On Nonblocking Multicast Three-Stage Clos Networks

Frank K. Hwang and Sheng-Chyang Liaw

Abstract—We give necessary and sufficient conditions for strictly nonblocking three-stage Clos networks with two types of multicast traffic. Then, we extend the results to other models where the routing is under certain restrictions. Finally, we compare our results with existing literature to mark the differences and similarities.

Index Terms—Multicast traffic, strictly nonblocking, switching networks, wide-sense nonblocking.

I. INTRODUCTION

MASSON and Jordan [6] first considered nonblocking networks for multicast, or f_2-cast, traffic in which a request consists of a simple idle inlet and f idle outlets where $1 \leq f \leq f_2$ and f_2 is a given upper bound. Actually there are two different types of multicast traffic. In the *bundled* type, all f outlets of the request are specified simultaneously. A possible scenario is a conference call. In the *incremental* type, the f outlets can join the request one by one. A possible scenario is the video program ordering service. For the second type, we should redefine a multicast request as a pair of inlet, outlet such that the outlet must be idle while the inlet can be already connected to up to $f_2 - 1$ outlets. Although most of the literature do not mention this difference, one can usually detect which type they have in mind by inspecting the proofs. We highlight this difference as we will show they can cause not only quantitative but also qualitative differences.

In this paper we will only be concerned with the three-stage Clos network $[X_{n_1 m}, X_{r_1 r_2}, X_{m n_2}]$ (in Cantor's notation [1]) where the input stage consists of r_1 crossbars of size $n_1 \times m$, the center stage m crossbars of size $r_1 \times r_2$, the output stage r_2 crossbars of size $m \times n_2$, and the connection of the crossbars between the input (output) stage and the center stage is a complete bipartite graph. Fig. 1 shows $[X_{24}, X_{42}, X_{43}]$.

A *state* of a network is a routing of a set of requests. A state is a *blocking* state if there exists a new request which cannot be routed. A network is *strictly nonblocking* (SNB) if it has no blocking state; it is *wide-sense nonblocking* (WSNB) if it can avoid getting into a blocking state, usually by following a routing algorithm.
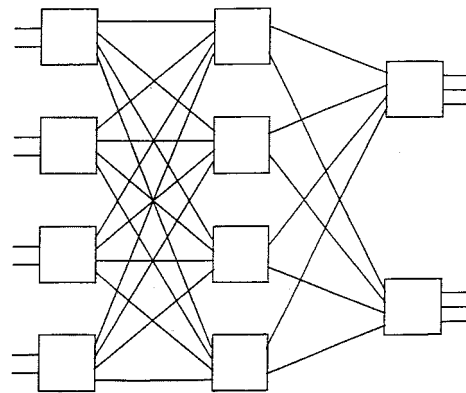
Fig. 1. Three-stage Clos network.

Suppose a request contains two outlets of the same output switch. Under the *no-split* assumption, the two paths to these two outlets cannot split before the output stage. This is a very natural assumption and was widely adopted in early literature. Masson and Jordan [6] gave a sufficient condition on m such that the network is SNB with incremental multicast traffic under the no-split assumption. As the no-split assumption can be considered as a rule of a routing algorithm, their result is now classified as WSNB instead of SNB. Recently, Giacomazzi and Trecordi [3] gave necessary and sufficient conditions for an SNB network with bundled multicast traffic. In this paper we give different necessary and sufficient conditions (under the same assumption) which are also more explicit and tighter than the conditions in [3]. We will compare these two sets of conditions in a later section. We also give analogous results for the incremental multicast traffic and under various assumptions on the fan-out capability of crossbars in a certain stage, as well as on no-split. In the following text we assume that the crossbar has fan-out capacity except otherwise specified.

II. MAIN RESULTS

We first show a relation between bundled traffic and incremental traffic.

Lemma 1: Let m be the number of center switches for a three-stage Clos network. If m is not enough to guarantee SNB for bundled traffic, it is not enough for incremental traffic.

Proof: Assume that m center switches are not enough to guarantee SNB for bundled traffic, but enough for incremental traffic. We can simulate a sequence of bundled traffic with a sequence of incremental traffic by turning an f-request into f 1-requests. Then the routing of the incremental sequence, which exists by our assumption, constitutes a routing of the bundled sequence. Hence the bundled sequence is not blocked, a contradiction to our assumption. ∎

An intuitive way suggested by a reviewer to interpret Lemma 1 is that the bundled traffic gets all information about a multicast request at one time, and hence should be easier to route.

Let $N_1 = n_1 r_1$ and $N_2 = n_2 r_2$ denote the number of network inlets and outlets. An f-request is a request from an inlet to f outlets, $1 \leq f \leq f_2$.

Theorem 2: A three-stage Clos network is SNB for bundled f_2-cast traffic if and only if the number of center switches is at least $m^* = \min\{(N_1 - 1)f_2 + 1, (n_1 - 1)f_2 + n_2, N_2\}$.

    Proof:

1) "if." Since $(N_1-1)f_2+1$ and N_2 are trivially sufficient for nonblocking, we show that $(n_1 - 1)f_2 + n_2$ is also enough. Consider an arbitrary f-request, say, from an inlet x lying on an input switch I to f outlets lying on t output switches, say, $O_1, O_2, \ldots, O_t$. Call the other (n_1-1) inlets lying on I *co-inlets*, and the other outlets lying on O_i *co-outlets*, $1 \leq i \leq t$. Clearly, the connection from x to O_i (for a fixed i) can be blocked only by existing connections involving either the co-inlets, or the co-outlets on O_i. The co-inlets can take at most $(n_1 - 1)f_2$ center switches. If we have $(n_1 - 1)f_2 + n_2$ center switches, then the remaining n_2 center switches can be used to connect x. For each O_i, the co-outlets of O_i can take at most $n_2 - 1$ center switches, thus x can always find one center switch to connect with O_i. So, $(n_1 - 1)f_2 + n_2$ is enough.

2) "only if." We construct a blocking state for $m^* - 1$ center switches with a new 1-request from an inlet lying on an input switch I to an outlet lying on an output switch O. Since SNB means nonblocking under arbitrary routing, we may assume that the inlets which have been routed do not share the center switches in the worst-case scenario. Let $x = \min\{(N_1 - n_1)f_2, n_2 - 1\}$. Then, the other multicasts originated from inlets on the other input switches and destined to O can take as many as x center switches. If in the meantime the $(n_1 - 1)$ co-inlets take $\min\{(n_1 - 1)f_2, N_2 - 1 - x\}$ center switches on way to some output switches other than O, then no accessible center switch is left for the new request. The second term $N_2 - 1 - x$ says that since there can be at most $N_2 - 1$ existing requests, the number of center switches taken by co-inlets is at most $N_2 - 1 - x$ (x already taken by other input switches).■

For incremental multicast traffic, an f-request is treated as f 1-requests. So the current request is always a 1-request while the engaged inlet is already connected to $f - 1$ distinct outlets in the network.

Corollary 3: Theorem 2 remains true if the f_2-cast traffic is incremental.

    Proof: By Lemma 1, it suffices to prove that m^* is enough to guarantee SNB for incremental traffic. But the argument given in part 1) of the Proof of Theorem 2 is independent of the type of traffic, hence Corollary 3.   ■

III. SOME OTHER MODELS

In ref. [6] the case was considered that input switches have no fan-out capacity and argued that the nonblocking result obtained
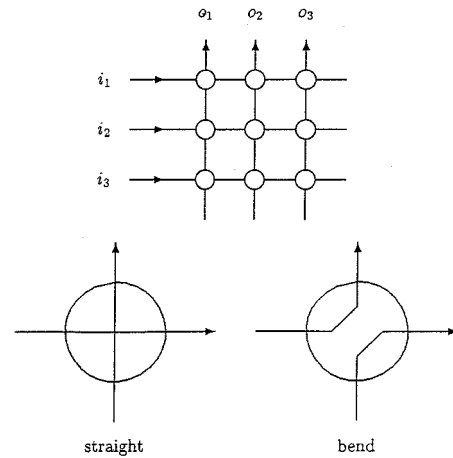

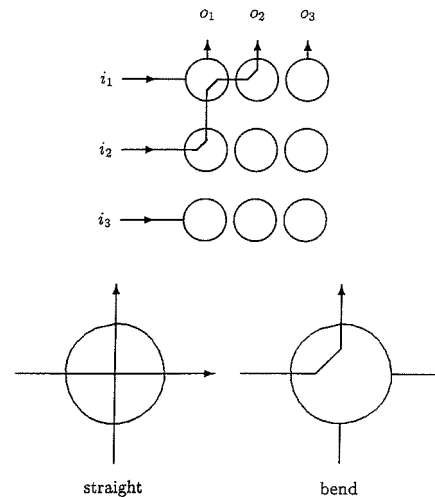Fig. 2.   Crossbar and crosspoint.

Fig. 3.   New crosspoint guaranteeing SNB for crossbar.

under this assumption should be classified as WSNB rather than SNB. We think a difference should be made here whether the incapacity is due to hardware handicap or software (meaning a routing algorithm). If it is the former, then the result should still be classified as SNB because the blocking states do not exist in the first place. Even if we are not sure of the source of incapacity, or do not want to get into that level of details, we may still assume it is the hardware as long as we know the hardware can be so implemented. This view is consistent with the general acceptance of a crossbar switch as SNB. Typically, a crossbar is represented by a grid of input rows and output columns with a crosspoint at each intersection, and the crosspoint is assumed to have two states "straight" and "bend" (see Fig. 2).

However, such a crossbar is not SNB since the connection from i_2 to o_2 blocks the request from i_1 to o_3, or i_3 to o_1 (as far as we know, this observation has not been made before). Note that this blocking can be removed either by wiring the crosspoint differently (see Fig. 3) or by the rule of a routing algorithm to disallow a path making a right turn, both having the effect of avoiding the $(i_2, 0_2)$ path as shown in Fig. 3. (Sometimes the two states of a crossbar are called "open" and "closed," and for each column only one crosspoint is allowed to be closed. Again, the latter statement represents an algorithm rule.)

Nevertheless, the literature continues to assume that crossbars are SNB, since the implementation is there. Inheriting this philosophy, we will treat the nonblocking results obtained under the assumption of no fan-out capacity as SNB. The assumption of no-split is a different issue, since no-split is clearly an algorithm rule. Granted, it is conceptually possible that someone is ingenious enough, or stupid enough, to hardwire no-split. But the burden is on that person to specify the no-split as hardwired before claiming SNB.

Lemma 4: Assuming the input switch has no fan-out capacity, then an f-request cannot be blocked in bundled f_2-cast traffic if and only if the number of center switches is at least

$$m_f^* = \begin{cases} \min\{N_1, n_1 + f(n_2 - 1)\}, \\ \quad \text{for } (n_1 - 1) + \min\{N_1 - n_1, \\ \quad f(n_2 - 1)\} \le N_2 - f; \\ N_2 - f + 1, \\ \quad \text{for } (n_1 - 1) + \min\{N_1 - n_1, \\ \quad f(n_2 - 1)\} \ge N_2 - f. \end{cases}$$

Proof:

1) "if." Since N_1 and $N_2 - f + 1$ are trivially sufficient for nonblocking, we show that $n_1 + f(n_2 - 1)$ is also enough. In any time, an f-request is from an inlet x lying on an input switch I on way to f outlets lying on t output switches, say, $O_1, O_2, \cdots, O_t, 1 \le t \le f$. Since the input switch has no fan-out capacity, every busy inlet must have one and only one center switch to route its request. The co-inlets can take at most $(n_1 - 1)$ center switches. For each O_i, the co-outlets of O_i can take at most $n_2 - 1$ center switches and t such output switches take at most $t(n_2 - 1)$ center switches. So, there must be one center switch that is not taken by the co-inlets and co-outlets in $n_1 + f(n_2 - 1)$ center switches. Then the f-request of x can be routed.

2) "only if." We construct a blocking state for $m_f^* - 1$ center switches with a new f-request from an inlet lying on an input switch I to f outlets lying on f output switches, say, $O_1, O_2, \cdots, O_f$. Let $x = \min\{N_1 - n_1, f(n_2 - 1)\}$. Namely, the inlets lying on the other input switches can support at least x requests. The blocking state is that $\min\{n_1 - 1, N_2 - f - x\}$ co-inlets take that many center switches on way to some outlets and the other inlets take another x center switches on way to the co-outlets. ∎

Theorem 5: Assuming the input switch has no fan-out capacity, then for an SNB three-stage Clos network with bundled f_2-cast traffic

$$m^* = \begin{cases} \min\left\{N_1, N_2 - \left\lceil \dfrac{N_2 - n_1 + 1}{n_2} \right\rceil + 1\right\}, \\ \quad \text{for } \left\lceil \dfrac{N_2 - n_1 + 1}{n_2} \right\rceil \le f_2; \\ \min\{N_1, n_1 + f_2(n_2 - 1)\}, \\ \quad \text{for } \left\lceil \dfrac{N_2 - n_1 + 1}{n_2} \right\rceil > f_2. \end{cases}$$
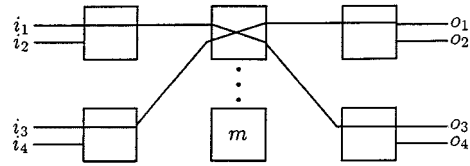

Fig. 4.   Blocking network.

Proof: Let $x = \lceil N_2 - n_1 + 1/n_2 \rceil$. Then $(n_1 - 1) + (x - 1)(n_2 - 1) \le N_2 - x \le (n_1 - 1) + x(n_2 - 1)$. Consider an f-request which cannot be blocked in bundled f_2-cast traffic. We claim that $m^* \ge m_f^*$ and there exists an f'-request such that $m_{f'}^* = m^*$.

Case 1. $f \ge x$. Then $(n_1 - 1) + f(n_2 - 1) \ge (n_1 - 1) + x(n_2 - 1) \ge N_2 - x \ge N_2 - f$ and $m^* = \min\{N_1, N_2 - x + 1\} = m_x^*$. If $N_1 - 1 \ge N_2 - f$, then by Lemma 4 $m_f^* = N_2 - f + 1 \le N_2 - x + 1 = m_x^* = m^*$. If $N_1 - 1 < N_2 - f$, then by Lemma 4 $m_f^* = N_1 < N_2 - f + 1 \le N_2 - x + 1$, that is $m^* = N_1 = m_f^*$.

Case 2. $f < x$. Then $(n_1 - 1) + f(n_2 - 1) \le (n_1 - 1) + (x - 1)(n_2 - 1) \le N_2 - x < N_2 - f$. By Lemma 4, $m_f^* = \min\{N_1, n_1 + f(n_2 - 1)\} \le \min\{n_1 + f_2(n_2 - 1), N_1, N_2 - x + 1\} \le m^*$. Moreover, $m^* = m_{f'}^*$ where $f' = \min\{x, f_2\}$. ∎

We show that the bundled and incremental multicast traffic can have fundamental differences. Assuming the input switch has no fan-out capacity, then an SNB three-stage Clos network with incremental f_2-cast traffic does not exist if f_2, r_1, r_2 all $>$ 1. Fig. 4 illustrates this fact by an example. Suppose the network has connected the two requests (i_1, o_3) and (i_3, o_1). Then the new request (i_1, o_2) is blocked since it has to be routed by the first center switch, which has already been taken by the (i_3, o_1) connection.

Theorem 6: Assuming the center switch has no fan-out capacity, then for an SNB three-stage Clos network with bundled f_2-cast traffic

$$m^* = \min\{(N_1 - 1)f_2 + \min\{f_2, r_2\}, \\ (n_1 - 1)f_2 + n_2 - 1 + \min\{f_2, r_2\}, N_2\}.$$

Proof: In the Proof of Theorem 2, we obtained the worst-case by not using the center switch fan-out, except the current f-request was routed through one center switch. When the center switch has no fan-out capacity, then the f-request will take $\min\{f, r_2\}$ center switches, since if $f > r_2$, then r_2 center switches suffice by using the output switch fan-out. Therefore, the only adjustment needed from Theorem 2 is to replace 1 for the current request by $\min\{f, r_2\}$. Moreover, since m^* is increasing in f, we can replace f by f_2. ∎

Corollary 7: Theorem 6 remains true if the f_2-cast traffic is incremental.

Proof: By Lemma 1 and the fact that the "sufficiency" argument does not depend on the traffic type. ∎

Theorem 8: Assuming the output switch has no fan-out capacity, then for an SNB three-stage Clos network with bundled f_2-cast traffic, $m^* = \min\{N_1 f_2, (n_1 - 1)f_2 + n_2, N_2\}$.

Proof: The only difference from Theorem 2 is that the new request will take more than one center switch if it goes to more than one outlet in an output switch.

1) "if." Let f be the maximum number of outlets on an O_i, $1 \leq i \leq t$. Then the sufficiency condition $(N_1 - 1)f_2 + 1$ in Theorem 2 is replaced by $(N_1 - 1)f_2 + f \leq (N_1 - 1)f_2 + f_2 = Nf_2$.

2) "only if." We construct a blocking state for $m^* - 1$ center switches with a new f'-request, $f' = \min\{f_2, n_2\}$, from an inlet lying on an input switch I to an outlet lying on an output switch O. Let $x = \min\{(N_1 - n_1)f_2, n_2 - f'\}$ (if $x = (N_1 - n_1)f_2$, then $f' = f_2$). Namely, the inlets lying on other input switches can take x center switches. The blocking state is that the $(n_1 - 1)$ co-inlets take $\min\{(n_1 - 1)f_2, N_2 - f' - x\}$ center switches on way to some output switches and the other inlets take another x center switches on way to the output switch O. So, $\min\{(n_1 - 1)f_2, N_2 - f' - x\} + x = m^* - f'$ center switches have been used. Then no accessible center switch is left for the new request and the new request needs f' fresh center switches. ∎

Corollary 9: Theorem 8 remains true if the f_2-cast traffic is incremental.

Proof: By Lemma 1 and the fact that the "sufficiency" argument does not depend on the traffic type. ∎

Theorem 10: A three-stage Clos network with bundled f_2-cast traffic is WSNB under the no-split routing, if

$$m^* = \min\{N_1 + (n_1 - 1)(r_2 - 1), (n_1 - 1)(r_2 - 1) + n_2, N_1 + (n_1 - 1)(f_2 - 1), (n_1 - 1)f_2 + n_2, N_2\}.$$

Proof: Theorem 10 is obtained from Theorem 2 by replacing $x = \min\{(N_1 - n_1)f_2, n_2 - 1\}$, the total number of fan-outs from "other" inlets, with $x' = \min\{N_1 - n_1, n_2 - 1\}$, since each inlet needs only one center switch to engage with any number of co-outlets from a fixed O_i. Also replace $\min\{(n_1 - 1)f_2, N_2 - 1 - x\}$, the total number of fan-out from the $(n_1 - 1)$ co-inlets, with $\min\{(n_1 - 1)(r_2 - 1) + n_2 - 1 - x, (n_1 - 1)r_2, (n_1 - 1)f_2, N_2 - 1 - x\}$. Note that each co-inlet can take at most $\min\{f_2, r_2\}$ center switches under the no-split routing and the worst-case scenario is the same as in Theorem 2, that is, $f = 1$. Let O be the output switch of co-outlets. If $r_2 \leq f_2$, then the $(n_1 - 1)r_2$ maximum connections from the co-inlets are supported by the $(n_2 - 1 - x)$ co-outlets and $(r_2 - 1)$ output switches. If $n_1 - 1 \geq n_2 - 1 - x$, then $(n_2 - 1 - x)$ co-inlets can go to r_2 output switches, but the other can go to only $r_2 - 1$, as O has no more co-outlets available. Thus the co-inlets can take $(n_2 - 1 - x)r_2 + [(n_1 - 1) - (n_2 - 1 - x)](r_2 - 1) = (n_1 - 1)(r_2 - 1) + n_2 - 1 - x$ center switches. If $n_1 - 1 < n_2 - 1 - x$, the $(n_1 - 1)$ co-inlets can take $(n_1 - 1)r_2$ center switches. ∎

Corollary 11: Theorem 10 remains true if the f_2-cast traffic is incremental.

Proof: By Lemma 1 and the fact that the "sufficiency" argument does not depend on the traffic type. ∎

IV. COMPARISONS WITH EXISTING RESULTS

Under the same assumption as Theorem 2, except that f also has a lower bound, Giacomazzi and Trecordi [3] gave a different result. For comparison, we set that lower bound to 1. Then the formula given in [3] is

$$m^* = \max_{f, f_b, r}\{f + f_b + \min\{N_2 - f - f_b, f_2 r\}\}$$

where f_b is the total fan-out of the c connections engaged by the co-inlets, and r is the maximum number of connections engaged by the co-outlets. They also specified

$$cf \leq f_b \leq \min\{cf_2, N_2 - f\},$$
$$1 \leq c \leq \min\{n_1 - 1, N_2 - f\},$$
$$\gamma = \min\{N_2 - f - f_b, (n_2 - 1)f\}.$$

First of all, since c is a variable, m^* should also maximize over c. Thus m^* is computed over the ranges of four variables, with three of them possibly substantial. The m^* given in Theorem 2 is chosen from three values. But the differences between these two results are more than expressions. For example, when $\gamma = (n_2 - 1)f$, then m^* is increasing in f. Hence

$$m^* = \min\{N_2, [1 + f_2(n_2 - 1)]f_2\}$$

which contains an f_2^2 term never appeared in Theorem 2.

Hwang and Jajszczyk [5] considered a WSNB three-stage Clos network with bundled f_2-cast traffic. (Actually, they allowed a request from f_1 inlets to f_2 outlets. We set $f_1 = 1$ for comparison.) Their "strategy 2" forbids using input switch fan-out. Thus their result

$$m^* \geq \max_{1 \leq f \leq f_2} \min\{n_1 + f(n_2 - 1), N_1, N_2 - f + 1\}$$

is comparable to Theorem 5. In fact, Theorem 5 solves the maximization problem in m^* by showing that m^* achieves maximum at

$$f = \min\left\{f_2, \left\lceil \frac{N_2 - n_1 + 1}{n_2} \right\rceil\right\}.$$

Thus m^* can be computed by comparing two, instead of $2f_2$, values.

Their "strategy 3" forbids using center switch fan-out. Thus their result

$$m^* = \min\{n_2 - 1 + (n_1 - 1)\min\{f_2, r_2 - 1\} + f_2, N_2, N_1 - n_1 + n_1 f_2\}$$

is comparable to Theorem 6. We notice that their term $N_1 - n_1 + n_1 f_2$ is quite different from our term $(N_1 - 1)f_2 + \min\{f_2, r_2\}$. The difference is because they define $x = \min\{N_1 - n_1, n_2 - 1\}$. Note that by multiplying $(N_1 - n_1)$ with f_2, as we did in Theorem 2, then $(N_1 - n_1)f_2 + n_1 f_2 = N_1 f_2$, which is the same as our term when $f_2 \leq r_2$.

Masson and Jordan [6] considered the model as assumed in Corollary 11 except they ignored the boundary effects of N_1 and N_2 and assumed $f_2 = r_2$. They obtained a sufficient condition $m \geq n_1 r_2 + n_2 - 1$, while Corollary 11 yields, under

the same premise as theirs, $(n_1 - 1)(r_2 - 1) + n_2$. Thus the two results differ by at least $n_1 + r_2 - 2$. Hwang [4] added in the boundary effects on N_1 and N_2 to obtain another sufficient condition $m \geq \min\{n_1 r_2, N_2\} + \min\{n_2, N_1\} - 1$, which we have also tightened.

Finally, for $f_2 = 1$, m^* in every theorem and corollary is reduced to $m^* = \min\{N_1, n_1 + n_2 - 1, N_2\}$, the Clos SNB theorem [2] for the point-to-point traffic.

V. CONCLUSIONS

We gave necessary and sufficient conditions for strictly nonblocking three-stage Clos network under both bundled and incremental multicast traffic. We also extended these results to some other models studied in the literature, namely, when switches in a certain stage have no fan-out capacity, and when a no-split algorithm is used. It is interesting to note that while the network is definitely doing worse by losing the fan-out capacity of either the center stage or the output stage, it is not necessarily so with respect to the input stage.

Our results should be helpful in the design of three-stage multicast networks. For example, all results except Theorem 5 tell us that for fixed N_1, N_2, and f_2, we should keep n_1 and n_2 small, but more importantly, to keep n_1 small. On the other hand, Theorem 5 tells us that if the input switch has no fan-out capacity, or this capacity is banned by our routing algorithm, then it is more important to keep n_2 small. We also compared our results with existing results to unify the theory.

ACKNOWLEDGMENT

REFERENCES

[1] D. Cantor, "On nonblocking switching networks," *Networks*, vol. 2, pp. 367–377, 1972.

[2] C. Clos, "A study of nonblocking switching networks," *Bell Syst. Tech. J.*, vol. 32, pp. 406–424, 1953.

[3] P. Biacomazzi and V. Trecordi, "A study of nonblocking multicast switching networks," *IEEE Trans. Commun.*, vol. 43, no. 2,3,4, pp. 1163–1168, 1995.

[4] F. K. Hwang, "Rearrangeability of multiconnection three-stage Clos networks," *Networks*, vol. 2, pp. 301–306, 1972.

[5] F. K. Hwang and A. Jajszczyk, "On nonblocking multiconnection networks," *IEEE Trans. Commun.*, vol. COM-34, no. 10, pp. 1038–1041, 1986.

[6] G. M. Masson and B. W. Jordan, "Generalized multistage connection networks," *Networks*, vol. 2, pp. 191–209, 1972.

Frank K. Hwang received the B.A. degree from National Taiwan University, Taipei, Taiwan, R.O.C., in 1960, and the Ph.D. degree from North Carolina State University, Raleigh, in 1968.

He worked at the Mathematics Center at Bell Labs from 1967-1996. He is now a Chair-Professor at National Chiao Tung University, Taiwan. He has published about 350 papers and written or coauthored the following books: *The Steiner Tree Problem*, North-Holland, 1992; *Combinatorial Group Testing and Its Applications*, Singapore: World Scientific, 1993, 2nd edition, 2000; *The Mathematical Theory of Nonblocking Switching Networks*, Singapore: World Scientific, 1998; and *Reliabilities of Consecutive-k Systems*, Norwell, MA: Kluwer, 2000.

Sheng-Chyang Liaw received the B.S., M.S., and Ph.D. degrees in applied mathematics from National Chiao Tung University, Hsinchu, Taiwan, R.O.C., in 1989, 1991, and 1998 respectively.

He was a postdoctoral with the Department of Applied Mathematics, National Sun Yat-Sen University, Kaohsiung, Taiwan. He is currently an Assistant Professor in the Department of Mathematics, National Central University, Chung-li, Taiwan. His research interests include network analysis, graph algorithm and combinatorial mathematics.