Nonblocking Conditions of Multicast Three-stage Interconnection Networks

Achille Pattavina and Gino Tesei
Department of Electronics and Information, Politecnico di Milano, Piazza Leonardo da Vinci 32, 20133 Milano, Italy

This article considers three-stage switching networks able to support multicast traffic, that is, connections in which one inlet is connected to more than one output at the same time. The nonblocking conditions for this network are studied under the assumption of absence of any optimized routing of the connections inside the structure (the so-called strict-sense nonblocking networks). The theoretical nonblocking condition for such network under point-to-point traffic is the well known Clos condition. We give here the necessary and sufficient conditions for such network to be strict-sense nonblocking under multicast traffic.

Keywords: switching; three-stage network; Clos network; multicast traffic; nonblocking conditions

1. INTRODUCTION

The evolution of communication networks is determined by the combined effect of three different factors: transmission, switching, and protocols. It is well known that transmission and switching usually rely on hardware devices, whereas protocols are based on software resources. The development of worldwide high-speed networking that matches the growing user needs requires an ever-improving capability in each of the three factors. For a long time the main technology governing transmission and switching was electronics. The optical technology opened this monopoly by showing the economical convenience of optical transmission systems over the corresponding electronic ones. As a result, optical point-to-point transmission links carrying a very large bandwidth were made available. However, the building of a flexible high-speed network requires that such huge transmission capacity is made available on demand by users: this is the task of switching. The development of more and more powerful network protocols has allowed the support of packet switched services: today’s success of Internet is the best witness. Tremendous progresses in processor technologies and storage devices have made possible all of this. In this scenario it seems that the bottleneck to the provision of a worldwide on-demand network connectivity for the support of any service is switching.

This article addresses a very important issue in the area of switching, which is how to make available input-output connections within a node of a communication network so as to set up connection paths between end-users crossing several network nodes. The set of devices in the node performing this task is known as a connecting network or interconnection network. It is shown here how to guarantee that any connection request can be satisfied by the node without blocking in the case of multicast connections, in which one input must be connected to several outputs at the same time.

In the following, Section 2 provides the backgrounds in the technical literature on the issue analyzed here, while Section 3 prepares the mathematical formulation needed to prove the non blocking conditions of multicast connecting networks. These conditions are stated and proved in Section 4, whereas Section 5 discusses the new result.

2. BACKGROUNDS ON NONBLOCKING NETWORKS

The understanding of what switching is about relies on the basic definition of a crossbar matrix: it is a structure capable of connecting any of its idle input lines to any of its idle output lines by means of crosspoints. A crossbar matrix with size $N_1 \times N_2$ includes $N_1 N_2$ crosspoints, each dedicated to the connection of a specific input to a specific output (point-to-point or unicast request). Apparently a control is needed in the crossbar matrix to make the single crosspoint “operate” for establishing the requested input/output connection. Supporting a multicast connection between an idle inlet and more than one idle outlet through a crossbar matrix implies only the operation of as many crosspoints as the number of connection outputs. The crossbar matrix is intrinsically nonblocking for both unicast and multicast traffic, in that any connection request between an idle inlet and idle outlet(s) is always accommodated.
The building of crossbar matrices using electronic components has not been a problem in the last decades; larger and larger crossbar matrices have been required as the switching needs kept increasing together with the growth of network users and correspondingly of the communication network connecting them. The order of magnitude of users and correspondingly of the communication network needs kept increasing together with the growth of network components has not been a problem in the last decades; larger and larger crossbar matrices have been required as the switching stages, each composed by a stack of crossbar matrices. The interstage connection pattern is such that each matrix has access to any matrices of the following stage, thus providing full internal connectivity. Clos provided the condition to design a three-stage connecting network so that it is always nonblocking for point-to-point (unicast) connections [2].

The single/multiple-stage connecting network described insofar represents the basic switch model for accomplishing the interconnection between inputs and outputs of a given switch fabric to transfer data. This kind of communication service is referred to as circuit switching, as a circuit is established at a given time between two users through a communication network, by thus crossing a certain number of network nodes in which the interconnection function is performed by a single/multiple-stage connecting network. Only a proper technology can enable the building of connecting networks with almost arbitrary size. Time division multiplexing is the key technique enabling to carry thousands of circuits on a digital link, which implies that the switching techniques has to take into account the type of signal to be switched. This consideration led to the point that digital switching is the only viable solution for building large switching fabrics.

To build a circuit-switched network capable of supporting large flows of data streams without requiring switching on a circuit-per-circuit basis, a different type of switch architecture has been devised called digital cross connect (DXC) [12]. This type of network is configured by telecommunication operators as their basic transport infrastructure so as to make available high-speed point-to-point paths between ending nodes in a complex communication structure. The DXCs are switching machines directly controlled by network operators who change the DXC configuration so that the needed point-to-point paths are configured and changed whenever needed. The internal structure of a DXC node does not differ much from that of a standard circuit-switched node; hence, similar architectural solutions are adopted for the connecting network. However, the key difference is that a DXC does not handle signalling, which is the means used to setup and teardown connections upon demand in a classical circuit-switched node.

The scalability property of Clos networks has made this network type important also for packet switched networks; in fact, the nonblocking concept is maintained given that we scale down the duration of a connection to the switching time of a packet (IP network) or of a cell (ATM network) [1].

The importance of Clos-connecting networks goes well beyond the electronic technology domain. In fact, when optical technology is being used, the concept of Clos nonblocking in three-stage networks remains as a landmark in the design process of switching fabrics. In the simplest cases the three-stage network is built using optical devices with straightforward construction (see, e.g. [7, 13, 15]). In the most complex cases in which different types of optical devices are being used, theoretical approaches have been developed, enabling us to prove if an optical interconnection network satisfies the nonblocking conditions [8].

Interest in this kind of networks has been growing in the last 2 decades due to the increasing need for supporting multicast communication services; a clear example is represented by the Internet in which multicast communications services have been provided since the early nineties (see, eg., [14]).

It is worth remembering that the nonblocking Clos network does not set any constraint on how to select the new connection out of the multiple paths available through the three-stage network. Such a class of networks, called “strict-sense nonblocking network” (SSNB), must be distinguished from other two types of nonblocking networks, that is, the “wide-sense nonblocking networks” (WSNB) and the “rearrangeable nonblocking networks” (RNB). In both cases a less complex nonblocking network is built compared to the SSNB one by allowing, at connection setup time, either an optimized path selection or a possible rearrangement of connections already setup.

In this article we will only consider the case of SSNB networks when each idle inlet can address a number of idle outlets ranging from a minimum \( f_1 \) up to a maximum \( f_2 \). Not many previous contributions are found in the technical literature to address the problem of dimensioning nonblocking multicast SSNB networks. The first contribution by Masson and Jordan [9] defines a sufficient condition for multicast SSNB networks that, as pointed out later by Feldman et al. [3], applies actually to WSNB networks. The work by Masson and Jordan has also been used by Hwang [5] to derive the nonblocking conditions of multicast rearrangeable three-stage RNB (a survey of main results on unicast RNB and SSNB can be found in [10]). Later, Yang and Masson [16] derived new results for WSNB networks, and more recently, Giacomazzi and Trecordi [4] claimed to have obtained necessary and sufficient conditions for SSNB multicast networks. We have already shown [11] how such conditions are actually only sufficient. We focus here only on SSNB networks and disregard any other types of nonblocking network. In this article we define the nonblocking properties for a multicast three-stage network, by providing as well the full proof of necessity and sufficiency.
3. MATHEMATICAL PRELIMINARIES

3.1. Definition of Multicast Clos Switching Network

In this section we shall give some fundamental definitions.

Definition 1. A three-stage network is defined as represented in Figure 1: it has $N_1$ inlets, $N_2$ outlets, and is composed of three stages of crossbar matrices with size $n_1 \times m$ at stage 1, $r_1 \times r_2$ at stage 2, $m \times n_2$ at stage 3, where $r_1 = N_1/n_1$ and $r_2 = N_2/n_2$. This class of networks, known as the Clos network, is SSNB if \[ m \geq n_1 + n_2 - 1. \]

Definition 2. An $N_1 \times N_2$ multicast switching network is now introduced, that is, in which each inlet can address a nonnegative integer number of outlets in the set $\{1,f_2\}$. Obviously, we assume that all matrices are multicast, that is, in which each inlet can be connected to more than one outlet at the same time. We denote this class of networks by $M$-CLOS($N_1,N_2,n_1,n_2,m,f_1,f_2$).

Definition 3. In an $M$-CLOS($N_1,N_2,n_1,n_2,m,f_1,f_2$) the set of inlets and outlets are denoted by $I = \{i_{k,i} \mid i \in \{1,\ldots,n_1\} \land k \in \{1,\ldots,r_1\}\}$ and $O = \{o_{h,j} \mid j \in \{1,\ldots,n_2\} \land h \in \{1,\ldots,r_2\}\}$, respectively, the set of inlets of the $k$-th matrix at first stage is denoted by $I_k = \{i_{k,i} \mid i \in \{1,\ldots,n_1\}\}$, where $k \in \{1,\ldots,r_1\}$, the set of outlets of the $h$-th matrix at third stage is denoted by $O_h = \{o_{h,j} \mid j \in \{1,\ldots,n_2\}\}$, where $h \in \{1,\ldots,r_2\}$. For ease of exposition, and with a slight abuse of notation, we have used $i_{k,i}$ to denote an inlet with index $(k,i)$.

Definition 4. According to the above definitions, in an $M$-CLOS($N_1,N_2,n_1,n_2,m,f_1,f_2$) the fanout of the inlet $i_{k,i}$ is denoted by $f_{k,i}$, where

\[
\begin{cases} 
  f_1 \leq f_{k,i} \leq f_2 & \text{if } i_{k,i} \text{ is busy} \\
  f_{k,i} = 0 & \text{else}
\end{cases}
\]

Definition 5. According to the above definitions, in an $M$-CLOS($N_1,N_2,n_1,n_2,m,f_1,f_2$) the multicast connection request from the inlet $i_{k,i}$, if $f_{k,i} \in \{f_1,\ldots,f_2\}$, to the set of free outlets $O_{k,i} \subseteq O$, is denoted by $C(i_{k,i};O_{k,i})$. Notice that $|O_{k,i}| = f_{k,i}$. Moreover, we denote by $c(i_{k,i},o_{h,j})$, where $o_{h,j} \in O_{i_{k,i}}$, one of the point-to-point connection requests in which $C(i_{k,i};O_{k,i})$ can be ideally decomposed.

Definition 6. Given an inlet $i_{k,i}$, let $C_{1,k,i}$ be the set of inlets belonging to the same matrix of $i_{k,i}$, except $i_{k,i}$, that is,

\[
C_{1,k,i} = \{i_{k,i'} \mid i' \in \{1,\ldots,n_1\} \land i' \neq i \land k' = k\}
\]

and, given an outlet $o_{h,j}$, let $C_{2,h,j}$ be the set of outlets belonging to the same matrix of $o_{h,j}$, except $o_{h,j}$, that is,

\[
C_{2,h,j} = \{o_{h,j'} \mid j' \in \{1,\ldots,n_2\} \land j' \neq j \land h' = h\}
\]

Connection requests between idle inlets and idle outlets can be blocked due to pre-established connections that make unavailable all the possible internal paths. An example of blocking state is represented in Figure 2, where an $M$-CLOS(24, 30, 4, 5, 10; 1, 5) is shown. For the sake of clarity in this figure and in the following only the interstage links actually engaged by the established connections are shown, despite the full connection pattern between matrices of adjacent stages. It can be easily verified that the point-to-point connection $c(i_{1,1},o_{5,5})$ cannot be set up due to network blocking, because no middle-stage matrix is left usable to connect the two tagged matrices at first and third stage.

Definition 7. In the condition of an unloaded network and given a tagged point-to-point connection request $c(i_{k,i},o_{h,j})$, the maximum number of second-stage matrices blocked by connections from inlets belonging to $C_{1,k,i}$, and thus unavailable for $c(i_{k,i},o_{h,j})$, is denoted by $m_{1,k,i}$. In detail, such conspiring connections must be addressed to outlets not
belonging to \( C_{2,h,j} \) first, and only when these outlets are exhausted, to outlets belonging to \( C_{2,h,j} \). Then, if \( r_1 > 1 \) and if there are outlets belonging to \( C_{2,h,j} \) not already addressed, we define \( m_{2,h,j} \) as the maximum number of second-stage matrices blocked for \( c(i_{k,i}, o_{h,j}) \) by connections addressed to outlets belonging to \( C_{2,h,j} \).

3.2. Counting the Number of Connections

In this section we find the maximum number of connection requests in multicast switching networks. Let us introduce the following notations

- \( x \mod y = x - \left\lfloor \frac{x}{y} \right\rfloor y \) if \( y \neq 0 \), for each \( x, y \in \mathbb{N} \);
- \( x \mod y = \left\lfloor \frac{x}{y} \right\rfloor y \) if \( y = 0 \), for each \( x, y \in \mathbb{N} \);
- \( x | y \iff x > 0 \land \exists \text{integer } k \mid y = kx \), for each \( x, y \in \mathbb{N} \);
- for each boolean predicate \( P \) (Iverson’s convention),
  \[ [P] = \begin{cases} 1 & \text{if } P \text{ is true} \\ 0 & \text{else} \end{cases} \]

**Lemma 1.** Let \( n \) be an integer and \( m \) be a positive integer; then

\[
 n = \sum_{0 \leq i \leq m-1} \left\lceil \frac{n+i}{m} \right\rceil = \sum_{0 \leq i \leq m-1} \left\lfloor \frac{n-i}{m} \right\rfloor
\]

**Proof.** There are two possible cases: \( m \mid n \) or \( m \not\mid n \).

If \( m \mid n \), then an integer \( k \) exists such that \( k = n/m \), which implies

\[
 \sum_{0 \leq i \leq m-1} \left\lceil \frac{n+i}{m} \right\rceil = \sum_{0 \leq i \leq m-1} k + \left\lceil \frac{i}{m} \right\rceil = km = n
\]

and

\[
 \sum_{0 \leq i \leq m-1} \left\lfloor \frac{n-i}{m} \right\rfloor = \sum_{0 \leq i \leq m-1} k - \left\lfloor \frac{i}{m} \right\rfloor = km = n
\]

(remember that, for each real number \( x \), \( \lceil -x \rceil = -\lfloor x \rfloor \)).

In the other case, because \( m \not\mid n \), then

\[
 \sum_{0 \leq i \leq m-1} \left\lceil \frac{n+i}{m} \right\rceil = \sum_{0 \leq i < \text{com } m} \left\lceil \frac{n}{m} \right\rceil + \sum_{\text{com } m \leq i \leq m-1} \left\lceil \frac{n}{m} \right\rceil
\]

which implies

\[
 \sum_{0 \leq i \leq m-1} \left\lceil \frac{n+i}{m} \right\rceil = \left\lceil \frac{n}{m} \right\rceil \text{com } m + \left\lceil \frac{n}{m} \right\rceil n \text{mod } m
\]

\[
 = n \left( \left\lceil \frac{n}{m} \right\rceil - \left\lfloor \frac{n}{m} \right\rfloor \right) = n
\]

having used the definitions of \( x \mod y \) and \( x \mod y \). Moreover,

\[
 \sum_{0 \leq i \leq m-1} \left\lfloor \frac{n-i}{m} \right\rfloor = \sum_{0 \leq i < \text{mod } m} \left\lfloor \frac{n}{m} \right\rfloor + \sum_{\text{mod } m \leq i \leq m-1} \left\lfloor \frac{n}{m} \right\rfloor
\]

which implies

\[
 \sum_{0 \leq i \leq m-1} \left\lfloor \frac{n-i}{m} \right\rfloor = \sum_{0 \leq i < \text{mod } m} \left\lfloor \frac{n}{m} \right\rfloor + \sum_{\text{mod } m \leq i \leq m-1} \left\lfloor \frac{n}{m} \right\rfloor
\]

\[
 = n \left( \left\lceil \frac{n}{m} \right\rceil - \left\lfloor \frac{n}{m} \right\rfloor \right) = n
\]

3.3. The Worst Traffic Pattern in SSNB Switching Networks: The Conspiracy Principle

In multicast switching networks, as in the unicast switching networks, (necessary and sufficient) SSNB conditions can be derived from the worst-case scenario. This can be constructed by the **conspiracy principle:** given a multicast “conspired” connection request \( \zeta \) the worst traffic pattern for \( \zeta \) is derived by assuming a set of multicast “conspiring” connection requests selected so as to maximize the network resources that would be used by \( \zeta \). The application of the conspiracy principle to a particular class of networks depends on the interconnection pattern and on the range of possible fanouts.

Let us consider an \( M-\text{CLOS} (N_1, N_2, n_1, n_2, m; f_1, f_2) \) network, in which \( C(i_{k,i}; O_{h,j}) \) is the multicast conspired connection request and \( c(i_{k,i}, o_{h,j}) \), where \( o_{h,j} \in O_{k,i} \), is one of the point-to-point connection requests in which \( C(i_{k,i}; O_{h,j}) \) can be ideally decomposed. To obtain the worst conspiring traffic pattern for \( C(i_{k,i}; O_{h,j}) \), we have to assume that the fanout of the conspiring inlet must be minimal, that is, \( f_{k,i} = f_1 \). In fact, if the fanout of \( i_{k,i} \) is greater than the minimum \( f_1 \), then the conspiring inlets could not address the maximum number \( C_{c,\text{max}} \) of conspiring outlets. This last number is simply given by Theorem 1, considering that the conspired connection engages one inlet and \( f_1 \) outlets, that is

\[
 C_{c,\text{max}} = \min \left\{ (N_1 - 1)f_2, N_2 - f_1, \left( \left\lceil \frac{N_2}{f_1} \right\rceil - 1 \right) f_2 \right\}
\]  

(2)

Initially, we are going to identify the worst case scenario for \( c(i_{k,i}, o_{h,j}) \) and then we shall show that this is also the worst-case scenario for \( C(i_{k,i}; O_{h,j}) \).

**Fact 1.** The connection request \( c(i_{k,i}, o_{h,j}) \) can be blocked only by connection requests from inlets belonging to the same matrix of \( i_{k,i} \) or by connection requests directed to outlets belonging to the same matrix of \( o_{h,j} \).

In fact, each connection request \( c(i_{k,i}, o_{h,j}) \), where \( (k' \neq k) \land (h' \neq h) \), uses links at the first interstage not available
for \( k,l \) and links at the second interstage not directed to the same matrix of \( o_{k,j} \). On the contrary, each connection request \( c(i_{k,j}, o_{k,j}) \), where \((k' = k) \lor (h' = h)\), uses links at the first interstage available for \( i_{k,j} \), or links at the second interstage directed to the same matrix of \( o_{k,j} \).

The conspiracy to \( c(i_{k,j}, o_{k,j}) \) starts in the condition of the unloaded network requiring that inlets belonging to \( C_{1,k,j} \) use the maximum number of matrices at the second stage \( (m_{1,k,j}) \). Then, if there are enough free outlets and if \( r_1 > 1 \), the conspiracy to \( c(i_{k,j}, o_{k,j}) \) continues requiring that the maximum number of connections directed to outlets belonging to \( C_{2,h,j} \) is established from inlets \( i_{k,j} \), where \((k' \neq k)\), using second-stage matrices not yet busy \((m_{2,h,j})\). Notice that

\[
m_{1,k,j} \leq (n_1 - 1)f_2, \forall k, i \quad \text{and} \quad m_{2,h,j} \leq n_2 - 1, \forall h, j
\]

**Example 1.** Figure 3 shows an M-CLOS \((24, 40, 4, 5, 13; 1, 3)\), in which the conspired point-to-point connection request is \( c(i_{1,j}, o_{8,1}) \). The number of matrices blocked at the second stage by inlets belonging to \( C_{1,1,1} \) is maximal (see Fig. 3a), that is, \((n_1 - 1)f_2 = (4 - 1) \cdot 3 = 9\). Moreover, the number of matrices blocked by outlets of \( C_{2,8,1} \) is also maximal (see Fig. 3b), that is, \((n_2 - 1) = (5 - 1) = 4\). Hence, no more second-stage matrix is left to satisfy \( c(i_{1,j}, o_{8,1}) \).

We have seen how to construct the worst traffic pattern for a single connection request \( c(i_{k,j}, o_{k,j}) \). Now, we have to find the worst traffic pattern when the conspired connection is multicast. Moreover, if the number of second-stage matrices is enough to satisfy \( c(i_{k,j}, o_{k,j}) \) for each network state, can every other connection request be satisfied?

Every other connection request \( c(i_{k,j}, o_{k,j}') \) from \( i_{k,j} \) can share \( l \) interstage links with \( c(i_{k,j}, o_{k,j}) \), where \( 0 \leq l \leq 2 \). If \( l = 2 \) (hence, \( h = h' \)), then \( c(i_{k,j}, o_{k,j'}) \) can be trivially realized using the same links as those of \( c(i_{k,j}, o_{k,j}) \) by operating the connection splitting at the third stage. We assume then \( l = 0 \lor l = 1 \), which is absence of fanout capability at the third stage, which corresponds to selecting the worst-case conditions. The examples shown in Figures 2 and 3 are consistent with such hypothesis.

**Remark 1.** In the conspiracy of \( c(i_{k,j}, o_{k,j'}) C_{1,k,j} \) is the same as in the conspiracy of \( c(i_{k,j}, o_{k,j}) \). But all the inlets belonging to \( C_{1,k,j} \) have already been involved in the conspiracy of \( c(i_{k,j}, o_{k,j}) \), hence, no more matrices at the second stage can be further blocked by inlets belonging to \( C_{1,k,j} \) in the conspiracy of \( c(i_{k,j}, o_{k,j'}) \).

On the contrary, it is not true that \( C_{2,h,j} = C_{2,h',j} \); rather, \( m_{2,h,j} \leq m_{2,h,j'} \leq n_2 - 1 \) holds, because a part of the outlets in the same matrix of \( o_{k,j'} \) could already be required in the conspiracy of \( c(i_{k,j}, o_{k,j}) \), which is the first to be satisfied, by construction, and not vice versa. It follows that the maximum number \( m \) of second-stage matrices necessary to satisfy \( c(i_{k,j}, o_{k,j}) \) is given by

\[
m = m_{1,k,j} + m_{2,h,j} + 1
\]

and the maximum number \( m' \) of second-stage matrices necessary to satisfy \( c(i_{k,j}, o_{k,j'}) \) is given by

\[
m' = m_{1,k,j} + m_{2,h,j'} + 1
\]

hence, \( m \geq m' \).

In conclusion, the maximum number of matrices blocked at the second stage by conspiring connections for a tagged conspired inlet \( i_{k,j} \) is given in the condition of the unloaded

![FIG. 3. Example showing the conspiracy principle in a blocking network.](image-url)
network, provided that \( f_{k,i} = f_1 \). Moreover, the worst-case scenario for one of the point-to-point connection requests of the multicast connection coincides with the worst-case scenario for the multicast connection itself.

Thus, remembering that \( m_1, k, i \leq (n_1 - 1) f_2, \forall k, i \) and \( m_2, k, j \leq n_2 - 1 \), the following lemma holds.

**Lemma 2.** An \( M\text{-CLOS}(N_1, N_2, n_1, n_2, m; f_1, f_2) \) is SSNB if

\[
  m \geq (n_1 - 1) f_2 + n_2
\]

**Example 2.** Figure 4a shows an \( M\text{-CLOS}(24, 40, 4, 5, 14; 1, 3) \) where the conspiracy is for \( c(i_{1,1}, o_{8,1}) \). The number of matrices blocked at the second stage by the inlets of \( C_{1,1,1} \) is maximal, that is, \((n_1 - 1) f_2 = (4 - 1) \cdot 3 = 9\). The number of matrices blocked at the second stage by outlets of \( C_{2,8,1} \) is also maximal, that is, \((n_2 - 1) = (5 - 1) = 4\).

Because \( m = 14 \), only one matrix is thus left to satisfy \( c(i_{1,1}, o_{8,1}) \). Figure 4b assumes that the fanout of the conspired inlet \( i_{1,1} \) is equal to 2 (instead of 1, like in the previous example); hence, two are the conspired point-to-point connection requests in which the multicast conspired connection request \( C(i_{1,1}; \{o_{8,1}, o_{7,1}\}) \) can be ideally decomposed, that is, \( c(i_{1,1}, o_{8,1}) \) and \( c(i_{1,1}, o_{7,1}) \). So, we have \( m_{1,1,1} = 9 \) and \( m_{2,8,1} = m_{2,7,1} = 4 \). In this example the requested multicast connection is satisfied crossing in the second stage both an idle matrix and a partially used matrix.

This concludes the application of the conspiracy principle in the case of \( M\text{-CLOS}(N_1, N_2, n_1, n_2, m; f_1, f_2) \) switching networks.

4. NECESSARY AND SUFFICIENT CONDITION OF STRICT-SENSE NONBLOCKING

We give here the necessary and sufficient conditions that make a three-stage multicast network strictly nonblocking.

**Theorem 2.** An \( M\text{-CLOS}(N_1, N_2, n_1, n_2, m; f_1, f_2) \) is SSNB if and only if

\[
  m \geq \min\left\{ (n_1 - 1) f_2 + n_2, (N_1 - 1) f_2 + 1, N_2 - f_1 + 1, \left\lfloor \frac{N_2}{f_1} \right\rfloor - 1 \right\} + 1
\]

**Proof.** By lemma 2, a sufficient condition for an \( M\text{-CLOS}(N_1, N_2, n_1, n_2, m; f_1, f_2) \) to be SSNB is

\[
  m \geq (n_1 - 1) f_2 + n_2
\]

We have also found that the maximum number of conspiring outlets \( C_{c,\text{max}} \) is given by Equation (2). Therefore, a multicast network is SSNB if

\[
  m \geq \min\{ (n_1 - 1) f_2 + n_2, C_{c,\text{max}} + 1 \}
\]

which proves the sufficiency condition of Equation (3).
To prove the necessity we assume that the number of middle-stage matrices \( m \) is at most that provided by Equation (3) decreased by 1. Therefore,

\[
m \leq \min \left\{ (n_1 - 1)f_2 + n_2, (N_1 - 1)f_2 + 1, N_2 - f_1 + 1, \left( \frac{N_2}{f_1} - 1 \right)f_2 + 1 \right\} - 1
\]

\[
\leq (n_1 - 1)f_2 + n_2 - 1
\]

It is easy to find network instances where \( m_{1,k,i} + m_{2,h,j} = m \), determining a blocking condition for the connection \( C(i_k,j;O_{h,j}) \). One of these instances is given by the network configuration and occupancy represented in Figure 3b.

Before concluding this section we give the application of theorem 2 in the case of unicast switching networks. If \( f_1 = f_2 = 1 \), then

\[
m \geq \min \left\{ (n_1 - 1)f_2 + n_2, (N_1 - 1)f_2 + 1, N_2 - f_1 + 1, \left( \frac{N_2}{f_1} - 1 \right)f_2 + 1 \right\}
\]

becomes

\[
m \geq \min\{n_1 + n_2 - 1, N_1, N_2\}
\]

This equation clearly reduces to the Clos condition [Equation (1)] of point-to-point networks in all nonlimiting cases (when \( n_1 + n_2 < N_1 \)) and \( n_1 + n_2 < N_2 \).

The particular case of a multicast three-stage network, in which the minimum fanout \( f_1 \) is set to 1 has been analyzed in [6]: the result of Theorem 2 is consistent with the findings of such an article.

5. DISCUSSION

We discuss now the application of the new result to two subcases.

The first one concerns the scenario in which a SSNB \( M\text{-CLOS}(N, n, n, n, n, n, m; f_1, f_2) \) has \( N, n \) and \( f_1 \) fixed (see Fig. 5a), while \( f_2 \) can range from a minimum \( f_1 = 1 \) up to a maximum given by Equation (3). Notice that, given \( f_2 \), \( m \) increases for greater values of \( n \) unless a large maximum fanout is selected. The physical meaning of this fact could be understood by the conspiracy principle. In fact, given all the other constraints, greater values of \( n \) imply, on one hand, a greater cardinality of \( C_{1,k,i} \) (i.e., a greater size of the first set of conspiring connections), and on the other hand, a greater cardinality of \( C_{2,h,j} \) (i.e., a greater size of the second set of conspiring connections).

The second application refers to the scenario in which a SSNB \( M\text{-CLOS}(N, n, n, n, m; f_1, f_2) \) has \( N, n \) and \( f_1 \) fixed (see Fig. 5b), while \( f_1 \) can range from 1 up to the maximum \( f_2 \). Notice that the number \( m \) of second-stage matrices required to provide nonblocking decreases for greater values of \( f_1 \). Mathematically, it depends on the fact that Equation (3) has several terms decreasing for greater values of \( f_1 \). But the (more interesting) physical reason is that, \( f_2 \) being fixed, the range of a (general) fanout decreases for greater values of \( f_1 \). In other words, increasing \( f_1 \) gives a less general kind of network and, hence, a smaller amount of network resources should be necessary to obtain the SSNB property.

6. CONCLUSIONS

In this article we have found the theoretical conditions to build nonblocking multicast three-stage connecting networks. These conditions refer to the case of strict-sense nonblocking, and therefore do not set any constraint on the algorithm to select the internal path for any new multicast.
connection. Three-stage nonblocking networks loaded by multicast traffic can thus be designed based on the fanout parameters of the requested connections. We have shown how large values of the minimum and maximum fanout affect the overall network structure.

ACKNOWLEDGMENTS

The authors are grateful to anonymous Associate Editor and reviewers for providing very useful suggestions that improved this article, with particular regard to the proof of Theorem 2, which is now much simpler than in the original draft.

APPENDIX

Proof of Theorem 1

Theorem. For each M-CLOS($N_1, N_2, n_1, n_2, m; f_1, f_2$) the maximum number of addressed outlets is

$$C_{o, max}(N_1, N_2; f_1, f_2) = \min \left\{ N_1 f_2, N_2, \left\lceil \frac{N_2}{f_1} \right\rceil f_2 \right\}$$

Proof. It is obvious that $C_{o, max}(N_1, N_2; f_1, f_2) \leq N_2$ and $C_{o, max}(N_1, N_2; f_1, f_2) \leq N_1 f_2$, thus

$$C_{o, max}(N_1, N_2; f_1, f_2) \leq \min \{N_1 f_2, N_2\}$$

(4)

So, in the case of $N_1 f_2 \leq N_2$, we have the maximum number of connection requests if each inlet requests $f_2$ connections, that is,

$$C_{o, max}(N_1, N_2; f_1, f_2) = N_1 f_2$$

(5)

which proves the theorem in the case of $N_1 f_2 \leq N_2$.

On the contrary, if $N_1 f_2 > N_2$ then the maximum number of busy inlets is given by $\left\lceil \frac{N_2}{f_1} \right\rceil f_2$ and, because the maximum inlet fanout is $f_2$, if $\left\lceil \frac{N_2}{f_1} \right\rceil f_2 < N_2$ then

$$C_{o, max}(N_1, N_2; f_1, f_2) = \left\lfloor \frac{N_2}{f_1} \right\rfloor f_2,$$

or if $\left\lceil \frac{N_2}{f_1} \right\rceil f_2 \geq N_2$ then

$$C_{o, max}(N_1, N_2; f_1, f_2) = N_2.$$ 

In fact, if $\left\lceil \frac{N_2}{f_1} \right\rceil f_2 \geq N_2$, then

$$\left\lfloor \frac{N_2}{f_1} \right\rfloor f_1 \leq N_2 \leq \left\lfloor \frac{N_2}{f_1} \right\rfloor f_2$$

which implies $N_2 - \left\lfloor \frac{N_2}{f_1} \right\rfloor f_1 \geq 0$. But $N_2 - \left\lfloor \frac{N_2}{f_1} \right\rfloor f_1 = N_2 \mod f_1$ which, by Lemma 1, becomes

$$N_2 \mod f_1 = \sum_{0 \leq i \leq \left\lfloor \frac{N_2}{f_1} \right\rfloor - 1} \left\lfloor \frac{N_2 \mod f_1 + i}{f_1} \right\rfloor$$

or

$$N_2 \mod f_1 = \sum_{0 \leq i \leq \left\lfloor \frac{N_2}{f_1} \right\rfloor - 1} \left\lfloor \frac{N_2 \mod f_1 - i}{f_1} \right\rfloor$$

In other words, the $N_2$ connections are set up by requiring that each of the $\left\lfloor \frac{N_2}{f_1} \right\rfloor$ busy inlets requests $f_1 + \frac{N_2 \mod f_1}{f_1}$ connections, if $\left\lfloor \frac{N_2}{f_1} \right\rfloor \mid N_2 \mod f_1$; or, if $\sim (\left\lfloor \frac{N_2}{f_1} \right\rfloor \mid N_2 \mod f_1)$, by requiring that each of the $\left(\frac{N_2 \mod f_1}{f_1} \right) \mod \left\lfloor \frac{N_2}{f_1} \right\rfloor$ busy inlets requests $f_1 + \frac{N_2 \mod f_1}{f_1}$ connections, and each of the other $(N_2 \mod f_1) \com \left\lfloor \frac{N_2}{f_1} \right\rfloor$ busy inlets requests $f_1 + \left\lfloor \frac{N_2 \mod f_1}{f_1} \right\rfloor$ connections. Notice that, if $\left\lfloor \frac{N_2}{f_1} \right\rfloor \mid N_2 \mod f_1$, then $f_2 - f_1 \geq \frac{N_2 \mod f_1}{f_1}$. In fact, if $f_2 - f_1 < \frac{N_2 \mod f_1}{f_1}$ then, by definition of a mod operator, we have $\left\lfloor \frac{N_2}{f_1} \right\rfloor f_2 < N_2$, which is in contrast with the assumptions. By similar reasoning, if $\sim (\left\lfloor \frac{N_2}{f_1} \right\rfloor \mid N_2 \mod f_1)$, it can be shown that $f_2 - f_1 \geq \left\lfloor \frac{N_2 \mod f_1}{f_1} \right\rfloor$. In conclusion,

$$C_{o, max}(N_1, N_2; f_1, f_2) = \min \left\{ N_2, \left\lfloor \frac{N_2}{f_1} \right\rfloor f_2 \right\}$$

(6)

which proves the theorem in the case of $N_1 f_2 > N_2$.

By Equations (5) and (6) the theorem is proved.

REFERENCES


