ARM Cortex® Processors

…driving the pace of multicore innovation

Chris Turner
ESSEI TecDay, October 13th 2015
Markets for ARM Processor IP

Mobile / Consumer / Wearables

Enterprise / Networking / Server

IoT / Embedded

Automotive / Industrial
ARM Cortex® Processor Profiles

Three architecture variants profiled for the different application sectors

- **Cortex-A processors**
  - Computation, robotics, computer-vision
  - Higher performance
  - Cortex-A processors
  - ARMv8-R
  - Linux®, QNX

- **Cortex-R processors**
  - Actuation, fast control
  - Extended Functional Safety
  - Fast response / Real-time control
  - ARM Cortex-R processors
  - DSP
  - RTOS

- **Cortex-M processors**
  - MCUs, IoT, sensors, motors
  - Smallest footprint / lowest power
  - Cortex-M processors
  - ARM DSP

---

- **Actuation, fast control**
- **Extended Functional Safety**
- **Fast response / Real-time control**
- **MCUs, IoT, sensors, motors**
- **Smallest footprint / lowest power**
Cortex Architecture Profiles

- **Cortex-A**, ARMv7A and v8A: Lower power, smaller area
- **Cortex-R**, ARMv7-R and v8-R: Higher performance
- **Cortex-M**, ARMv6-M, v7-M and v8-M: Lower power, smaller area

**Operating System**
- Cortex-A, Cortex-R: Linux/Rich OS
- Cortex-M: RTOS only

**Instruction set**
- Cortex-A, Cortex-R: 32/64b ARM and Thumb ISA
- Cortex-M: 32b Thumb ISA

**Interrupts**
- Cortex-A: SW managed interrupts
- Cortex-R: Deterministic SW managed
- Cortex-M: HWV managed interrupt

**CPU Memory**
- Cortex-A: Multiple level cache
- Cortex-R: Caches including TCMs
- Cortex-M: TCMs in Cortex-M7

**Safety support**
- Cortex-A: ASIL B capable
- Cortex-R: ASIL D capable
- Cortex-M: ASIL-D capable
Processor Clusters

- Cortex-A processor cores can operate in a coherent cluster up to MP4
  - Synthesis-time choice of cores per cluster
- Each core runs a process thread
  - e.g. Linux/Android kernel has support built in
- ARM’s Generic Interrupt Controller (GIC) distributes to the cores
  - OS re-programs distribution on-the-fly
  - Provision for inter-core interrupts
- Automated data cache coherency
  - Snoop Control Unit (SCU) includes level 2 cache, tag RAM copies and Accelerator Coherency Port (ACP)
Cortex-A72: High-end MP4 Cluster

- Highest performance core from ARM – applicable to enterprise infrastructure, automotive, mobile, consumer and beyond
- Significant boost in power and area efficiency
- Highly scalable across many different segments and price points
- Different Implementation possibilities enable optimal solutions for different markets
- Wide range of core counts possible through advanced interconnect (CCN/CCI)
- Advanced ARMv8-A feature set: 64 bit, ECC, AMBA 5 CHI, high performance FP and cryptography
- Safety documentation package support for Automotive and Industrial markets
Cortex-A CPU Portfolio

All can be configured as 1, 2, 3 or 4 MPCore clusters

High Performance
‘big’

ARMV7-A
High performance 32bit CPU with enterprise class feature set

ARMV7-A
Premium performance with mid-range area & power

ARMV8-A, 64bit
High single thread performance CPU

ARMV8-A, 64bit
Highest single thread performance CPU

Cortex-A9

High Efficiency
‘LITTLE’

Cortex-A5
Smallest & lowest power v7-A CPU

Cortex-A7
Highest efficiency V7-A CPU

Cortex-A53
Highest efficiency V8-A CPU, 64bit
ARM big.LITTLE Technology

Saving yet more energy by using the right core for the right task

- Heterogeneous Computing
  - More than 40% higher User Experience*
  - 45% to 65% CPU power savings**

- Architecturally Identical Processors
  - High performance tuned ‘big’ cores
  - High efficiency tuned ‘LITTLE’ cores

- Hardware Coherency
  - Automatically managed for all cache levels

- Seamless & Automatic Task Allocation
  - Global Task Scheduling (big.LITTLE MP)

* Compared to LITTLE-only platforms; ** Compared to big-only platforms
† Average power across high-end (Epic Citadel) gaming and low-utilisation (Audio playback) workloads
Fine Grain Market Segmentation through different CPU combinations

SoC Configuration

Device Tier

- Octacore big.LITTLE
- Hexacore big.LITTLE
- Quadcore big.LITTLE
- Octacore
- Quad core
- Dual core
- Single core

Latest features, advanced spec

Lowest Power & footprint

big core – A15, A17, A57, A72
LITTLE core – A7, A53
big.LITTLE employs AMBA Coherency Extensions

**First Generation big.LITTLE**
- All coherency snoops sent to all processors

**Next Generation System Coherency**
- Integrated Snoop Filter

**Higher Efficiency and Performance**
- One central snoop vs many
- Lower snoop latency
Extensible Architecture for Heterogeneous Multi-core Solutions

- Up to 4 cores per cluster
- Up to 12 coherent clusters
- Integrated L3 cache
- Up to Quad channel DDR3/4 x72
- Virtualized Interrupts
- Heterogeneous processors – CPU, GPU, DSP and accelerators
- Up to 24 I/O coherent interfaces for accelerators and I/O

CoreLink™ CCN-512 Cache Coherent Network

- I-32MB L3 cache
- Snoop Filter
- Peripheral address space

Memory Controllers (DMC-520)
- x72 DDR4-3200

Network Interconnect
- NIC-400
- SATA
- USB
- PCIe
- SATA
- USB
- PCIe

Cortex CPU or CHI master

GIC-500

AHB Snoop Filter

I/O Virtualisation CoreLink MMU-500

Peripheral address space

Virtualized Interrupts

Heterogeneous processors – CPU, GPU, DSP and accelerators
Cortex-A processors combined in big.LITTLE clusters deliver high performance and save energy
ARM Cortex-R and Cortex-M Processor Portfolio

**Cortex-R**

- **Cortex-R7**: High performance 4G modem and storage
- **Cortex-R4**: Real-time standard
- **Cortex-R5**: Functional safety package

**Cortex-M**

- **Cortex-M4**: Mainstream Control & DSP
- **Cortex-M7**: Maximum Performance Control & DSP
- **Cortex-M0**: Low power with maximum cost efficiency
- **Cortex-M0+**: Highest energy efficiency
- **Cortex-M3**: Performance efficiency
What Makes a Real-Time Processor Microarchitecture

Typically Cortex-R5

- Superscalar / dual issue execution throughput
  - Cache line buffers minimise stalling while waiting for L2 memory system
  - ECC and its RMW timing is mostly transparent

- Low Interrupt Latency pipeline mode
  - Fast interrupt response abandons any pending and re-startable memory operations

- Tightly Coupled Memory
  - Level-1 memory system for fast access to code and data, e.g. Interrupt Service Routines

- Low Latency Peripheral Port
  - Introduced in Cortex-R5
  - Direct paths to LSU and store queue avoid delays in caches and AXI-Main

![Diagram of processor microarchitecture](image)
Real-Time Memory Address Map

**VMSA - PMSA**

- **Cortex-A**
  - For applications
  - Single flat memory
  - Virtual memory, MMU
  - Address translation
  - TLB acceleration
  - Page table manager
  - Memory protection
  - Program relocation
  - ‘Open’ systems

- **Cortex-R**
  - For real-time
  - Specialised memories e.g. TCM, LLRAM, Main
  - Position-dependent code
  - Deterministic behaviour
  - Memory protection
  - ‘Composed’ systems
Cortex-R5 Processor

Dependable and proven real-time performance

- High performance 1.68 DMIPS/MHz
  - 8-stage dual-issue pipeline, pre-fetch, branch prediction
  - HDIV, SIMD, SP/DP FPU, ARMv7R Thumb2/ARM instructions
- Deterministic response to hard deadlines
  - Low Interrupt Latency microarchitecture
  - Tightly Coupled Memory (TCM)
- Reliable with fault detection and control features
  - MPU, ECC/Parity on L1 memories, Dual-Core Lock-Step
  - ECC and Parity also on AXI bus port interfaces
  - Support for safety-related applications
- Cost Effective – synthesis configurable for optimum PPA
- Low Latency Peripheral Port (LLPP)
  - Non-blocking access to I/O registers and GIC
- Accelerator Coherency Port (ACP)
  - Performance boosting data cache maintenance
Dual-core Cortex-R5

Support for safety-related applications

- Two cores can be used either in lock-step or ‘performance’ mode
  - In performance mode both cores act as bus masters
  - In lock-step one core provides a redundant copy whilst a single instance of the cache and TCM RAMs is protected with ECC

- i/o coherency but no inter-processor coherency
  - An external data source can write through the SCU and coherency is maintained simply by invalidating cache lines holding addresses being written
  - Such hardware automated cache maintenance is very beneficial in many real-time applications
Cortex-R5 Fault Detection & Control Features

Error Correcting Code, Cache & TCM
- Single Error Correct – Double Error Detect
- 64-bit scheme is most efficient for I-side
- 32-bit scheme is best for D-side to minimize Read-Modify-Write cycles
- RMWs to re-calculate ECC when writing a quantity smaller than memory chunk size
- RMWs performed automatically with minimal, or even zero, performance impact

Bus ECC
- ECC and Parity are generated, detected and corrected
- Interconnect 'veneered' with same ECC/P functionality

Hard Errors in Cache and TCM
- Hard errors cannot be corrected by writing back corrected data and repeat when memory is read again
- 'Live-lock' scenario when uncorrected instructions or data are continuously re-fetched

ARMv7-R Architecture
- Protected Memory System Architecture. Precise aborts

Dual Core Lock Step
- Both spatial (also orientation) and temporal separation
- Avoiding common cause failures, i.e. reduced probability of both CPUs seeing the same failure at the same time and still checking OK
Research Project: TCLS ARM for Space

- Collaborative project funded by European Commission H2020 Space program
  - Start in Feb 2015 with a two year duration

- Project Objectives
  - Investigate feasibility of a fail-functional ARM CPU using the triple core lockstep (TCLS) principle
  - Target rad-tolerant space and safety-critical terrestrial applications
  - Assess the fail functional design using rad-tolerant STM65nm technology

- Concept
  - Three ARM CPUs execute in lockstep
  - Fail functional – Resynchronize upon divergence
  - Shared ICache, DCache and memory

http://www.tcls-arm-for-space.eu/
The latest Cortex-M7 processor from ARM

- High-performance processor with DSP capabilities
  - Six-stage superscalar pipeline with powerful DSP and SP/DP FP
  - Best-in-class core for high-end MCU or replace MCU+DSP

- Flexible powerful memory system
  - Tightly-Coupled Memories for real-time determinism
  - 64-bit AXI AMBA4 memory interface with I and D cache
  - Next-gen MCU with more memories and peripherals

- ARMv7-M architecture and CMSIS support
  - 100% binary compatibility from Cortex-M4
  - Cortex-M family ease-of-use and very low interrupt latency
  - Reuse code and system design from existing products

- Fault detection and control features
  - MPU, memory ECC (SEC-DED), on-line MBIST, DCLS
Thank You

The trademarks featured in this presentation are registered and/or unregistered trademarks of ARM Limited (or its subsidiaries) in the EU and/or elsewhere. All rights reserved. Any other marks featured may be trademarks of their respective owners.