

## A

- abstrações
  - definição, 13
  - interface de hardware/software, 13–15
  - princípio, 15
- acerto sob perda, 436
- acertos de cache, 409
- acrônimos, 5
- add (Add), 61
- adição, 181–185
  - binária, 181–182
  - instruções, B-667
  - operandos, 182
  - ponto flutuante, 202–205, 209, B-684–B-685
  - significandos, 202
  - velocidade, 185
  - Ver também* aritmética
- adição de ponto flutuante, 202–205
  - binária, 203, 204
  - diagrama de blocos de unidade aritmética, 205
  - etapas, 202–203
  - ilustração, 203
  - instruções, 209, B-684–B-685
  - testando a associatividade, 219
- adição sem sinal, instrução, 182
- Advanced Technology Attachment (ATA), discos, 465, 494, 495
- AGP, A-574
- algoritmo de divisão, 193
- algoritmo de multiplicação, 189
- algoritmos de classificação, 127
- aliasing, 409
- All-pairs N-body, algoritmo, A-620
- alocação de espaço
  - na heap, 94–97
  - na pilha, 94
- Alpha, arquitetura
  - definição, 424
- ALU, bloco de controle, 257
- ALU, controle, 254–256
  - bits, 255
  - Ver também* Arithmetic Logic Unit (ALU)
- ALUOp, 254
  - bits, 255, 256
  - sinal de controle, 257
- AMD Opteron X4 (Barcelona), 13, 34–39, 242
  - benchmark de CPU SPEC, 36–38
  - benchmark de potência SPEC, 38–39
- cache L3 compartilhado, 437
- caches, 436
- características, 546
- CPI, taxas de falta e acessos à DRAM, 437
- definição, 546
- desempenho básico versus totalmente otimizado, 551
- desempenho LBMHD, 551
- desempenho SpMV, 550
- hardware TLB, 435
- hierarquias de memória, 435–437
- ilustração de pipeline, 327
- ilustração, 547
- microarquitetura, 325, 326
- modelo rooﬂine, 548
- pipeline, 325–327
- registradores arquiteturais, 325
- técnicas de redução de penalidade de falta, 436–437
- tradução de endereço, 435
- AMD64, 136
- American Standard Code for Information Interchange. *Ver* ASCII
- AND, operação, 81–82, B-668
- Annual Failure Rate (AFR), 462, 494
- antidependência, 320
- Application Binary Interface (ABI), 15
- Application Programming Interfaces (APIs)
  - definição, A-570
  - gráficos, A-578
- Arithmetic Logic Unit (ALU)
  - antes do forwarding, 296
  - caminho de dados de desvio, 252
  - entrada imediata com sinal, 298
  - hardware, 182
  - operações em formato R, 250
  - para valores de registrador, 248
  - uso de instrução de referência de memória, 243
  - Ver também* ALU controle; unidades de controle
- aritmética, 179–228
  - adição, 181–185
  - divisão, 191–197
  - multiplicação, 186–191
  - para multimídia, 183–184
  - ponto flutuante, 197–219
  - subtração, 181–185
- aritmética de multimídia, 183–184
- aritmética de ponto flutuante (GPUs), A-601
  - básica, A-601
  - desempenho, A-603
  - especializada, A-601–A-603
  - formatos admitidos, A-601
  - operações de textura, A-603
  - precisão dupla, A-604, A-627
- ARM, instruções, 130–133
  - cálculos, 130–132
  - campo de condição, 308
  - campo imediato de 12 bits, 132
  - características, 132–133
  - compare e desvio condicional, 132
  - formatos, 132
  - loads e stores em bloco, 133
  - lógicas, 133
  - modos de endereçamento, 130–132
  - registrador-registrador, 131
  - semelhanças do MIPS, 131
  - transferência de dados, 131
- armadilhas
  - acessos de disco pelo sistema operacional, 496–497
  - associatividade, 439
  - avaliação de processador fora de ordem, 439
  - backups de fita magnética, 496
  - definição, 39
  - desempenho da taxa de transferência de pico, 497
  - desenvolvimento de software com multiprocessadores, 553
  - endereços de word sequenciais, 142
  - extensão do espaço de endereços, 439
  - GPUs, A-627–A-628
  - hierarquias de memória, 437–441
  - ignorando o comportamento do sistema de memória, 438
  - implementação de VMM, 439–441
  - movendo funções para processador de E/S, 496
  - pipelining, 328–329
  - ponteiro para variáveis automáticas, 142
  - provisão de recurso da rede, 495–496
  - simulando cache, 437–438
  - subconjunto de equação de desempenho, 40–41
  - Ver também* falácias
- armazenamento
  - disco, 464–467
  - flash, 468–469
  - não volátil, 464

- armazenamento de disco, 464–467
    - características, 467
    - densidades, 465
    - interfaces, 465–467
    - latência de rotação, 465
    - não volátil, 464
    - setores, 464
    - tempo de busca, 464
    - tempo de transferência, 465
    - trilhas, 464
  - armazenamento não volátil, 464
  - arquitetura de GPU unificada, A-575–A-576
    - array de processador, A-575–A-576
    - ilustração, A-575
  - arquitetura de multiprocessador
    - multithreaded, A-587
    - comparação de multiprocessador, A-596
    - conclusão, A-596
    - gerenciamento de threads/blocos de threads, A-591
    - instruções de thread, A-592
    - ISA, A-592–A-595
    - multiprocessador, A-588–A-589
    - multithreading massivo, A-587–A-588
    - processador streaming (SP), A-595
    - SIMT, 589tunidades de função especial (SFUs), A-596
  - arquitetura do conjunto de instruções
    - ARM, 130–133
    - cálculo de endereço de desvio, 250
    - definição, 15, 41
    - história, 145
    - mantendo, 41
    - proteção, 425–426
    - suporte para máquina virtual, 424–425
    - thread, A-592–A-595
  - arquivos de registradores
    - definição, 248
    - únicos, 253
  - arquivos executáveis, B-630
    - definição, 113
    - produção do link-editor, B-643
  - arquivos fonte, B-630
  - arquivos-objeto, 113, B-630
    - cabeçalho, 113, B-638
    - definição, B-636
    - formato, B-638–B-639
    - informação de depuração, 113
    - informação de relocação, 113
    - link-edição, 114–116
    - segmento de dados estático, 113
    - segmento de texto, 113
    - tabela de símbolos, 113
  - arrays
    - dimensão múltipla, 215
    - ponteiros versus, 127–130
    - procedimentos para definir como zero, 128
  - arredondamento
    - bits, 217
    - com dígitos de guarda, 216
    - definição, 215
    - modos IEEE, 754, 217
    - preciso, 215
  - ASCII
    - definição, 97
    - número binários versus, 98
    - representação de caractere, 97
    - símbolos, 100
  - Assembly, linguagem
    - definição, 7, 111
    - desvantagens, B-635–B-636
    - ilustração, 8
    - linguagens de alto nível *versus*, B-637
    - MIPS, 61, 78–79, B-663–B-690
    - ponto flutuante, 210
    - produção, B-634–B-635
    - programas, 111
    - quando utilizar, B-634–B-635
    - traduzindo para linguagem de máquina, 78–79
  - associatividade
    - aumentando o grau, 387, 417
    - aumentando, 391–392
    - conjunto, tamanho de tag versus, 391–392
    - em caches, 387–389
    - teste de adição de ponto flutuante, 219
  - atalho de negação, 72–73
  - atalho de verificação de limites, 87
  - atraso no pior caso, 265
  - Average Memory Access Time (AMAT), 385
    - calculando, 385
    - definição, 385
- B**
- backpatching, B-638
  - backups, 496
  - balanceamento de carga, 514–515
  - bancos de registradores, 248, 253
  - barramento backplane, 469
  - barramento processador-memória, 469
  - barramento síncrono, 470
  - barramentos, 471, 472
    - backplane, 469
    - processador-memória, 469
    - síncronos, 470
  - benchmark de servidor de arquivos (SPECFS), 481
  - benchmark de servidor Web (SPECWeb), 481
  - benchmarks
    - definição, 36
    - E/S, 480–482
    - Linpack, 537
    - multicóres, 531–552
    - multiprocessador, 537–539
    - NAS paralelo, 539
    - paralelos, 538
    - PARSEC, 539
    - SPEC CPU, 36–38
    - SPEC power, 38–39
    - SPECrate, 537
    - SPLASH/SPLASH, 2, 537–539
    - Stream, 546
  - bibliotecas de programas, B-630
  - big-endian, ordem de byte, 66, B-661
  - bit de erro, 475
  - bit de modificação, 404
  - bit de pronto, 475
  - bit de referência, 402
  - bit de sinal, 71
  - bit de validade, 369
  - bit mais significativo
    - definição, 69
  - bits
    - ALUOp, 255, 256
    - arredondamento, 217
    - definição, 7
    - erro, 475
    - guarda, 215–216
    - modificados, 404
    - padrões, 217
    - prontos, 475
    - referência, 402
    - sinal, 71
    - sticky, 217
    - válidos, 369
  - bits menos significativos
    - definição, 70
  - bits sticky, 217
  - bloco básico, 85–86
  - blocos
    - dados válidos, 369
    - definição, 366
    - encontrando, 418–419
    - estratégias de posicionamento, 387
    - estratégias de substituição, 419–420
    - exploração de localidade espacial, 374
    - loads/stores, 133
    - locais de posicionamento, 417–418
    - localizando na cache, 389–391
    - multiword, mapeando endereços para, 373–374
    - posicionamento flexível, 385–389
    - seleção de substituição, 391
    - taxa de perda, 374
    - usados menos recentemente (LRU), 391
  - blocos de thread, 533
    - compartilhamento de memória, A-584
    - criação, A-585
    - definição, A-582
    - gerenciando, A-591
    - sincronização, A-584
  - bolhas, 301
  - branch delay, slots
    - definição, 306
    - escalonamento, 308
  - branch-on-equal, instrução, 262
  - Bubble Sort, 126
  - buffer de frame, 12
  - buffer de renovação de rastreamento, 12
  - buffers de armazenamento, 325
  - buffers de escrita
    - cache write-back, 377
    - definição, 376
    - stalls, 383
  - buffers de reordenação, 321, 324, 325
  - busca-e-incremento atômico, 111
  - bytes
    - endereçamento, 66
    - ordem, 66, B-661

## C

### C, linguagem

algoritmos de classificação, 127  
atribuição, compilando no MIPS, 62–63  
compilando loops while na, 84–85  
compilando, 130  
hierarquia de tradução, 112  
tarefa de compilação com registradores, 64–65  
tradução para linguagem assembly do MIPS, 62  
variáveis, 93

### caches, 368–382

acessando, 370–374  
associativas em conjunto, 385  
associatividade, 387–389  
bits necessários, 370  
bits, 373  
campo de tag, 370  
controlador de disco, 467  
definição, 13, 368  
divisão, 378  
escritas, 375–377  
esvaziando, 479  
fisicamente endereçadas, 409  
fisicamente indexadas, 408  
fisicamente marcadas, 408  
FSM para controlar, 426–434  
GPU, A-598  
ilustração do conteúdo, 371  
inconsistentes, 375  
índice, 370  
Intrinsity FastMATH, exemplo, 377–378  
locais, 369  
localizando blocos, 389–391  
mapeadas diretamente, 368, 370, 373, 385  
memória virtual e integração de TLB, 406–409  
multinível, 382, 392–395  
não bloqueantes, 436  
primárias, 393, 396  
projeto de sistema de memória, 379–382  
resumo, 382  
secundárias, 393, 396  
simulando, 437–438  
tamanho, 372  
totalmente associativas, 385  
vazias, 370  
virtualmente endereçadas, 409  
virtualmente indexadas, 409  
virtualmente marcadas, 409  
write-back, 376, 377, 420  
write-through, 376, 377, 420  
*Ver também* blocos

### caches associativos em conjunto, 385–386

definição, 385  
duas vias, 387  
escolha, 419  
estratégias de substituição de bloco, 420  
falhas, 387–389  
local de bloco da memória, 386  
*n*-vias, 385  
partes de endereço, 389

quádruplas, 387, 391

*Ver também* caches

### caches de mapeamento direto

comparador único, 391  
definição, 368, 385  
escolha, 419  
faltas, 387  
ilustração, 370  
local de bloco de memória, 386  
número total de bits, 373  
partes de endereço, 389  
*Ver também* caches

### caches divididas, 378

### caches endereçadas fisicamente, 409

### caches multinível

complicações, 393  
definição, 382, 393  
desempenho, 392–393  
penalidade de falta, reduzindo, 392–395  
resumo, 395–396  
*Ver também* caches

### caches sem bloqueio, 325, 436

### caches totalmente associativas

definição, 385  
escolha, 419  
estratégias de substituição em bloco, 420  
local de bloco da memória, 386  
perdas, 389  
*Ver também* caches

### caches virtualmente endereçados, 409

### caches write-back

buffers de escrita, 377  
complexidade, 377  
definição, 376, 420  
stalls, 383  
vantagens, 420  
*Ver também* caches

### caixas de pizza, 489

### cálculo de execução/endereço

instrução load, 281  
instrução store, 283  
linha de controle, 288

### callee, 89, 91

### caller, 89

### caminho de dados de desvio

ALU, 252  
operações, 251

### caminhos de dados

definição, 12  
desvio, 251, 252  
dois despachos estáticos, 318  
em operação para instrução  
branch-on-equal, 262  
em operação para instrução load, 261  
em operação para instrução tipo R, 260  
montagem, 247–254  
operação, 258–262  
para arquitetura MIPS, 254  
para instrução de salto, 264  
para instruções de busca, 249  
para instruções de memória, 253  
para instruções tipo R, 253, 260  
para resolução de hazard via forwarding, 297

pipeline, 276–288

projeto, 247

tratamento de exceção, 311

único ciclo, 277

únicos, criando, 252–254

unidade de controle, 259

### caminhos de dados de ciclo único

execução de instrução, 278

ilustração, 277

*Ver também* caminhos de dados

### caminhos de dados em pipeline, 276–288

com sinais de controle conectados, 291

com sinais de controle, 288

corrigidos, 285

ilustração, 279

nos estágios da instrução load, 285

### campo de condição, 308

### campos

definição, 75

MIPS, 76–77

nomes, 77

registrador Cause, B-654, B-655

registrador de Status, B-654, B-655

### caracteres

em Java, 100–101

representação ASCII, 97

### carga, B-643

### carregadores, 116

### cartões de memória removíveis baseados em

flash, 16

### caso comum rápido, 143

### Central Processor Unit (CPU)

co-processor 0, B-654

definição, 12

desempenho, 22–23

equação de desempenho clássica, 26–28

medições de tempo, 23

tempo de execução, 22, 23, 24

tempo do sistema, 22

tempo do usuário, 22

tempo, 382

*Ver também* processadores

### centros de dados, 2

### Cg, programa shader de pixel, A-579–A-580

### chamadas de procedimento

convenção, B-645–B-653

exemplos, B-647–B-653

frame, B-646

preservação por, 93

### chamadas do sistema, B-661–B-663

código, B-662

definição, 410

carga, B-661

### chips. *Ver* circuitos integrados (ICs)

### ciclos de clock

atraso no pior caso, 265

definição, 23

número de registradores, 64

stall de memória, 382, 383

### ciclos de clock de stall da memória, 382, 383

### ciclos de clock por instrução (CPI), 25–26, 274

dois níveis de caching, 393

um nível de caching, 393

ciclos de stall de escrita, 383  
 ciclos de stall de leitura, 383  
 circuitos integrados (ICs)  
   custo, 35  
   definição, 19  
   processo de manufatura, 34  
   *Ver também chips específicos*  
 clusters  
   definição, 510, 517  
   desvantagens, 518  
   isolamento, 520  
   organização, 509  
   overhead na divisão da memória, 518  
 co-processadores  
   co-processador 0, B-654  
   definição, 215  
   instruções move, B-682–B-683  
 codificação  
   instrução de ponto flutuante, 211  
   instrução MIPS, 78, 107, B-667  
   instrução x86, 141  
 código de correção, 485  
 código de função, 77  
 código de máquina, 75  
 coerência de cache, 430–433  
   coerência, 430  
   consistência, 431  
   esquemas de imposição, 432  
   migração, 432  
   problema, 430, 431, 433  
   protocolo snooping, 432–433  
   protocolos, 432  
   replicação, 432  
 comandos para dispositivos de E/S,  
   475–475  
 commit na ordem, 322  
 Compact Disks (CDs), 16, 17  
 compactação de estrutura de dados, 549  
 comparações, 85–86  
   com sinal *versus* sem sinal, 87  
   operandos constantes, 86  
 compare e swap atômicos, 111  
 compartilhamento falso, 433  
 compilação  
   C, instruções de atribuição, 62–63  
   C, linguagem, 84–85, 130  
   if-then-else, 83  
   loops while, 84–85  
   procedimentos recursivos, 92–93  
   procedimentos, 89, 92–93  
   programas de ponto flutuante, 209–214  
 compiladores, 111  
   criação de desvio, 84  
   definição, 7  
   especulação, 316–317  
   função, 9, 111, B-632  
   Just In Time (JIT), 119  
   otimização, 130  
   produção de linguagem de máquina,  
     B-634–B-635, B-636  
 complemento a um, 74  
 Compressed Sparse Row (CSR), matriz, A-612,  
   A-613

computação GPU  
   aplicações visuais, A-571–A-572  
   definição, A-570  
   *Ver também* Graphics Processing Units (GPUs)  
 computação visual, A-569  
 computadores  
   aplicações, 2  
   aritmética, 179–228  
   classes de aplicação, 2–4  
   componentes, 9, 180, 458  
   desktop, 2, 10  
   embutidos, 2–4, B-634  
   laptop, 12  
   medida de projeto, 42  
   montagem em rack, 488  
   organização de componente, 9  
   princípios, 79  
   representação da instrução, 74–80  
   revolução da informação, 2  
   servidores, 2  
 computadores desktop  
   definição, 2  
   ilustração, 11  
 computadores embutidos  
   definição, B-634  
   design, 5  
   requisitos da aplicação, 4  
 computadores laptop, 12  
 Compute Unified Device Architecture. *Ver*  
   CUDA, ambiente de programação  
 comunicação, 17–18  
   reduzindo o overhead, 33  
   thread, A-595  
 conceito de programa armazenado, 60  
   ilustração, 80  
   princípio do computador, 79  
   princípios, 143  
 confiabilidade, 462  
 conjunto de trabalho, 416  
 conjuntos de instruções  
   ARM, 308  
   MIPS-32, 227  
   MIPS, 60, 144, 225  
   NVIDIA GeForce, 8800, A-607  
   projeto para pipelining, 269  
   Pseudo MIPS, 227  
   x86, crescimento, 143  
 consoles de jogo, A-574  
 contador de instruções, 26, 27  
 controladores de programa (PCs), 247  
   atualizações de instrução, 279  
   definição, 89, 247  
   exceção, 410, 412  
   incrementando, 247, 249  
   mudando com desvio condicional, 308  
 Content Addressable Memory (CAM), 391  
 controladores de cache, 433  
 controladores de canal, 478  
 controladores de disco  
   caches, 467  
   definição, 465  
   tempo, 465

controle  
   ALU, 254–256  
   desafio, 309  
   forwarding, 294  
   para instrução de salto, 264  
   pipeline, 288–292  
   terminando, 262  
 controle em pipeline, 288–292  
   especificando, 288  
   ilustração do esboço, 302  
   linhas de controle, 288, 289  
   *Ver também* controle  
 Cooperative Thread Arrays (CTAs), A-591  
 copy back. *Ver* write-back  
 cores  
   definição, 31  
   número por chip, 32  
 CUDA, ambiente de programação, 533, A-571  
   abstrações de chave, A-582  
   definição, A-571  
   desenvolvimento, A-580, A-582  
   hierarquias de grupos de threads, A-582  
   implementação plus-reduction, A-619  
   kernels, A-582, A-586  
   memória compartilhada por bloco, A-615  
   memórias compartilhadas, A-582  
   paradigma, A-582–A-585  
   programação paralela escalável com,  
     A-580–A-585  
   programas, A-571, A-586  
   SDK, 141  
   sincronização de barreira, A-582, A-595  
   template paralela plus-scan, A-617  
   threads, A-596

## D

dados estáticos  
   dados dinâmicos, B-644  
   definição, B-644  
   segmento, 94  
 data race, 109  
 decisão adiada, 275  
 decodificando linguagem de máquina, 106  
 decomposição do problema paralelo de dados,  
   A-580, A-582  
 dependência de nome, 320  
 dependências  
   detecção, 293  
   entre registradores de pipeline e entradas  
     da ALU, 294  
   entre registradores de pipeline, 295  
   inserção de bolha, 301  
   nome, 320  
   sequência, 292  
 dependências em pipeline, 293  
 desafio de speed-up, 512–515  
   balanceando a carga, 514–515  
   problema maior, 513–514  
 desdobramento de loop  
   definição, 320  
   para pipelines de despacho múltiplo, 320  
   renomeação de registrador, 320

- desempenho, 19–29
    - armazenando, A-611–A-612
    - avaliando, 19–20
    - componentes, 28
    - CPU, 22–23
    - definição, 20–22
    - equação clássica da CPU, 26–28
    - instrução, 25–26
    - medição de tempo, 22
    - medindo, 22–23
    - melhorando, 23–25
    - programa, 28
    - razão, 22
    - relativo, 21
    - tempo de resposta, 20, 21
    - throughput, 20
    - usando equação, 26
  - desempenho da cache, 382–396
    - calculando, 384
    - impacto sobre desempenho do processador, 383–384
    - tempo de acerto, 385
  - desempenho de classificação, A-611–A-612
  - desempenho do programa
    - elementos afetando, 28
    - entendendo, 77
  - desempenho máximo de ponto flutuante, 540
  - desempenho relativo, 21
  - despacho de threads, 533
  - despacho múltiplo, 314–322
    - definição, 314
    - desdobramento de loop, 320
    - dinâmico, 316, 320–322
    - escalonamento de código, 319
    - estático, 316, 317–320
    - pacotes de despacho, 317
    - processadores, 314, 316
    - vazão, 323
  - destino do desvio
    - buffers, 308
    - endereços, 250
  - desvio igual, 302
  - desvio não tomado
    - definição, 251
    - suposição, 302
  - desvio tomado
    - definição, 251
    - redução de custo, 302
  - desvios
    - adiados, 87, 252, 275, 302–305, 306, 308
    - condição, 252
    - criação de compilador, 84
    - decisão, subindo, 302
    - endereçamento, 103–105
    - endereço de destino, 304
    - execução no estágio ID, 304
    - incondicionais, 83
    - pipeline, 304
    - terminando, 85
    - Ver também* desvios condicionais
  - desvios adiados, 87
    - definição, 252
    - limitações de escalonamento, 306
  - pipelines de cinco estágios, 308
    - reduzindo, 302–305
    - solução do hazard de controle, 275
    - Ver também* desvios
  - desvios condicionais
    - alterando o contador de programa com, 308
    - ARM, 132
    - compilando if-then-else em, 83
    - definição, 83
    - em loops, 103
    - endereçamento relativo ao PC, 103
    - implementação, 88
  - desvios incondicionais, 83
  - desviosem pipeline, 304
  - deteção de erro, 485
  - diagramas de pipeline de múltiplo ciclo de clock, 286
    - cinco instruções, 286
    - definição, 286
    - ilustração, 286
  - diagramas de pipeline de único ciclo de clock, 286
    - definição, 286
    - ilustração, 288
  - dicing, 35
  - dies, 35
  - Digital Video Disks (DVDs), 16, 17
  - dígitos binários. *Ver* bits
  - dígitos de guarda
    - arredondamento com, 216
    - definição, 215
  - Direct Memory Access (DMA)
    - configuração, 478
    - definição, 477
    - múltiplos dispositivos, 478
    - transferências, 478, 479
  - Direct3D, A-577
  - diretivas de leiaute de dados, B-639
  - diretivas do montador, B-632
  - discos magnéticos. *Ver* discos rígidos
  - discos óticos
    - definição, 16
    - tecnologia, 17
  - discos rígidos
    - cabeça de leitura-escrita, 15
    - definição, 15
    - diâmetros, 16
    - ilustração, 15
    - tempos de acesso, 16
  - discos rígidos híbridos, 468
  - disponibilidade, 462
  - dispositivos de entrada, 10
  - dispositivos de saída, 10
  - dividendo, 192
  - divisão, 191–197
    - algoritmo, 193
    - com sinal, 193–195
    - dividendo, 192
    - divisor, 192
    - hardware, 192–193
    - hardware, versão melhorada, 194
    - instruções, B-668–B-669
  - mais rápida, 195
  - no MIPS, 195–197
  - operandos, 192
  - ponto flutuante, 209, B-686
  - quociente, 192
  - resto, 192
  - SRT, 195
  - Ver também* aritmética
  - divisor, 192
  - don't cares
    - termo, 256
  - Double Data Rate RAMs (DDR RAMs), 381
  - DRAM síncrona (SRAM), 381
  - Dynamic Random Access Memory (DRAM), 365, 379
    - crescimento de capacidade, 20
    - custo, 16
    - definição, 12
    - Double Date Rate (DDR), 381
    - GPU, A-597–A-598
    - largura de banda externa, 382
    - síncrona (SDRAM), 381
    - tamanho, 382
    - velocidade, 16
  - Dynamically Linked Libraries (DLLs), 116–117
    - definição, 117
    - versão da link-edição de procedimento tardio, 117, 118
- ## E
- E/S controlada por interrupção, 475
  - E/S mapeada na memória
    - definição, 475
    - uso, B-658
  - E/S, B-658–B-659
    - chip sets, 473
    - comunicação do processador, 475–476
    - controlada por interrupção, 475
    - controladores, 478, 496
    - desempenho, 461
    - direções futuras, 498
    - impacto no desempenho do sistema, 483–484
    - instruções, 475
    - largura de banda, 498
    - mapeada na memória, 475, B-658
    - medidas de desempenho, 480–482
    - padrões, 471
    - paralelismo, 483–488
    - problema de coerência, 479
    - sistemas, 459
    - solicitações, 461, 498
    - taxa, 480, 492, 493
    - transações, 470
  - E/S, benchmarks, 480–481
    - processamento de transação, 480–481
    - sistema de arquivos, 481–482
    - Web, 481–482
    - Ver também* benchmarks



- E/S, dispositivos
  - características, 460
  - comandos, 475
  - diversidade, 460
  - expansibilidade, 461
  - ilustração, 459
  - interfaces, 473–479
  - leituras/escritas, 461
  - múltiplos caminhos, 498
  - número máximo, 497
  - prioridades, 476–477
  - transferências, 472, 477–478
- E/S, interconexões
  - função, 470
  - processadores x86, 471–473
- E/S, sistemas
  - projeto, 482–483
  - avaliação de potência, 493
  - desempenho, 498
  - elo mais fraco, 482
  - exemplo de projeto, 491–493
  - história, 498
  - organização, 472
  - responsabilidades do sistema operacional, 474
  - taxa de transferência máxima, 497
- Electrically Erasable Programmable Read-Only Memory (EEPROM), 468
- elementos
  - caminho de dados, 247, 252
  - combinacionais, 245
  - estado, 246, 247, 248
- elementos combinacionais, 245
- elementos de estado
  - armazenando/acessando instruções, 248
  - clock, 247
  - definição, 246
  - entradas, 246
  - lógica combinacional, 247
- elementos do caminho de dados
  - compartilhamento, 252
  - definição, 247
- endereçamento
  - base, 105
  - deslocamento, 105
  - imediatos de 32 bit, 102–109
  - intermediário, 105
  - modos do MIPS, 105
  - modos x86, 136, 138
  - pseudodireto, 105
  - registrador, 105
  - relativo ao PC, 103, 105
  - saltos e desvios, 103–105
- endereçamento de base, 66, 105
- endereçamento de registradores, 105
- endereçamento intermediário, 105
- endereçamento por deslocamento, 105
- endereçamento pseudodireto, 105
- endereçamento relativo ao PC, 103, 105
- endereço de retorno, 89
- endereços
  - base, 66
  - byte, 66
  - definição, 65
  - imediatos de 32 bits, 102–109
  - memória, 72
  - virtuais, 397–398, 414
- endereços físicos, 397
  - definição, 396
  - espaço, 515, 517
  - mapeando, 397
- endereços virtuais
  - causando falhas de página, 414
  - definição, 397
  - mapeamento, 397
  - tamanho, 398
- entradas, 256
- escalonamento de pipeline dinâmico, 321–322
  - buffer de reordenação, 321
  - conceito, 322
  - definição, 321
  - especulação baseada em hardware, 322
  - estação de reserva, 321
  - unidade de commit, 321
  - unidades primárias, 321
- escritas
  - cache write-back, 376, 377
  - cache write-through, 376, 377
  - complicações, 376
  - custo, 416
  - esquemas, 376
  - memória virtual, 404
  - tratamento, 375–377
  - tratamento da hierarquia de memória, 420
- espaço de endereços plano, 439
- espaço de endereços, 396, 399
  - compartilhado, 516–517
  - estendendo, 439
  - físico único, 515
  - ID (ASID), 411
  - não mapeado, 414
  - plano, 439
  - virtual, 411
- especulação baseada em hardware, 322
- especulação, 316–317
  - baseada em hardware, 322
  - definição, 316
  - desempenho, 317
  - implementação, 316
  - mecanismo de recuperação, 317
  - problemas, 317
- espelhamento, 485
- estações de reserva
  - definição, 321
  - operandos de buffering, 322
- estado
  - componentes lógicos, 246
  - especificação, 399
  - exceção, salvando/restaurando, 415
  - no esquema de previsão de 2 bits, 306
- estágio de acesso à memória
  - instrução load, 281
  - instrução store, 283
  - linha de controle, 291
- estágio de busca de instrução
  - instrução load, 279
  - instrução store, 283
  - linha de controle, 288
- estágio de decodificação de instrução/leitura de banco de registradores
  - instrução load, 279
  - instrução store, 283
  - linha de controle, 288
- estágio do cálculo de execução ou endereço, 281, 283
- Ethernet, 17, 18
- EX, estágio
  - deteção de exceção de overflow, 311
  - instruções load, 281
  - instruções store, 283
- exceções, 309–314, B-655–B-656
  - associação, 313
  - caminho de dados com controles para tratamento, 311
  - definição, 183, 310
  - detectando, 310
  - estágio para salvar/restaurar, 415
  - estouro, 311
  - exemplo de computador em pipeline, 312
  - imprecisas, 313
  - instruções, B-689
  - interrupções *versus*, 309–310
  - motivos, 310–311
  - na arquitetura MIPS, 310–311
  - na implementação em pipeline, 311–314
  - PC, 410, 412
  - precisas, 313
  - resultado devido a estouro na instrução add, 387
  - tipos de evento, 310
- exception enable, 412
- Exception Program Counters (EPCs), 310
  - captura de endereço, 313
  - copiando, 183
  - definição, 183, 311
  - determinando o reinício, 310
  - transferindo, 185
- exclusão mútua, 109
- exclusive OR (XOR), instruções, B-672
- execução fora de ordem
  - complexidade do desempenho, 393
  - definição, 322
  - processadores, 325
- exemplo de shader de pixel, A-579–A-580
- expansão
  - forte, 514, 515
  - fraco, 514
- expansão forte, 514, 515
- expansão fraca, 514
- expoentes, 197–198
- extensão de sinal, 250
  - atalho, 73–74
  - definição, 98
- extensões de multimídia
- vetor *versus*, 528

## F

facilidades, B-639–B-641  
falácias  
  adição imediata sem sinal, 223  
  baixa utilização usa pouca potência, 40  
  definição, 39  
  desempenho de pico, 552–553  
  deslocamento à direita, 222–223  
  GPUs, A-626–A-627, A-628  
  importância da compatibilidade binária comercial, 142  
  instruções poderosas significam maior desempenho, 141  
  lei de Amdahl, 552  
  linguagem assembly para desempenho, 141–142  
  MTTF, 494  
  pipelining, 328  
  taxas de falha de disco, 494–495  
falhas  
  motivos para, 463  
  taxas de disco, 494–495  
  tempo médio entre (MTBF), 462  
  tempo médio para (MTTF), 462, 463, 494  
falta sob falta, 436  
faltas de cache  
  cache associativo em conjunto, 387–389  
  cache de mapeamento direto, 387  
  cache totalmente associativa, 389  
  capacidade, 421  
  ciclos de clock de stall da memória, 382  
  compulsórias, 421  
  conflito, 421  
  definição, 375  
  etapas, 375  
  na cache write-through, 376  
  reduzindo com posicionamento de bloco flexível, 385–389  
  substituição de bloco, 419–420  
  tratamento, 375  
faltas de página, 401  
definição, 397  
  endereço virtual causando, 414  
  para acesso a dados, 413  
  tratamento, 398, 411–416  
  *Ver também* memória virtual  
Fast Fourier Transforms (FFT), A-609  
Filebench, 481  
fitas magnéticas, 496  
  definição, 16  
  histórico de uso, 496  
fluxo de instruções da esquerda para a direita, 278  
formato de instrução tipo J, 103  
formato R, 257  
  definição, 77  
  operações da ALU, 250  
formato-I, 77  
formatos de instrução  
  ARM, 132  
  definição, 75  
  instrução de salto, 263

MIPS, 132  
  tipo I, 77  
  tipo J, 103  
  tipo R, 77, 257  
  x86, 141  
forwarding, 292–302  
  ALU antes, 296  
  caminho de dados para resolução de hazard, 297  
  com duas instruções, 269–270  
  controle, 294  
  definição, 269  
  funcionamento, 293  
  multiplexadores, 297  
  múltiplos resultados, 272  
  registradores de pipeline antes, 296  
  representação gráfica, 270  
frações, 197, 198, 199  
função do próximo estado, 427  
  definição, 427  
funções de controle  
  definindo, 258  
  para implementação de único ciclo, 262  
Fused-Multiply-Add (FMA), operação, 217, A-604

## G

General Purpose GPUs (GPGPUs), 530, A-571  
gigabytes, 16  
GPU, arquiteturas do sistema, A-572–A-576  
  heterogêneas, A-572–A-574  
  implicações, A-586  
  interfaces e drivers, A-574  
  pipeline lógico gráfico, A-574  
  unificadas, A-575–A-576  
grades, A-582  
Graphics Processing Units (GPUs), 528–534  
  aplicações N-body, A-620–A-626  
  aritmética de ponto flutuante, A-580, A-601, A-627  
  arquitetura NVIDIA, 530–533  
  caches multinível, 530  
  como aceleradoras, 529  
  computação geral, A-627  
  definição, 34, 512, A-569  
  dobrando o desempenho, A-570  
  evolução, A-571  
  falácias e armadilhas, A-626–A-628  
  General Purpose (GPGPUs), 530, A-571  
  geração GeForce série, 8, A-571  
  gráficos de tempo real, A-577  
  história, A-569  
  interfaces de programação, 529, A-580  
  interpolação de atributo, A-602–A-603  
  mapeando aplicações, 612t–683f  
  memória principal, 530  
  memória, 530  
  modo gráfico, A-571  
  paralelismo, 530, A-629  
  perspectiva, 533–534  
  pipeline gráfico lógico, A-577–A-578  
  programando, A-576–A-586

resumo, A-629  
sistema de memória paralela, A-596–A-600  
software de driver, 530  
tendências gráficas, A-570  
*Ver também* computação GPU

## H

halfwords, 100  
handlers  
  definição, 413  
  perda de TLB, 414  
handlers de interrupção, B-654  
hardware  
  como camada hierárquica, 6  
  linguagem, 7–9  
  operações, 60–63  
  procedimentos de suporte, 88–97  
  virtualizável, 424  
hardware virtualizável, 424  
hazard de dados de uso de load, 271, 302  
hazards, 269–275  
  controle, 272–275, 302–309  
  dados, 269–272, 292–302  
  definição, 269  
  estruturais, 269, 283  
  forwarding, 298  
  *Ver também* pipelining  
hazards de controle, 272–275, 302–309  
  definição, 272, 302  
  previsão de desvio como solução, 275  
  previsão de desvio dinâmico, 306–308  
  processadores estáticos de despacho múltiplo, 317  
  redução branch delay, 302–305  
  resumo de pipeline, 308–309  
  simplicidade, 302  
  soluções, 273  
  stalls de pipeline como solução, 273  
  suposição de desvio não tomado, 302  
  técnica de decisão adiada, 275  
hazards de dados, 269–272, 292–302  
  definição, 269  
  forwarding, 269, 292–302  
  uso de load, 271, 302  
  stalls, 298–301  
  *Ver também* hazards  
hazards de desvio. *Ver* hazards de controle  
hazards estruturais, 269, 283  
heap  
  alocando espaço, 94–97  
  definição, 94  
hierarquias de memória  
  armadilhas, 437–445  
  bloco (ou linha), 366  
  caches, 368–382  
  definição, 365  
  dependência de, 366  
  desafios de projeto, 423  
  desempenho de cache, 382–396  
  diagrama de estrutura, 367  
  estrutura comum, 417–423  
  estrutura, 366

hierarquias de memória (*cont.*)  
 explorando, 361–546  
 inclusão, 437  
 memória virtual, 396–416  
 múltiplos níveis, 366  
 operação geral, 408  
 paralelismo, 430–433  
 parâmetros de projeto quantitativos, 417  
 pares de nível, 366  
 tempo de execução de programa, 395  
 variância, 395  
 hot-swapping, 487

## I

IBM Cell QS20  
 características, 546  
 definição, 549  
 desempenho básico versus totalmente otimizado, 551  
 desempenho SpMV, 550  
 ilustração, 547  
 LBMHD performance, 551  
 modelo roofline, 548  
 ID, estágio  
 execução de desvio, 304  
 instrução store, 280  
 instruções load, 280  
 identificadores de processo, 411  
 identificadores de tarefa, 411  
 IEEE, 754, padrão de ponto flutuante, 199, 200  
 modos de arredondamento, 217  
*Ver também* ponto flutuante  
 If-then-else, 83  
 If, instruções, 103  
 implementação de único ciclo  
 definição, 262  
 desempenho com pipeline *versus*, 266–267  
 execução sem pipeline *versus* execução com pipeline, 268  
 função de controle, 262  
 não uso da, 263–265  
 penalidade, 265  
 índice fora dos limites, verificação, 87  
 informação de depuração, B-638  
 informação de relocação, B-638, B-639  
 instruções, 58–178  
 acesso à memória, A-594–A-595  
 add imediato, 67  
 adição, 182, B-667  
 ARM, 130–133  
 assembly, 63  
 básicas, 228  
 bloco básico, 85–86  
 busca, 249  
 campos, 75  
 cientes da cache, 441  
 codificação, 78  
 como words, 60  
 comparação, B-672–B-674  
 conversão, B-685–B-686  
 definição, 7, 60  
 desempenho, 25–26  
 desvio condicional, 83  
 desvio, B-674–B-676  
 divisão, B-668–B-669  
 E/S, 475  
 exceção e interrupção, B-689  
 flushing, 302, 304, 313  
 fluxo da esquerda para a direita, 278  
 imediatas, 67  
 introdução, 60  
 load ligado, 110  
 load, 65, B-679–B-680  
 lógicas aritméticas, 248, B-667–B-672  
 manipulação de constante, B-672  
 move condicional, 308  
 movimentação de dados, B-682–B-683  
 multiplicação, 190, B-669  
 negação, B-670  
 nop, 300  
 operações lógicas, 80–83  
 OR exclusivo, B-672  
 ponto flutuante (x86), 221  
 ponto flutuante, 209–211, B-684–B-690  
 PTX, A-592, A-593  
 referência à memória, 243  
 reiniciáveis, 413  
 representação no computador, 74–80  
 resto, B-671  
 retomando, 416  
 salto, 87, 89, B-677  
 sequência de pipeline, 298  
 shift, B-671  
 sinais eletrônicos, 74  
 store condicional, 110–111  
 store, 67, B-680–B-682  
 subtração, 182, B-671–B-672  
 thread, A-592  
 tipo R, 248–249  
 tomada de decisão, 83–88  
 transferência de dados, 65  
 trap, B-677–B-679  
 vetor, 527  
 x86, 133–141  
*Ver também* instruções aritméticas; MIPS;  
 operandos  
 instruções aritméticas  
 lógicas, 248  
 MIPS, B-667–B-672  
 operandos, 63  
*Ver também* instruções  
 instruções cientes da cache, 441  
 instruções de acesso à memória, A-594–A-595  
 instruções de comparação, B-672–B-674  
 lista, B-672–B-674  
 ponto flutuante, B-685  
 instruções de conjunto, 86  
 instruções de conversão, B-685–B-686  
 instruções de deslocamento, 80, B-671  
 instruções de desvio, B-674–B-676  
 impacto do pipeline, 302  
 instrução de salto *versus*, 263  
 lista, B-674–B-676  
 instruções de flushing, 302, 304

definição, 302  
 exceções, 313  
 instruções de manipulação de constante, B-672  
 instruções de máquina, 75  
 instruções de movimentação de dados, B-682–B-683  
 instruções de movimento condicional, 308  
 instruções de negação, B-670, B-688–B-689  
 instruções de ponto flutuante, B-684–B-690  
 adição, B-684–B-685  
 comparação, B-685  
 conversão, B-685–B-686  
 divisão, B-686  
 load, B-686–B-687  
 move, B-687–B-688  
 multiplicação, B-688  
 negação, B-688–B-689  
 raiz quadrada, B-689  
 store, B-689  
 subtração, B-689–B-690  
 truncamento, B-689  
 valor absoluto, B-684  
 instruções de jump, 252  
 controle e caminho de dados, 264  
 formato de instrução, 263  
 implementando, 263  
 instrução de desvio *versus*, 263  
 lista, B-677  
 instruções de tipo R, 248–249  
 caminho de dados em operação para, 260  
 caminho de dados para, 260  
 instruções de tomada de decisão, 83–88  
 instruções de transferência de dados  
 definição, 65  
 load, 66  
 offset, 66  
 store, 67  
*Ver também* instruções  
 instruções de trap, B-677–B-679  
 instruções imediatas, 67  
 instruções load  
 acesso, A-601  
 bloco, 133  
 caminho de dados em pipeline, 285  
 caminho de dados na operação, 261  
 com sinal, 98  
 compilando, 67  
 definição, 66  
 detalhes, B-679–B-680  
 estágio EX, 281  
 estágio ID, 280  
 estágio IF, 280  
 estágio MEM, 282  
 estágio WB, 282  
 halfword sem sinal, 100  
 interligadas, 110, 111  
 lista, B-679–B-680  
 load byte sem sinal, 98  
 load half, 100  
 load upper immediate, 102, 103  
 ponto flutuante, B-686–B-687  
 registrador de base, 257  
 sem sinal, 98



unidade para implementação, 251  
*Ver também* instruções store  
 instruções por ciclo de clock (IPC), 314  
 instruções reiniciáveis, 413  
 Intel Nehalem  
   caches, 436  
   foto do processador, 434  
   hierarquias de memória, 435–437  
   técnicas de redução de penalidade de falha, 435–437  
   TLB hardware for, 435  
   tradução de endereço, 435  
 Intel Threading Building Blocks, A-615  
 Intel Xeon e5345  
   básico *versus* totalmente otimizado  
   características, 546  
   definição, 546  
   desempenho do LBMHD, 551  
   desempenho do SpMV, 550  
   desempenho, 551  
   ilustração, 546  
   modelo roofline, 548  
 intensidade aritmética, 540  
 intercalação, 380, 382  
 interconexão assíncrona, 470  
 interpolação de atributos, A-602–A-603  
 interrupções  
   definição, 183, 310  
   exceções *versus*, 309–310  
   imprecisas, 313  
   instruções, B-689  
   precisas, 313  
   tipos de evento, 310  
   vetorizadas, 311  
 interrupções imprecisas, 313  
 interrupções precisas, 313  
 interrupções vetorizadas, 311  
 interrupt enable, 412  
 Interrupt Priority Levels (IPLs), 476–477  
   definição, 477  
   mais altos, 477  
 Intrinsic FastMATH, processador,  
   377–378  
   caches, 378  
   definição, 377  
   processamento de leitura, 408  
   processamento write-through, 408  
   taxas de perda de dados, 378, 389  
   TLB, 406

## J

Java  
   algoritmos de classificação, 127  
   bytecode, 118  
   caracteres, 100–101  
   hierarquia de tradução, 119  
   interpretando, 119, 130  
   objetivos, 117  
   programas, iniciando, 117–119  
   strings, 100–101  
 Java Virtual Machine (JVM), 118  
 Just In Time (JIT), compiladores, 119, 554

## K

kernels  
   CUDA, A-582, A-586  
   definição, A-582

## L

LAPACK, 219  
 largura de banda  
   bisseção, 535  
   cache L2, 546  
   E/S, 498  
   externa à DRAM, 382  
   memória, 379, 380  
   rede, 535  
 largura de banda de bisseção, 535  
 latência  
   instrução, 329  
   memória, A-628  
   pipeline, 276  
   restrições, 482  
   rotacional, 465  
   uso, 318, 319  
 latência de instrução, 329  
 latência de uso  
   definição, 318  
   uma instrução, 319  
 latência rotacional, 465  
 Lattice Boltzmann Magneto-Hydrodynamics  
   (LBMHD), 549–551  
   definição, 549  
   otimizações, 550–551  
   performance, 551  
 lei de Amdahl, 384, 512  
   corolário, 40  
   definição, 39  
   falácia, 552  
 Lei de Gresham, 228  
 Lei de Moore, 528, 626t–683f  
 linguagem de máquina  
   offset de desvio, 104–105  
   decodificação, 106  
   definição, 7, 75, B-630  
   ponto flutuante, 210  
   ilustração, 8  
   MIPS, 79  
   SRAM, 13  
   traduzindo linguagem assembly MIPS para,  
     78–79  
 linguagem fonte, B-632  
 linguagens de alto nível, 7–9, B-632  
   benefícios, 9  
   definição, 8  
   importância, 8  
 linguagens de programação  
   orientadas a objeto, 130  
   variáveis, 64  
   *Ver também linguagens específicas*  
 linguagens de shading, A-578  
 linguagens orientadas a objeto  
   definição, 130  
   *Ver também* Java

linhas. *Ver* blocos  
 linhas de controle  
   acesso à memória, 291  
   ativadas, 260  
   busca de instrução, 288  
   configuração, 258, 260  
   decodificação de instrução/leitura de banco  
     de registradores, 288  
   execução/cálculo de endereço, 288  
   no caminho de dados, 257  
   três estágios finais, 288  
   valores, 289  
   write-back, 291  
 link-editores, 113–116, B-642–B-643  
   arquivos executáveis, 113, B-643  
   definição, 113, B-630  
   etapas, 113  
   ilustração da função, B-643  
   usando, 114–116  
 linkagem de arquivos objeto, 114–116  
 Linpack, 537  
 Liquid Crystal Displays (LCDs), 11  
 little-endian, ordem de bytes, B-661  
 load word, 65, 67  
 localidade  
   espacial, 364–365, 367  
   princípio, 364, 365  
   temporal, 364, 365, 367  
 localidade espacial, 364–365  
   definição, 364  
   exploração de bloco grande, 374  
   tendência, 367  
 localidade temporal, 365  
   definição, 364  
   tendência, 367  
 locks, 516  
 lógica  
   combinacional, 247  
   componentes, 246  
 lógica combinacional, 247  
 loops, 84–85  
   definição, 84  
   desvios condicionais, 103  
   for, 127  
   previsão, 306  
   teste, 128  
   while, compilando, 84–85  
 loops for, 127

## M

macros  
   definição, B-630  
   exemplo, B-640–B-641  
   uso, B-640  
 mapa de MIP, A-603  
 mapas de bits, 12  
   armazenamento, 12  
   definição, 11, 68  
   objetivo, 12  
 mapeando aplicações, A-612–A-626  
 máquinas de estados finitos (FSMs), 426–430  
   controladores, 429

máquinas de estados finitos (FSMs) (*cont.*)  
 definição, 427  
 estilo, 429  
 função do próximo estado, 427  
 implementação, 427  
 Mealy, 429  
 Moore, 429  
 para controlador de cache simples, 429  
 máquinas virtuais (VMs), 423–426  
 benefícios, 424  
 definição, B-660  
 ilusão, 426  
 melhoria do desempenho, 425  
 para melhoria da proteção, 424  
 simulação, B-660  
 suporte à arquitetura do conjunto de  
 instruções, 424–425  
 matrizes esparsas, A-612–A-614  
 Mealy, máquina, 429  
 Mean Time Between Failures (MTBF), 462  
 Mean Time To Failure (MTTF), 462, 463  
 avaliações, 484  
 falácias, 494  
 Mean Time To Repair (MTTR), 462, 463  
 meia precisão, A-601  
 memória  
 afinidade, 549, 550  
 atômica, A-585  
 cache, 13, 368–396  
 CAM, 391  
 compartilhada, A-585, A-599–A-600  
 constante, A-600  
 definição, 12  
 DRAM, 12, 365, 379, 381  
 eficiência, 518  
 endereços, 72  
 espaços, A-599  
 flash, 15, 16, 468–469  
 global, A-585, A-599  
 GPU, 530  
 instruções, caminho de dados para, 253  
 largura de banda, 379, 380  
 layout, B-644  
 local, A-585, A-600  
 não volátil, 15  
 operandos, 65–66  
 principal, 15  
 SDRAM, 381  
 secundária, 15  
 sistema paralelo, A-596–A-600  
 stalls, 385  
 tecnologias para criação, 18–19  
 textura, A-600  
 uso, B-644–B-645  
 virtual, 396–416  
 volátil, 15  
 memória compartilhada  
 bancos de SRAM, A-600  
 caching, A-615  
 CUDA, A-615  
 definição, A-585  
 memória de baixa latência, A-585  
 n-body, A-622

por CTA, A-599  
*Ver também* memória  
 memória constante, A-600  
 memória de textura, A-600  
 memória física. *Ver* memória principal  
 memória flash, 468–469  
 características, 16, 468  
 definição, 15, 468  
 EEPROM, 468  
 nivelamento de desgaste, 468  
 NOR, 468  
 memória global, A-585, A-599  
 memória local, A-585, A-600  
 memória não volátil, 15  
 memória primária. *Ver* memória principal  
 memória principal, 397  
 definição, 15  
 endereços físicos, 396, 397  
 tabelas de página, 404  
*Ver também* memória  
 memória secundária, 15  
 memória virtual, 396–416  
 definição, 396  
 escritas, 404  
 faltas de página, 397, 401  
 implementação da proteção, 409–411  
 integração, 406–409  
 mecanismo, 416  
 motivações, 396–397  
 resumo, 416  
 segmentação, 398  
 tradução de endereço, 397, 404–406  
 virtualização, 426  
*Ver também* páginas  
 memória volátil, 15  
 metodologia de clocking disparada  
 por transição, 246, 247  
 metodologia de clocking, 246–247  
 definição, 246  
 disparada por transição, 246, 247  
 por previsibilidade, 246  
 métodos  
 estáticos, B-644  
 microarquitecturas  
 AMD Opteron X4 (Barcelona), 326  
 definição, 325  
 migração, 432  
 milhões de instruções por segundo (MIPS), 41  
 MIPS-32, conjunto de instruções, 227  
 MIPS, 61, 78–79, B-663–B-690  
 alocação de memória para programa  
 e dados, 94  
 campos, 76–77  
 classes de instruções, 145  
 codificação de instrução, 78, 107, B-667  
 compilando atribuição C complexa  
 para, 62–63  
 compilando instruções de atribuição C  
 para, 62  
 conjunto de instruções, 60, 144, 225  
 convenções de registrador, 96  
 CPU, B-663  
 despacho múltiplo estático, 317–320

divisão, 195–197  
 endereçamento para imediatos de 32 bits,  
 102–109  
 endereços de memória, 66  
 exceções, 310–311  
 formatos de instrução, 109, 132, B-667  
 FPU, B-663  
 instruções aritméticas, 60, B-667–B-672  
 instruções de comparação, B-672–B-674  
 instruções de desvio, B-674–B-676  
 instruções de manipulação de constante,  
 B-672  
 instruções de ponto flutuante, 209–211  
 instruções de salto, B-677–B-679  
 instruções lógicas, B-667–B-672  
 linguagem de máquina, 79  
 mapa de opcode, B-666  
 mapeamento de instrução assembly, 75  
 modos de endereçamento, B-663–B-664  
 multiplicação, 190  
 núcleo aritmético, 226  
 operandos, 61  
 pseudo, 226, 227  
 registradores de controle, 412  
 semelhanças do ARM, 131  
 sintaxe do assembler, B-664–B-667  
 suporte a diretiva do assembler,  
 B-664–B-667  
 MIPS, conjunto de instruções básico,  
 228  
 ilustração da implementação, 245  
 implementação, 242–244  
 subconjunto, 242–243  
 visão abstrata, 243  
 visão geral, 243–244  
*Ver também* MIPS  
 MIPS, core  
 arquitetura, 196  
 conjunto de instruções, 228, 242–244  
 mix de instruções, 28  
 modelo de consistência de memória,  
 433  
 modelo dos três Cs, 421  
 modo kernel, 410  
 modos de endereçamento, B-663–B-664  
 módulos, B-630  
 montadores, 112–113, B-636–B-641  
 aceitação de número, 113  
 arquivo objeto, 113  
 definição, 7, B-630  
 função, 113, B-636  
 informação de relocação, B-638, B-639  
 macros, B-630, B-640–B-641  
 montagem de código condicional, B-641  
 pseudoinstruções, B-641  
 tabela de símbolos, B-637  
 velocidade, B-638  
 Moore, máquinas, 429  
 mouse, anatomia, 11  
 move, instruções, B-682–B-683  
 co-processadores, B-682–B-683  
 detalhes, B-682–B-683  
 ponto flutuante, B-687–B-688

Multiple Instruction Multiple Data (MIMD), 533  
definição, 524  
Multiple Instruction Single Data (MISD), 525  
multiplexadores  
controle de seletor, 253  
controles, 427  
definição, 243  
forwarding, valores de controle, 297  
no caminho de dados, 257  
multiplicação com sinal, 189  
multiplicação de ponto flutuante, 205–209  
binária, 206–208  
etapas, 205–206  
ilustração, 207  
instruções, 209  
significandos, 205  
multiplicação-adição (MAD), A-601  
multiplicação, 186–191  
assinada, 189  
hardware, 187–189  
instruções, 190, B-669  
mais rápida, 190  
multiplicador, 187  
multiplicando, 187  
no MIPS, 190  
operandos, 187  
ponto flutuante, 205–207, B-688  
primeiro algoritmo, 188  
produto, 187  
rápida, hardware, 191  
*Ver também* aritmética  
multiplicador, 187  
multiplicando, 187  
multiprocessador Tesla, 532  
multiprocessadores  
arquitetura multithreaded, A-588–A-589, A-596  
benchmarks, 537–539  
definição, 510  
desempenho, 553–554  
memória compartilhada, 511, 515–517  
organização, 509, 517  
passagem de mensagens, 517–521  
perspectiva histórica, 555  
software, 510  
UMA, 516  
multiprocessadores  
mudança de projeto, 511  
multicore, 5, 31, 510  
multiprocessadores de memória compartilhada (SMP), 515–517  
definição, 511, 515  
espaço único de endereços físicos, 515  
sincronização, 516  
multiprocessadores multicore, 31  
benchmarking com modelo roofline, 546–552  
características, 546  
definição, 5, 510  
dois soquetes, 547  
organização do sistema, 547  
multithreading coarse-grained, 521–522

multithreading do hardware, 521–524  
coarse-grained, 521–522  
definição, 521  
fine-grained, 521, 523  
opções, 522  
simultâneo, 522–524  
multithreading fine-grained, 521, 523  
multithreading simultâneo (SMT), 522–524  
definição, 522  
paralelismo em nível de thread, 523  
slots de despacho não usados, 524  
suporte, 523  
multithreading, A-587–A-588  
coarse-grained, 521–522  
definição, 512  
fine-grained, 521, 523  
hardware, 521–524  
simultâneo (SMT), 522–524

## N

n-body  
algoritmo all-pairs, A-620  
comparação de desempenho, A-623–A-624  
GPU, simulação, A-625  
matemática, A-620  
otimização, A-622  
resultados, A-624–A-626  
uso da memória compartilhada, A-622  
NAS (NASA Advanced Supercomputing), 539  
Newton, iteração, 215  
níveis de prioridade, 476–477  
nivelamento de desgaste, 468  
Nonuniform Memory Access (NUMA), 516  
nops, 300  
NOR, memória flash, 468  
NOR, operação, 82–83, B-670  
north bridge, 471  
NOT, operação, 82, B-671  
notação científica  
definição, 197  
para reais, 197  
somando números, 202  
notação imparcial, 74, 200  
números  
binários, 68  
com sinal, 68–74  
computador *versus* mundo real, 217  
decimais, 68, 71  
desnormalizados, 218  
hexadecimais, 75–76  
sem sinal, 68–74  
números binários  
ASCII *versus*, 98  
conversão para números decimais, 71  
conversão para números hexadecimais, 76  
definição, 68  
números com sinal, 68–74  
sinal e magnitude, 70  
tratando como sem sinal, 87  
números decimais  
conversão de números binários para, 71  
definição, 68

números desnormalizados, 219  
números hexadecimais, 75–76  
conversão de números binários para, 76  
definição, 75  
números não sinalizados, 68–74  
NVIDIA GeForce, 8800, A-605–A-612  
algoritmo all-pairs N-body, A-625  
cálculos de álgebra linear densa, A-609  
conjunto de instruções, A-607  
desempenho FFT, A-609  
desempenho na classificação, A-611–A-612  
desempenho, A-609  
escalabilidade, A-609  
estatísticas de aproximação de função especial, A-602  
processador streaming, A-608  
rasterização, A-608  
ROP, A-608  
Special Function Unit (SFU), A-608  
Streaming Multiprocessor (SM), A-607  
Streaming Processor Array (SPA), A-605  
Texture/Processor Cluster (TPC), A-606–A-607  
NVIDIA, arquitetura de GPU, 530–533

## O

opcodes  
definição, 77, 257  
definição de linha de controle, 260  
OpenGL, A-577  
OpenMP (Open MultiProcessing), 539  
operação de memória atômica, A-585  
operações  
atômicas, implementando, 110  
hardware, 60–63  
inteiros do x86, 136–140  
lógicas, 80–83  
operações lógicas, 80–83  
AND, 81–82, B-668  
ARM, 133  
definição, 80–83  
MIPS, B-667–B-672  
NOR, 82–83, B-670  
NOT, 82, B-671  
OR, 82, B-671  
shifts, 80  
operandos constantes, 67–68  
em comparações, 86  
ocorrência frequente, 68  
operandos, 63–68  
compilação quando na memória, 65  
constante, 67–68  
deslocamento, 132  
divisão, 192  
imediatos de 32 bits, 102–103  
instruções aritméticas, 63  
memória, 65–66  
MIPS, 61  
multiplicação, 187  
ponto flutuante, 210  
somando, 182  
*Ver também* instruções

OR, operação, 82, B-671  
 ordem principal de linha, 214  
 otimização  
   compilador, 130  
   manual, 130  
 overflow  
   definição, 70, 198  
   detecção, 182  
   exceções, 311  
   ocorrência, 71  
   ponto flutuante, 198  
   saturação, 183–184  
   subtração, 182

## P

P + Q, redundância, 486  
 Packed, formato de ponto flutuante, 222  
 pacotes de despacho, 317  
 páginas  
   definição, 397  
   localizando, 399  
   LRU, 402  
   modificadas, 404  
   número físico, 397  
   número virtual, 397  
   offset, 397  
   posicionando, 399  
   tamanho, 398  
   *Ver também* memória virtual  
 páginas modificadas, 404  
 paralelismo, 31, 314–325  
   benefícios de desempenho, 33  
   E/S, 483–488  
   emissão múltipla, 314–322  
   GPUs e, 530, A-629  
   hierarquias de memória, 430–433  
   multicore, 524  
   multithreading, 524  
   nível de dados, 525  
   nível de instrução, 31, 314, 324  
   nível de processo, 510  
   nível de tarefa, 510  
   tarefa, A-586  
   thread, A-585  
 paralelismo de tarefas, A-586  
 paralelismo de threads, A-585  
 paralelismo em nível de dados, 525  
 paralelismo em nível de instrução (ILP)  
   definição, 31, 314  
   exploração, aumentando, 324  
   *Ver também* paralelismo  
 paralelismo em nível de processo, 510  
 paralelismo em nível de tarefa, 510  
 parâmetros formais, B-641  
 paravirtualização, 441  
 paridade, 485  
   disco, 486  
   intercalada por bit, 485  
   intercalada por bloco, 485–486  
   intercalada por bloco distribuído, 486–487  
 PARSEC (Princeton Application Repository for Shared Memory Computers), 539

passagem de mensagens  
   definição, 517  
   multiprocessadores, 517–521  
 PCI-Express (PCIe), A-573  
 penalidade de falta  
   caches multinível, reduzindo, 392–395  
   definição, 366  
   determinação, 374  
   técnicas de redução, 435–437  
 Pentium, jogada de moralidade do bug, 223–225  
 perdas compulsórias, 421  
 perdas de capacidade, 421  
 perdas de conflito, 421  
 petabytes, 2  
 pilhas  
   alocando espaço, 94  
   definição, 89  
   para argumentos, 126  
   pop, 89  
   procedimentos recursivos, B-650–B-651  
   push, 89, 91  
 pipeline de projeto digital, 327–328  
 pipeline lógico gráfico, A-575  
 pipelines  
   AMD Opteron X4 (Barcelona), 325–327  
   cinco estágios, 267, 279–281, 288  
   despacho duplo estático, 317  
   diagramas de múltiplos ciclos de clock, 286  
   diagramas de único ciclo de clock, 286  
   estágio de acesso à memória, 281, 283  
   estágio de busca de instrução, 279, 283  
   estágio de decodificação de instrução e leitura de arquivo de registrador, 279, 283  
   estágio de execução e cálculo de endereço, 281, 283  
   estágio de write-back, 281, 283  
   estágios, 267  
   gargalos de desempenho, 324  
   impacto da instrução de desvio, 302  
   latência, 276  
   representação gráfica, 270, 286–288  
   sequência de instruções, 298  
 pipelining, 265–276  
   analogia da lavanderia, 265  
   armadilha, 328–329  
   avançado, 324–325  
   benefícios, 265  
   definição, 265  
   exceções, 311–314  
   falácias, 328  
   fórmula de speed-up, 267  
   hazards de controle, 272–275  
   hazards de dados, 269–272  
   hazards estruturais, 269, 283  
   hazards, 269–275  
   instruções de execução simultânea, 276  
   melhoria de desempenho, 269  
   paradoxo, 265  
   projeto do conjunto de instruções, 269  
   resumo, 275  
   tempo de execução, 276

vazão, 276  
 visão geral, 265–276  
 placas-mãe, 12  
 polling, 475  
 ponteiros  
   arrays *versus*, 127–130  
   aumentando, 128  
   frame, 94  
   globais, 93  
   pilha, 89, 91  
 ponteiros de frame, 94  
 ponteiros globais, 93  
 ponto flutuante, 197  
   arquitetura SSE2, 222  
   arredondamento, 215–216  
   cálculos imediatos, 215  
   codificação de instruções, 211  
   conversão binário para decimal, 201  
   definição, 197  
   desafios, 226  
   desvio, 209  
   dígitos de guarda, 215–216  
   divisão, 209  
   forma, 198  
   formato empacotado, 222  
   frequência de instrução MIPS, 228  
   instruções MIPS, 209–211  
   linguagem assembly, 210  
   linguagem de máquina, 210  
   multiplicação e adição reunidas, 217  
   no x86, 220–222  
   operandos, 210  
   overflow, 198  
   padrão IEEE, 754, 199, 200  
   precisão, 219  
   procedimento com matrizes bidimensionais, 212–214  
   programas, compilando, 209–214  
   registradores, 214  
   representação, 197–202  
   sinal e magnitude, 198  
   subtração, 209  
   underflow, 198  
   unidades, 216  
   variação de operandos no x86, 222  
 pop, 89  
 potência  
   eficiência, 324–325  
   natureza crítica, 42  
   relativa, 30  
   taxa de clock e, 29  
 potência relativa, 30  
 precisão dupla  
   definição, 198  
   FMA, A-604  
   GPU, A-604, A-627  
   representação, 201  
   *Ver também* precisão simples  
 precisão simples  
   definição, 198  
   representação binária, 200  
   *Ver também* precisão dupla  
 prefetching, 441, 549

- previsão
    - desvio dinâmico, 306–308
    - esquema de 2 bits, 306
    - estado fixo, 306
    - loops, 306
    - precisão, 306, 307
  - previsão de desvio
    - buffers, 306, 307
    - como solução do hazard de controle, 275
    - definição, 274
    - dinâmica, 274, 275, 306–308
    - estática, 317
  - previsão de desvio dinâmica, 306–308
    - buffer de previsão de desvio, 306
    - definição, 306
    - loops, 306
    - Ver também* hazards de controle
  - previsão de desvio estática, 317
  - previsão de estado fixo, 306
  - previsor de correlação, 308
  - previsores de desvio
    - correlação, 308
    - informação, 308
    - precisão, 306
    - torneio, 308
  - previsores de desvio de torneio, 308
  - previsores de hardware dinâmicos, 274
  - primeira word crítica, 375
  - primeira word requisitada, 375
  - procedimentos, 88–97
    - aninhados, 91–93
    - compilando, 89
    - compilando, mostrando ligação de
      - procedimento aninhado, 92–93
    - cópia de string, 98–100
    - definição, 88
    - etapas de execução, 88
    - folha, 91
    - frames, 94
    - para definir arrays em zero, 128
    - recursivos, 96, B-647–B-649
    - sort, 120–125
    - strcpy, 99, 100
    - swap, 119–120
  - procedimentos aninhados, 91–93
    - compilando procedimento recursivo
      - mostrando, 92–93
    - definição, 91
  - procedimentos de folha
    - definição, 91
    - exemplo, 100
    - Ver também* procedimentos
  - procedimentos recursivos, 96, B-647–B-649
    - definição, B-647
    - invocação de clone, 91
    - pilha, B-650–B-651
    - Ver também* procedimentos
  - processadores de despacho múltiplo estático,
    - 316, 317–320
    - com MIPS ISA, 317–320
    - conjuntos de instruções, 317
    - hazards de controle, 317
    - Ver também* despacho múltiplo
  - processadores de vetor, 525–528
    - comparação de código convencional, 526
    - escalares versus, 527
    - extensões de multimídia, 528
    - instruções, 527
    - Ver também* processadores
  - processadores dinâmicos de despacho
    - múltiplo, 316, 320–322
    - escalonamento de pipeline, 321–322
    - superescalar, 320
    - Ver também* despacho múltiplo
  - processadores streaming, 531, A-595
    - array (SPA), A-601, A-605
    - GeForce, 8800, A-608
  - processadores, 239–329
    - caminho de dados, 12
    - como cores, 31
    - comunicação de E/S com, 475–476
    - controle, 12
    - crescimento do desempenho, 32
    - definição, 9, 12
    - despacho duplo, 318
    - despacho múltiplo dinâmico, 316
    - despacho múltiplo estático, 316, 317–320
    - despacho múltiplo, 314, 316
    - especulação, 316–317
    - execução fora de ordem, 325, 393
    - ROP, A-576, A-601
    - streaming, 531, A-595
    - superescalar, 320, 321–322, 522
    - tecnologias para criação, 18–19
    - vetor, 525–528
    - VLIW, 317
  - produto, 187
  - programas
    - iniciando, 111–119
    - Java, iniciando, 117–119
    - linguagem assembly, 111
    - processamento paralelo, 512–515
    - traduzindo, 111–119
  - programas de processamento paralelo,
    - 512–515
    - definição, 510
    - dificuldade na criação, 512–515
    - para espaço de endereços compartilhado,
      - 516–517
    - para passagem de mensagens, 518–519
    - uso, 553
  - programas de shader gráficos, A-578
  - projeto
    - caminho de dados, 247
    - comprometimentos, 143
    - conjuntos de instruções de pipeline, 269
    - digital, 327–328
    - hierarquia de memória, desafios, 423
    - lógico, 244–247
    - sistema de E/S, 482–483
    - unidade de controle principal, 256–262
  - proteção
    - definição, 396
    - grupo, 485
    - implementando, 409–411
    - VMs para, 424
  - protocolo de handshaking, 471
  - protocolo snooping, 432–433
  - protocolos de invalidação de escrita, 432, 433
  - pseudo MIPS
    - conjunto de instruções, 227
    - definição, 226
  - pseudoinstruções
    - definição, 112
    - resumo, 112
  - Pthreads (POSIX threads), 539
  - PTX, instruções, A-592, A-593
  - push
    - definição, 89
    - usando, 91
- ## Q
- quad words, 136
  - quantidade de deslocamento, 77
  - Quicksort, 393, 394
  - quociente, 192
- ## R
- Radix sort, 393, 394, A-618–A-620
    - CUDA, código, A-620
    - implementação, A-618–A-620
  - RAID. *Ver* Redundant Arrays of Inexpensive Disks
  - raiz quadrada, instruções, B-689
  - Raster Operation (ROP), processadores, A-576, A-601
    - função fixa, A-601
    - GeForce, 8800, A-608
  - rasterização, A-608
  - redes crossbar, 536
  - redes locais (LANs)
    - definição, 18
    - Ver também* redes
  - redes multiestágio, 536
  - redes totalmente conectadas, 535, 536
  - redes, 17–18, 493–494
    - crossbar, 536
    - largura de banda, 535
    - locais (LANs), 18
    - multiestágios, 536
    - remotas (WANs), 18
    - totalmente conectadas, 535, 536
    - vantagens, 17
  - redução paralela, A-618
  - redução, 517
  - Redundant Arrays of Inexpensive Disks (RAID), 484–488
    - cálculo, 487
    - controlador PCI, 493
    - definição, 484
    - ilustração de exemplo, 485
    - popularidade, 484
    - RAID 0, 485
    - RAID 1 + 0, 488
    - RAID 1, 485
    - RAID 2, 485
    - RAID 3, 485



## Redundant Arrays of Inexpensive Disks (RAID) (*cont.*)

- RAID 4, 485–486
- RAID 5, 486
- RAID 6, 486
- resumo, 487
- referências
  - absolutas, 113
  - forward, B-636
  - não resolvidas, B-630, B-642
- referências absolutas, 113
- referências de forwarding, B-636
- referências não resolvidas
  - definição, B-630
  - link-editores e, B-642
- registrador Cause, 476
  - campos, B-654, B-655
  - definição, 311
  - ilustração, 477
- registrador contador, B-654
- registrador de controle do receptor, B-659
- registrador de dados do receptor, B-658, B-659
- registrador de status, 476
  - campos, B-654, B-655
  - ilustração, 477
- registrador salvo pelo callee, B-645
- registrador salvo pelo caller, B-645
- registradores
  - arquiteturais, 325
  - base, 66
  - Cause, 311, 476, 477, B-655
  - compilando a atribuição C com, 64–65
  - contador, B-654
  - convenção de uso, B-645, B-646
  - convenções MIPS, 96
  - definição, 63
  - destino, 78, 257
  - especificação numérica, 249
  - mapeando, 74
  - metade direita, 279
  - metade esquerda, 279
  - pipeline, 294, 295, 296, 298
  - ponto flutuante, 214
  - primitivos, 63–64
  - Receiver Control, B-659
  - Receiver Data, B-658, B-659
  - renomeando, 320
  - salvo pelo callee, B-646
  - salvo pelo caller, B-646
  - spilling, 67
  - Status, 311, 476, 477, B-655
  - tabela de página, 400
  - tempo de ciclo de clock, 64
  - temporários, 64, 90
  - Transmitter Control, B-659
  - Transmitter Data, B-659
  - variáveis, 64
  - x86, 136
- registradores arquiteturais, 325
- registradores de base, 66
- registradores de pipeline
  - antes do forwarding, 296

- dependências, 294, 295
- seleção de unidade de forwarding, 298
- registradores temporários, 64, 90
- reinício antecipado, 375
- replicação, 432
- representação em complemento a dois, 70, 71
  - atalho de extensão de sinal, 73–74
  - atalho de negação, 72–73
  - definição, 70
  - regra, 74
  - vantagem, 71
- reservas em standby, 487
- restaurações, 462
- resto
  - definição, 192
  - instruções, B-671
- restrição de alinhamento, 66
- retorno de exceção (ERET), 410
- RISC. *Ver* RISCs de desktop e servidor; RISCs embutidas; Reduced Instruction Set Computer (RISC), arquiteturas
- roofline, modelo, 539–546
  - benchmarking de multicóres com, 546–552
  - com áreas sobrepostas sombreadas, 545
  - com ceilings, 544, 545
  - com dois kernels, 545
  - desempenho de pico da memória, 541
  - desempenho de ponto flutuante máximo, 540
  - IBM Cell QS20, 548
  - ilustração, 541
  - Intel Xeon e5345, 548
  - kernel intensivo de E/S, 546
  - Opteron, gerações, 542
  - roofline computacional, 543
  - Sun UltraSPARC T2, 548
- rotina receber mensagem, 517
- rótulos
  - externos, B-636
  - globais, B-636
  - locais, B-636

## S

- saturação, 183–184
- SCALAPAK, 219
- segmentação, 398
- segmento de dados, B-638
- segmento de pilha, B-645
- segmento de texto, B-638
- seletores de dados, 244
- sem alocação de escrita, 376
- semicondutores, 34
- send message, rotina, 517
- serialização da escrita, 431–432
- servidores
  - custo e capacidade, 2
  - definição, 2
  - Ver também* RISCs de desktop e servidor
- setores, 464
- shaders
  - aritmética de ponto flutuante, A-578
  - definição, A-578
  - exemplo de pixel, A-579–A-580
  - gráficos, A-578
  - adição, 202
  - multiplicação, 205
  - significandos, 199
- silício
  - como tecnologia básica de hardware, 41
  - definição, 34
  - lingote de cristal, 34
  - wafers, 34
- SIMD (Single Instruction Multiple Data), 525, 533
  - arquitetura de vetor, 525–528
  - no x86, 525
  - vetor de dados, A-596
- simplicidade, 143
- sinais
  - ativados, 246
  - controle, 247, 257, 258, 259
  - desativados, 246
- sinais de controle
  - ALUOp, 257
  - bits múltiplos, 259
  - caminhos de dados em pipeline, 288
  - definição, 247
  - efeito, 258
- sinais desativados, 246
- sinal e magnitude, 198
- sincronização, 109–111
  - barreira, A-582, A-584, A-595
  - definição, 516
  - lock, 109
  - overhead, reduzindo, 33
  - unlock, 109
- sincronização de barreira, A-582
  - definição, A-584
  - para comunicação de threads, A-595
- sincronização de desbloqueio, 109
- sincronização de lock, 109
- Single Instruction Single Data (SISD), 524
- Single-Instruction Multiple-Thread (SIMT), 589t–651f
  - arquitetura de processador, A-590
  - definição, A-589
  - escalonamento de warp multithreaded, A-590
  - execução e divergência de warp, A-591
  - overhead, A-596
- Single-Program Multiple Data (SPMD), 524, A-585
- sistema de memória paralelo, A-596–A-600
  - acesso de load/store, A-601
  - caches, A-598
  - considerações de DRAM, A-597–A-598
  - espaços de memória, A-599
  - memória compartilhada, A-599–A-600
  - memória constante, A-600
  - memória de textura, A-600
  - memória global, A-599
  - memória local, A-600
  - MMU, A-598–A-599

- ROP, A-601
  - superfícies, A-601
  - Ver também* Graphics Processing Units (GPUs)
  - sistemas heterogêneos, A-570
    - arquitetura, A-572–A-574
    - definição, A-569
  - sistemas operacionais
    - armadilha de escalonamento do acesso ao disco, 496–497
    - definição, 6
    - encapsulamento, 15
  - Small Computer Systems Interface (SCSI), discos, 465, 494
  - software
    - camadas, 6
    - como serviço, 488, 553
    - driver GPU, 530
    - multiprocessador, 510
    - paralelo, 511
    - sistemas, 6
  - software de sistemas, 6
  - software paralelo, 511
  - somadores carry save, 190
  - Sort, procedimento, 120–125
    - alocação de registrador, 121
    - chamada de procedimento, 123
    - código para o corpo, 121–123
    - definição, 120
    - passando parâmetros, 124
    - preservando registradores, 124
    - procedimento completo, 124–125
    - Ver também* procedimentos
  - south bridge, 471
  - Sparse Matrix-Vector Multiply (SpMV), 549, 550, A-612, A-614, A-615
    - código serial, A-614
    - CUDA, versão, A-614
    - versão de memória compartilhada, A-616
  - SPEC
    - benchmark de CPU, 36–38
    - benchmark de potência, 38–39
    - SPEC2006, 228
    - SPECPower, 481
    - SPECrate, 537
    - SPECratio, 36
  - Special Function Units (SFUs), A-596
    - definição, A-602
    - GeForce, 8800, A-608
  - spilling de registradores, 67, 90
  - SPIM, B-659–B-663
    - chamadas do sistema, B-661–B-663
    - definição, B-659
    - iniciação, B-661
    - ordem de byte, B-661
    - recursos, B-661
    - simulação de máquina virtual, B-660
    - suporte a diretivas do montador MIPS, B-664–B-667
    - velocidade, B-660
    - versões, B-661
  - SPLASH/SPLASH 2 (Stanford Parallel Applications for Shared Memory), 537–539
  - stack pointers
    - ajuste, 91
    - definição, 89
    - valores, 91
  - stalls de uso de load, 302
  - stalls do pipeline, 271–272
    - definição, 271
    - evitando com reordenação de código, 271–272
    - hazards de dados, 298–301
    - inserção, 301
    - solução para hazards de controle, 273
    - uso de load, 302
  - stalls, 271–272
    - buffer de escrita, 383
    - como solução para hazard de controle, 273
    - definição, 271
    - esquema write-back, 383
    - evitando com reordenação de código, 271–272
    - hazards de dados, 298–301
    - inserção em pipeline, 301
    - memória, 385
    - uso de load, 302
  - Static Random Access Memories (SRAMs)
    - definição, 13
  - store word, 67
  - store, instruções
    - acesso, A-601
    - bloco, 133
    - compilando com, 67
    - condicional, 110–111
    - definição, 67
    - dependência de instrução, 298
    - detalhes, B-680–B-682
    - estágio EX, 283
    - estágio ID, 280
    - estágio IF, 280
    - estágio MEM, 284
    - estágio WB, 284
    - lista, B-680–B-682
    - ponto flutuante, B-689
    - registrador de base, 257
    - unidade para implementar, 251
    - Ver também* instruções load
  - Strcpy, procedimento, 99
    - definição, 98
    - ponteiros, 100
    - procedimento de folha, 100
    - Ver também* procedimentos
  - Stream, benchmark, 546
  - Streaming Multiprocessor (SM), A-607
  - Streaming SIMD Extension 2 (SSE2)
    - arquitetura de ponto flutuante, 222
  - strings
    - definição, 98
    - em Java, 100–101
    - representação, 98
  - striping, 485
  - subnormais, 219
  - subtração, 181–186
    - binária, 181–182
    - instruções, B-671–B-672
  - número negativo, 182
  - overflow, 182
  - ponto flutuante, 209, B-689–B-690
  - Ver também* aritmética
  - subtrilhas, 488
  - Sun Fire x4150, servidor, 488–493
    - conexões lógicas e larguras de banda, 491
    - ilustração de posterior/anterior, 490
    - memória mínima, 493
    - potência ociosa e máxima, 493
  - Sun UltraSPARC T2 (Niagara 2), 523, 532
    - básico *versus* totalmente otimizado
    - características, 546
    - definição, 546
    - desempenho, 551
    - desempenho do LBMHD, 550
    - desempenho SpMV, 550
    - ilustração, 547
    - modelo roofline, 548
  - supercomputadores, 2
  - superescalares
    - definição, 320
    - escalonamento de pipeline dinâmico, 321, 321–322
    - opções de multithreading, 522
  - superfícies, A-601
  - swap, espaço, 401
  - Swap, procedimento, 119–120
    - alocação de registradores, 119–120
    - código do corpo, 120
    - completo, 120, 121
    - definição, 119
    - Ver também* procedimentos
  - System Performance Evaluation Cooperative. *Ver* SPEC
- T**
- tabelas de histórico de desvio. *Ver* previsão de desvio, buffers
  - tabelas de página, 419
    - atualizando, 399
    - definição, 399
    - ilustração, 402
    - indexando, 400
    - invertidas, 402
    - memória principal, 404
    - níveis, 402–404
    - registrador, 400
    - técnicas de redução de armazenamento, 402–404
    - VMM, 426
  - tabelas de símbolos, 113, B-637, B-638
  - tabelas verdade
    - definição, 255
    - par bits de controle, 256
  - tags
    - definição, 369
    - localizando bloco, 389
    - tabelas de página, 401
    - tamanho, 391–392
  - tail call, 96
  - taxa de acerto, 366

taxa de clock  
definição, 23  
frequência comutada como função da, 30  
potência, 29

taxa de dados, 480

taxa de transferência máxima, 497

taxas de falhas locais, 393

taxas de falta  
cache de dados, 418  
cache repartido, 378  
definição, 366  
global, 393  
Intrinsity FastMATH, processador, 378  
locais, 393  
melhoria, 374  
origens de falta, 422  
tamanho de bloco versus, 375

taxas de perda globais, 393

telas gráficas  
LCD, 11  
suporte de hardware do computador, 12

Telsa PTX ISA, A-592–A-595  
instruções aritméticas, A-594  
instruções de acesso à memória, A-594–A-595  
instruções de thread de GPU, A-593  
sincronização de barreira, A-595

tempo de acerto  
definição, 366  
desempenho de cache, 385

tempo de busca, 464

tempo de execução  
como medida de desempenho válida, 41  
CPU, 22, 23, 24  
pipelining, 276

tempo de leitura de disco, 465

tempo de resposta, 20, 21

tempo de transferência, 465

tempo médio entre falhas (MTBF), 462

tempo médio para falha. *Ver* Mean Time To Failure (MTTF)

tempo médio para reparo. *Ver* Mean Time To Repair (MTTR)

terabytes, 2

Texture/Processor Cluster (TPC), A-606–A-607

thrashing, 416

threads  
criação, A-585  
CUDA, A-596  
gerenciando, A-591  
ISA, A-592–A-595  
latências de memória, A-628  
warps, A-589

TLB, falhas, 405  
handler, 414  
minimização, 550  
ocorrência, 411  
ponto de entrada, 414  
problema, 416  
tratamento, 411–416  
*Ver também* Translation-Lookaside Buffer (TLB)

topologias de rede, 534–537  
implementando, 536–537  
multiestágios, 537

tradução de endereço  
AMD Opteron X4, 435  
definição, 397  
Intel Nehalem, 435  
rápida, 404–406  
TLB para, 404–406

Transaction Processing (TP)  
benchmarks de E/S, 480–481  
definição, 480

Transaction Processing Council (TPC), 480

transistores, 19

Translation-Lookaside Buffer (TLB), 404–406  
associatividades, 405  
definição, 404  
ilustração, 404  
integração, 406–409  
Intrinsity FastMATH, 406  
valores típicos, 405  
*Ver também* TLB, falhas

Transmitter Control, registrador, B-659

Transmitter Data, registrador, B-659

tratadores de exceção, B-656–B-657  
definição, B-655  
retorno dos, B-658

trilhas, 464

troca atômica, 109

troca de contexto, 411

tubos de vácuo, 19

## U

underflow, 198

Unicode  
alfabetos, 100  
alfabetos de exemplo, 101  
definição, 100

unidades  
commit, 321, 324  
controle, 244, 254–255  
definição, 216  
detecção de hazard, 298, 300  
função especial (SFUs), A-596, A-602, A-608  
para implementação de load/store, 251  
ponto flutuante, 216  
rank, 488, 489

unidades de avaliação, 488, 489

unidades de commit  
buffer, 321  
definição, 321  
no controle de atualização, 324

unidades de controle, 244  
ilustração, 259  
principais, projetando, 256–262  
saída, 254–255  
*Ver também* Arithmetic Logic Unit (ALU)

unidades de detecção de hazard, 298  
conexões de pipeline, 300  
funções, 300

Uniform Memory Access (UMA), 515–516, A-574  
definição, 515  
multiprocessadores, 516  
usado menos recentemente (LRU)  
definição, 391  
estratégia de substituição de bloco, 420  
páginas, 402

## V

variáveis  
classe de armazenamento, 93  
estáticas, 93  
linguagem C, 93  
linguagem de programação, 64  
registrador, 64  
tipo, 93

variáveis estáticas, 93

varredura paralela, A-615  
baseada em árvore, A-618  
definição, A-615  
inclusiva, A-615  
modelo CUDA, A-617

varredura paralela baseada em árvore, A-618

vazão  
definição, 20  
despacho múltiplo, 323  
pipelining, 276, 323

Very Large-Scale Integrated (VLSI), circuitos, 19

Very Long Instruction Word (VLIW)  
definição, 317  
processadores, 317

Video Graphics Array (VGA), controladores, A-569

Virtual Machine Monitors (VMMs)  
atitude *laissez-faire*, 440  
definição, 424  
implementando, 439–441  
na melhoria do desempenho, 425  
requisitos, 424  
tabelas de página, 426

## W

wafers, 34  
defeitos, 34  
definição, 34  
dies, 34  
yield, 34

warps, 531, A-589

while, loops, 84–85

Wide Area Networks (WANs)  
definição, 18  
*Ver também* redes

words  
acessando, 65  
definição, 64  
duplas, 136  
load, 65, 67  
quad, 136  
store, 67

words duplas, 136  
write-back, estágio  
    instrução load, 281  
    instrução store, 283  
    linha de controle, 291  
write-through, caches  
    vantagens, 420  
    definição, 376, 420  
    divergência de tag,  
        377  
    *Ver também* caches

## X

X86, 133–141  
    codificação de instruções, 140–141  
    codificação do especificador do primeiro  
        endereço, 141  
    conclusão, 141  
    crescimento do conjunto de instruções, 143  
    evolução, 133–136  
    formatos de instrução, 141  
    instruções de ponto flutuante, 221

instruções/funções típicas, 140  
interconexões de E/S, 471–473  
linha de tempo histórica, 133–136  
modos de endereçamento de dados,  
    136, 138  
operações com inteiros, 136–140  
operações típicas, 141  
ponto flutuante, 220–222  
registradores, 136  
SIMD, 525  
tipos de instruções, 137