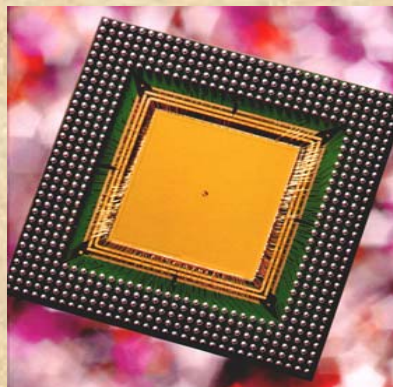


Tarnów i nowoczesne technologie

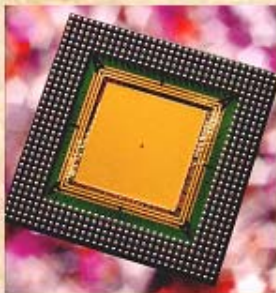


Tarnów - 27 marca 2007 r.

Akceleracja obliczeń w ultraszybkich zastosowaniach

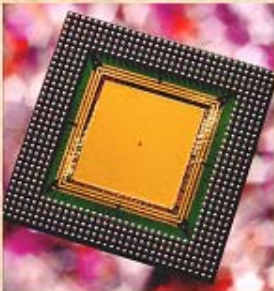
Prof. Kazimierz Wiatr

- Senator RP
- Przewodniczący Komisji Nauki, Edukacji i Sportu
- Dyrektor ACK Cyfronet AGH
- Przewodniczący Rady Konsorcjum Pionier



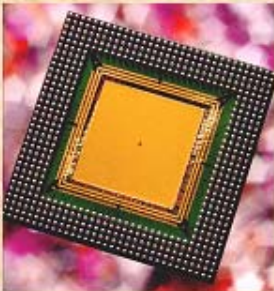
Dzisiejsze wymagania systemów obliczeniowych

- **szybkość !!!!**
- **cena**
- **wielkość**
- **niezawodność**



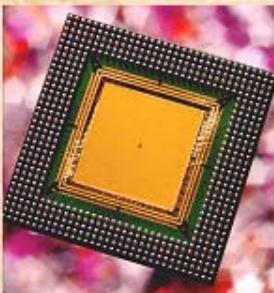
Szybkość komputerów

- wczoraj
- dziś
- jutro



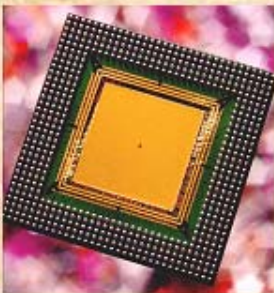
Metody zwiększania szybkości pracy komputerów

- ◆ szybszy zegar
- ◆ optymalizacja kodu
- ◆ wiele procesorów
- ◆ architektura wewnętrzna
- ◆ sprawne otoczenie procesora



Rodzaje komputerów

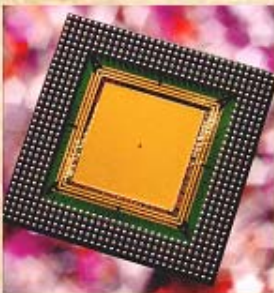
- komputery osobiste
- superkomputery
- architektury specjalizowane



Akceleracja obliczeń

- zbyt długie czasy obliczeń w procesorach GPP i DSP oraz komputerów HPC
- poszukiwanie przyspieszenia-akceleracji obliczeń
- ograniczenie rodzaju obliczeń i ich aplikacji
- poszukiwanie nowych jakościowo metod szybkiej realizacji algorytmów

$$S = t_{\text{sys_klasycznego}} / t_{\text{sys_specjalizowanego}}$$



Akceleracja obliczeń w systemach wieloprocesorowych

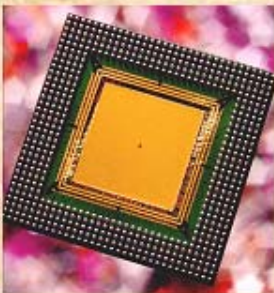
Dla systemów wieloprocesorowych (Prawo Amdahl'a)

$$S = (s + p) / (s + p/n)$$

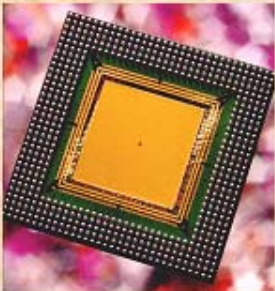
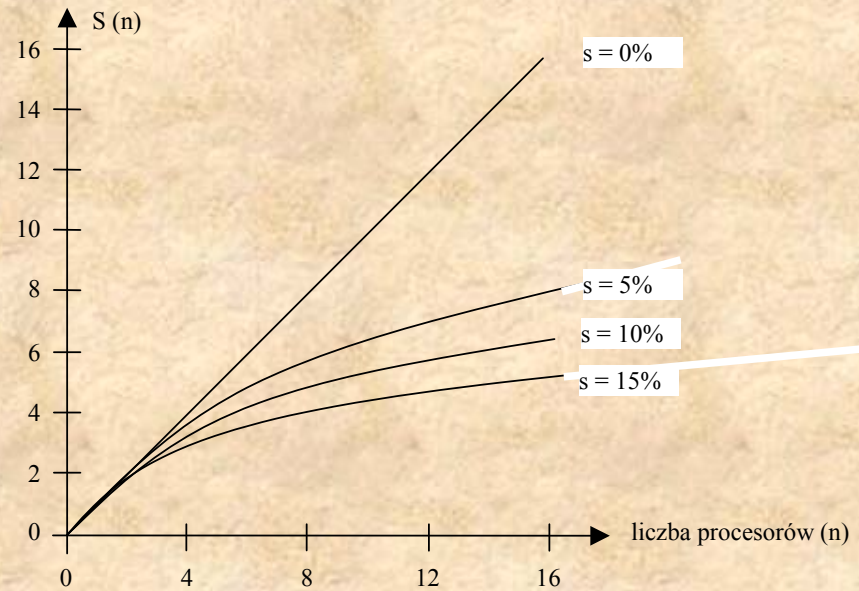
gdzie:

s - czas algorytmu w części szeregowej programu

p - czas algorytmu w części równoległej



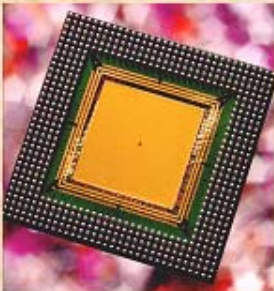
Akceleracja obliczeń w systemie wieloprocessorowym



Potrzeby obliczeniowe

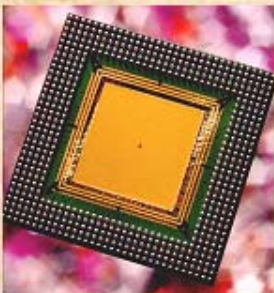
- ◆ systemy on-line
- ◆ systemy off-line

- ◆ np. obliczenia symulacyjne DNA (ok.100 procesorów Itanium2):
 - ☞ 20 atomów - 3 godziny
 - ☞ 200 atomów - 34 lata
- ◆ np. medycyna z interakcją itp.



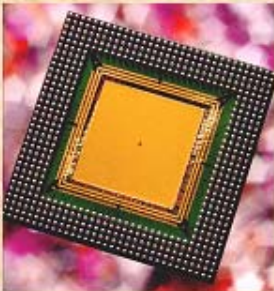
Sprzętowe metody akceleracji obliczeń

- **Systemy obliczeniowe**
- **Wpływ architektury systemu na szybkość obliczeń**
- **Specjalizowane procesory sprzętowe**
- **Implementacja w układach FPGA**
- **Parametryzacja implementacji**
- **Rekonfigurowalne systemy obliczeniowe**
- **Dedykowane użytkownikowi platformy CCM**
- **Hardware/Software CoDesign**
- **Elementy IP**
- **Szybkość pracy systemów obliczeniowych**



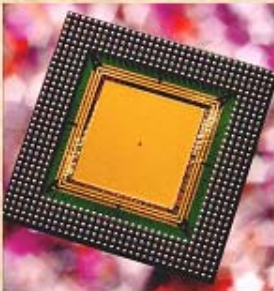
Kierunki badawcze

- poszukiwanie przyspieszenia obliczeń
- ograniczenie rodzaju obliczeń i ich aplikacji
- poszukiwanie nowych jakościowo metod szybkiej realizacji algorytmów
- poszukiwania w 2 kierunkach:
 - specjalizowane elementy obliczeniowe - procesory
 - dedykowane połączenia tych elementów - architektura
- prace badawcze oprzeć na realnych możliwościach implementacji w dostępnych technologiach



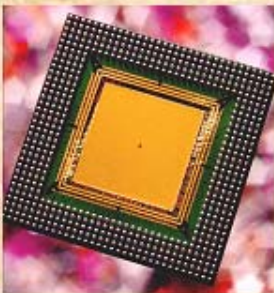
Sprzętowa akceleracja obliczeń

- wielkie ilości przetwarzanych danych
- bardzo szybkie systemy obliczeniowe
 - algorytmy zdominowane przez dane
 - algorytmy zdominowane przez instrukcje



Liczba przesłań operandów w operacjach przetwarzania obrazów

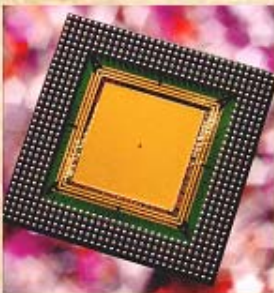
OPERACJA	1 PIKSEL	1 OBRAZ	1 SEKUNDA
■ LUT	1	262 144	6 553 600
■ MEDIANA-5	5	1 510 720	37 768 000
■ KONWOLUCJA 3x3	9	4 718 592	117 964 800



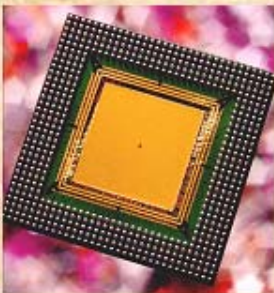
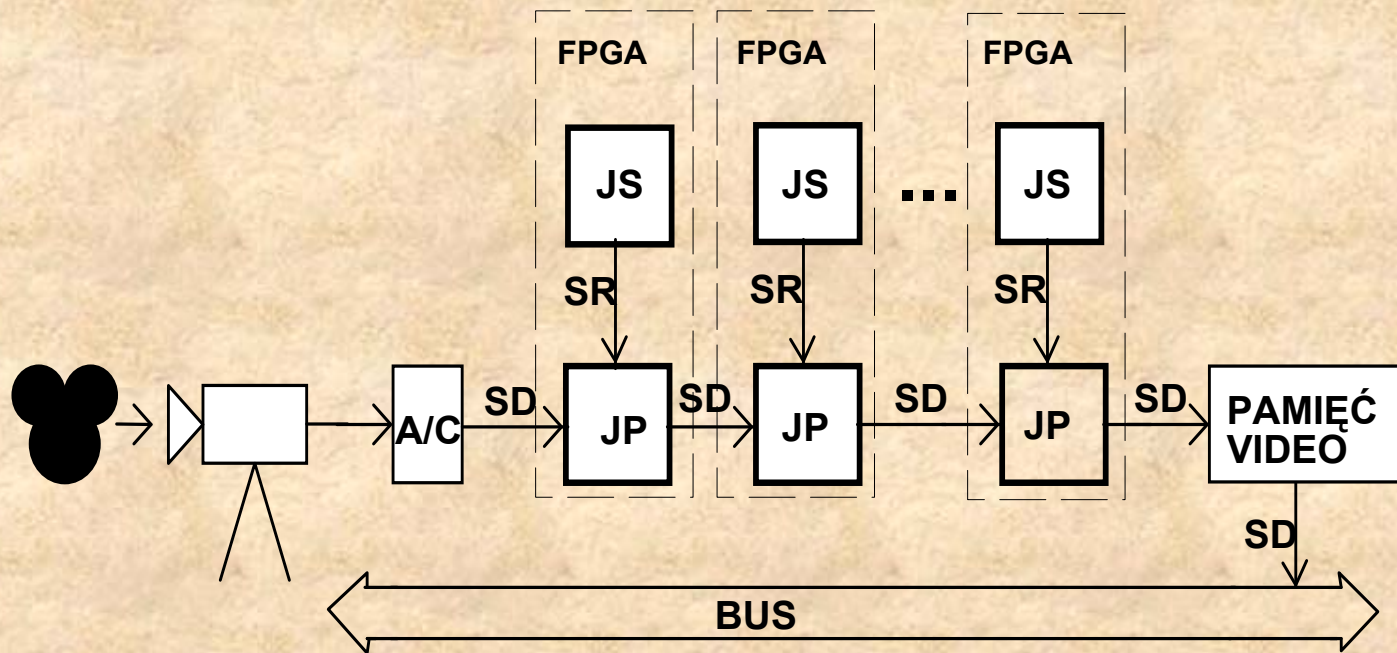
Liczba operacji w przetwarzaniu obrazów

(w czasie 1 s)

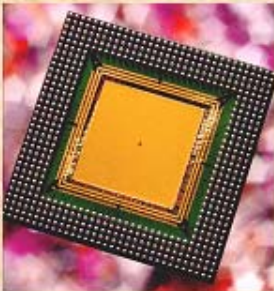
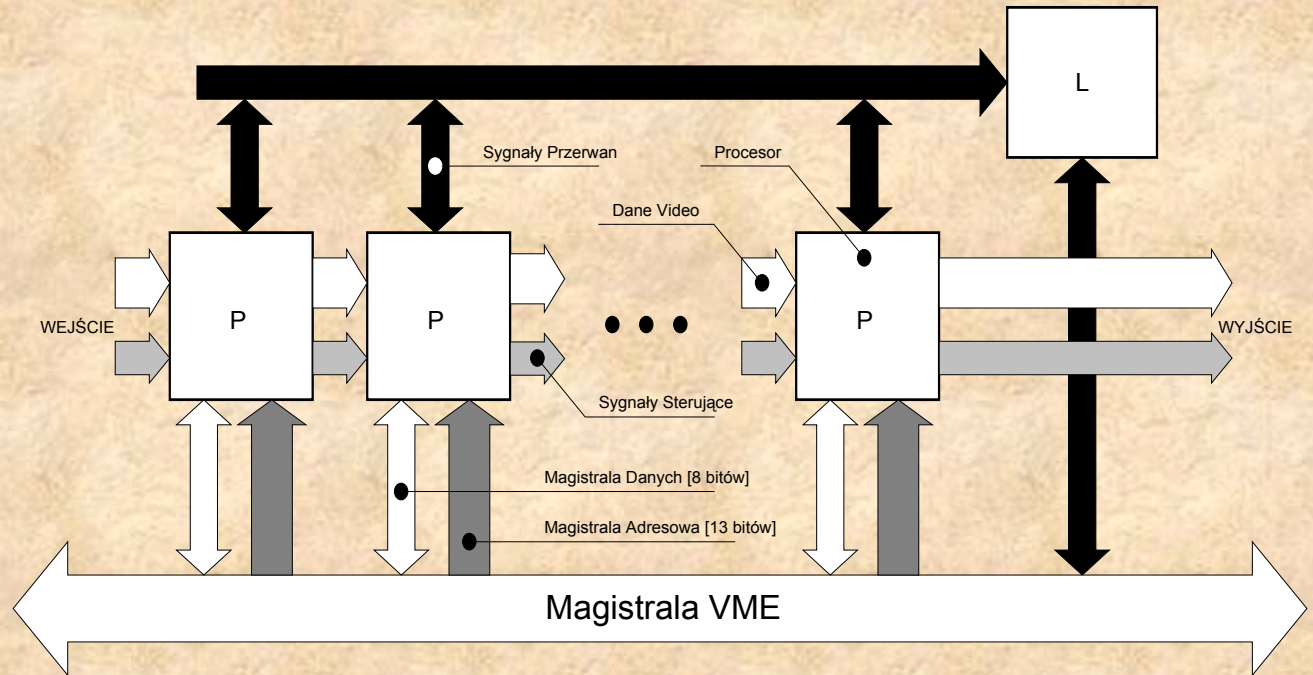
OPERACJA	MNOŻENIA	DODAWANIA PORÓWNANIA	PRZESŁANIA
■ LUT	0	6 553 600	13 107 200
■ MEDIANA-5	0	65 536 000	44 321 600
■ MEDIANA-9	0	294 912 000	124 518 400
■ KONWOLUCJA 3x3	117 964 800	4 718 592	124 518 400



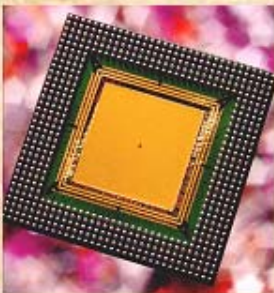
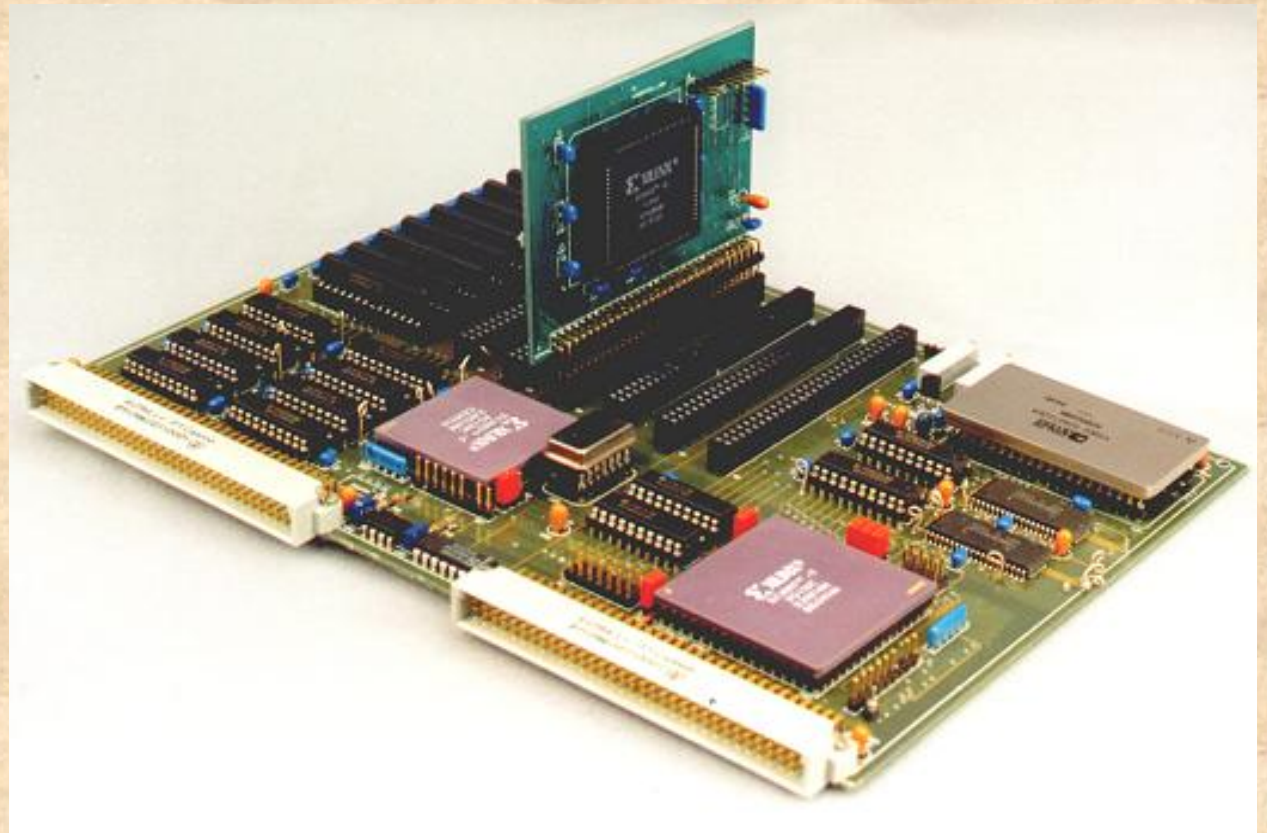
Architektura MISD specjalizowanych procesorów sprzętowych



Schemat blokowy architektury potokowej DePiAr



Moduł z architekturą potokową DePiAr



Kazimierz Wiatr, ACK CYFRONET AGH, Zespół Akceleracji Obliczeń

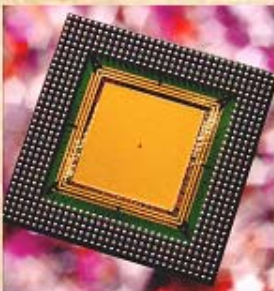
Architektura potokowa DePiAr

■ Zalety:

- ◆ minimalizacja czasu transmisji
- ◆ minimalizacja czasu obliczeń przez dedykowane procesory sprzętowe
- ◆ procesory są autonomiczne:
 - ☞ elastyczna zmiana kolejności operacji
 - ☞ niezależne projektowanie procesorów
 - ☞ implementacja w dowolnym otoczeniu sprzętowym
- ◆ dostęp do procesorów z zewnętrznej magistrali:
 - ☞ dynamiczna zmiana parametrów przetwarzania
 - ☞ dynamiczna rekonfiguracja kolejności operacji

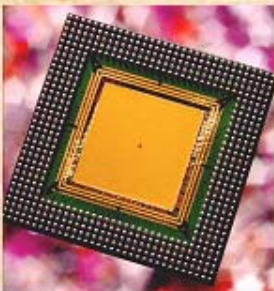
■ Wymagania:

- ◆ kolejnoliniowy sygnał wizyjny
- ◆ szybka transmisja na wyjściu potoku
- ◆ konieczność zachowania wymogów czasowych:
 - ☞ jednakowe opóźnienia dla danych i sygnałów sterujących
 - ☞ opóźnienie wielokrotnością czasu trwania piksela

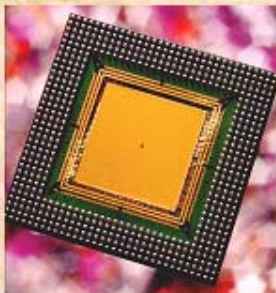
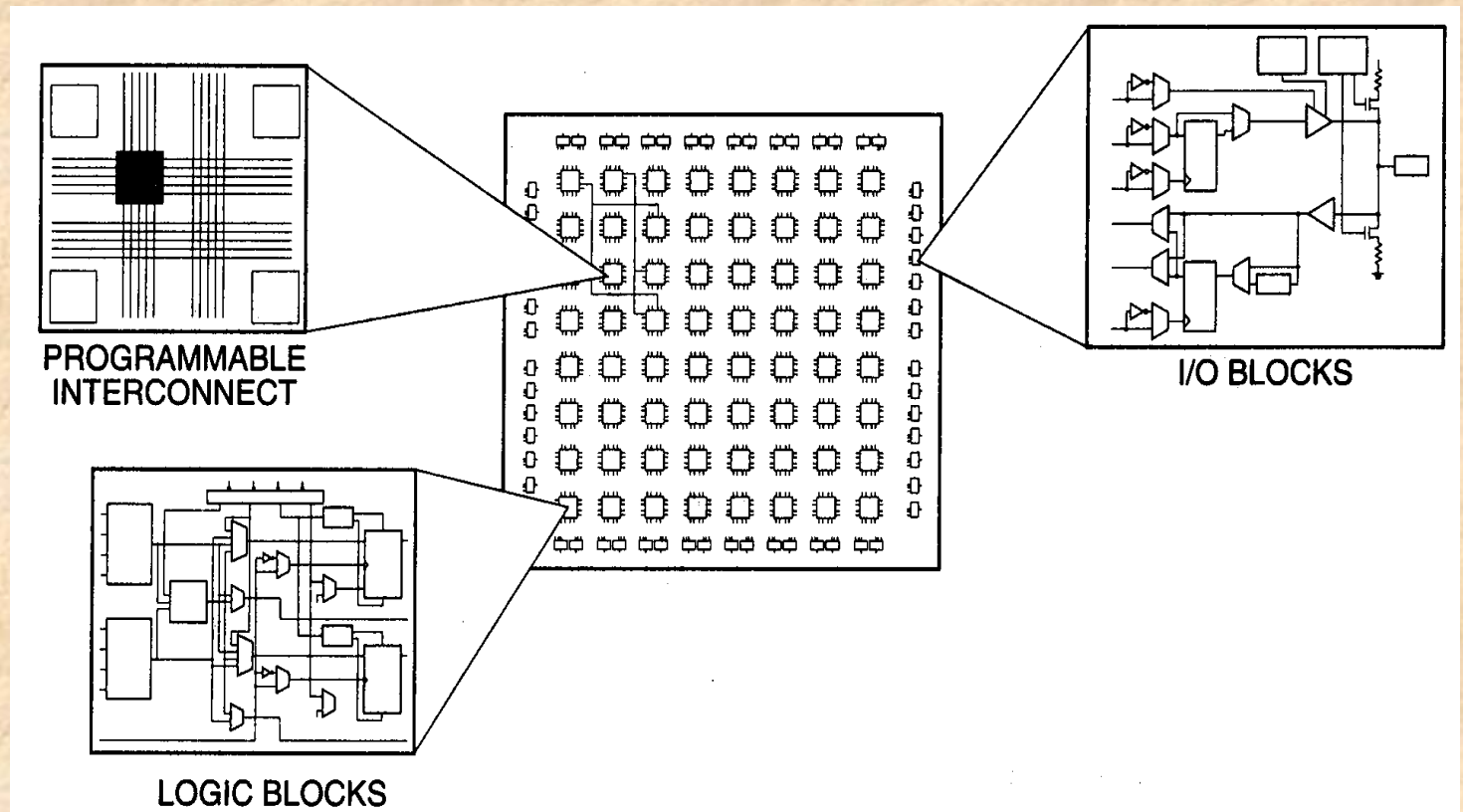


Czasy realizacji operacji wizyjnych: DePiAr, DSP i GPP

System	Histogram	Odejmowanie	Konwolucja 3×3	Mediana	LUT
■ DePiAr – 15 MHz	66,0 ns	132,0 ns	102,0 μ s 198,0 ns	68,0 μ s 198,0 ns	66,0 ns
■ TMS320C80	4,3 ms	5,4 ms	19,4 ms	10,7 ms	–
■ Dedykowany moduł Pentium II – 466 MHz	4,0 ms	2,7 ms	6,8 ms	–	3,0 ms
■ Akceleracja	60 000	20 000	34 000	54 000	45 000

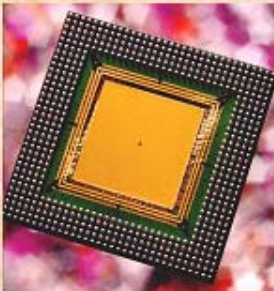


Implementacja w układach FPGA



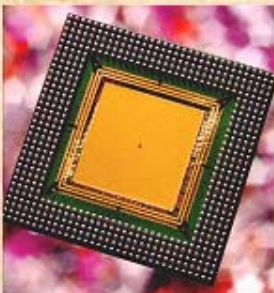
Parametryzacja implementacji w układach FPGA

- parametryzacja szybkości pracy
- parametryzacja użytych zasobów układu FPGA
- automatyczna generacja kodu VHDL
- optymalizacja implementacji struktur obliczeniowych DA

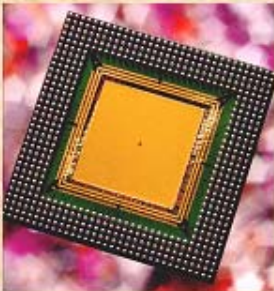
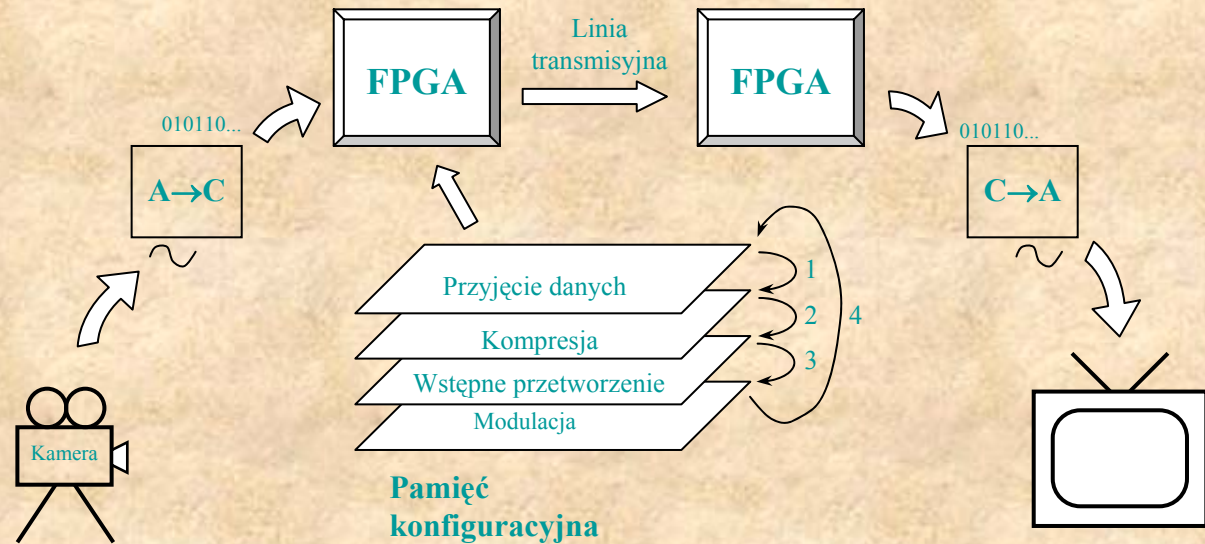


Dobór struktury elementów obliczeniowych

- układy mnożące bezmnożne MM
- układy mnożące LM z pamięcią LUT
- układy ze stałym współczynnikiem KCM
- układy ze zmiennym współczynnikiem VCM
- układy z dynamiczną rekonfiguracją DKCM

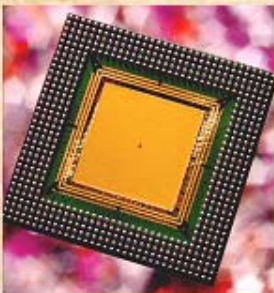


System z wielokrotnym wykorzystaniem do obliczeń jednego układu FPGA



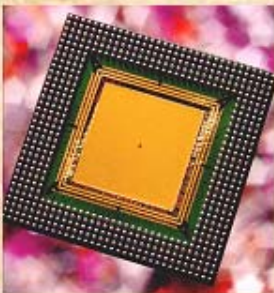
Parametry rekonfiguracji układów FPGA

Seria układu	Komórki CLB	Bramki [×1000]	Czas rekonfig.	Rekonfig. częściowa
Flex6000	1 960	24	100 [ms]	–
Flex8000	1 296	16	100 [ms]	–
Flex10000	12 160	250	320 [ms]	–
XC3000	484	7,5	10 [ms]	–
XC4000	7 448	85	240 [ms]	–
XC4000EX/XL	7 448	180	50 [ms]	–
XC6200	16 384	100	0,2 [ms]	Y
Spartan	1 862	40	32 [ms]	–
Virtex	27 648	1 124	0,1 [ms]	–
Virtex II	122 880	10 000	1 [ms]	Y
AT6000	6 400	30	8 [ms]	Y
AT40K	2 304	50	5 [ms]	Y
QL4000	1 584	90	40 [ms]	–
DY8000	6 272	105	200 [ms]	–



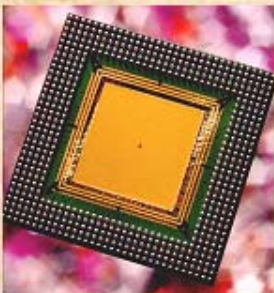
Systemy CTR i RTR

- 📖 systemy rekonfigurowane przed wykonaniem całości obliczeń CTR (ang. *Compile Time Reconfigurable*),
- 📖 systemy rekonfigurowane w trakcie obliczeń w celu wielokrotnego wykorzystania tego samego sprzętu dla realizacji różnych fragmentów realizowanego algorytmu RTR (ang. *Run Time Reconfigurable*).



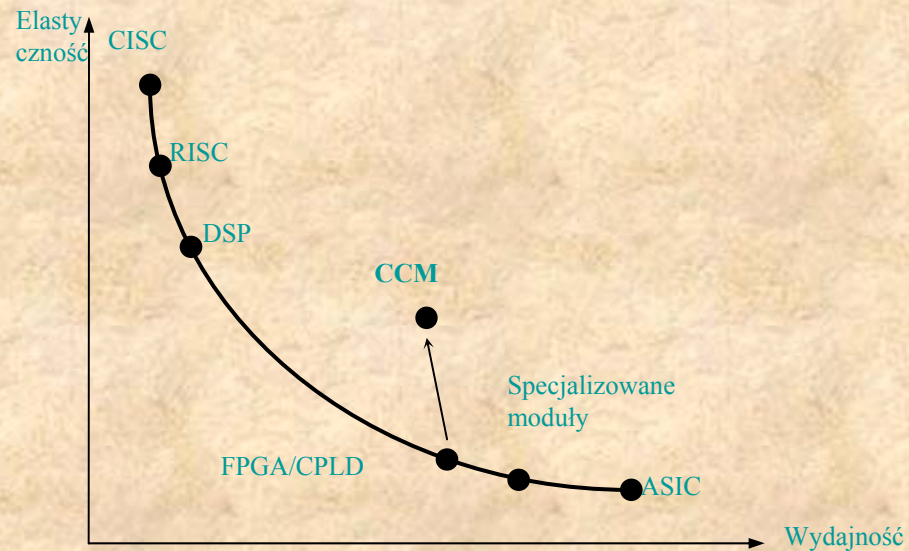
Efektywność stosowania systemów CTR i RTR

$$f = \frac{TR}{TO} \quad \frac{ERTR_{\max}}{ECTR} - 1 \geq f$$

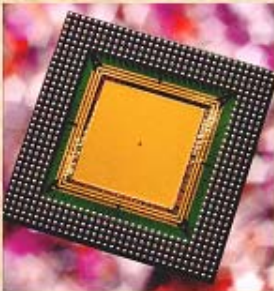


Dedykowane użytkownikowi platformy obliczeniowe CCM

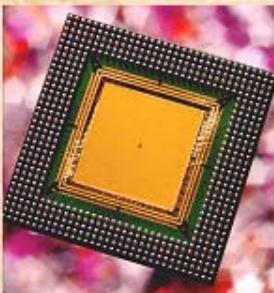
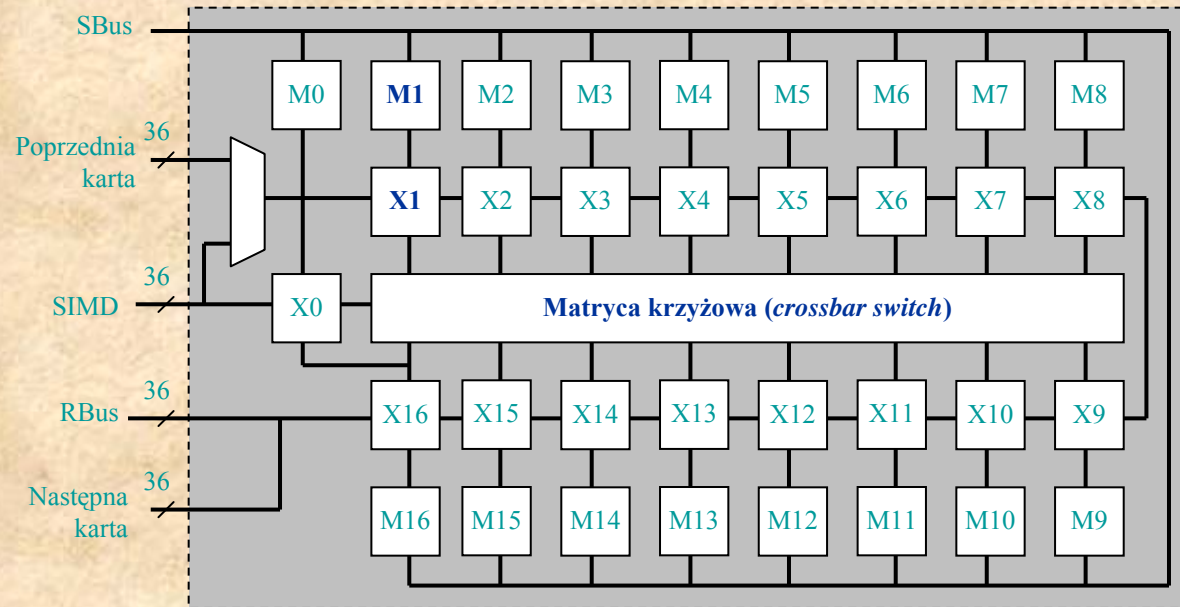
Custom Computing Machines



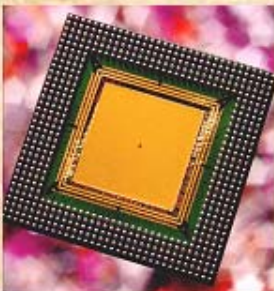
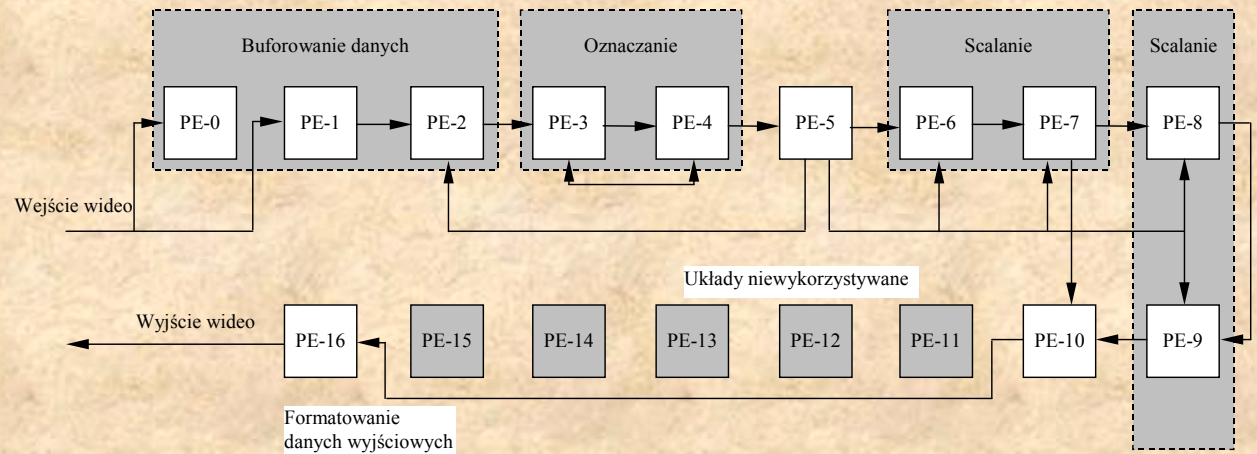
**Wydajność a elastyczność
układów obliczeniowych**



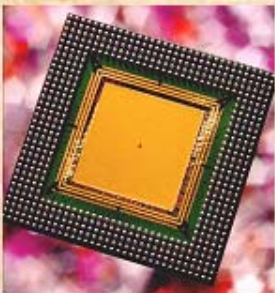
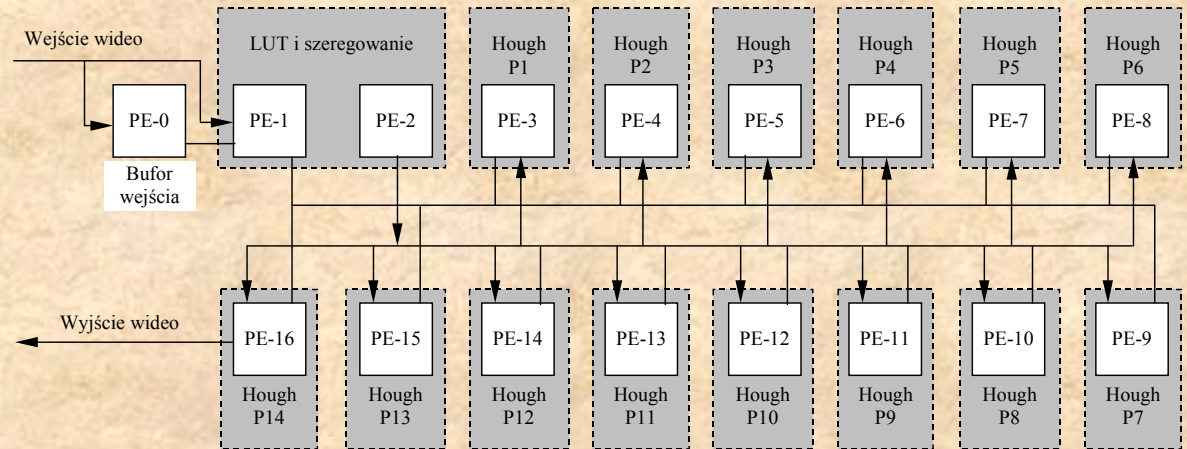
Architektura karty procesorowej w systemie *Splash2*



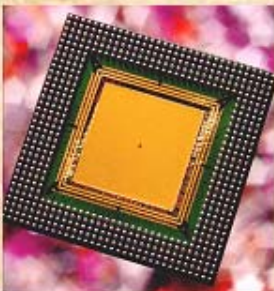
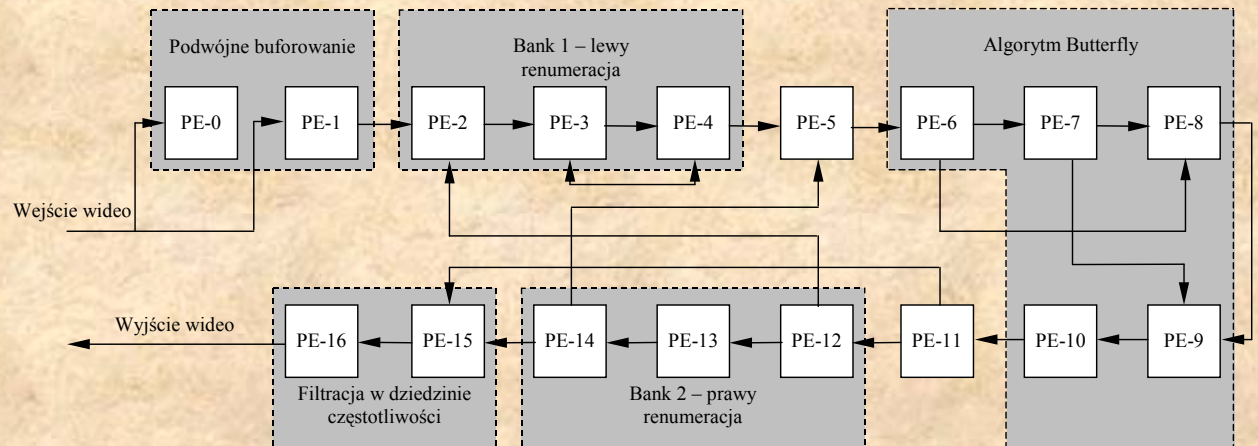
Implementacja detekcji obiektów i etykietowania



Implementacja transformacji Hough'a



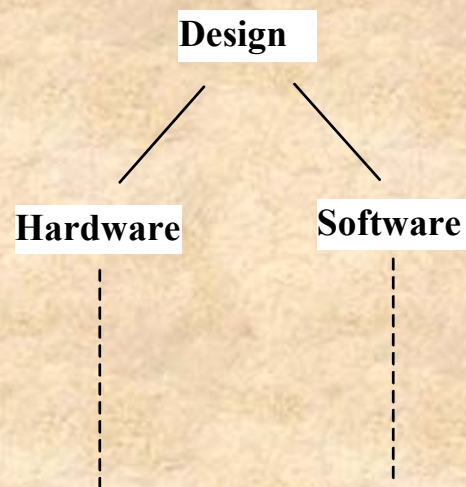
Implementacja szybkiej transformaty Fouriera FFT



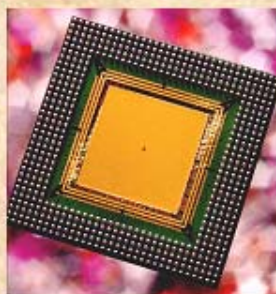
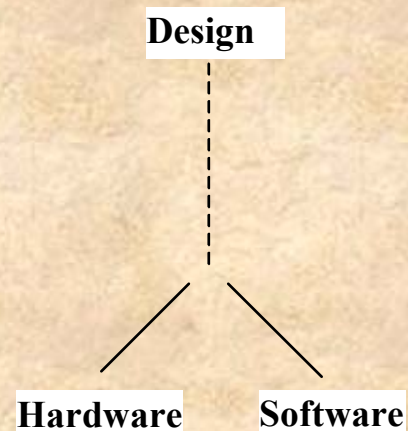
Projektowanie metodą Hardware/Software CoDesign

Wczesne i późne rozdzielanie hardware'u i software'u
w projektowaniu systemów

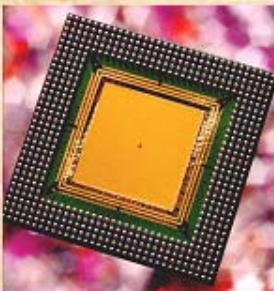
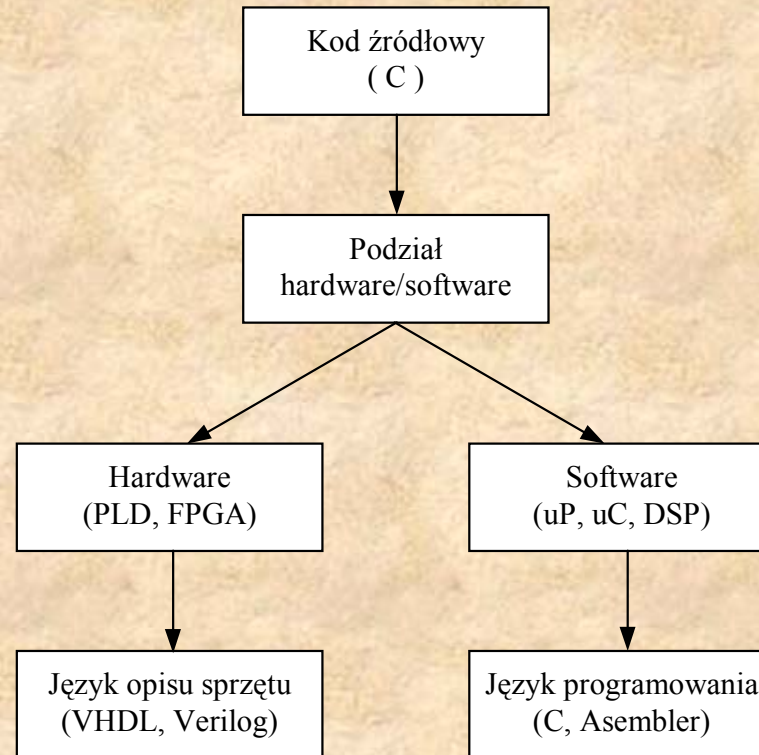
a)



b)

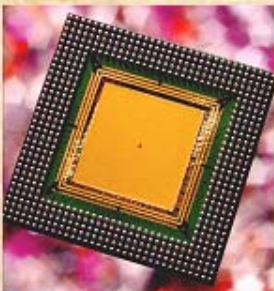


Hardware/Software CoDesign

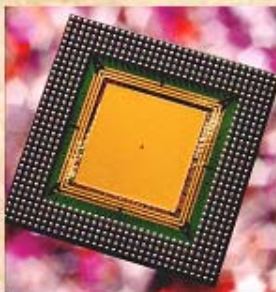
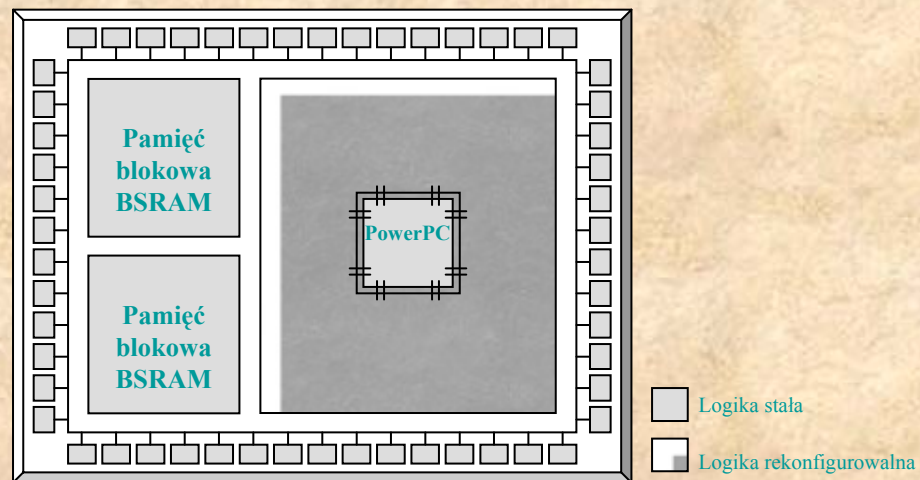


Wzrost pojemności układów programowalnych PLD

RODZAJ	FIRMA	BRAMKI	CLB	I/O	REJESTRY	KONFIG.
Classic	Altera	900	48	48	48	EEPROM
Max5000	Altera	3 800	192	64	192	EPROM
Flex8000	Altera	24 000	1 296	204	1 500	SRAM
APEX 20k	Altera	526 000	8 320	376		SRAM
APEX II	Altera	5 250 000	67 200	1 060		SRAM
ACT2	Actel	20 000		140	998	OTP
A500k	Actel	473 000	26 880	446	26 880	
AX2000	Actel	2 000 000	21 504	684	21 504	
ATV5100	Atmel	5 000		60	128	EPROM
ATK40k	Atmel	50 000	2 304	384	2 304	SRAM
XC4000	Xilinx	250 000	8 464	448	18 400	SRAM
XC9500	Xilinx	12 800	576	232	576	ISP(EEPROM)
SPARTAN	Xilinx	200 000	1 176	284	4 704	SRAM
VIRTEX	Xilinx	1 124 022	6 144	512	24 576	SRAM (2,5V)
VIRTEX II	Xilinx	10 000 000	15 360	1 108	61 440	SRAM (2,5V)
VIRTEX Pro	Xilinx		125 136	1 200		SRAM (1,5V)

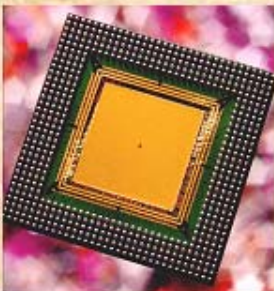


Struktura układu FPGA serii Virtex-II Pro

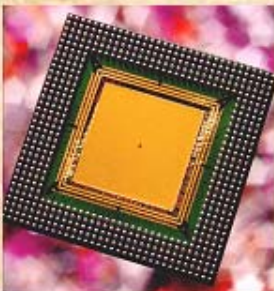
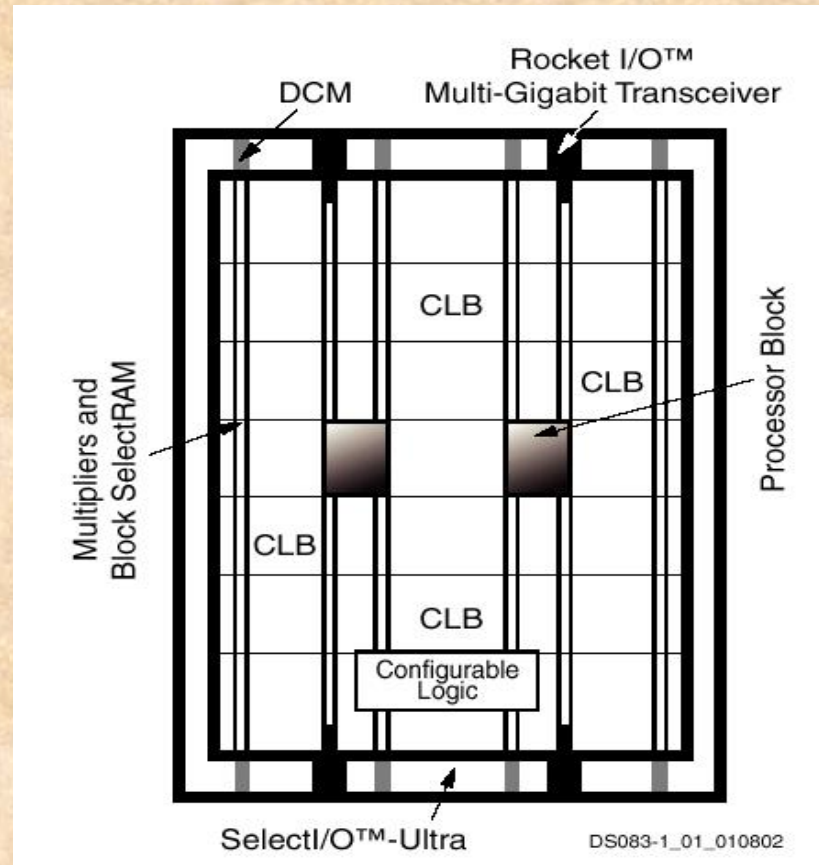


Zasoby układów FPGA serii Virtex-II Pro

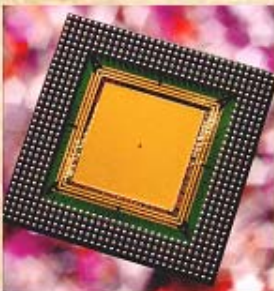
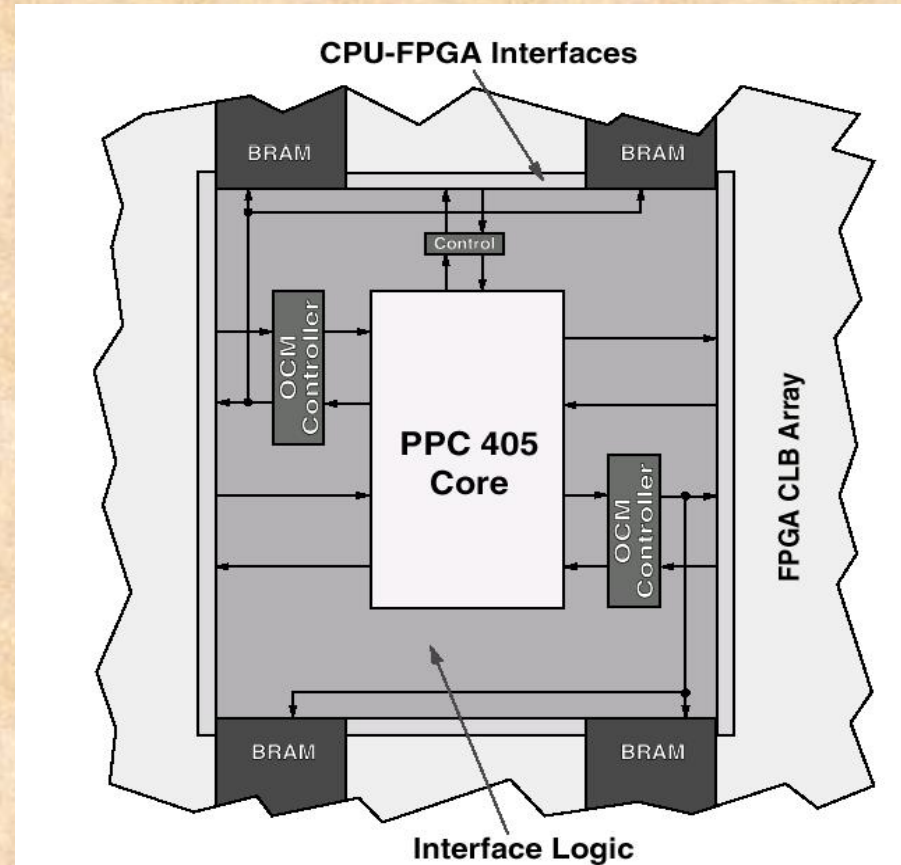
Układ	Rocket I/O	Power PC	LC	Slice	Distr. RAM kb	Mnożarki 18x18	BSR kb	DCM	I/O
XC2VP2	4	0	3 168	1 408	44	12	216	4	204
XC2VP4	4	1	6 768	3 008	94	28	504	4	348
XC2VP7	8	1	11 088	4 928	154	44	792	4	396
XC2VP20	8	2	20 880	9 280	290	88	1 584	8	564
XC2VP30	8	2	30 816	13 696	428	136	2 448	8	692
XC2VP40	12	2	43 632	19 392	606	192	3 456	8	804
XC2VP50	16	2	53 136	23 616	738	232	4 176	8	852
XC2VP70	20	2	74 448	33 088	1 034	328	5 904	8	996
XC2VP100	20	2	99 216	44 096	1 378	444	7 992	12	1164
XC2VP125	24	4	125 136	55 616	1 738	556	10 008	12	1200



Lokalizacja zasobów w układzie Virtex-II Pro

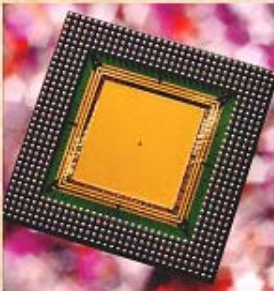
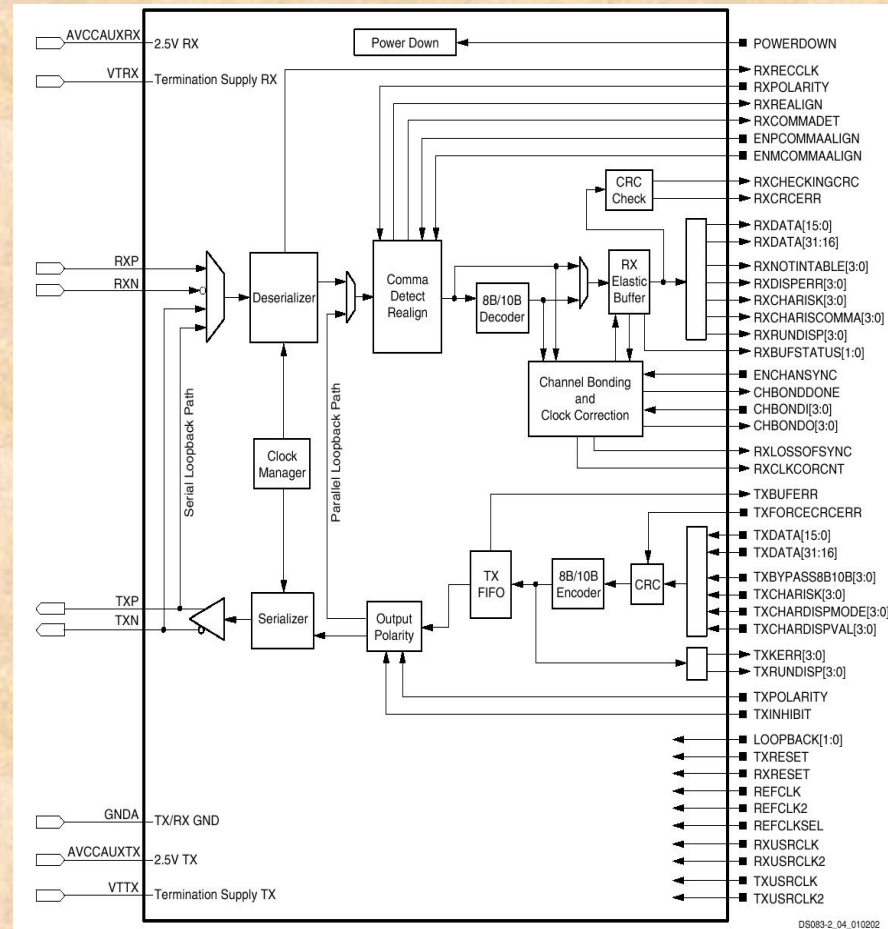


Przyłączenie wbudowanego procesora PowerPC do zasobów układu Virtex-II Pro

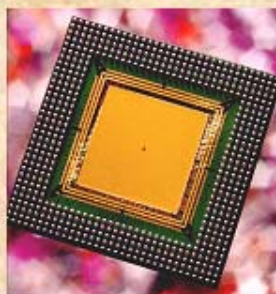
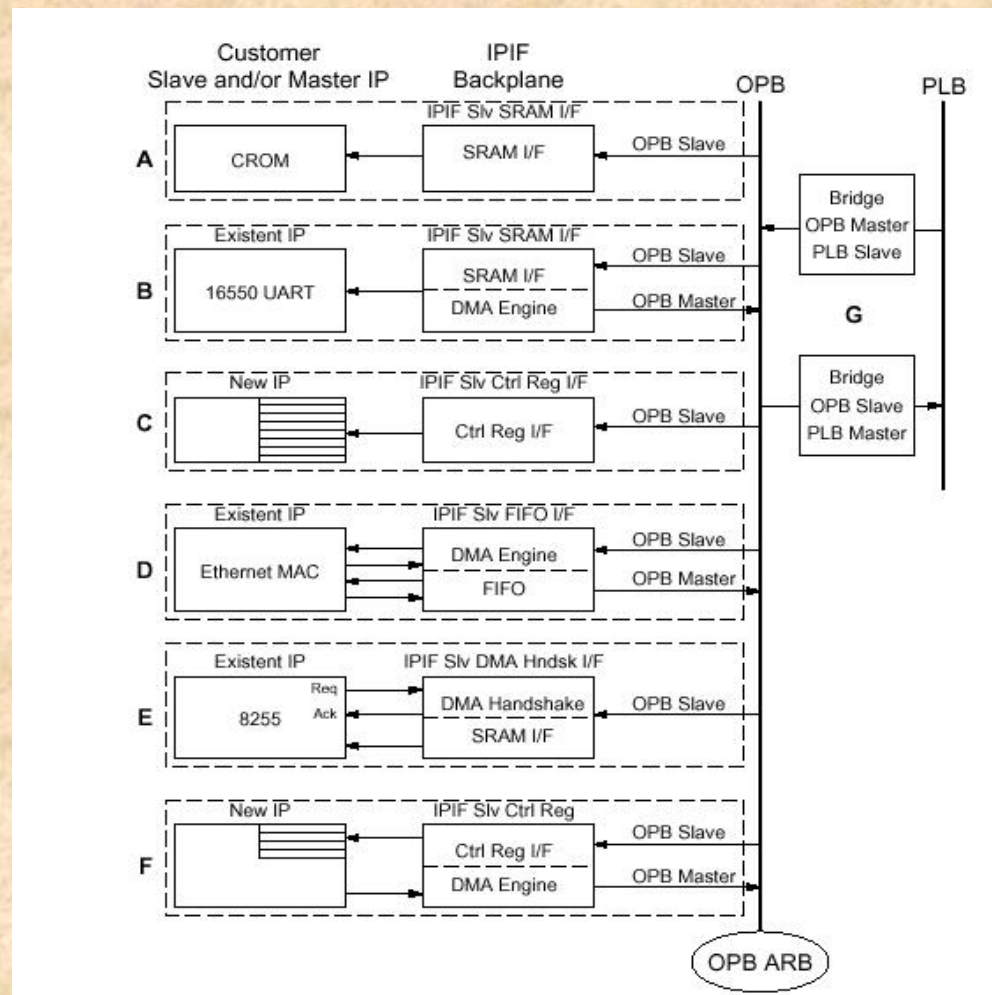


Blok łącz szeregowych Rocket I/O

o przepustowości 3,125 Gb/s

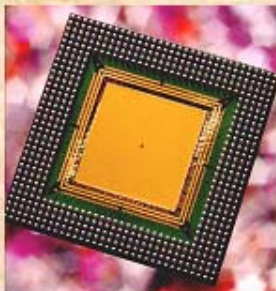
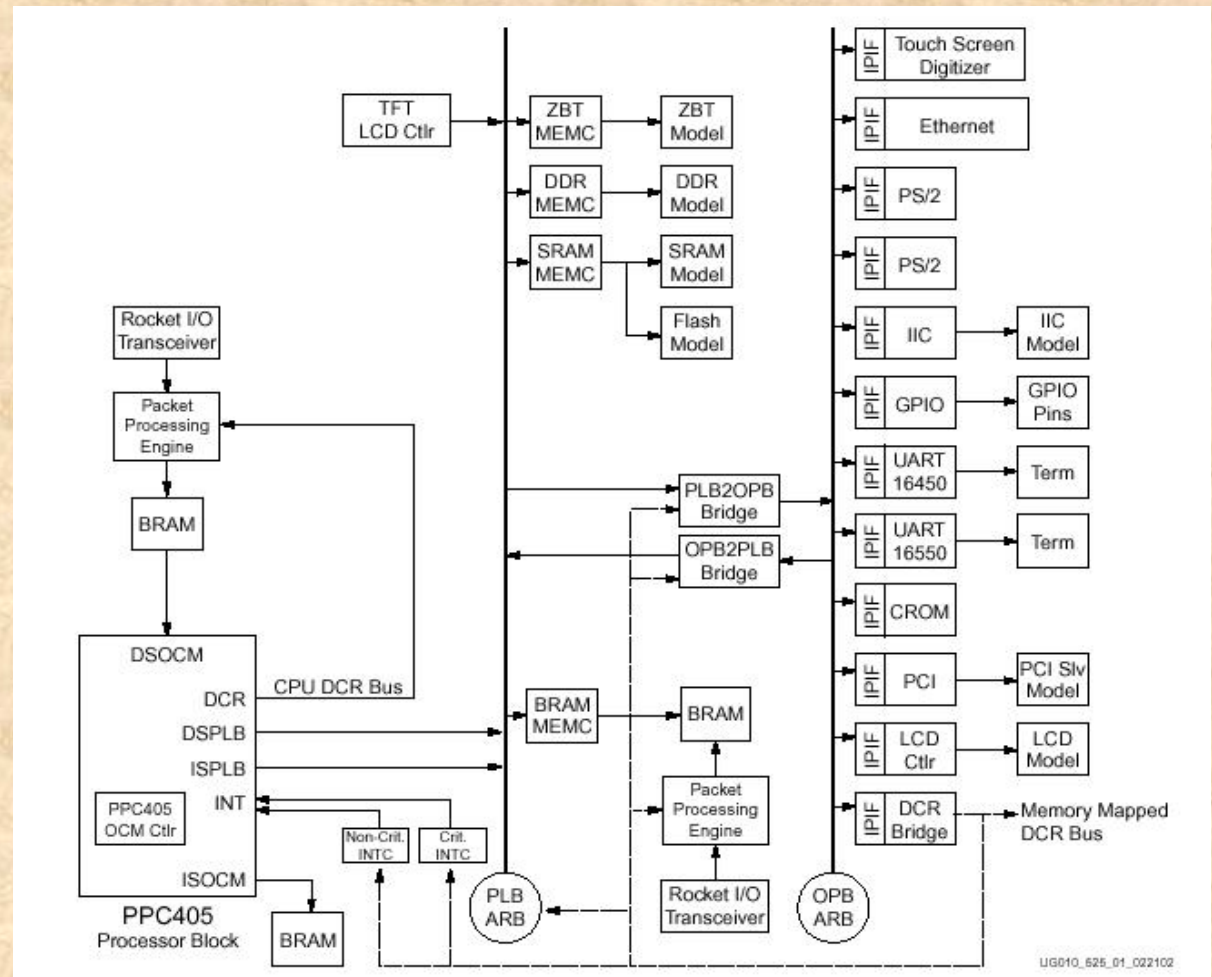


Połączenie elementów IP poprzez interfejs IPIF z magistralą peryferyjną OPB i magistralą lokalną procesora PLB



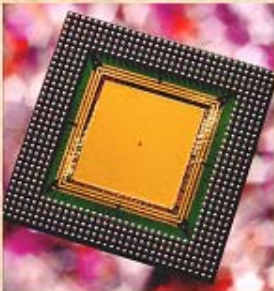
Przyłączenie jednego z procesorów PowerPC 405

do magistrali lokalnej procesora PLB oraz układów peryferyjnych

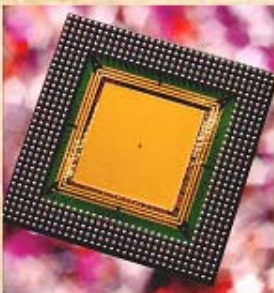
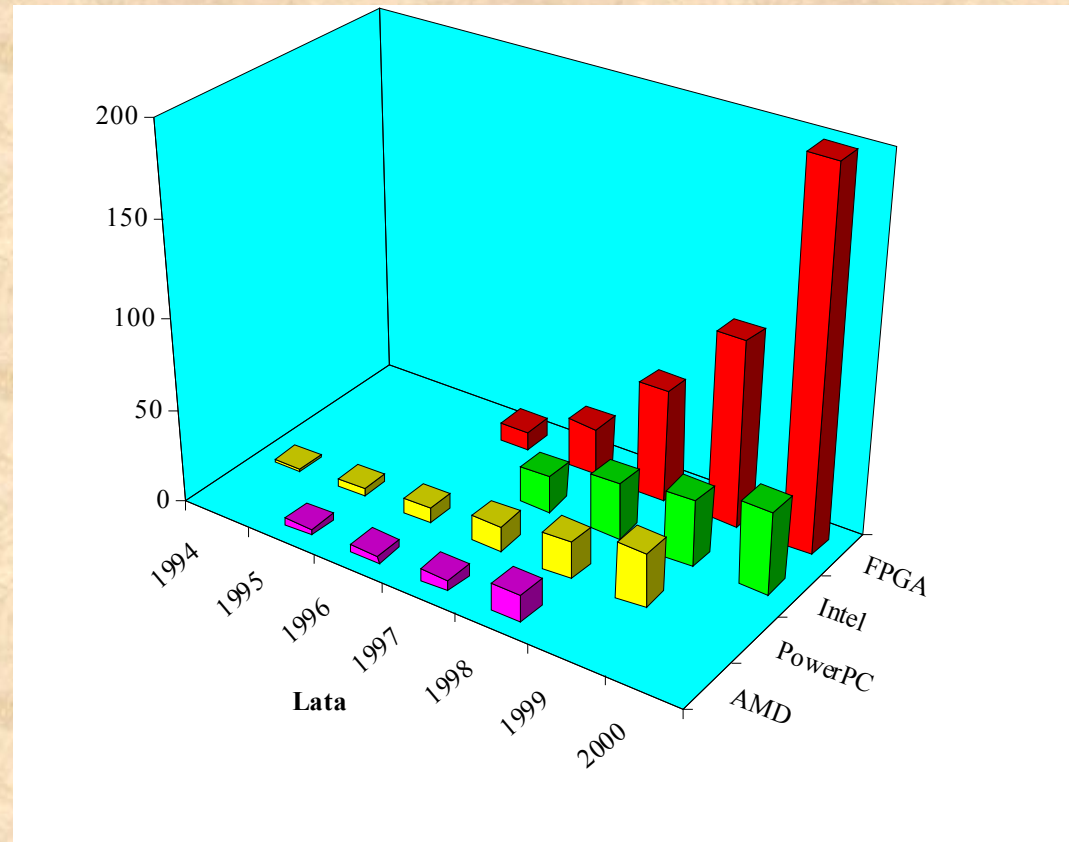


Elementy IP - operacje przetwarzania obrazów

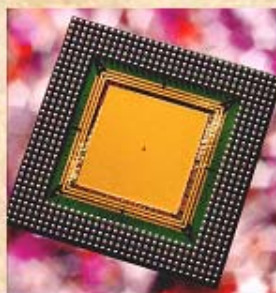
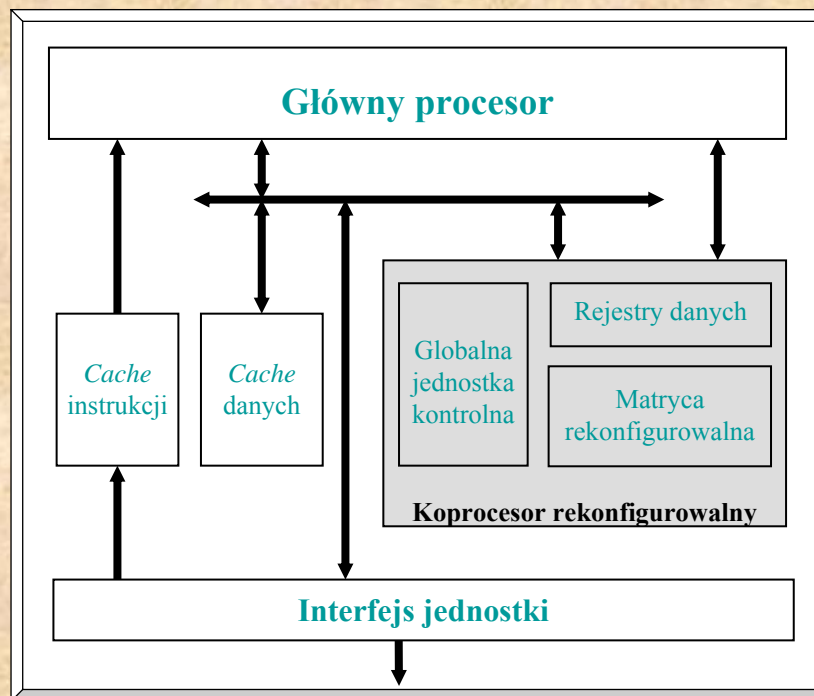
Funkcja	Biblioteka	Firma	Układ	Użyte zasoby	f [MHz]
FFT/IFFT 1024 p-pty	LogiCORE	Xilinx	XC2V500	62 %	100
FFT/IFFT 64 p-pty	LogiCORE	Xilinx	XC2V500	38 %	100
FFT/IFFT 32 p-pty	LogiCORE	Xilinx	XC2V500	29 %	110
FFT/IFFT 16 p-pty	LogiCORE	Xilinx	XC2V500	37 %	130
2D DCT/IDCT	AllianceCORE	Barco-Silex	XC2V250	77 %	133
2D FDCT	AllianceCORE	CAST	XC2V500	42 %	83
2D DWT	AllianceCORE	CAST	XC2V250	62 %	52
FIDCT	AllianceCORE	Telecom	XCV200	77 %	78
MAC FIR	LogiCORE	Xilinx	XCV250	16 %	
Dekoder Huffmana	AllianceCORE	CAST	XC2V1000	22 %	25
Dekoder Reed Solomon	AllianceCORE	Telecom	XC2V500	97 %	61
TMS32025 DSP core	AllianceCORE	CAST	XC2V500	66 %	63

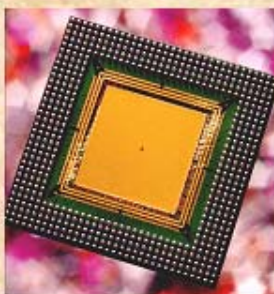
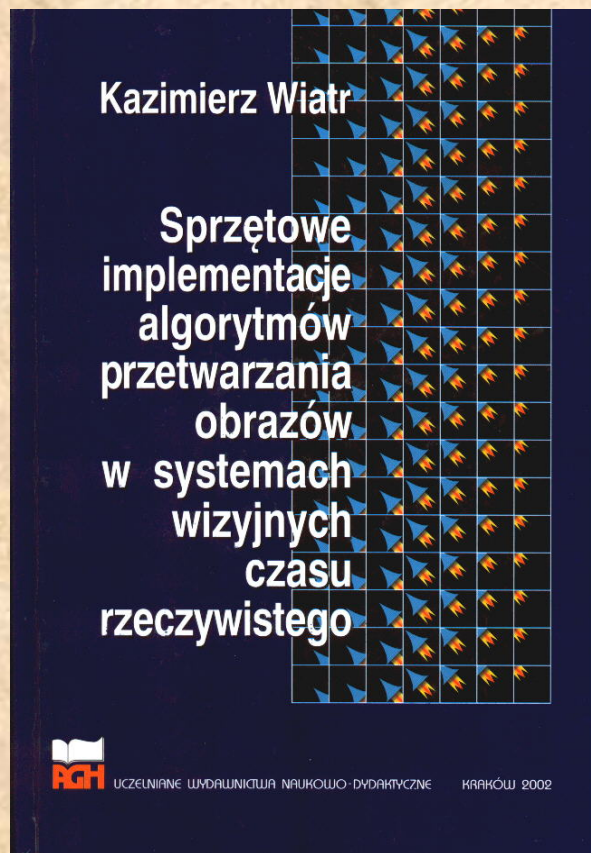


Wzrost liczby tranzystorów w układach różnego typu



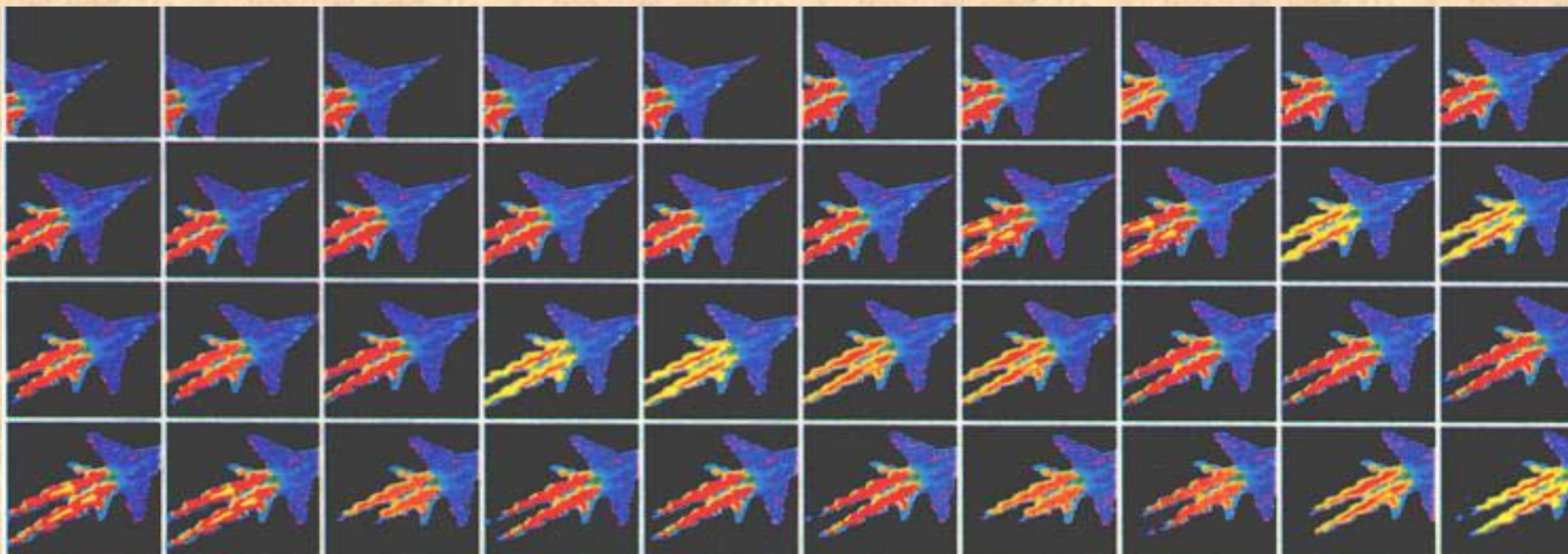
Mikroprocesor z rekonfigurowalnym koprocesorem



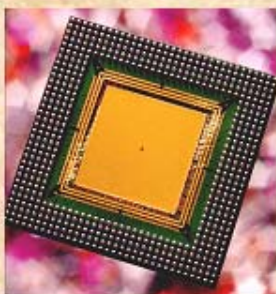


Kazimierz Wiatr, ACK CYFRONET AGH, Zespół Akceleracji Obliczeń

Potrzeby obliczeniowe systemów nieustannie rosną



- szybkość akwizycji 1000 do 5000 obrazów na sekundę
- rozdzielczość 128x128 i 256x256
- napływ pikseli do 327,68 MHz (3,1 ns)
- standardowo 14,75 MHz (67,8 ns)



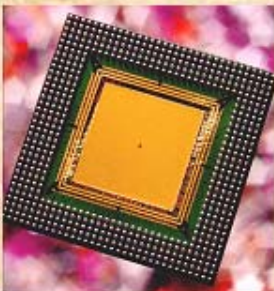
1.1. Introduction to the X-Stream™ Camera



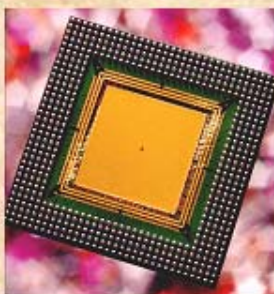
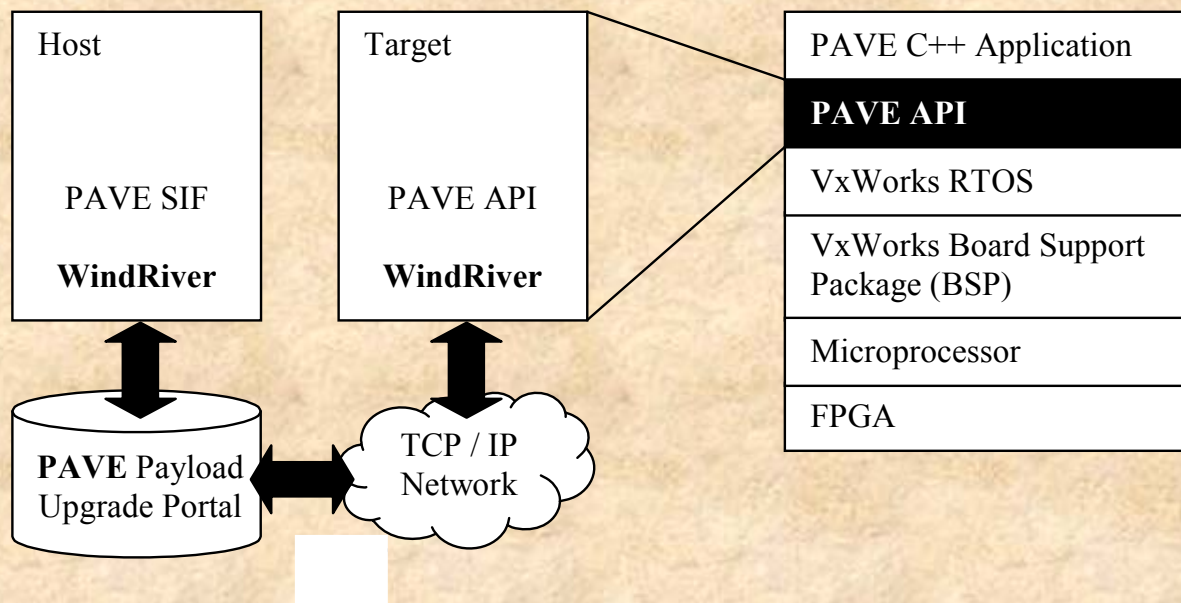
X-Stream VISION Framing Rates

Resolution	Frame Rate [Hz]
1280H x 1024V	628
1280H x 640V	1000
1280H x 320V	2000
1280H x 160V	4000
1280H x 80V	8000
1280H x 40V	16000
1280H x 20V	32000

- **1280 x 1024 x 628 = 823 MHz (823 132 160 Hz)**
- **10 bitów/piksel**



Logika rekonfigurowana za pomocą internetu wykorzystywana do akceleracji obliczeń



Infrastruktura informatyczna ACK

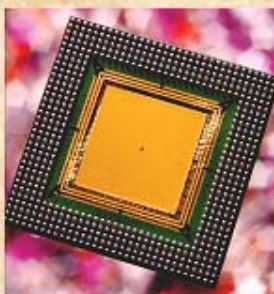


2 x 10 Gbps

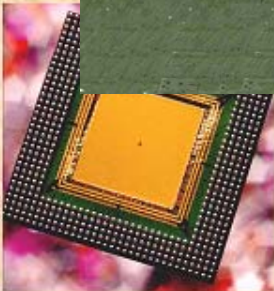


- Klaster PC RackSaver – **2 TFLOPS**
- Klaster IBM BladeCenter HS21 - **1,2 TFLOPS**
- SGI Altix 3700 – 768 GFLOPS
- Klaster HP Integrity rx2600 - **291 GFLOPS**
- SGI Altix 4700 – 212 GFLOPS
- HP Integrity SuperDome - **48 GFLOPS**
- SunFire 6800 - **43,2 GFLOPS**
- SunFire V490 – **24 GFLOPS**

4,6 TFLOPS



Komputery Dużej Mocy Obliczeniowej KDMO



Kazimierz Wiatr, ACK CYFRONET AGH, Zespół Akceleracji Obliczeń

system operacyjny:

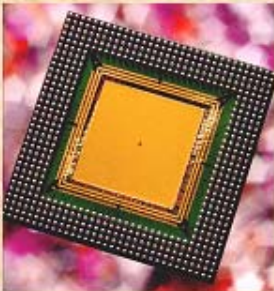
Suse Linux Enterprise Server 9

konfiguracja:

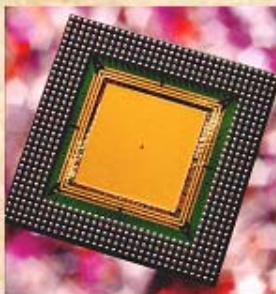
- 128 procesorów
Intel Itanium 2
z zegarem 1.5 GHz
- pamięć operacyjna 256 GB
- pamięć dyskowa 1,46 TB

- **moc obliczeniowa
768 Gflops**

SGI Altix 3700



Kazimierz Wiatr, ACK CYFRONET AGH, Zespół Akceleracji Obliczeń



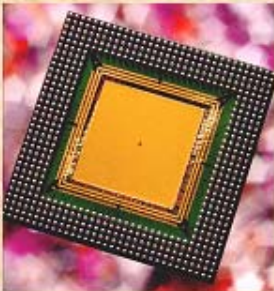
Kazimierz Wiatr, ACK CYFRONET AGH, Zespół Akceleracji Obliczeń

Klaster komputerów PC

- 384 procesorów Xeon
- 440 GB pamięci operacyjnej
- 17.4 TB pamięci dyskowej
- **moc obliczeniowa 2071 Gflops**

Klaster serwerów HP Integrity rx2600

- 56 procesorów Intel Itanium2
- 56 GB pamięci operacyjnej
- 2 TB pamięci dyskowej
- **moc obliczeniowa 291 Gflops**



Klaster obliczeniowy

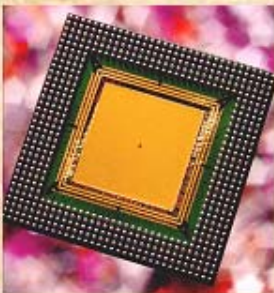
IBM BladeCenter HS21

system operacyjny:

Linux RedHat

konfiguracja:

- 112 procesorów Intel Xeon Dual Core 2.66 GHz
- pamięć operacyjna: 448 GB
- pamięć dyskowa: 5TB
- **moc obliczeniowa 1192 Gflops**



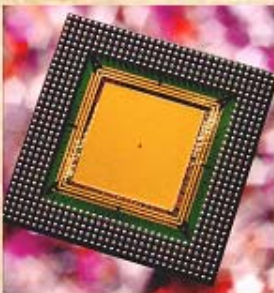
SGI Altix 4700

system operacyjny:

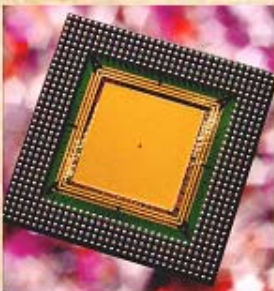
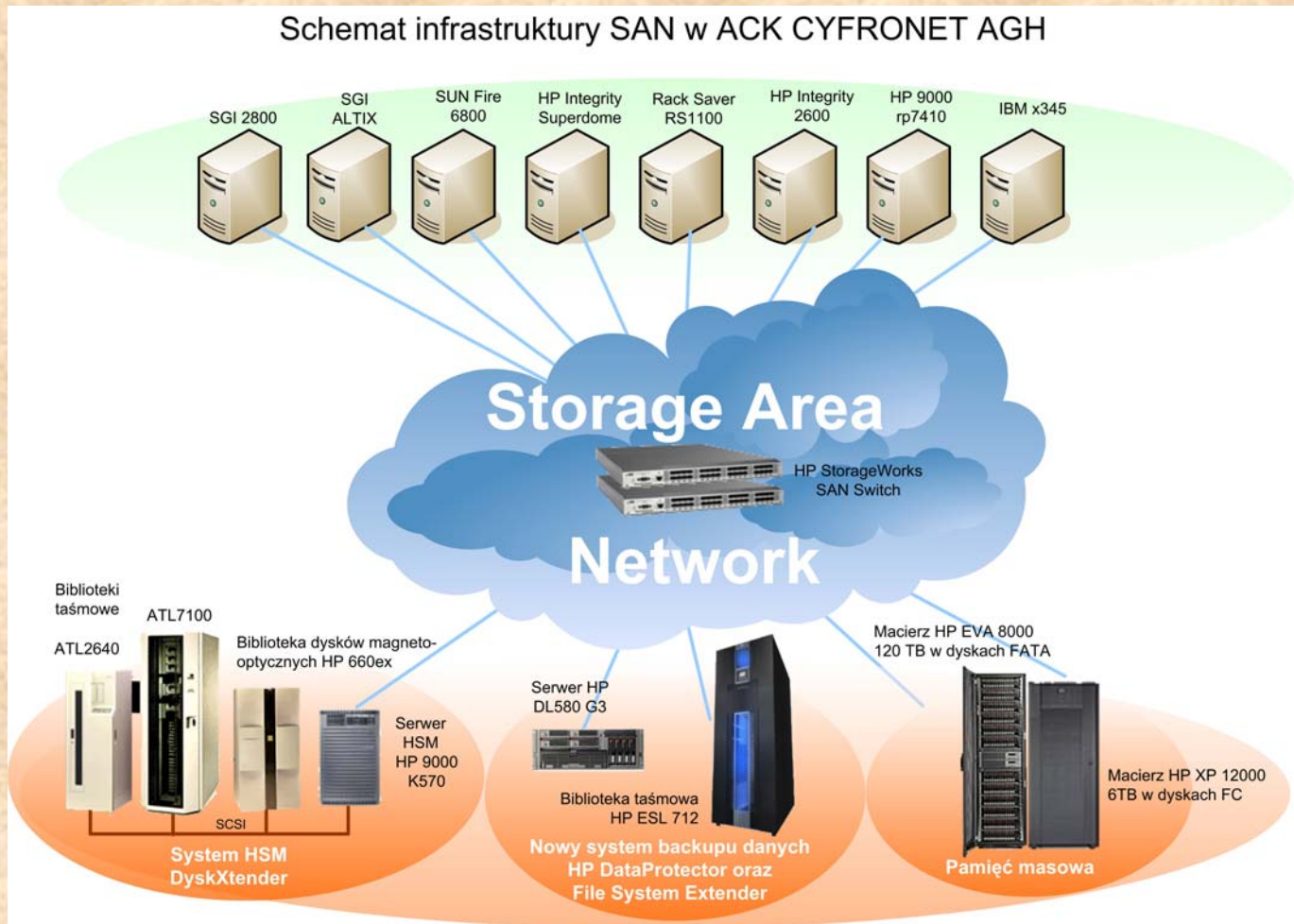
Suse Linux Enterprise Server 9

konfiguracja:

- 32 procesory Intel Itanium 2 z zegarem 1.66 GHz
- pamięć operacyjna 64 GB
- pamięć dyskowa 1,8 TB
- **moc obliczeniowa 212 Gflops**



Hierarchiczny system składowania danych



Hierarchiczny system składowania danych

W ACK CYFRONET AGH
serwery obliczeniowe i sieciowe
przyłączone są do centralnego zasobu
pamięci dyskowej

o pojemności 130 TB

Macierze dyskowe:

HP XP 12000 - 6 TB

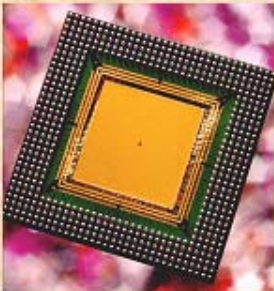
HP EVA 8000 - 120 TB

Biblioteki taśmowe:

ATL 2640 - pojemność 10 TB

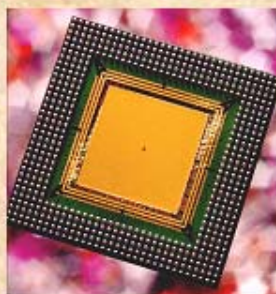
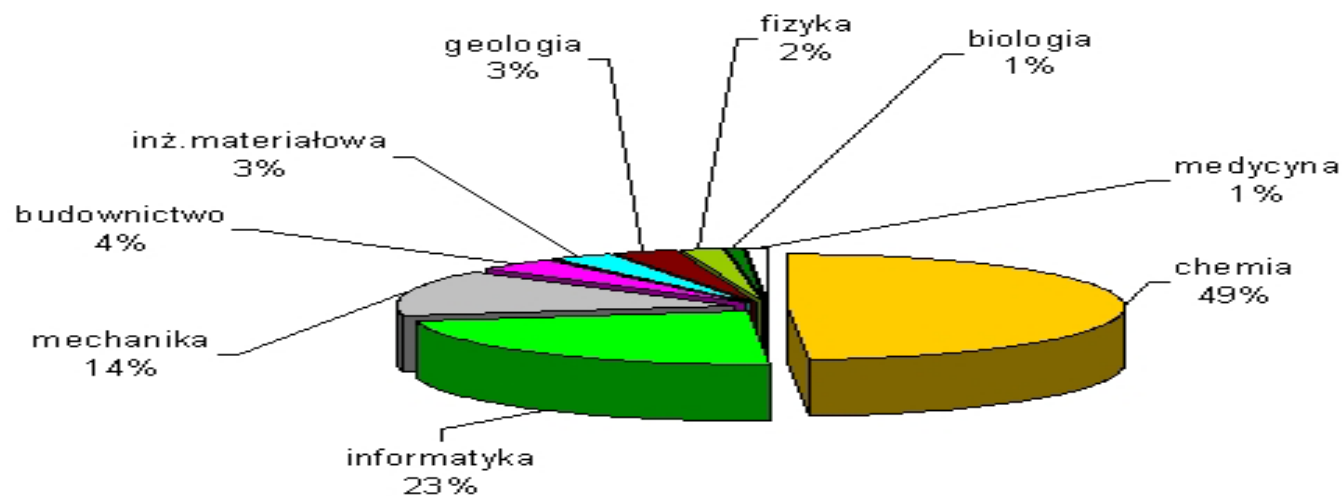
ATL 7100 - pojemność 7 TB

HP ESL712e - **pojemność 280 TB**



Wykorzystanie SGI Altix 3700

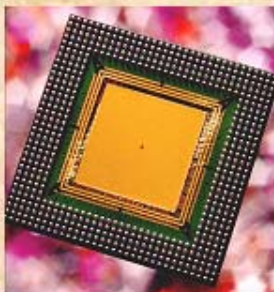
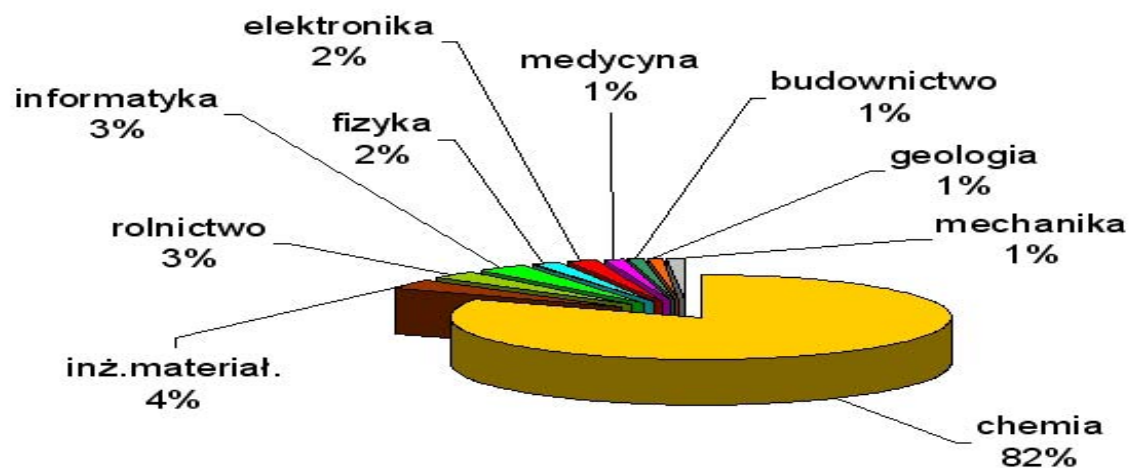
w roku 2006 w/g dyscyplin naukowych



Kazimierz Wiatr, ACK CYFRONET AGH, Zespół Akceleracji Obliczeń

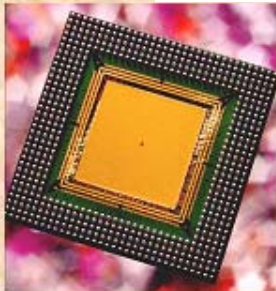
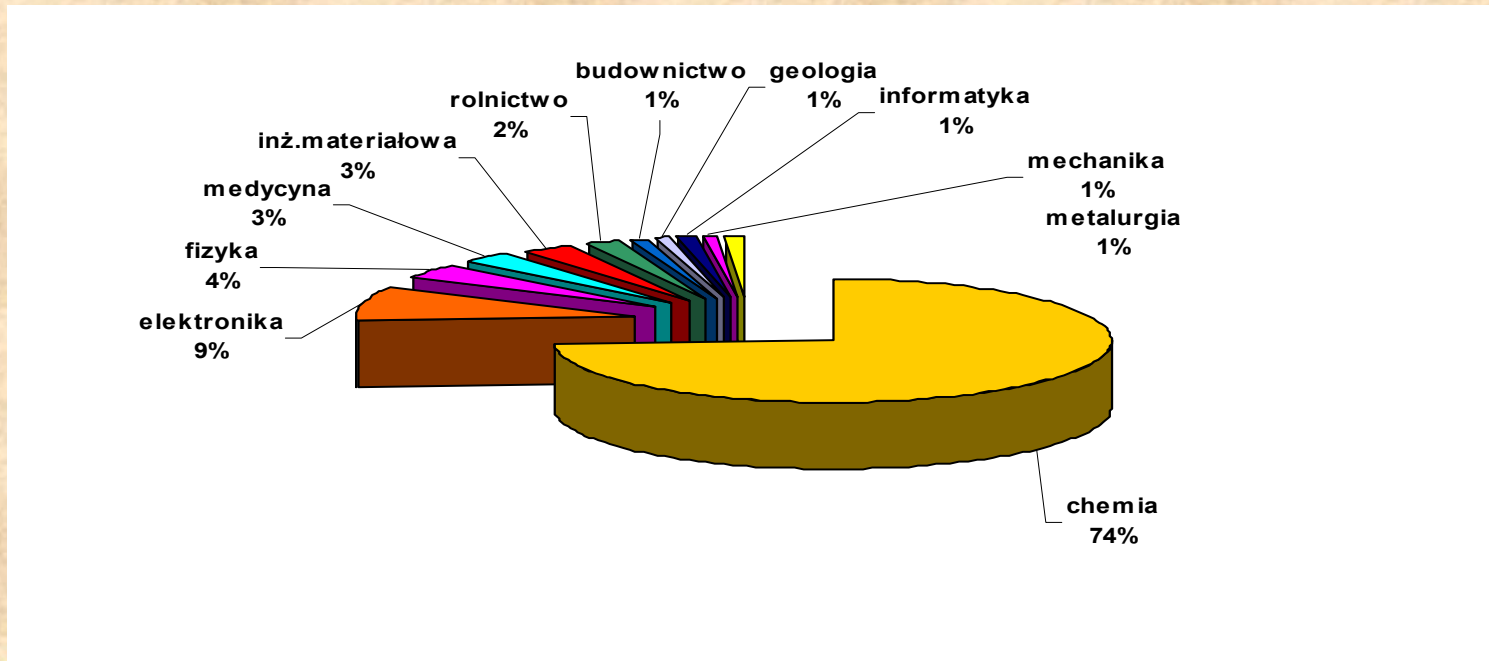
Wykorzystanie SGI 2800

w roku 2006 w/g dyscyplin naukowych



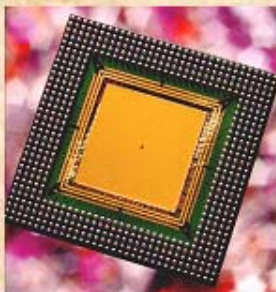
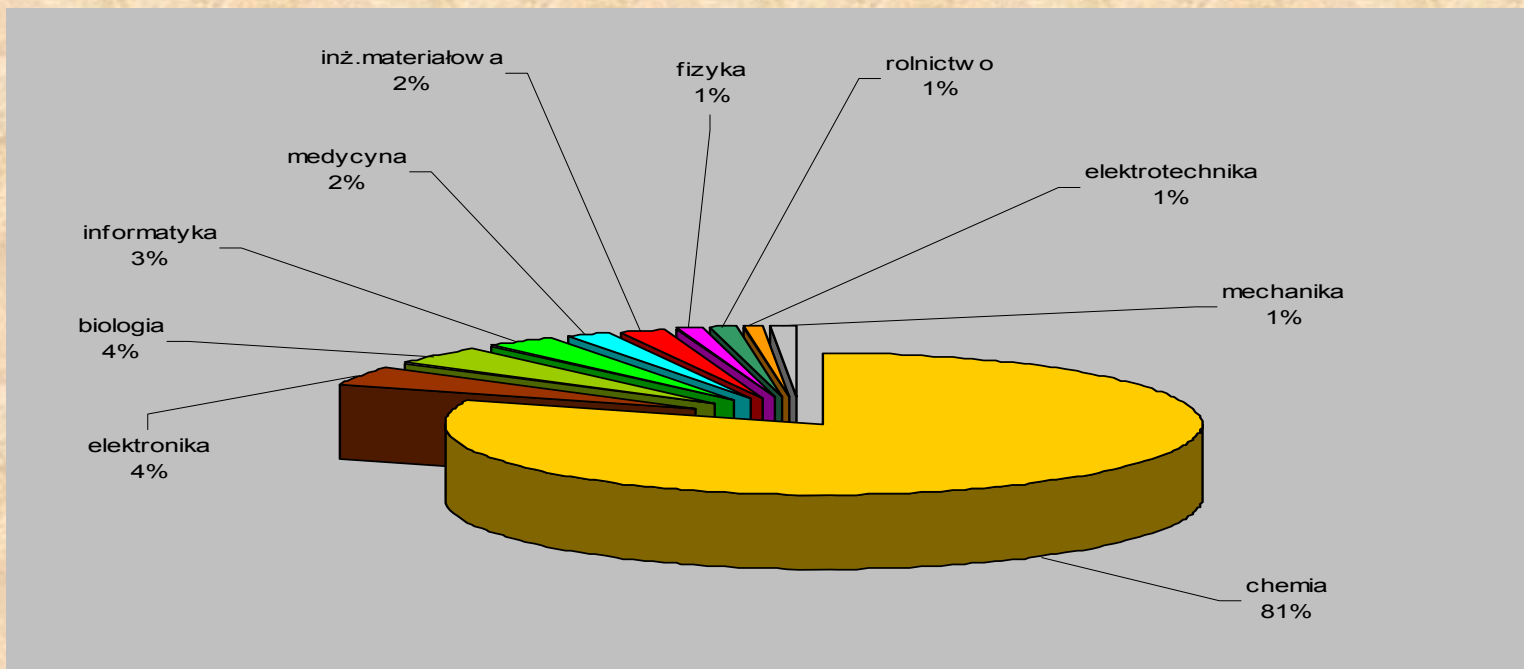
Wykorzystanie SGI 2800

w roku 2005 w/g dyscyplin naukowych



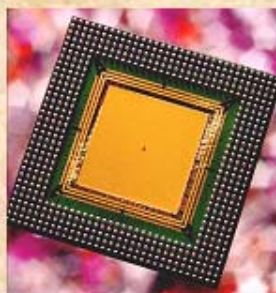
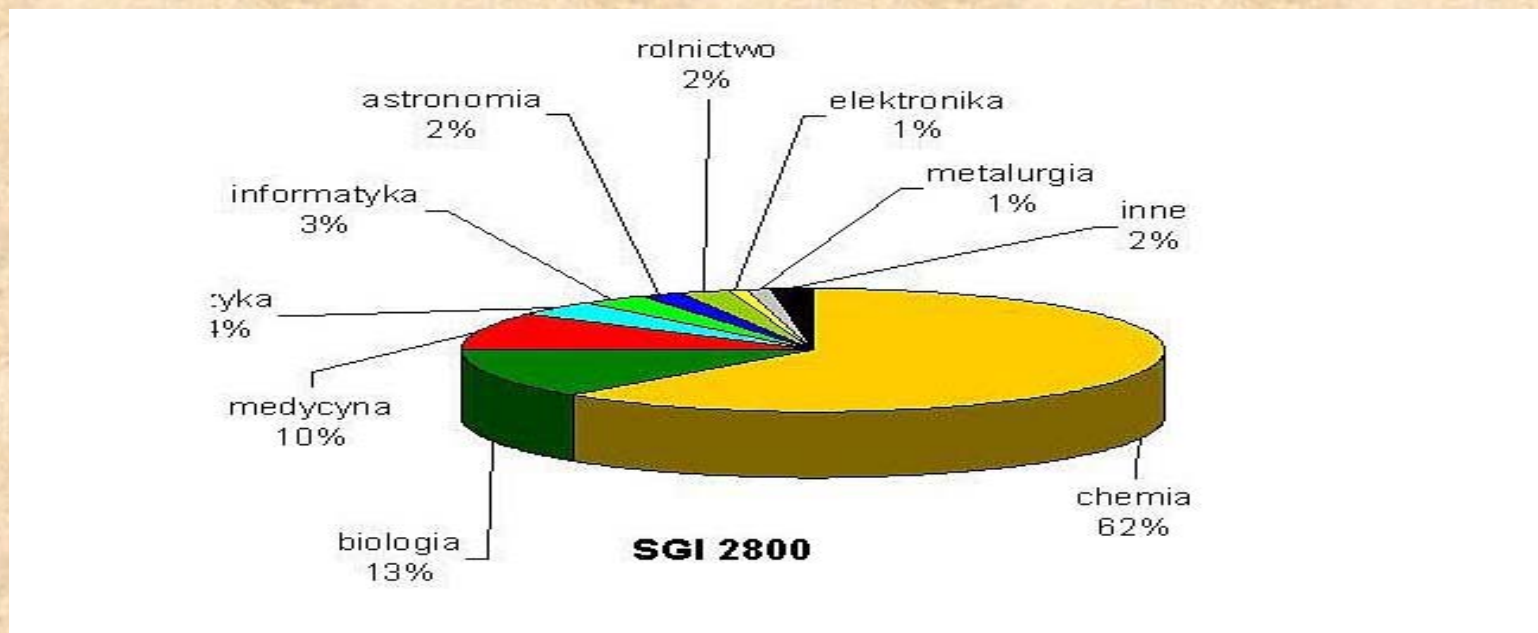
Wykorzystanie SGI 2800

w roku 2004 w/g dyscyplin naukowych

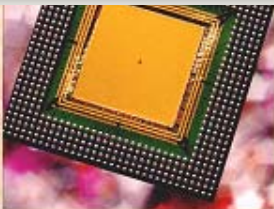
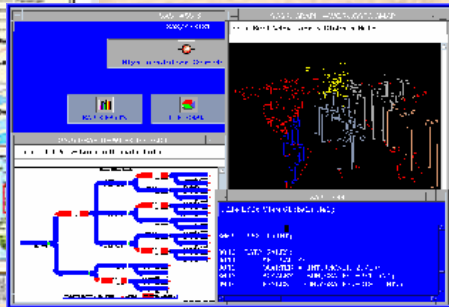
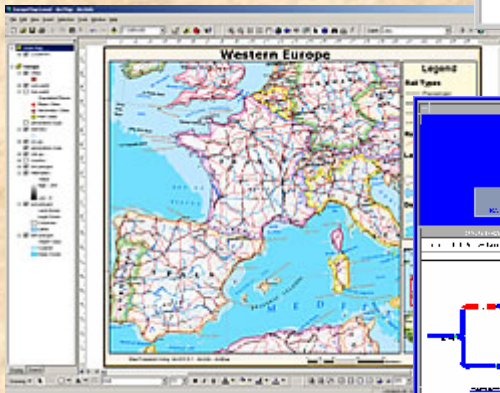
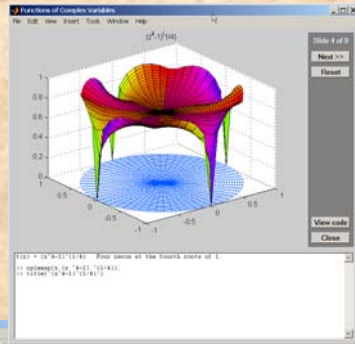
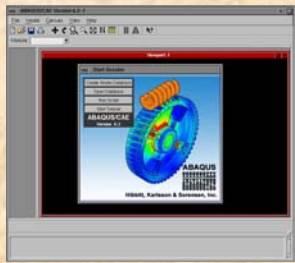


Wykorzystanie SGI 2800

w roku 2003 w/g dyscyplin naukowych

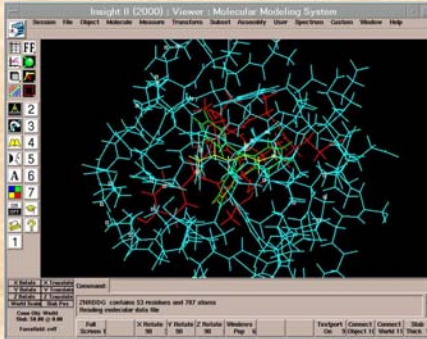


Zestawienie oprogramowania aplikacyjnego



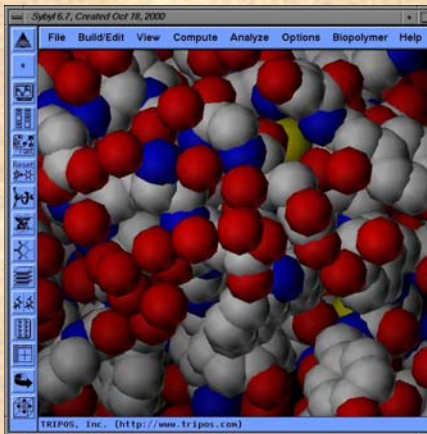
Program	SGI 2800	SGI Altix 3700	SUN Fire 6800	HP Super Dome	SGI Onyx 300, SGI Octane/SE
ABAQUS	x	x	x		x
Accelrys/Insight II					x
Accelrys/Cerius2					x
Accelrys/Catalyst					x
Accelrys/Quanta					x
ANSYS	x	x		x	
ARC/INFO					x
ERDAS/IMAGINE					x
FIDAP	x		x		
FLUENT	x	x	x	x	
GAMSS		x			
GAUSSIAN 03	x	x			
MAPLE			x	x	
MATHEMATICA			x		
MATLAB	x		x		
MSC/FATIGUE	x		x		
MSC/NASTRAN	x		x		
MSC/PATRAN	x		x		
OPERA-2d	x				x
ORACLE			x		
SAS				x	
SYBYL					x

Aplikacje chemiczne



ACCELRYS

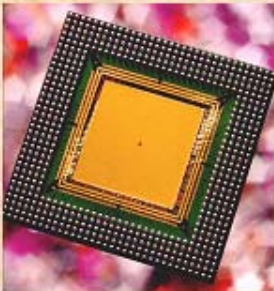
- **InsightII** - do modelowania dużych molekuł biologicznych
- **Cerius2** - do modelowania małych molekuł i ciała stałego
- **Quanta** - do modelowania molekularnego przeznaczony z zakresu krytalografii
- **Catalyst** - pakiet do projektowania leków



SYBYL - pakiet programów do modelowania i analizy struktur molekularnych. Celem oprogramowania jest budowanie, analiza i manipulacja molekułami.

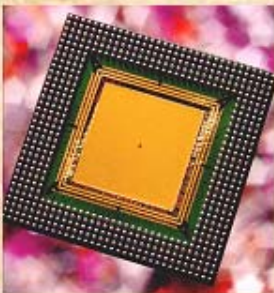
GAUSSIAN - system przeznaczony do obliczeń orbitali molekularnych przy użyciu metod półempirycznych i ab initio.

GAMESS - wszechstronny pakiet do obliczeń chemii kwantowej.



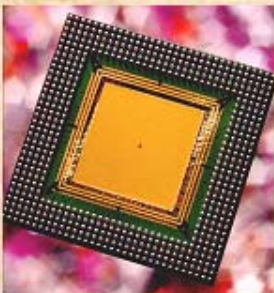
SGI Altix 4700 i projekt Gaussian

- Pakiet Gaussian (82% obliczeń)
- Przyspieszenie obliczeń kwantowo-chemicznych
- Operacje algebraiczne na macierzach

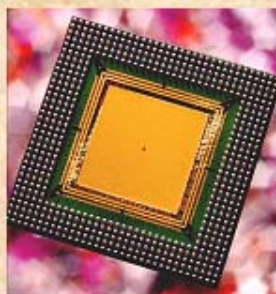
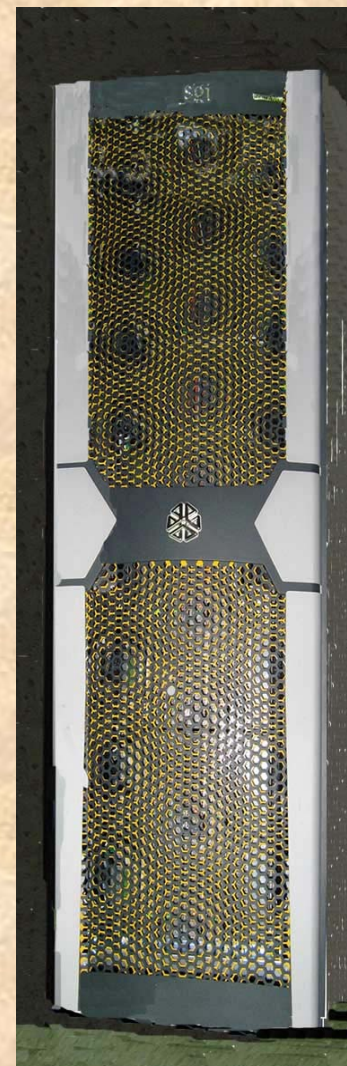
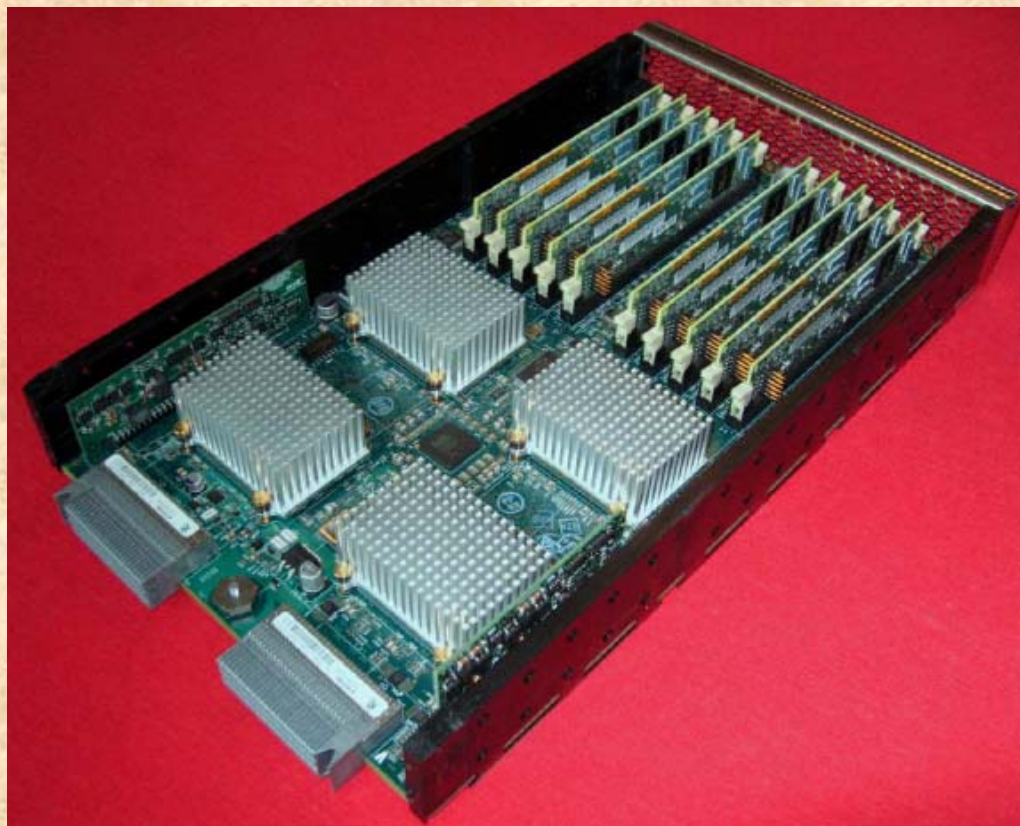


Problemy do rozwiązania

- Funktory FP
- Implementacja bibliotek BLAS i LAPACK
- Równoległe algorytmy operacji na macierzach
- Biblioteki softwer'owe



SGI RASC Blade

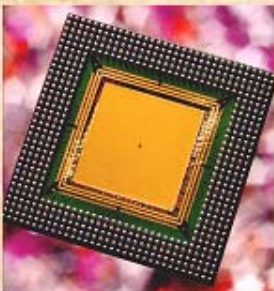
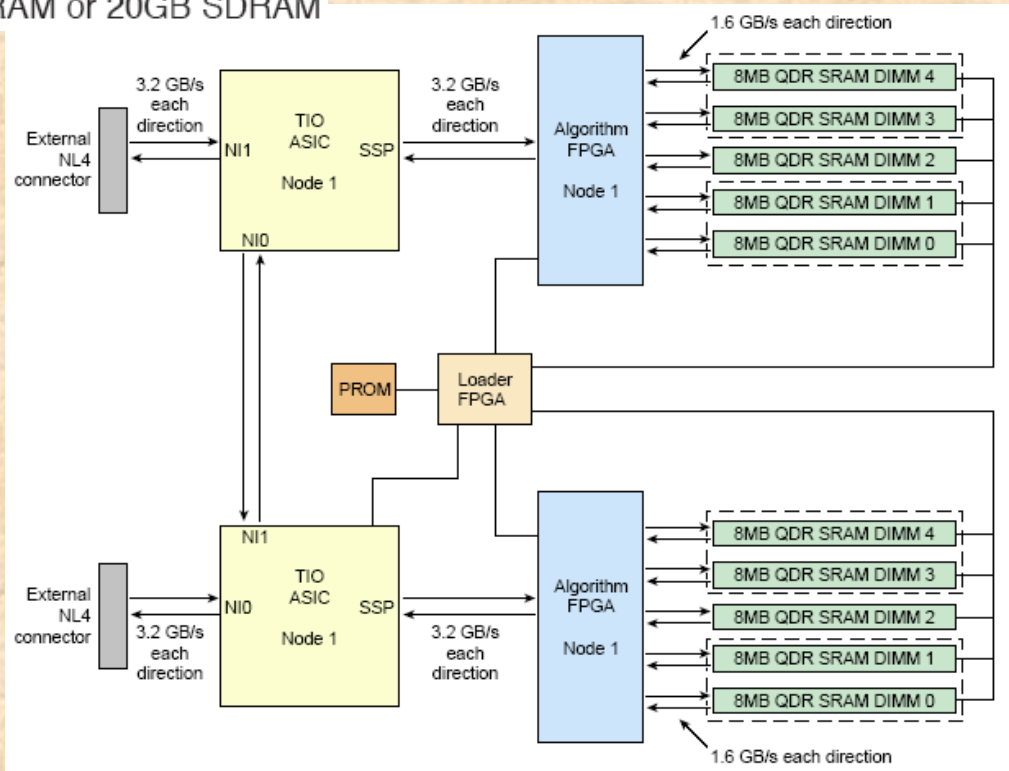


Kazimierz Wiatr, ACK CYFRONET AGH, Zespół Akceleracji Obliczeń

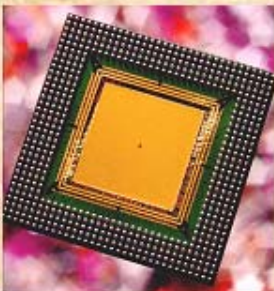
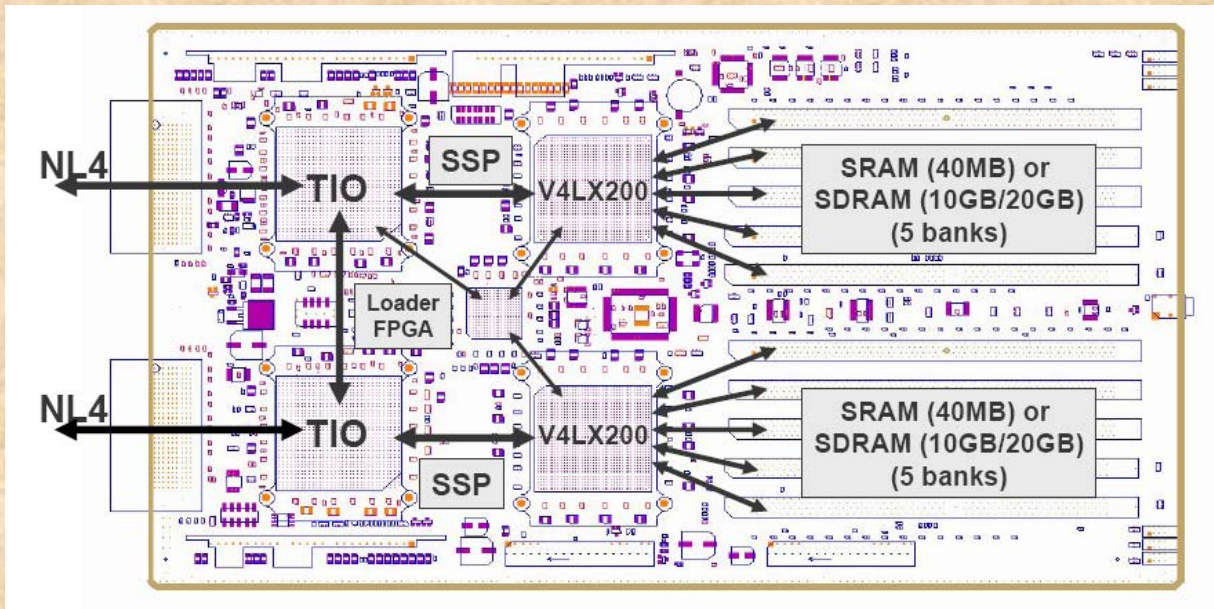
SGI RASC Blade

Pojedynczy moduł rekonfigurowany

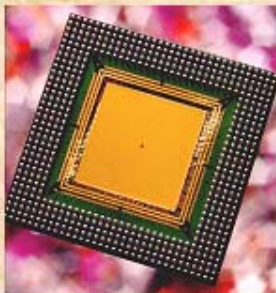
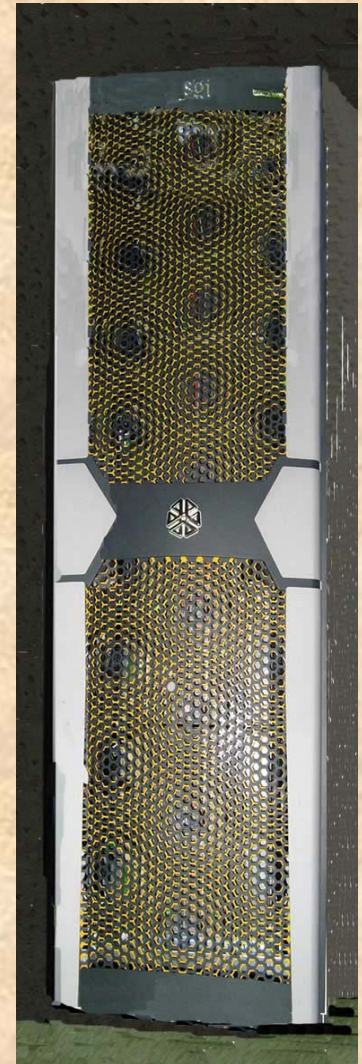
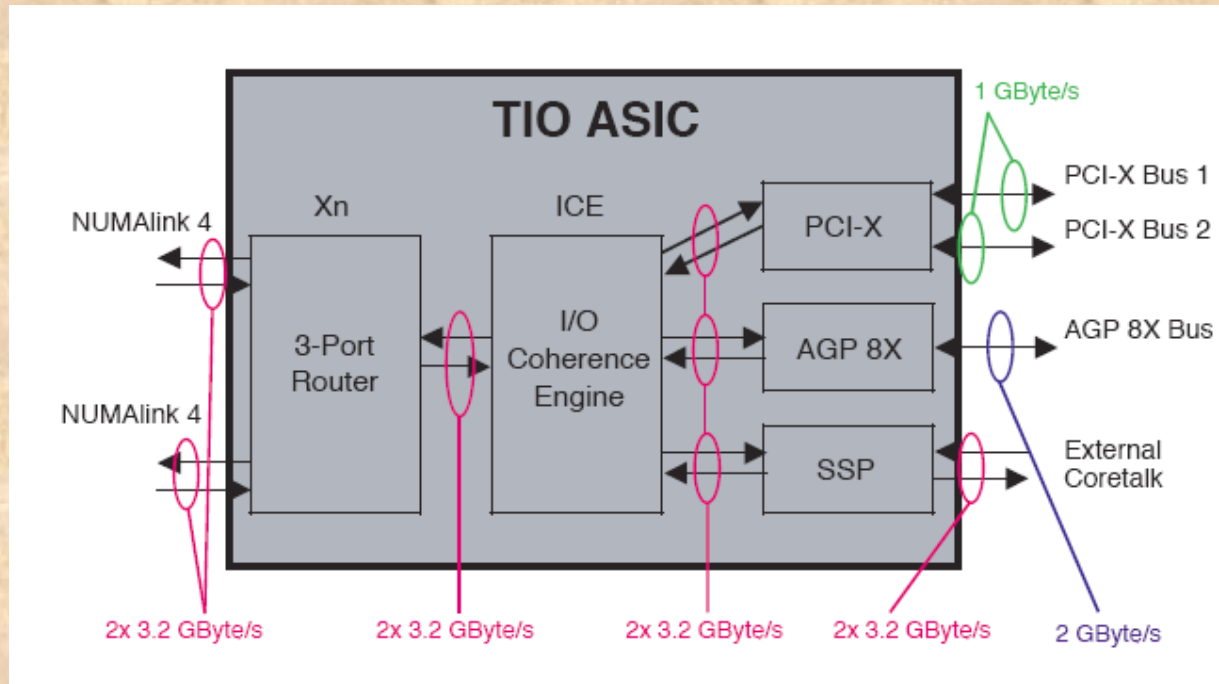
- Dual SGI NUMalink 4 ports
- Dual Xilinx Virtex 4 LX200 FPGAs
- 200,000 Logic Cells
- 80MB QDR SRAM or 20GB SDRAM



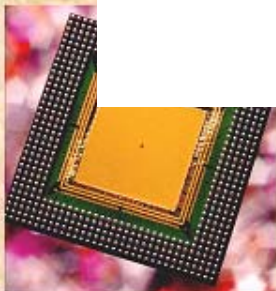
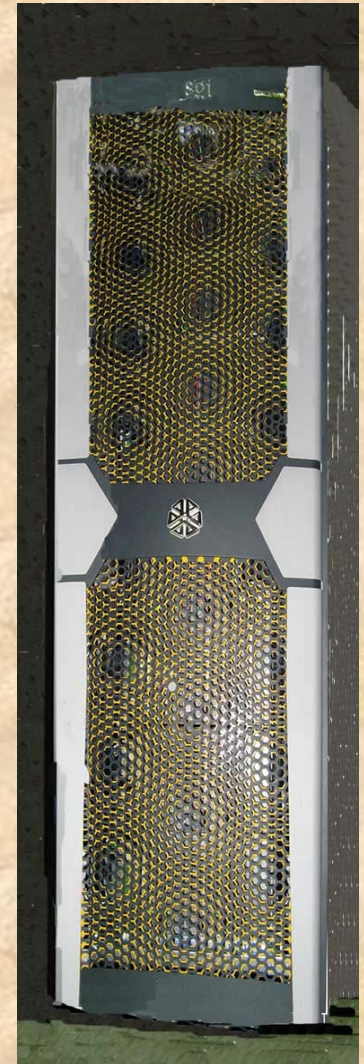
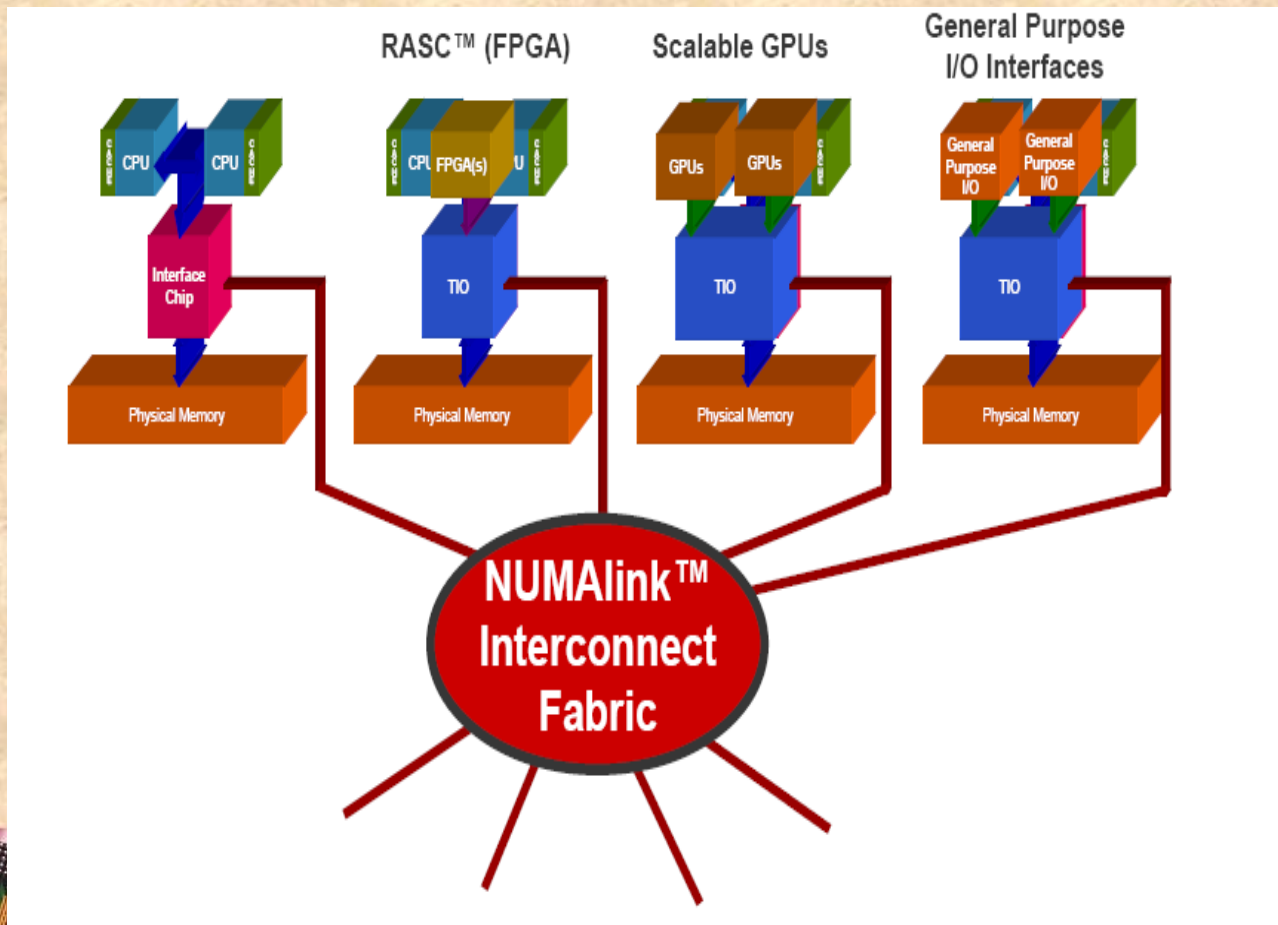
SGI RASC Blade



RASC Technology

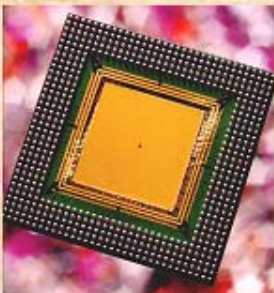
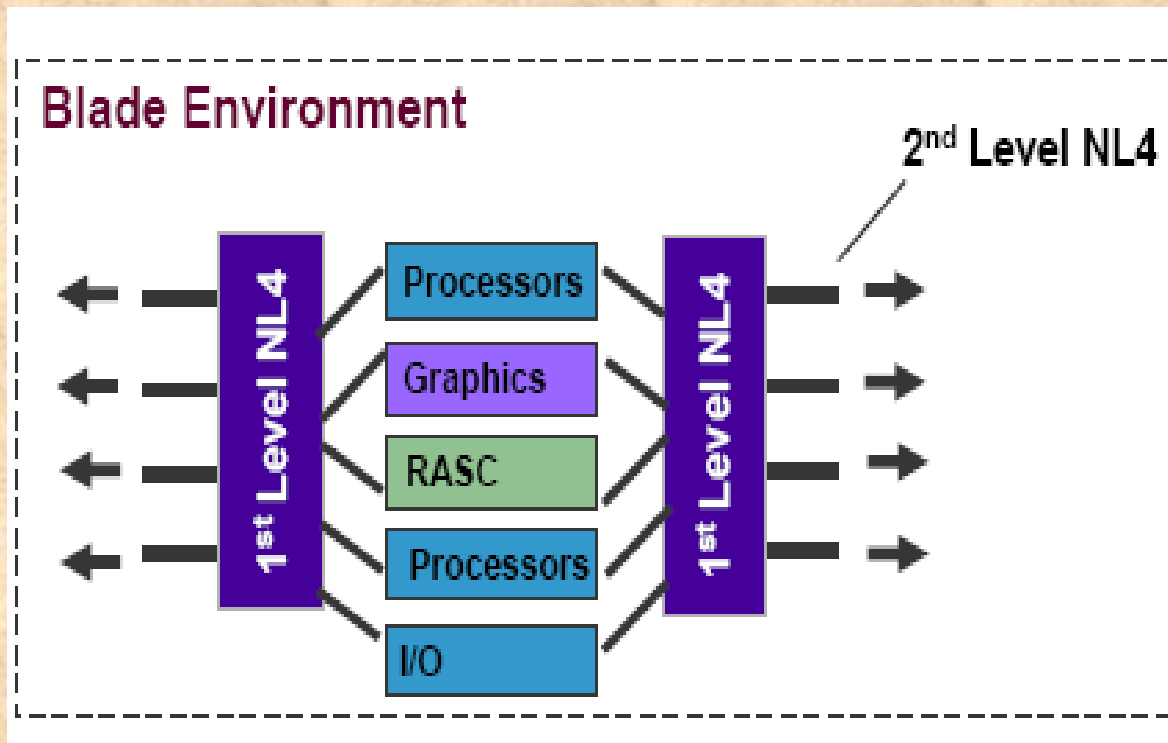


RASC Technology - NUMAlink



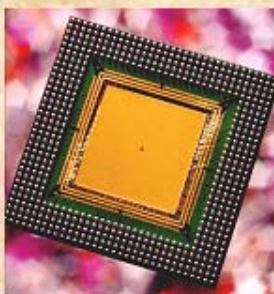
RASC Technology

Różne konfiguracje pracy



Zagadnienia realizowane w ramach projektu

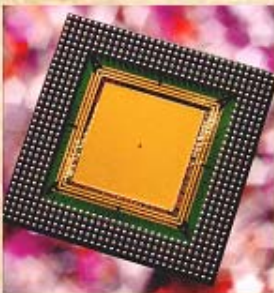
- Profiling
- Implementacja wytypowanych w profilingu funkcji
- Stworzenie własnego środowiska automatyzującego projektowanie oraz testowanie aplikacji HPTC



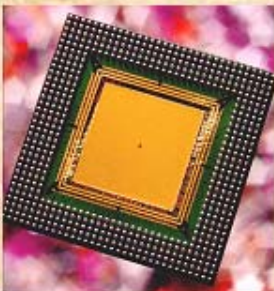
Prace badawczo-rozwojowe prowadzone obecnie przez Zespół Akceleracji Obliczeń

Współpraca z wydziałem Chemii Teoretycznej UJ

- Rozwiązywanie równania Schroedingera metodą Hartree-Focka
- Potrzeba zrównoleglenia obliczeń funkcji najbardziej czasochłonnych oraz najczęściej pojawiających się, takich jak:
 - Pierwiastek kwadratowy
 - Eksponenta
 - Logarytm
 - Funkcja power
- Argumenty funkcji oraz uzyskane wyniki są zgodne ze standardem IEEE 754 double



- **Dziękuję za uwagę**



Kazimierz Wiatr, ACK CYFRONET AGH, Zespół Akceleracji Obliczeń