

Título do Trabalho:
Arquitetura Raw

Ricardo M Nishihara

RA 936161

Razões para o sucesso dos ASICs na exploração de paralelismo

- **1. Especificação de operadores**
 - ASICs especializam operações da aplicação no nível de gate.
- **2. Maior disponibilização de recursos em paralelo.**
 - Aceleradores gráficos em ASICs: grande número de ops de baixa granularidade / ciclo
- **3. Gerenciamento dos *wire delays***
 - *Minimização de latências:*
placement adequado de operadores, canais de comunicação dedicados
- **4. Gerenciamento de pinos**
 - *Inexistência de gargalos devido a sistema de memória*
 - *E/S mais eficiente e aderente à aplicação para diferentes dispositivos: conversores A/D, CCD, array de sensores, etc*

Objetivo da Arquitetura Raw

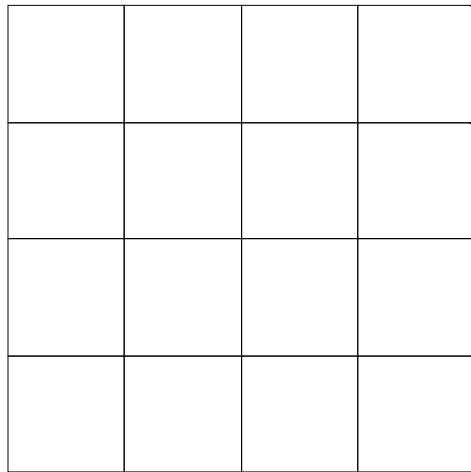
O objetivo principal do projeto Raw:

Arquitetura de propósito geral com bom desempenho:

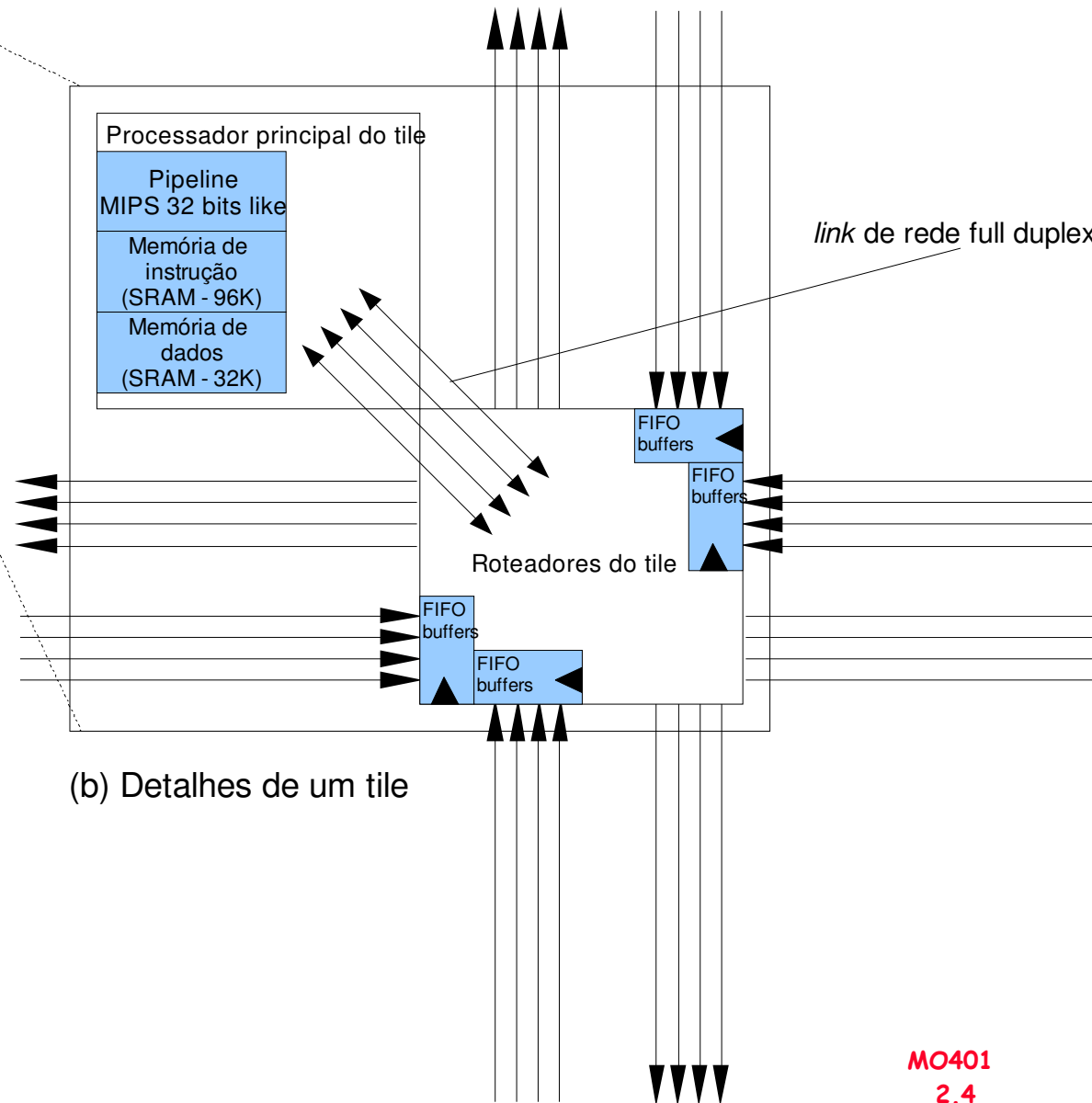
- aplicações sequenciais típicas (via exploração de ILP)
- aplicações “streaming” com alto grau de paralelismo (até aqui implementadas principalmente através de hardware dedicado).

Arquitetura Raw

Array Mesh 2D de tiles



(a) Processador Raw – conjunto de tiles idênticos



(b) Detalhes de um tile

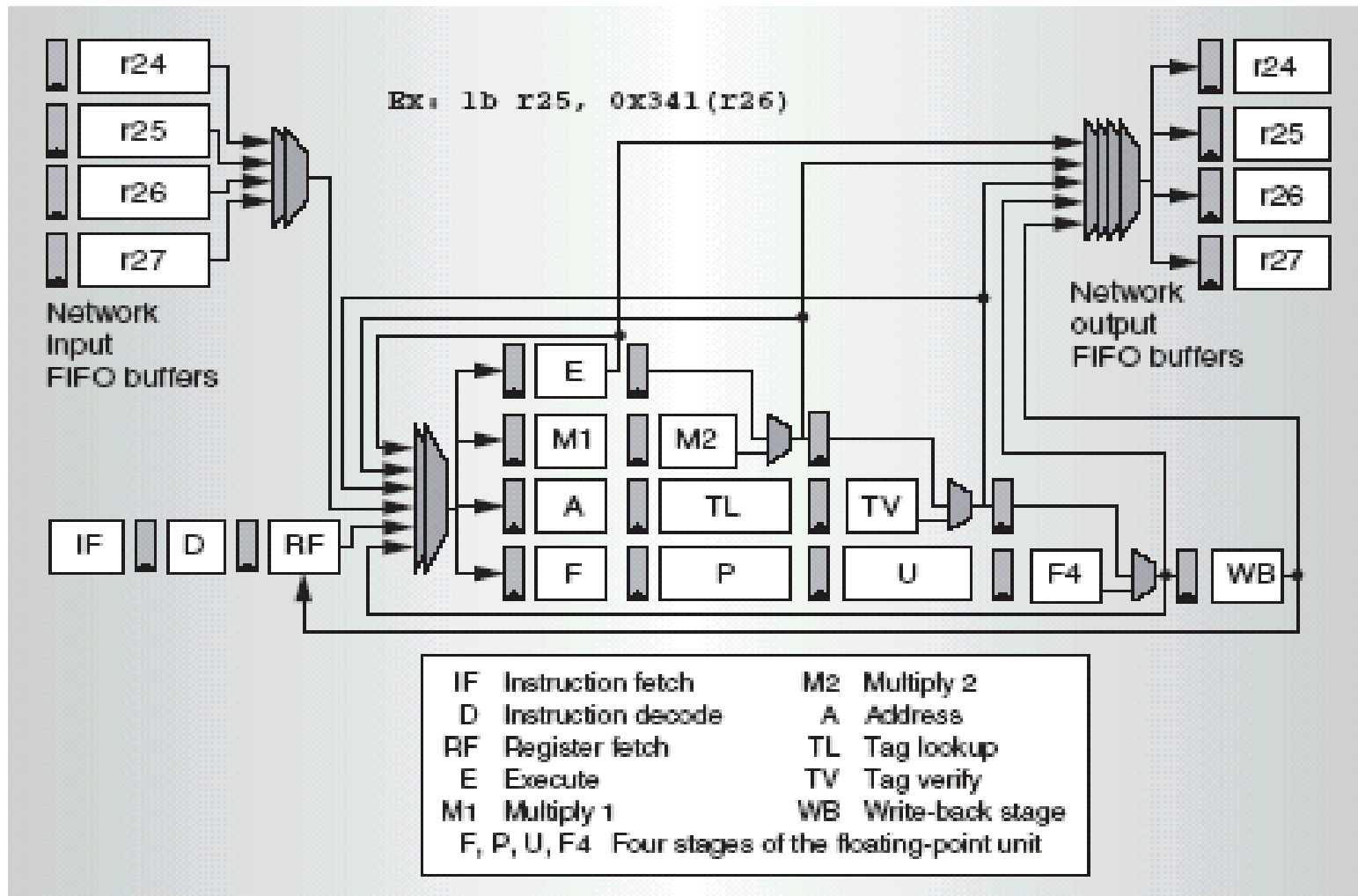
Como Raw atende os quatro fatores ?

- **1. Especialização:**
 - Implementação de operadores requeridos para a exploração de ILP e/ou paralelismo "streaming".
- **2. Disponibilidade de recursos em paralelo:**
 - replicar tais operadores, e expô-los ao software através do ISA.
- **3. Gerenciamento do wire delay:**
 - expondo ao software operadores relativos aos canais de comunicação interligando estas unidades.
- **4. Gerenciamento de pinos:**
 - As abstrações expostas pelo ISA referentes aos "pinos", permitem o gerenciamento via software de sistemas de memória cache e de interfaces E/S de alto desempenho.

Estrutura do tile

- Cada tile contém:
 - Um Processador *MIPS like* de 32 bits
 - » Pipeline de 8 estágios, in-order, single issue
 - » Unidade ponto-flutuante com pipeline de quatro estágios
 - 32 Kbytes data cache
 - 96 Kbytes memória de instruções (cache via software)
 - Um roteador estático programável
 - Dois roteadores dinâmicos programáveis
- Cada tile é conectado apenas a seus quatro tiles vizinhos mais próximos (norte, sul, leste e oeste)

Processador Pipeline do Tile



Exposição dos ports da rede de comunicação entre tiles

Registrador	Alias	Descrição
\$24	\$csti	Port de entrada para a rede estática
\$25	\$csgn[i/o]	Port de E/S para a rede dinâmica de uso geral
\$26	\$csti2	Segundo port de entrada para a rede estática
\$27	\$cmn[i/o]	Port de E/S para a rede dinâmica reservada à memória

Tabela 1 – Mapeamento de registradores aos ports das redes estáticas e dinâmicas

```
# XOR register 2 with 15,  
# and put result in register 31  
xori $31,$2,15  
# get two values from switch,  
# add to register 3, and put  
# result in register 9  
addu $9,$csti2,$csti  
# an ! indicates that the result  
# of the operation should also  
# be written to $csto  
and! $0,$3,$2  
# load from address at $csti+25  
# put value in register 9 AND  
# send it through $csto port  
# to static switch  
ld! $9,25($csti)  
# jump through value specified  
# by $csti2  
jr $csti2
```

Figura 3 – Trecho de código assembly ilustrando o acesso às interfaces com as redes para comunicação entre tiles

Redes Estáticas

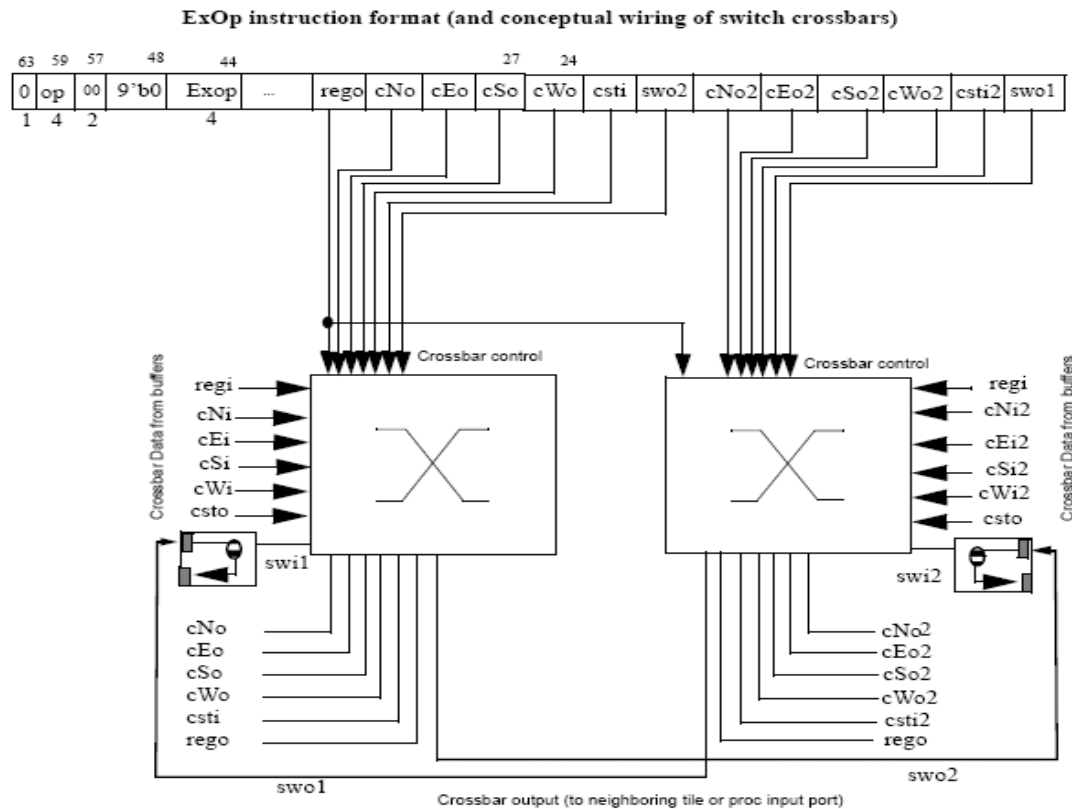
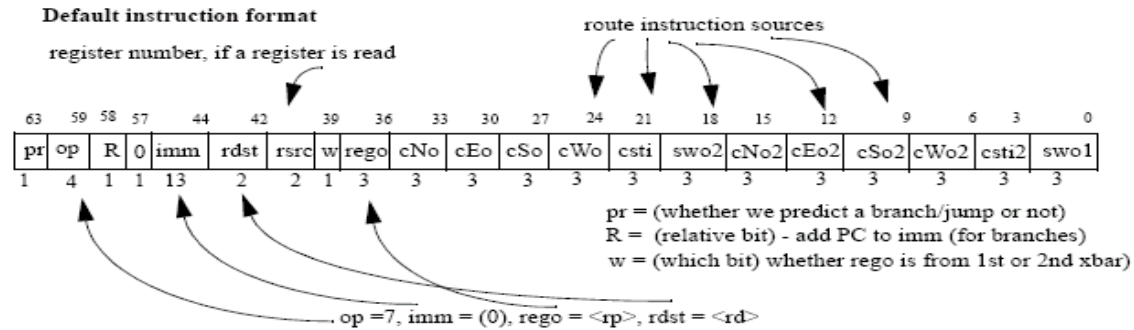
Roteador estático:

- **Processador pipeline de 5-estágios**
 - *instruções com palavras de 64 bits. Cada instrução: um pequeno comando (ex.: branch) + 13 rotas (uma para cada saída de crossbar)*
 - **Memória de instruções com 8096 palavras**
- **crossbar switchers:**
 - **o processador do roteador estático, os quatro tiles vizinhos, o processador principal do tile, e o outro crossbar**
- **Flow control:**
 - **O roteador estático prossegue para a próxima instrução somente depois que todas as rotas de uma dada instrução são completadas**
- **Redes de comunicação entre tiles**

Estática = rotas estabelecidas em tempo de compilação

Exemplo de instrução do roteador estático

MOVE \$2, \$3 ROUTE \$2->\$csti, \$2->\$cNo, \$2->\$cSo, ...



Redes dinâmicas

- Cada tile conta também com 2 roteadores dinâmicos: um par de redes dinâmicas para comunicação entre tiles.
- **Dinâmica = rotas estabelecidas em tempo de execução**
- **Mensagem:**
 - palavra de cabeçalho especificando:
 - » **o tile destino (ou port de E/S),**
 - » um campo de usuário e o
 - » **comprimento da mensagem.**
 - O usuário pode enviar até 31 palavras de dados em uma mensagem.

Redes dinâmicas - Deadlock

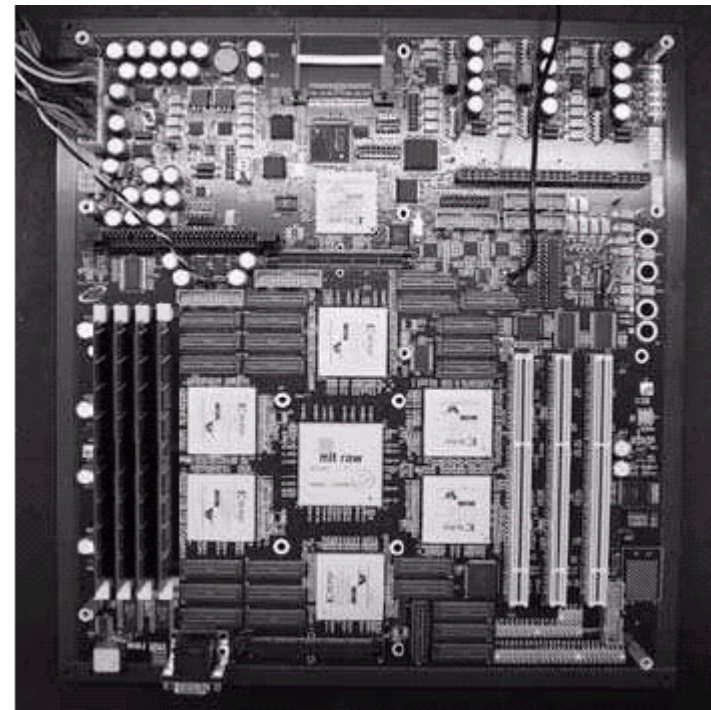
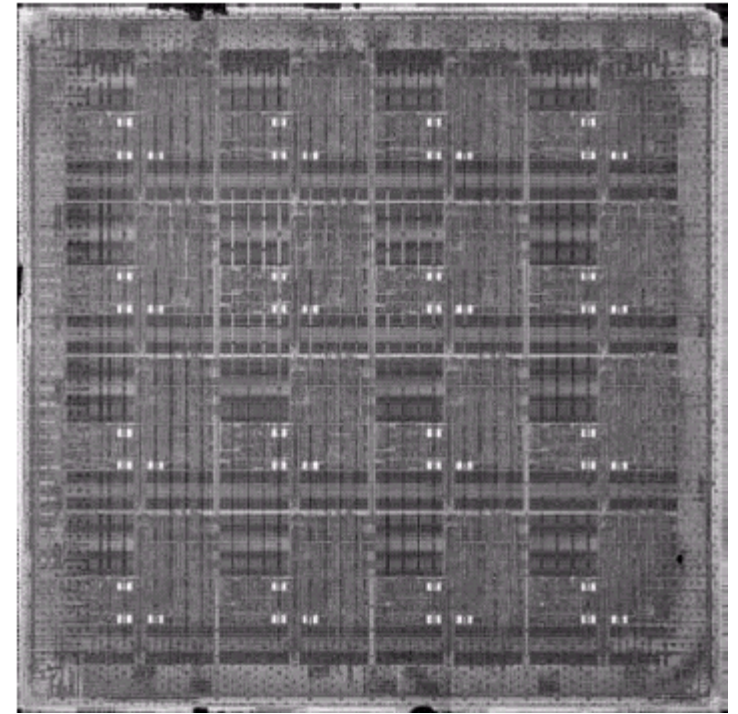
- Podem ocorrer deadlocks nos acessos aos buffers da rede (roteadores):
 - **Eliminação**
 - **Recuperação**
- **Solução Raw**
 - Uma rede é dedicada a acesso a memória (e E/S), e
 - » Eliminação de deadlock
 - » Apenas clientes c/ privilégio podem utilizá-la: SO, Interrupção, cache de software,...
 - A outra rede de uso geral e acesso irrestrito
 - » Usa recuperação de deadlock: contador no processador do tile + interrupções + rede dedicada à acesso a memória

Suporte de software - Compilador Raw

- **Alocação de recursos - analogia com placement dos ASICs**
 - programa dividido em regiões paralelas. Cada região associada a um conjunto de tiles (processador lógico único)
 - o número de tiles por região: paralelismo de granularidade mais fina nesta região
 - Comunicação entre tiles:
 - » Estática (dentro de uma região)
 - » Dinâmica: entre regiões
- **Exploração de paralelismo de granularidade fina**
 - Partição de um fluxo único de instruções em múltiplos fluxos de instruções
 - mapea cada fluxo para um tile, e
 - escalona a comunicação estática entre os fluxos.
- **Escalonamento de comunicação**

Raw Chip

- Array de 16 tiles
- Processo ASIC da IBM SA-27E (0.15-micron, 6 níveis, cobre).
- Die: 18.2 x 18.2 mm
- encapsulamento CCGA de 1657 pinos, com 1080 pinos de E/S HSTL
- O chip consome em média 18.2 watts
425MHz/500Mhz c/
1.8V/2.2V.



Comparação Raw x Pentium 3

- Comparação direta com Pentium 3 não esconde ineficiências do compilador e da arquitetura.
- Facilita a comparação indireta da arquitetura Raw com outras alternativas desenvolvidas

Parameter	Raw (IBM ASIC)	P3 (Intel)
Lithography Generation	180 nm	180 nm
Process Name	CMOS 7SF (SA-27E)	P858
Metal Layers	Cu 6	Al 6
Dielectric Material	SiO ₂	SiOF
Oxide Thickness (T _{ox})	3.5 nm	3.0 nm
SRAM Cell Size	4.8 μm ²	5.6 μm ²
Dielectric k	4.1	3.55
Ring Oscillator Stage (FO1)	23 ps	11 ps
Dynamic Logic, Custom Macros (SRAMs, RFs)	no	yes
Speedpath Tuning since First Silicon	no	yes
Initial Frequency	425 MHz	500-733 MHz
Die Area ²	331 mm ²	106 mm ²
Signal Pins	~ 1100	~ 190
Vdd used	1.8 V	1.65 V
Nominal Process Vdd	1.8 V	1.5 V

- Pentium 3 era a implementação de processador Intel mais próxima:
 - > processo com a mesma geração de litografia 180 nm
 - > latências associadas às unidades funcionais muito próximas

Resultados

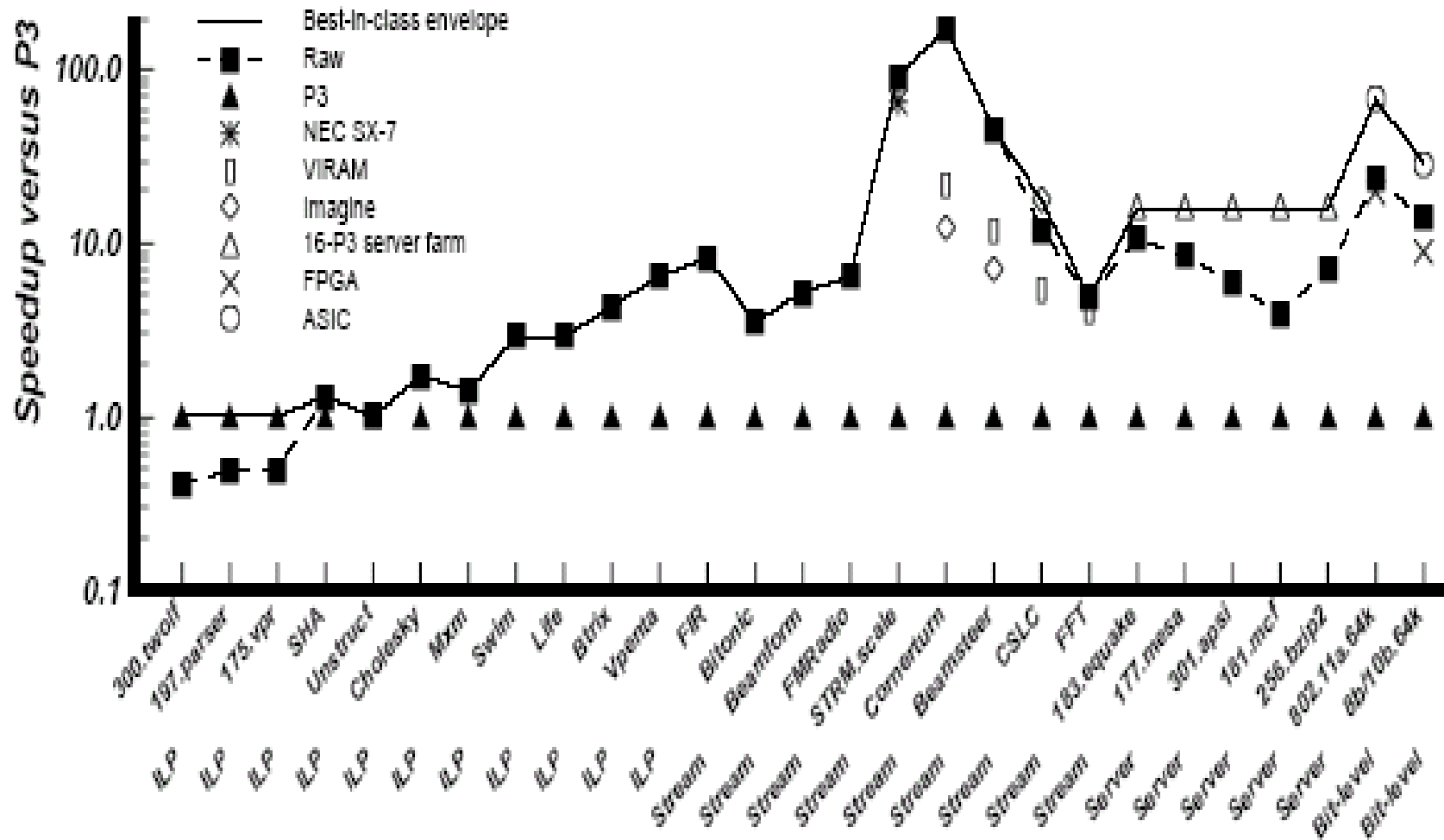


Figura 5 – Desempenho comparativo para várias aplicações

Razões para o speedup

Classe	Benchmark	S	R	W	P
ILP	Swim, Tomcatv, Btrix, Cholesky, Vpenta, Mxm, Life, Jacobi, Fpppp-kernel, SHA, AES, Encode, Unstructured, 172.mgrid, 173.aplu, 177.mesa, 183.equake, 188.ammp, 301.apsi, 175.vpr, 181.mcf, 197.parser, 256.bzip2, 300.rwolf	X	X	X	
Stream	Beamformer, Bitonic Sort, FFT, Filterbank, FIR, FMRadio, Mxm, LU fact., Triang. solver, QR fact., Conv., Copy, Scale, Add, Scale & Add, Acoustic Beamforming, FIR, FFT, Beam Steering, Corner Turn, CSLC	X	X	X	X
Server	172.mgrid, 173.aplu, 177.mesa, 183.equake, 188.ammp, 301.apsi, 175.vpr, 181.mcf, 197.parser, 256.bzip2, 300.rwolf		X		X
Bit-level	802.11a ConvEnc, 8b/10b Encoder	X	X	X	

Utilização de funcionalidades Raw:

S = Especialização

R = Utilização de Recursor Paralelos

W = Gerenciamento dos atrasos dos condutores

P = Gerenciamento de Pinos