

Wielowątkowość a wielordzeniowość

Wielowątkowość i wielordzeniowość to niezależne, ale dopełniające się, techniki projektowania aplikacji. Korzyści z ich stosowania zależą od zadań powierzonych systemowi, ale może się zdarzyć, że żadna z nich nie będzie odpowiednia w danej sytuacji. Ponieważ pracujące wielowątkowo oraz wielordzeniowe procesory stały się normą i są dostępne także w komputerach SBC, projektanci układów cyfrowych powinni znać ich poszczególne ich cechy, wady i zalety.

Przetwarzanie współbieżne

Zwiększenie wydajności systemu przez podniesienie jego częstotliwości taktowania powoduje większe zużycie mocy i jest ograniczone technologiami elektronicznymi. Granica ta zwiększa się co roku, jednak wydajność tanich i energooszczędnych rozwiązań wciąż pozostawia wiele do życzenia. Na szczęście, prawie wszystkie aplikacje komputerowe reprezentują pewien stopień współbieżności, ponieważ często istnieje potrzeba wykonywania wielu niezależnych zadań jednocześnie. Pozwala to na zastosowanie technik przetwarzania wielowątkowego, a w tym np. rozłożonego na wiele rdzeni procesora. Zadanie to nie zawsze jest proste, ale i tak łatwiejsze niż przekraczanie ograniczeń fizycznych, na które natrafiają twórcy procesorów.

Cechą systemów wieloprocessorowych lub jednoprocessorowych ale wielordzeniowych jest wykorzystanie przetwarzania współbieżnego do podziału zadań. W optymalnej sytuacji liczba wątków oprogramowania wykonywanych w tym samym czasie, powinna odpowiadać liczbie dostępnych w systemie procesorów.

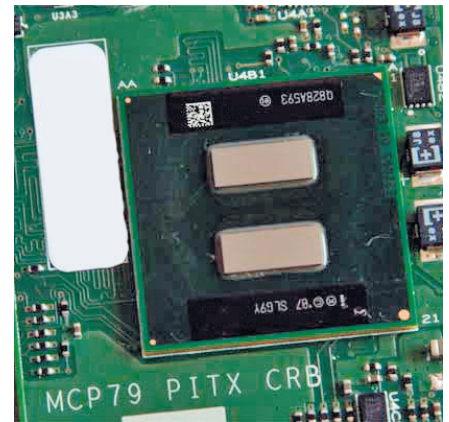
W specjalizowanych systemach odpowiednie wątki mogą być przydzielane do różnych procesorów (a niekiedy i ich części), z których każdy ma inne przeznaczenie i może zostać optymalnie skonfigurowany do konkretnych operacji. Pozwalają na to np. najnowsze wersje QNX-a.

Wielowątkowy sprzęt

Istnieją także procesory wielowątkowe, które mogą wykonywać przetwarzanie współbieżne, ale w nieco inny sposób niż układy wielordzeniowe. Zamiast wykorzystania technik podziału obciążenia na poziomie systemu, ich działanie polega na opty-

malizacji wykorzystania powierzchni procesora oraz zwiększenia efektywności energetycznej. W przeciwieństwie do układów wielowątkowych, rdzenie jednowątkowe, zanim wykonają zaplanowane operacje, muszą czekać na przesłanie danych z pamięci.

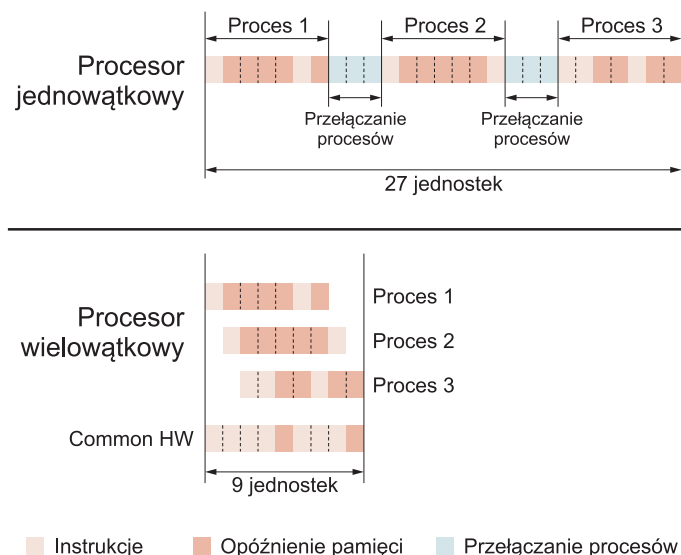
Wielowątkowość może być więc zdefiniowana następująco: jeśli ze względu na opóźnienia w komunikacji między procesorem a pamięcią, pojedyncze zadanie nie będzie w stanie całkowicie wykorzystać zasobów procesora, jednoczesne przetworzenie wielu zadań trwałoby krócej niż wykonanie ich sekwencyjne. Innymi słowy, wielowątkowość oznacza możliwość uruchomienia więcej niż jednego ciągu instrukcji (wątku) w tym samym czasie. W tym celu procesor powinien mieć więcej niż jeden licznik i zbiór programowalnych rejestrów. Powielenie tych zasobów jest dużo mniej kosztowne niż wymiana całej jednostki obliczeniowej. Przykładowo, dla procesora MIPS32 34K, wykorzystującego architekturę wielowątko-



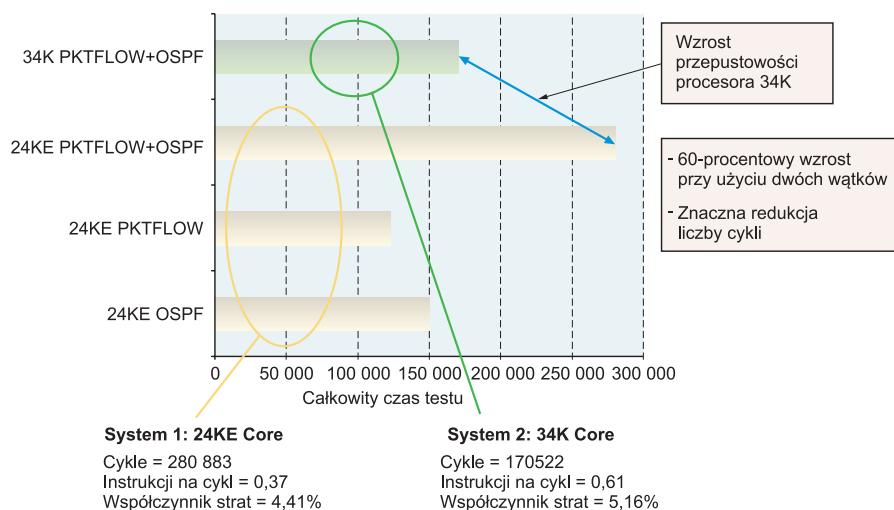
wą, wzrost powierzchni układu scalonego o 14% odpowiada 60-procentowemu wzrostowi przepustowości w stosunku do rdzenia jednowątkowego.

Skalowalność wielordzeniowości

Teoretycznie, wieloprocessorowe architektury są nieskończenie skalowalne, jednak w praktyce tylko ograniczona liczba aplikacji mogłaby wykorzystać moc obliczeniową tysięcy procesorów. Z kolei uruchomienie wielu wątków na pojedynczym rdzeniu przyniesie wzrost wydajności tylko do momentu maksymalnego obciążenia jego jednostek wykonawczych. Chociaż sposoby i cele stosowania wielordzeniowości i wielowątkowości są różne, wspólnym wymogiem ich skutecznego wykorzystania jest użycie odpowiedniego oprogramowania oraz wielozada-



Rys. 1. Wydajność dwuwątkowego procesora MIPS 34K



Rys. 2. Wyniki testów

niowego systemu operacyjnego. Natomiast jednowątkowe, nieobsługujące tych technologii aplikacje powinny zostać napisane od nowa i zdekomponowane do postaci podprogramów lub oddzielnych wątków oprogramowania.

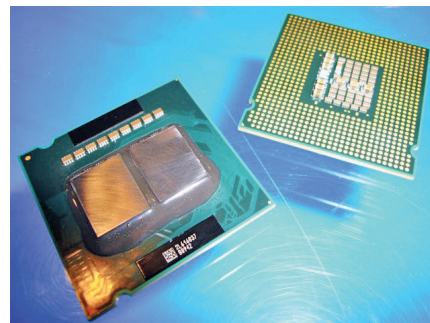
Kiedy wielordzeniowość ma sens

Zastosowanie wielordzeniowości w projektach jest korzystne, jeśli funkcje systemu mogą być zdekomponowane do postaci niezależnych zadań. Wykorzystanie wielu wolniejszych rdzeni zamiast jednego bardzo wydajnego procesora przekłada się przeważnie na oszczędność mocy i ewentualne także powierzchni, jeśli mniejsza przepustowość pozwala na użycie mniejszych komórek pamięci.

Jeśli funkcje układu, które mają być wykonywane jednocześnie, nie mogą zostać zdekomponowane na etapie projektowania elektroniki systemu, można zastosować

podjęcie alternatywne, polegające na budowie symetrycznego klastra obliczeniowego, składającego się z wielu procesorów. Poszczególne procesory klastra wykonują przydzielone im w locie zadania, dzięki czemu realizowane jest przetwarzanie współbieżne. Kosztem takiej elastyczności jest konieczność budowy złożonej sieci połączeń między poszczególnymi procesorami a współużytkowaną szybką i pojemną pamięcią. To zwiększa powierzchnię całego układu i zużycie mocy, jednak wynikiem jest funkcjonalny odpowiednik systemu wielordzeniowego.

Każdy dodatkowy rdzeń, niezależnie czy w postaci oddzielnego procesora, czy też jako część jednego układu scalonego, powoduje wzrost sumarycznej powierzchni zastosowanych układów scalonych i nawet w trybie oszczędzania energii zużywa pewną ilość prądu upływowego. W związku z tym cały



system powinien zawierać minimalną liczbę rdzeni, niezbędną dla danej aplikacji. Budowanie systemu wielordzeniowego nie ma sensu, jeśli określone zadanie może być wydajnie zrealizowane przez pojedynczy rdzeń.

Kiedy wielowątkowość ma sens

Zastosowanie wielowątkowości przynosi korzyści w sytuacji, w której procesor musiałby długo czekać na instrukcje. Czas oczekiwania rośnie wraz ze wzrostem częstotliwości rdzenia i zależy także od technologii oraz sposobu zarządzania pamięcią przez daną aplikację. Typowe złożone oprogramowanie może być nieefektywne w jednowątkowym środowisku przez 30% czasu, natomiast programy gorzej zoptymalizowane pod kątem pamięci – nawet przez 75% czasu pracy. Systemy, w których szybkości procesora i pamięci są dobrze dopasowane, tzn. nie występuje spadek wydajności spowodowany opóźnieniami w komunikacji między nimi, w ogóle nie zyskują na wielowątkowości. Dodatkowe zasoby procesora wynikające z zastosowania wielowątkowości mogą być wykorzystane także do innych zadań, których nie można zrealizować w systemie wielordzeniowym, opartym na procesorach konwencjonalnych. Jeśli, na przykład, w zwykłym procesorze wystąpi żądanie przerwania, następuje wstrzymanie i zapamiętanie stanu aktualnego programu i wykonanie kodu procedury przerwania. Natomiast wielowątkowy procesor może sprzętowo dokonywać przełączania wątków, co zapewnia wykonywanie przerwania bez opóźnienia obecnego w przypadku ich programowej obsługi.

Podsumowanie

Z punktu widzenia efektywnego wykorzystania powierzchni i mocy, optymalny procesor SoC powinien zawierać wielowątkowy rdzeń jako podstawowy element przetwarzający i replikować go w konfiguracji wielordzeniowej, jeśli aplikacja wymagałaby większej wydajności niż może zapewnić pojedynczy rdzeń.

Grzegorz Michałowski
artykuł ukazał się
w magazynie **Elektronika**

