

Analyse Énergétique et Conception Optimisée RISC-V pour l'IA

Projet de Recherche (PRe)
mai - juillet 2022

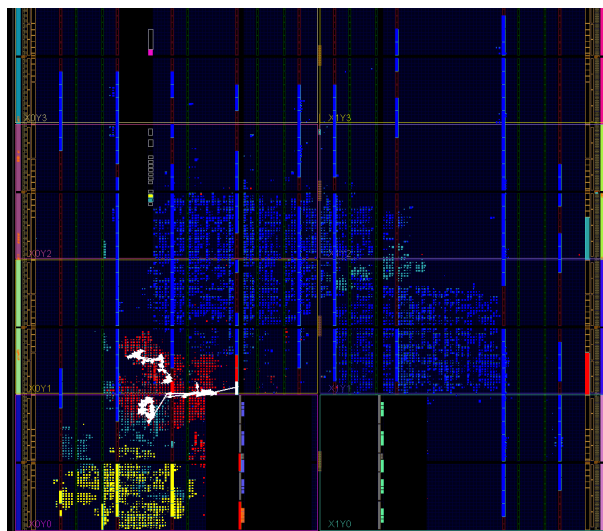
Bastien HUBERT - Promotion 2023

Ecole Nationale Supérieure des Techniques Avancées de Paris

Tuteur de Stage : Omar HAMMAMI

Organisme d'accueil : U2IS

Adresse : 91762 Palaiseau - France



Remerciements :

Je tiens à remercier toute l'équipe de l'U2IS pour leur accueil chaleureux et la qualité du temps que j'ai passé avec eux.

Je remercie tout particulièrement Omar Hammami pour son encadrement, sa confiance et la qualité de ses conseils tout au long de mon stage.

Un grand merci à Hervé Le Provost et Farhat Thabet, qui ont suivi mon stage avec attention et bienveillance, et qui m'ont fourni des conseils précieux pour me permettre de mener à bien.

Merci à François Pessaux qui a su m'aider à résoudre bon nombre de problèmes informatiques épineux.

Merci enfin à Lorraine Brisacier-Porchon, Roua Chaouachi et Wael Sabri sans la compagnie desquels mon stage aurait été bien plus monotone.

Résumé :

L'architecture RISC-V fait l'objet d'une diffusion et d'une exploitation croissante dans de nombreux domaines d'applications. Les prédictions sur l'exploitation du RISC-V dans l'IA indiquent par ailleurs que cette architecture sera exploitée de manière intensive dans un futur proche.

Pour ces raisons, nous avons analysé les performances énergétiques du processeur RISC-V ARIANE/CV32A6, écrit dans le langage SystemVerilog, par simulation avec l'outil Ques-tasim et tenté de modifier sa conception pour y introduire des techniques de réduction de la consommation énergétique.

Des comparaisons entre les résultats obtenus et l'état de l'art de la recherche nous ont permis de proposer des optimisations au niveau du compromis performances/énergie/ressources, que nous avons implémenté sur le processeur sur circuit FPGA Xilinx Zynq sur carte Zybo Z7. Des mesures de performances spatiales ont enfin été réalisées afin de compléter le jeu de données obtenu.

Mots-clés : RISC-V, basse consommation, planification de circuit, simulation, FPGA, recherche opérationnelle

Abstract :

The RISC-V architecture is becoming increasingly popular and exploited in many application areas. Predictions about the use of RISC-V architectures in AI also indicate that they will be intensively exploited in the near future.

For these reasons, we have analysed the energy performance of the ARIANE/CV32A6 RISC-V processor, written in SystemVerilog, by simulation with the Questasim tool, and tried to modify its design to introduce low-power techniques.

Comparisons between the obtained results and the state of the art of research allowed us to suggest optimisations regarding the performance/energy/resource trade-off, which we then implemented on the processor, using a Xilinx Zynq FPGA on Zybo Z7 board. Finally, spatial performance measurements were performed in order to complete the obtained dataset.

Keywords : RISC-V, low-power design, floorplanning, simulation, FPGA, optimisation

Table des matières :

1	Introduction	6
1.1	Un peu d'histoire	6
1.2	RISC-V	6
2	La recherche	6
2.1	Puissance statique, puissance dynamique	6
2.2	Pourquoi optimiser la consommation énergétique ?	9
2.3	Les paramètres de l'équation de la puissance dynamique	9
2.4	Le Clock Gating	9
2.5	Le Power Gating	9
2.6	Dynamic Frequency Management	9
2.7	Dynamic Power Management	9
2.8	À propos de la fréquence d'horloge	9
3	Les outils utilisés	9
3.1	SystemVerilog	9
3.2	Vivado	9
3.3	Vitis	9
3.4	Questasim	9
3.5	Zybo Z7	9
3.6	Zynq 7020	9

4	Méthodologie de résolution	9
4.1	9

Table des figures :

Table des tableaux :

1 Introduction

1.1 Un peu d'histoire

1.2 RISC-V

2 La recherche

2.1 Puissance statique, puissance dynamique

$$P_{tot} = P_{dyn} + P_{CS} + P_{leak}$$

Puissance de fuite :

Puissance de court-circuit :

Calcul de la puissance dynamique :

On considère un circuit électrique alimenté à tout instant par une tension V et parcouru par un courant i . Ce circuit est contrôlé par une horloge de fréquence f_{CK} , et changeant d'état en moyenne α fois par unité de temps. α est appelé facteur d'activité, et permet de définir une fréquence f , appelée fréquence effective, telle que le circuit étudié change d'état en moyenne autant de fois sur une même période que s'il était simplement cadencé par un signal d'horloge de fréquence f . On en déduit donc que $f = \alpha \times f_{CK}$.

La puissance dynamique instantanée à tout instant t d'un tel circuit est alors donnée par la formule :

$$P_{dyn, instant}(t) = i(t) \times V(t) \tag{1}$$

La puissance dynamique est la fraction de la puissance totale qui provient de la charge et de la décharge des condensateurs après un changement d'état des transistors du circuit. On suppose ici que $t_2 - t_1 \gg 1/f$ de telle sorte à négliger les effets de bords et permettre l'approximation selon laquelle le circuit n'est soumis qu'à une horloge de fréquence f . À

chaque front montant de l'horloge, une charge $Q = C \times V$ est nécessaire pour charger un condensateur de capacité C , lequel est déchargé à chaque front descendant. On en déduit donc que l'intégrale du courant qui circule dans le circuit entre deux instants t_1 et t_2 est donnée par :

$$\begin{aligned} \int_{t_1}^{t_2} i(t) dt &= \int_{t_1}^{t_2} \frac{dQ}{dt} dt \\ &= [Q]_{t_1}^{t_2} \\ &= C \times V \times (t_2 - t_1) \times f \end{aligned} \quad (2)$$

L'énergie consommée entre l'instant t_1 et l'instant t_2 est définie comme étant l'intégrale de la puissance entre les deux instants :

$$E^{t_1, t_2} = \int_{t_1}^{t_2} P_{dyn, instant}(t) dt \quad (3)$$

Enfin, on définit la puissance dynamique (moyenne) comme étant la moyenne de la puissance dynamique instantanée sur l'intervalle de temps $[t_1, t_2]$:

$$\begin{aligned} P_{dyn}^{t_1, t_2} &= \frac{E^{t_1, t_2}}{t_2 - t_1} \\ &= \frac{1}{t_2 - t_1} \int_{t_1}^{t_2} P_{dyn, instant}(t) dt && \text{d'après (3)} \\ &= \frac{1}{t_2 - t_1} \int_{t_1}^{t_2} i(t) \times V(t) dt && \text{d'après (1)} \\ &= \frac{V}{t_2 - t_1} \int_{t_1}^{t_2} i(t) dt && \text{V supposé constant entre } t_1 \text{ et } t_2 \\ &= \frac{V}{t_2 - t_1} C \times V \times (t_2 - t_1) \times f && \text{d'après (2)} \\ &= C \times V^2 \times f \\ &= C \times V^2 \times \alpha \times f_{CK} \end{aligned} \quad (4)$$

En particulier, on voit que $P_{dyn}^{t_1, t_2}$ est indépendant de t_1 et t_2 , et on écrira simplement P_{dyn} pour parler de la puissance dynamique :

$$P_{dyn} = C \times V^2 \times \alpha \times f_{CK} \quad (5)$$

Ce résultat est essentiel pour la suite car c'est en jouant sur les paramètres du membre de droite que nous allons pouvoir réduire la consommation énergétique du circuit (ou d'une partie du circuit) sans en impacter visiblement le besoin en ressources ou les performances.

2.2 Pourquoi optimiser la consommation énergétique ?

2.3 Les paramètres de l'équation de la puissance dynamique

2.4 Le Clock Gating

2.5 Le Power Gating

2.6 Dynamic Frequency Management

2.7 Dynamic Power Management

2.8 À propos de la fréquence d'horloge

3 Les outils utilisés

3.1 SystemVerilog

3.2 Vivado

3.3 Vitis

3.4 Questasim

3.5 Zybo Z7

3.6 Zynq 7020

4 Méthodologie de résolution

4.1