‘01’ 时分别代表“输出目标端口为 0 的有效数据分组”, “输出目标端口为 1 的有效数据分组”, “空分组”, “空分组”。

自路由模型的基本 2×2 单元(如图 2(a)~图 2(c))按如下规定的线性排序关系实现自路由: $10 < 00 < 11$, 具体控制方式可参见表 1。

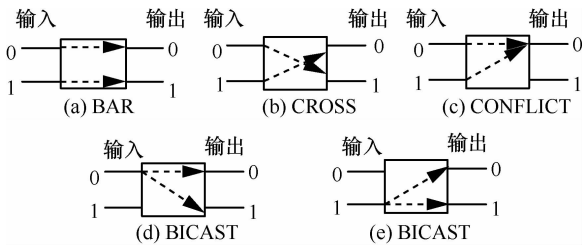


图 2 2x2 基本自路由排序单元处于 BAR、CROSS、CONFLICT、BICAST 的状态

表 1 中, CONF(CONFLICT)表示冲突, 由其他方式比如优先级来决定哪个信号会被输出。

表 1 单播 2x2 基本自路由单元以两位带内信令自路由的交换控制

单元状态	INPUT-1 带内信令: A_1D_1		
	10	00	11
INPUT-0 带内信令: A_1D_1	10	00	11
	CONF	BAR	BAR
	00	CROSS	EITHER
	11	CROSS	CROSS
			CONF

自路由模型的基本单播 2x2 单元显然不具备硬件电路拷贝的线速多播特性。为了达到线速多播交换的目的, 需要将图 2 中只有 BAR 和 CROSS 2 种单播路由方式的排序单元进行重新定义, 使其成为一种全新的满足多播要求的排序交换单元, 称为多播交换单元 (如图 2(d)、图 2(e)所示)。基本交换单元在加入 BICAST 状态后, 需要满足何种条件才能达到线速多播的要求, 为了能够方便描述和证明多播单元以及由其构成的多播网络特性, 本文将从分配格的理论角度来说明。

2.2 布尔多播交换单元

在表 1 中定义输入控制的基础上, 可以进一步定义 $\Omega_{route} = \{0-bound, 1-bound, idle\}$, 即 $0-bound=10$; $1-bound=11$, $idle=00$ 。故 $10 < 00 < 11$ 的线性排序等于如下规则的顺序。

规则 1 $0-bound < idle < 1-bound$

路由单元按照规则的序列排序即可将尽可能多的有效分组转发到预定的输出端口。当出现 2 个 $0-bound$ 分组或 2 个 $1-bound$ 分组的输出争用时, BAR/CROSS 状态的选择依赖于特定的应用, 如信号的优先级。这样的操作不影响信号自路由的最优性。

当排序单元应用于多播时, 必须定义新的带内信令和相应的交换控制方式, 例如表 2 所示。

表 2 支持多播以线速扇出的 2x2 多播排序单元的信号控制

单元状态	Input-1 带内信令				
	0	1	B	I	
Input-0 带内信令	0	CONF	BAR	BAR	BAR
	1	CROSS	CONF	CROSS	CROSS
	B	CROSS	BAR	EITHER	BICAST
	I	CROSS	BAR	BICAST	EITHER

表 2 中, B 表示 BICAST 多播分组 (如图 2(d)、图 2(e)所示); I 表示 IDLE 空分组。

可见支持多播的自路由带内信令控制表必须扩展为 $\Omega_{bicast} = \Omega_{route} \cup \{bicast\}$ 。

这个扩展集合由规则 1 和规则 2 各自部分排序得到。多播单元对属于集合 Ω_{bicast} 的信号执行排序, 当 idle 遇到 bicast 时交换控制遵循规则 3。

规则 2 $0-bound < bicast < 1-bound$

规则 3 output-0 端口输出信号值 $0-bound$, output-1 端口输出信号值 $1-bound$, 其后都跟随着相同的负载。

这样, 分组的负载在通过多播单元时可以是 BAR 状态、CROSS 状态或者 BICAST 状态的硬件电路拷贝的线速多播方式。按这种方式, 多播单元能将尽可能多的分组负载转发到其预定目标端口。

为了能够方便分析多播单元的特性, 排序规则 1、2 与交换规则 3 需要寻求一种可以同时满足 3 种排序规则的理论, 而分配格理论刚好能够满足需求。格是一个有二元操作数布尔和 \vee 与布尔积 \wedge , 且遵守布尔交换律、布尔结合律、布尔幂等律、布尔吸收率的集合^[15]。

常见的格是遵守布尔操作中并集和交集运算的集合的一组子集。

规则 4 $a \leq b \Leftrightarrow a \wedge b = a$

根据规则 4, 每一个格都可以导出一个部分有序序列。如果一个部分有序集合 (简称偏序集) 中任意两元素存在一个最大下界和一个最小上界, 则称这个偏序集为偏序格, 用布尔和与布尔积表示这 2 个界限, 偏序格逻辑上等同于格。图 3 所示为 2 种格的举例, 其中图 3(a)是有 3 个变量的自由格, 图 3(b)为分配格。

根据规则 1 和规则 2, 图 3(b)描述了 Ω_{bicast} 的部分排序, 很明显偏序集 Ω_{bicast} 具有偏序格的特性, 因此被称为格。至此多播单元运行的 3 个规则 1 到规则 3 能够被统一于以下布尔规则 (规则 5)。

规则 5 输出端口 0 的值为 2 个输入的布尔积，输出端口 1 的值为 2 个输入的布尔和。

如果一个格满足分配律

$$a \wedge (b \vee c) = (a \wedge b) \vee (a \wedge c)$$

$$a \vee (b \wedge c) = (a \vee b) \wedge (a \vee c)$$

则此格为分配格。 Ω_{bicast} 就是一个简单的例子，如图 3(b)所示。因此有时也把该多播单元称为布尔单元。

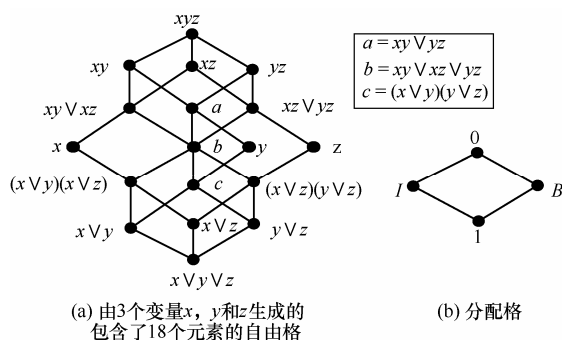


图 3 2 种格的举例

本研究采用分配格来构建多播排序和多播交换的自路由带内信号表。到目前为止，常用交换结构信号表使用的都是线性有序集，尤其当信号值为数字表示的时候更是如此。对所有带内信号进行完全排序的要求不仅限制了结构的应用范围，而且完全阻碍了多播操作。把 2×2 多播单元与 2×2 单播单元从分配格的角度来对比可以清楚地看出这一点。

2.3 多播集线器到布尔多播集线器一般化

定理 1 (多播集线器定理^[16]) 以排序单元多级互联网络构建的 n -to- m 的集线器 (n 个输入， m 个输出的集线器) 为参考，将其中的排序单元用多播单元替换。记信号表为 Ω_{bicast} ，不妨设，输入 V_0 个 0-bound， V_1 个 1-bound，以及 V_B 个多播信号。可得到以下结论。

结论 1 上面 $n-m$ 个输出端口最大可能产生共 $\min\{n-m, V_0 + V_B\}$ 个 0-bound 和多播信号。

结论 2 下面 m 个输出端口最大可能产生共 $\min\{m, V_1 + V_B\}$ 个 1-bound 和多播信号。

2.2 节把基本交换排序单元统一到布尔单元并用布尔代数来描述，此处由基本单元多级互联构成的多播集线器同样用布尔代数来描述。

定义 1 令 V_j ($0 \leq j \leq n$) 为 n 个变量 $X_0, X_1, \dots, X_{(n-1)}$ 的合取范式。若 $1 \leq m < n$ ，当

$$V_j \leq \sigma_{m+1}(X_0, X_1, \dots, X_{n-1}), r_0 \leq j < n-m \quad (1)$$

$$\sigma_m(X_0, X_1, \dots, X_{n-1}) \leq V_k, n-m \leq k < n \quad (2)$$

成立时，相应的布尔交换器被称为是 n -to- m 布尔集线器。

$n \times n$ 布尔排序器在一定条件下等同于布尔集线器。如 2×2 布尔排序器和 2-to-1 布尔集线器，其实和布尔单元一样。

推论 1 当 V_j 和 σ_k 应用在任意分配格中的 n 元素而不是变量 $X_0, X_1, \dots, X_{(n-1)}$ 的时候，式(1)、式(2)在定义 1 中仍是正确的。

在推论 1 中分配格的假设是关键，因为分配率的应用在合取范式中的作用是非常重要的。当一个有序集合是分配格时，布尔集线器相对于普通集线器是一个高级版本。

定理 2 (布尔集线器的 0-1 原理^[16]) 当且仅当输入任意包含 $n-m$ 个 0 和 m 个 1 的序列后，上面的 $n-m$ 个端口输出值为 0，下面 m 个端口输出值为 1 时， $n \times n$ 布尔交换器就是 n -to- m 布尔集线器。

3 大规模扩展线速多播交换结构模型

为了构建大规模扩展多播交换结构模型，本文把“统计线组”^[15]技术应用到 2.1 节的分治网络中。将自路由分治网络中的每一个 2×2 节点放大为 $2G \times 2G$ 节点，并用 $2G$ -to- G 布尔多播集线器替换该节点，交换网络中每一根连接线用一束共 G 根线替换，从而构造了一个带有统计复用特点的多播自路由结构，每 G 根输出线共享一个群组地址。由于通信波动和突发性引发的分组丢失率随 G 值的增大呈指数级减小^[13]，在可允许的输入流量下，“统计线组”技术可以达到几乎无阻塞交换。

在实际应用中 G 表示为群组，将 G 根线组成的一束记为一个群组，这个值一般比较大， K 表示群组数， N 表示交换结构的输入/输出的总线数 $N = KG$ 。

如图 4(a)所示为一个 $N=128, K=16, G=8$ 的多播路由交换结构。该结构是通过将 16×16 的 banyan 网络中每一个 2×2 节点替换为 $2G$ -to- G 的由布尔单元组成的布尔集线器来实现。这里 $G=8$ ，在实际应用中 G 本应该为一个大的数，所以一个 $2n \times 2n$ banyan-type 的 G -line 版本的网络构建成了一

个 $N \times N$ 的几乎无阻塞的多播交换器。根据 fast knockout^[17]集线器构造算法或利用 bitonic circular 再配合一级半清器 (half-cleaner)^[15], 可构造 G 为任意大小的群组集线器, 从而保证了交换结构的大规模可扩展性能。

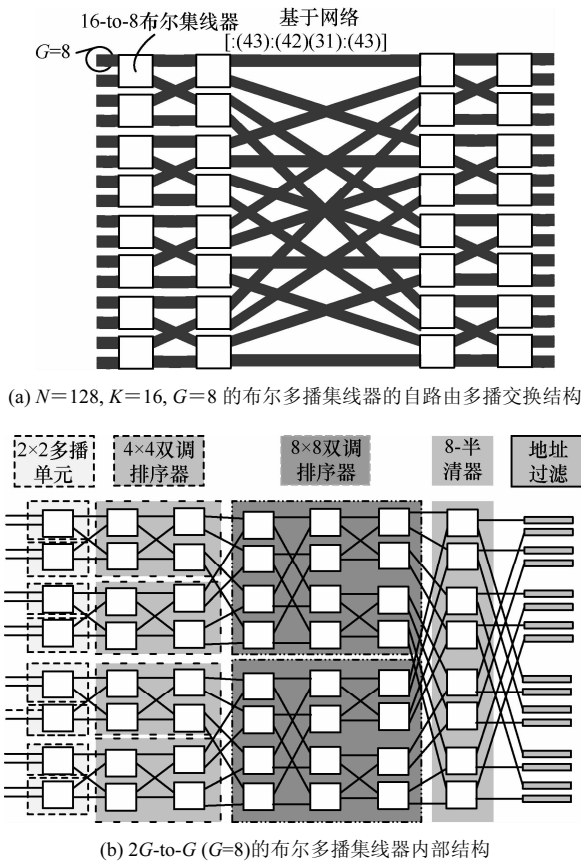


图 4 布尔多播集线器交换结构及内部结构

图 4(b)所示为图 4(a)中任一 $2G$ -to- G 的布尔多播集线器的内部结构图。图 4(b)中的每一级 (除了最后一级地址过滤器) 中最小的单元即为 2.2 节的布尔多播交换单元。由图 4 可以对本文提出线速多播交换的内部结构有一个更详细的理解。

本文提出的线速多播交换模型所采用的完全自路由分治网络结构和布尔多播集线器都具有良好的大规模可扩展性, 交换结构中每个最小的交换单元使用硬件电路拷贝的方式实现多播, 从而保证了交换结构的线速多播以及大规模可扩展特性。此外, 交换结构模型中无排队缓存, 更不需要调度从而保证了多播时延不会高于某一个定值。

对于定理 1, 当有一个适当的多播信号表 Ω_{bicast} 为输入时, 多播集线器就能够达到最优的多播交换。而多播交换网络则是由多播集线器递归构造而

成, 当用布尔单元代替集线器网络中的排序单元时, 对任意的一个格或分配格结构的信号表, 该定理是否也成立。由此深入考虑 Ω_{bicast} 格结构的支持多播集线器定理的本质属性, 仔细观察发现 Ω_{bicast} 在结论 1 中分为上部理想 $\{0, B\}$ 和下部理想 $\{1, 1\}$, 在结论 2 中类似地分为 $\{1, B\}$ 和 $\{0, 1\}$ 。

定理 3 如果格 Ω 的非空子集 S 满足条件 $x \in S, y \in S \Rightarrow x \wedge y \in S, x \vee y \in S$, 则 S 为子格; 如果 $x \in S, y \in S \Rightarrow x \wedge y \in S$, 则子格 S 是一个上部理想; 如果 $x \in S, y \in S \Rightarrow x \vee y \in S$, 则子格 S 是一个下部理想。

定义 2 如果 2 个格之间映射保持布尔运算不变, 则称该映射为格同态。

如从格 Ω 到 Ω_2 的格同态等同于格 Ω 划分为一个上部理想和一个下部理想。

定理 4 设 μ 是格 Ω 映射到 Ω_2 的一个同态, 那么格 Ω 上部理想为 $\mu^{-1}(0)$ 和下部理想为 $\mu^{-1}(1)$ 。

若将格 Ω 划分为上部理想 U 和下部理想 L , 对 $s \in U, \mu(s) = 0$ 和 $s \in L, \mu(s) = 1$, 则可以得到 μ 是格 Ω 到 Ω_2 的同态。

定理 5 对于由排序单元多级互连网络构成的 n -to- m 集线器, 用布尔单元替换多级互连网络中所有的排序单元, 所得到一个 n -to- m 布尔集线器网络便称之为布尔集线器。

将任意的分配格 Ω 划分为上部理想 U 和下部理想 L 。从 U 中输入 u 个值, $0 \leq u \leq m$, 从 L 中 $n-u$ 个值进入集线器网络之后, 有如下结论。

结论 3 上部 $n-m$ 个端口输出 $\min\{n-m, u\}$ 属于 U 的值。

结论 4 下部 m 个端口输出 $\min\{m, n-u\}$ 属于 L 的值。

证明 根据定理 5 可知上述的单元为一个 n -to- m 布尔集线器。为了证明定理剩下的部分, 令 μ 表示和 $s \in L, \mu(s) = 1$ 中 Ω 到 Ω_2 的同态。

用 $\mu(s)$ 替换每一个输入信号 s 。这就将 n 个输入信号值转变为 u 个 0 的和 $n-u$ 个 1 的组合。从集线器网络的性质可知, 上部 $n-m$ 端口输出为 $\min\{n-m, u\}$ 个 0 和下部 n 端口输出个 1。因为 μ 是一个格同态, 级间用同样的信号替代。当替代后集线器的一个输出端口输出值 0 或 1 时, 则替代前其输出值分别属于 $\mu^{-1}(0) = U$ 或 $\mu^{-1}(1) = L$ 。证毕。

将定理 5 应用到 $\Omega = \Omega_{\text{bicast}}$, $U = \{0, B\}$, $L = \{I, 1\}$, 那么结论 1 和结论 3 变为完全相同。类似地, $U = \{0, 1\}$ 和 $L = \{B, I\}$ 结论 2 将和结论 4 一致。

定理 6 针对大规模可扩展多播结构中的 n -to- m 集线器, 将所有的排序单元用 *bicast* 单元来替换, 则当信号表是 Ω_{bicast} 时, 以下结论中有一项成立。

结论 5 上面的 $n-m$ 个端口输出仅为 0-bound 信号。

结论 6 下面的 m 个端口输出仅为 1-bound 信号。

结论 7 没有端口输出为 *idle* 信号。

结论 8 没有端口输出为 *bicast* 信号。

证明 信号值 0-bound, 1-bound, *bicast* 和 *idle* 分别为 0, 1, B 和 I 。bicast 单元组成的多级互连网络的精确输入组合为: V_0 个 0, V_1 个 1, V_B 个 B , V_I 个 I 。由图 3(b) 可看到该信号表满足分配格的结构, 若将多级互连网络中的 *bicast* 单元用布尔单元替换, 则布尔集线器定理得以应用, 同时也保证了布尔集线器网络构成及式(1)、式(2)的应用。证毕。

将信号表 Ω_{bicast} 表示的格划分为上部理想 $U = \{0, B\}$ 和下部理想 $L = \{I, 1\}$ 。可以得到由 n 元组成的对称多项式 σ_{m+1} 的如下性质。

性质 1 如果 $V_1 + V_I \geq m + 1$, 可得 σ_{m+1} 的值不是 1 就是 I , 因为 σ_{m+1} 中的一些单项式只与 n 个变量中赋值为 1 和 B 的值的变量有关。

性质 2 如果 $V_1 + V_I \leq m + 1$, 可得 σ_{m+1} 的值不是 0 就是 B , 因为 σ_{m+1} 中的一些单项式只与 n 个变量中赋值为 0 或 B 的值的变量有关。

另一种是, 将理想的分为上部分 $U = \{1, B\}$ 和下部分 $L = \{0, I\}$ 。同理可得以下性质结论。

性质 3 如果 $V_1 + V_B \geq m + 1$, 可得 σ_{m+1} 的值是 1 或 B 。

性质 4 如果 $V_1 + V_B \leq m + 1$, 可得 σ_{m+1} 的值是 0 或 I 。

假设性质 2 和性质 4 成立, σ_{m+1} 唯一可能的值就是 0, 再由式(1)可以得出结论 4 成立。

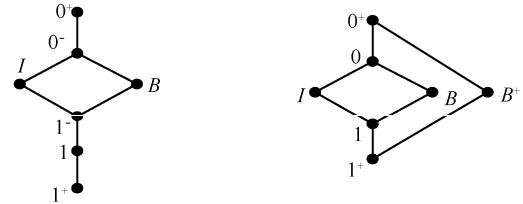
对称地, 假设性质 1 和性质 3 成立, σ_{m+1} 唯一可能的值就是 1, 因此 $\sigma_m = 1$, 再由式(2)可以得出结论 6 成立。

假设性质 2 和性质 3 成立, σ_{m+1} 唯一可能的值就是 B , 因此 $\sigma_m \geq \sigma_{m+1} = B$ 。从式(1)可以得出, 上部分 $n-m$ 个端口输出不可能为 I 。从式(2)可以得出, 下部

分 m 个端口输出也不可能为 I 。所以结论 7 可以成立。

由对称的特性可得, 在性质 1、性质 4 假设下结论 8 也成立。

在定理 6 中, 0-bound 和 *bicast* 信号总数在多级互连网络被逐级保留, 1-bound 和 *bicast* 信号总数亦然。同样可以看出结论 5 到结论 8 的每一个声明都暗示着结论 1 和结论 2 成立。因此定理 6 应该可以说是布尔多播集线器原理的一个详细版本。从定理 6 可以得出一个更通用的结论: 布尔集线器理论不仅仅可以把尽可能多的信号路由到目的输出群组, 而且可以依据“优先级”的不同而达到最佳路由。此处的“优先级”要求集线器的信号表必须是分配格。这个定理适合通过各种可能的方式去划分分配格的上部理想和下部理想。通过以下例子来说明。



(a) 附加优先级多播信号表的分配格 (b) 不可分配

图 5 优先级多播信号表示例

设 $\Omega = \{0^+, 0^-, I, B, 1^-, 1, 1^+\}$ 为在图 5(a) 中的分配格。 Ω 中元素的命名规则为: 上标 ‘+’ 为到达期望集线器的目的地址 0 或 1 的最高优先级, ‘-’ 表示最低优先级。 Ω 中所有通过一个 n -to- m 布尔集线器的路由信号的优先级涉及了所有可能的对 Ω 的划分为上部理想 U 集合和下部的理想 L 集合。这里可以组合

$$\begin{cases} U = \{0^+\}; L = \{0^-, I, B, 1^-, 1, 1^+\} \\ U = \{0^+, 0^-, I\}; L = \{B, 1^-, 1, 1^+\} \\ U = \{0^+, 0^-, B\}; L = \{I, 1^-, 1, 1^+\} \\ U = \{0^+, 0^-, I, B, 1^-\}; L = \{1, 1^+\} \\ U = \{0^+, 0^-, I, B, 1^-, 1\}; L = \{1^+\} \end{cases}$$

从定理 5 可以看出, 在去往输出群组 0 的元素中集合 U 中的每一个元素都比集合 L 中的元素的优先级要高; 同样, 信号路由到输出群组 1 也是如此。应用这个理论到上面所列出的 5 个划分方法可以得到以下结论。

结论 9 在有序子集 $\{0^+, 0^-, B, 1^-, 1, 1^+\}$ 的信号中去往输出群组 0 的较小的元素被赋予较高优先

级，然而去往输出群组 1 的信号被赋予较高的优先级。布尔集线器的路由最佳性是和这个优先级的策略相一致的。

结论 10 类似的优先级处理同样可以应用在有序子集 $\{0^+, 0^-, 1, 1^-, 1, 1^+\}$ 中。

结论 11 同时， B 和 I 这 2 个信号在路由向任意输出群组时被赋予同样的优先级，这就意味着 B 和 I 不能同时出现在彼此对立的输出群组中。

如图 6(a)实例说明分配格 Ω 的信号通过一个 8-to-4 布尔集线器网络的传播过程，其中有上标 ‘+’ 的信号是具有高优先级处理的。0, 1, B , I 分别表示 0-bound, 1-bound, *bicast* 和 *idle*。图中阴影部分单元表示发生多播的信号，在这里 *bicast* 和 *idle* 信号被 0-bound 和 1-bound 信号所替换。这样输出群组 0 得到尽可能多的 0-bound 和 *bicast* 信号，同时输出群组 1 得到尽可能的 1-bound 和 *bicast* 信号。此外，这个路由最优是和结论 9 和结论 11 相一致的。

例如，图 6(b)例证了图 5(b)中非分配格 $\Omega = \{0^+, 0^-, 1, B, 1^-, 1, 1^+\}$ 中信号值通过同一个 8-to-3 布尔集线器网络的传播过程，其理想的交换是得不到保障的。因为在这个格中 $B^+ \wedge I = 0^+$ 和 $B^+ \vee I = 1^+$ ，一个高优先级的 *bicast* 信号在一个布尔单元中遇到一个 *idle* 信号后输出高优先级 0-bound 和 1-bound 信号，从而导致输出群组 0 丢失了一个有效信号，同时也只有一个有效信号出现在输出群组 0。这个次优的路由结果反应了在定义布尔交换，集线器或排序器中输入信号需要满足分配格的结构才能达到最优的多播交换。

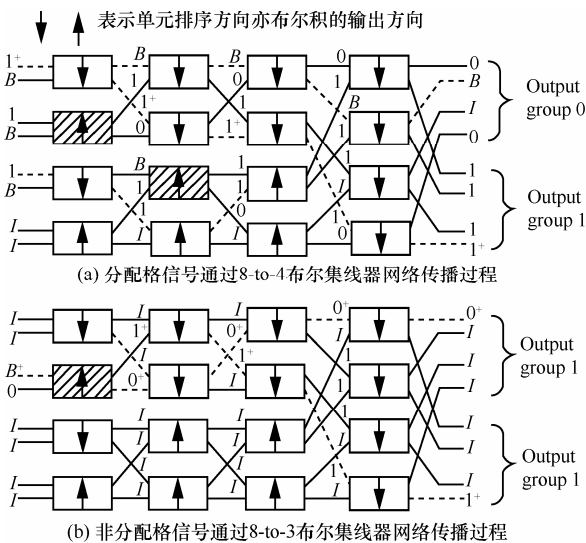


图 6 信号通过布尔集线器网络传播过程

在实际应用中非分配格有时会被用作信号表，因为高优先级多播通信在整个通信量中占很小一部分，那么布尔集线器理论依旧能够被统计应用，这是一个特殊情况。

4 线速多播结构阻塞率模型分析

从上述布尔多播单元 *bicast* 的状态可以看出，该多播交换结构的特点是在交换网络中线速扇出。当扇出为 1 时即是单播，否则就是多播，所以该交换多播交换结构支持单多播的混合输入。分组的扇出率（多播扇出的个数）直接影响了其阻塞率。由于交换的规模是由其级数决定的，交换规模越大，多播分组扇出的可能性也就越大，交换的级数对阻塞率和扇出率也有直接的影响。

单播时自路由路径 $\text{trace}(\text{guide})^{[15]}$ 信息表示为 $d_{\phi(g)} \cdots d_{\phi(2)} d_{\phi(1)} 1$ ，其中，下角标 $\phi(1), \phi(2), \dots, \phi(g)$ 表示级间置换信息^[13]。多播分组头信息序列为 $\cdots Q_{110} Q_{010} Q_{100} Q_{000} Q_{01} Q_{10} Q_{00} Q_1 Q_0 1$ ，多播分组头信息的第一位 1 为多播有效位，此分组序列包含了该多播中所有单播的信息，从右向左依次排序，可以分解为多个单播分组的信息头，并且都满足单播的级间置换规则。

不妨设第 $i(1 \leq i \leq m)$ 级布尔集线器内单播分组控制信令为 $d_{\phi(i)}$ ，多播分组控制信令为 $Q_{q_i} 1$ ， q_i 是二进制数，其长度为 $i-1$ 。每当空分组 I 与多播分组 B 在同一个 2×2 的布尔单元相遇时，若该级满足多播扇出控制要求，其两端口的输出控制信息则采用 *cut-through*^[15] 的方法将多播信息头进行分割，否则多播头信息不发生变化。

下面具体分析线速多播交换结构在不同网络负载下的分组阻塞率情况。假定各输入群组所到达的分组满足独立同分布的要求，在某一时刻到达交换结构中第 i 级 $2G$ -to- G 布尔集线器中某交换单元某输入/输出群组的分组数目表示为 $X_{I_i} / X_{O_i} (1 \leq i \leq m)$ ，其中， $0 \leq X_{I_i} \leq G, 0 \leq X_{O_i} \leq G$ 。

设交换结构到达的业务数据总额为 1，单播业务所占比例参数 $\lambda = P_1$ ，表示在单个时隙内单播业务到达的概率为 P_1 ，多播业务所占比例参数为 $\lambda = P_2$ ，表示单个时隙里多播业务到达的概率 P_2 ，用 X_{I_1} 表示进入交换结构的第一级中某输入群组的单播分组个数，则

$$P(X_{I_{1s}} = x_{1s}) = \binom{G}{x_{1s}} P_1^{x_{1s}} (1 - P_1)^{G - x_{1s}} \quad (3)$$

其中, $X_{I_{1m}}$ 表示进入交换结构的第一级某输入群组的多播分组个数。

$$P(X_{I_{1m}} = x_{1m}) = \binom{G}{x_{1m}} P_1^{x_{1m}} (1 - P_1)^{G - x_{1m}} \quad (4)$$

$P = P_1 + P_2, x_1 = x_s + x_m$, 其中, x_s 表示单播分组个数, x_m 表示多播分组个数。如果 $x_m \neq 0$, 表示存在多播分组, 由于多播分组扇出的特点, 那么在交换结构中的某集线器最后输出的分组个数在一定概率上要比输入的分组多一些。假设每个多播分组在交换结构的每一级都要求发生多播, 即其带内多播控制信号全是由 $2^k - 1$ 个 B 构成, 由于布尔集线器的输入信号表为严格的分配格结构, 即当单播与多播在集线器内争用同一个出口时, 根据交换规则单播优先级高而占用输出端口, 导致其多播会被误路由, 所以该多播只能以一定的概率扇出。 X_{I_1}, Y_{I_1} 为交换结构中第一级集线器的输入, 满足独立同分布要求。用 A_{I_1} 表示第一级布尔集线器输入分组的总和。为了简化表达式用 $D_I(d_i)$ 表示 $D_I = d_i$ 时的概率 $P(D_I = d_i)$ 。

$$A_{I_1}(a_1) = \sum_{x_{1s} + x_{1m} + y_{1s} + y_{1m} = a_1} X_{I_{1s}}(x_{1s}) X_{I_{1m}}(x_{1m}) Y_{I_{1s}}(y_{1s}) Y_{I_{1m}}(y_{1m}) \quad (5)$$

$\forall a_1 \leq 2G, \exists I \wedge B = 0, I \vee B = 1$, 使 $a_1 + x_{1m_2} + y_{1m_2} \leq 2G$, 其中, $x_{1m} = x_{1m_1} + x_{1m_2}, y_{1m} = y_{1m_1} + y_{1m_2}$, x_{1m_1}, y_{1m_1} 为扇出为 1 的多播分组 (即为单播分组), x_{1m_2}, y_{1m_2} 为扇出 2 的多播分组。从而可得实际的输出分组总和 a'_1

$$A_{O_1}(a'_1) = \begin{cases} \sum_{x_{1s} + 2x_{1m} + y_{1s} + 2y_{1m} = a'_1} X_{I_{1s}}(x_{1s}) X_{I_{1m}}(x_{1m}) Y_{I_{1s}}(y_{1s}) Y_{I_{1m}}(y_{1m}), a'_1 < 2G \\ \sum_{x_{1s} + y_{1s} = 0}^{2G} (X_{I_{1s}}(x_{1s}) Y_{I_{1s}}(y_{1s})) \sum_{2x_{1m} + 2y_{1m} = a'_1 - x_{1s} - y_{1s}}^{2G} X_{I_{1m}}(x_{1m}) Y_{I_{1m}}(y_{1m}), a'_1 = 2G \end{cases} \quad (6)$$

因为多播信令全是有 B 成的, 所以若 $a'_1 < 2G$, 则表示输出端口依然有空余, 即多播分组已经完全被扇出; 若 $a'_1 = 2G$ 时, 表示所有的输出端口都已被使用, 那么就存在部分多播分组可能没有被扇出的情

况。若在某输出群组有 $z'_1 < G$ 的分组到达, 那么到达该集线器输入群组的分组不少于 $z'_1 < a'_1 < 2G$, 且都被正确路由, 若到达某输出群组的分组为 $z'_1 = G$, 那么有 $a'_1 = G, a'_1 = G + 1, \dots, a'_1 = 2G$ 个分组出现竞争的情况, 即导致分组阻塞。交换结构中第一级布尔集线器的某输出群组单播和多播分组总和 z'_1 分布为

$$X_{O_1}(z'_1) = \begin{cases} \sum_{a'_1 = z'_1}^{2G} A_{I_1}(a'_1) 2^{-a'_1} \binom{a'_1}{z'_1}, z'_1 < G \\ \sum_{a'_1 = z'_1}^{2G} A_{I_1}(a'_1) 2^{-a'_1} \sum_{k=G}^{a'_1} \binom{a'_1}{k}, z'_1 = G \end{cases} \quad (7)$$

那么在第一级布尔集线器的输出中单播分组所占的概率分布 $X_{O_{1s}}(x'_{1s})$ 为

$$X_{O_{1s}}(x'_{1s}) = \begin{cases} \sum_{a'_1 = x'_{1s}}^{2G-1} \lambda_{1s} A_{I_1}(a'_1) 2^{-a'_1} \binom{a'_1}{x'_{1s}} + \sum_{a'_1 = 2G} \lambda'_{1s} A_{I_1}(a'_1) 2^{-a'_1} \binom{a'_1}{x'_{1s}}, x'_{1s} < G \\ \sum_{a'_1 = x'_{1s}}^{2G-1} \lambda_{1s} A_{I_1}(a'_1) 2^{-a'_1} \sum_{k=G}^{a'_1} \binom{a'_1}{k} + \sum_{a'_1 = 2G} \lambda'_{1s} A_{I_1}(a'_1) 2^{-a'_1} \sum_{k=G}^{a'_1} \binom{a'_1}{k}, x'_{1s} = G \end{cases} \quad (8)$$

$$\text{其中, } \lambda_{1s} = \frac{X_{I_{1s}}(x'_{1s})}{X_{I_{1s}}(x'_{1s}) + X_{I_{1m}}\left(\frac{1}{2}(a'_1 - x'_{1s})\right)},$$

$$\lambda'_{1s} = \frac{X_{I_{1s}}(x'_{1s})}{X_{I_{1s}}(x'_{1s}) + X_{I_{1m}}\left(\frac{1}{2}(G - x'_{1s})\right)}$$

多播分组所占的概率 $X_{O_{1m}}(x'_{1m})$ 为

$$X_{O_{1m}}(x'_{1m}) = \begin{cases} \sum_{a'_1 = x'_{1m}}^{2G-1} \lambda_{1m} A_{I_1}(a'_1) 2^{-a'_1} \binom{a'_1}{x'_{1m}} + \sum_{a'_1 = 2G} \lambda'_{1m} A_{I_1}(a'_1) 2^{-a'_1} \binom{a'_1}{x'_{1m}}, x'_{1m} < G \\ \sum_{a'_1 = x'_{1m}}^{2G-1} \lambda_{1m} A_{I_1}(a'_1) 2^{-a'_1} \sum_{k=G}^{a'_1} \binom{a'_1}{k} + \sum_{a'_1 = 2G} \lambda'_{1m} A_{I_1}(a'_1) 2^{-a'_1} \sum_{k=G}^{a'_1} \binom{a'_1}{k}, x'_{1m} = G \end{cases} \quad (9)$$

其中, $\lambda_{1m} = 1 - \lambda_{1s}, \lambda'_{1m} = 1 - \lambda'_{1s}$, 已知多播交换结构的第 i 级集线器的输出是第 $i+1$ 级集线器的输入, 即 $X_{O_{1s}}, X_{O_{1m}}, X_{O_i}$ 与 $X_{I_{(i+1)s}}, X_{I_{(i+1)m}}, X_{I_{(i+1)}}$ 满足相同分布, 由此可得集线器实际输入的分组总数 a_{i+1} 服

从以下分布。

$$A_{I_1}(a_{i+1}) = \sum_{x_{(i+1)s} + x_{(i+1)m} + y_{(i+1)s} + y_{(i+1)m} = a_{i+1}} X_{I_{(i+1)s}}(x_{(i+1)s}) \cdot X_{I_{(i+1)m}}(x_{(i+1)m}) Y_{I_{(i+1)s}}(y_{(i+1)s}) Y_{I_{(i+1)m}}(y_{(i+1)m}) \quad (10)$$

集线器实际输出的分组总数 a'_{i+1} 服从以下分布。

$$A_{O_1}(a'_{i+1}) = \begin{cases} \sum_{x_{(i+1)s} + 2x_{(i+1)m} + y_{(i+1)s} + 2y_{(i+1)m} = a'_{i+1}} X_{I_{(i+1)s}}(x_{(i+1)s}) \cdot X_{I_{(i+1)m}}(x_{(i+1)m}) Y_{I_{(i+1)s}}(y_{(i+1)s}) Y_{I_{(i+1)m}}(y_{(i+1)m}), \\ a'_{i+1} < 2G \\ \sum_{x_{(i+1)s} + y_{(i+1)s} = 0}^{2G} (X_{I_{(i+1)s}}(x_{(i+1)s}) Y_{I_{(i+1)s}}(y_{(i+1)s})), \\ a'_{i+1} = 2G \\ \sum_{2x_{(i+1)m} + 2y_{(i+1)m} = a'_{i+1} - x_{(i+1)s} - y_{(i+1)s}}^{2G} X_{I_{1m}}(x_{1m}) Y_{I_{1m}}(y_{1m}), \\ a'_{i+1} = 2G \end{cases} \quad (11)$$

同样可得出第 $i+1$ 级布尔集线器的某输出群组单多播分组总数 z'_{i+1} 的分布：

$$X_{O_{i+1}}(z'_{i+1}) = \begin{cases} \sum_{a'_i = z'_{i+1}}^{2G} A_{I_1}(a'_{i+1}) 2^{-a'_{i+1}} \binom{a'_{i+1}}{z'_{i+1}}, z'_{i+1} < G \\ \sum_{a'_{i+1} = z'_{i+1}}^{2G} A_{I_1}(a'_{i+1}) 2^{-a'_{i+1}} \sum_{k=G}^{a'_{i+1}} \binom{a'_{i+1}}{k}, z'_{i+1} < G \end{cases} \quad (12)$$

重复上述步骤可以推导出第 $i+1$ 级布尔集线器输出的单播分组个数 $X_{O_{(i+1)s}}$ 分布，以及多播分组个数 $X_{O_{(i+1)m}}$ 分布。通过的迭代的方法，可得第 k 级布尔集线器某输出群组单多播分组个数的分布情况 $X_{O_{ks}}$, $X_{O_{km}}$ 。

单播阻塞率指平均输入单播分组减去输出单播分组数的平均值后再与平均输入单播分组数的比，得式(13)。

$$P_r = 1 - \frac{E(X_{O_{ks}})}{Gp_1} = 1 - \frac{\sum_{x'_{ks}=0}^G x'_{ks} X_{O_{ks}}(x'_{ks})}{Gp_1} \quad (13)$$

多播的阻塞率要考虑扇出的因素，多播的输出分组会远大于输入分组，所以多播的阻塞率应该为平均输出多播分组数与理想最大输出多播分组数的比，得式(14)。其中， $(k+1)Gp_2 \bmod(G-Gp_1)$ 的含义为：因带内控制信令由全 B 组成，所以理想最

大可扇出分组数可表示为 $(k+1)Gp_2$ ，但受到输出端口数 G 的限制和单播输入分组数 Gp_1 的与多播分组争用端口的影响，最大可输出多播分组数表示为 $(k+1)Gp_2 \bmod(G-Gp_1)$ 。

$$P_r = 1 - \frac{E(X_{O_{km}})}{(k+1)Gp_2 \bmod(G-Gp_1)} = 1 - \frac{\sum_{x'_{km}=0}^G x'_{km} X_{O_{km}}(x'_{km})}{(k+1)Gp_2 \bmod(G-Gp_1)} \quad (14)$$

5 仿真结果

本节采用 C++ 编程实现在均匀分布业务源下对线速多播交换结构的多播阻塞率和时延的仿真。阻塞率是指分组在经过交换系统时没能被成功交换到目的端口的概率。时延是指分组第一个比特进入交换系统到分组第一个比特从交换系统输出时所需要的时间。时延在交换系统中一般分为传播时延和信号处理时延。仿真过程要求数据源满足伯努利均匀分布，集线器的群组输入业务要求服从二项分布，到达的业务强度用归一化负载表示。

1) 多播阻塞率仿真

负载强度对阻塞率有直接的影响，本节选取一个代表性的负载，不妨设单播业务所占负载的比例为 $P_1 = 0.2$ ，则可得多播业务所占负载的比例为 $0 \leq P_2 \leq 1 - P_1$ 。在多播业务强度相对固定的情况下，通过改变交换网络的级数参数 K 以及群组参数 G 的大小仿真多播阻塞率的情况。选取在 $K=4$ 的条件下仿真并比较 $G=8, 16, 32$ 时多播的阻塞率情况。从图 7 可以看出，随着 G 增大其阻塞率而减小。因为参数 K 决定了交换结构网络的级数，当 K 不变且单播业务负载强度相同时，则单播与多播在集线器中发生争用的情况基本相同。然而若 G 变大，分组在布尔集线器内部发生争用的概率就变小，阻塞率也就越小。

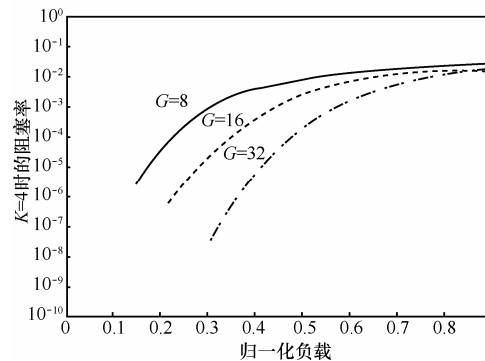


图 7 G 不同时的多播阻塞率

图 8 的显示的是群组 $G=32$ 时, $K=4、8、16$ 时得到的多播阻塞率仿真结果。从图中曲线可看出 K 变大其阻塞率也变大, 然而在同一负载点上, K 的不同, 其阻塞率区别不明显。这是因为由于多播的扇出特性, 当交换结构网络级数不小于 4, 归一化负载 $P_2 \geq 0.2$ 时, 得到扇出的多播分组数接近 $G - 0.2G = 0.8G$ 所以在最后的几级布尔多播集线器中多播分组被扇出的概率会很小, 所以在 $G=32$ 归一化负载下, 虽然 k 不同, 但其阻塞率差别并不明显。

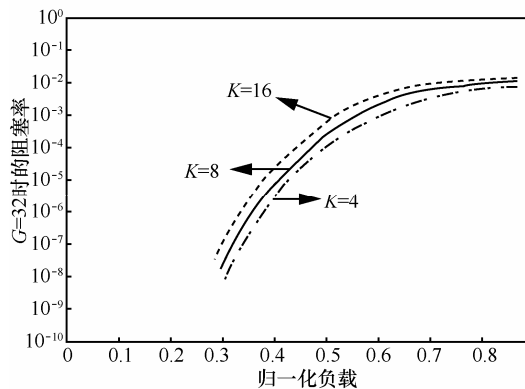


图 8 K 不同时的多播阻塞率

2) 多播分组时延仿真

多播分组在经过交换结构网络的每一个布尔集线器中的每一个 2×2 布尔单元时, 需要一定的时间完成对分组数据的处理, 从而产生了时延。所以多播分组的时延主要由分组所经过基本布尔单元个数决定, 设分组在经过任一个布尔单元时对数据处理的时间即时延为 Δd 。

多播交换结构网络的规模由 k 决定, 每一个布尔集线器的规模则是 G 决定, 交换结构中布尔单元的个数则是由 K 和 G 共同决定。为了仿真多播时延数据, 不妨先固定一个参数, 设 $K=4$, 可以得到 $G=8, 16, 32$ 的时延, 如图 9 所示。

从图 9 中可看出 G 越小时延也越小。从图中还可以看出 G 不同时, 各时延增长速度 T_v 的关系为 $T_{v(G=32)} < T_{v(G=16)} < T_{v(G=8)}$, 这是因为 $G=32$ 的阻塞率是最小的, 所以分组丢失的概率较小, 统计分组时延也就会最低。随着负载的增大, 当负载大于 0.6 时, 时延增加的速度加快。

固定参数 $G=32$ 时, 得到 $K=4、8、12$ 的时延仿真结果, 如图 10 所示, K 越小其时延越小, K 不同。从图 8 可以得知, K 不同, 其阻塞率的变化不明显, 所以该时延曲线的变化也不明显。该多播交

换结构采用的多级互联的结构, 无论 G, K 参数如何变化, 其时延也总会小于某一个定值。

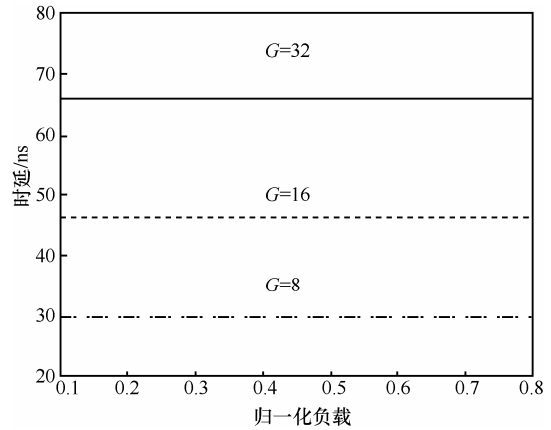


图 9 $K=4, G$ 不同时的多播时延

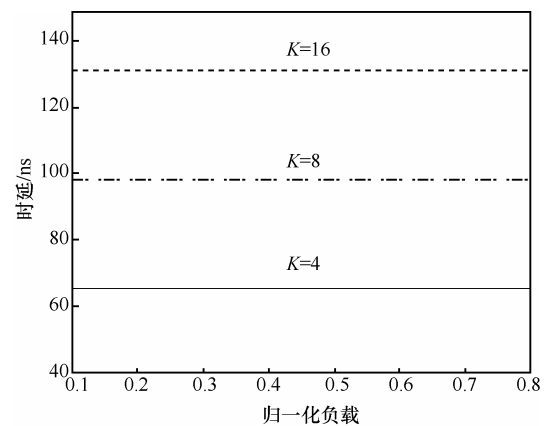


图 10 $G=32, K$ 不同时的多播时延

6 结束语

本文提出的基于分配格的线速多播交换模型, 具有低时延、无抖动、可大规模扩展优点, 当多播业务负载强度控制在一定可允许的范围内时, 其多播业务的阻塞率可以被接受, 其多播时延总会小于某个定值, 且该定值会在百纳秒量级, 所以该交换结构能够为到达业务提供时延上限保障。该交换结构模型可以解决某些网络环境中对业务的特殊需求应用的问题, 如时延、阻塞率、线速多播等。

参考文献:

[1] HUANG A. Starlite: a wideband digital switch[A]. Proc Globecom'84[C]. Atlanta, GA, 1984.1-5.
 [2] SAITO H, YAMANAKA H, YAMADA H, et al. Multicast function and its LSI implementation in a shared multibuffer ATM switch[A].

- Networking for Global Communications, 13th Proceedings IEEE[C]. Toronto, Ont, 1994.315-322.
- [3] YEH Y S, HLUCHYJ M, ACAMPORA A. The knockout switch: a simple, modular architecture for high-performance packet switching[J]. Selected Areas in Communications, 1987, 5(8):1274-1283.
- [4] CHAO H J, CHOE B S, PARK J S, *et al.* Design and implementation of abacus switch: a scalable multicast ATM switch[J]. Selected Areas in Communications, 1997,15(5):830-843.
- [5] LEE T T. Nonblocking copy networks for multicast packet switching[J]. Selected Areas in Communications, 1988, 6(9): 1455-1467.
- [6] CHAO H J, LIU B. High performance switches and routers[M]. Wiley-IEEE Press, 2007.
- [7] MARCHOK D J, ROHRS C E, SCHAFER R M. Multicasting in a growable packet (ATM) switch[A]. INFOCOM'91. Proceedings. Tenth Annual Joint Conference of the IEEE Computer and Communications Societies Networking [C]. BalHarbour, FL, 1991. 850-858.
- [8] JASTRZEBSKI A, KUBALE M. Rearrangeability in multicast clos networks is np-complete[A]. Information Technology(ICJT)[C]. Gdansk, 2010, 183-186.
- [9] HUI J Y, RENNER T. Queueing analysis for multicast packet switching[J]. IEEE Transaction Communications, 1994, 42(234): 723-731.
- [10] ALI M K M, YANG S. Performance analysis of a random packet selection policy for multicast switching[J]. IEEE Transactions Communications, 1996, 44(3):388-398.
- [11] MARSAN M A, BIANCO A, GIACCONE P, *et al.* Multicast traffic in input-queued switches: optimal scheduling and maximum throughput[J]. IEEE/ACM Transactions on Networking (TON), 2003, 11(3): 465-477.
- [12] PAN D, YANG Y. FIFO-based multicast scheduling algorithm for virtual output queued packet switches[J]. Computers, IEEE Transactions, 2005, 54(10):1283-1297.
- [13] HE W, LI H, WANG B, *et al.* Load-balanced multipath self-routing switching structure by concentrators[A]. IEEE International Conference on Communications, ICC'08[C]. Beijing, China,2008.5935-5939.
- [14] LI H, LI S Y R. Layout complexity of bit-permuting exchange in multi-stage interconnection networks[M]. Boston: Kluwer Academic Publishers, 2001.259-276.
- [15] LI S Y R. Algebraic switching theory and broadband applications[M]. Academic Press, 2001.
- [16] LI S Y R. Unified algebraic theory of sorting, routing, multicasting, and concentration networks[J]. IEEE Transactions Communications, 2010, 58(1):247-256.
- [17] LI S Y R, LI H, KOO G M. Fast knockout algorithm for self-route concentration[J]. Computer Communications, 1999, 22(17):1574-1584.
- [18] LI H, LI S Y R. On the Complexity of Concentrators and Multi-stage Interconnection Networks in Switching System[D]. Hong Kong: Chinese University of Hong Kong, 2011.

作者简介:



尘福兴(1986-), 男, 山东单县人, 北京大学博士生, 主要研究方向为网路编码、代数交换、路由交换结构及下一代互联网。



李挥[通信作者](1964-), 男, 广东汕头人, 博士, 北京大学深圳研究生院教授、博士生导师, 主要研究方向为三网合一、流媒体云计算网络视频关键技术、下一代网络体系结构、网络路由和宽带交换结构、网络编码理论及其应用、嵌入式系统开发。E-mail: lih64@pkusz.edu.cn。



崔凯(1985-), 男, 山西晋中人, 北京大学硕士生, 主要研究方向为路由和宽带交换结构、网络编码、嵌入式系统。



张博(1982-), 男, 河北张北人, 国家数字交换系统工程技术研究中心博士生, 主要研究方向为可重构路由理论与技术、可重构交换理论与技术。