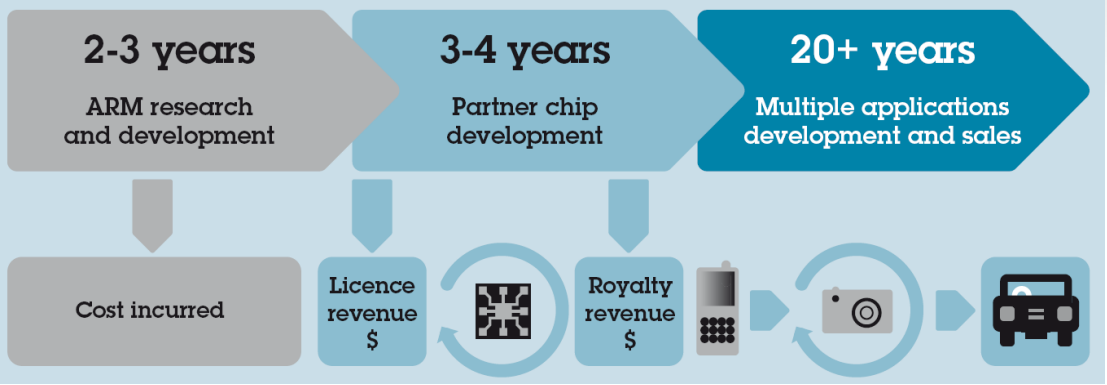
Integrating CPU and GPU, The ARM Methodology

Edvard Sørgård, Senior Principal Graphics Architect, ARM
Ian Rickards, Senior Product Manager, ARM

The ARM Business Model

- Global leader in the development of semiconductor IP
  - R&D outsourcing for semiconductor companies
- Innovative business model yields high margins
  - Upfront license fee – flexible licensing models
  - Ongoing royalties – typically based on a percentage of chip price
  - Technology reused across multiple applications
- Long-term, secular growth markets



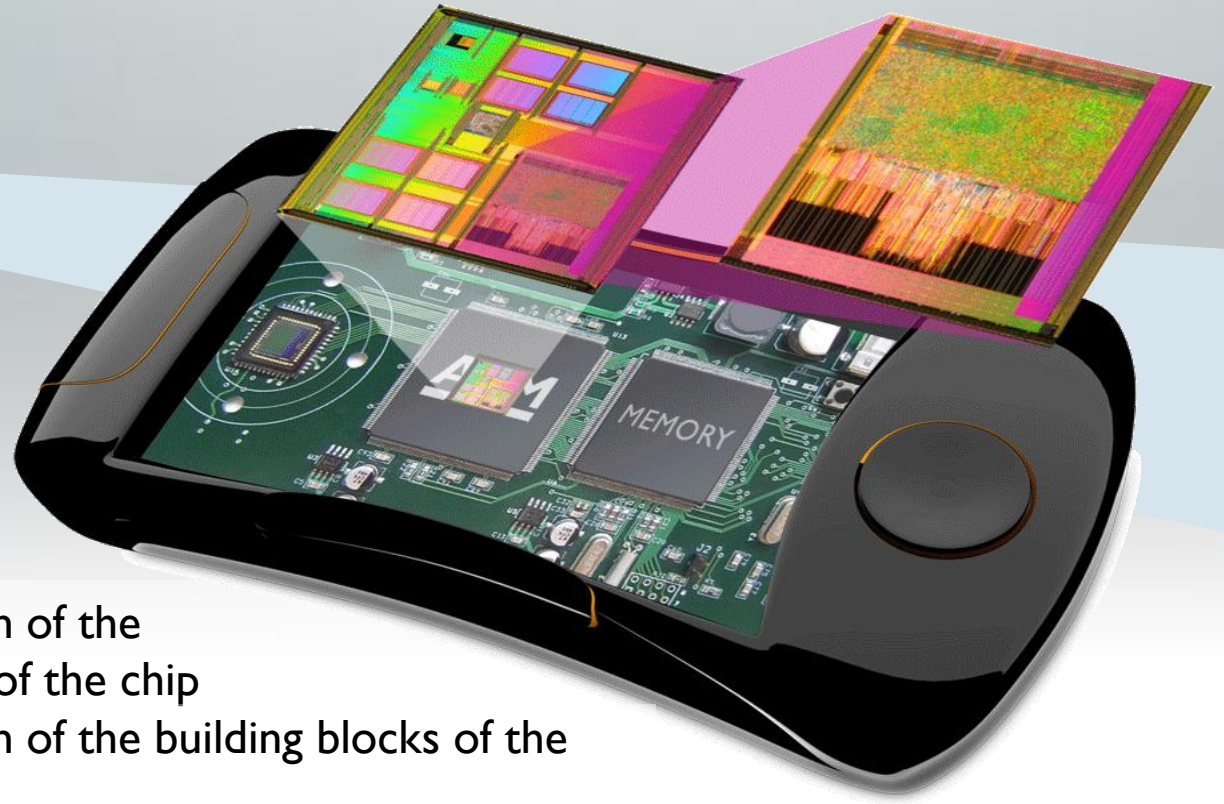
Approximately 850 licenses
Grows by 60-90 every year

More than 250 potential royalty payers

>8bn ARM-based chips in '11
>25% CAGR over last 5 years

ARM Technology

- Advanced consumer products are incorporating more and more ARM technology – from processor and multimedia IP to software



Processor IP – Design of the brain of the chip

Physical IP – Design of the building blocks of the chip

Software development tools

ARM® Mali™ GPU Momentum

#1 graphics in Android™ tablets (>50% market share)

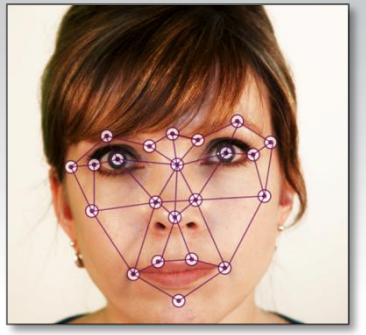
#1 in Smart TVs (>70% market share)

>20% Android™4 smartphones

Mali GPU shipments outpacing industry growth rate & gaining market share

Comprehensive GPU Compute Support

- ARM's best-in-class CPU know-how combined with expertise in graphics technology enabling complex use-cases
  - Computational photography: Panorama stitching
  - Image recognition: Face, smile, landmark, context
  - Image improvement, stabilization, editing, filtering



- By moving GPU Compute tasks onto the GPU will enable lower power consumption and faster response over being solely run on the CPU

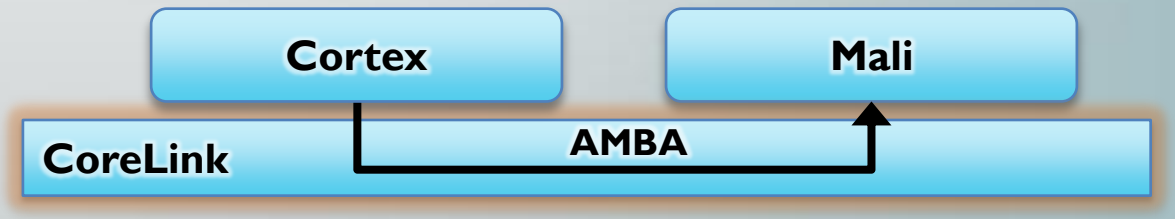
Mali GPU Compute: No FUD... Facts

- Passed Khronos Conformance
  - OpenCL™ 1.1 Full Profile on Linux and Android™
- Proven in Silicon
  - Samsung Exynos 5 Dual, implements Full Profile
  - OpenCL and Renderscript DDK available now
- Mali-T604 shipping in real products
  - Google Chromebook
  - Google Nexus 10
  - InSignal Arndale Community Board
- API exposed for developers
  - OpenCL on Linux for Arndale platform
  - Renderscript computation on Android for Nexus 10

Compute Use Case Example

- ARM Seemore demo
- OpenCL 1.1 FP accelerated world
  - Interactive items and lights
  - Bullet physics broad-phase fully OpenCL accelerated on GPU
- Performance boost
  - GPU Kernel speedup >10x
  - But system speedup is less
- ARM integration goal
  - *Take the system cost out!*

Integration: Coherency



- SoCs are heterogenous systems
- But sharing data can still be costly
  - Cache flushes, locks, syncs reduces the heterogeneous benefit

- HW coherency makes sharing data cheap and automatic
- ARM is in leading position with full technology coverage
  - Cortex™ CPUs
  - Mali GPUs
  - CoreLink™ system IP
  - AMBA™ bus protocols

Integration: Address Space Alignment

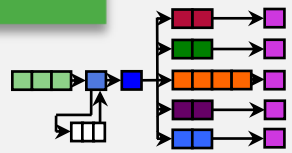
- The 32-bit address space is running out, even in mobile
- Midgard architecture built for full 64-bit addresses
- Embedded distributed Mali MMU for VA to PA/IPA translation
  - Mali-T604: 48-bit VA and 40-bit PA/IPA
  - Uses ARMv7 LPAE page table format, just like Cortex-A15 & Cortex-A7
- Multiple simultaneous address spaces supported
  - Mali GPUs run many threads in parallel
  - Independent processes may execute on GPU simultaneously
  - Seamless process transitions ensures maximum utilization/efficiency

Fine-Tuned to Different Performance Points

LITTLE

Most energy-efficient applications processor from ARM

- Simple, in-order, 8 stage pipelines
- Performance better than mainstream, high-volume smartphones (Cortex-A8 and Cortex-A9)



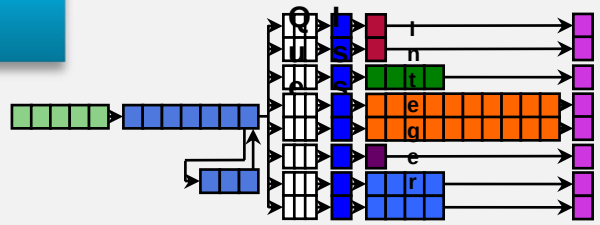
Cortex-A7

Cortex-A53

big

Highest performance in mobile power envelope

- Complex, out-of-order, multi-issue pipelines
- Up to 2x the performance of today's high-end smartphones



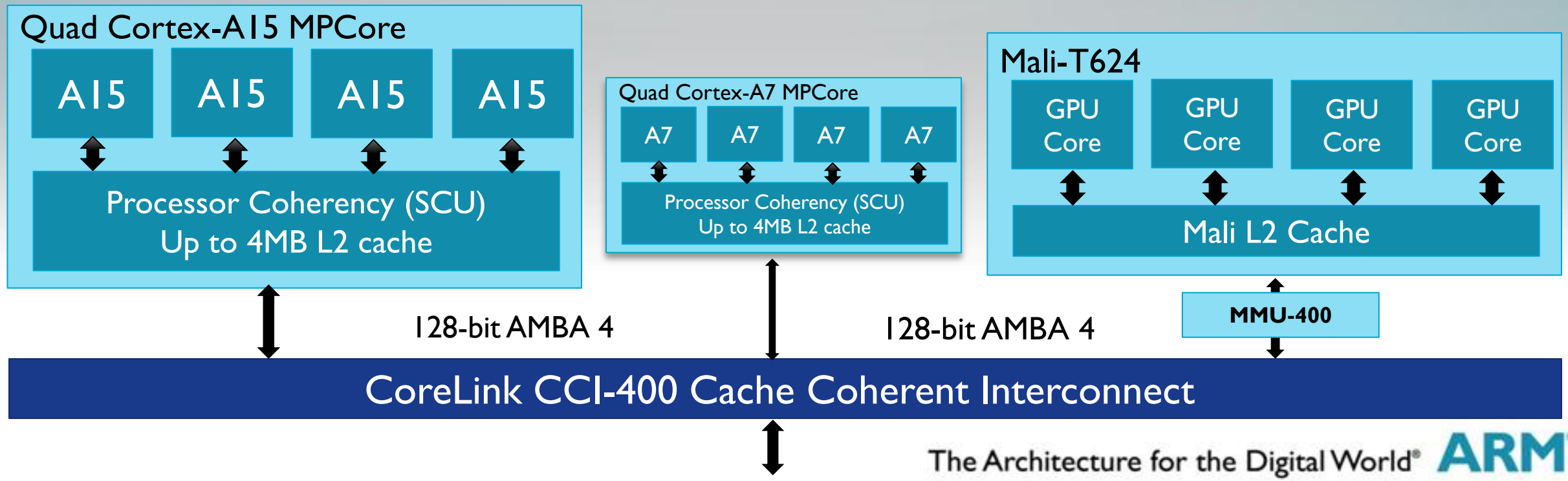
Cortex-A15

Cortex-A57

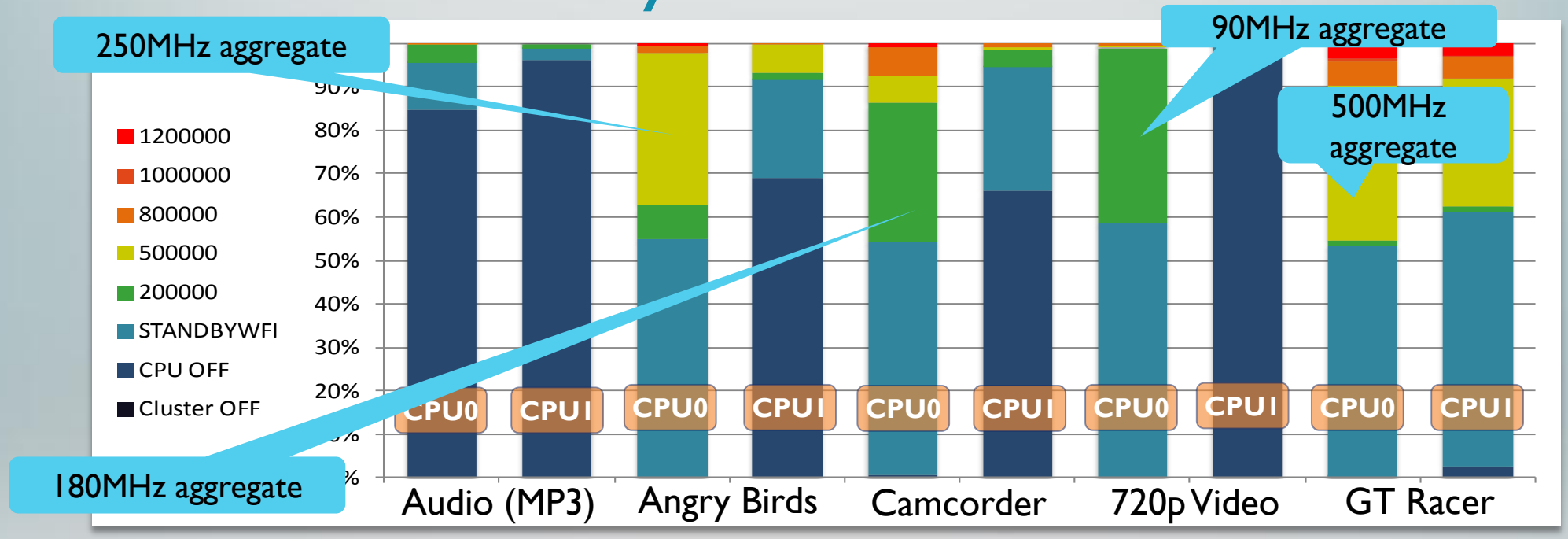
ARM System Scalability

Introducing CCI-400 Cache Coherent Interconnect

- Processor to Processor Coherency and I/O coherency
- Memory and synchronization barriers
- Virtualization support with distributed virtual memory signaling



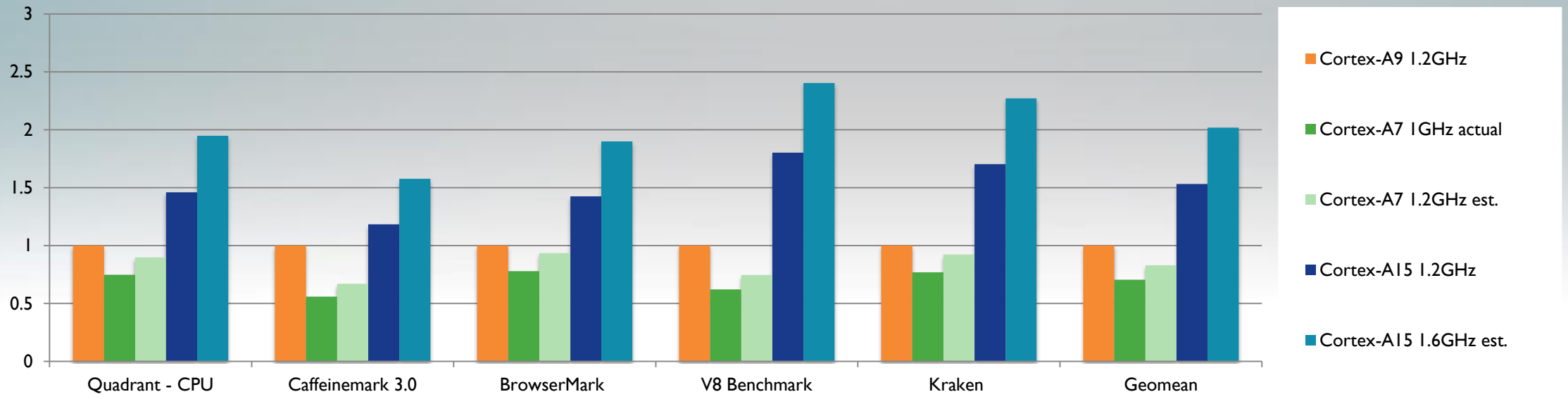
Low to Medium Intensity Use Cases



- DVFS profiles from leading Dual Cortex-A9 Smartphone
- Demonstrates that many common applications require only low to moderate processing power
- All of these use cases will run predominantly on the LITTLE cores
- The small % peaks to high MHz don't necessarily require migration to big cores
  - The short term DVFS system response to any increase in average load is to go to the highest MHz

big and LITTLE CPU Performance

- Cortex-A9 powers high end mobile devices today
- Cortex-A7 delivers comparable performance
  - …at lower power and area
- Cortex-A15 delivers significantly higher performance



Note: Cortex-A15 and Cortex-A7 results are from a test platform, with lower memory performance than production systems will deliver

Summary

- Getting the maximum efficiency out of modern SoCs is highly complex
  - Interactions between many sub-system to optimize
  - Requires new innovations and technology focus
- ARM Cortex-A15 / coherent Mali / big.LITTLE enable highest performance and scalability from mobile through to console class gaming.
- ARM continues to drive the development for better system integrations
  - Cortex™ CPUs, Mali™ GPUs and CoreLink™ fabric leading the way
  - Future v8 AArch64 with multi-cluster for next-generation gaming

Thank you