

Processador ARM Cortex-A9

Erick Nogueira Nascimento (032483)

Franz Pietz (076673)

Lucas Watanabe (134068)

11 de junho de 2012

- Alto desempenho e economia de energia
- Arquitetura ARMv7-A
- Características do processador
 - super-escalável multi-issue
 - pipeline de 8 estágios com especulação
 - execução out-of-order
 - caches L1 com tamanho configurável
 - versões single e multicore
- Aplicações: celular, tablets, video-games, etc.

Versões

- Cortex-A9: single core
- Cortex-A9 MPCore: multicore (2 ou 4 núcleos)

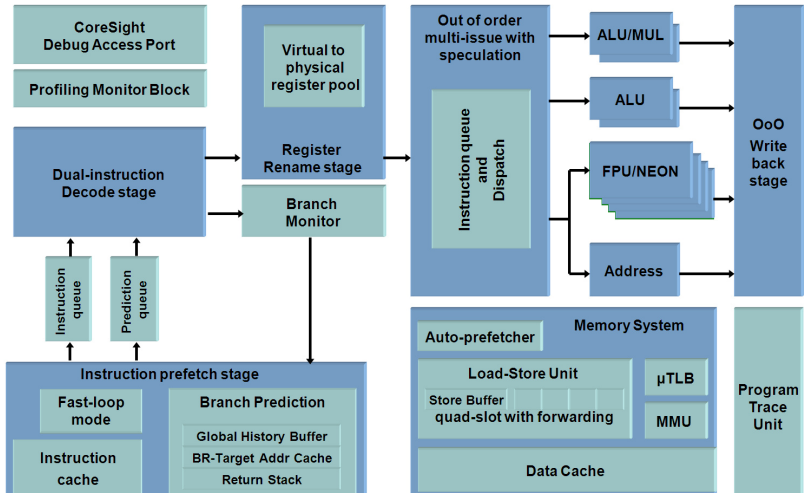
Detalhes

- Superescalar com escalonamento dinâmico
- Multi-issue
- Pipeline de 8 estágios com especulação
- Out-of-order
- Cache L1 é quatro-associativa de 16, 32 ou 64KB

Consumo de Energia

- Clock variável (Apple A5)
- Desligamento de núcleos (OMAP4400 da Texas Instruments)

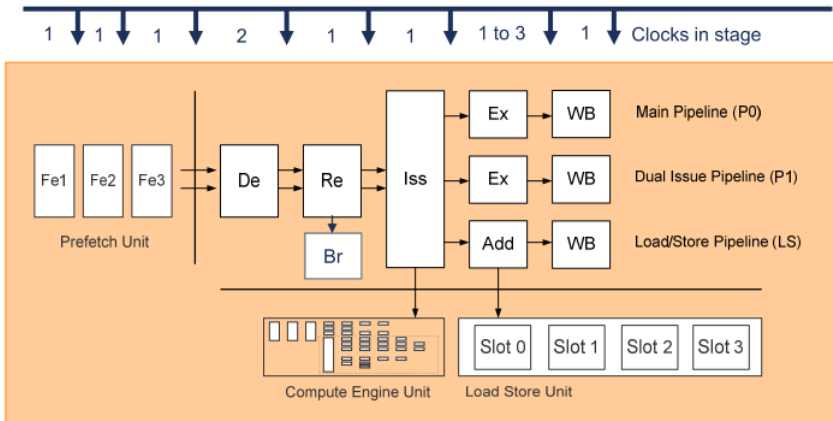
Cortex-A9 Single Core - Arquitetura



Pipeline

- Avançado processamento de fetch e predição de branch
- Prefetching de até quatro instruções
- Fast-loop
- Pipeline out-of-order
- Executa até duas instruções aritméticas completas e duas de load-store,
- Resolve um branch a cada ciclo.
- Em média
 - instruções comuns: 9 ciclos
 - instruções complexas: 11 ciclos

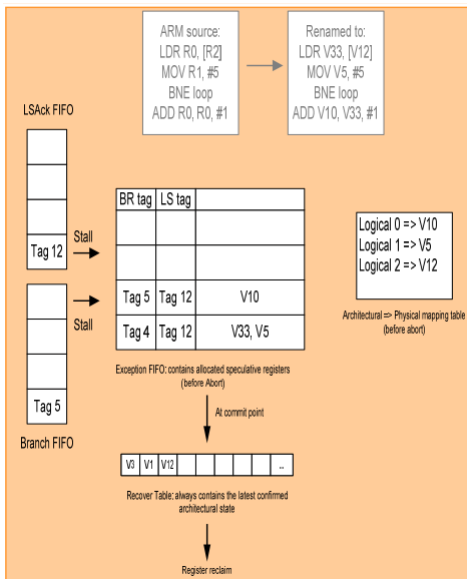
Cortex-A9 Single Core - Pipeline



Renomeação de registradores

- Explícita: mapeamento dos 32 registradores arquiteturais (visíveis ao programador) para os 56 registradores físicos (não visíveis ao programador).
- Vantagens:
 - remoção dos hazards WAW (write-after-write) e WAR (write-after-read) no pipeline.
 - desdobramento completo de loops pequenos por hardware.

Cortex-A9 Single Core - Renomeação



Hierarquia de memória

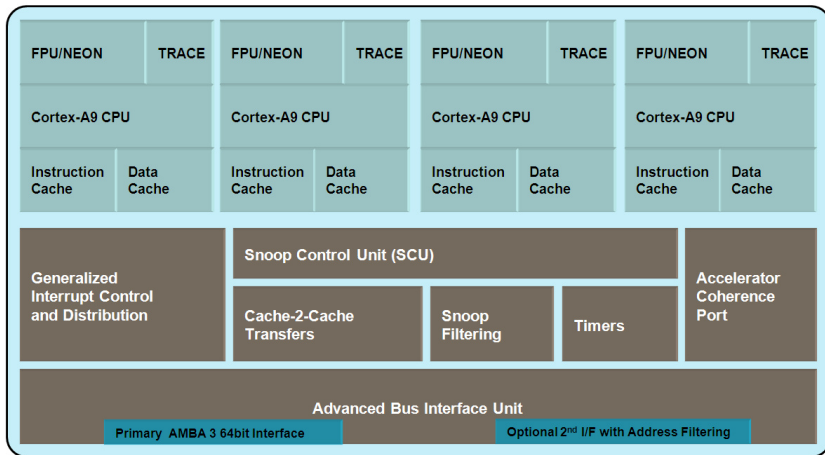
- Cache L1 dividida em I e D (Instruções e Dados)
- 16KB, 32KB ou 64KB
- Cache L2: controlador externo (PrimeCell PL310)
- Razão síncrona de $\frac{1}{2}$ clock, reduzindo a latência
- Até 8 MB
- 4-associativa e 16-associativa
- Suporta paridade e memórias RAMs com correção de erros (ECC)

TLB

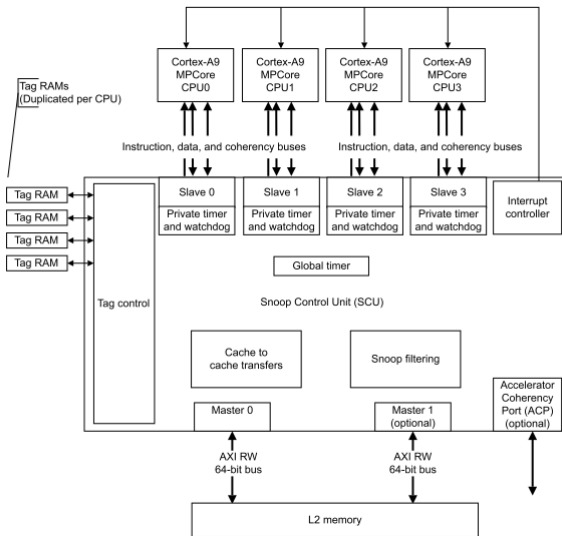
- dois níveis
- Micro TLB
 - duas TLBs, I e D, a primeira tem 32 entradas, e a segunda tem 32 ou 64 entradas.
 - completamente associativa
 - consulta sua tabela de endereços virtuais para endereços físicos em um único ciclo de clock.
- Main TLB
 - array de 4 elementos completamente associativo
 - estrutura 2-associativa com 2x32, 2x64, 2x128 e 2x256 entradas.

- Cada processador pode ter uma quantidade de cache L1 específica, FPU ou qualquer outro componente diferente do outro núcleo
- Snoop Control Unit: responsável pela interconexão, como árbitro, comunicação, transferência entre os caches e a memória e todas as funções necessárias para garantir o funcionamento coerente e correto
- Accelerator Coherence Port (ACP) para coerência de cache

Cortex-A9 MPCore - Arquitetura



Cortex-A9 MPCore - Snoop Control Unit



Gerenciamento de Energia

- Previsão de desvio e de retorno de função precisos
- Uso de caches fisicamente endereçáveis
- Uso de micro TLBs (um dos dois níveis de TLB)
- Loops de instrução que são menores do que 64B podem ser executados sem acessos adicionais a cache (além do misses compulsórios)
- Modos: Run Mode; Standby, Dormant, Shutdown
- Cada núcleo pode estar em um modo de energia diferente

Advanced Bus Interface Unit

- MPCore
- Até 16 transações de barramento pendentes por processador (núcleo)
- Transferências de cache L1 para cache L1 diretas
- Suporte a razão de barramento de meio-clock, síncrona e de baixa latência

Características gerais

- Antecessor do Cortex-A9
- Vendido como IP:
 - hard IP cores: alto desempenho (tempo ou energia), baixa flexibilidade de configuração
 - soft IP cores: desempenho em geral inferior ao hard, maior flexibilidade
- Duas instruções por ciclo de clock
- Clock rate máximo de 1GHz

Hierarquia de Memória - Cache L1

- Par de caches, I e D (Instrução e Dados)
- 16KB ou 32KB, com blocos de 64B
- 4-associativa
- Way-prediction
- Random replacement
- Virtualmente indexada
- Fisicamente tagged
- Latência: 1 ciclo para hazards do tipo load-use.

Cache L2

- Opcional
- 128KB a 1MB, com blocos de 64B
- Organizada em 1 a 4 bancos, permitindo transferências concorrentes
- Fisicamente indexada e tagged
- O barramento entre cache L2 e memória pode ser de 64 ou 128bits

TLB

- Um nível, e é composta por duas TLBs, I e D
- Completamente associativas, com 32 entradas
- Política de substituição round-robin
- Tamanhos de página para o sistema de memória virtual: 4KB, 16KB, 64KB, 1MB ou 16MB.

Pipeline

- 2 instruções por ciclo
- Superscalar
- Escalonamento estático
- CPI médio de 2 (pipeline stall cycles são o fator preponderante)
- Previsão de desvio em dois níveis
- Branch target buffer com 512 entradas, história global de 4K
- Return stack com 8 entradas

Comparação

- Escalonamento estático X escalonamento dinâmico
- Cortex-A9 é mais rápido por um fator de 1.28 (em média)
- As medidas (simulações) de desempenho de ambos os processadores foram realizadas com a versão single-core do Cortex-A9 com a mesma taxa de clock e o mesmo tamanho de cache
- A cache L1 do A8 é de 32KB e a L2 tem 1MB, sendo 8-associativa no A8 e 16-associativa no A9
- Melhor desempenho do Cortex-A9 devido ao escalonamento dinâmico, pipeline de multiplicação mais sofisticado, além de um previsor de desvio, uma cache L1 não bloqueante e um prefetch de instrução mais poderosos.

Evolução desde o ARM1136 até o Cortex-A9

- Single-issue, in-order, single bus transaction, no L2
 - Nokia N810 using OMAP2420 with **ARM1136** at 400MHz \Rightarrow 500 DMIPS
- Single-issue, in-order, multiple decoupled L2 transactions
 - NVIDIA Tegra with **ARM11 MPCore** at 600-800MHz \Rightarrow 1000 DMIPS
- Dual-issue, in-order, multiple integrated L2 transactions
 - T.I. OMAP3430 with **Cortex-A8** at 600-800MHz+
 - (previously demo'ed at 1GHz) \Rightarrow 2000 DMIPS
- Multi-issue, out-of-order, multi-out-of-order bus transaction
 - **Cortex-A9** multicore capable \Rightarrow 8000+ DMIPS