

## → Choisir un FPGA

# Utiliser habilement un FPGA pour traiter 200 Mpixels/s

Pour la réalisation d'une fonction de convolution en deux dimensions, 9x9, ayant des performances en vitesse de traitement satisfaisantes, la société MVD a tiré parti de sa connaissance approfondie de l'architecture des FPGA de Xilinx.

Le traitement d'images est un domaine souvent exigeant en puissance de calcul. C'est en particulier le cas pour la réalisation de filtres en deux dimensions. Ainsi, les convolutions en temps réel sont habituellement limitées à la taille 3x3 du fait de la quantité importante de ressources logiques nécessaires à l'implantation de tels algorithmes. Par ailleurs, les capteurs d'image les plus rapides affichent maintenant des vitesses très élevées. Un

Par **Edgard Garcia,**

**Xilinx**

Edgard Garcia a créé la société MVD, spécialisée dans le traitement d'images et référence française en matière de FPGA Xilinx. Depuis de nombreuses années, MVD conçoit des FPGA en langage VHDL pour tous types d'applications, y



compris le prototype d'Asic.

débit de 200 Mpixels/s permet de capturer une image de dimension 1Kx1K en quelque 5 ms. Or, un tel débit est parfois requis pour l'acquisition d'images dans les systèmes d'imagerie médicale, militaire ou industrielle. Une solution consiste à réaliser des filtres plus larges pour l'implantation de ces algorithmes afin d'obtenir un traitement en temps réel plus efficace. La société MVD a

relevé ce défi pour une fonction de convolution 2D de taille 9x9 en utilisant les astuces architecturales offertes par les matrices FPGA Virtex-II de Xilinx. Rappelons la formule mathématique d'une convolution 9x9 :

$$C(x, y) = \sum_{i=0}^{i=8} \sum_{j=0}^{j=8} D(x, y) (i, j) \times M(i, j),$$

où C est le résultat de la convolution, D(x, y) sont les données d'entrée et M(i, j) sont les coefficients du masque de convolution.

En fait, une telle fonction peut être décomposée en 9 filtres FIR (de 9 cellules chacun). La somme des sorties des 9 filtres donne le résultat de la convolution étudiée. Remarquons que certaines applications requièrent des coefficients dynamiques, d'autres des coefficients fixes; dans un cas comme dans l'autre, l'étude de la société MVD montre que la fonction peut être implantée sur un FPGA Virtex-II. Et ceci pour des vitesses atteignant les 200 Mpixels/s. La figure 1 présente l'architecture fonctionnelle d'un module de

convolution 2D, 9x9, travaillant en temps réel avec 81 coefficients signés et indépendants. La matrice ciblée pour l'implantation étant du type Virtex-II (voir encadré de la page suivante), il est possible d'utiliser une structure « Block SelectRam » en mode « Read\_First » pour réaliser chaque buffer de ligne. Jusqu'à 2 Kpixels de flux vidéo peuvent être gérés ainsi, car un module « Block SelectRam » de 18 Kbits, utilisé en mode double port, matérialise deux buffers de ligne d'1 K pixels chacun.

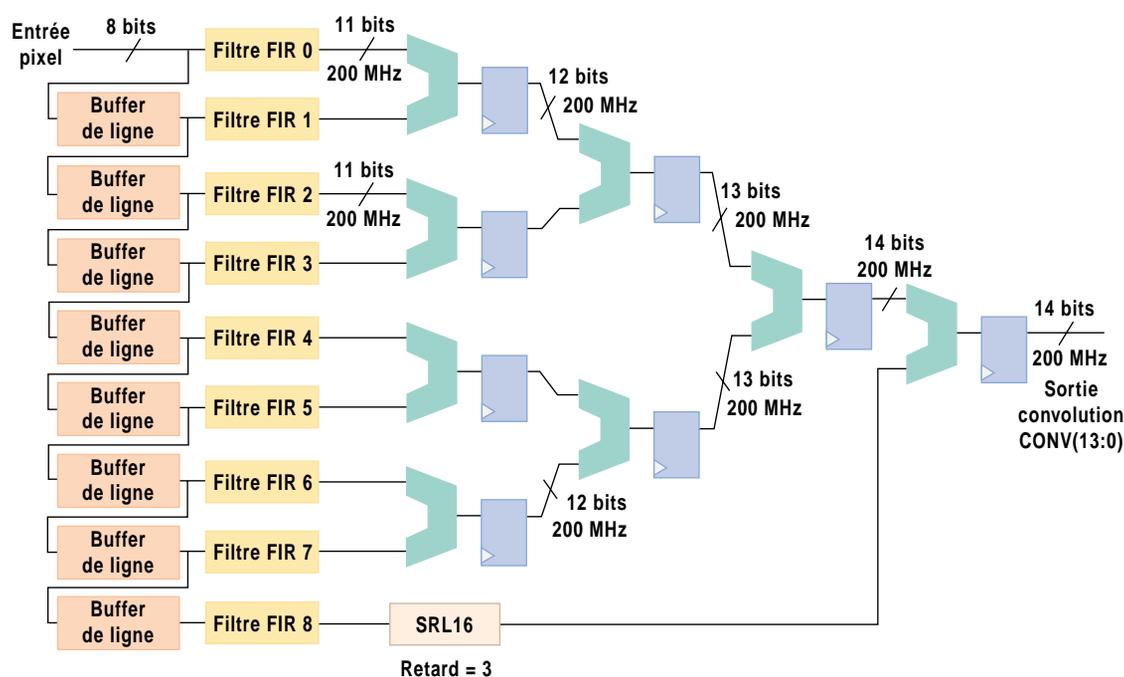
## Pour un filtre à 200 Mpixels/s

L'architecture Virtex-II met à disposition des multiplieurs 18x18 bits caractérisés par un temps de propagation de l'ordre de 5 ns. Cependant, dans le cas du module de convolution, la précision nécessaire est bien plus faible. Il est donc possible de

### Architecture d'une convolution 2D, 9x9

FIGURE 1

La fonction de convolution peut être structurellement décomposée en 9 filtres FIR de 9 pôles suivis d'un arbre d'additionneurs.



réduire la quantité de ressources requises en employant, à nouveau, une structure « Block SelectRam » de 18Kbits pour stocker les résultats de 9 multiplications prises en charge par chacun des 9 filtres. Et ceci pour chaque pixel entrant.

Cette organisation engendre d'autres avantages, notamment un routage plus simple. Utiliser moins de ressources d'interconnexion facilite le respect des contraintes sur les temps de propagation, donc sur la fréquence de fonctionnement globale. Enfin, une telle configuration tend à réduire la puissance dissipée.

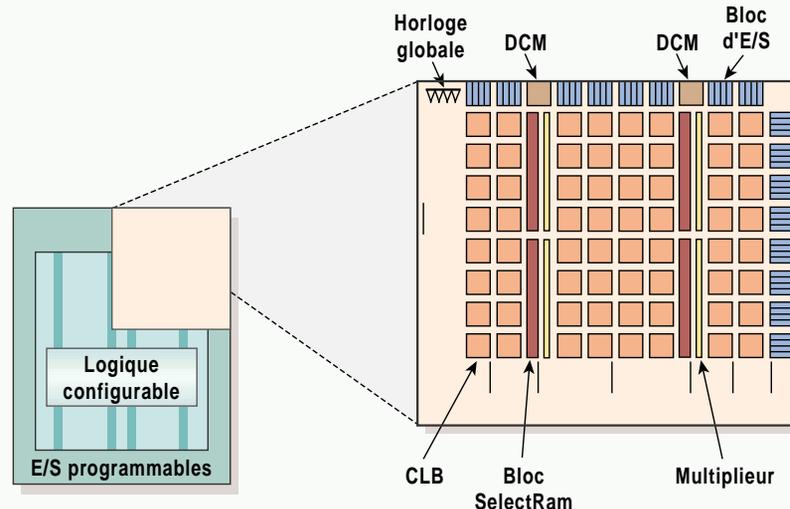
Revenons sur l'utilisation des blocs « Select-Ram » en tant que multiplieurs. Ces blocs peuvent devenir de réelles mémoires Rom double port, chaque port étant organisé suivant la configuration 512x36 bits. En reliant respectivement les signaux ADDRA(8) au VCC et ADDR8(8) au GND, nous obtenons une Rom synchrone de 256x72 bits (figure 2). Etudions maintenant les simplifications de calcul acceptables pour cette application. En arrondissant correctement les résultats des multiplications 8x8 bits, l'erreur sera d'un demi-LSB sur chaque résultat de 8 bits. A la fin de la convolution, l'erreur accumulée sera donc de  $81 \times 1/2$  LSB, soit 40 LSB sur la sor-

## Les blocs mémoire dans Virtex-II

→ Les blocs « SelectRam » dans l'architecture Virtex-II sont des mémoires Ram double port de 18 Kbits, configurables de différentes manières entre 18Kx1 bit et 512x36 bits. Chaque port est totalement synchrone et indépendant. Les blocs « SelectRam »

peuvent être mis en cascade pour réaliser de larges zones de stockage enfouies dans le FPGA. Un module multiplieur 18x18 bits est disposé à proximité de chaque bloc « Select-Ram » et est optimisé pour opérer sur le contenu d'un port de la mémoire. Les fonctions mémoire

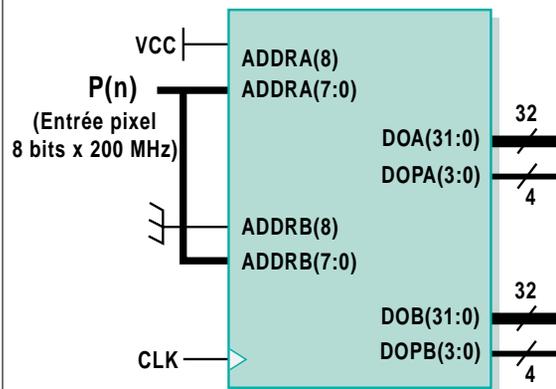
« SelectRam » et multiplieur sont toutes deux connectées aux matrices d'interconnexions leur ouvrant l'accès aux ressources de routage globales de la matrice FPGA. La figure ci-après présente un détail de l'architecture d'un composant Virtex-II.



### Utilisation des blocs de Ram

FIGURE 2

Sur les FPGA Virtex-II, un bloc « SelectRam » de 18Kbits peut être organisé en Rom synchrone de 256x72 bits.



tie de 14 bits de la convolution. Ceci signifie que les 5 bits de plus faible poids de ce résultat ne seront pas significatifs. En conclusion: si on enlève le bit de signe (soit le bit 13) inutilisé pour l'affichage vidéo, les 8 bits correspondant réellement à l'image après convolution seront les bits 12 à 5 de la sortie.

Ainsi, en considérant que le résultat de chaque multiplication peut être limité à 8 bits, à chaque cycle d'horloge, un bloc « Select-Ram » fournira les résultats de 9 multiplications. Ce qui veut dire qu'à 200Mpixels/s, le taux de traitement sera de 1,8 Gigamultipli-

Suite p.60

➤ cations/seconde pour chacun des 9 filtres, soit 16,2 Gigamultiplications/seconde pour l'ensemble du « convolveur ».

Etant donné que, sur les matrices Virtex-II, ces blocs « SelectRam » supportent sans problème une fréquence de fonctionnement de 200MHz, la performance de cet étage de filtres sera quasiment indépendante du placement et/ou du routage. Finalement, les 81 multiplieurs pourront être réalisés avec uniquement 9 blocs « SelectRam ». L'initialisation du contenu de la mémoire donnera le résultat de la multiplication des pixels par chacun des 9 coefficients (voir tableau).

Pour compléter l'architecture d'un filtre FIR à 9 pôles, il faut rajouter au bloc « Select-

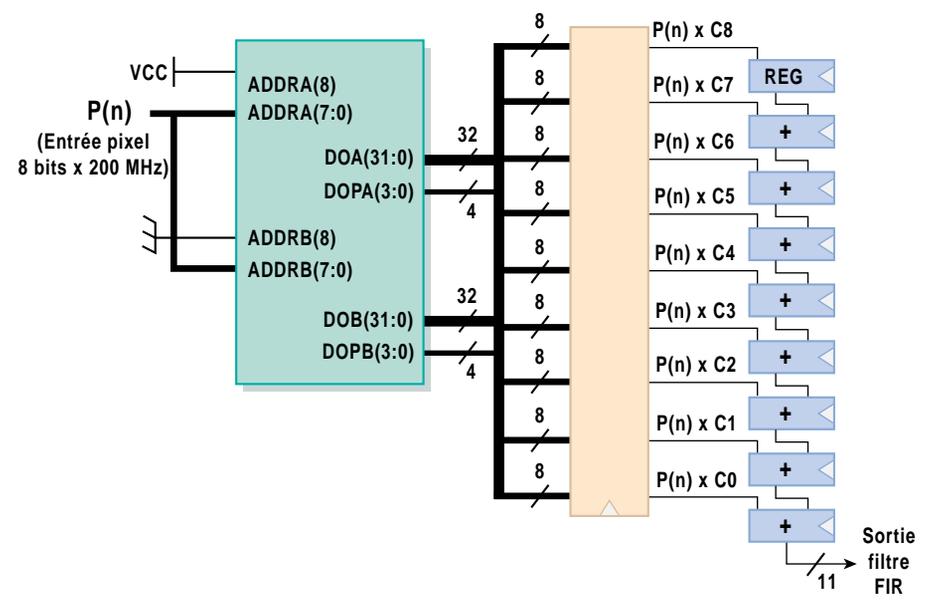
### Représentation de la succession des résultats de multiplication en sortie d'un bloc SelectRam

Cycle 0	Cycle 1	Cycle 2	Cycle 3	Cycle 4	Cycle 5	Cycle 6	Cycle 7	Cycle 8	Cycle 9
P(n0).C0	P(n1).C0	P(n2).C0	P(n3).C0	P(n4).C0	P(n5).C0	P(n6).C0	P(n7).C0	P(n8).C0	P(n9).C0
P(n0).C1	P(n1).C1	P(n2).C1	P(n3).C1	P(n4).C1	P(n5).C1	P(n6).C1	P(n7).C1	P(n8).C1	P(n9).C1
P(n0).C2	P(n1).C2	P(n2).C2	P(n3).C2	P(n4).C2	P(n5).C2	P(n6).C2	P(n7).C2	P(n8).C2	P(n9).C2
P(n0).C3	P(n1).C3	P(n2).C3	P(n3).C3	P(n4).C3	P(n5).C3	P(n6).C3	P(n7).C3	P(n8).C3	P(n9).C3
P(n0).C4	P(n1).C4	P(n2).C4	P(n3).C4	P(n4).C4	P(n5).C4	P(n6).C4	P(n7).C4	P(n8).C4	P(n9).C4
P(n0).C5	P(n1).C5	P(n2).C5	P(n3).C5	P(n4).C5	P(n5).C5	P(n6).C5	P(n7).C5	P(n8).C5	P(n9).C5
P(n0).C6	P(n1).C6	P(n2).C6	P(n3).C6	P(n4).C6	P(n5).C6	P(n6).C6	P(n7).C6	P(n8).C6	P(n9).C6
P(n0).C7	P(n1).C7	P(n2).C7	P(n3).C7	P(n4).C7	P(n5).C7	P(n6).C7	P(n7).C7	P(n8).C7	P(n9).C7

### Détails de l'architecture d'un filtre FIR

FIGURE 3

Le filtre FIR à 9 pôles peut être réalisé avec un bloc « SelectRam », un étage de multiplexeurs et une chaîne d'additionneurs.



Ram » 8 registres additionneurs pour obtenir les 11 bits du résultat (figure 3). Enfin, pour assurer les 5 ns de période au niveau des multiplexeurs et des divers routages associés, il est préférable d'insérer des bascules flip-flops pour réduire le Tco (Tco : temps de propagation « Clock to Out », plus important sur les blocs de Ram que sur les bascules, 0,4 ns au lieu de 2,9 ns).

En conclusion, l'implantation sur une matrice Virtex-II d'une fonction de convolution 2D, 9x9, fonctionnant réellement à 200 Mpixels/s sur une image de 1 K x 1 K pixels demande moins de 600 « slices » (1 bloc logique configurable comprend 4 « slices » et 2 buffers trois-états ; un « slice » réunit, entre autres, des générateurs de fonctions, des éléments de stockage et des multiplexeurs) et 12 blocs « SelectRam », les buffers de ligne compris. Cela représente la moitié des blocs logiques configurables (« CLB ») et des blocs Ram d'un FPGA 2V250. Enfin, bien d'autres astuces sont possibles avec cette architecture proposée par Xilinx, pour des versions de plus petites tailles ou plus lentes. ■