

# Partitionnement 3D de Circuits Intégrés

François Galea<sup>1</sup>, Lilia Zaourar<sup>1</sup>, Guillaume Berhaut<sup>2</sup>, Sébastien Thuriès<sup>2</sup>

<sup>1</sup> CEA, LIST, 91191 Gif-sur-Yvette Cedex, France.  
first.last@cea.fr

<sup>2</sup> CEA, LETI, 38054 Grenoble, France.  
first.last@cea.fr

**Mots-clés :** *partitionnement, circuits intégrés, hypergraphes, recuit simulé, heuristiques.*

## 1 Introduction

Dans le flot de conception de circuits intégrés, le problème de placement/routage de réseau de cellules interconnectées (*netlist*) est une étape déterminante dans l'efficacité de fonctionnement de la puce. L'optimisation de la localité des différentes cellules permet la diminution des longueurs des fils entre cellules, ce qui a un impact positif sur la consommation d'énergie, la sensibilité aux interférences et la fréquence d'horloge maximale du circuit.

De récentes innovations technologiques d'intégration permettent de superposer plusieurs couches de silicium. Cela permet l'augmentation de la densité surfacique du circuit et apporte la possibilité de réduire les longueurs de fils par l'utilisation de *vias 3D*, liens très courts permettant de relier directement deux cellules situées sur deux couches différentes.

Dans ce contexte, nous présentons un modèle de bi-partitionnement spécifique issu d'un framework de partitionnement 3D pour la technologie CoolCube<sup>TM</sup> du CEA [1]. La fonction objectif à minimiser tend à équilibrer la répartition sur les deux couches des cellules interconnectées par un même fil, en privilégiant l'utilisation d'un *via 3D* pour relier les deux plus grandes cellules et réduire la longueur totale du fil. Par rapport au problème de partitionnement *min-cut* classique, nous ne cherchons pas à minimiser la coupe entre les deux partitions, mais une répartition qui tend au contraire à couper les sections les plus longues de chaque fil.

Nous résolvons ce problème sur quelques cas réels à l'aide d'une méthode de recuit simulé.

## 2 Modélisation

Nous représentons la *netlist* à partitionner par un hypergraphe  $(V, E)$  dont les sommets correspondent aux cellules du circuit, et les hyperarêtes sont les fils (ou *nets*) reliant les cellules entre elles. Cet hypergraphe est défini par sa matrice d'incidence  $H$  dont chaque composante  $h_{ij}$  vaut 1 si la cellule  $v_i \in V$  est connectée à l'hyperarête  $e_j \in E$ . À chaque cellule  $v_i$  est associé un couple de dimensions  $(w_i, h_i)$  correspondant aux dimensions rectangulaires (largeur, hauteur) de la cellule.

Les variables de décision  $x_{ik} \in \{0; 1\}$  sont des variables booléennes, telles que  $x_{ik}$  vaut 1 si la cellule  $v_i$  est placée dans le tiers (la partition)  $k$ , et 0 sinon.

Pour évaluer la qualité d'un partitionnement nous utilisons la métrique de Total Wire Length (TWL). Comme nous nous situons dans le cas où les cellules ne sont pas encore placées sur le circuit (et donc nous ne connaissons pas encore la longueur physique des fils), la longueur d'un fil (Wire Length) est estimée par le modèle Half-Perimeter Wire Length (HPWL), qui correspond à la somme des demi-périmètres (largeur+hauteur) des cellules connectées par ce fil. Nous en déduisons une fonction de poids HW pour chaque hyperarête, dont les cellules sont

réparties entre les tiers  $k = 1$  et  $k = 2$  :

$$\text{HW}(e_j) = \left( \sum_{i=1}^{|V|} (w_i + h_i) h_{ij} x_{i1} \right)^2 + \left( \sum_{i=1}^{|V|} (w_i + h_i) h_{ij} x_{i2} \right)^2$$

La fonction objectif à minimiser est donc la somme des poids de toutes les hyperarêtes.

La contrainte est l'équilibrage entre les surfaces occupées sur les deux tiers, telle que le ratio de surface entre les tiers reste dans l'intervalle paramétré par la constante strictement positive  $\epsilon$  :

$$\frac{1}{1 + \epsilon} \leq \frac{A_2}{A_1} \leq 1 + \epsilon$$

tel que :

$$A_k = \sum_{i=1}^{|V|} (w_i \times h_i) x_{ik}$$

### 3 Résolution

Le modèle ayant une fonction objectif quadratique et des contraintes non-linéaires, nous avons développé une méthode de résolution par recuit simulé [2]. Une solution initiale est générée à l'aide d'une heuristique gloutonne.

Les résultats obtenus sur un Xeon E5620 à 2.4 GHz sont présentés dans la table 1.

Instance	$ E $	$ V $	Coût initial	Coût final	Temps (s)
FFT	26089	25086	1,51.10 <sup>12</sup>	1,33.10 <sup>12</sup>	107
LDPC	41295	42031	2,85.10 <sup>12</sup>	2,52.10 <sup>12</sup>	163
AES	141823	141951	4,33.10 <sup>12</sup>	3,59.10 <sup>12</sup>	631

TAB. 1 – Résultats d'expérimentation sur quelques instances réelles

Les partitionnements obtenus sont considérés comme de qualité satisfaisante par les personnes du métier, comparativement à des résultats préliminaires obtenus selon des méthodologies plus empiriques.

### 4 Conclusions et perspectives

Nous proposons une modélisation originale d'un problème de partitionnement spécifique à la conception de circuits en technologie 3D. Ce problème constitue une première étape vers l'automatisation du placement/routage en technologie 3D. En effet, à notre connaissance il n'existe pas à l'heure actuelle de méthode de placement/routage tenant compte de la répartition 3D des cellules.

### Références

- [1] Guillaume Berhaut, François Galea, Lilia Zaourar, Mélanie Brocard, and Sebastien Thuriès. 3DIP : An Iterative Partitioning Tool For Monolithic 3D IC. In *Proceedings of 2016 IEEE International 3D Systems Integration Conference (3DIC 2016)*. IEEE, 2016.
- [2] Scott Kirkpatrick, C. Daniel Gelatt, and Mario P. Vecchi. Optimization by Simulated Annealing. *Science*, 220(4598) :671–680, 1983.