

KLASYFIKACJA SYSTEMÓW WSPOMAGAJĄCYCH PROCES PRZETWARZANIA I STEROWANIA

Andrzej Stasiak

**Instytut Informatyki i Elektroniki, Uniwersytet Zielonogórski
65-246 Zielona Góra, ul. Podgórna 50**

e-mail: A.Stasiak@iie.uz.zgora.pl

STRESZCZENIE

Klasyfikacja oraz omówienie poszczególnych działów przeprowadzonego w artykule podziału systemów informatycznych wspierających proces przetwarzania i sterowania. Praca daje pogląd na stan bieżący akceleratorów cyfrowych oraz pozwala sprecyzować usytuowanie własnych prac w różnych domenach informatyki.

1. WPROWADZENIE

Na szerokim polu zastosowań użytkowych, inżynierskich i przemysłowych, bezustannie występuje zapotrzebowanie na coraz bardziej wydajne systemy obliczeniowe. Jest to wynik postępu technologicznego oraz globalnej informatyzacji społeczeństwa. Opracowywane są wciąż nowe metody projektowania złożonych systemów cyfrowych w celu uproszczenia procesu poznawczego, korekty błędów oraz zarządzania.

W kolejnych etapach rozwoju technologicznego opracowano tranzystor, zbudowano bramkę logiczną, mikroprocesor, koprocessor matematyczny, opracował technologię MMX, 3DNOW!, powstały konstrukcje specjalizowanych procesorów graficznych i dźwiękowych (firmy: 3dfx, Ati, nVidia, Aureal), systemy akceleracji specjalizowanych obliczeń matematycznych (Alatek-HES[15]) oraz dedykowane jednostki symulacyjno-emulacyjne (MentorGraphics IKOS[1], Cadence QuickTurn[14]). Opracowano systemy wieloprocessorowe (Sun Enterprise 10K, komputery macierzowe), systemy typu klastrer oraz farmy komputerowe. Cały wysiłek był i jest skupiony na zwiększeniu mocy obliczeniowej budowanych maszyn i systemów cyfrowych.

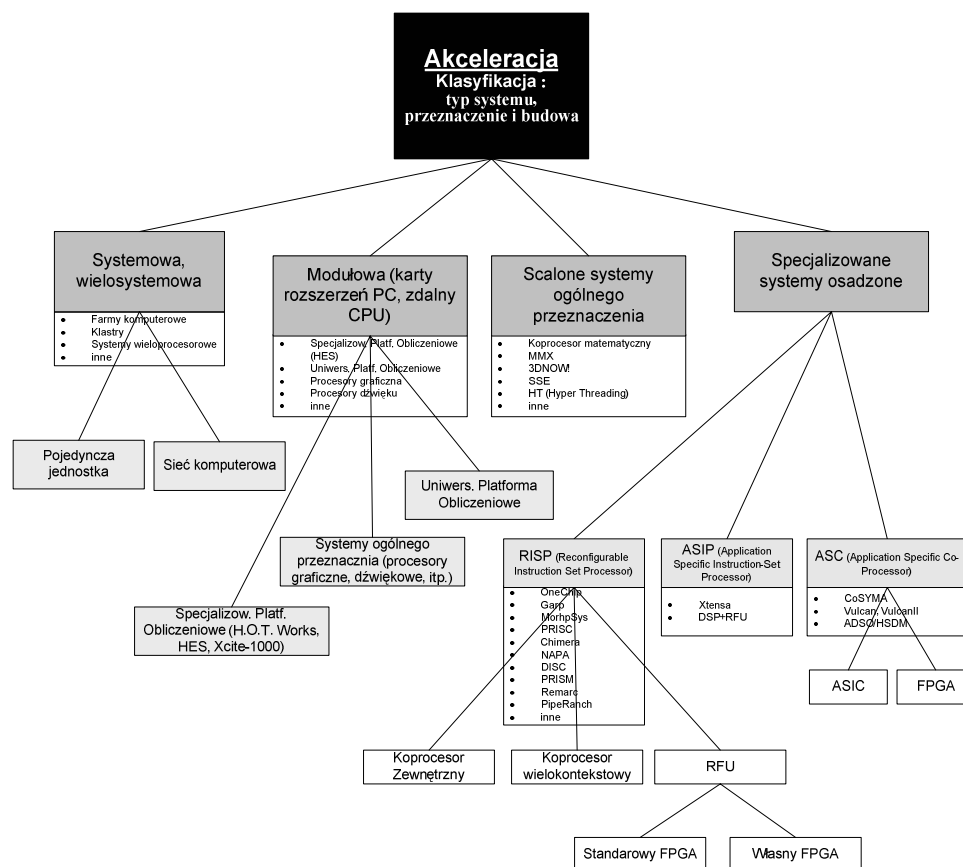
Potrzebę przyspieszenia wykonywania operacji przez maszyny cyfrowe zaobserwowano już na samym początku rozwoju techniki cyfrowej. Do systemu zawierającego procesor ogólnego przeznaczenia wprowadzono specjalizowany koprocessor matematyczny, którego zadaniem było wspomaganie głównej jednostki przetwarzania danych podczas wykonywania złożonych operacji matematycznych (operandy - duże liczby całkowite i rzeczywiste). Obecnie na całej płaszczyźnie badań i zastosowań informatycznych oraz telekomunikacyjnych, obserwuje się intensywne prace, których wspólnym celem jest osiągnięcie coraz bardziej wydajnych

systemów obliczeniowych. Jest to naturalne zjawisko obserwowane we wszystkich dziedzinach życia człowieka.

W pracy klasyfikacji poddano znane działy i technologie systemów cyfrowych, gdzie poprzez rozwiązania inżynieryjno-naukowe dąży się do przyspieszenia operacji przetwarzania danych i sterowania. Omawiając sklasyfikowane dziedziny, autor w głównej mierze skupił się na dziale dotyczącym specjalizowanych systemów osadzonych, które są przedmiotem badań w Instytucie Informatyki i Elektroniki Uniwersytetu Zielonogórskiego.

2. KLASYFIKACJA

Klasyfikację przeprowadzono ze względu na typ rozwiązania/systemu (systemy rozproszone, jednostka robocza, system operacyjny, układ cyfrowy), jego budowę (karta rozszerzeń PC, procesor/koprocessor) i przeznaczenie (ogólnego przeznaczenia, specjalizowane). Rysunek nr 1 przedstawia ogólną klasyfikację systemów akceleracji obliczeń.



Rys. 1. Klasyfikacja systemów

Akceleracja systemowa i wielosystemowa.

Rozwiązania zawarte w tej grupie dotyczą między innymi systemów wieloprocessorowych, których właściwości coraz częściej są dostrzegane i efektywnie wykorzystywane w zastosowaniach biurowych (wydajne stacje robocze) oraz biznesowych (serwery aplikacji i baz danych). Jednymi z najprostszych rozwiązań są komputery klasy PC z dwoma procesorami x86, obsługiwane przez typowe systemy operacyjne (Linux, Windows). Do technologii znacznie bardziej rozwiniętych można zaliczyć komputery firmy IBM klasy iSeries o architekturze TIMI[16]. Rozwiązania specjalizowane o architekturze wieloprocessorowej to superkomputery: IBM Blu Sky (Narodowe Centrum Badań Atmosfery USA, 7 bilionów (7×10^{12}) obliczeń na sekundę), Blue Mountain[17] (Los Alamos National Laboratory, Nowy Meksyk, trzy biliony (3×10^{12}) operacji zmiennoprzecinkowych na sekundę), 192-processor SGI Origin2000[18] (Boston University), Cray Inc. Cray Supercomputer 3-128 CPU.

Rozwiązania bazujące na sieciach komputerowych to klastry komputerowe, które charakteryzują się połączeniem wielu jednostek komputerowych w jeden wydajny system obliczeniowy. Ta technologia umożliwia osiągnięcie bardzo dużej mocy obliczeniowej przy niskich nakładach budżetowych. Zadanie obliczeniowe dzielone jest na wiele N procedur, które w sposób zsynchronizowany i równoległy wykonywane są przez N jednostek obliczeniowych. Przykładem może być projekt Clusterix [19] (ang. *National CLUSTER of LInuX Systems*) – rozproszony klaster PC.

Scalone systemy ogólnego przeznaczenia

Koprocessory o stałej architekturze (ogólnego przeznaczenia) odgrywają bardzo istotną rolę w przetwarzaniu danych większości systemów komputerowych, w tym również systemów osadzonych. Stanowią one układy z rodziny ASIC (ang. *Application Specific Integrated Circuit*) o dużym nakładzie produkcyjnym, charakteryzujące się niskimi kosztami produkcji oraz uniwersalnością zastosowań. Wadą uniwersalnych koprocessorów arytmetycznych jest przystosowanie ich architektury do obliczeń o bardzo dużej złożoności. Efektem tego jest znaczny nadmiar elementów scalonych w strukturze koprocessora w stosunku do bieżących wymagań obliczeniowych. Sztywna architektura uniwersalnych koprocessorów powoduje, że średnie wykorzystanie ich możliwości wynosi niewiele ponad połowę [11], np.: 55% – w procesorze Pentium, 53% – w Pentium II czy zaledwie 44% w procesorze RISC AX3000 firmy IDT. Względnie niska efektywność koprocessora jest spowodowana koniecznością zapewnienia uniwersalności struktury, w której ścieżki obliczeniowe są projektowane i optymalizowane przy spełnieniu kompromisu różnych warunków. W sprzeczności powszechnego użytku taki spadek mocy obliczeniowej nie jest wyraźnie dostrzegalny, jak w przypadku obliczeń w zadaniach o bardzo dużej złożoności, dla których może on mieć znaczące znaczenie, a nawet decydować o ich wykonalności.

Wśród propozycji poprawy efektywności systemów obliczeniowych przy użyciu dodatkowego sprzętu są interesujące dwa sposoby realizacji akceleratora:

- wykonanie specjalizowanego koprocatora realizującego ściśle określone zadanie lub jego najbardziej czasochłonne fragmenty. Rozwiązanie to jest kosztowne i wymaga wprawnego łączenia umiejętności inżynierskich i teoretycznych;
- zaprojektowanie rekonfigurowalnego koprocatora, który zmienia swoją architekturę sprzętową stosownie do realizowanego zadania; zaletą tego rozwiązania jest możliwość wykorzystywania tej samej wersji koprocatora w różnych aplikacjach, przy czym właściwe rozłożenie algorytmu (*software/hardware codesign*) na podzadania pomiędzy sprzęt (koprocator) i oprogramowanie (główny procesor systemu) pozwoli znacznie zwiększyć jego szybkość w porównaniu z koprocatorami uniwersalnymi o stałej architekturze.

Za wyborem drugiego rozwiązania przemawiają bardzo obiecujące wyniki badań i udane eksperymenty przeprowadzone w wielu ośrodkach akademickich oraz naukowych-badawczych przemysłu i wojska.

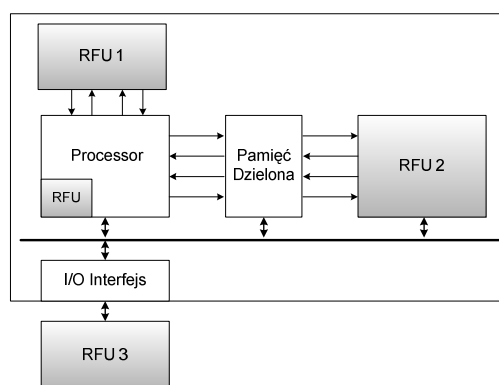
Moduły (karty rozszerzeń PC, zdalny CPU)

Kolejna grupa „modułowa” zawiera klasyfikację specjalizowanych systemów obliczeniowych, które głównie znajdują zastosowanie w pracach inżynierskich i badawczych. Można tu wymienić symulatory układów cyfrowych: H.O.T. firmy VCC (karta PC)[20], HES Alatek [15]; oraz symulatory hybrydowe przeznaczony do przyspieszenia symulacji systemów sprzętowo-programowych: Xcite-1000 Axis System [21]. Powstają również uniwersalne systemy obliczeniowe [11] dla komputerów klasy PC w celu przygotowania jednorodnej platformy obliczeniowej dla wielu algorytmów i zastosowań. Do rozwiązań specjalizowanych, które wykorzystywane są na co dzień w zastosowaniach domowych i profesjonalnych, można zaliczyć akceleratory graficzne, układy przetwarzania dźwięku, itp. Dzięki dynamicznemu rozwojowi IT jest to jedna z najbardziej rozwijających się branż na świecie.

Specjalizowane Systemy Osadzone

Dzisiejsze systemy osadzone składają się z wielu komponentów sprzętowych i programowych, które pozostają w ścisłej interakcji ze sobą. Określenie właściwego podziału (balansu) systemu pomiędzy obiema częściami determinuje sukces rozwiązania. Rozważając rozwiązanie programowe można łatwo zaobserwować, że komponenty programowe są prostsze do modyfikacji, niż ich odpowiedniki sprzętowe. Dzięki tej właściwości, programowe elementy systemu, przetwarzane przez programowalne procesory, zapewniają prostotę i szybkość podczas eliminacji błędów, łatwą zmianę zastosowania danego programu, ponowne użycie funkcji/procedur do budowy innych programów, lub w celu eliminacji czasu realizacji projektu. Jednak w porównaniu z komponentami sprzętowymi, elementy

programowe są dużo wolniejsze i zużywają znacznie więcej energii podczas wykonywania tych samych funkcji. Rozwiązania wyłącznie sprzętowe używane są w systemach, gdzie czas reakcji i pobór mocy są parametrami krytycznymi. Niestety, czas projektowania takich systemów jest długi i drogi. Dodatkowo, procedury zaimplementowane w strukturach sprzętowych (ASIC) nie mogą być modyfikowane. Zadaniem projektanta systemu jest znalezienie właściwego balansu pomiędzy komponentami programowymi i sprzętowymi w celu zapewnienia wymagań pracy systemu. Interakcja pomiędzy elementami sprzętowymi i programowymi jest ścisła, szczególnie w systemach osadzonych. Procesory (ASIP – ang. Application Specific Instruction-Set Processor)[2][3] o ograniczonej liście instrukcji oraz procesory ze zmienną/programowalną listą instrukcji (RISP – Reconfigurable Instruction-Set Processor)[5] stanowią przykład konstrukcji, gdzie wymiana informacji pomiędzy częściami sprzętowymi i funkcjami programowymi – interakcja – jest bardzo ścisła. Procesory ogólnego przeznaczenia nie mają tak krytycznych zależności, co wynika z ich uniwersalnej budowy. Projektowanie systemów specjalizowanych, a w szczególności procesorów specjalizowanych (ASIP, RISP), wymaga metodologii, która uwzględni wszelkie aspekty projektowania architektury części sprzętowej, jak i programu. Procesory ASIP można określić jako specjalizowane procesory o zdefiniowanej liście instrukcji. W procesie projektowym, na podstawie analizy programu, który będzie wykonywany przez ten procesor, ustala się listę instrukcji, jakie zostały użyte w kodzie programu. W procesorze ASIP implementowane są te bloki sprzętowe, które są niezbędne do wykonania wybranych instrukcji programu. W ten sposób budowany jest procesor zoptymalizowany pod kątem wydajności obliczeniowej o minimalnym poborze mocy. Wadą tego rozwiązania jest produkcja procesora dla jednego programu, jednorazowe zastosowanie.

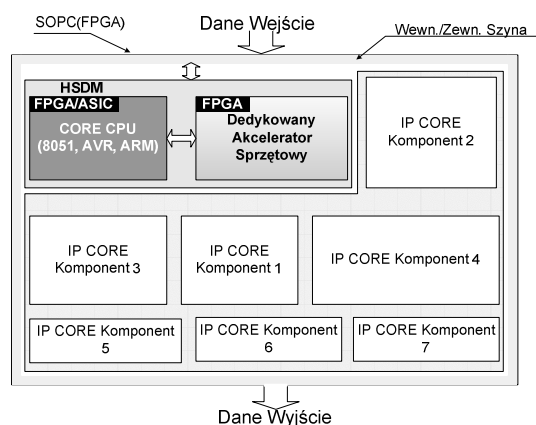


Rys. 2. Możliwe połączenia bloku RFU z procesorem głównym

Bardziej uniwersalnym rozwiązaniem, z jednoczesnym zachowaniem wydajności i zminimalizowanym poborze mocy, są procesory typu RISP. Charakteryzują się one tym, że zawierają rekonfigurowany blok funkcjonalny RFU (ang. *Reconfigurable Functional Unit*),

który spełnia rolę akceleratora podczas przetwarzania wybranej instrukcji. Blok RFU może zostać zintegrowany z główną jednostką przetwarzania programu (zazwyczaj procesor RISC) na wiele sposobów. Rysunek 2 przedstawia kilka najbardziej typowych połączeń. Coraz częściej również układy klasy DSP (ang. *Digital Signal Processing*)[4] zostają doposażone w bloki RFU w celu przyspieszenie przetwarzania danych i sterowania poprzez zrównoleglenie wykonywania operacji.

Do konstrukcji najbardziej interesujących ze względu na swą uniwersalność aplikacji, a w szczególności ze względu na proces projektowy, należą układy typu ASC (ang. *Application Specific Co-processors*). W procesie projektowym, nazywanym dekompozycją systemową (ang. *Hardware Software Codesign*) program, który ma zostać uruchomiony na docelowym procesorze, lub specyfikacja kontroli systemu cyfrowego dedykowana dla wybranego procesora, poddawane są analizom formalnym i symulacyjnym. W rezultacie, program lub specyfikacja dzielone są na dwie części: sprzętową i programową. Część sprzętowa realizuje zbiór zadań, które stanowią największe obciążenie operacyjne dla procesora. Natomiast procesor przetwarza rzadko wykonywane proste instrukcje sterowania i przetwarzania. Wynikiem procesu dekompozycji funkcjonalnej jest konfigurowalny koprocesor sprzętowy, który wspomaga procesor pracujący sekwencyjnie. W oparciu o założenia projektowania zintegrowanego opracowano wiele algorytmów dekompozycji funkcjonalnej, między innymi: Cosyma[12], Vulcan, Polis[13], Camad, Chinook, Tosca, SpecSyn. Domeną wymienionych rozwiązań są architektury typu SoB (ang. System on Board) [6] i SoC (ang. System on Chip). Trudno, lub jest to wykluczone, zastosować powyższe technologie do akceleracji systemów typu SOPC (ang. System on Programmable Chip), gdzie kompletny system cyfrowy (i/lub mikroprocesorowy) rezyduje w jednym układzie FPGA (ang. Field Programmable Gates Array)[22][23].



Rys. 3. Architektura HSDM

W celu rozwiązania tego problemu opracowana została architektura HSDM (ang. Hardware Software Digital Microsystem)[7][10] przedstawiona na rysunku nr 3 oraz algorytm ADSO[9][8], które charakteryzują się:

- wykorzystanie technik projektowania zintegrowanego;
- współbieżna praca części programowej i sprzętowej;
- sieci Petriego jako model pośredni specyfikacji systemu;
- weryfikacja formalna i symulacyjna systemu;
- analiza DFG i CFG;
- automatyzacja procesu dekompozycji funkcjonalnej.

Prace nad architekturą HSDM i algorytmem ADSO wciąż trwają. Aktualne przyspieszenie plasuje się na poziomie 300-400%, natomiast spodziewane rezultaty powinny sięgać przyspieszenia rzędu dziesiątek (10-20 razy).

3. PODSUMOWANIE

Zaproponowana klasyfikacja systemów informatycznych i architektur cyfrowych ma za zadanie uszeregowanie działu systemów wspierających proces przetwarzania danych i sterowania, w szczególności specjalizowanych systemów osadzonych. Cztery główne działy zostały sklasyfikowane na podstawie oceny subiektywnej autora, dostępnych materiałów oraz opinii i wiedzy inżynierów informatyków.

LITERATURA

- [1] H. Krupnova, G. Saucier: *FPGA-Based Emulation: Industrial and Custom Prototyping Solutions*, Lecture Notes in Computer Science, Springer-Verlag GmbH, ISSN: 0302-9743, June 2003
- [2] G. Dittmann, A. Herkersdorf: *Multi-Layer Intermediate Representation for ASIP Design and Critical Path Optimization*, Electrical Engineering, IBM Research Report RZ 3484, 02.2003
- [3] J. Lee, K. Choi, N. D. Dutt: *Automatic Instruction set design through efficient instruction encoding for ASIP*, ACES 2003, University of California, Irvine 2003
- [4] S. Ramanathan, S. K. Nandy, V. Visvanathan: *Reconfigurable filter coprocessor architecture for DSP applications*, Journal of VLSI Signal Processing 26, p.333-359, 2000 Kluwer Academic Publishers
- [5] F. Barat, R. Lauwereins, G. Deconinck: *Reconfigurable Instruction Set Processors from a Hardware/Software Perspective*, IEEE Transactions on software engineering, vol. 28, no. 9, 2002

- [6] A. A. Jerraya, J. Mermet: *System-Level Synthesis*, Kulwer Academic Publishers, 1999, Dordrecht/Boston/ London, Published in cooperation with NATO Scientific Affair Division
- [7] A. Stasiak: *Zintegrowany mikrosystem sprzętowo-programowy jako główna jednostka przetwarzania w systemach SOPC*, III Krajowa Konferencja Elektroniki. Kołobrzeg, Polska, 2004 .- Koszalin : Wydaw. Uczelniane Politechniki Koszalińskiej, 2004 .- T. 1/2, s. 279-284 .- ISBN: 83-7365-057-1
- [8] A. Stasiak, M. Michalczak, Z. Skowroński: *Algorytm dekompozycji systemowej dla potrzeb projektowania zintegrowanego*, RUC 2003 : materiały VI krajowej konferencji naukowej. Szczecin, Polska, 2003 .- Szczecin : Politechnika Szczecińska, 2003, s. 113-122 .- ISBN: 83-87362-56-5
- [9] Z. Skowroński, A. Stasiak: *Automatyczne dekompozycja systemów osadzonych*, Reprogramowalne Układy Cyfrowe - RUC 2003 : materiały VI krajowej konferencji naukowej. Szczecin, Polska, 2003
- [10] A. Stasiak, Z. Skowroński: *An universal coprocessor for SOC systems*, Synergies between information procesing and automation : 49. Internationales Wissenschaftliches Kolloquium. Ilmenau, Niemcy, 2004 .- Aachen : Shaker Verlag, 2004 .- Vol. 2, s. 208-213 .- ISBN: 3-8322-2824-1
- [11] K. Jasiński, P. Zbysiński: *Rekonfigurowany koprocesor systemowy: uniwersalna platforma obliczeniowa*, Przegląd Telekomunikacyjny, nr 10/2000
- [12] <http://www.ida.ing.tu-bs.de/research/projects/cosyma/overview/cosyma.html>
- [13] <http://www-cad.eecs.berkeley.edu/Respep/Research/hsc/abstract.html>
- [14] <http://www.cadence.com/quickturn/>
- [15] <http://www.alatek.com/>
- [16] www.ibm.pl
- [17] <http://www.lanl.gov/asci/bluemtn/>
- [18] <http://archive.ncsa.uiuc.edu/SCD/Hardware/Origin2000/>
- [19] <http://clusterix.pcz.pl/>
- [20] <http://www.vcc.com/prod1.html>
- [21] <http://www.verisity.com/products/xcite.html>
- [22] www.xilinx.com
- [23] www.altera.com