Appears in *IEEE Communications Magazine*, October 2003

Multirate Clos Networks

Jonathan S. Turner
Washington University in St. Louis
jst@cse.wustl.edu

Riccardo Melen
Università dell'Insubria, Varese, Italy
riccardo.melen@uninsubria.it

Abstract

The Clos networks are a class of multistage switching network topologies that provide alternate paths between inputs and outputs, making it possible to minimize or eliminate the blocking that can otherwise occur in such networks. In his seminal paper in the Bell System Technical Journal in 1953, Charles Clos showed how these networks could be configured to make them nonblocking and effectively launched the systematic study of switching system performance, a field that has developed a rich technical literature and which continues to be very active and of continuing practical importance. This paper describes how Clos' results have been generalized to systems that support connections with varying bandwidth requirements. These generalizations have extended the application of Clos networks well beyond their original technological context and have led to a number of interesting new results, especially in connection with systems that support multicast communication.

1. Introduction

In 1953, Charles Clos [1] showed that the class of switching networks that now bears his name was immune to the phenomenon of *blocking* that was the key performance limitation of the electromechanical telephone switching systems of that era. This was the first class of networks with sub-quadratic complexity to exhibit *nonblocking* performance. Clos' seminal paper sparked the development of the theory of interconnection networks and the *Clos networks* still maintain a central role in the design of practical switching systems in applications ranging from telephone switching, to digital cross connects, to video production switchers, to IP routers.

That fact that the Clos networks have maintained their relevance in the face of 50 years of rapid technological change reflects the fundamental nature of the performance benefits they confer. What is perhaps more remarkable than the continuing relevance of the networks is the enduring role of the theory that developed around them. Although that theory was developed to address the performance issues of telephone switches constructed from bulky electro-mechanical relays (called *crosspoints*), it retained its applicability through more than thirty years of technology development, including the introduction of large-scale integrated circuits and time-division digital telephone switches. While the original cost metric of *crosspoint count* lost much of its direct relevance in the integrated circuits era, the performance results of the original theory could be directly applied to the new technology through a simple transformation from modern time-division switch designs to the classical space-division switches.

A more fundamental departure from the original framework was required with the introduction of integrated services networks and packet switching technology. Networks supporting integrated services had to consider connections with different bandwidth requirements, including voice, data, video and multimedia traffic streams. The broadband nature of many of these new services also had the effect of

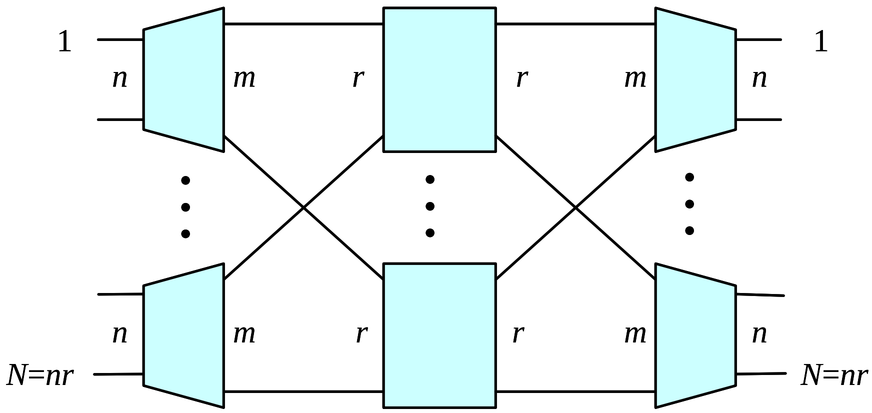

Figure 1. Symmetric three stage Clos network, $C(n,m,r)$

ruling out ad-hoc adaptations of existing networks (such as ISDN), making necessary a rethinking of the technical foundations of networks. Initially both multirate circuit switching and connection-oriented packet switching were considered as potential technologies for broadband networks, but connection-oriented packet switching, in the form of Asynchronous Transfer Mode (ATM) technology was soon chosen as the most promising technical solution. The development of a theory of multirate interconnection networks stemmed from the necessity of modelling a new generation of switches for broadband services. While the first studies were motivated by multirate circuit switches [3], the advent of ATM, re-oriented the research towards packet switching.

This paper describes how the classical theory of nonblocking networks was generalized to accommodate multirate switching systems, with a particularly emphasis on the application of the new theory to Clos networks. What is perhaps most remarkable about the results reviewed here is their similarity to the classical results of the 1950s and '60s. While 50 years of technological development has given us systems with many millions of times the capability of those early telephone switches, the theoretical foundations have proved remarkably resilient. This is a testament to the enduring nature of Charles Clos' insightful observations and the utility of the theory those observations spawned.

2. Elementary Multirate Nonblocking Conditions for Clos Networks

A symmetrical 3-stage Clos network is shown in Figure 1. There are three key parameters for this network: the number of switch modules in the first and third stages, the number of switch modules in the middle stage and the number of inputs (outputs) to the first (third) stage switch modules. These parameters are commonly denoted by r, m and n respectively and completely characterize the network. We use the notation $C(n,m,r)$ to denote such a network and we let $N=nr$ be the number of network inputs (and outputs). While we shall focus our attention in this paper on three stage networks, Clos networks with 5, 7 or more stages are also possible and most of the results described can be extended to cover these cases.

In the classical theory of interconnection networks, a network is said to be *strictly nonblocking* if, there is no configuration of connections that can prevent the addition of a new connection between an idle input and an idle output. In 1953, Charles Clos showed that if $m>2(n-1)$, $C(n,m,r)$ is strictly nonblocking. The reason for this is that any first stage switch with an idle input has at most $n-1$ busy links connecting it to the middle stage, so from any idle input, there are at most, $n-1$ "unreachable" middle stage modules. When $m>2(n-1)$, this is fewer than half the number of the middle stage switch modules. Similarly, fewer than half of the middle stage switch modules are unreachable from any idle output, meaning that there must be some middle stage switch module that can be reached from both sides. What made Clos' observation important was that it showed, for the first time, that one could construct strictly nonblocking networks with less than quadratic complexity. If one chooses $n=r=(n/2)^{1/2}$ and $m=2(n-1)$ the complexity of $C(n,m,r)$ (measured in terms of *crosspoint count*) is just under $6N^{3/2}$, compared to N^2 for a single stage switch.

The development of time-division telephone switches in the 1970s was a major technological advance, allowing the data paths within switching systems to be "time-shared" among a large number of voice circuits. Important as time-division switching was, it had little impact on the theory of interconnection networks, since the different "timeslots" in time-division switches could be treated as parallel space-division switches, allowing earlier results to be re-interpreted in the new technological context. In the 1980s, the growing interest in integrated voice, data and video communication led researchers to consider *multirate* circuit switches in which multiple timeslots in a time-division switch could be used for a single application. This made things more interesting, since in this context an application session could now be blocked if there was no path through the switch with the number of timeslots needed by the application. The varying needs of different applications created more complex sets of conditions that could cause blocking.

In the late 1980s and early 1990s, Niestegge [2] and Jajszczyk [3] derived nonblocking conditions for multirate time-division switching systems based on a three stage Clos network. In such systems, each of the links has k distinct timeslots, meaning that it can carry k circuits operating at a basic rate (typically 64 Kb/s) or any equivalent combination of circuits, so long as the total number of timeslots required by all the circuits sharing a given link is no more than k. Niestegge and Jajszczyk showed that if we constrain circuits to use no more than B timeslots, then no circuit can be blocked so long as

$$m > 2\left\lfloor \frac{nk-B}{k-B+1} \right\rfloor$$

The argument that leads to this result is a natural generalization of Clos' original argument. Any first stage switch module that can accommodate a new session using w timeslots has at most $nk-w$ timeslots that are in use on the links connecting it to switch modules in the second stage. A link cannot be used by a new circuit with w timeslots only if at least $k-w+1$ of its timeslots are already in use. This means that for any first stage switch that that can accommodate a new circuit with w timeslots there are at most $\lfloor (nk-w)/(k-w+1) \rfloor$ links connecting it to middle stage switches that are "too busy" to accommodate a new circuit using w timeslots. If m satisfies the inequality above, then this number is fewer than half the number of middle stage switches. By the same argument, fewer than half of the middle stage switches are unreachable from any output that can accommodate a new connection using w timeslots. Hence, there must be some middle stage switch that is reachable from both sides. Note that if $B=1$, the bound reduces to the original bound derived by Clos. On the other hand, if we allow single circuits that use $k/2$ timeslots, the number of middle stage switches needed to guarantee nonblocking operation doubles.

At the same time that some researchers were studying the potential of multirate synchronous time-division switching, others were studying an alternative form of switching that came to be known as *Asynchronous Transfer Mode* or ATM. ATM switching is a simplified form of packet switching, in which short fixed length packets are labeled with a *virtual circuit identifier*. The so-called *labeled multiplexing* scheme used in ATM allows more flexible transmission rates than is possible in synchronous switching. In particular, there is no intrinsic minimum rate in ATM, and rates need not be multiples of a fixed minimum rate. In addition, ATM switches are often designed with internal data paths that operate at higher speeds than the externals links. The ratio of the internal data rate to the external rate is referred to as the speed advantage or *speedup* of the system.

In 1989, Melen and Turner [4] generalized the model for multirate switching to accommodate ATM switching systems. They explicitly introduced the notion of speedup for multirate switching systems and derived nonblocking conditions for the multirate Clos network and other networks. The results for

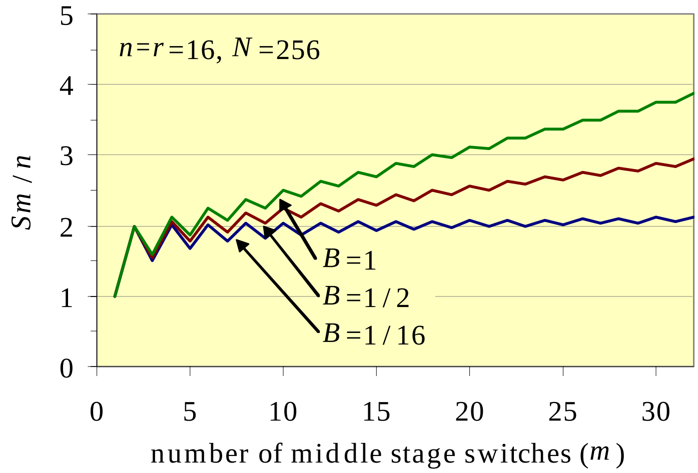

Figure 2. Cost tradeoff for $C(n,m,r)$

$C(n,m,r)$ were later improved by Chung and Ross [5] and by Liew et. al. [6]. If there is no lower bound on the rate of individual circuits, then this bound can be expressed as

$$m > 2\left\lceil \frac{n-B}{S-B} - 1 \right\rceil$$

where S is the speedup of the internal links and B is the maximum virtual circuit rate, which is expressed as a fraction of the bandwidth of the external links (so, $0<B\leq 1$). The argument is similar to the one used by Niestegge. The amount of bandwidth in use at any first stage switch that can accommodate a new virtual circuit at rate w is at most $n-w$. In order for a link connecting the first stage switch to a middle stage switch to be too busy to accept a new connection of bandwidth w, the amount of traffic on the link must be greater than $S-w$. Thus, the number of inaccessible middle stage switches must be strictly less than $(n-w)/(S-w)$. If the stated nonblocking condition holds, then this number is less than half the number of middle stage switches, meaning that blocking cannot occur. If there is a lower bound b on the virtual circuit rate and if $b>S-B$, then the switch is nonblocking if $m > 2\lfloor (n-b)/b \rfloor$. Note that if $b=B=1$, this reduces to the original bound derived by Clos. We note that Kabacinski and Liotopoulos [7] have generalized these results to asymmetric Clos networks with non-homogeneous link speeds.

It's also interesting to note that since the cost of a multirate Clos network is proportional to S and to m/n, there is a tradeoff between m/n and S, as shown in Figure 2. Increasing m allows us to decrease S and vice versa, and the lowest cost is obtained when $m=1$ and $S=n$. Practical constraints often prevent the use of a single very fast middle stage switch, but there is clear advantage to using the largest speedup that is practical, especially when individual connections can use a large fraction of the capacity of the external links.

3. Wide-Sense and Rearrangeably Nonblocking Variants

Melen and Turner noted in [4] that in systems with $b=0$ and $S=B=1$, the multirate nonblocking condition for the Clos networks implies that the number of middle stage switches required for the Clos network to be strictly nonblocking is infinite. That is, there is no strictly nonblocking Clos network under these conditions. Indeed, we can make the much stronger statement that under these conditions every strictly nonblocking multirate network has a cost that grows as the product of the number of network inputs to the number of network outputs. This holds because under the given conditions, a connection with a very small bandwidth is sufficient to block a new connection with a bandwidth of B joining an input x to an

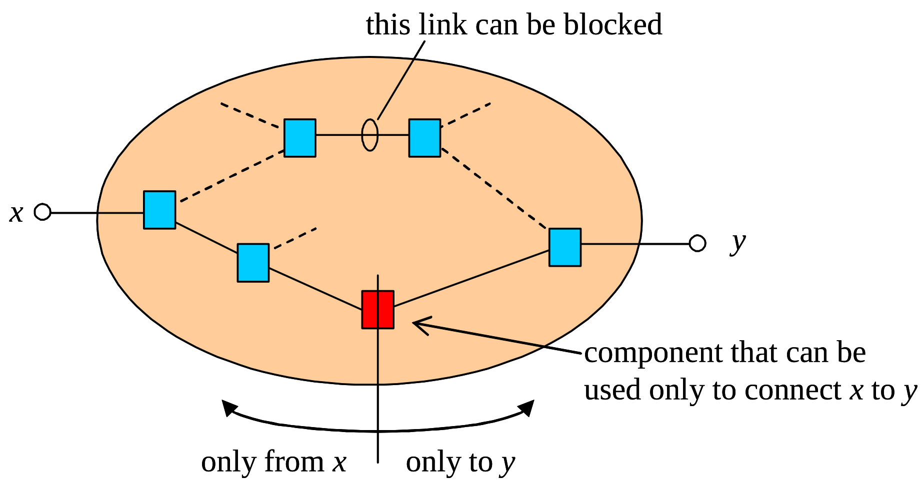

Figure 3. Quadratic complexity of strictly nonblocking mulitrate networks

output *y*. This means that any link that can be reached by an input other than *x* and an output other than *y* can easily become too busy to accommodate the new connection. Consequently, any path from *x* to *y* must consist of one or more links reachable only from *x* followed by one or more links that are reachable only from *y*. At the boundary between these two path segments, there is a portion of a switch module that can be used only for connections between *x* and *y* (see Figure 3). So, for the network to be nonblocking, there must be such a dedicated path segment for every (input, output) pair.

In spite of this, there are Clos networks that can be operated in a nonblocking fashion even when $b=0$ and $S=B=1$. The key to this result is to carefully select the routes used by connections to avoid the situations that can cause blocking. A network in which blocking can be avoided through the judicious selection of routes is called a *wide-sense nonblocking* network. In this case, the key idea is to divide connections into two subsets according to their bandwidth and segregate the routes used by these two subsets. Melen and Turner [4] showed that if the middle stage switches were divided into one group that carries only *small* connections (those with rates no more than 0.5) and another that carries only *large* connections (those with rates greater than 0.5), that blocking could be avoided if the number of middle stage switches is at least $8n$. Chung and Ross [5] showed that $6n$ middle stage switches were sufficient. A more sophisticated routing algorithm, called the *quota scheme*, was proposed by Gao and Hwang [8] in 1997. In its simplest form, it consists of letting large connections use any middle stage switch, while routing small connections only through a subset of them. In [8] it was shown that, by first dividing the connections into two separate sets, as in [4 and 6], and then applying the quota scheme to the small ones, the number of middle stage switches needed to ensure nonblocking operation could be reduced to $5.75n$. The possibility of a small further improvement was noted in [7].

In many practical contexts, the cost of strictly nonblocking and wide-sense nonblocking networks exceeds the practical benefits they confer. In such situations, switch designers are often interested in whether a given network is *rearrangeably nonblocking*. In the classical theory of interconnection networks, a network is said to be rearrangeably nonblocking if it is always possible to add a new connection linking an idle input to an idle output, by rearranging existing connections. Or equivalently, a rearrangeably nonblocking network is one where any set of compatible connection requests can be routed, all at once, rather than one-at-a-time. While rearrangement is rarely used in practice, this fundamental topological property is considered an important figure of merit that generally contributes to good performance.

In the late fifties, Slepian [9] showed that three stage Clos networks were rearrangeably nonblocking so long as $m \geq n$. This can be shown most simply be reformulating the problem of routing a set of

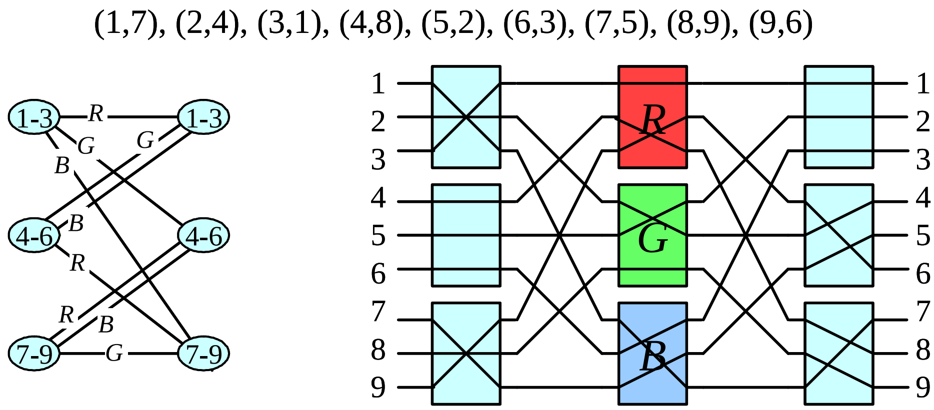

Figure 4. Simultaneous connection routing in $C(n,n,r)$ using graph edge coloring

connections through a Clos network as a graph edge coloring problem, as illustrated in Figure 4. The graph used to determine the set of routes has one vertex for every first stage switch in the network and one vertex for every third stage switch in the network. An edge is added between a vertex in the first set and a vertex in the second set for every connection that needs to be routed from an input of the switch corresponding to the first vertex to an output of the switch corresponding to the second vertex. This is illustrated in Figure 4, where the (input, output) pairs at the top are used to construct the graph shown at left. Given this graph, the next step is to assign *colors* to the edges, in such a way that no two edges incident to the same first stage switch are assigned the same color. The colors assigned to the edges correspond to the middle stage switches used to carry the connections and the constraint on the colors corresponds to the constraint that no two connections in the same first (or third) stage switch can pass through the same middle stage switch. In Figure 4, the letters R, G and B denote the colors assigned to the edges and the diagram at right shows the corresponding set of routes. By a classical result from graph theory, the graph coloring can be done using no more than n colors, implying that the given set of connections can be routed using at most n middle stage switches.

Slepian's result was extended to the multirate environment by Melen and Turner in [4]. Their generalization uses the same bipartite graph to represent the set of connections, but now the edges of the bipartite graph are assigned *weights* between 0 and 1 that represent the bandwidth used by each of the connections. As before, we want to color the edges of the graph, but now we allow edges incident to the same vertex to have the same color, so long as the total weight of the edges incident to a single vertex does not exceed a specified constant S, corresponding to the speedup of the switch. This is illustrated in Figure 5, where a set of multirate connections is listed at the top and the corresponding weighted graph is shown at left. This weighted graph coloring problem can be converted to an ordinary graph coloring problem by *splitting* each of the vertices and associating different subsets of the edges incident to the vertex with different "sub-vertices". In particular, the n heaviest edges are all assigned to the same sub-vertex, the next n heaviest edges are assigned to another sub-vertex, and so forth. This is illustrated in the middle part of Figure 5. When this splitting procedure has been applied to all vertices, the resulting graph has at most n edges incident to each vertex and so can be colored in the ordinary way (only one edge of each color) using just n colors. The coloring is shown in the middle part of Figure 5 and some of the corresponding routes are shown at the right. The assignment of edges to the sub-vertices ensures that the sum of the weights of like colored edges from the same original vertex is less than $B+(n–B)/m$. Consequently, the system is rearrangeably nonblocking if $S \geq B+(n–B)/m$ or if $m \geq \lceil (n-B)/(S-B) \rceil$. Note the similarity to the previous result for the strictly nonblocking case. Note also that when $m=n$, a speedup of about $1+B$ is sufficient to make the network rearrangeably nonblocking. Melen and Turner also studied the case of Clos networks with more than three stages and showed that when $m=n$, these networks are also

(1,7,.5), (1,4,.5), (2,3,.7), (3,1,.4), (3,6,.3), (4,8,1), (5,2,.8), (6,3,.3), (7,5,.5), (8,9,.9), (9,6,.6)

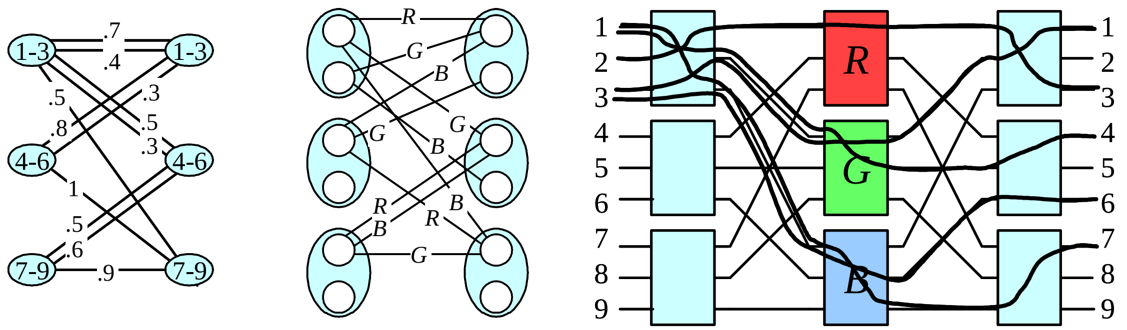

Figure 5. Simultaneous routing of multirate connections in $C(n,n,r)$

rearrangeably nonblocking if the speedup is allowed to grow logarithmically with the number of stages. It remains an open question whether a constant speedup can suffice for networks with an unbounded number of stages.

4. Routing Multicast Connections

Up to this point, we have considered only point-to-point or unicast connections. It's also interesting to consider the case of *multicast* in which connections can be made from an input to a set of outputs. Routes in multicast switches form trees, rooted at the inputs from which they originate. In the classical theory of interconnection networks, these trees must be edge disjoint. In the extension to multirate switching, the trees have an associated weight (representing the bandwidth of the connection) and are permitted to share links, so long as the sum of the weights on any given link does not exceed its capacity.

Masson and Jordan [10] derived conditions under which a three stage Clos network is nonblocking for multicast for single rate networks. When specialized to symmetric networks, their nonblocking condition can be written $m > (f_{\max}+1)(n-1)$, where $f_{\max}$ is the maximum number of distinct third stage switches with outputs in the same multicast connection (for unrestricted multicast $f_{\max}=r$). When generalized to multirate networks, the condition becomes

$$m > \left\lceil F\frac{n-B}{S-B} - 1 \right\rceil + \left\lceil \frac{n-B}{S-B} - 1 \right\rceil$$

where $F = \min\{f_{\max}, m\}$. The inclusion of m in the definition of F allows us to account for systems where a large speedup is used to allow m to be smaller than $f_{\max}$. Note that when $f_{\max}=1$, the result reduces to the standard unicast nonblocking condition. For systems that must handle large multicasts, the result is of limited practical use, because the nonblocking multicast network is usually impractically expensive (roughly $(r+1)/2$ times as expensive as the nonblocking unicast network if multicasts are permitted to reach all outputs and $m \geq r$). However the result does point the way to more practical results. In [11], Yang and Masson showed that a much less expensive network could provide a limited form of nonblocking operation if a restriction is placed on the amount of branching a multicast connection can have in the first stage. Specifically, they showed that it is always possible to add a new multicast connection to $C(n,m,r)$ if multicast branching in the *first stage* is limited to no more than f and

$$m > (f + r^{1/f})(n-1)$$

By choosing f to make the right-hand side as small as possible, one obtains a nonblocking network for which the cost may not be impractically large. For example, for $C(16,m,16)$, the best choice of f is 3, giving $m>82.5$, which results in a network whose cost is 2.7 times as large as for the nonblocking unicast network. However, it must be noted that the nonblocking property obtained in this way is less than ideal. While we can always add a new multicast connection, we may not be able to add a new output to a pre-existing multicast connection without rerouting the connection. According to the usual classification of nonblocking networks, this network is only rearrangeably nonblocking. However, since the required rearrangement is limited to the one connection being extended, this is a somewhat less onerous form of rearrangement. We call a multicast network in which it is always possible to add a new multicast connection, *reroutably nonblocking*, since in these networks any existing connection can be extended if we allow it to be re-routed. Zegura [12] extended Yang and Masson's result to the multirate context. Here, the nonblocking condition becomes

$$m > \left\lceil f \frac{n-B}{S-B} - 1 \right\rceil + \left\lceil \frac{n-B}{S-B} - 1 \right\rceil r^{1/f}$$

Later, Kim and Du [13] combined Yang and Masson's approach with the quota scheme [8] to obtain a wide-sense nonblocking multicast network.

Another way to obtain a reroutably nonblocking multicast network is to cascade two Clos networks, connecting the outputs of the first to the inputs of the second [14]. In this system, "branching" of multicast connections is limited to the second network. The first network is used to route connections to an entry point of the second network from which the multicast connection can be completed. Specifically, when routing a new multicast connection we route the connection through the *most lightly loaded* first stage switch in the second network. For single rate networks, the resulting nonblocking condition is $m>2(n-1)$. In the multirate case, the network is reroutably nonblocking if

$$m > \left\lceil \frac{n-B/r}{S-B} - 1 \right\rceil + \left\lceil \frac{n-B}{S-B} - 1 \right\rceil$$

and $S \geq 1+(1-1/rn)B$. This bound is just slightly larger than the bound that applies to the unicast case, showing that the cost of a reroutably nonblocking multicast network is only twice as large as the cost of a strictly nonblocking unicast network. A similar result can be obtained for a system with a single Clos network in which multicast connections are routed in *two passes* through the network, with multicast branching occurring only during the second pass. In systems where multicast is a relatively small part of the total traffic, this can be far more economical, since the extra system capacity required to make the system nonblocking for multicast is roughly the same as the fraction of the total outgoing traffic which belongs to multicast connections.

5. Some Practical Considerations

The results described here are based on a particular abstraction of multirate switching that applies to varying degrees in different system contexts. The key assumption is that the rates of different connections can simply be added together. This is completely appropriate for multirate circuit switches and for ATM switches carrying *constant bit rate traffic*. It has certain limitations when applied in contexts where connections have time varying rates. In such systems, there is not even a simple answer to the question of what the rate is. In situations where connections have a pre-specified *peak rate* (which is very common) one can take the connection rate to be the peak rate, but this can lead to under-utilization and is typically not an acceptable choice. On the other hand, in most system contexts, the connection peak rates are much smaller than the link capacities. In data networks for example, most applications have peak rates less than 100 Mb/s, which is just 4% of an OC-48 link and 1% of an OC-192. In such situations the connection rate can be taken to be the *average rate,* with a small safety factor added. Of course, this applies only if the connection's average rate is known in advance, or can be estimated with acceptable accuracy.

The problem of deciding when a set of connections with time varying rates can be multiplexed together with acceptable performance has been studied extensively. These studies have led to the

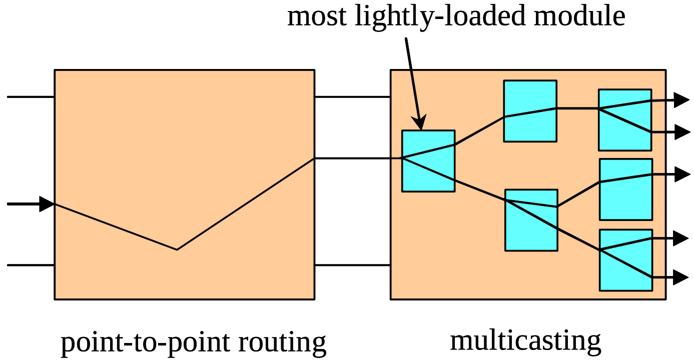

Figure 6 Adding multicast connection to cascaded Clos networks

formalization of the concept of an *effective rate* for a time varying connection. Effective rates can be added together and compared to a fixed capacity bound, in order to make connection admission and routing decisions. As one might expect, a connection's effective rate lies between its average and its peak, and is close to the average when the peak rate is much smaller than the link capacity, and close to the peak when the peak is a large fraction of the link rate. The results described here can be applied using effective rates, but there are some subtleties that can affect the accuracy of the results. Specifically, because the effective rate is a function of the link capacity, a connection's effective rate is smaller in systems with a large speedup than in systems with a more modest speedup. This effect tends to make the nonblocking conditions discussed above conservative, in the sense that they overestimate how large m must be to obtain nonblocking performance.

Another practical concern is how to evaluate the cost of different system configurations. In the classical theory of nonblocking networks, the cost metric was the *crosspoint* count, which measured the number of elementary switching components that were needed to implement a switch module. The crosspoint count of a switch module is simply the product of the number of inputs and outputs it has. This metric is completely appropriate for systems constructed from electromechanical switches, where the crosspoints dominate the system cost, but it is not appropriate for systems constructed using integrated circuits with tens of millions of logic gates. In such systems, there are no crosspoints in the classical sense, but each switch component does still contain subsystems for which the cost grows in proportion to the product of the number of inputs and outputs. However, each component also contains subsystems for which the cost grows linearly with the number of inputs and outputs. This leads to a somewhat more complex, but still useful cost metric. The key issue when applying it, is to appropriately assign relative weights to the two components of the cost. For modern integrated circuits, the linear component has a much larger weight, but the quadratic component can still become important as the number of inputs and outputs of single switch modules gets large. The relative weights of these two components are inherently technology-dependent and must be adjusted with continuing advances in technology. Coppo et al. [15] has developed a generalized cost metric for ATM switching systems and used it to compare alternative Clos network configurations.

While the lack of a simple, fixed cost metric is a little disturbing, from a theoretical standpoint, it's important to recognize that one can still draw fairly broad conclusions about the effect of key system parameters on cost. For example, in $C(n,m,r)$ one can still conclude that in any given technology, doubling m will double overall system cost, since it doubles the number of middle stage switches, the crosspoint counts of the first and third stage switches, the number of outputs of the first stage switches and the number of inputs of the third stage switches. Similarly, in any given technology, doubling S will double the cost, since the larger speedup typically must be obtained through increased parallelism.

6. Summary

The development of the theory of multirate interconnection networks has attracted the interest of researchers both for its intrinsic intellectual appeal and its practical application to the development of broadband switching systems. In the late 1980s and 1990s, most major telecommunication equipment manufacturers developed ATM switching systems based on the intellectual framework provided by this theory.

The recent shift in the telecommunications marketplace from ATM to IP technology has also been accompanied by a shift in how switching systems are designed. Most modern high performance routers are now implemented with a single high bandwidth switch surrounded by multiplexing stages that collect traffic from a large number of external links. This architecture can be viewed as a degenerate case of a Clos network with m=1 and as has been noted above, this is the best possible choice when the technology is available to move all the traffic in the system through a single switching stage. The rapid growth of the Internet has also led router vendors to develop multistage designs to allow the construction of systems with capacities exceeding 10 Tb/s. Current multistage router designs mostly use *dynamic routing* in which traffic is distributed through the multistage topology on a packet-by-packet basis, rather than a connection basis. This better fits the best-effort nature of IP networks, but makes it more difficult to deliver consistently high quality of service, to applications that require it. As the Internet's role in modern life continues to grow, one can expect quality of service to become a more central concern for IP router vendors, possibly leading to renewed interest in systems that route on the basis of connections (or "flows") for those application sessions that require consistently high performance.

While the winds of technological change are hard to predict with any certainty, the enduring role of telecommunications networks in our modern information society and the role of switching systems in those networks seems secure. We can also expect that the Clos networks will retain their central place in the design of high performance switching systems of all kinds, and that the intellectual framework created to model their performance will continue to develop and evolve to meet the needs of new technologies and applications.

References

[1] C. Clos: "A Study of Non-Blocking Switching Networks," *Bell System Technical Journal*, March 1953, pp.406-424.

[2] G. Niestegge: "Nonblocking Multirate Switching Networks," *Proc. 5th ITC Seminar*, Como Lake, Italy, May 1987.

[3] Jajszczyk: "On Nonblocking Switching Networks Composed of Digital Symmetrical Matrices," *IEEE Trans. on Communications*, Vol. COM-31, No.1, Jan.1983, pp.2-9.

[4] R. Melen, J. S. Turner: "Nonblocking Multirate Networks," *SIAM J. Comput.*, Vol.18, No.2, April 1989, pp.301-313.

[5] S.-P. Chung, K. Ross: "On Nonblocking Multirate Interconnection Networks," *SIAM J. Comput.*, Vol.20, No.4, Aug. 1991, pp.726-736.

[6] S. Liew, M.-H. Ng, C. Chan: "Blocking and Nonblocking Multirate Clos Switching Networks," *IEEE/ACM Trans. on Networking*, Vol.6, No.3, June 1998, pp.307-318.

[7] W. Kabacinski, F. Liotopoulos: "Multirate Non-blocking Generalized Three-Stage Clos Switching Networks," *IEEE Trans. on Communications*, Vol. 50, No.9, Sept.2002, pp.1486-1494.

[8] B. Gao, F. K. Hwang: "Wide-sense Nonblocking for Multirate 3-stage Clos Networks," *Theor. Comp. Science*, Vol.182, 1997, pp.171-182.

[9] V. E. Benes: "Mathematical Theory of Connecting Networks and Telephone Traffic," Academic Press, New York, 1965.

[10] G. Masson, B. Jordan: “Generalized Multistage Connection Networks,” *Networks*, Vol.2, 1972, pp.191-209.

[11] Y.Yang, G.Masson: “Nonblocking Broadcast Switching Networks,” *IEEE Trans. on Computers*, Vol.40, No.9, Sept.1991, pp.1005-1015.

[12] E. Zegura: *Design and Analysis of Practical Switching Systems*, Doctoral Dissertation, Washington University Computer Science Department, 6/93.

[13] D. Kim, D.-Z. Du: “Multirate Broadcast Switching Networks Nonblocking in a Wide Sense,” *Advances in Switching Networks, DIMACS Series in Discrete Mathematics and Theoretical Computer Science*, Vol 42, July 1998, pp. 59-74.

[14] R. Melen, J. S. Turner: “Nonblocking Multirate Distribution Networks,” *IEEE Trans. on Communications*, Vol.41, No.2, Feb.1993, pp.362-369.

[15] P. Coppo, M. D'Ambrosio, R. Melen: “Optimal Cost/Performance Design of ATM Switches,” *IEEE Trans. on Networking*, Vol.1, No.5, Oct.1993, pp.566-575.