3-dimensional Routing

András Recski
Department of Computer Science and Information Theory
Budapest University of Technology and Economics
H-1521 Budapest, Hungary
recski@cs.bme.hu

Dávid Szeszlér*
Department of Computer Science and Information Theory
Budapest University of Technology and Economics
H-1521 Budapest, Hungary
szeszler@cs.bme.hu

Abstract: Consider a single planar grid, or two parallel planar grids of size $w \times n$. The vertices of the grids are called terminals and pairwise disjoint sets of terminals are called nets. We aim at routing all nets in a cubic grid (above the single grid, or between the two grids holding the terminals) in a vertex-disjoint way. However, to ensure solvability, it is allowed to extend the length and the width of the original grid to $w' = sw$ and $n' = sn$ by introducing $s-1$ pieces of empty rows and columns between every two consecutive rows and columns containing the terminals. Hence the routing is to be realized in a cubic grid of size $(s \cdot n) \times (s \cdot w) \times h$. The objective is to minimize the height h.

The above problems are motivated by VLSI design, where technological improvements of the past two decades motivated the research of routing problems with a "real" 3-dimensional flavour. In this paper we survey a few recent results.

Keywords: VLSI design, 3-dimensional routing

1 Introduction

Traditionally, the detailed routing phase of the design of VLSI (Very Large Scale Integrated) circuits was considered as a **2-dimensional problem**, gradually extended to $2, 3, \ldots$ layers. Even within this problem, single row routing and channel routing are the better understood subproblems, where the terminals are placed on one side, or two opposite sides, respectively, of a rectangular circuit board, and the routing is to be realized on a few planar layers. Since the length n of the board is fixed by the row(s) of terminals and the number of layers k is fixed, the objective is to minimize the *width* w of the routing.

The specification of a (2-dimensional) single row routing or channel routing problem instance involves a family of pairwise disjoint subsets of the terminals, called *nets*. By a *routing* we mean an assignment of pairwise vertex-disjoint Steiner-trees in the 3-dimensional grid (of size $w \times n \times k$) to each net, such that the assigned tree connects the terminals of the corresponding net.

Thus in single row routing and channel routing the inputs are essentially one-dimensional (one or two lists of terminals) and the output is essentially two-dimensional (a fixed number of

*Research is supported by grant Nr. OTKA 67651 of the Hungarian National Science Fund

planar layers). However, as technology permits more and more layers, a "real" **3-dimensional** approach becomes reasonable. The research of 3-dimensional routing started in the 1980s and there are plenty of deep results in this area, see [1, 2, 4, 5, 6, 8, 9, 10, 15, 17], for example. However, most of these results embed certain "universal-purpose" graphs (like n-permuters, n-rearrangeable permutation networks, shuffle-exchange graphs) into the 3-dimensional grid, ensuring that *pairs* of terminals can be connected, moreover, in some papers along *edge-disjoint* paths.

The results surveyed in this paper allow *multiterminal nets* as well, and ensure *vertex disjoint* paths (or Steiner-trees) for the interconnections of the terminals within each net. We consider the 3-dimensional analogues of single row routing and channel routing. That is, terminals occupy certain vertices of a $w \times n$ planar grid (*single active layer routing problem*) or two parallel $w \times n$ grids (*3-dimensional channel routing*) and the third dimension (above the single grid or between the two grids) is for interconnections only. The objective is to minimize the height h of the routing.

One can easily see that a routing can be impossible with an arbitrary height (even in small instances, like a single active layer problem of size 4×1 or 2×2). Therefore it is allowed to extend the length and the width of the original grids containing the terminals to $w' = s_w \cdot w$ and $n' = s_n \cdot n$, respectively. s_w and s_n, called the *spacing*, are given constants. Extending the dimensions of the grids is done by introducing empty rows and columns into the (one or two) planar grids holding the terminals. Hence the final routing is realized in a cubic grid of size $(s_w \cdot w) \times (s_n \cdot n) \times h$.

As mentioned above, the results of this paper are motivated by the research of VLSI design, where 3-dimensional routing has been a much investigated topic in the past two decades. However, as it is very often the case with algorithms having provably good performance in VLSI routing, our constructions are not intended for designing specific, ready-to-use routing patterns for given specifications. Contrarily, we rather aim at understanding the nature of the problem, exploring the capacity of the 3D grid for solving such problems and thus helping heuristics to be designed and tested.

2 Basic Definitions

Assume that two parallel grids of size $w \times n$ (consisting of w rows and n columns) are given. The vertices of these grids are called *terminals*. A *net* N is a set of terminals. A *3-dimensional channel routing problem* (or *3DCRP* for short) is a set $\mathcal{N} = \{N_1, N_2, \ldots, N_t\}$ of pairwise disjoint nets. n and w are the *length* and the *width* of the routing problem, respectively. By a *single active layer routing problem* (or *SALRP* for short) we mean the special case of the 3DCRP where one of the two grids (say, the top grid) is empty (that is, nets do not contain terminals of the top grid).

By a *spacing of s_w in direction w* we are going to mean that we introduce $s_w - 1$ pieces of extra columns between every two consecutive columns (and also to the right hand side of the rightmost column) of the original grids. This way the width of the grids is extended to $w' = s_w \cdot w$. A *spacing of s_n in direction n* is defined analogously.

A *solution with a given spacing s_w and s_n* of a routing problem $\mathcal{N} = \{N_1, N_2, \ldots, N_t\}$ is a set $\mathcal{T} = \{T_1, T_2, \ldots, T_t\}$ of pairwise vertex-disjoint Steiner-trees in the cubic grid of size $(w \cdot s_w) \times (n \cdot s_n) \times h$ (between the two parallel planar grids containing the terminals) such that the terminal set of T_i is N_i for every $1 \leq i \leq t$. The Steiner-trees T_i are called *wires*. The *height* h of the routing is to be minimized.

3 Single Active Layer Routing

We start by two straightforward observations.

Lemma 1 *For any given n there exists a routing problem that cannot be solved with height h smaller than $\frac{n}{2s_w}$.*

PROOF: Let, for simplicity, the width and the length be even, let $w = 2a$ and $n = 2b$. Consider the following example (the idea is very similar to that in [3, 7]). Suppose that each net consists of two terminals in central-symmetric position as shown in Figure 1.

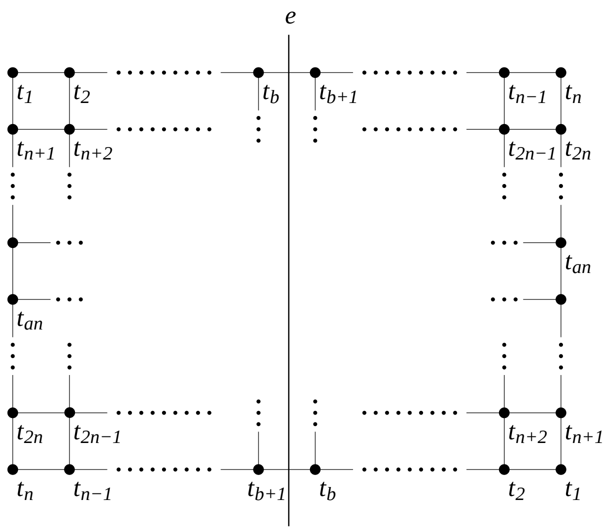

Figure 1:

The number of nets is an. Since each net is cut into two by the central vertical line e, any routing with width $w' = s_w \cdot w$ and height h must satisfy $w'h \geq an$. Therefore $h \geq (w/2w')n$, hence $h \geq \frac{n}{2s_w}$. $\square$

Lemma 2 *If $s_w \geq 2$ and $s_n \geq 2$ then every routing problem can be solved with height $h \leq \frac{wn}{2}$.*

PROOF: We assign a separate layer to each net. For every terminal we introduce a vertical (that is, parallel with the height) wire segment to connect the terminal with the layer of its net. The interconnection of the terminals of each net can now be performed trivially on its layer using the extra rows and columns guaranteed by the spacing in both directions.

Since 1-terminal nets can be disregarded, the number of nets is at most $\frac{1}{2}nw$ thus $h \leq \frac{wn}{2}$ follows immediately. $\square$

It seems that the nature of the single active layer routing problem depends fundamentally on whether the assumptions of the above lemma are true (that is, both $s_w \geq 2$ and $s_n \geq 2$) or not.

An alternative interpretation of Lemma 2 is that if we fix w then there is a routing of height $h = O(n)$, provided that $s_w, s_n \geq 2$. The truth of the same statement is not at all obvious in the $s_n = 1$ case. However, the following result shows that it is still essentially true.

Theorem 3 (A. Recski and D. Szeszlér, 2001)[12] *If $s_w \geq 8$ then for any fixed value of w and for any n a single active layer routing problem can always be solved in time $t = O(n)$ and with height $h = O(n)$ such that the length n is preserved or increased by at most one.*

We emphasize that the above result considers w as fixed and obtains a bound for the height as a function of n only. The proof of Theorem 3 is highly technical, it is based on decomposing the SALRP instance into $\binom{w}{2}$ (2-dimensional) channel routing instances. However, the proof involves an algorithm that gives $t = O(w^3 n)$ and $h = O(wn)$; details can be found in [12].

We will assume $s_w, s_n \geq 2$ in the sequel. Thus the trivial lower and upper bounds of Lemmata 1 and 2 provide the gap to be filled. For example, in the $n = w$ case we have the lower bound of $h = \Omega(n)$ in the worst case and the upper bound of $h = O(n^2)$. The first major step towards narrowing this gap was the following result, which was proved by elaborate probabilistic methods.

Theorem 4 (A. Aggarwal, J. Kleinberg and D. P. Williamson, 2000) [2] *If each net consists of two terminals only then the nets of an $n \times n$ SALRP can be partitioned into $O(n \log^2 n)$ classes such that each class of nets can be routed on a copy of the grid (of size $n \times n$).*

An easy corollary of this theorem is that if $s_w = s_n = 2$ and each net consists of two terminals only then every SALRP can be solved with height $h = O(n \log^2 n)$. To show this, we assign a layer to each partition class and we place the routings of the partition classes into the "new" rows and colums provided by the spacing (in both directions); now the terminals can be connected to the appropriate layer by a "long" vertical wire segment, and (at most) two additional 1-unit-long segments are needed for each terminal to connect this segment to the routing of its net.

Figure 2 illustrates the above idea. The nets of the SALRP with $w = n = 3$ on the left hand side are first partitioned into 2 classes and each class of nets is routed on a single layer. Then the two routings are combined to give a solution of the original problem with $s_w = s_n = 2$ and height $h = 2$. (This illustration also shows that this construction works for multiterminal nets as well.)

If we want to improve further on the upper bound of $h = O(n \log^2 n)$, we could try to reduce the necessary number of partition classes in Theorem 4; it is an open problem whether this is possible. However, the following result shows that we can still achieve a height of $h = O(n)$ if we allow two layers (intstead of one) for the routing of each partition class.

Theorem 5 (A. Recski and D. Szeszlér, 2007)[13] *Assume that a SALRP instance is given such that each net contains two terminals only. Then*

(1) the nets can be partitioned into $\lfloor \frac{3}{2}n \rfloor$ classes such that each class of nets can be routed as a separate SALRP with $s_n = \lceil \frac{w}{2n} \rceil$, $s_w = 1$ and height $h = 2$;

(2) there is a routing with $s_n = \lceil \frac{w}{2n} \rceil + 1$, $s_w = 2$ and height $h = 3n$.

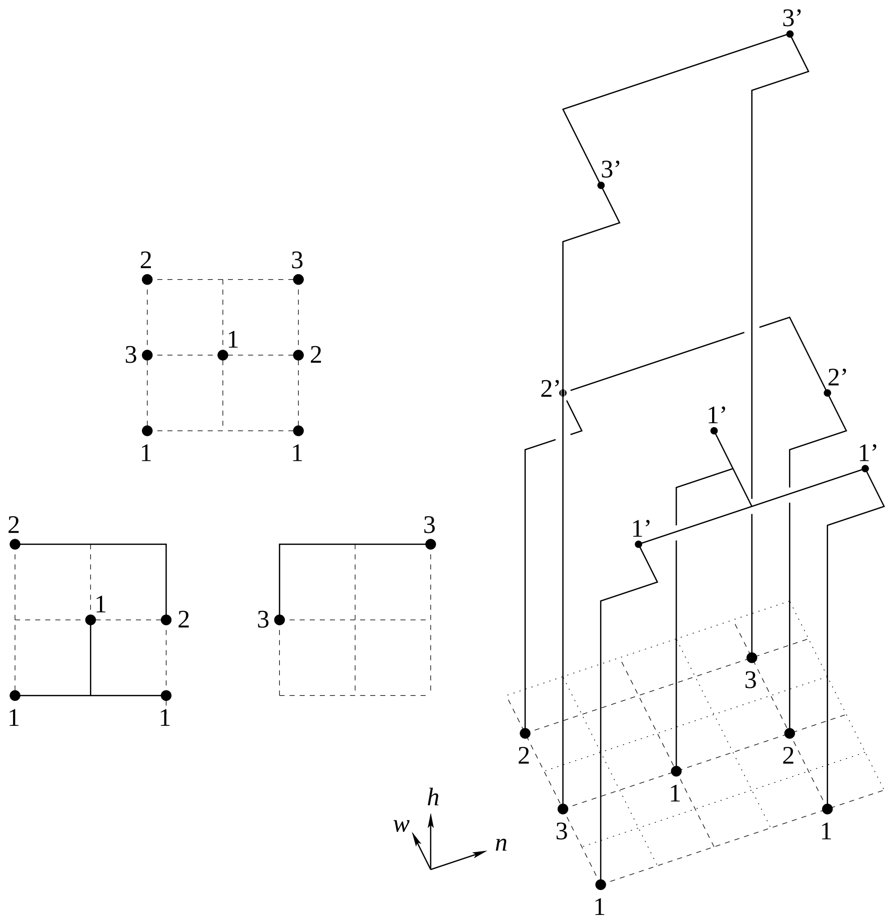

Figure 2:

As mentioned above, the second statement is a corollary of the first one. To prove the first statement, we have to observe that a SALRP is trivially solvable on 2 layers if the increased length $n \cdot s_n$ is at least the number of nets and each row contains either one terminal or two terminals belonging to the same net. From this we can obtain the necessary partition by edge-coloring a suitably defined conflict graph (whose edges correspond to the nets). The coefficient of $\frac{3}{2}$ comes from Shannon's classic theorem [16] claiming that the chromatic index of a multigraph is at most $\frac{3}{2}\Delta$. Details can be found in [13].

Since the role of n and w is symmetric, we can also formulate the following corollary of Theorem 5 for the $s_w = s_n = 2$ case.

Corollary 6 (A. Recski and D. Szeszlér, 2007)[13] *Assume that a SALRP instance is given such that each net contains two terminals only.*

(1) if $n \le w \le 2n$ then there is a routing with $s_n = s_w = 2$ and height $h = 3n$;

(2) there is a routing with $s_n = s_w = 2$ and height $h = 3\max(n, w)$.

In Theorem 5 and Corollary 6 we only considered SALRP instances where each net is a terminal pair. However, allowing multiterminal nets is a crucial point in VLSI design. If we wanted to apply the proof method of Theorem 5 for the multiterminal case in a straightforward way, we would come across the problem of edge-coloring hypergraphs, where an analogue of Shannon's theorem does not exist. However, if we break up each t-terminal net into $t-1$ terminal pairs then the construction of Theorem 5 works with the only difference being that the maximum degree in the corresponding conflict graph, and thus the height of the obtained routing is doubled.

Theorem 7 (A. Recski and D. Szeszlér, 2007)[13] *Any SALRP can be solved with*

(1) $s_n = \lceil \frac{w}{2n} \rceil + 1$, $s_w = 2$ *and height* $h = 6n$*;*

(2) $s_n = s_w = 2$ *and height* $h = 6\max(n, w)$*.*

We mention that the constructions in the proof of Theorems 5 and 7 can be realized by a polynomial algorithm. More precisely, the running time is $O(t \cdot (w + n))$, where t is the number of nets (and w and n denote the width and the length, respectively). The size of the input is $A = w \cdot n$, thus if we make the (sensible) assumption that $w = \Theta(n)$ holds then the running time is $O(A^{\frac{3}{2}})$.

4 3-dimensional Channel Routing

The constructions in the proof of Theorems 5 and 7 have the following, very useful property:

> for each terminal t the vertical (that is, parallel with the height) line of the grid intersecting t is occupied by a single "long" vertical wire segment (which, of course, ends in t). $(*)$

This property implies the following observation:

Lemma 8 *If* $s_w \geq 2$ *and* $s_n \geq 4$ *then every 3DCRP can be solved with height* $h = 6\max(n, w)$*.*

Proof: Let us move the terminals of the top grid by two units in a direction parallel with the length n and then project these terminals on the bottom grid. This way we obtain a SALRP with $s_w, s_n \geq 2$, which can be solved with height $h = 6\max(n, w)$ by Theorem 7. This solution can be modified in a straightforward way to give a solution of the original 3DCRP by the above property $(*)$. □

Therefore we will only consider the 3DCRP problem with $s_w = s_n = 2$ in the sequel. Property $(*)$ also enables us to reduce the problem to a very special case: a 3DCRP instance is called *bipartite* if each net contains two terminals only, one on the top grid and the other one on the bottom grid. Indeed, if we solve the two SALRP instances defined by the top grid and the bottom grid separately, the remaining task is nothing but a bipartite 3DCRP instance. (This bipartite 3DCRP instance is obtained by choosing one terminal on the top grid and one on the bottom grid from every net that has terminals on both grids arbitrarily.) Again, property $(*)$ enables us to combine any solution of the bipartite 3DCRP instance with the two SALRP soultions to give a routing of the original 3DCRP instance.

Theorem 9 (A. Reiss and D. Szeszlér, 2005)[14]

(1) Every bipartite 3DCRP can be solved with $s_w = s_n = 2$ *and height* $h = 3\max(n, w)$.

(2) Every 3DCRP can be solved with $s_w = s_n = 2$ *and height* $h = 15\max(n, w)$.

The second statement is a corollary of the first one, using Theorem 7 and the above observation. The first statement is proved by a careful construction that again uses the edge-coloring of a conflict graph (which is, in this case, bipartite) and a (2-dimensional) channel routing algorithm of Recski and Strzyzewski [11]. Details can be found in [14].

Again, the construction of the proof of Theorem 9 can be realized by a polynomial algorithm with a running time of $O(t \cdot (w + n))$ (which is $O(A^{\frac{3}{2}})$ if $w = \Theta(n)$ is assumed).

References

[1] Aggarwal, A., M. Klawe, D. Lichtenstein, N. Linial and A. Wigderson, A lower bound on the area of permutation layouts, *Algorithmica* (1991) **6**, 241-255.

[2] Aggarwal, A., J. Kleinberg and D. P. Williamson, Node-disjoint paths on the mesh and a new trade-off in VLSI layout, *SIAM J. Computing* (2000) **29**, 1321-1333.

[3] Boros, E., A. Recski and F. Wettl, Unconstrained multilayer switchbox routing, *Ann. Oper. Res.* (1995) **58**, 481-491.

[4] Cutler, M. and Y. Shiloach, Permutation layout, *Networks* (1978) **8**, 253-278.

[5] Enbody, R.J., G. Lynn and K. H. Tan, Routing the 3-D chip, *Proc. 28th ACM/IEEE Design Automation Conf.* (1991) 132-137 .

[6] Games, R. A., Optimal book embeddings of the FFT, Benes, and barrel shifter networks, *Algorithmica* (1986) **1**, 233-250.

[7] Hambrusch, S. E., Channel routing in overlap models, *IEEE Trans. Computer-Aided Design of Integrated Circ. Syst.* (1985) **CAD-4**, 23-30.

[8] Leighton, T., Complexity issues in VLSI: Optimal layouts for the shuffle-exchange graph and other networks, *The MIT Press, Cambridge, MA.* (1983)

[9] Leighton, T. and A. L. Rosenberg, Three-dimensional circuit layouts, *SIAM J. Computing* (1986) **15**,793-813 .

[10] Leighton, T., S. Rao and A. Srinivasan, New algorithmic aspects of the Local Lemma with applications to routing and partitioning, *Proc. Tenth Annual ACM-SIAM Symp. on Discrete Algorithms* ACM/SIAM, New York and Philadelphia (1999) 643-652.

[11] Recski A. and F. Strzyzewski, Vertex-disjoint channel routing on two layers, *Integer programming and combinatorial optimization*, 397-405, Ravi Kannan and W.R. Pulleyblank, eds., University of Waterloo Press (1990).

[12] Recski, A. and D. Szeszlér, Routing vertex disjoint Steiner-trees in a cubic grid — an application in VLSI, *Discrete Applied Mathematics* (2007) **155**, 44-52.

[13] Recski, A. and D. Szeszlér, 3-dimensional single active layer routing, *Discrete and Computational Geometry*, 318-329, Lecture Notes in Computer Science **2098**, Springer, Berlin (2001).

[14] Reiss, A. and D. Szeszlér, 3-dimensional Channel Routing, *Proc. 4th Japanese-Hungarian Symposium on Discrete Mathematics and Its Applications* (2005), 409-415.

[15] Rosenberg, A. L., Three-dimensional VLSI: a case study, *J. ACM.* (1983) **30**, 397-416.

[16] Shannon, C., A theorem on coloring the lines of a network, *J. Math. Phys.* (1949) **28**, 148-151.

[17] Shirakawa, I., Some comments on permutation layout, *Networks* (1980) **10**, 179-182.