Memòria del projecte de tesi:

The subscalar microarchitecture for ultra-low power

Doctorand:	Ramon Canal Corretger
Director de tesi:	Antonio González Colás (Universitat Politècnica de Catalun
Codirector de tesi:	James E. Smith (University of Wisconsin-Madison, USA)
Tutor:	Antonio González Colás

UPC UNIVERSITAT POLITÈCNICA DE CATALUNYA

Departament d'Arquitectura de Computadors

Dades d'identificació:

Doctorand:	Ramon Canal Corretger
Director de tesi:	Antonio González Colás (Universitat Politècnica de Catalu
Codirector de tesi:	James E. Smith (University of Wisconsin-Madison, USA)
Tutor:	Antonio González Colás

Títol provisional de la tesi:

The subscalar architecture for ultra-low power.

1. Motivació

Des de fa poc temps, els sistemes "embedded" s'han convertit en un dels objectius preferits les grans companyies productores de processadors. El mercat potencial per aquests dispositiu molt gran: des de petits sistemes com telèfons mòbils o "handhelds" als ja tradicionals ordina portàtils. Més enllà d'aquests límits, ja ens trobaríem per una part amb processador d' prestacions i per l'altre amb sistemes específics. Tot i això, i cada vegada més, els límits e aquestes categories són difosos.

Dins d'aquestes arquitectures, té una importància especial el reduir el consum de potè De fet, podem dir que aquestes arquitectures estan dissenyades amb un èmfasis especial sob consum. Si tenim en compte que aquests dispositius funcionen amb piles o bateries -que ha tenir un tamany compacte i reduït, per raons d'espai- el sistema en si ha de consumir p eficientment l'energia. Al mateix temps, ha d'intentar mantenir el rendiment que tindri processador sense les restriccions de consum.

El problema del consum es pot atacar des de diferents perspectives. En un sistema *embed* el consum de potència pot ser regulat pel sistema operatiu -apagant alguna secció del dispo perquè no es fa servir. També es pot controlar el consum mirant que els components tingui ús eficient i mínim de la potència, en el nostre cas, aquest component és el processado processador, com a nucli central de processament és una de les part més actives i que pot ar a consumir gran part de l'energia disponible. Per exemple, l'alpha 21264 consumeix en r 75Watts [8]. Tot i no ser un processador per a sistemes *embedded* (que consumeixen entre 3 Watts), veiem que el consum d'energia pot arribar a ser un problema.

De fet, el problema del consum no prové només de la necessitat de fer durar unes bat sinó també del problema que suposa el refredament del processador. És ben conegut qu velocitat de transmissió dels electrons es veu deteriorada per altes temperatures. Per tant, cal en compte el consum d'energia també com a font de l'escalfament general del processa

Aquest escalfament fa perillar tan el rendiment del processador com la fiabilitat d'aquest ja la diferència de temperatura entre diferents parts del processador pot portar a trencaments material no pot suportar aquests pics de temperatura (*hot spots*) del processador.

El consum es pot dividir en dos components: el consum estàtic i el dinàmic. El consum es (*leakage power*) prové de la pèrdua d'energia degut al disseny dels transistors i que fa que hi un petit corrent entre la font (*source*) i el terra (*drain*). El consum dinàmic prové d capacitància dels components (*capacitive power*) i dels curt-circuits en activar/desactiva transistors (*short-circuit power*). Per les futures tecnologies [6], més enllà de les 0.25µm component estàtic tindrà una importància creixent ja que fins ara, la major part del consum p de la conmutació dels circuits. Per tant, és interessant desenvolupar processadors que minim el número de transistors o, en altres paraules, que utilitzin eficientment tots els transistors ja altrament seran una font de consum estàtic.

Per altra banda, s'ha demostrat que en les aplicacions que típicament executen aq processadors, gran part del números utilitzats no necessiten tota l'amplada de bits que ofere processador per poder ser representats. En altres paraules, molts dels processadors utilitze bits per representar els seus valor (registres dins del processador). Amb 32 bits podem represe fins a 2^{32}-1 però realment, els números que se solen fer servir són més petits. A la taula 1, po veure la distribució de les instruccions executades (en mitja pels Mediabench [9], un recu programes típicament executats per aquest tipus de processadors) segons el número de significatius dels seus operands. O sigui, el màxim número de bits significatius que tene operands (d'entrada i de sortida).

.

Taula 1: Distribució de les instruccions d'enters segons l'amplada mínima de datapath que necessiten (Avg Mediabench)

#bits significants	percentatge	acumulat
4 bits	19.3566	19.3566
8 bits	13.2336	32.5902
12 bits	7.0438	39.6340
16 bits	7.7309	47.3649
20 bits	4.9191	52.2840
24 bits	5.9346	58.2186
28 bits	3.2753	61.4939
32 bits	38.5061	100

En conclusió, donada la tendència a reduir el consum del processador i vistes estadístiques sobre el tamany dels operands, en aquesta tesi es proposa estudiar el disseny processador que contempli aquestes característiques. Aquest processador estarà basat en tècniques de segmentació presents en tots els processadors actuals [12] i haurà d'incl mecanismes que permetin explotar el fet de que els operands no sempre són de 32 bits. A això suposa un nou concepte per reduir el consum de potència ja que tradicionalment els est sobre consum han estat encarats a desconnectar parts del processador [11] i no en realitza disseny des de zero del processador pensant en el consum.

2. Resum del projecte

El projecte consistirà en el desenvolupament lògic d'un processador destinat per a tecnolo futures tenint en compte les restriccions que imposaran l'ús d'aquesta tecnologia. Aque restriccions són les següents:

1. Només un percentatge de l'àrea del xip és accesible en un sol cicle.
2. El component estàtic del consum té un pes considerable dins del consum total.
3. El consum total del disseny és un paràmetre crític per la seva avaluació.

La proposta per aquesta tesis és la de realitzar una organització basada en un processador es que intenti reduir al màxim el consum mantenint el rendiment. L'arquitectura base serà la processador 1-way en-ordre de 32 bits. Aquest processador tindrà un pipeline de 5 etapes (*Fe Decode*, *Execute*, *Memory* i *Writeback*). A l'etapa de *fetch* el processador accedirà la c d'instruccions per llegir la següent instrucció a executar. Al *decode*, la instrucció es decodifi i llegirà els operands que tingui al banc de registres. La tercera etapa, l'execució, consistir l'ús de l'unitat funcional que necessiti l'instrucció. Després vindrà l'accés a memòria (instrucció ho requereix) i les instruccions acabaran fent l'escriptura del registre destí al ban registres.

L'arquitectura proposada tindrà un pipeline de 8 bits, per tant, podríem dir que l'arquitec de sortida (8 bits) tindrà un rendiment pròxim a 4 vegades pitjor que l'arquitectura de 32 bits tal de millorar aquest dèficit, l'arquitectura proposada només es realitzarà el càlcul dels b significatius. Per tant, dependrà del número de bytes que s'hagin de calcular que l'instrucció més o menys. Tot i aquesta arquitectura base, també s'avaluaran altres alternatives com *pipelines* de 4 bits i 16 bits.

Un cop realitzada la base, s'intentarà expandir el coneixements apresos per aqu arquitectura per aplicar-los a altres microarquitectures ja conegudes i, també, estudiar l'influè que pot tenir el compilador de cara a minimitzar el consum d'energia dins del processador.

3. Objectius

En aquesta tesis es pretén desenvolupar una arquitectura per tecnologies que en els propers seran a l'abast per veure quines característiques imposa en el disseny dels processadors futu

L'arquitectura dissenyada s'avaluarà i es buscaran les avantatges i desavantatges qu respecte el rendiment (mesurat en instruccions per cicle), el consum (mesurat tan en ter d'energia com en activitat) i en la complexitat (mesurant la lògica adicional).

Es pretén també, aprofundir en el coneixement de com modelitzar, mesurar i analitz consum de potència a nivell de microarquitectura. També es volen proporcionar eines realitzin aquests treballs per així facilitar la feina que es pugui fer després de realitzar aqu tesis i que tracti temes de consum.

4. Estat de l'art

Només hi ha el treball de Brooks i Martonosi [2] que faci referència al tamany dels operands aquest treball, els autors intenten aprofitar els operands petits (*narrow operands*) per enviar d'una instrucció a una sola unitat funcional. En altres paraules, intenten simular dinàmicamen funcionament similar al que fa l'extensió MMX als processadors Intel. La diferència és qu l'arquitectura Intel les instruccions són detectades a nivell de compilació i tenen un codi espe En el treball de Brooks i Martonosi [2], les instruccions són detectades dinàmicament.

El treball que es realitzarà en aquesta tesis difereix molt del treball descrit al paràgraf ant ja que tot i usar una mateixa característica, els propòsits són totalment diferents. De fet, l'obj d'aquesta tesis és molt més ambiciós que el desenvolupat en aquell treball i, a més, en aqu tesis es pretén fer un estudi de la complexitat que requereix detectar i utilitzar la llargada operands com a factor per reduir el consum. En aquell treball, els autors no considere complexitat que pot tenir a l'hora de fer l'*issue* el detectar el tamany dels operands i decidir s el seu destí a les unitats funcionals.

A part d'aquest treball, n'hi ha d'altres que desenvolupen tècniques per reduir el con [1][11][13]. Aquestes tècniques es limiten a estudiar el funcionament d'una part del process i tracten d'optimitzar-lo. En la tesis proposada, es pretén dissenyar un processador partint de que tingui com a principal característica el mínim consum. Altres treballs mostren difer mètodes per estimar el consum que poden tenir diferents parts del processador [3][5][10]. aquests treballs són complementaris amb el treball que es realitzarà. Apart, caldrà tenir en co les publicacions sobre processadors actuals i que són una mostra del que es pot fer am tecnologia actual. Aquestes informacions seran de gran valor per conèixer les característi previsibles dels processadors *embedded* en els propers anys.

5. Pla de treball

Primerament, s'ha realitzat una recerca bibliogràfica sobre el tema de baix consum (*low po* i les últimes propostes sobre les arquitectures de processadors *embedded*. Al mateix temps, s desenvolupat les tècniques per aprofitar el tamany dels operands.

A partir d'aquesta base de coneixement, es realitzarà una proposta d'una arquitectura mí on s'apliquin les tècniques realitzades.

Tenint l'arquitectura base es faran simulacions per veure el seu rendiment/consum i ta per veure les penalitzacions que tenen les instruccions en la nova arquitectura. Aquest es permetrà realitzar una proposta d'una arquitectura on el pipeline estigui equilibra proporcionat). O sigui, per exemple, que si resulta que si el *pipeline* es bloqueja molt so perquè li falten instruccions per executar llavors caldrà millorar el *fetch*.

Un cop s'ha obtingut un pipeline equilibrat es farà una proposta d'arquitectura per p veure el màxim rendiment que pot arribar a obtenir aquesta arquitectura.

Després d'aquest treball s'estudiarà com implantar mecanismes dissenyats per processador d'altes prestacions però a un cost més baix (predictor de salts, execució fora d'or D'aquesta manera l'arquitectura base s'anirà transformant per fer un processador *embedded* ja amb més prestacions.

Després d'aquest pas s'introduiran les tècniques desenvolupades per als processa *embedded* als processadors d'altes prestacions (superescalars, *multithread* microarquitectures *clustered*, entre altres). Creiem que és important també realitzar contribuc en aquest camp ja que aviat també en aquest mercat pot ser crític el factor del consum. Aquí un paper molt important tant l'experiència del grup d'investigació com la del doctorand ja aquest ja ha realitzat una sèrie de treballs en aquestes arquitectures [14][15][16][17].

Finalment, s'estudiarà com el compilador i, per tant, el codi generat pot tenir una influè en el consum de potència. En aquest punt, es pretén aprofundir en la reducció en el consu potència però des de tècniques dissenyades en temps de compilació.

6. Recerca bibliogràfica

[1] S.G. Abraham and S.A. Mahlke, "*Automatic and Efficient Evaluation of Mem Hierarchies for Embedded Systems*", 32nd Annual Int. Symp. on Microarchitec (MICRO-32), 1999, pp. 114-125.

[2] D. Brooks and M. Martonosi, "Dynamically Exploiting Narrow Width Operand Improve Porcessor Power and Performance", in Proc. of 5th. Int. Symp. on H Performance Computer Architecture (HPCA-5), Orlando (USA).

[3] Z. Chen and K. Roy, “A Power Macromodeling Technique Based on Power Sensitiv in the Proceedings of the 35th ACM/IEEE conference on Design Automation Confer (DAC 98), San Francisco (USA), 1998, pp. 678-683.

[4] D. Burger, T.M. Austin, S. Bennett, “Evaluating Future Microprocessors: SimpleScalar Tool Set”, Technical Report CS-TR-96-1308, University of Wisco Madison, 1996.

[5] G. Z.N. Cai, K. Chow, T. Nakanishi, J. Hall and M. Barany, “Multivariate Po Performance Analysis for High Performance Mobile Microprocessor Design”,

[6] D.J. Eaglesham, “*0.18µm CMOS and beyond*”, Proceedings of the 36th ACM/I conference on Design Automation Conference (DAC 99), New Orleans (USA), 1999 703-708.

[7] J.A. Fisher, “Customized Instruction-Sets for Embedded Processors”, in Proceedin the 36th ACM/IEEE conference on Design Automation Conference (DAC 99), Orleans (USA), 1999, pp. 253-257.

[8] M.K. Gowan, L.L. Biro and D.B. Jackson, “Power Considerations in the Design o Alpha 21264 Microprocessor”, in the Proceedings of the 35th ACM/IEEE conferenc Design Automation Conference (DAC 98), San Francisco (USA), 1998, pp. 726-731

[9] C. Lee, M. Potkonjak and W. H. Mangione-Smith, “Mediabench: A Tool for Evalua and Synthesizing Multimedia and Communications Systems”, *Proc. of the IEEE/A Int. Symposium on Microarchitecture* (Micro 30), December 1997, pp. 330-335.

[10] E. Macii, M. Pedram and F. Somenzi, “Hihg-Level Power Modelling, Estimation, Optimization”, in Proceedings of the 34th ACM/IEEE conference on Design Automa Conference (DAC 97), Anaheim (CA-USA), 1999, pp. 504-511.

[11] E. Musoll, “*Predicting the usefulness of a block result: a micro-architectural techn for high-performance low-power processors*”, 32nd Annual Int. Symp. Microarchitecture (MICRO-32), 1999, pp. 238-247.

[12] D.A. Patterson, J.L. Hennessy, “Computer Organization and Design: The Hardv Software Interface”, Morgan Kaufmann Publishers, 1994, pp.362-451.

[13] W.T. Shiue and C. Chakrabarti, “Memory Exploration for Low Power, Embe Systems”, in Proceedings of the 36th ACM/IEEE conference on Design Automa Conference (DAC 99), New Orleans (USA), 1999, pp. 140-145.

7. Publicacions

Treball previ de microarquitectures “clustered”.

[14] Ramon Canal, Joan Manuel Parcerisa and Antonio González, "*Dynamic Cl*
Assigment Mechanisms", in Proc. of 6th. Int. Symp. on High-Performance Comp
Architecture (HPCA-6), Toulouse (France), Jan. 10-12, 2000 (Best student paper)

[15] Ramon Canal, Joan Manuel Parcerisa and Antonio González, "*A Cost-Effective Clust*
Architecture", in Proc. of Int. Conf. on Parallel Architectures and Compilation Techni
(PACT-99), New Port Beach (USA), Oct. 12-16, 1999

[16] Ramon Canal, Joan Manuel Parcerisa and Antonio González, "*Dynamic C*
Partitioning for Clustered Architectures", To appear in the International Journa
Parallel Programming.

Treball previ “Issue logic for highly pipelined processors”.

[17] Ramon Canal and Antonio González, "*A Low-Complexity Issue Logic*", in Proc. o
International Conference on Supercomputing (ICS-00). Santa Fe (USA). May 8-11, 2