AN EFFICIENT SIGNAL REDISTRIBUTION ALGORITHM FOR MCM

*Ming-Fu Shiao, Chieh Changfan, Sao-Jie Chen, *Chia-Chun Tsai*
Dept. of Electrical Engineering
National Taiwan University
Taipei, Taiwan, R.O.C.

*Dept. of Electronic Engineering
National Taipei Institute of Tech.
Taipei, Taiwan, R.O.C.

ABSTRACT

In a Multichip Module (MCM), the I/O pins of dies are redistributed to signal layers before routing. In this paper, a new model based on congestion control was proposed to solve this problem. A 2-layer global and detailed router was employeed to finish the routing. And experimental results show that this model can produce very routable pin distribution in each layer for Multichip Module.

1 INTRODUCTION

In a Multichip Module (MCM), several bare dies are mounted on the top layer of a laminated multilayer ceramic or a thin substrate. The substrate provides metal wiring for the signal as well as for the power and ground connections. The elimination of individual chip package and direct connections of bare dies can significantly reduce the inter-chip delay, weight, size, and hence promotes the entire system performance.

This paper deals with the signal redistribution of MCM-C (multilayer ceramic), but it can also be applied to MCM-D(thin film module) and PCBs. In Section 2, the signal redistribution problem is formulated. In Section 3, a congestion control model is derived to evaluate the congestion of each layer. In Section 4, a redistribution algorithm based on the congestion model is used to solve this problem. Experimental results and conclusion are described in Section 5.

2 PROBLEM FORMULATION

In Multichip Modules, the I/O pins of mounted dies have to be redistributed to the signal layers to facilitate routing. A good redistribution of I/O pins can significantly reduce the signal layers needed for routing as well as increase the wirability of each layer. Typically, the layer structure of a multilayer MCM is arranged as shown in Figure 1. Neighboring X- and Y-signal layers construct a local routing region. The redistribution problem is formulated as follows:

†This Paper is supported in part by the National Science Council R.O.C. under Grants NSC 82-0404-E002-166, and NSC 82-0404-E027-020

Given a set of nets N_1, N_2, ..N_n, and a set of pins P_i which belongs to N_1, N_2, ..N_n, a substrate which consists of $2m$ signal layers is used to connect all these nets. We want to find a transformation function

$$T(P_i) = j$$
$$\text{where } 1 \leq j \leq m$$

such that
(1) all nets can be connected in the $2m$ layers,
(2) the use of m is minimized.

There are many possible ways to solve this problem. One popular way is to pre-generate two-dimensional wiring routes for all nets, and then, assign all the wires to layers such that intersecting wires of different nets are assigned to distinct layers [1] [2]. The other possible way is to partition all the nets into groups based on some criteria, and then route each group of nets in a local routing region independently. This method does not generate the topologies of nets in advance, and hence has more flexibility in wiring topologies.

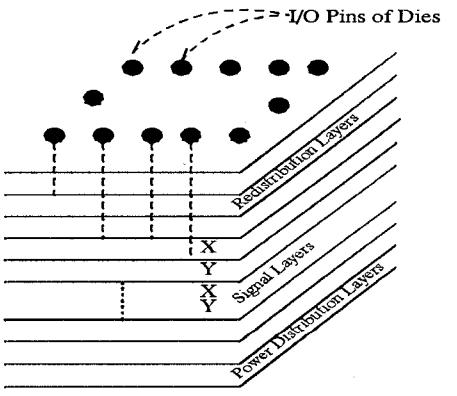


Figure 1: Layer structure of a multilayer Multichip Module, where neighboring X- and Y-signal layers construct a local routing region.

2.1 The MCM Multilayer Routing Model

Multilayer routing problem has been studied extensively since the age of PCB. The routing model of M-

0-7803-0826-3/93 $3.00 © 1993 IEEE

CM is analogous to that of PCB except the following variations: First, most PCB technologies place chips using drilled through hole, i.e., all signal layers can access the I/O pins of chips. While in MCM, only the top layer bears the mounted dies. The I/O pins of dies have to be redistributed to the signal layers before routing. Second, recent MCM technologies support different types of vias. While PCBs only supports vias of Plated-Through-Hole (PTH).

Generally, there are three different types of vias used in MCM [3]. A *pass through via* connects signal from the top layer to the bottom one. This kind of vias usually connect I/O pins of dies to the I/O pads of the entire module. A *segmented via* defined as a via pass through a subset of wiring planes, serves the role of distributing signals. A *programmable via* interconnects wires on two adjacent wiring planes, and can be placed at any grid location. That is, vias of different nets can be placed on the same grid location of different layers. Figure 2 shows these three types of vias.

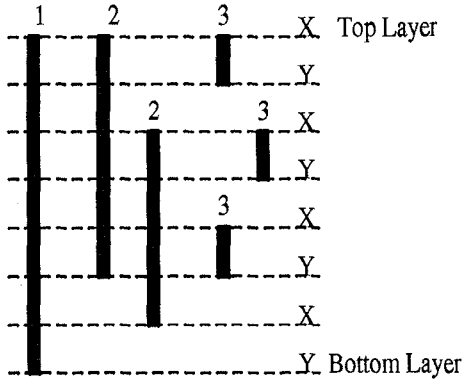


Figure 2: Three types of vias: Type 1, pass through via; Type 2, segmented via; Type 3, programmable via

3 CONGESTION MODEL

3.1 Net Split

If a net consists of too many pins, it could cause potential congestion in many layers. Since this kind of nets are usually split into subnets and assigned to distinct layers, the pins of a subnet must be localized in a small area in order to minimize the wiring length and potential congestion. A practical split procedure based on constrained clustering is proposed as follows.

First, in order to localize the pins of a subnet. We define a threshold value Max_Mass_Pin_Ratio (MMPR), such that the mass_pin_ratio of all subnets must not exceed this value. Second, a pair of pins which has the farest distance is selected as the seeds. These seed pins start to merge the pins nearest to it one by one until all the pins are merged. Another cycle is started while the mass_pin_ratio of some split subnets still exceed the MMPR. The cycle will stop until no subnet has a mass_pin_ratio greater than the MMPR.

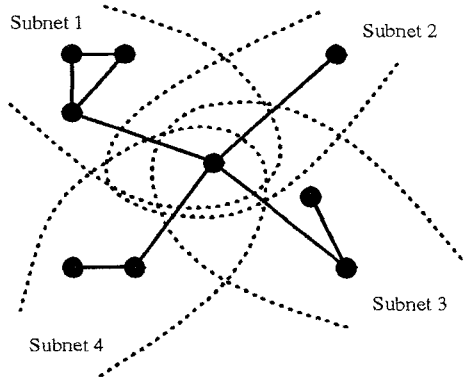


Figure 3: A large net is split into 4 subnets

After the net split, all subnets are regarded as independent nets. Figure 3 shows an example of a split net.

3.2 Force Model

After the net split, all subnets are limited in a local region. In the next phase, subnets are partitioned into several groups such that each group can be routed in a local routing region. To achieve a routable group, each subnet of the same group must have the minimum influence on each other. To capture this scenario, a force model is proposed as the following:

A net n which contains P pins is regarded as a mass point with mass and center of mass point where the center of this mass point is defined as:

$$(\overline{X}_n, \overline{Y}_n) = (\frac{\sum_{pi \in n} X_{pi}}{P}, \frac{\sum_{pi \in n} Y_{pi}}{P})$$

And the mass is defined as:

$$M_n = \sum_{pi \in n} ((X_{pi} - \overline{X}_n)^2 + (Y_{pi} - \overline{Y}_n)^2)$$

where X_{pi}, Y_{pi} are the X- and Y-coordinate of the i'th pin of net n respectively.
The induced force of two nets N_i, N_j is defined as

$$Force(N_i, N_j) = \frac{M_i * M_j}{Distance(N_i, N_j)^2}$$

where M_i, M_j are the masses of net i, net j respectively.

$$Distance(N_i, N_j) = \sqrt{(\overline{X}_i - \overline{X}_j)^2 + (\overline{Y}_i - \overline{Y}_j)^2}$$

Before the assignment, the induced force of a subnet to a routing region is calculated in advance. If a net N

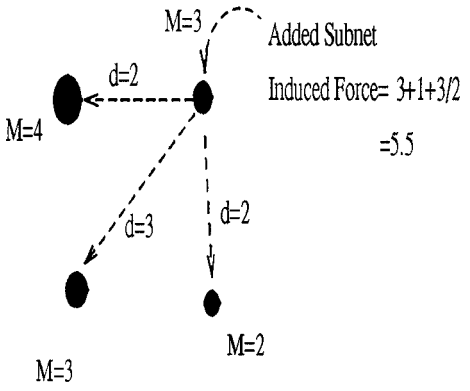


Figure 4: The force induced by an added subnets

is assigned to a routing region RR, the induced force would be:

$$Force(N, RR) = \sum_{j \in RR} Force(N, N_j)$$

Obviously, if the induced force of a net is strong, this net is potentially congested in this routing region. Hence the pending subnet prefers being assigned to a layer which is the most sparse. Figure 4 shows an example of induced force by a newly added subnet.

3.3 Capacity Estimation

A capacity estimation function is used to evaluate the available tracks while a subnet is assigned to a routing region. And this estimation function provides a penalty in the cost function. This estimation function is initially defined as the following:

$$C_{x,0} = E_x * (AVX)$$

$$C_{y,0} = E_y * (AVY)$$

where AVX, AVY is the available X tracks and Y tracks of this routing region respectively. E_x, E_y which range from 0 to 1 are the user defined X-track and Y-track densities for each routing region. If this value is low, that is, the cost in capacity penalty is high, such that all subnets are not likely to reside on this layer. Normally, the top layers have higher value of E_x, E_y than the bottom layers. If a subnet n is assigned to a routing region RR, the penalty will be:

CP(Net n, RR)=

$$\left[\frac{1}{C_{x,n} - \sum_{pi \in n}(X_{pi} - \overline{X}_n)} + \frac{1}{C_{y,n} - \sum_{pi \in n}(Y_{pi} - \overline{Y}_n)}\right]^{W1}$$

After assignment, the capacity will become:

$$C_{x,n+1} = C_{x,n} - \sum_{pi \in n}(X_{pi} - \overline{X}_n)$$

$$C_{y,n+1} = C_{y,n} - \sum_{pi \in n}(Y_{pi} - \overline{Y}_n)$$

4 REDISTRIBUTION ALGORITHM

The cost function of a pending subnet corresponding to a routing region RR is defined as the sum of the induced force and capacity penalty. The force cost balances the congestion of layers, and the capacity penalty balances the track usage of each layer. Therefore, for a given subnet and a routing region, the induced cost is:

$$Cost(N_i, RR) = CP(N_i, RR) + W2 * Force(N_i, RR)$$

where $W2$ is the weight of the induced force. The redistribution procedure is listed below.

```
Procedure Signal_Redistribution
  Net_Split(all nets)
  For(i=1 to no. of unassigned nets)
    For(j=1 to no. of routing regions)
     Calculate the cost of net i in routing region j
    End For
    Select a routing region with the minimum cost
    Assign net i to the selected region
  End For
End Signal_Redistribution
```

5 RESULTS AND CONCLUSION

5.1 Experimental Results

Three circuits which contain pins from 400 to 3200 are randomly generated and used to test the algorithm. The results of signal redistribution algorithm are routed by an integrated routing tool [4][5], which combines the 2-layer global routing and detailed routing. Table 1 lists the results, and Figure 5 shows the final routing of a sample circuit, which has 900 pins in the top layer and where six layers are used to complete the final routing.

Table 1: Experimental Results

Circuit	MCM 400	MCM 900	MCM 3200
Net	80	200	800
Pin	400	900	3200
Grid	90X90	170X170	320X180
Layers	6	6	12
Vias	2027	4882	19108
Wire Length	12696	55806	160078

5.2 Conclusion

A new concept of signal redistribution strategy is introduced in this paper. And this method has the following advantages. (1) By using our net split and redistribution algorithm, a $2m$-layer routing problem is reduced to an m 2-layer routing problem, which will significantly reduce the complexity of MCM routing problem. (2) The high routability pin distribution in each layer is derived by the congestion control which reduces the total number of signal layers needed to complete the final routing. (3) After the signal redistribution, the final routing of local routing regions can be done independently. This is a good characteristic for parallel processing.

References

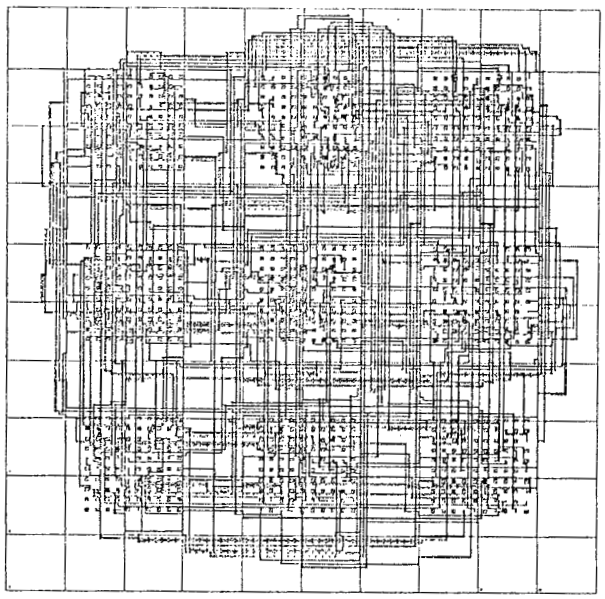
[1] J.M.Ho, M.Sarrafzadeh, G.Vijayan, C.K.Wong, "Layer Assignment for Multichip Modules," *IEEE Trans. CAD*, CAD-9(12), pp.1272-1277, Dec. 1990.

[2] M.Sriram and S.M.Kang, "Detailed Layer Assignment for MCM Routing," in *Proceedings of the International Conference on Computer Aided Design*, pp.386-389, Nov. 1992.

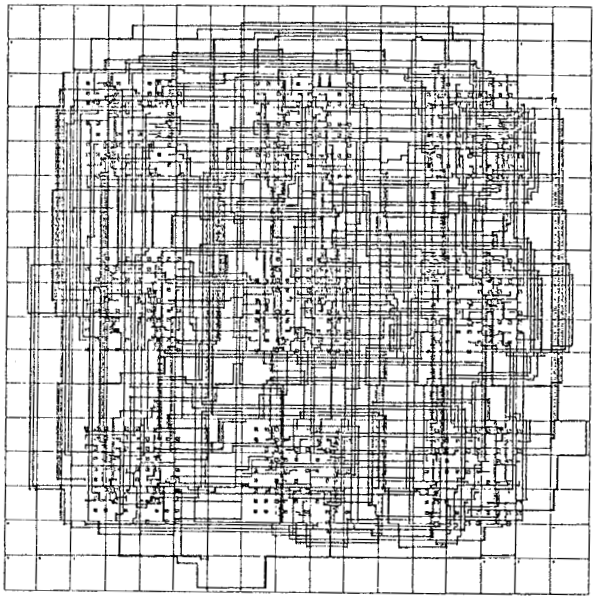
[3] R.R.Tummala, "Ceramic Packaging", in *Microelectronics Packaging Handbook*, edited by R.R.Tummala and E.J.Rymaszewski, van Nostrand Reinhold, New York, 1989, pp.455-522.

[4] T.M.Parng and R.S.Tay, "A new approach to sea-of-gate global routing," in *Proceedings of the International Conference on Computer-Aided Design*, pp.52-55, 1989.

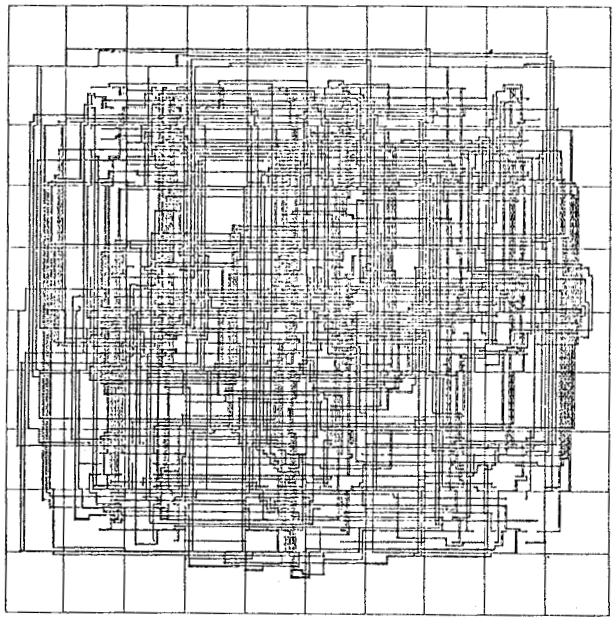
[5] W.C.Kao, *A tool bridging the gap between global and detailed routing*, M.S. thesis, Dept. of Electrical Engineering, National Taiwan University, R.O.C., 1992.



(a)



(b)



(c)

Figure 5: The final routing of a sample circuit which has 900 pins and six signal layers. (a) the top routing region (b) middle. (c) the bottom routing region.