Segmented Channel Routing Is Nearly as Efficient as Channel Routing (and Just as Hard)

Abbas El Gamal
Information Systems Laboratory
Stanford University
Stanford, CA 94305

Jonathan Greene
Actel Corporation
955 E. Arques Avenue
Sunnyvale, CA 94086

Vwani Roychowdhury
Information Systems Laboratory
Stanford University
Stanford, CA 94305

Abstract

Novel problems concerning the design and routing for the segmented routing channel, are introduced. These problems are fundamental to the design and design automation for Field Programmable Gate Arrays (FPGAs), a new type of electrically programmable VLSI. The problems introduced may also be applicable to configurable multi-processors. The paper demonstrates using a probabilistic model for connections that a segmented channel with judiciously chosen segment lengths can be nearly as efficient (in terms of the number of tracks required) as a freely customized routing channel. This is corroborated by experimental data from actual designs. The paper also presents the first known theoretical results on the combinatorial complexity and algorithm design for segmented channel routing. In particular it is shown that the segmented channel routing problem is in general NP-Complete, and that efficient polynomial time algorithms can be designed for a number of important special cases.

1 Introduction

Conventional channel routing [1] concerns the assignment of a set of connections to tracks within a rectangular region. The tracks are freely customized by the appropriate mask layers. Even though the channel routing problem is in general NP-Complete [4], efficient heuristic algorithms exist. In a channeled gate array, which is customized by metal masks, the routing channels are of fixed width and routing must be completed within the allotted number of tracks. The channel width is selected such that routing can be successfully completed for most connection instances likely to occur in practice. Using probabilistic models for connections (e.g., [11]) the channel width is chosen so that successful routing is achieved with high probability.

In this paper we consider the more restricted channel routing problem where the routing is constrained to use fixed wiring segments of predetermined lengths and positions within the routing channel. Such segmented channels have been incorporated in Field Programmable Gate Arrays (FPGAs) [2, 3], a new type of electrically customized VLSI used to drastically reduce the time and cost for prototyping/implementing application-specific integrated circuits. We investigate the routing problem for segmented channels and show using a probabilistic model for connections that with judiciously chosen segment lengths the channel width needed to achieve high probability of routing completion is not much greater than that for a comparable size gate array. This result is further corroborated by experimental data from actual designs. Thus leading us to believe that the integration limit for FPGAs is nearly as high as that for mask programmed gate arrays, a truly surprising finding. We also investigate the computational complexity of and algorithms for the segmented routing problem for given segmented channels. Our results show that although the problem is in general NP-complete, several special cases admit polynomial time algorithms.

1.1 FPGA Architecture

The FPGA architecture of [3] is much like that of conventional (mask programmed) gate arrays: comprising rows of cells (logic modules) separated by segmented routing channels (Fig. 1). The inputs and outputs of the modules each connect to a dedicated vertical segment. Programmable switches are located at each crossing of vertical and horizontal segments and also between pairs of adjacent horizontal segments in the same track. By programming a switch, a low resistance path is created between the two crossing or adjoining segments.

Different logic circuits are implemented in an FPGA by assigning the gates to logic modules (placement) and then connecting them via programmed switches and segments (routing). A typical example of that is shown in Fig. 1. The vertical segment connected to the output of module 3 is connected by a programmed switch to a horizontal segment, which in turn is connected to the input of module 4 through another programmed switch. In order to reach the inputs of modules 1 and 2, two adjacent horizontal segments are connected to form a longer one.

The choice of the wiring segment lengths in a segmented channel is driven by tradeoffs involving the number of tracks, the resistance of the switches, and the capacitances of the segments. These tradeoffs are illustrated in Fig. 3.

Fig. 3A shows a set of connections to be routed. With the complete freedom to configure the wiring afforded by mask programming, the *left edge* algorithm [5] will always find a routing using a number of tracks equal to the density of the connections (Fig. 3B). This is the case since there are no 'vertical constraints' in the problems we consider.

In an FPGA, achieving this complete freedom would require switches at every cross point. Furthermore, switches would be needed between each two cross points along a wiring track so that the track could be subdivided into segments of arbitrary length (Fig. 3C). Since all present technologies offer switches with significant resistance and capacitance, this would cause unacceptable delays through the routing. Another alternative would be to provide a number of continuous tracks large enough to accommodate all nets (Fig. 3D). Though the resistance is limited, the capacitance problem is only compounded, and the area is excessive.

A segmented routing channel offers an intermediate approach. The tracks are divided into segments of varying lengths (Fig. 3E), allowing each connection to be routed using a single segment of the appropriate size. Greater routing flexibility is obtained by allowing limited numbers of adjacent segments in the same track to be joined end-to-end by switches (Fig. 3F). Enforcement of simple limits on the number of segments joined or their total length guarantees that the delay will not be unduly increased. Our results apply to the models of Figs. 3E and 3F.

The segmented channel routing scheme may also be considered as a model for a communication network in a multi-processor architecture. The logic modules in Fig 1. can be replaced by Processing Elements (PEs); the segmented routing network can then be used for dynamically reconfiguring interconnections among the PEs (by programming the appropriate switches as described for the FPGAs). In [8] a preliminary network model that uses specially segmented channels (referred to as express channels) has already been proposed. Tradeoffs similar to those discussed above also appear to hold for such multi-processor communication networks.

1.2 Definitions and Summary of Results

A segmented channel routing problem, as depicted in Fig. 2, comprises a set $\mathcal{C}$, of M connections and a set $\mathcal{T}$, of T tracks. The tracks are numbered from 1 to T. Each track extends from column 1 to column N, and is divided into a set of contiguous segments separated by switches. The switches are placed between two consecutive columns.

For each segment s, we define $left(s)$ and $right(s)$ to be the leftmost and rightmost column in which the segment is present, $1 \leq left(s) \leq right(s) \leq N$. Each connection c_i, $1 \leq i \leq M$, is characterized by its left-most and right-most column: $left(c_i)$ and $right(c_i)$. Without loss of generality, we assume throughout that the connections have been sorted so that $left(c_i) \leq left(c_j)$ for $i < j$.

A connection c may be *assigned* to a track t, in which case the segments in track t that are present in the columns spanned by the connection are considered *occupied.* More precisely, a segment s in track t is occupied by the connection c if $right(s) \geq left(c)$ and $left(s) \leq right(c)$. In Fig. 2 for example, connection c_3 would occupy segments s_{21} and s_{22} in track 2 or segment s_{31} in track 3.

Definition 1 A routing, R, of a set of connections consists of an assignment of each connection to a track such that no segment is occupied by more than one connection.
A K**-segment** routing is a routing that satisfies the additional requirement that each connection occupies at most K segments.

We can now define the following *routing* problems:

Problem 1 **[Unlimited Segment Routing]** Given a set of connections and a segmented channel, find a routing.

For technological reasons, mentioned above, it may be desirable to limit the number of segments used for each connection.

Problem 2 **[K-Segment Routing]** Given a set of connections and a segmented channel, find a K-segment routing.

It is often desirable to determine a routing that is optimal with respect to some criterion. We may thus specify a weight $w(c,\ t)$ for the assignment of connection c to track t, and define:

Problem 3 **[Optimal Routing]** Given a set of connections and a segmented channel, find a routing which assigns each connection c_i to a track t_i such that $\sum_{i=1}^{M} w(c_i, t_i)$ is minimized.

For example, a reasonable choice for $w(c,t)$ would be the sum of the lengths of the segments occupied when connection c is assigned to track t. Note also that with appropriate choice of $w(c,t)$, Problem 3 subsumes Problem 2.

The problems defined above address the issue of routing connections using *given* segmented-channels. Next we address the issue of designing such segmented channels. In a conventional routing channel the wiring segments are designed for the given set of connections using the left edge algorithm (when no vertical constraints are present). The number of tracks required (i.e. channel width) equals the channel density, i.e. the maximum number of connections present in any column. Therefore, the width of a gate array channel must be at least the density for most connection instances likely to occur in practice. The channel width can be estimated [11] using a probabilistic model for the connection lengths and their positions in the channel to model the likely connection instances. In the segmented routing channels considered in this paper, the segment lengths as well as the channel width must be chosen such that a routing can be found for most connection instances likely to occur. We assume a probability distribution $p(c)$ over the set of all possible connections, $\tilde{\mathcal{C}}$, and define the following:

Problem 4 [**Segmentation Design Problem**]:
Given: M, N, K, $\epsilon > 0$, and a probability distribution $p(c)$ over the set of all possible connections, $\vec{C}\vec{C}$, with $1 \leq left(c) \leq right(c) \leq N$.
Find: a segmented channel with the minimum number of tracks T such that with probability at least $1-\epsilon$ there is a K-segment routing for M connections drawn randomly according to $p(c)$.

In this paper we establish the following results.

Theorem 1 Determining a solution to Problem 1 is strongly NP-complete.

Theorem 2 Determining a solution to Problem 2 is strongly NP-complete even when $K = 2$.

The reductions used to prove these theorems are rather tricky, and may have applications to problems in the area of task-scheduling on non-uniform processors. A proof of Theorem 1 is presented in Section 2, and a proof of Theorem 2 is given in [10].

Although Theorems 1 and 2 show that the routing problem is in general NP-complete, several special cases of the problem are tractable. We have developed polynomial-time algorithms for the following special cases:
Identically Segmented Tracks: The *left edge* algorithm used for conventional channel routing can be applied to solve Problems 1, 2, and 3.
1-Segment Routing: A routing can be determined by a linear time ($O(MT)$) greedy algorithm that exploits the geometry of the problem.

The connections are assigned in order of increasing left ends as follows. For each connection, find the set of tracks in which the connection would occupy one segment. Eliminate any tracks where this segment is already occupied. From among the remaining tracks, choose one where the occupied segment's right end is farthest to the left, and assign the connection to it. In the example of Fig. 2, the algorithm assigns c_1 to s_{11}, c_2 to s_{21}, c_3 to s_{31}, c_4 to s_{32}, and c_5 to s_{13}. It can be shown that if a connection cannot be assigned to any track, then no complete routing is possible. The time required is $O(MT)$.

The corresponding optimization problem (Problem 3) can be also solved in polynomial time by reducing it to a weighted maximum bipartite matching problem.
At most 2-Segments Per Track: If each track is segmented into at most two segments then also a greedy linear time algorithm (similar to the one for 1-Segment routing) can be designed to determine a routing.

We have also developed a general $O(T!M)$-time algorithm using dynamic programming for solving Problems 1, 2, and 3. This general algorithm can be adapted to yield more efficient algorithms for the following cases:
Fixed Number of Tracks: If the number of tracks is fixed then the general algorithm directly yields a polynomial time algorithm.
K-segment Routing: The general algorithm can be modified to yield an $O((K+1)^T M)$-time algorithm. Note that for small values of K the modified algorithm performs better than the general one.

Fixed types of Tracks: If the number of tracks is *unbounded* but the tracks are chosen from a fixed set, where T_i is the number of tracks of type i, then an $O((\prod_{1}^{l} T_i^{K+2})M)$ time (hence, a polynomial-time) algorithm can be designed.
Furthermore, we have developed a heuristic algorithm based on linear programming for solving Problems 1 and 2 that appears to work surprisingly well in practice.

The general algorithm and the above mentioned special cases are described in [9, 10].

Assume that the connection lengths are selected independently from an exponential distribution and the left end points of the connections are chosen uniformly from N columns; for an exact description of the probability distribution, $p(c)$, see Section 3. Anyway, in Section 3, using the above mentioned $p(c)$ we prove the following result which shows that asymptotically the width of a segmented routing channel is within a small constant factor of that for a conventional gate array:

Theorem 3 For any N, M, $0 < \rho < 1$, $0 < \alpha < 1$, and $\beta > 1$, a segmented channel can be designed such that with probability $O(N^{-\rho})$ there is no 1-segment routing of M connections chosen independently. The number of tracks in the channel is $[C(1+\alpha)\beta]\ N^{\rho}M/N$, where $C \approx 2.5$.

The proof of Theorem 3 (see Section 3) is constructive, and may be generalized to a large class of distributions. Such results give theoretical justification for the favorable experimental results concerning 1- and 2-segment routing discussed in Section 4.

2 Complexity of the Segmented Routing Problem

In this section we prove Theorem 1, i.e. determining a solution to Problem 1 is strongly NP-complete. The proof of Theorem 2, i.e. determining a solution to Problem 2 is strongly NP-complete even when $K = 2$, is presented in the Appendix. The NP-Complete reductions for both the theorems is from the *Numerical Matching Problem with Target Sums*, which has been shown to be strongly NP-Complete [7].

Numerical Matching with Target Sums [7]: Given a set $S = \{1, \cdots, n\}$, and positive integers $x_1, \cdots, x_n,\ y_1, \cdots, y_n,\ z_1, \cdots, z_n$ with $\sum_{i \in S}(x_i + y_i) = \sum_{i \in S} z_i$, do there exist permutations α and β of S such that $x_{\alpha(i)} + y_{\beta(i)} = z_i$, for all $i \in S$?

We assume without loss of generality that $x_1 < x_2 < \cdots < x_n$, $y_1 < y_2 < \cdots < y_n$, and $z_1 < z_2 < \cdots < z_n$. Furthermore, we assume that for any instance of the problem, we have $x_{i+1} - x_i \geq n$ and $x_1 + y_1 \geq x_n + n$. If these conditions are not met for an instance of the problem then one can define an equivalent problem (i.e. the modified problem has a solution if and only

if the original problem has a solution) for which the conditions are met by performing the following transformations:
1. *Scaling:* Define $m = \lceil n/min(x_i - x_{i-1}) \rceil$. If $m > 1$ then set $x_i \leftarrow mx_i$, $y_i \leftarrow my_i$, and $z_i \leftarrow mz_i$.
2. *Translation:* Define $p = x_n + n - (y_1 + x_1)$. If $p > 0$ then set $y_i \leftarrow y_i + p$, and $z_i \leftarrow z_i + p$.

Given an instance of the Numerical matching problem $\mathcal{N}$, we now show how to construct an instance of the segmented routing problem $\mathcal{Q}$ in polynomial time.
The set of connections $\mathcal{C}$ is defined as follows:
1. For each x_i we define a connection a_i such that left(a_i)= 4, right(a_i)= $x_i + 3$. Thus, each connection a_i is of length $x_i - 1$, and starts at column number 4.
2. For each y_k, we define n connections $b_{k1}, \cdots, b_{kn}$ (one for each x_j) such that left(b_{kj})= $x_j + 4 + (n-k)$ and right(b_{kj})= $(y_k + x_j) + 4$.
3. n connections $d_1, \cdots, d_n$ are defined with left(d_i)= 1, and right(d_i)=3.
4. $n^2 - n$ connections $e_1, \cdots, e_{n^2-n}$ are defined with left(e_i) = 1, and right(e_i) = 5.
5. n^2 connections $f_1, \cdots, f_{n^2}$ are defined with left(f_i)= $x_n + y_n + 5$ and right(f_i)= $x_n + y_n + 7$.

Set the number of columns in the construction to $N = x_n + y_n + 7$.
The set $\mathcal{T}$ of n^2 tracks are then defined as follows:
1. For the first n tracks $t_1, \cdots, t_n$ each track t_i begins with a segment $(1, 3)$ followed by unit length segments that span the region from column 4 to column $z_i + 4$ (*i.e.* there is a switch between every two columns between column 4 and column $z_i + 4$), followed by a single segment of the form $(z_i + 5, N)$.
2. The rest of the $n^2 - n$ tracks are best described by dividing them into n blocks, each consisting of $n - 1$ tracks. Each such track comprises 3 segments.

The first block of $n - 1$ tracks, i.e. tracks $t_{n+1}, t_{n+2}, \cdots, t_{2n-1}$, are constructed using the definitions of the connections b_{1j}, $1 \leq j \leq n$. The segments in the track t_{n+1} are (1, left(b_{11}) $-$ 1), (left(b_{11}), right(b_{12})), and (right(b_{12}) $+ 1, N$). That is, the middle segment in the track t_{n+1} is defined such that the connections b_{11} or b_{12} can be assigned to it. In general, the segments in each track t_{n+j}, $1 \leq j \leq n-1$, are defined as (1, left(b_{1j}) $-$ 1), (left(b_{1j}), right($b_{1(j+1)}$)), and (right($b_{1(j+1)}$) $+ 1, N$). That is, the middle segment in the track t_{n+j}, $1 \leq i \leq n-1$, is designed such that the connections b_{1j} or $b_{1(j+1)}$ can be assigned to it.

The i^{th} block of $n - 1$ tracks (i.e. tracks $t_{n+(i-1)(n-1)+1}, \cdots, t_{n+i(n-1)}$) is constructed using the definitions of the connections b_{ij}, $1 \leq j \leq n$. The segments in the track $t_{(n+(n-1)(i-1))+j}$ (i.e. the j^{th} track in the i^{th} block) are (1, left(b_{ij}) $-$ 1), (left(b_{ij}), right($b_{i(j+1)}$)), and (right($b_{i(j+1)}$) $+ 1, N$). That is the middle segment in the track $t_{(n+(n-1)(i-1))+j}$ is designed such that the connections b_{ij} or $b_{i(j+1)}$ can be assigned to it.

The following example illustrates this construction.

Example 1: Consider the segmented channel routing problem (see Fig. 4) corresponding to the instance of the Numerical matching problem with Target Sums:
$x_1 = 2,\ x_2 = 5,\ x_3 = 8,\quad y_1 = 9,\ y_2 = 11,\ y_3 = 12$ and
$z_1 = 11,\ z_2 = 17,\ z_3 = 19.$ □.

We might note here that our proof of the NP-Complete reduction is geometric in nature and it is helpful to use the above example in understanding the statement and the proof of each of the following propositions and lemmas. Before we proceed, however, let us define the following:

Two connections c_1 and c_2 will be said to *overlap* if they are present in the same column(s) i.e. $left(c_2) \leq left(c_1) \leq right(c_2)$ or $left(c_1) \leq left(c_2) \leq right(c_2)$.

A connection c_1 is said to *fit* in a segment S_1 if $left(c_1) \geq left(S_1)$ and $right(c_1) \leq right(S_1)$.

A segment is said to be *available* for a set of connections if it is unoccupied by the rest of the connections in $\mathcal{C}$.

Proposition 1 In any routing R of $\mathcal{Q}$:
(a) the connections f_i, $1 \leq i \leq n^2$, are assigned to n^2 different tracks.
(b) the connections d_i, $1 \leq i \leq n$, and a_i, $1 \leq i \leq n$, are assigned to tracks $t_1, \cdots, t_n$, and connections e_i, $1 \leq i \leq n^2 - n$, are assigned to tracks t_{n+1} through t_{n^2}.

Proof: Claim (a) follows directly from the construction; *i.e.* , the connections f_i, $1 \leq i \leq n^2$ are all identical and overlapping.
Claim (b) follows from the following observations that are based on the above construction:
1. Each connection e_i, $1 \leq i \leq (n^2 - n)$, overlaps with every other e_i. Each e_i also overlaps with every connection d_j, $1 \leq j \leq n$ and every connection a_k, $1 \leq k \leq n$.
2. In tracks t_1 through t_n a d_i and a_j can be assigned to the same track, and such assignment is not possible for tracks t_i, where $i > n$.
3. Finally, it follows from 1., 2. and (by the pigeon-hole principle) that if any e_i is assigned to a track t_j, $j \leq n$, then there would not be a sufficient number of tracks so as to assign all the connections d_i, $1 \leq i \leq n$, a_j, $1 \leq j \leq n$, and e_k, $1 \leq k \leq n^2 - n$. □

Proposition 2 In any routing R of $\mathcal{Q}$, the segments available for assigning the connections a_i, $1 \leq i \leq n$, and b_{ij}, $1 \leq i, j \leq n$ are as follows:
(a) In any track t_i, $1 \leq i \leq n$, the segments in columns 4 through $z_i + 4$ (*i.e.* , the portion that is fully segmented) are available.
(b) In any track t_i, $n + 1 \leq i \leq n^2$, only the middle segment is available.

Proof: Follows from Proposition 1: (a) In any track t_i, $1 \leq i \leq n$, the first segment is always occupied by a d_j (for some $1 \leq j \leq n$), and the last segment is occupied by an f_k. Hence, the only available portion is the fully segmented part of the track. (b) Every track t_i, $n + 1 \leq i \leq n^2$, has only three segments, and from Proposition 1 we know that the left segment is occupied by a connection e_j (for some $1 \leq j \leq n^2$) and the right segment by another connection f_k, $1 \leq k \leq n^2$. □

The following proposition shows that in any routing R of $\mathcal{Q}$, every track has exactly one b_{ij} assigned to it.

Proposition 3 All connections b_{ij}, $1 \leq i, j \leq n$, overlap; hence they have to be assigned to different tracks.

Proof: Given the geometry of our construction, it suffices to show that b_{11} and b_{1n} overlap. Now right(b_{11})$= x_1 + y_1 + 4$, and left(b_{1n})$= x_n + y_1 + n + 3$. Hence, right(b_{1n})$-$left(b_{11}) $= x_1 + y_1 - (x_n + n - 1)$ which is strictly greater than 0 by our assumptions. □
We can now show one direction of the reduction procedure.

Lemma 1 If the given Numerical Matching problem with target sums has a solution then there exists a routing for the segmented channel problem $\mathcal{Q}$.

Proof: Suppose there exist permutations α and β such that $x_{\alpha(i)} + y_{\beta(i)} = z_i$ for all $1 \leq i \leq n$. Then we can define a routing as follows:
1. Connections d_i, $1 \leq i \leq n$, e_i, $1 \leq i \leq n^2 - n$, and f_i, $1 \leq i \leq n^2$, are assigned according to Proposition 1.
2. For every i, $1 \leq i \leq n$, connections $a_{\alpha(i)}$ and $b_{\beta(i)\alpha(i)}$ are assigned to track t_i. Since $x_{\alpha(i)} + y_{\beta(i)} = z_i$, one can easily show that the connections can be appropriately assigned in the available segments (also see Proposition 2).
At this stage, for every i, $1 \leq i \leq n$, all except one connection among the connections b_{ij}, $1 \leq j \leq n$, need to be routed.
3. Consider the connections b_{1j}, $1 \leq j \leq n$. Let b_{1k} be the connection that has been assigned to one of the tracks t_i, $1 \leq i \leq n$. Recall that the tracks t_{n+1} through $t_{n+(n-1)}$ were designed using the definitions of b_{1j}, $1 \leq j \leq n$, and that the middle segment in track t_{n+1} can accommodate either connection b_{11} or connection b_{12}. So assign b_{11} to track t_{n+1} and repeat this procedure by assigning connections b_{12} through $b_{1(k-1)}$ to tracks t_{n+2} through $t_{n+(k-1)}$. Now, b_{1k} has already been assigned, hence one has to assign connections $b_{1(k+1)}$ through b_{1n}. By construction, however, $b_{1(k+1)}$ can be assigned to track t_{n+k}, and this assignment procedure can be continued by assigning $b_{1(k+2)}$ to track $t_{n+(k+1)}$, and so on.

In general for any i, the unassigned $n-1$ connections among b_{ij}, $1 \leq j \leq n$ can be assigned to the i^{th} block of tracks (i.e. tracks $t_{n+(i-1)(n-1)+1}, \cdots, t_{n+i(n-1)}$) by following the same procedure as above. □

Next we show that if the routing problem Q has a valid routing then there is a solution for the numerical matching problem $\mathcal{N}$. The following definitions that capture the geometry of the routing problem, Q, will be helpful:
It is clear from Propositions 1, 2, and 3 that each track t_l, $i \leq l \leq n$ has one connection from a_i, $1 \leq i \leq n$ and one connection from b_{kj}, $1 \leq k, j \leq n$

assigned to it. Also, note that since the parts of the first n tracks that are available for the connections a_i, $1 \leq i \leq n$, and b_{kj}, $1 \leq k, j \leq n$ are fully segmented, two connections a_i and b_{kj} can be assigned to the same track only if they do not overlap.

We define the *length* or *space* occupied by the connections a_i and b_{kj} assigned to some track t_l ($1 \leq l \leq n$) as equal to right(b_{kj})-left(a_i). That is, the length (or space) occupied by the two connections is the geometrical length from the left end of the connection a_i to the right end of the connection b_{kj}.

Note that it follows from Proposition 2 that the total length (or space) available in the first n tracks for assigning the connections a_i, $1 \leq i \leq n$ and b_{ij}, $1 \leq i, j \leq n$, is $\sum_1^n z_i$.

Proposition 4 Connections a_i and b_{kj} cannot be assigned to the same track if $j < i$.

Proof: left(b_{kj})$= x_j + 4 + (n-k)$, thus right(a_i)$-$left(b_{kj})$= x_i - (x_j + n - k)$. However, by our assumptions $x_i - (x_j + n - 1) > 0$, for all $j < i$. Hence, a_i and b_{kj} overlap for $j < i$. □

Proposition 5 If a_i and b_{kj} ($j \geq i$) are assigned to the same track t_l ($1 \leq l \leq n$), then the length occupied in the track t_l is $x_j + y_k$ ($\geq x_i + y_k$).

Proof: left(a_i)$= 4$, and right(b_{kj})$= x_j + y_k + 4$. Hence, right(b_{kj})$-$left(a_i)$=$ $x_j + y_k \geq x_i + y_k$ (because by our assumption $j \geq i$ implies that $x_j \geq x_i$). □

Proposition 6 None of the connections b_{kj} for $k > 1$ can be assigned to tracks t_{n+1} through $t_{n+(n-1)}$.

Proof: Recall that the tracks t_{n+1} through $t_{n+(n-1)}$ were constructed using the connections b_{1j}, $1 \leq j \leq n$. Now consider any track t_{n+l}. From Proposition 1 we know that its end segments are already occupied. Hence, for any b_{kj} to be assigned to this track it must fit within the middle segment (left(b_{1l}), right($b_{1(l+1)}$)).

First consider the case where $k > 1$ and $j \leq l$. Recall that left(b_{kj})$=$ $x_j + (n-k) + 4$; since $j \leq l$, we have $x_j \leq x_l$ and since $k > 1$, we can write left(b_{kj})$= x_j + (n-k) + 4 < x_l + (n-1) + 4 =$ left(b_{1l}). Hence, b_{kj} cannot be assigned to track t_{n+l}.

Next consider the case where $k > 1$ and $j > l$. Recall that right(b_{kj})$=$ $x_j + y_k + 4$; since $j \geq (l+1)$, we have $x_j \geq x_{l+1}$; furthermore, $k > 1$ implies that $y_k > y_1$. Hence, right(b_{kj})$= x_j + y_k + 4 > x_{l+1} + y_1 + 4 =$ right($b_{1(l+1)}$). Therefore, b_{kj} cannot be assigned to track t_{n+l}. □

Proposition 7 In general, none of the connections b_{kj} for $k > i$ can be assigned to tracks $t_{n+(n-1)(i-1)+1}$ through $t_{n+(n-1)i}$.

Proof: Recall that these tracks were constructed using the definitions of b_{ij}, $1 \leq j \leq n$. The proof then follows along the lines of the previous proposition. □

Let R be any routing of $\mathcal{Q}$, then we define m_i as follows:

$$m_i = |\{b_{ij} : 1 \leq j \leq n, \text{ and } b_{ij} \text{ is assigned to } t_l, \ 1 \leq l \leq n, \text{ in } R\}|.$$

In other words, m_i is the number of connections from the set $\{b_{i1}, b_{i_2}, \cdots, b_{in}\}$ that are assigned to the first n tracks (*i.e.* , $t_1, t_2, \cdots, t_n$). The following Propositions 8, 9, and 10 show that in any valid routing R of $\mathcal{Q}$ $m_i = 1$ for all $1 \leq i \leq n$.

Proposition 8 $\sum_1^k m_i \leq k, \forall\ 1 \leq k \leq n$ and $\sum_1^n m_i = n$.

Proof: Each track has exactly one connection b_{ij} (for some i and j) assigned to it. Hence, by definition $\sum_1^n m_i = n$.

To show that $\sum_1^k m_i \leq k$ for every $1 \leq k \leq n$, first consider $k = 1$. Suppose that $m_1 > 1$, then exactly $n - m_1$ connections from among the connections b_{1j}, $1 \leq j \leq n$ are assigned to tracks t_{n+1} through t_{n^2}. Even if all of them were assigned to tracks in the first block (*i.e.* , among $t_{n+1}, \cdots, t_{n+n-1}$), there would be $(m_1 - 1) \geq 1$ tracks in the block that are left unassigned. However, by Proposition 6, no connection b_{ij}, when $i > 1$, can be assigned to any track among $t_{n+1}, \cdots, t_{n+n-1}$. Thus, at least $(m_1 - 1)$ tracks among $t_{n+1}, \cdots, t_{n+n-1}$ have no connection b_{ij} assigned to it. This leads to a contradiction.

Using Proposition 7, the same arguments can be applied for any $k > 1$. That is for $k = 2$, one can show (using Proposition 7) that if $m_1 + m_2 > 2$, then some tracks among t_{n+1} through $t_{n+2(n-1)}$ do not have any connection b_{ij} assigned to them. □

Proposition 9 Let $w_1 < w_2 < \cdots < w_n$ be a sequence of positive integers and let non-negative integers m_i, $1 \leq i \leq n$, satisfy the following relations: $\sum_1^k m_i \leq k$, $\forall\ 1 \leq k \leq n$, and $\sum_1^n m_i = n$. Then $\sum_1^n m_i w_i > \sum_1^n w_i$ if and only if some of the m_i are 0.

Proof: First we observe that if there exists an $m_j > 1$, then there exists $l = m_j - 1$ distinct variables $m_{j_1}, \cdots, m_{j_l}$, such that all of them are 0 and $j_i < j$. If not, then one can easily show that $\sum_1^j m_i > j$, which is a contradiction. Thus, if any of the variables $m_i > 1$, then it always forces some m_k to equal 0 such that $k < i$. Hence, $\sum_1^n m_i w_i > \sum_1^n w_i$ if and only if some of the m_i are 0. □

Proposition 10 In any routing R, $m_i = 1\ \forall\ 1 \leq i \leq n$, *i.e.* , in every routing only one connection from the set $\{b_{i1}, \cdots, b_{in}\}$ is assigned to one of the first n tracks.

Proof: If a_i and b_{kj} are assigned to the same track then from Proposition 5 we know that the length occupied is $\geq x_i + y_k$. Now by definition, m_k connections from among b_{kj}, $1 \leq j \leq n$, appear in the first n tracks. Hence, the total length occupied by the connections a_i, $1 \leq i \leq n$, and the connections b_{ij}, $1 \leq i, j \leq n$ that are assigned in the first n tracks is $\geq \sum_1^n x_i + \sum_1^n m_k y_k$.

If at least one m_k is 0, then Proposition 9 implies $\sum_1^n m_k y_k > \sum_1^n y_k$ (because $y_1 < y_2 < \cdots y_n$). Hence, the total length occupied by the connections a_i and b_{ij} in the first n tracks is $> \sum_1^n x_i + \sum_1^n y_k = \sum_1^n z_i$. This leads to a contradiction because Proposition 2 shows that the total space available is equal to $\sum_1^n z_i$. Hence, $m_i = 1 \; \forall \; 1 \leq i \leq n$. □

Lemma 2 If there is a routing for the segmented channel problem $\mathcal{Q}$, then there exists a solution to $\mathcal{N}$.

Proof: Proposition 10 shows that for all i, $1 \leq i \leq n$, only one connection among $\{b_{i1}, \cdots, b_{in}\}$ is assigned to one of the first n tracks. By Proposition 5, if a_i and b_{kj} $(j \geq i)$ are assigned to the same track then the length occupied is $x_j + y_k$ $(\geq x_i + y_k)$. Hence, the total length occupied by the connections is $\geq \sum_1^n x_i + \sum_1^n y_k = \sum_1^n z_i$.

Claim 1. A connection a_i can only be assigned to the same track with some b_{ki}.

Proposition 4 shows that if a_i and b_{kj} are assigned to the same track then $j \geq i$. Now if a_i is matched with some b_{kj} and $j > i$, then the length occupied is $x_j + y_k > x_i + y_k$. Hence, the total length occupied by all the connections in the first n tracks is greater than $\sum_1^n x_i + \sum_1^n y_k = \sum_1^n z_i$. However, this leads to a contradiction since the total space available in the first n tracks is $\sum_1^n z_i$ (Proposition 2). Hence, a_i can only be assigned to the same track as some b_{ki}.

It follows then that if we define the connections assigned to track t_i, $1 \leq i \leq n$, as $a_{\alpha(i)}$ and $b_{\beta(i)\alpha(i)}$, then α and β are permutations of the set $\{1, \cdots, n\}$. Also by our convention, the total length occupied in track t_i by $a_{\alpha(i)}$ and $b_{\beta(i)\alpha(i)}$ is $= x_{\alpha(i)} + y_{\beta(i)}$.

Claim 2. $x_{\alpha(i)} + y_{\beta(i)} = z_i$.

Suppose this is not the case for some i, $1 \leq i \leq n$. Then it implies that in the track t_i, the length occupied by the connections $a_{\alpha(i)}$ and $b_{\beta(i)\alpha(i)}$ is $< z_i$. Now by Proposition 2, in any track t_k $(1 \leq j \leq n)$ the space available for assigning the connections a_i and b_{ij} is z_k. Hence, the length occupied by the connections a_i and the connections b_{ij} in the first n tracks is $< \sum_1^n z_i$. However, this leads to a contradiction because we showed that the length occupied is $\geq \sum_1^n z_i$.

Thus, the assignment of connections to the first n tracks defines permutations α and β such that $\forall\, i$, $x_{\alpha(i)} + y_{\beta(i)} = z_i$. □

Theorem 1 Determining a solution to Problem 1 is strongly NP-complete.

Proof: Follows from Lemmas 1 and 2. □

3 On the Segmentation Design Problem

Suppose that M connections are chosen randomly from a distribution with average length N^ρ for some ρ, $0 < \rho < 1$. The expectation of the total length of the connections is MN^ρ. Thus successful routing, even in the case of an unconstrained channel, is likely to require at least MN^ρ/N tracks.

In this section, we show how to construct a segmented channel for arbitrarily large N and M such that with probability approaching one there is a 1-segment routing of the M randomly chosen connections. The number of tracks in the channel exceeds MN^{ρ}/N by only a small constant factor.

To keep the mathematics simple, we assume that the channel is circular; *i.e.*, column 1 and column N are made adjacent. (Our results may easily be extended to a normal straight channel). We further assume that the connection lengths are selected independently with geometrically distributed lengths and left end points uniformly chosen from among the N columns. More precisely the M connections are chosen independently according to the distribution

$$p(c) = \frac{1}{N} \cdot \frac{\overline{\gamma}^{[right(c)-left(c)]}\gamma}{1-\overline{\gamma}^{N}}$$

where $\gamma = N^{-\rho}$, and $\overline{\gamma} = 1 - \gamma$.

The strategy for constructing the channel is to divide the possible connection lengths l into a number of ranges $L_j < l \leq U_j$ for various lower and upper bounds L_j and U_j. We then provide a dedicated set of tracks for connections with lengths in each range. Define the ranges of connection length as follows. Range $j = 0$ is defined by $L_0 = 0$ and $U_0 = N^{\rho}$. Range $j > 0$ is defined by $U_j = \exp(j)N^{\rho}$ and $L_j = U_j/e$. Since the lengths can be at most N, we need only consider ranges where $L_j < N$. Thus we limit j to $0 \leq j \leq j_{max}$, where $j_{max} = \lceil (1-\rho)\log N \rceil$.

Fig. 5 shows the tracks set aside for connections whose length l is in the range $L_j < l \leq U_j$. The columns are divided into $N/(\alpha U_j)$ groups of αU_j columns, for some constant $0 < \alpha < 1$. Each track is divided into segments of length $(1+\alpha)U_j$, and sets of m_j identically segmented tracks are placed in a staggered manner according to the column groups as shown. (The precise values of m_j and α are discussed below). The number of tracks is $m_j(1+\alpha)/\alpha$. It is apparent that if at most m_j of the connections in this length range have their left end in any particular column group, then a 1-segment routing of these connections exists.

Theorem 3 For any N, M, $0 < \rho < 1$, $0 < \alpha < 1$, and $\beta > 1$, the probability is $O(N^{-\rho})$ that there is no 1-segment routing of M connections chosen independently according to $p(c)$ in the above described channel. The number of tracks in the channel is $[C(1+\alpha)\beta]\ N^{\rho}M/N$, where $C \approx 2.5$.

Proof: For each range j, $0 \leq j \leq j_{max}$ and its corresponding bounds L_j and U_j, let us choose the parameter

$$m_j = max\{\frac{\log N}{\beta^{-1} - 1 + \log\beta}, \beta M(\alpha U_j/N)\exp(-L_j/N^{\rho})\}.$$

Next let us define q_j as the probability that a randomly chosen connection has its left end in a particular group of αU_j columns and has length l in the range $L_j < l \leq U_j$. Then

$$q_j = (\alpha U_j/N)\overline{\gamma}^{\,L_j}(1-\overline{\gamma}^{\,U_j-L_j})\ /\ (1-\overline{\gamma}^{\,N})\ <\ (\alpha U_j/N)\exp(-L_j/N^{\rho}).$$

Let the random variable X_j be the number of connections, out of all M chosen connections, that meet these requirements. Note that X_j is the sum of M independent Bernoulli random variables with parameter q. The probability that $X_j > m_j$ may be bounded as follows, using the Chernoff bound [12]:

$$Pr(X_j > m_j) < \exp(-sm_j)(q_j \exp(s) + (1 - q_j))^M$$

for any $s > 0$. Choosing s so that $\exp(s) = m(1 - q_j)/q_j(M - m)$, we have:

$$\begin{aligned} Pr(X_j > m_j) &< (qM/m_j)^{m_j}((1 - q_j)/(1 - m_j/M))^{M-m_j} \\ &\leq (q_j M/m_j)^{m_j} \exp(m_j - q_j M) \\ &\simeq exp(-m_j(\log(z_j) - 1 + 1/z_j)) \end{aligned}$$

where $z_j = m_j/(q_j M)$. Note that the quantity $(\log(z_j) - 1 + 1/z_j)$ is positive for $z_j > 1$ and increases monotonically with z_j. Since our choice of m_j guarantees that $z_j = m_j/q_j M \geq \beta > 1$, we have:

$$Pr(X_j > m_j) < \exp(-m_j(\log(\beta) - 1 + 1/\beta)).$$

Since our choice of m_j also guarantees that $m_j \geq \log(N)/(\log(\beta) - 1 + 1/\beta)$, we conclude that

$$Pr(X_j > m_j) < 1/N.$$

Thus the probability of running out of segments in one group of αU_j columns for connections with length l, $L_j < l \leq U_j$, is at most $1/N$. The probability of running out of tracks for connections with length l, $L_j < l \leq U_j$, in any of the $N/(\alpha U_j)$ column groups is at most $(1/N)(N/(\alpha U_j)) = 1/(\alpha U_j)$. Summing this for all ranges $0 \leq j \leq j_{max}$, we find that the probability of not being able to successfully route the M connections is at most

$$\sum_{j=0}^{j_{max}} \frac{1}{\alpha U_j} = (N^{-\rho}/\alpha) \sum_{j=0}^{j_{max}} e^{-j} \leq N^{-\rho}(e/(e-1))/\alpha$$

It remains to calculate the total number of tracks required by the scheme. Recall that there are $m_j(1 + \alpha)/\alpha$ tracks set aside for each length range. The number of tracks for range $j = 0$ is at most:

$$(1 + \alpha)\beta M N^{\rho-1} + O(\log N)$$

The number of tracks for each range $j > 0$ is at most:

$$(1 + \alpha)\beta M N^{\rho-1} \exp(j - e^{j-1}) + O(\log N)$$

The $O(\log N)$ term accounts for the possibility that the first term in the definition of m may be the maximum. Summing over j, $0 \leq j \leq j_{max}$ we have:

$$\begin{aligned} &(1 + \alpha)\beta M N^{\rho-1}(1 + \textstyle\sum_{j=1}^{j_{max}} \exp(j - e^{j-1})) + j_{max} O(\log N) \\ \leq\ &(1 + \alpha)\beta M N^{\rho-1}(1 + \textstyle\sum_{j=1}^{\infty} \exp(j - e^{j-1})) + O(\log^2 N) \end{aligned}$$

The series evaluates to approximately 1.5. □

If the connections are chosen according to $p(c)$ defined above then the expected density (which is the number of tracks required in conventional channel routing) is easily shown to be a constant factor greater than $N^{\rho}M/N$; hence, the constant factor penalty incurred in segmented routing could be much less than 2.5.

The method of constructing the segmentation is quite general, and the result may be extended to other distributions. It may also be possible to further reduce the constant C in Theorem 3 below 2.5. Moreover, by using the greedy algorithm for 1-segment routing (Section 3.1), routing can be done in the above designed channels in linear time.

4 Concluding Remarks

We have introduced novel problems concerning the design and routing for segmented channels. We also presented the first known theoretical results on the efficiency, algorithm-design, and combinatorial-complexity of segmented channel routing. In particular, we showed that: 1. The segmented channel routing problem is in general NP-Complete, 2. efficient polynomial time algorithms can be designed for several special cases, and 3. segmented channels with judiciously chosen segment lengths near the efficiency of conventional channels.

The theoretical results on the efficiency of segmented channels is also corroborated by experimental data involving actual designs as presented in Fig. 6. Channel segmentations were designed by a combination of trial-and-error and human judgement. Two segmentations, each with 32 tracks and 40 columns, were created. 'Segmentation 1' and 'Segmentation 2' are intended for 1- and 2-segment routing, respectively. They were each tuned to achieve the greatest likelihood of complete routing of randomly chosen sets of connections under the corresponding segment limitation. The distribution of connections was derived from actual placements of 510 channels from 34 designs, and gives the probability of occurrence as a function of the length and starting point of the connection. The segmentation designed for 1-segment routing does fairly well, and the segmentation designed for 2-segment routing uses only a few more tracks than the density (which is the minimum number of required tracks). For full details of the experimental results see [9].

There are several open issues in this new area of routing. For example, although we have developed efficient algorithms for many special cases of the segmented routing problem (as listed in Section 1.2), several other interesting cases are yet to be solved; following are some relevant ones: 1. channel length (N) is bounded; 2. connection lengths are bounded; and 3. connections are non-overlapping. Also, no analytic methods for designing segmented channels that will perform well for practical problems are known. Moreover, as mentioned in the introduction, the segmented routing model could be applied to configurable multi-processors; this area needs further investigation.

5 Acknowledgements

The work of Abbas El Gamal and Vwani Roychowdhury was partially supported under contract J-FBI-89-101.

References

[1] M. Lorenzetti and D. S. Baeder. Routing. Chapter 5 in *Physical Design Automation of VLSI Systems,* B. Preas and M. Lorenzetti, eds. Benjamin/Cummings, 1988.

[2] H. Hsieh, et al. A Second Generation User Programmable Gate Array. *Proc. Custom Integrated Circuits Conf.*, May 1987, pp. 515-521.

[3] A. El Gamal, J. Greene, J. Reyneri, E. Rogoyski, K. El-Ayat, and A. Mohsen. An Architecture for Electrically Configurable Gate Arrays. *IEEE J. Solid-State Circuits,* Vol. 24, No. 2, April, 1989, pp. 394-398.

[4] T. Szymanski. Dogleg Channel Routing is NP-Complete. *IEEE Trans. CAD,* CAD-4(1), pp. 31-41, Jan. 1985.

[5] A. Hashimoto, J. Stevens. Wire Routing by Optimizing Channel Assignment Within Large Apertures. *Proc. 8th IEEE Design Automation Workshop*, 1971.

[6] C. H. Papadimitrou and K. Steiglitz. Combinatorial Optimization: Algorithms and Complexity. Ptrentice Hall, Inc., New Jersey,1982.

[7] M. R. Garey and D. S. Johnson. Computers annd Intractability– A guide to the Theory of NP-Completeness W. H. Freeman and Company, New York, 1979.

[8] W. J. Dally. Express Cubes: Improving the Performance of the k-ary n-cube Interconnection Networks. MIT, VLSI Memo No. 89-564, October 1989.

[9] J. Greene, V. Roychowdhury, S. Kaptanoglu, and A. El Gamal. Segmented Channel Routing. *Proc. 27th ACM/IEEE Design Automation Conf.*, June 1990, pp. 567-572.

[10] A. El Gamal, J. Greene, and V. Roychowdhury. Segmented Channel Routing. submitted to IEEE Trans. on CAD of IC and Systems, 1990.

[11] A. El Gamal. Two Dimensional Stochastic Model for Interconnections in Master Slice Integrated Circuits. *IEEE Trans. on Circuits and Systems,* CAS-28, 127-138, February, 1981.

[12] R. Gallager. Information Theory and Reliable Communication. John Wiley Press, 1968, p. 127.

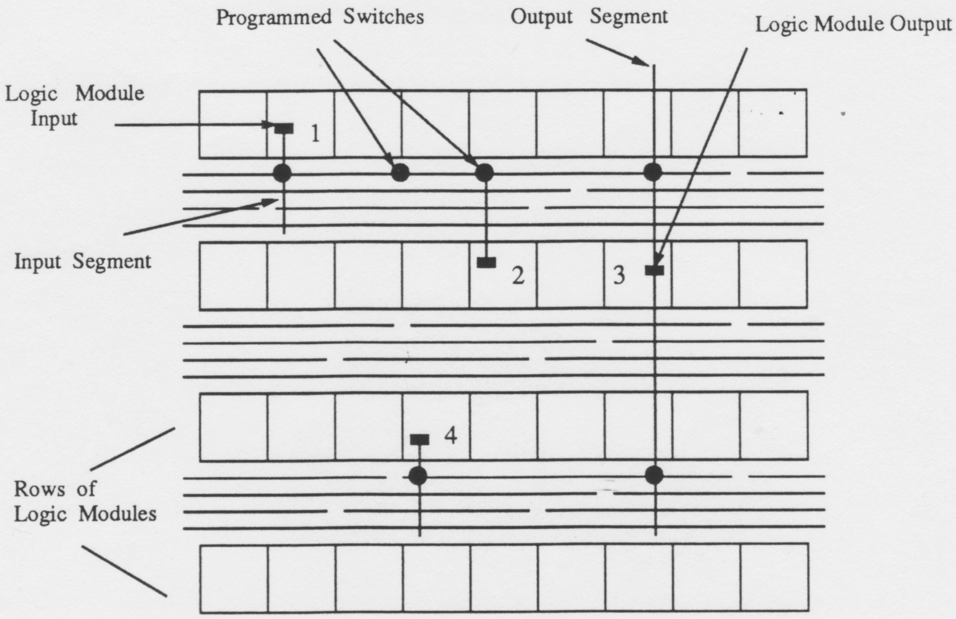


Figure 1: FPGA routing architecture. • denotes a programmed switch; unprogrammed switches are omitted for clarity.

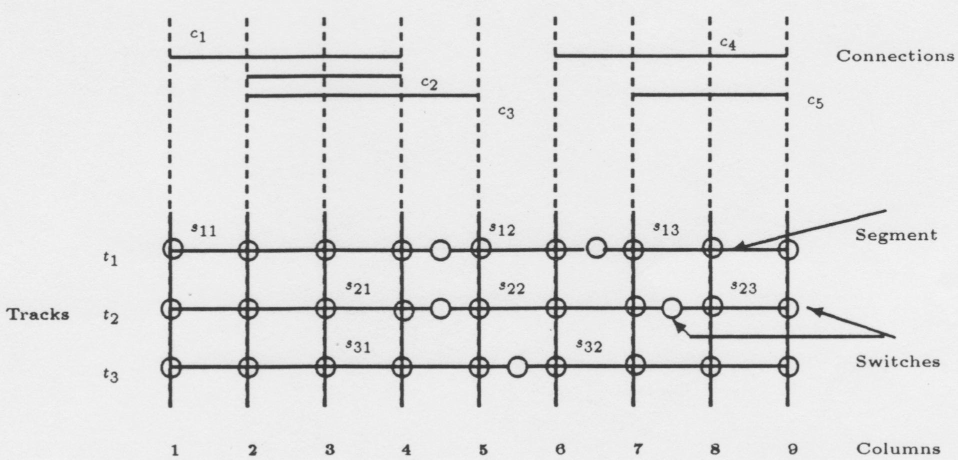
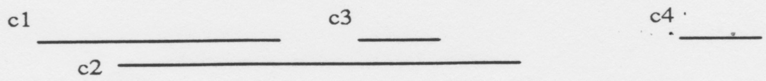
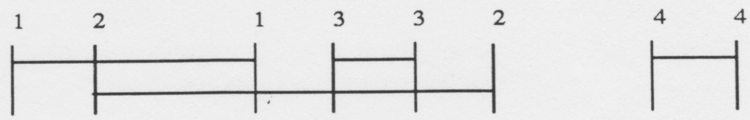


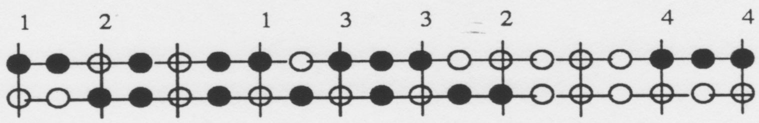
Figure 2: An instance of the segmented routing problem. $M = 5$, $T = 3$, $N = 8$. Connections: c_1, c_2, c_3, c_4, c_5. Segments: s_{11}, s_{12}, s_{13}, s_{21}, s_{22}, s_{23}, s_{31}, s_{32}.



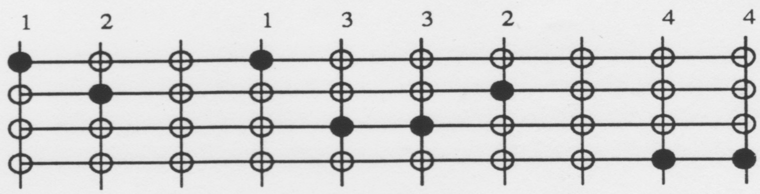
a. Set of connections to be routed.



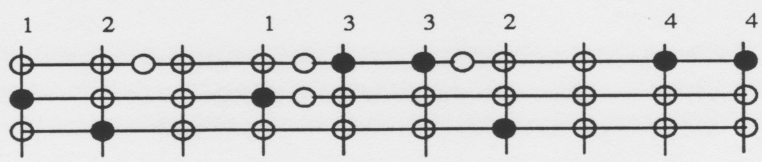
b. Routing in unconstrained channels.



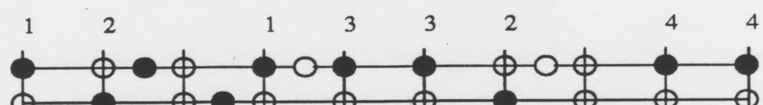
c. Routing in fully segmented channels.



d. Routing in unsegmented channel.



e. Segmented for 1-segment routing.



f. Segmented for 2-segment routing.

Figure 3: Examples of channel routing; ∘ denotes an open switch, and • a closed switch.

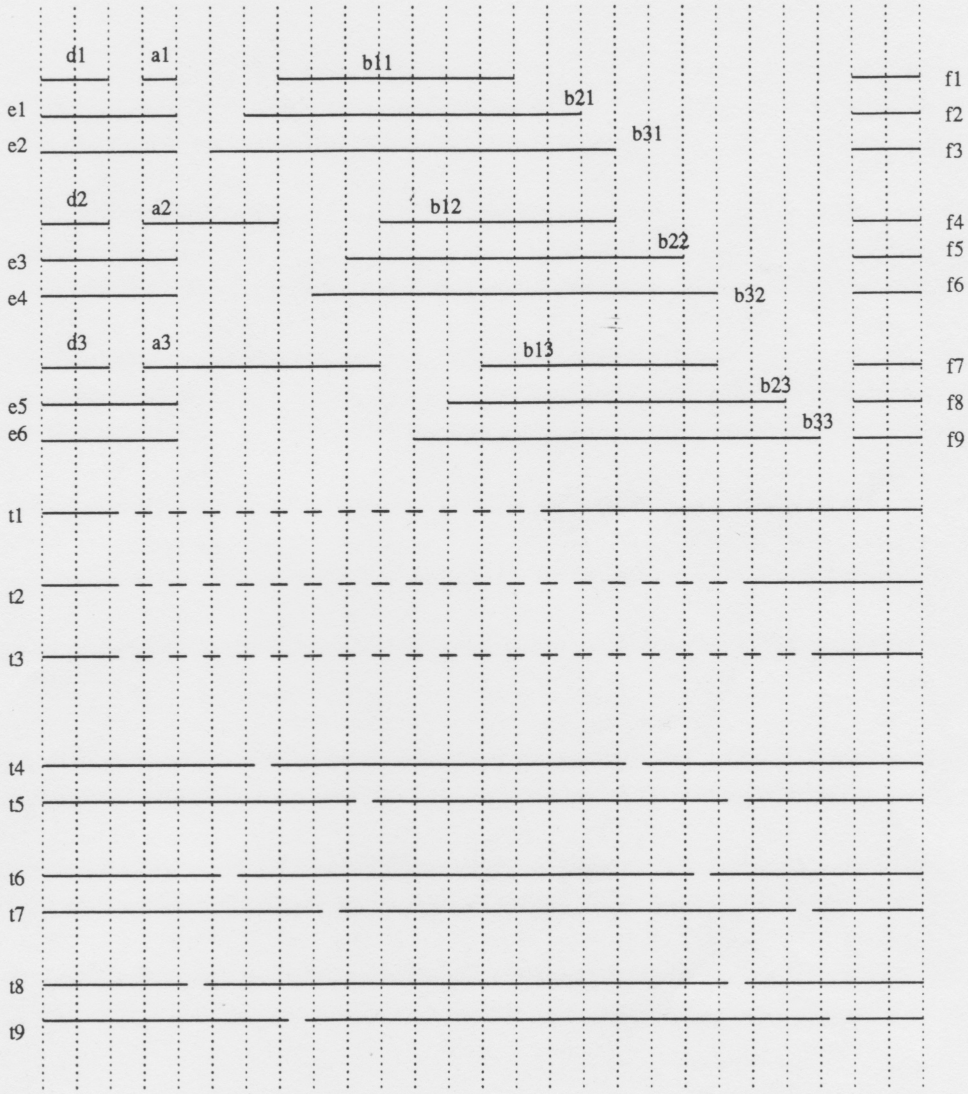


Figure 4: The segmented routing problem for Example 1.

Figure 5: Example of channel segmentation for Theorem 3. $\alpha = 1/2, m_j = 3$

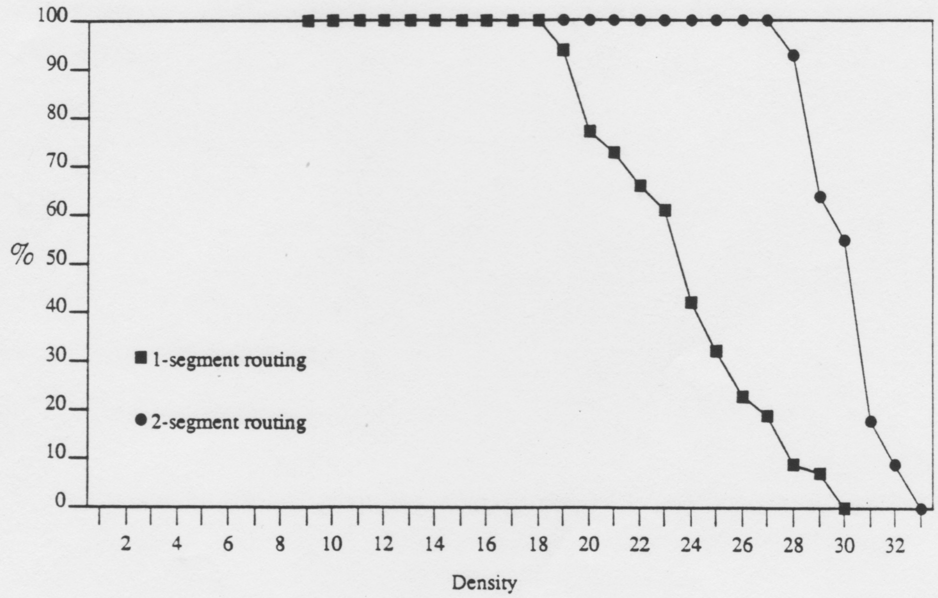


Figure 6: Probability of Routing Success vs. Density in a 32-track channel