Congestion Estimation for 3-D Circuit Architectures

Lerong Cheng, William N. N. Hung, Guowu Yang, and Xiaoyu Song

Abstract—Three-dimensional (3-D) routing is an important step in deep submicrometer very large-scale integrated design. Given a 3-D grid graph and a set of two-terminal nets to be routed, we propose a probabilistic model to calculate the routing density (congestion) on each edge of the grid graph. The routing density provides a direct congestion estimation. Our experimental results demonstrate the effectiveness of the method on routing benchmarks.

Index Terms-Estimation, probabilistic methods, routing.

I. INTRODUCTION

R OUTING plays a critical role in the high performance deep submicrometer very large-scale integrated (VLSI) design process [10], [11]. Under the current technology, application-specific integerated circuit (ASIC) designs are frequently implemented with four to six layers of metals [10]. Most printed circuit boards (PCBs) and multichip modules (MCMs) use multiple layers where their logic cells occupy the bottom layers and their interconnects are routed in upper metal layers [2], [10]. Three-dimensional (3-D) layouts [2], [5], [8], [12], [13] and placement [3] were considered for VLSI. Three-dimensional architectures have also been proposed for field-programmable gate arrays (FPGAs) [7], [9] and reconfigurable computing [1].

A probabilistic model of interconnect routing will enable designers to quickly compute estimates of the routing congestion. There has been some work on probabilistic models. Lou *et al.* [6] presented a fast probabilistic estimation for congestion. They assumed each net uses the shortest route and each possible route has the same usage probability. Based on such assumption, they estimate the congestion for the routing area. The previous probabilistic approaches are not sufficiently accurate as the shortest routes are merely considered in their restricted models. A more precise prediction should incorporate congestion-related detouring in the model to reflect the actual practice.

In this brief, we consider the following problem: Given a 3-D grid graph and a set of two-terminal nets to be routed, we propose a probabilistic model to calculate the routing density (congestion) on each edge of the grid graph. The routing density provides a direct congestion estimation. Our method is the first attempt to model 3-D congestion. The experimental results demonstrate the effectiveness of the method on routing benchmarks.

Digital Object Identifier 10.1109/TCSII.2004.838548

II. PROBLEM DEFINITION

Without loss of generality, we model a 3-D routing region by $R = \{0, 1, \ldots, W\} \times \{0, 1, \ldots, L\} \times \{0, 1, \ldots, H\}$ where W, L, and H are nonnegative integers. Within the discrete routing space, the set of lines including the rectangle boundaries is called the grid (or the grid graph) and the lines are called grid lines. An intersection between grid lines is referred to as a grid point. We denote the set of all grid points by G. A subset of G is referred to as terminal points. A terminal t is denoted by a 3-tuple (x(t), y(t), z(t)) of the x, y, z coordinate values of t. An edge is a small interval connecting two adjacent grid points along one of the three dimensions. In this brief we also refer to an edge as a unit line. A route is a set of consecutively adjacent edges connecting two grid points.

The set of terminal points is partitioned into N sets, $n_0, n_1, \ldots, n_{N-1}$. Each set n_i is called a net. Hence, a net is a collection of terminals to be interconnected together. Nets containing exactly two terminals are called two-terminal nets; otherwise, they are called multiterminal nets. In this brief, we restrict our discussion on two-terminal nets. Our model can be extended to handle the case for multiterminal nets using rectilinear Steiner trees or minimum spanning trees.

Given any two-terminal net, with terminals (x_a, y_a, z_a) and (x_b, y_b, z_b) , we can apply a coordinate transformation on the entire routing region R so that the two terminals become (0, 0, 0) and (p, q, r), where $(p, q, r) = (|x_b - x_a|, |y_b - y_a|, |z_b - z_a|)$. Terminal (0, 0, 0) is the start terminal and terminal (p, q, r) is the end terminal of the net. The direction of the route is from the start terminal to the end terminal. A route segment is a interconnect connecting two grid points (not necessarily adjacent to each other) along one of the three dimensions. The direction of the route segment is a route segment is a route segment is a route segment is a segment is a route segment is a route segment which goes negatively along its dimension. The reverse segments are also called detours.

Since we do not restrict the net route with the shortest length, the routes may not be monotonic, i.e., may not always use forward segments. If a route contains reverse segments, it will increase the total interconnect length, thus increasing path delay. To limit the interconnect length, we assume the route is either shortest path (monotonically increasing route) or have a limit length detour (detour length $l \leq d$) and the detour can go in only one dimension. By allowing detours, the legal routing space for each net is actually extended beyond the bounding cube. Fig. 1 shows an extended routing cube. For each net, the detour can happen many times, but they all have to be in the same dimension for that net. However, different nets can have detours in different dimensions. We define the density of an edge as the number of possible routes that go through this edge. Given the

Manuscript received November 6, 2003; revised May 21, 2004. This paper was recommended by Associate Editor X. Hong.

The authors are with the Department of Electrical and Computer Engineering, Portland State University, Portland, OR 97229 USA (e-mail: whung@ece.pdx.edu.)



Fig. 1. Extended routing cube with detour of length d.

above routing model, our problem is to calculate the density of all the edges in the entire routing region R.

III. DENSITY ESTIMATION

For a two-terminal net n_i , assume there are M_i possible routes, and for a unit line U_j , assume there are B_{ji} such routes passing though it. Therefore, the density of the unit line under the two terminal net n_i is $D_{ij} = B_{ji}/M_i$. The estimated density for each unit line U_i is: $D_j = \sum_{n_i \in N} D_{ij}$.

Let us define u(p,q,r) as the number of routes that go monotonically from (0,0,0) to (p,q,r). Since there are p unit lines in x-dimension, q unit lines in y-dimension, and r unit lines in z-dimension. We use X, Y, Z to represent such unit lines. Each route is correspondent to a combination that we put p X's, qY's, and r Z's in a series. The number of such combination is $C_{p+q+r}^{p,q,r}$, where $C_{p+q+r}^{p,q,r} = C_{p+q+r}^{p} \times C_{q+r}^{q}$. Hence

$$u(p,q,r) = C_{p+q+r}^{p,q,r} = \begin{cases} \frac{(p+q+r)!}{p!q!r!}, & p \ge 0, \quad q \ge 0, \\ 0, & \text{otherwise.} \end{cases}$$

Let us define $f_x(p,q,r,l)$ as the number of routes that: 1) go from (0, 0, 0) to (p,q,r); and 2) have detour only in the x-dimension and the total detour length is l.

Lemma 1:

$$f_x(p,q,r,l) = \begin{cases} u(p,q,r), & l = 0\\ C_{q+r}^r \sum_{j=1}^l C_{l-1}^{j-1} C_{q+r+1}^j C_{p+q+r+l-j}^{q+r-j}, & l > 0 \end{cases}$$

Proof:

Case 1) (l = 0): The route is monotonic, hence the lemma. Case 2) (l > 0): In this case, the route is nonmonotonic, as shown in Fig. 2. There are p + l unit lines in x-dimension, q unit lines in y-dimension and r unit lines in z-dimension, and r unit lines in detour. We use X, Y, Z, and R to represent such unit lines respectively. Each route corresponds to a combination such that X, Y, Z, and R are put in a series and X and R can not be adjacent to each other. For instance, the route presented in Fig. 2 is

To compute the number of such combinations, we first put Y's and Z's in a series, there are C_{q+r}^r such combinations.

XXXZRYRYYXXX.



Fig. 2. Nonmonotonic route in f_x .

Later, we insert the X's and R's in the series. There are q+r+1 places for X's and R's. We pick *i* of them. There are C_{q+r+1}^i choices. In the chosen *i* places, we pick *j* to put R's and the other i - j to put X's. There are C_j^i choices.

For the example route in Fig. 2, there are three Y's and one Z. They are put as ZYYY. We pick four of the places to put R's and X's. In the four picked places, we choose two out of four to put R's and the other two to put X's. Such as $\bullet_X Z \bullet_R Y \bullet_R Y \circ Y \bullet_X$, where \circ denotes an unpicked place, \bullet_R denotes a place picked to put R and \bullet_X denotes a place picked to put X.

And then, we have to divide p + l X's into i - j groups and also divide l R's into j groups. For the first task, dividing p + lX's into i-j groups is the same as inserting i-j-1 breaks into p+l-1 places. Therefore, there are C_{p+l-1}^{i-j-1} such combinations. For the second task, there are C_{l-1}^{j-1} combinations that l R's are divided into j groups.

Moreover, it is obvious that the range of i and j are $2 \le i \le q+r+1, 1 \le j \le l$. If we multiply all of the above choices, we arrive at $C_{q+r}^r \sum_{i=2}^{q+r+1} \sum_{j=1}^{l} C_{q+r+1}^i C_i^j C_{l-1}^{j-1} C_{p+l-1}^{i-j-1}$. This is the total number of routes, and can be simplified as

$$C_{q+r}^{r} \sum_{j=1}^{l} C_{l-1}^{j-1} \sum_{i=2}^{q+r+1} C_{i}^{j} C_{q+r+1}^{i} C_{p+l-1}^{i-j-1}.$$
 (1)

Among the terms of (1), $C_i^j C_{q+r+1}^i$ can be simplified to: $C_{q+r+1}^j C_{q+r+1-j}^{i-j}$. We can substitute it to (1), and simplify to

$$C_{q+r}^{r} \sum_{j=1}^{l} C_{l-1}^{j-1} C_{q+r+1}^{j} \sum_{i=2}^{q+r+1} C_{q+r+1-j}^{i-j} C_{p+l-1}^{i-j-1}.$$
 (2)

We now show that the last summation term of (2) can be reduced.

First we want to show that

$$\sum_{k=0}^{u} C_{u}^{k} C_{v}^{s-i} = C_{u+v}^{s}.$$
(3)

Consider the term a^s in the polynomial $(a + 1)^{u+v}$ where its coefficient is C_{u+v}^s . We also consider the same term in the polynomial $(a + 1)^u (a + 1)^v$. It is known that

$$(a+1)^{u} = \left(C_{u}^{0}a^{0} + C_{u}^{1}a^{1} + C_{u}^{2}a^{2} + \dots + C_{u}^{u}a^{u}\right)$$
$$(a+1)^{v} = \left(C_{v}^{0}a^{0} + C_{v}^{1}a^{1} + C_{v}^{2}a^{2} + \dots + C_{v}^{v}a^{v}\right)$$

So, the coefficient of the term a^s in the polynomial $(a + 1)^u (a+1)^v$ is $\sum_{k=0}^u C_u^k C_v^{s-k}$. Notice that $(a+1)^u (a+1)^v = (a+1)^{u+v}$, we obtain: $\sum_{k=0}^u C_u^k C_v^{s-k} = C_{u+v}^s$. Hence the (3) holds. Now, let q + r + 1 - j = u, p + l - 1 = v, i - j = k, p + l = s = v + 1. For the last summation term of (2), we

know that $j \ge 1$. Also, when $k \le 0$, $C_u^k C_v^{v+1-k} = 0$ Therefore, the last summation term of (2), $\sum_{i=2}^{q+r+1} C_{q+r+1-j}^{i-j} C_{p+l-1}^{i-j-1}$, can be reduced to $C_{p+q+r+l-j}^{p+l}$. Thus, formula (2) can be reduced to

$$C_{q+r}^{r} \sum_{j=1}^{l} C_{l-1}^{j-1} C_{q+r+1}^{j} C_{p+q+r+l-j}^{q+r-j}$$

This is precisely the total number of routes that we stated in our lemma.

In the same manner, we define $f_y(p,q,r,l), f_z(p,q,r,l)$ as

$$f_y(p,q,r,l) = f_x(q,p,r,l)$$

$$f_z(p,q,r,l) = f_x(r,q,p,l)$$

Let us define total_number(p, q, r, l) as the total number of routes going from (0, 0, 0) to (p, q, r) with the detour length l. When l = 0, total number is u(p, q, r). When l > 0, it is equal to: $f_x(p, q, r, l) + f_y(p, q, r, l) + f_z(p, q, r, l)$.

Let us define $f_{x,xr}(p,q,r,l)$ as the number of routes that: 1) go from (0,0,0) to (p,q,r); 2) have detour in the x-dimension and the total detour length is l; and 3) have a segment connecting the point (0,0,0) in the x-dimension and it is the detour segment. Based on ideas similar to Lemma 1, we can see that $f_{x,xr}(p,q,r,l)$ is 0 when l = 0. When l > 0, it is equal to $C_{q+r}^r \sum_{j=0}^{l-1} C_{l-1}^j C_{q+r}^{j} C_{p+q+r+l-j-1}^{q+r-j-1}$. We now define $f_{x,xf}(p,q,r,l), f_{x,yf}(p,q,r,l), f_{x,zf}(p,q,r,l),$

We now define $f_{x,xf}(p,q,r,l)$, $f_{x,yf}(p,q,r,l)$, $f_{x,zf}(p,q,r,l)$, and $f_{x,yz}(p,q,r,l)$. They all count the number of routes that go from (0, 0, 0) to (p,q,r) and has a detour in the *x*-dimension and the total detour length is *l*. Their differences are in the segment that connect to (0, 0, 0). For $f_{x,xf}$, this segment is in the *x*-dimension and is NOT the detour segment. For $f_{x,yf}$ and $f_{x,zf}$, this segment is in the *y* and *z* dimension respectively. For $f_{x,yz}$, this segment is simply NOT in the *x*-dimension. Thus, the set of routes counted by $f_{x,yz}$ is the union of the set of routes counted by $f_{x,yf}$ and $f_{x,zf}$. Also, the sets of routes counted by $f_{x,yf}$ and $f_{x,zf}$ are disjoint.

We define $f_{y,xf}(p,q,r,l)$ as the number of routes that: 1) go from (0,0,0) to (p,q,r); 2) have detour in y-dimension and the total detour length is l; and 3) have a segment connecting the point (0,0,0) in the x-dimension. We also define $f_{z,xf}$ in a way similar to $f_{y,xf}$.

Given $f_x(p,q,r,l)$ as evaluated in Lemma 1, we can compute the values of the following by exchanging x, y, z coordinates:

$$\begin{split} f_{x,yf}(p,q,r,l) &= f_x(p,q-1,r,l) \\ f_{x,zf}(p,q,r,l) &= f_x(p,q,r-1,l) \\ f_{y,xf}(p,q,r,l) &= f_{x,yf}(q,p,r,l) \\ f_{z,xf}(p,q,r,l) &= f_{x,zf}(r,q,p,l). \end{split}$$

We have

$$\begin{split} f_{x,yz}(p,q,r,l) &= f_{x,yf}(p,q,r,l) + f_{x,zf}(p,q,r,l) \\ &= \begin{cases} u(p,q-1,r) + u(p,q,r-1), & l = 0 \\ C_{q+r}^r \sum_{j=1}^l C_{l-1}^{j-1} C_{q+r}^j C_{p+q+r+l-j-1}^{q+r-j-1}, & l > 0. \end{cases} \end{split}$$

We can also deduce that $f_{x,xf}(p,q,r,l)$ equals u(p,q,r)when l = 0. In case l > 0, it is equal to $f_x(p,q,r,l) - f_{x,xr}(p,q,r,l) - f_{x,yz}(p,q,r,l)$, which is the same as

$$C_{q+r}^{r} \sum_{j=1}^{l} C_{l-1}^{j-1} C_{q+r}^{j} C_{p+q+r+l-j-1}^{q+r-j}.$$

Definition 1: $n_x(x, y, z, p, q, r, l)$ represent the number of routes that have the following features: 1) it goes from (0, 0, 0) to (p, q, r); 2) the total detour length is l; and 3) it passes through the unit line connecting the point (x, y, z) and (x - 1, y, z).

Now let us consider x, y, z in the following range: $1 - l \le x \le p + l, -l \le y \le q + l, -l \le z \le r + l$. Theorem 1:

$$\begin{split} n_x(x,y,z,p,q,r,l) &= \begin{cases} 0, & cn_0(x,y,z,p,q,r,l) \\ U(x,y,z,p,q,r), & l=0 \\ F_{xf}(p,q,r,l), & cn_{xf}(x,y,z,p,q,r), \\ f_{x,xr}(p,q,r,l), & cn_{x,xr}(x,y,z,p,q,r), \\ on_x(x,y,z,p,q,r,l) & \text{otherwise} \end{cases} \end{split}$$

where $cn_0(x, y, z, p, q, r, l)$ equals

$$\begin{aligned} (x < 1 - l) \lor (y < -l) \lor (z < -l) \\ \lor (x > p + l) \lor (y > q + l) \lor (z > r + l) \\ \lor (((x > p) \lor (x < 1)) \\ \land ((y > q) \lor (y < 0) \lor (z > r) \lor (z < 0))) \\ \lor (((x > q) \lor (y < 0)) \\ \land ((x > p) \lor (x < 1) \lor (z > r) \lor (z < 0))) \\ \lor (((z > r) \lor (z < 0)) \\ \land ((y > q) \lor (y < 0) \lor (x > p) \lor (x < 1))). \end{aligned}$$

U(x, y, z, p, q, r) is equal to $u(x-1, y, z) \times u(p-x, q-y, r-z)$. $F_{xf}(p, q, r, l)$ is equal to $f_{x,xf}(p, q, r, l) + f_{y,xf}(p, q, r, l) + f_{z,xf}(p, q, r, l)$.

 $cn_{xf}(x, y, z, p, q, r)$ is equal to

$$\begin{aligned} ((x=1)\wedge(y=0)\wedge(z=0))\\ \lor((x=p)\wedge(y=q)\wedge(z=r)). \end{aligned}$$

 $cn_{x,xr}(x,y,z,p,q,r)$ is equal to

$$\begin{aligned} ((x=0)\wedge(y=0)\wedge(z=0))\\ \vee((x=p+1)\wedge(y=q)\wedge(z=r)). \end{aligned}$$

$$on_x(x, y, z, p, q, r, l)$$
 is equal to

$$\sum_{k=0}^{l} (f_y(x-1,y,z,k) \times f_y(p-x,q-y,r-z,l-k) + f_z(x-1,y,z,k) \times f_z(p-x,q-y,r-z,l-k) + (f_{x,yz}(x-1,y,z,k) + f_{x,xf}(x-1,y,z,k)) \times (f_{x,yz}(p-x,q-y,r-z,l-k) + f_{x,xf}(p-x,q-y,r-z,l-k))) + \sum_{k=0}^{l-1} ((f_{x,yz}(x,y,z,k) + f_{x,xr}(x,y,z,k)) \times (f_{x,yz}(p-x+1,q-y,r-z,l-k-1) + f_{x,xr}(p-x+1,q-y,r-z,l-k-1))).$$

Proof: The evaluation of n_x is divided into five cases.

- The unit line is not in the usage area.
- The route is monotonic. The route will first reach the point (x 1, y, z), passes through the unit line and then goes from (x, y, z) to (p, q, r).
- The unit line is connecting the start point or end point. The start point and end point cases are symmetric, so it is the sum of $f_{x,xf}$, $f_{y,xf}$, and $f_{z,xf}$.
- The unit line is connecting the start point or end point, and it is the detour segment. The start point and end point cases are symmetric, so it is equal to $f_{x,xr}$.
- Otherwise, we divide to four subcases: (1) The detour goes y-dimension. (2) The detour goes z-dimension. (3) The detour goes x-dimension and the unit line is not in the detour segment. (4) The detour goes x-dimension and the unit line is in the detour segment.

All of the above cases have been evaluated in previous lemmas. Hence we prove Theorem 1.

Similarly, we can define n_y and n_z by coordinate transformations from n_x

$$n_y(x, y, z, p, q, r, l) = n_x(y, x, z, q, p, r, l)$$

where $-l \le x \le p+l, 1-l \le y \le q+l, -l \le z \le r+l$

$$n_z(x, y, z, p, q, r, l) = n_x(z, y, x, r, q, p, l)$$

where $-l \le x \le p + l, -l \le y \le q + l, 1 - l \le z \le r + l.$

After computing the number of routes, we can compute the density of each unit lines.

Definition 2: $den_x(x, y, z, p, q, r)$ is the density of the unit line connecting the point (x, y, z) and (x - 1, y, z) under the two terminal net.

Theorem 2:

$$den_x(x, y, z, p, q, r) = \frac{\sum_{l=0}^{d} weight(x, y, z, p, q, r, l) \times n_x(x, y, z, p, q, r, l)}{\sum_{l=0}^{d} weight(x, y, z, p, q, r, l) \times total_number(p, q, r, l)}$$

Here, weight (x, y, z, p, q, r, l) is the weight factor of the route whose detour length is l. We add such weight factor because in practice, the longer route will have smaller usage probability. The details of such weight factor can be adjusted by users.

We can also define den_y and den_z and compute them in a similar way. The weight factor of den_z can be scaled by a factors less than 1 to account for different characteristics of wiring in the Z dimension. We have focused on congestion estimation of two-terminal nets. We can also handle multiterminal nets by transforming them into two-terminal nets using Steiner trees or other methods.

IV. EXPERIMENTS

We used a random number generator to randomly construct a set of 3-D routing benchmarks. The random data are generated using a random number generator where all terminals are randomly placed within a 3-D routing region. We then create

TABLE I THREE-DIMENSIONAL ROUTING DENSITIES

Dimension						Routed		Unrouted	
Х	Y	Z	d	Nets	d_T	$< d_T$	avg	$> d_T$	avg
10	10	0	0	22	.200	20%	.607	20%	.114
10	10	6	0	78	.202	35%	.344	35%	.180
15	15	0	0	28	.198	25%	.498	25%	.132
15	15	6	0	105	.173	37%	.308	37%	.156
20	20	0	0	45	.228	28%	.479	29%	.147
20	20	6	0	155	.171	37%	.295	37%	.159
40	40	0	0	152	.217	27%	.501	27%	.147
40	40	3	0	308	.151	36%	.308	36%	.145
100	100	0	0	498	.183	29%	.395	29%	.135
100	100	2	0	1547	.185	34%	.340	34%	.160
100	100	4	0	2497	.184	37%	.310	37%	.172
100	100	6	0	3757	.194	39%	.308	39%	.187
10	10	0	1	22	.083	24%	.286	23%	.058
10	10	6	1	78	.177	34%	.276	34%	.147
15	15	0	1	37	.083	24%	.284	26%	.068
15	15	6	1	116	.168	36%	.254	36%	.146
20	20	0	1	57	.097	24%	.284	24%	.075
20	20	6	1	163	.160	37%	.245	37%	.144
40	40	0	1	142	.092	23%	.255	23%	.069
40	40	3	1	297	.145	34%	.250	34%	.127
100	100	0	1	541	.078	24%	.230	24%	.063
100	100	2	1	1522	.146	32%	.259	32%	.124
100	100	4	1	2546	.176	38%	.259	38%	.158
10	10	0	2	20	.042	24%	.194	25%	.037
10	10	6	2	67	.139	30%	.228	30%	.106
15	15	0	2	33	.042	26%	.175	26%	.041
15	15	6	2	111	.139	32%	.222	32%	.112
20	20	0	2	53	.045	25%	.189	25%	.041
20	20	6	2	160	.143	33%	.220	33%	.114
40	40	0	2	138	.049	25%	.183	25%	.043
40	40	3	2	273	.117	31%	.210	31%	.094
100	100	0	2	500	.051	23%	.162	23%	.040
100	100	2	2	1446	.118	29%	.211	29%	.094
100	100	4	2	2463	.151	35%	.228	35%	.127

an extended routing cube with detour of length d as described in Fig. 1. The extended routing cube is then used for both the heuristic router and our congestion estimation. For each benchmark, we compute the routing density for each edge as described in previous section, using a weight function $(l + 1)^{-6}$. We then used a heuristic 3-D maze router [10] to draw out the routing paths for each benchmark. The router will attempt to connect terminals with a shortest path, and will also attempt to relax the constraints and route with a longer path. The capacity of all edges is 1 for detailed routing. Consequently, we obtained the densities of the edges that were routed (by the heuristic router) as well as the densities of the edges that were not routed.

The result is shown in Table I. The "Dimension" column shows the size of the 3-D routing space where X, Y, Z are the maximum index (counting from 0) for their corresponding axis. The "d" column shows the maximum number of detours allowed. The "Nets" column indicates the number of nets routed in each benchmark. Based on results from the heuristic router, we divide all the edges in the 3-D cube into two groups: routed edges and unrouted edges. We then choose a threshold density d_T to see how much percentage of routed edges have density lower than d_T , and how much percentage of unrouted edges have density higher than d_T (d_T was chosen such that these percentages are comparable). The above two percentages indicate routed and unrouted edges with overlapping densities, i.e., hard to differentiate. The worst case is roughly 39% from our experiments. For all benchmarks, the average density of routed edges is higher than the average density of unrouted edges. This is a good indication that the routing heuristic is in agreement with the density estimation.

To better understand the result, we take the largest benchmark (X = 100, Y = 100, Z = 6, d = 0 with 3757 nets)



Fig. 3. Density distribution for X = 100, Y = 100, Z = 6, d = 0 with 3757 nets.

and divide the edge densities into several consecutive ranges: $[0, 0.15), [0.15, 0.3), [0.3, 0.45), \ldots$ We tally the routed edges and the total number of edges for each density range, and plot the result in Fig. 3. The dashed exponentially decreasing line (with triangular markers) shows the percentage of edges in each sampling range out of total number of edges in the routing space. It is natural to see that the majority (highest percentage) of edges in the routing space have low densities, and this percentage decreases exponentially as we increase the density sampling range. The solid increasing solid line represents the percentage of routed edges out of all edges within each density sampling range. As the density increases, the percentage of routed edges (in that density range) also increases. However, the percentage fluctuates noticeably in high density sampling ranges. This is due to the diminishing (smaller and smaller) number of edges as the sampling range increases. So the statistical variation becomes more significant for these high density edges. This is because the heuristic router is choosing only one routing path for each net. The densities are computed considering all the possible routing paths of each net. The routing path chosen by the heuristic router does not necessarily contain the edge with the highest density. Nevertheless, the general trend from the graph indicates that there is a high likelihood that larger density edges are actually routed.

To further demonstrate the effectiveness of our approach, we used routing benchmarks from Profs. Sarrafzadeh and Kastner [4], and compared our congestion estimation with results from their global router [4]. The results are shown in Table II. The "Nets" column shows the number of nets for each benchmark. The "Grid" column indicates the grid size (dimension). The "Capacity" column shows the max. number of nets allowed for each vertical (V) and horizontal (H) edge. We compute the density of each edge (with or without detour) and compare them with the actual density (number of routed nets) created by the router. The absolute difference between our two estimation methods and the actual density is shown in the "without detour"

TABLE II GLOBAL ROUTING DENSITIES

		Capacity		Witho	ut Detour	With Detour	
Nets	Grid	V	Н	V	Н	V	H
9874	80×64	14	22	1.64	0.94	1.48	0.76
13484	80×64	20	30	2.33	1.48	1.90	1.18
15583	96 imes 64	20	23	2.27	1.57	1.82	1.28
19386	128×64	20	33	2.08	1.00	1.79	0.83
24195	192×64	14	24	2.13	1.02	1.98	0.87
26736	192×64	14	22	1.67	1.00	1.52	0.85
28796	256×64	14	28	1.74	0.89	1.59	0.76

and "with detour" columns for horizontal (H) and vertical (V) edges. In general, the absolute difference between estimation without detour and the actual routing is higher than the difference between estimation with detour (and the actual routing). This is a good indication that congestion estimation with detour is a better than without detour.

V. CONCLUSION

In this brief, we proposed a probabilistic model to calculate the routing density (congestion) on each edge of a 3-D grid graph. The routing density provides a direct congestion estimation. Our experimental results demonstrate the effectiveness of the method on routing benchmarks. Our work is the first method for density (congestion) estimation in 3-D routing.

References

- M. J. Alexander, J. P. Cohoon, J. L. Coleflesh, J. Karro, and G. Robins, "Three dimensional field-programmable gate arrays," in *Proc. IEEE Int. ASIC Conf.*, Austin, TX, Sept. 1995, pp. 253–256.
- [2] A. Hanafusa, Y. Yamashita, and M. Yasuda, "Three-dimensional routing for multilayer ceramic printed circuit boards," in *Proc. ICCAD*, Nov. 1990, pp. 386–389.
- [3] I. Kaya, M. Olbrich, and E. Barke, "3-d placement considering vertical interconnects," in *Proc. IEEE Int. SOC Conf.*, Sept. 2003, pp. 257–258.
- [4] Labyrinth. A global router and routing development tool. Univ. California at Santa Barbara, CA. [Online]. Available: http://www.ece.ucsb.edu/~kastner/labyrinth/
- [5] F. T. Leighton and A. Rosenberg, "Three-dimensional circuit layouts," SIAM J. Comput., vol. 15, no. 3, pp. 793–813, 1986.
- [6] J. Lou, S. Thakur, S. Krishnamoorthy, and H. S. Sheng, "Estimating routing congestion using probabilistic analysis," *IEEE Trans. Computer-Aided Design*, vol. 21, pp. 32–41, Jan. 2002.
- [7] W. M. Meleis, M. Lesser, P. Zavracky, and M. M. Vai, "Architectural design of a three dimensional FPGA," in *Proc. IEEE Conf. Advanced Research in VLSI*, Sept. 1997, pp. 256–268.
- [8] F. Preparata, "Optimal three-dimensional VLSI layout," *Math. Syst. Theory*, vol. 16, pp. 1–8, 1983.
- [9] A. Rahman, S. Das, A. P. Chandrakasan, and R. Reif, "Wiring requirement and three-dimensional integration technology for field programmable gate arrays," *IEEE Trans. VLSI Syst.*, vol. 11, pp. 44–54, Feb. 2003.
- [10] N. Sherwani, S. Bhingarde, and A. Panyam, *Routing in the Third Dimen*sion: From VLSI Chips to MCMs. New York: IEEE Press, 1995.
- [11] X. Song, Y. Tang, D. Zhou, and Y. Wang, "Wire space estimation and routability analysis for gate array chips," *IEEE Trans. Computer-Aided Design*, vol. 19, pp. 624–628, May 2000.
- [12] C. C. Tong and C. Wu, "Routing in a three-dimensional chip," *IEEE Trans. Comput.*, vol. 44, pp. 106–117, Jan. 1995.
- [13] R. Venkatesan, J. A. Davis, K. A. Bowman, and J. D. Meindl, "Optimal n-tier multilevel interconnect architectures for gigascale integration (gsi)," *IEEE Trans. VLSI Syst.*, vol. 9, pp. 899–912, Dec. 2001.