ARM Cortex-A* Series Processors

Haoyang Lu, Zheng Lu, Yong Li, James Cortese
ARM Cortex-A* Series Processors

- Applications
- Instruction Set
- Multicore
- Memory Management
- Exclusive Features
ARM Cortex-A* series: Applications

Ford Sync

Digital TV

Networking solutions
ARM Cortex-A* series: Applications

Smartphones and Tablets
## ARM Cortex-A* series: Applications

Processors of the Cortex-A series and their applications:

<table>
<thead>
<tr>
<th></th>
<th>A5</th>
<th>A7</th>
<th>A8</th>
<th>A9</th>
<th>A15</th>
<th>A53</th>
<th>A57</th>
</tr>
</thead>
<tbody>
<tr>
<td>Smart phones</td>
<td>*</td>
<td></td>
<td>*</td>
<td>*</td>
<td></td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>Home Computing</td>
<td>*</td>
<td></td>
<td>*</td>
<td>*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Smart TVs</td>
<td>*</td>
<td></td>
<td>*</td>
<td>*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Digital Cameras</td>
<td>*</td>
<td></td>
<td></td>
<td>*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Embedded Computing</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>Home Networking</td>
<td>*</td>
<td></td>
<td></td>
<td>*</td>
<td></td>
<td></td>
<td>*</td>
</tr>
<tr>
<td>Storage</td>
<td>*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
ARM: Instruction Set

• Two instruction set:
  - ARM instruction set (32-bit)
  - Thumb instruction set (mixed 16/32 bit)

• Thumb-2: Bit-field manipulation, table branches and conditional execution

• Unified Assembly Language (UAL): supports generation of either ARM and Thumb instructions from the same source code
ARM Cortex-A8 series: Pipeline

- Dual-issue
- Statically scheduled superscalar
- Dynamic issue detection – issue two instructions per clock
- Dynamic branch predictor – 512 entry branch target buffer
  - 4K-entry global history buffer
  - Mispredict penalty: 13 cycles
ARM Cortex-A8 series: Pipeline

13-stage pipeline
ARM Cortex-A8 series: Pipeline

5-stage Instruction Decode
ARM Cortex-A8 series: Pipeline

Instruction decode execution
ARM Cortex-A8 series: Pipeline

- Ideal CPI is 0.5 according to its dual-issue
- Stalls:
  - Functional hazards, which occur when two instructions selected for issue simultaneously use the same functional pipeline.
  - Data hazards, which are detected early in the pipeline and may stall either both instructions
  - Control hazards, which arise only when branches are mispredicted, the penalty is 13 cycle.
Arm Cortex-A series: Multicore

Cortex™-A9

ARM CoreSight™ Multicore Debug and Trace

ARMv7 32b CPU

NEON™ Data Engine

Floating Point Unit

16-64k I-Cache

16-64k D-Cache

Core 1 2 3 4

ACP

SCU

Dual 64-bit AMBA3 AXI
Multicore configurations are controlled and managed by the Snooping Control Unit (SCU). The SCU makes sure that level 1 cache coherence is achieved. Additional levels of coherence are achieved with an Accelerator Coherence Port (ACP).
Arm Cortex A series: Multicore

- big.LITTLE technology: a powerful processor is paired with a less powerful processor;
  eg. A15 and A7, or the A53 and A57
Arm Cortex A series: big.LITTLE

The processing is divided between the two processors to achieve increased efficiency but no decrease in performance.
Memory: A8 VS Intel i7

- **Cortex-A8**

<table>
<thead>
<tr>
<th></th>
<th>Size</th>
<th>Associativity</th>
<th>Latency</th>
</tr>
</thead>
<tbody>
<tr>
<td>L1</td>
<td>16 / 32KB</td>
<td>4-way</td>
<td>Two words per cycle</td>
</tr>
<tr>
<td>L2</td>
<td>0 / 128 / 256 / 512 / 1024KB</td>
<td>8-way</td>
<td></td>
</tr>
<tr>
<td>TLB</td>
<td>32</td>
<td>Fully associative</td>
<td></td>
</tr>
</tbody>
</table>

- **Intel i7**

<table>
<thead>
<tr>
<th></th>
<th>Size</th>
<th>Associativity</th>
<th>Latency</th>
</tr>
</thead>
<tbody>
<tr>
<td>L1</td>
<td>32KB</td>
<td>4-way I</td>
<td>4 cycles, pipelined</td>
</tr>
<tr>
<td></td>
<td></td>
<td>8-way D</td>
<td></td>
</tr>
<tr>
<td>L2</td>
<td>256KB</td>
<td>8-way</td>
<td>10 cycles</td>
</tr>
<tr>
<td>L3</td>
<td>2MB per core</td>
<td>16-way</td>
<td>35 cycles</td>
</tr>
<tr>
<td>ITLB</td>
<td>128</td>
<td>4-way</td>
<td>1 cycle</td>
</tr>
<tr>
<td>DTLB</td>
<td>64</td>
<td>4-way</td>
<td>1 cycle</td>
</tr>
</tbody>
</table>
Cortex-A8 Features

• L1 Caches
  ➢ physically tagged, and virtually indexed for instruction and physically indexed for data
  ➢ fixed line length of 64 bytes
  ➢ two words per cycle
  ➢ parity error detection

• L2 Cache
  ➢ physically indexed and tagged
  ➢ fixed line length of 64 bytes
  ➢ programmable preloading engine
  ➢ parity detection on the tag arrays
  ➢ Error Correction Code on data arrays
  ➢ partitioned into multiple banks to enable parallel operations
Cortex-A8 Features

Structure of L2 Cache
Cortex-A8 Performance

- simulated with 32 KB primary caches and a 1 MB eight-way set associative L2 cache using the integer Minnespec benchmarks
- instruction cache miss rates are close to zero for most and under 1% for all of them
- For the data cache test, there are significant L1 and L2 miss rates
Intel i7 Features

- L1 instruction cache, L1 data cache, and a L2 cache in each core
- Support up to three memory channels of bandwidth over 25 GB/sec
- 48-bit virtual addresses and 36-bit physical addresses, a maximum physical memory of 36 GB
I7 Level 1 Data Cache Features

- a write-back write-allocate cache
- Store Forwarding - forward data directly from the store operation to load
- Memory Disambiguation - predict that a load does not depend on a preceding store
- Data Prefetching
• evaluated by 19 of the SPECCPU2006 benchmarks

• L1 instruction cache miss rate varies from 0.1% to 1.8%, averaging just over 0.4% - Since the i7 does not generate individual requests for single instruction units, but instead prefetches 16 bytes of instruction data (between four and five instructions typically).

• L1 data cache misses are shown in two ways:
  ➢ relative to the number of loads that actually complete - graduation
  ➢ relative to all the L1 data cache accesses from any source.

• the miss rate when measured against only completed loads is 1.6 times higher (an average of 9.5% versus 5.9%)
Intel i7 Performance

The graph shows the performance metrics for various benchmarks on Intel i7 processors. The x-axis represents different benchmarks, and the y-axis shows the percentage of L1 D misses and L1 D cache references. The performance metrics are compared with L1 D misses graduated loads.
ARM Exclusive Features - NEON

• NEON technology is used in ARM Cortex™-A series processors to enhance user’s multimedia experiences.

• It can highly enhance the multimedia and signal processing algorithms which are frequently required by multimedia applications
ARM Exclusive Features - NEON

- The Advanced SIMD instructions perform packed SIMD operations:
  - Registers are considered as vectors of elements of the **same data type**.
  - Instructions perform the **same operation** in all lanes.
ARM Exclusive Features - NEON
ARM Exclusive Features - VFP

• ARM Floating Point architecture (VFP) provides hardware support for floating point operations ARM Cortex™-A series processors.

• VFP architecture v3 is an enhancement to v2:
  - Double the double-precision registers
  - Instructions of fixed-point and floating-point conversion
ARM Exclusive Features - VFP

• ARM Cortex-A8 has a cut down VFPLite module instead of a full VFP module, and require roughly ten times more clock cycles per float operation
Thanks