

# Micro-réseau sur puce compatible avec l'approche GALS

Ivan MIRO PANADES  
STMicroelectronics  
850, rue Jean Monnet  
38926 Crolles, France  
ivan.miro-panades@st.com

Alain GREINER  
Université Pierre et Marie Curie  
4, place Jussieu  
75252 Paris, France  
alain.greiner@lip6.fr

Abbas SHEIBANYRAD  
Université Pierre et Marie Curie  
4, place Jussieu  
75252 Paris, France  
abbas.sheibanyrad@lip6.fr

## Résumé

*Ce papier présente le micro-réseau DSPIN qui est une évolution de l'architecture SPIN. DSPIN est un micro-réseau à commutation de paquets pour architectures multiprocesseurs intégrés sur puce (MPSoC) utilisant l'approche GALS (Globalement Asynchrone, Localement Synchrones). L'architecture de DSPIN a une surface de silicium très petite et fournit des canaux de communication avec garanties de latence et de bande passante pour des applications soumises à des contraintes temps-réel.*

## 1. Introduction

L'architecture SPIN [1] est sans doute la première architecture de micro-réseau à commutation de paquets publiée [1]. Elle visait à résoudre le problème du goulot d'étranglement constitué par le bus système dans les architectures multiprocesseurs à mémoire partagée intégrées sur puce (MPSoC). Par la suite, un grand nombre d'architectures de type Network-on-Chip (NoC) ont été publiés [4,5,9,14,15]. En particulier, les concepteurs du micro-réseau ÆTHEREAL [4] insistent en la nécessité d'introduire des garanties de latence et de bande passante dans les communications de paquets pour des applications temps-réel [13].

Simultanément, l'implémentation matérielle d'un micro-réseau SPIN à 32 ports chez STMicroelectronics [11] a permis d'identifier les points faibles de cette architecture. Parmi les faiblesses de SPIN, on trouve l'approche complètement synchrone, qui n'est pas compatible avec les systèmes GALS, la topologie en arbre quaternaire élargi, qui est peu modulaire, et finalement la réalisation par macro-cellule optimisée, qui ne permet pas d'utiliser les bibliothèques de cellules précaractérisées fournies par le fondeur. Ceci nous a conduit à définir l'architecture DSPIN (Distributed, Scalable, Predictable, Integrated Network).

Comme SPIN, DSPIN supporte des architectures multiprocesseurs à mémoire partagée, mais fournit des services supplémentaires : il supporte l'approche GALS, il est synthétisable, et fournit des canaux de communication avec garanties de service (latence et la bande passante), pour une surface de silicium très petite.

Dans la section 2 nous présentons l'architecture de DSPIN, dans la section 3 nous présentons les résultats

expérimentaux des mesures de performances et l'estimation de la surface sur silicium.

## 2. Architecture de DSPIN

DSPIN est un micro-réseau destiné aux systèmes multiprocesseurs intégrés sur puce, possédant une architecture clustérisée, avec mémoire partagée. Le système complet est une composition de sous-systèmes synchrones appelés clusters. Chaque cluster peut contenir plusieurs processeurs, plusieurs mémoires, des contrôleurs d'entrée/sortie et un système d'interconnexion local. Chaque cluster peut avoir sa propre horloge qui peut être complètement asynchrone par rapport aux autres horloges. Ceci nous permet d'utiliser ce réseau dans des systèmes utilisant l'approche Globalement Asynchrone, Localement Synchrones (GALS) [2,3], où chaque cluster est localement synchrone mais asynchrone par rapport à ses voisins. Cette technique permet de réduire la puissance dissipée [12] et de simplifier la distribution des horloges [10].

### 2.1 Architecture distribuée et clustérisée

Pour éviter les situations d'interblocage entre les paquets requête/réponse, il y a deux réseaux physiquement indépendants pour les requêtes et pour les réponses. Ces deux réseaux partagent un unique contrôleur d'interface réseau (NIC). Les routeurs font l'aiguillage des paquets (requêtes ou réponses) sur le réseau et le NIC interconnecte le réseau avec les modules du cluster. La Figure 1 montre une architecture générique de cluster.

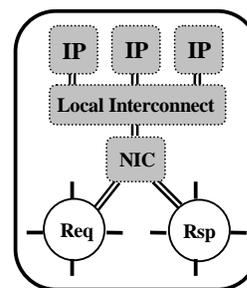


Figure 1. Architecture générique d'un cluster

Les IP sont connectés au NIC à travers le système d'interconnexion local, qui est l'unique porte d'accès au réseau. La topologie du réseau est une grille à deux dimensions où chaque case correspond à un cluster. La Figure 2 montre un exemple de réseau. Les clusters ont des liens bidirectionnels vers les clusters du Nord, Sud, Est et Ouest. Ces liens bidirectionnels sont composés par des FIFOs bi-synchrones qui ont une double fonctionnalité : tampons de stockage des paquets en cas de contention, et re-synchronisation lors du franchissement de la frontière entre deux domaines d'horloge.

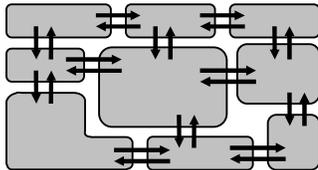


Figure 2. Topologie du réseau

## 2.2 Algorithme de routage

DSPIN utilise l'algorithme X-First [6] pour router les paquets. Cet algorithme a la particularité d'être déterministe et simple à implémenter. Avec cet algorithme, les paquets sont routés premièrement vers la direction X et puis vers la direction Y.

La décision du routage des paquets se fait grâce à l'entête des paquets qui contient l'adresse du cluster destinataire. Tout le paquet est retransmis de routeur à routeur jusqu'au destinataire en utilisant la technique de routage wormhole [7,8] (trou de ver).

## 2.3 Routeur DSPIN

Il y a deux niveaux de distribution dans DSPIN : Non seulement, les routeurs sont physiquement distribués dans le SoC, mais à l'intérieur d'un cluster, le routeur lui-même est décomposé en 5 modules (North, South, East, West et Local), qui sont physiquement disjoints : Les quatre premiers modules permettent la communication entre clusters voisins, et sont placés sur les 4 faces de la zone synchrone occupée par un cluster (Figure 3). Il y a donc deux types de fils d'interconnexion dans cette architecture : Les fils qui relient les modules appartenant à un même cluster sont appelés fils intra-cluster, et les fils qui interconnectent deux clusters voisins sont appelés fils inter-clusters.

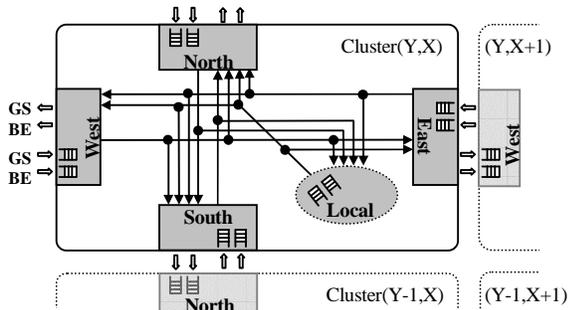


Figure 3. Architecture du routeur

Un des avantages d'utiliser l'algorithme X-First, est de réduire le nombre de connexions possibles dans un routeur. Par exemple les paquets qui proviennent du Nord, peuvent uniquement se diriger vers le Sud ou le port Local. Ceci simplifie le nombre de fils intra-cluster et aussi l'arité des multiplexeurs (1 parmi 2, pour les modules East et West)

## 2.4 Garantie de service

DSPIN admet deux types de trafic Best Effort (BE) et Guaranteed Service (GS). Le trafic GS a la particularité de fournir des garanties de latence bornée et de bande passante minimale. Il est utilisé principalement dans des applications avec contrainte temps-réel.

La cohabitation de ces deux types de trafic se fait grâce à l'utilisation des Canaux Virtuels (VC). Ces canaux permettent partager les mêmes ressources matérielles pour les deux types de trafic. Il existe cependant des FIFOs séparés pour les deux types de trafic.

Chaque module du routeur contient trois multiplexeurs, trois automates d'états et deux FIFOs. Ces FIFOs reçoivent les paquets BE et GS de façon séparée et leur contenu est transmis sur le canal virtuel, grâce à une technique de TDM (Time Division Multiplexing) pour garantir que chaque type de trafic peut obtenir au moins 50 % de la bande passante. Si une des FIFOs est vide, alors 100% du trafic est alloué à l'autre FIFO. La Figure 4 montre le détail du module Ouest.

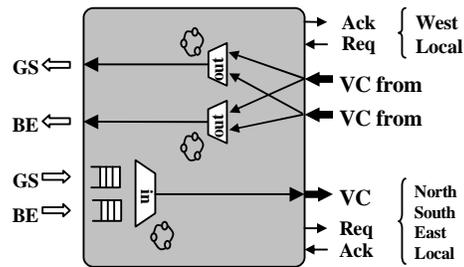


Figure 4. Détail du module Ouest du routeur

## 3. Résultats expérimentaux

Deux types de modèles ont été développés pour les composants matériels du réseau DSPIN. Des modèles de simulation écrits en SystemC précis au cycle près, et des modèles synthétisables de niveau RTL (Register Transfer Level), écrits en VHDL. Une plate-forme de simulation comportant 10\*10 clusters a permis d'analyser les performances du réseau. Chaque cluster contient un initiateur BE, une cible BE, une cible GS et optionnellement un initiateur GS. La latence moyenne des paquets a été analysée. Le paramètre variable est la charge offerte par chacun des initiateurs BE. La charge offerte par un initiateur est définie comme le pourcentage d'utilisation de la bande passante maximale du canal d'accès au réseau. C'est donc un paramètre compris entre 0% et 100%. Les conditions de simulation sont : distribution uniforme du destinataire du paquet et longueur des paquets distribuée uniformément entre 1 et 16 mots.

Les résultats montrent une latence déterministe pour les paquets de type GS, indépendante de la charge BE.

La latence des paquets BE dépend principalement de trois facteurs : la longueur des paquets BE, le nombre de clusters entre émetteur et récepteur et la charge du réseau. La Figure 5 montre la latence des paquets BE et GS en fonction de la charge injectée BE. Avec ces conditions de simulation, le point de saturation du réseau BE correspond à une charge maximale de 25%. A partir de ce seuil, la latence des paquets BE tend vers l'infini (ce qui correspond à une situation de saturation du réseau BE), tandis que la latence des paquets GS reste constante.

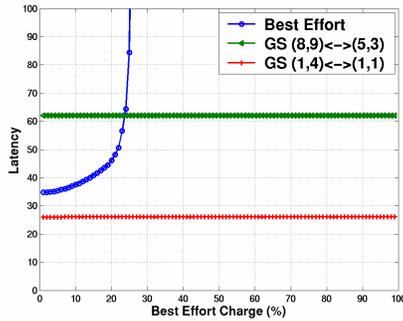


Figure 5. Latence des paquets BE et GS en fonction de la charge injectée BE

Les estimations de surface des routeurs et FIFOs ont été évaluées pour une technologie CMOS à 90nm. Tous les composants DSPIN sont synthétisables et peuvent être implémentés avec un flot de conception standard. La Table 1 montre les estimations de surface par routeur obtenue après synthèse.

Composant	Surface
5 FIFOs pour BE	0.036 mm <sup>2</sup>
5 FIFOs pour GS	0.018 mm <sup>2</sup>
Routeur (sans FIFOs)	0.028 mm <sup>2</sup>
<b>Total</b>	<b>0.082 mm<sup>2</sup></b>

Table 1. Estimation de surface par routeur

Les FIFOs pour le trafic BE ont une profondeur de 8 mots de 34 bits, celles pour le trafic GS ont une profondeur de 4 mots de 34 bits.

#### 4. Conclusions

L'expérience acquise, grâce à l'implantation sur silicium du réseau SPIN à 32 ports, nous a permis de définir l'architecture DSPIN. Cette architecture est adaptée aux systèmes Globalement Asynchrones, Localement Synchrones (GALS). La topologie en grille et l'algorithme de routage X-First nous permettent obtenir une surface de routeur très réduite. Nous avons démontré que la technique de canal virtuel (utilisé à l'intérieur d'un même cluster, et non entre deux clusters) permet de fournir des garanties de service qui assurent une latence et une bande passante strictement bornées pour certaines classes de trafic.

La surface de silicium par routeur après synthèse est de 0.082mm<sup>2</sup> avec une technologie CMOS à 90nm. Ceci est significativement plus petit que d'autres micro-réseaux publiés qui offrent le même type de service.

#### Références

- [1] P. Guerrier and A. Greiner. A generic architecture for on-chip packet-switched interconnections, Proc. Design Automation and Test in Europe (DATE'00), pp. 250-256, Mars 2000.
- [2] D. M. Chapiro. Globally-Asynchronous Locally-Synchronous systems. PhD thesis, Stanford University, 1984.
- [3] J. Muttersbach, T. Villiger, K. Kaeslin, N. Felber and W. Fichtner. Globally-Asynchronous Locally-Synchronous Architectures to Simplify the Design of On-CHIP Systems, Proc. 12th International ASIC/SOC Conference, pp. 317-321, Sept. 1999.
- [4] E. Rijpkema, K. Goossens, A. Radulescu, J. Dielssen, J. van Meerbergen, P. Wielage and E. Waterland. Trade-offs in the design of a router with both guaranteed and best-effort services for networks on chip, IEE. Proc.-Comput. Digit. Tech., Vol. 150, No 5, September 2003.
- [5] T. Bjerregaard and J. Sparsø. A router architecture for connection-oriented service guarantees in the MANGO clockless Network-on-Chip, IEEE Proc. Design Automation and Test in Europe (DATE'05), March 2005.
- [6] W. J. Dally and C. L. Seitz. Deadlock free message routing in multiprocessor interconnection networks, IEEE Transactions on Computers. C-36, 5, pp. 547-553, May 1987
- [7] W. J. Dally and B. Towles. Route packets, not wires: on-chip interconnection networks, Design Automation Conference (DAC 2001), June 2001.
- [8] L. M. Ni and P.K. McKinley. A survey of wormhole routing techniques in direct networks, IEEE Computer 2 (1993) 62-75.
- [9] E. Bolotin, I. Cidon, R. Ginosar and A. Kolodny. QNoC: QoS architecture and design process for network on chip, Journal of Systems Architecture, 50(2-3), pp. 105-128, February 2004.
- [10] D. Peiliang, Y. Rilong, X. Hongbo and Y. Chengfang. Multi-clock driven system: a novel VLSI architecture, Proc. 4th Int. Conf. ASIC, pp. 555-558, 2001.
- [11] A. Andriahantenaina and A. Greiner. Micro-network for SoC: Implementation of a 32-port SPIN network, DATE 2003, Munchen, Germany, pp. 1128-1129, March 2003
- [12] W. Qing, M. Pedram and X. Wu. Clock-gating and its application to low power design of sequential circuits, IEEE Trans. Circuits Syst. I, Fundam. Theory Applicat., vol. 47, no3, pp.414-420, Mars 2000.
- [13] K. Goossens, J. van Meerbergen, A. Peeters and P. Wielage. Networks on Silicon: Combining Best-Effort and Guaranteed Services, Design Automation and Test in Europe (DATE'02), 2002.
- [14] M. Millberg, E. Nilsson, R. Thid and A. Jantsch. Guaranteed bandwidth using looped containers in temporally disjoint networks within the Nostrum network on chip, IEEE Proc. Design Automation and Test in Europe (DATE'04), vol. 2, pp. 890 - 895, February 2004.
- [15] D. Bertozzi and L. Benini. Xpipes: A Network-on-Chip architecture for gigascale Systems-on-Chip, IEEE Circuits and Systems Magazine, Q2 2004.