

Systemy przetwarzania sygnałów

Procesory DSP

wykład №12

Podstawowym wymogiem zadań CPS jest czas przetwarzania, czyli **możliwość dokonania przewidzianej symulacji w dopuszczalnym dla danego zastosowania czasie.**

Główną cechą systemów CPS jest **możliwość ich pracy w trybie czasu rzeczywistego**, czyli konieczność wykonania wszystkich operacji obliczeniowych i sterujących w czasie wystarczającym na niezakłócony przebieg trwającego procesu.

Ekspansja telekomunikacji analogowej i cyfrowej oraz techniki internetowej jest czynnikiem stymulującym rozwój technologii mikroelektronicznej.

Z tego względu na potrzeby systemów telekomunikacyjnych zostały stworzone procesory sygnałowe (**DSP**).

Nieustanna ewolucja tych procesorów jest odpowiedzią na stale rosnące zapotrzebowanie rynku nowych generacji telefonii cyfrowej na bardziej wydajne platformy sprzętowe, a także odpowiedzią na konkurencję ze strony alternatywnych rozwiązań systemów cyfrowych.

Warto zauważyć, że klasyczna, sztywna architektura procesorów, pomimo licznych modyfikacji i usprawnień, traci powoli na sile i znaczeniu na rzecz nowych koncepcji.

Aby efektywnie wykonać wiele zadań z dziedziny telekomunikacji, układy scalone DSP wyposaża się w mechanizmy wydajnej obsługi, oparte na wbudowanych w strukturę krzemową **dedykowanych blokach** o szerokim zakresie funkcjonalnym oraz elastyczności. Dojrzałość tej techniki umożliwia osiągnięcie wydajności znacznie przekraczającej tę, którą oferują wiodące klasyczne układy procesorowe i nawet najnowsze usprawnienia budowy architektur procesorów uniwersalnych nie są w stanie skutecznie temu zjawisku przeciwdziałać.

Specjalizowana architektura rdzeni procesorów sygnałowych stanowi wsparcie dla cyfrowego przetwarzania sygnałów w systemy mogące efektywnie realizować algorytmy czasu rzeczywistego.

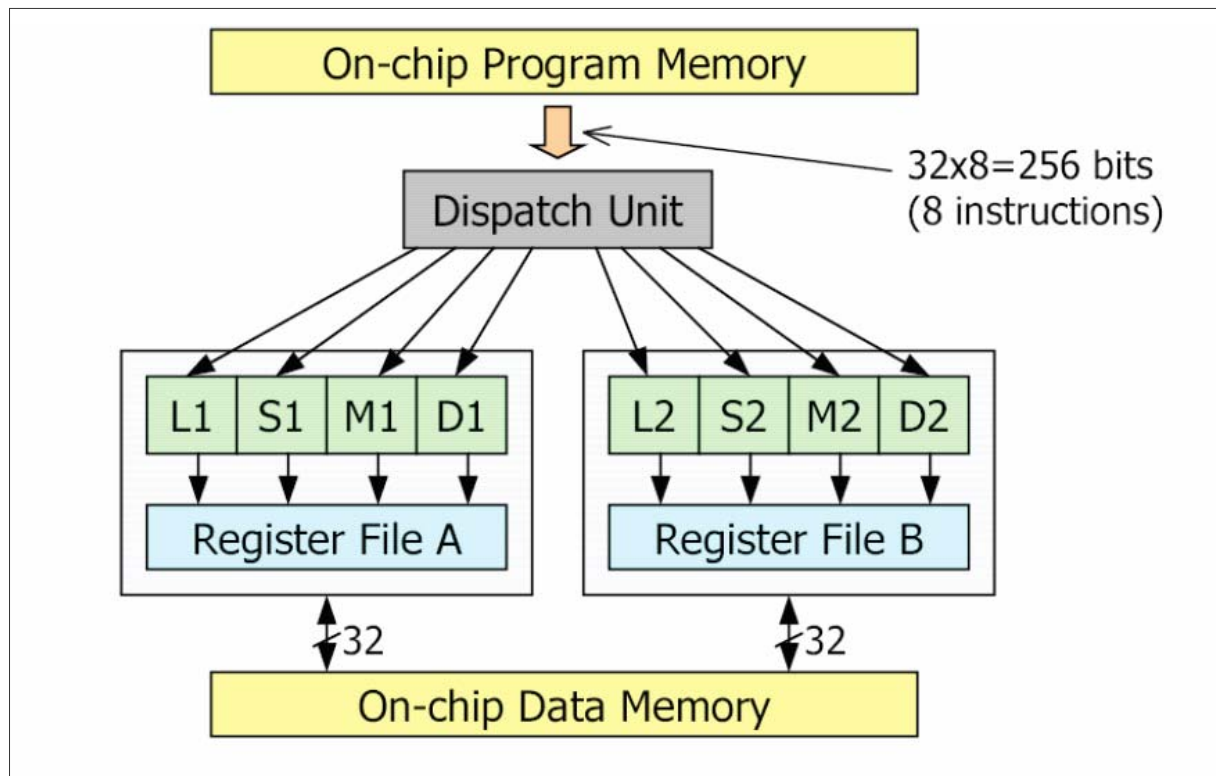
Poza tym nowoczesna technika umożliwia nie tylko zwiększanie częstotliwości pracy procesora, ale również rozszerzanie liczby wbudowanych w system jednostek obliczeniowych, potrafiących równolegle przetwarzać dane.

Przykładem takich koncepcji są architektury: **VelociTI** oraz zmodernizowana **VelociTI.2** (są one dostępne w procesorach rodziny **TMS320C6000** firmy Texas Instruments).

Ich cechą wspólną jest to, że mają **osiem** niezależnie pracujących **jednostek wykonawczych**, podzielonych na dwie grupy.

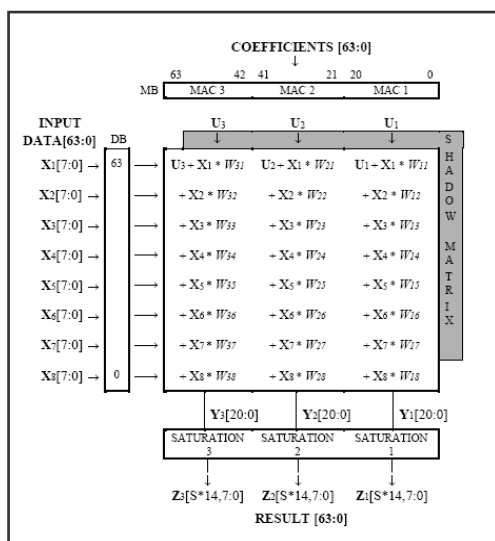
Dzięki specjalnej konstrukcji jest możliwe jednoczesne wykonanie do **ośmiu** 8-bitowych operacji **MAC** lub **czterech** (dla **VelociTI. 2**) albo **dwóch** (dla **VelociTI**) 16-bitowych mnożeń w jednym taktie zegara.

Architektura typu *VelociTI*

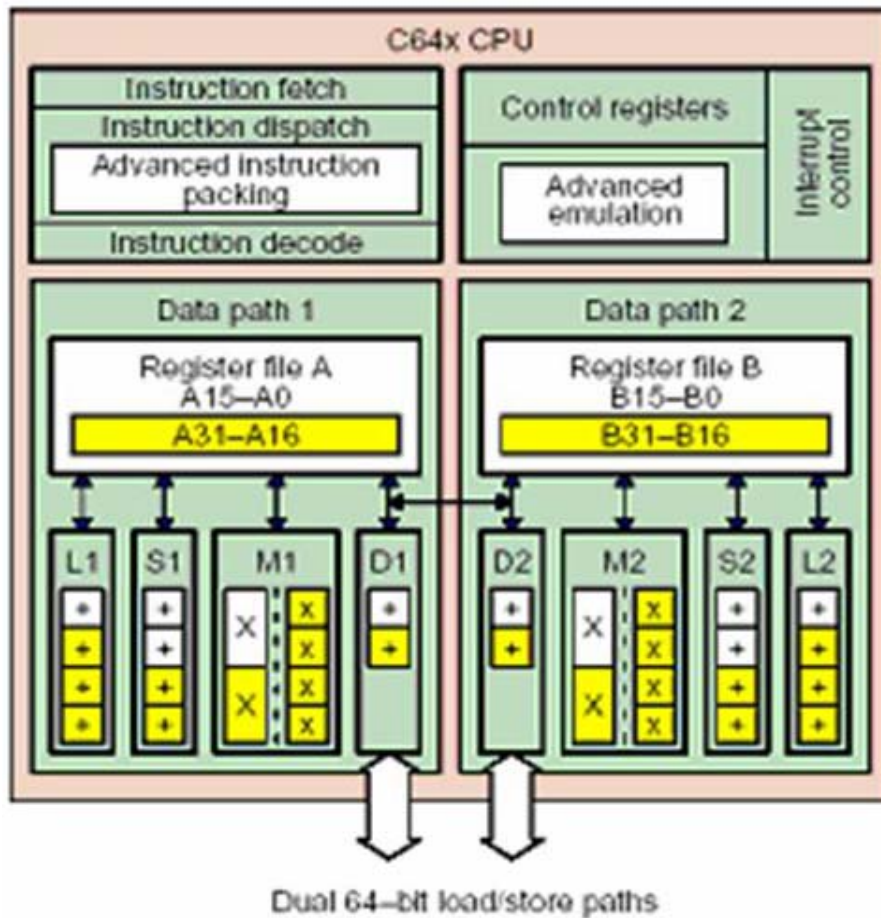


Wymienione rozwiązania są dostępne także w wielu innych platformach DSP, jak **Carmel (Infineon)**, **TigerSHARC** czy **Blackfin** firmy **Analog Devices**.

Neuroprocesor MN-6403



Architektura typu *VelociTI.2*



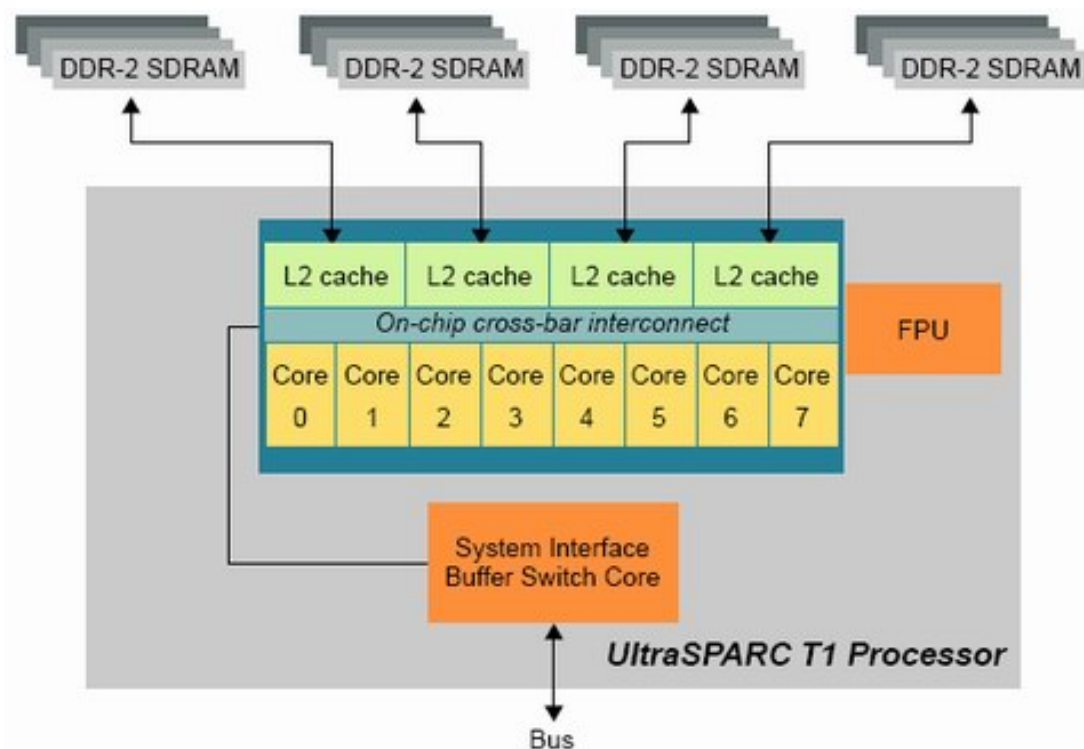
Jednorodne struktury wielordzeniowe

Jednoukładowy wielordzeniowy procesor UltraSPARC T1 (Niagara)

Sun Microsystems wypuszczyła na rynek procesor UltraSPARC T1 znanych pod nazwą Niagara.

UltraSPARC T1 zawiera 8 rdzeni taktowanych zegarem o częstotliwości 1.0 lub 1.2 GHz. Każdy rdzeń procesora UltraSPARC T1 może jednocześnie wykonywać do 4 niezależnych wątków (technologia CMT; Chip Multi-Threading). Układ zawiera też szynę JBUS, 4 układy współdzielonej pamięci podręcznej (3 MB) oraz kontroler pamięci RAM. Wymiana danych między procesorem a pamięcią RAM odbywa się za pośrednictwem 4 układów pamięci.

W zależności od życzenia klienta, producent aktywuje 4, 6 lub 8 rdzeni.

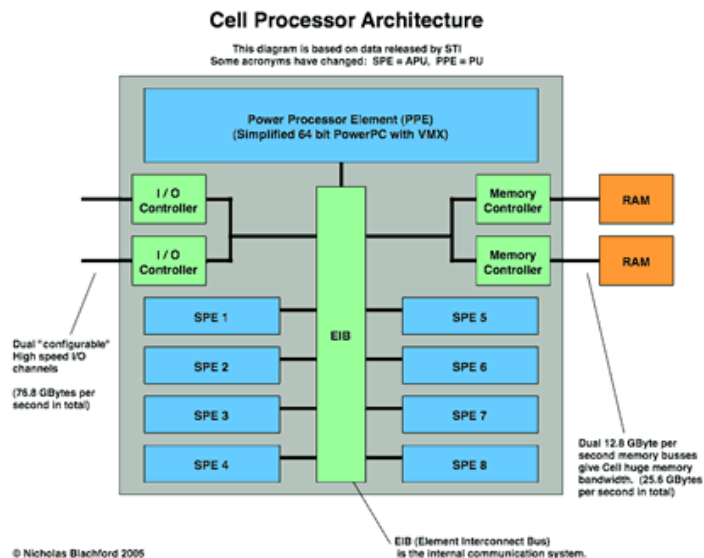


Wielordzeniowy procesor Cell Broadband Engine Architecture (Cell Bee)

IBM uruchomił produkcję nowej wersji procesora Cell. Jest to układ scalony opracowany wspólnie przez trzy firmy: IBM, Sony i Toshiba. Nowy procesor (noszący nazwę Cell Broadband Engine) jest produkowany w oparciu o technologię 65 nanometrów w East Fishkill (New York), gdzie IBM posiada fabrykę.

Układ Cell jest obecnie instalowany przez Sony Computer Entertainment w konsoli gier PlayStation 3, a IBM używa go w wielu swoich produktach. Firma wprowadziła na rynek swój pierwszy komputer oparty na układzie Cell we wrześniu 2006 r. Jest to serwer kasetowy BladeCenter QS20.

IBM wygrał ostatnio przetarg na dostawę superkomputera dla U.S. Department of Energy. Będzie to system obliczeniowy noszący nazwę Roadrunner, oparty na układach Cell, które będzie wykonywać do **milionu miliardów operacji na sekundę** (jedyńka i 15 zer; jeden petaflop).



Wielordzeniowy processor KiloCore

IBM i Rapport (mała firma kalifornijska) opracowały energooszczędny procesor, przeznaczony do instalowania w mobilnych urządzeniach, przetwarzających obrazy wideo o wysokiej jakości.

Firmy stworzyły swój nowy procesor Kilocore1025 bazujący na technologii Kilocore i Power Architecture.

Chip łączy w sobie **1024** 8-bitowe jednostki wykonawcze oraz jeden głównodowodzący układ PowerPC, stąd cyfra 1025 w nazwie oznaczająca ilość rdzeni.

Główne walory procesora, to wysoka wydajność i niewielki pobór mocy.

IBM podaje, że procesor **Kilocore 1025** pozwala transmitować obrazy wideo od 5 do 10 razy szybciej niż mogą to robić dostępne obecnie na rynku procesory.

Dalsza współpraca IBM i Rapport ma przynieść jak to określono "niszczące rozwiązania" na rynek komputerów o niskim poborze mocy, ich pierwszy wspólny produkt jest tego doskonałym przykładem.

Układy reprogramowalne

W przeciwieństwie do architektury procesorów sygnałowych macierz elementów logicznych struktur **FPGA** nie przewiduje ukierunkowania na stosowanie w konkretnych aplikacjach.

Dlatego dzięki uniwersalności tych układów jest możliwe odwzorowanie dowolnie złożonego systemu cyfrowego.

Jednak ogromna popularność tej klasy układów logicznych w dziedzinie telekomunikacji wymusiła i ukształtowała wiele rozwiązań konstrukcyjnych wzbogacających architekturę matryc reprogramowalnych FPGA oraz przyczyniła się do stworzenia nowych metodologii projektowych.

Wśród nich istnieją następujące trendy i koncepcje¹:

- nasilające się tendencje opracowywania bloków logicznych (będących elementarnym składnikiem struktury logicznej) o coraz większej uniwersalności, funkcjonalności i możliwościach konfigurowania,
- oferowanie zaawansowanych trybów pracy przez wbudowane moduły pamięci RAM (pamięć dwuportowa, kolejka **FIFO**),
- implementacja dedykowanych bloków obliczeniowych DSP, mających wydajnie zwiększyć prędkość wykonywania podstawowych operacji arytmetycznych oraz odciążać strukturę logiczną,
- rosnąca złożoność struktur **PLD** powodująca zacieranie się różnic w budowie między układami **CPLD** i **FPGA**,
- integrowanie wraz z matrycą reprogramowalną wydajnych procesorów RISC z architekturą wspierającą funkcje przetwarzania sygnałów DSP,
- rozwinięcie wielu mechanizmów statycznej i dynamicznej rekonfiguracji sprzętu,
- integracja w zasobach FPGA ultraszybkich interfejsów komunikacyjnych dla szerokopasmowej wymiany danych; takie rozwiązania są przeznaczone do implementacji sprzętowej gigabitowych protokołów teleinformatycznych,
- intensywny rozwój metod i narzędzi w dziedzinie nowoczesnej syntezy układów logicznych, umożliwiający efektywną implementację złożonych systemów cyfrowych programowalnych w strukturach bramkowych,

¹ Folejewski M., „Czy ewolucja układów reprogramowalnych przyczyni się do zagłady procesorów sygnałowych?” Przegląd Telekomunikacyjny, Rocznik LXXVII, №2-3/2004, s. 78-85.

- pojawienie się i rozwinięcie rynku własności intelektualnej **IP** w zakresie projektowania uniwersalnych komponentów wirtualnych (*IP-Core*), opisanych językami **HDL**.

Skomplikowane obliczenia arytmetyczne, algorytmy cyfrowego przetwarzania sygnałów oraz przetwarzania obrazów, działające w czasie rzeczywistym, wymagają dużej mocy obliczeniowej i dostępnej pamięci operacyjnej.

Wskutek tego ich implementacja w matrycach FPGA pochłania wiele zasobów sprzętowych i powoduje niezbyt wysoką wydajność wykonywania aplikacji oraz znaczne rozpraszanie mocy przez układ scalony.

Przeciwdziałając tym niekorzystnym zjawiskom oraz wykorzystując fakt, że aplikacje przetwarzające strumień danych multimedialnych są oparte na podstawowych operacjach arytmetycznych (którymi są dodawanie i mnożenie), producenci układów reprogramowalnych integrują w strukturach FPGA różnego rodzaju **sprzętowe moduły obliczeniowe DSP**, takie jak: ultraszybkie sumatory, mnożarki i rejestry przesuwające.

Dzięki nowoczesnej technologii submikronowej **możliwe jest scalenie kilkunastu (a nawet kilkuset) bloków obliczeniowych** z matrycą rekonfigurowalną w pojedynczej strukturze FPGA, co jest równoważne mocy obliczeniowej **systemu wieloprocesorowego** (np. macierzy procesorów sygnałowych) – zauważalne jest zmierzanie w kierunku architektury **SIMD** (*Single Instruction Multiple Data*).

Komponenty te wspomagają wykonywanie obliczeń i ułatwiają realizację funkcji z zakresu telekomunikacji i technik teleinformatycznych, takich jak: cyfrowa filtracja sygnałów, szybka transformata Fouriera **FFT**, dyskretna transformata kosinusowa **DCT** czy też kompresja obrazu ruchomego według standardu **MPEG**.

Zintegrowane bloki DSP, wspierające aplikacje cyfrowego przetwarzania sygnałów, można podzielić na dwie grupy.

Pierwszą stanowią **jednofunkcyjne, dedykowane moduły** realizujące proste i ściśle określone operacje arytmetyczne (jak **mnożenie** lub **sumowanie**) w ustalonej, stałoprzecinkowej reprezentacji danych.

Stosunkowo prosta struktura logiczna tych bloków, przekładająca się na niewielką zajmowaną powierzchnię struktury krzemowej, umożliwia umieszczanie wielu setek takich jednostek scalonych wraz z matrycą FPGA.

Najczęściej spotykanym rozwiązaniem są **mnożarki** wykonujące operacje w formacie 18×18 bitów (układy **Spartan-3**, **Virtex-II** i **Virtex-II Pro** firmy Xilinx).

W drugiej grupie znajdują się złożone, **uniwersalne bloki obliczeniowe DSP**, charakteryzujące się szerokim zakresem realizowanych operacji oraz dynamiczną konfiguracją architektury.

Ich złożoność oraz osiągnięta wydajność jest porównywalna z jednostkami przetwarzającymi procesorów sygnałowych.

Ze względu na wysoki stopień skomplikowania budowy bloków DSP, integracja wewnątrz struktur FPGA takich zasobów jest ograniczona i waha się od **kilku do kilkunastu modułów**. Przedstawicielami tej klasy rozwiązań są układy generacji **Eclipse** i **QuickMIPS** firmy **QuickLogic**, a także platformy **Stratix** produkowane przez firmę Altera.

Rodzina	Producent	Liczba bloków logicznych	Wbudowana pamięć RAM [kbit]	Bloki DSP	Liczba końcówek I/O	Dostępne kanały komunikacyjne	Wbudowany rdzeń procesora RISC
Stratix	Altera	10570–114140	898–9980	6–28	426–1314	–	–
Stratix GX	Altera	10570–41250	898–3343	6–14	328–542	4–20	–
EclipsePlus	QuickLogic	960–4032	45–81	10–18	250–506	–	–
QuickMIPS	QuickLogic	1152–2016	20–81	10–18	109–252	1–2	MIPS 4Kc
Virtex-II Pro	Xilinx	3168–125136	216–10008	12–5561)	204–1200	4–24	0–4, PowerPC
Spartan-3	Xilinx	1728–74880	12–2392	0–1041)	124–784	–	–
Excalibur	Altera	4160–38400	52–320	–	246–711	–	ARM922T
ORCA 4	Lattice	4992–16192	154–404	–	262–466	–	–
Mercury	Altera	4800–14400	48–112	–	303–486	8–18	–
Axcelerator	Actel	2016–32256	28–330	–	168–684	–	–
Delta39K	Cypress	512–3072	–	–	174–428	–	–

Uwagi: 1) Jako moduły dedykowanych sprzętowych multiplikatorów 18×18 bitów

Układy FPOA firmy MathStar (Field Programmable Object Array)

To są układy programowalne, zawierające do 400 obiektów, z których każdy jest 16-bitowym ALU. Obiekty zjednoczone ze sobą za pomocą zpatentowanych połączeń pracujących z częstotliwością 1 Ghz.

Układy FPOA firmy PicoChip

To też są układy obiektów programowalnych, zawierające do **kilkaset** 16-bitowych ALU i DSP.

Proć programowania w VHDLu są funkcje, które można programować za pomocą C oraz assemblera.

Układy MDSP rodziny CT3600 firmy Cradle Technologies

Firma Cradle Technologies proponuje jeszcze wyższy poziom granulacji.

W jednej strukturze zintegrowane są **8 32-bitowych** rdzeni ogólnego przeznaczenia oraz **16 DSP**.

Łączna wydajność układu 22, 000 16-bitowych MMACs'ów przy cyklu zegarowym 350 MHz.

Rdzenie konfigurowalne programowo

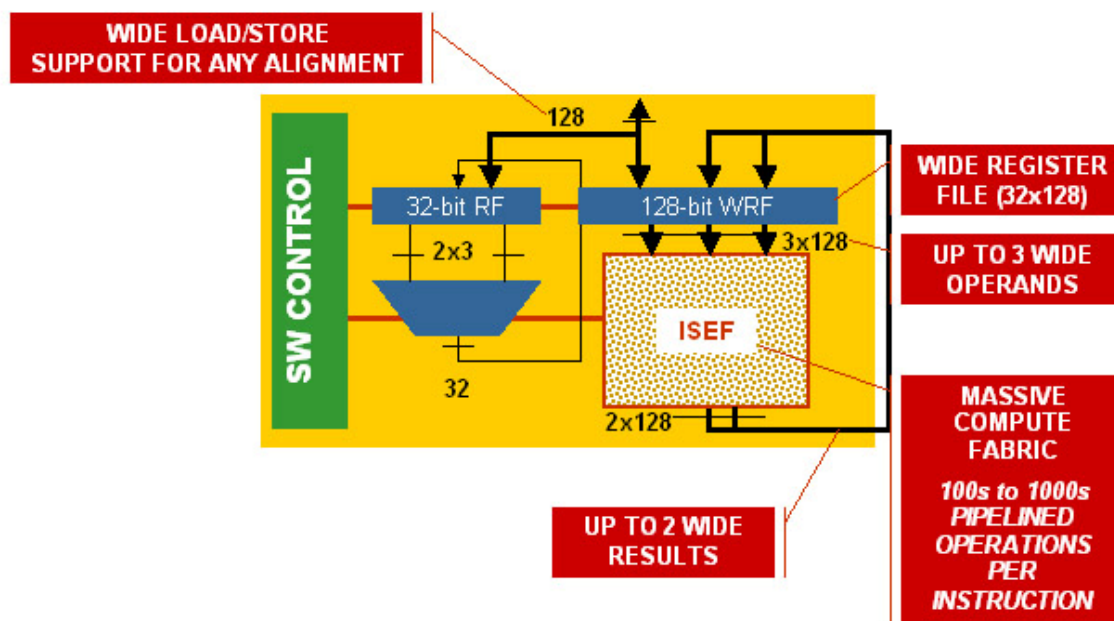
Firma Tensilica proponuje rekonfigurowalne programowo 32-bitowe rdzenie noszące nazwą Xtensa.

Korzystając z Xtensa Processor Generator projektanci SOC mogą tworzyć architekturę podsystemów procesorowych oraz kompletne środowiska programistyczne przystosowane do specyficznych potrzeb każdego projektu, w przeciągu kilku godzin.

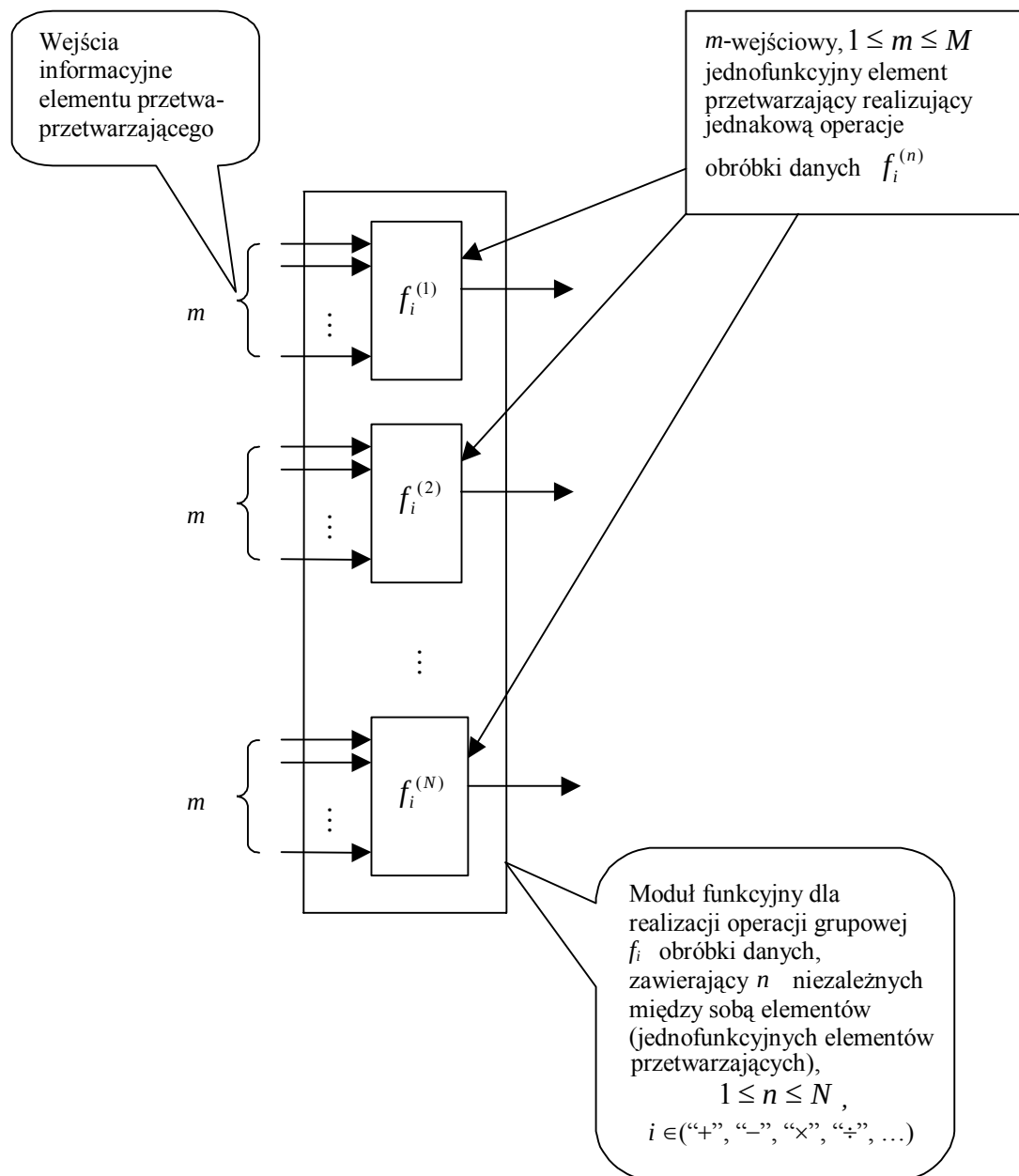
Niski koszt i wydajność to kluczowe aspekty dzisiejszego boomu na przenośne systemy elektroniczne.

Firma Stretch wykorzystuje rdzeń Tensilica Xtensa w połączeniu z własną rekonfigurowalną matrycą liczącą „Instruction Set Extension Fabric (ISEF)”.

Rekonfiguracja może być zrealizowana kilka tysięcy raz na sekundę, co zapewnia adaptację procesora do konkretnych zadań DSP.



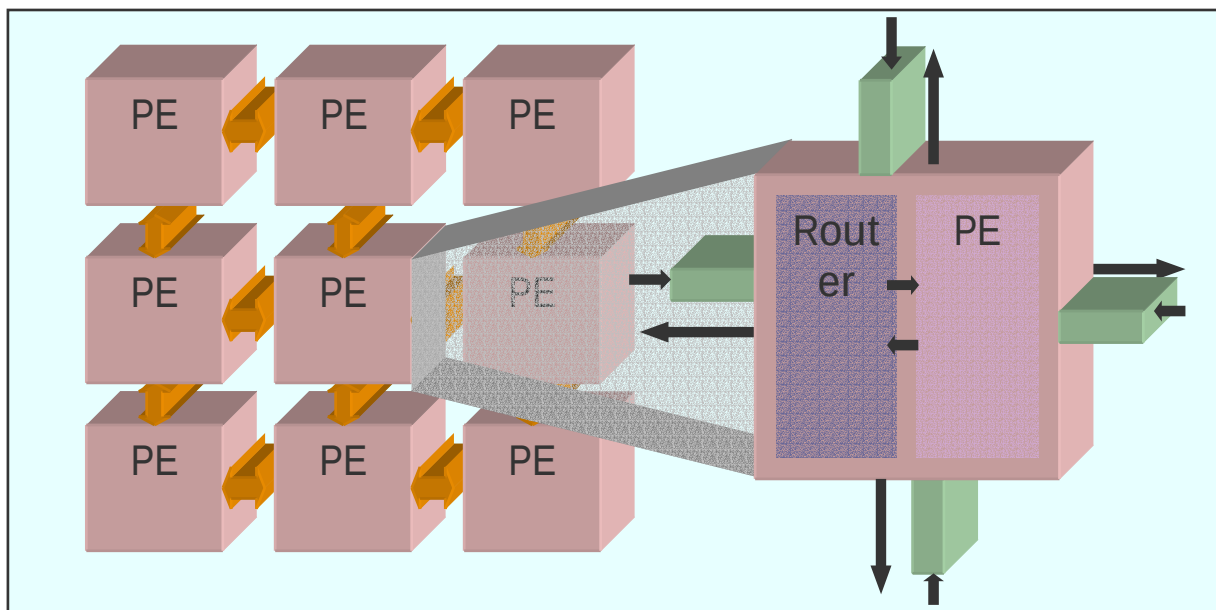
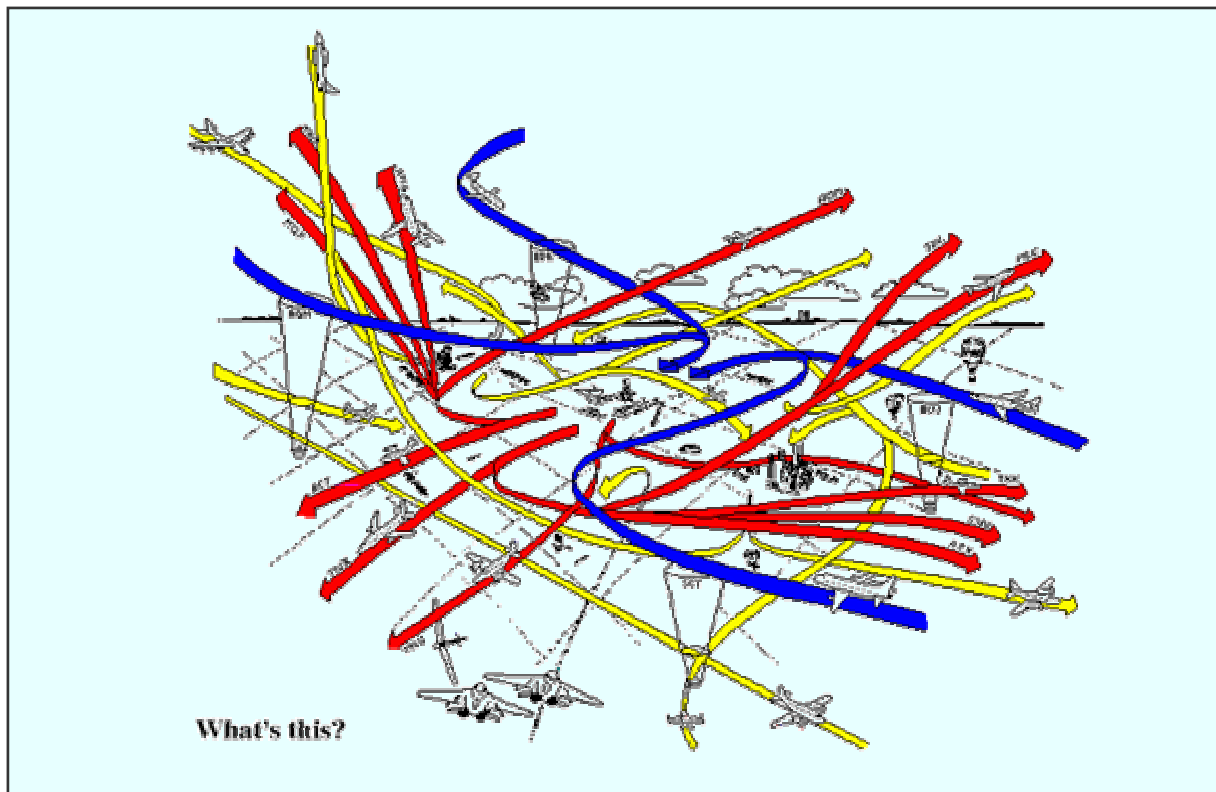
Stara teza	Spostrzerzenie aktualne
Straty energetyczne można nie uwzględniać, natomiast tranzystory są drogie	Energia kosztuje drogo, tranzystory zaś są tani: w jednym krysztale możemy zmieścić tyle tranzystorów, aż nie potrafimy je wykorzystać
Mnożenie jest wolnym, dostęp do pamięci zaś - szybkim	Mnożenie jest szybkim, zaś dostęp do pamięci stanowi problem dla wzrostu wydajności
Wzrost stopnia zrównoleglenia obliczeń zapewnia się poprzez udoskonalenie kompilatorów	Wzrost stopnia zrównoleglenia obliczeń zapewnia się poprzez udoskonalenie struktury sprzętowej



Dotychczasowe badania dowodzą, że stosując tradycyjną architekturę, opłaca się budować układy zawierające maks. 16 rdzeni.

Przy większej ilości rdzeni wydajność nie rośnie liniowo, ponieważ układy obsługujące komunikację między rdzeniami nie wyrabiają się i trzeba wtedy stosować w tym obszarze zupełnie nowe rozwiązania.

The Network-on-Chip (NoC) revolution



Intel zbudował i testuje 80-rdzeniowy układ scalony, który wykonuje ponad bilion operacji zmiennoprzecinkowych na sekundę (jeden teraflop).

Układ zajmuje powierzchnię 275 mm kwadratowych i pobiera mniej mocy niż standardowe procesory instalowane obecnie w pecetach.

Procesor jest taktowany zegarem 3,16 GHz, wykonuje dokładnie 1,01 bilionów operacji zmiennoprzecinkowych na sekundę i pobiera moc o wielkości 62 watów. Oznacza to, że na jeden wat poboru mocy przypada wydajność 16 gigaflopów.

Układ może pracować jeszcze szybciej, ale współczynnik wydajność/pobór mocy jest wtedy mniejszy (układy taktowane zegarami 5,1 GHz i 5,7 GHz pracują odpowiednio z szybkościami 1,63 i 1.81 teraflopów).

Układ może się pochwalić tak niskim poborem mocy dzięki temu, że niepracujące rdzenie znajdują się w stanie uśpienia.

Rdzenie takie wchodzi do akcji tylko wtedy, gdy zapotrzebowanie na moc obliczeniową wzrasta.

Każda warstwa zawierająca rdzeń obliczeniowy ma do dyspozycji swój własny router, dlatego układ reprezentuje swoistego rodzaju "sieć na skalaku".

Wydajność rzędu teraflopów uzyskano po raz pierwszy w 1996 roku w superkomputerze **ASCI Red** zbudowanym przez Intel dla Sandia National Laboratory.

Komputer ten zajmował ponad *180 metrów kwadratowych*, składał się z *niemal 10 000 procesorów Pentium® Pro* i zużywał ponad *500 kilowatów* energii.

Badawczy procesor Intela osiąga tę samą wydajność w pojedynczym wielordzeniowym układzie.

Pierwsze 80-rdzeniowe procesory będą gotowe do produkcji pomiędzy 2010 a 2015 rokiem.

Intel planuje wyposażyć je do tego czasu w rdzenie dedykowane do specjalistycznych zadań, jak np. kryptografia czy wirtualizacja.

Muszą też w międzyczasie pojawić się aplikacje, które będą w stanie skorzystać z mocy 80-rdzeniowego układu.