Mobile GPUs

Varun Sampath
University of Pennsylvania
CIS 565 - Spring 2012

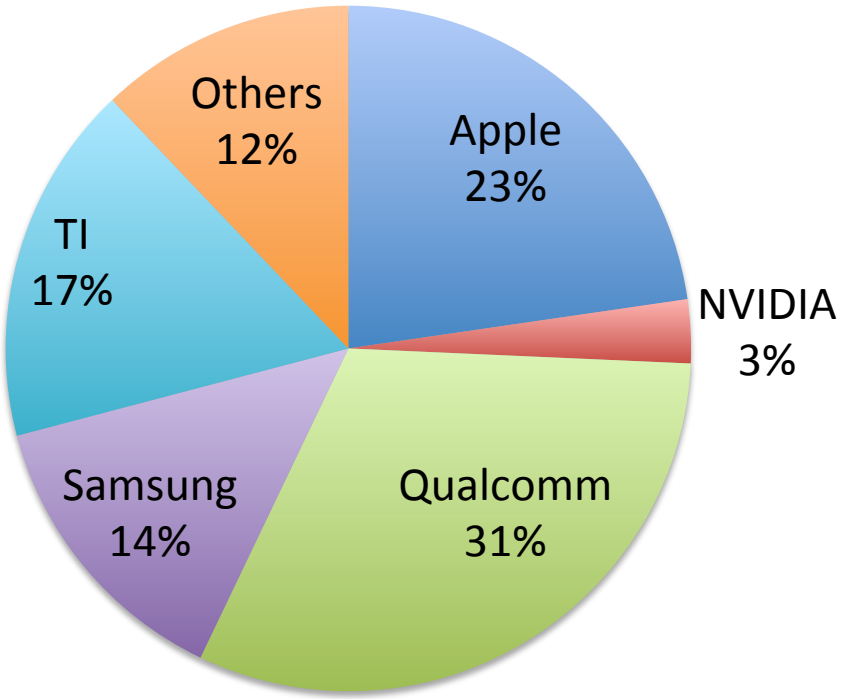
Agenda

- SoCs
- Case Studies
  - NVIDIA Tegra 2, Tegra 3
  - Imagination Technologies' PowerVR SGX Series5XT
  - Apple iPad (2012)
- Future
- Note about sources

What is an SoC?

- System-on-a-Chip
  - CPU, GPU, DSP, I/O
  - Single-chip solution
- Top mobile SoC vendors:
  - Qualcomm, Apple, TI, Samsung, NVIDIA
- Advantages of using SoCs?
- Disadvantages?
- We will see all consumer chips converge to SoCs

Mobile SoC Market Share 2011



Market Share Data from PC Perspective

What is an SoC?

- Apple A5 Dual-core Processor (more on this later)
- Qualcomm RTR8605 Multi-band/mode RF Transceiver. Chipworks has provided us with a die photo.
- Skyworks 77464-20 Load-Insensitive Power Amplifier (LIPA®) module developed for WCDMA applications
- Avago ACPM-7181 Power Amplifier
- TriQuint TQM9M9030 surface acoustic wave (SAW) filter
- TriQuint TQM666052 PA-Duplexer Module

Image from iFixit

Block Diagram of TI OMAP 4470

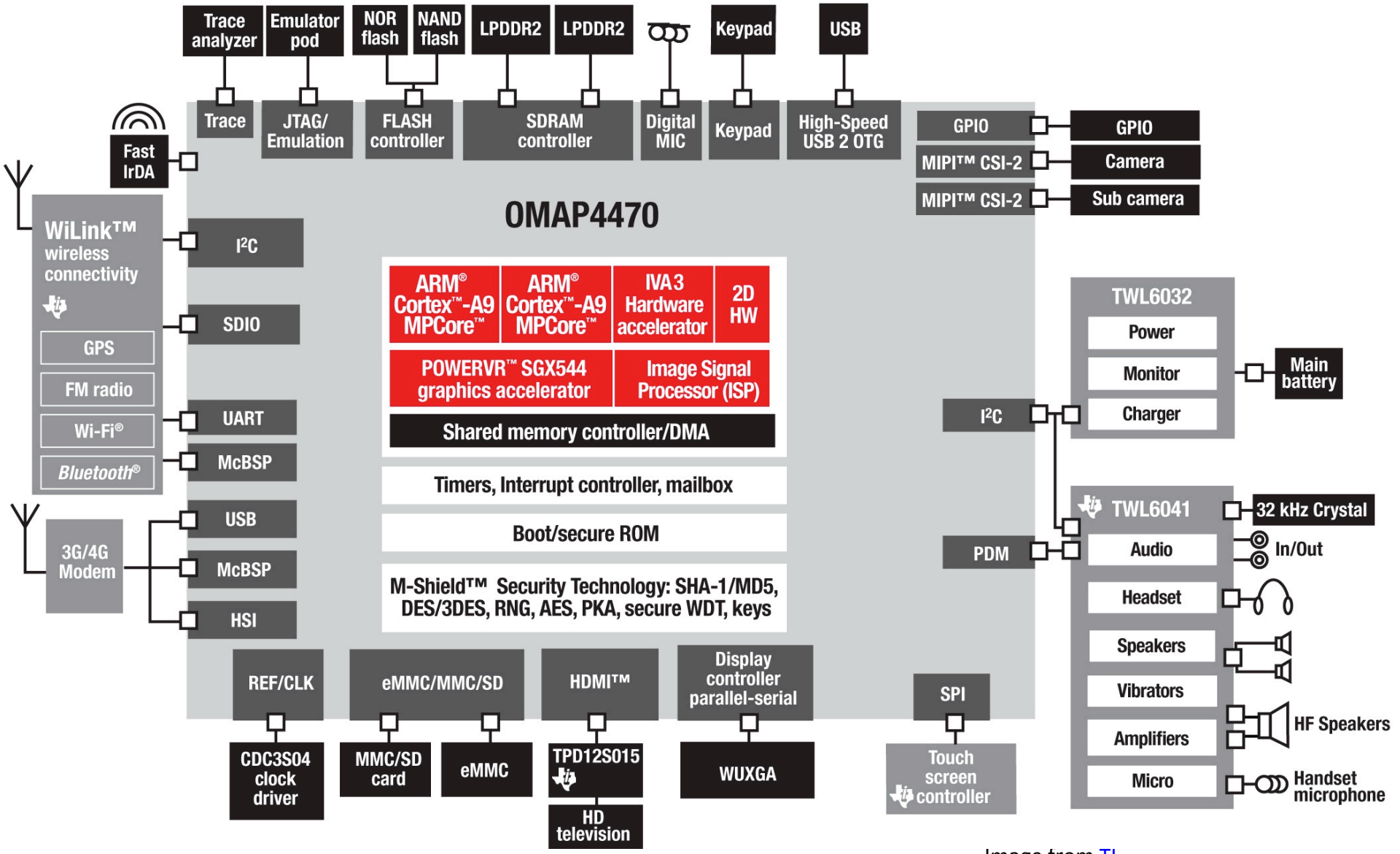


Image from TI

Brief Discussion of ARM

- RISC CPU vendor that currently dominates mobile
- Mobile Designs: Cortex-A8, A9, A15
- Fabless Designer
  - Core Design Licensees
  - Architecture Licensees
    - Qualcomm Scorpion/Krait
    - NVIDIA

The Constraints of Mobile

- Energy
  - Cell phone battery capacity of 5-7 Wh (tablets 20-40 Wh)
  - How much energy can our chips consume?
- Area
  - PCB size constraints
  - Cooling constraints

Some Energy Numbers

Power Consumption Comparison		
	Apple iPhone 4 (AT&T)	Apple iPhone 4S (AT&T)
Idle	0.7W	0.7W
Launch Safari	0.9W	0.9W
Load AnandTech.com	**1.0W**	1.1W
Maps (Determine Current Location via GPS/WiFi)	**1.3W**	1.4W

Power Consumption Comparison		
	Apple iPhone 4 (AT&T)	Apple iPhone 4S (AT&T)
Launch Infinity Blade	**2.2W**	2.6W
Infinity Blade (Opening Scene, Steady State)	**2.0W**	2.2W

Data from AnandTech

Some Contributors to Switching Energy

- Off-chip Interconnect (to DRAM)
  - Bandwidth is expensive
  - Minimize reasons to fire up memory bus
- High frequencies
  - Requires increased voltages

Some Theoretical Performance Numbers

	Apple iPad 2	ASUS Transformer Prime	Some Nice Desktop
CPU	A5 @ 1GHz	Tegra 3 @ 1.4GHz	Sandy Bridge @ 3.4GHz
GPU	POWERVR SGX543MP2 @ 250MHz	Mobile GeForce @ 500MHz	GTX680 @ 1GHz
Memory Interface	64-bit @ (maybe) 800MHz = 6.4GB/s	32-bit	256-bit @ 6GHz = 192GB/s
GPU GFLOPS	16 GFLOPS	12 GFLOPS	3 TFLOPS

Mobile Data from AnandTech
GTX680 Specs from Newegg

GeForce GPU in NVIDIA Tegra 2

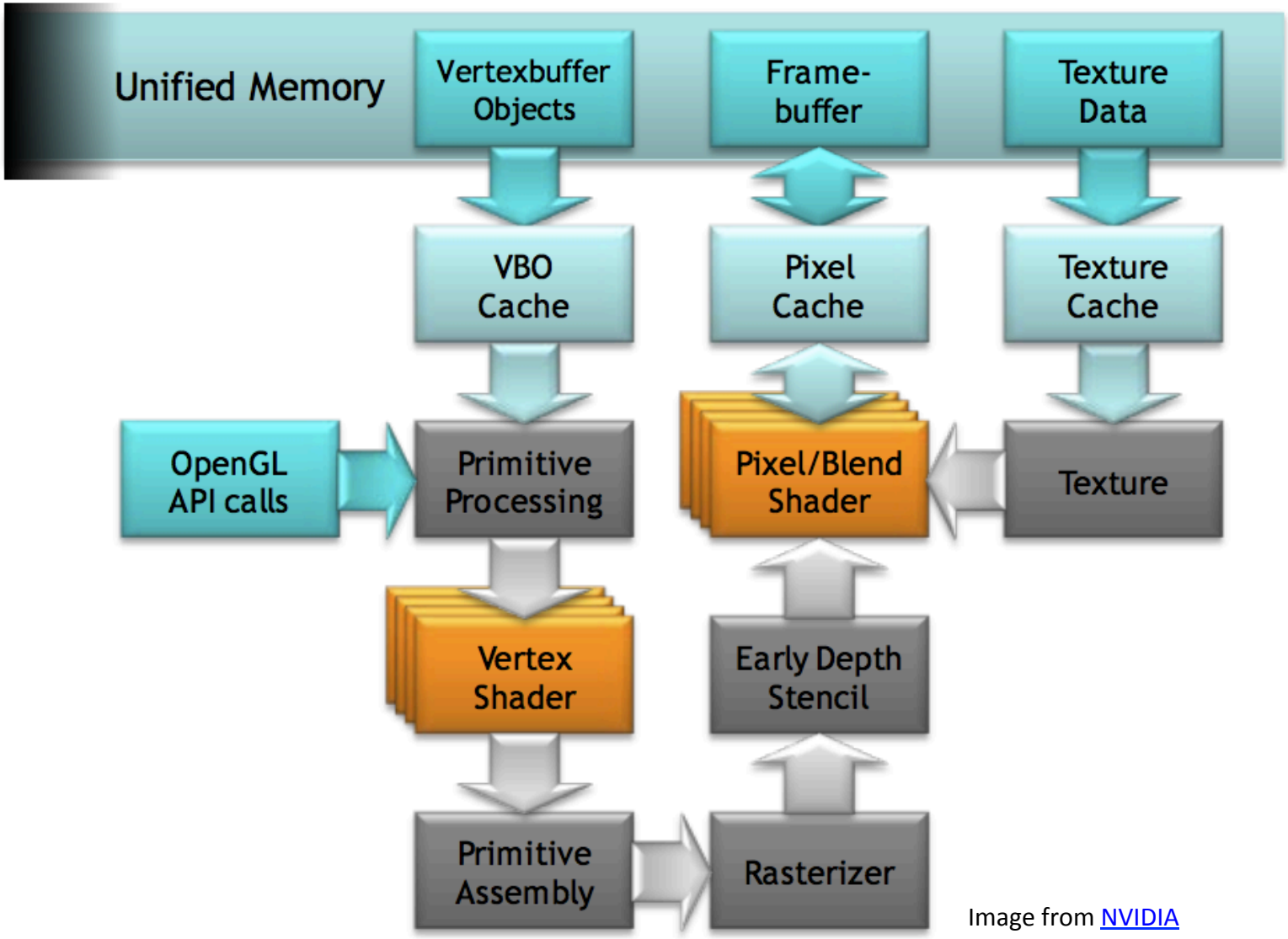


Image from NVIDIA

Tegra 2 Mobile GeForce

- Separate vertex and pixel shaders
  - 4 of each, each capable of 1 multiply-add /clock
- Pixel, texture, vertex, and attribute caches
  - Reduce memory transactions
  - Pixel cache useful for UI components
- Memory controller optimizations
  - Arbitrate between CPU & GPU requests
  - Reorder requests to limit bank switching

NVIDIA Tegra 3 (Kal-El)

- Expanded Mobile GeForce
  - 4 vertex and 8 pixel shaders
- 4-PLUS-1 architecture

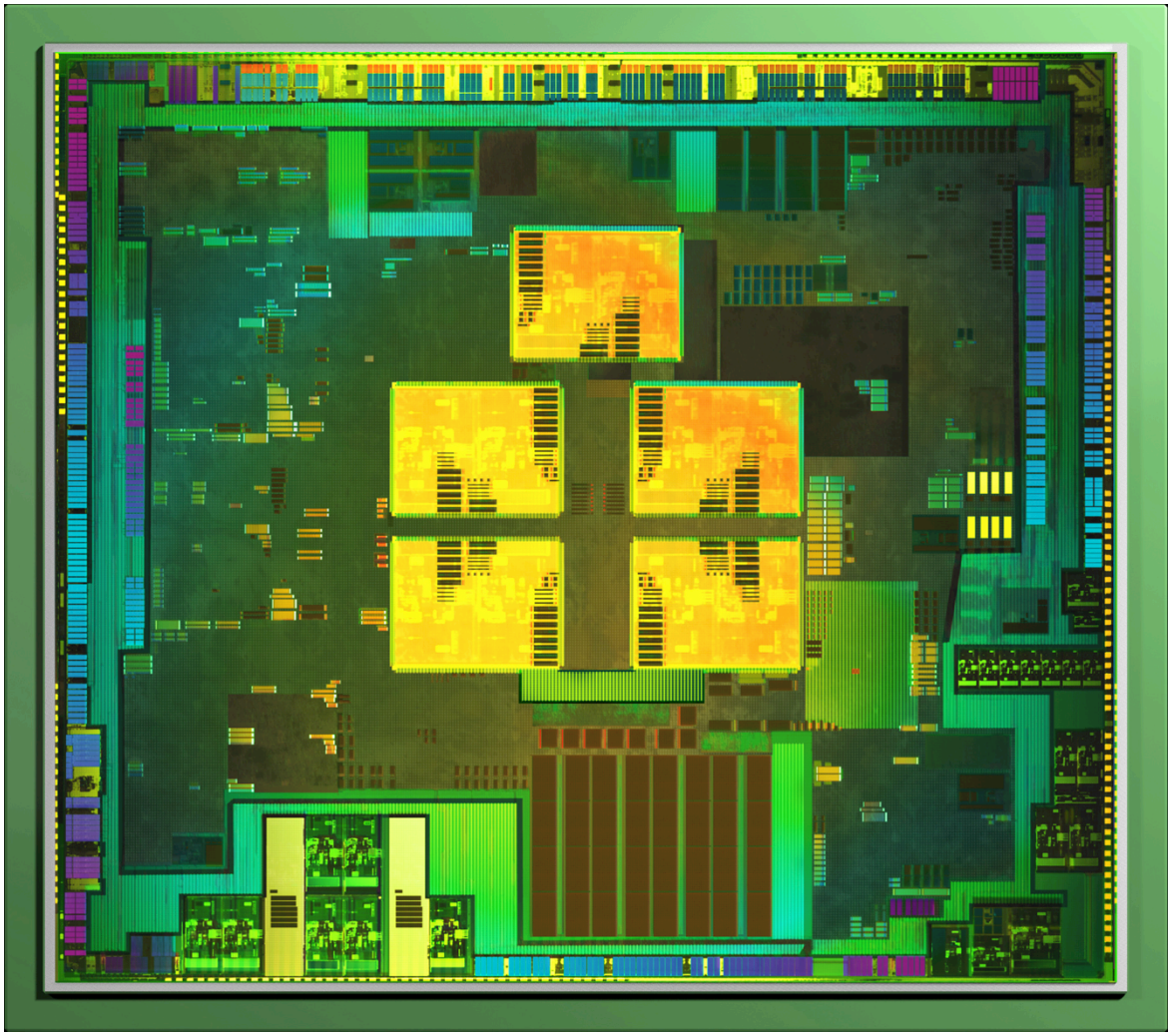
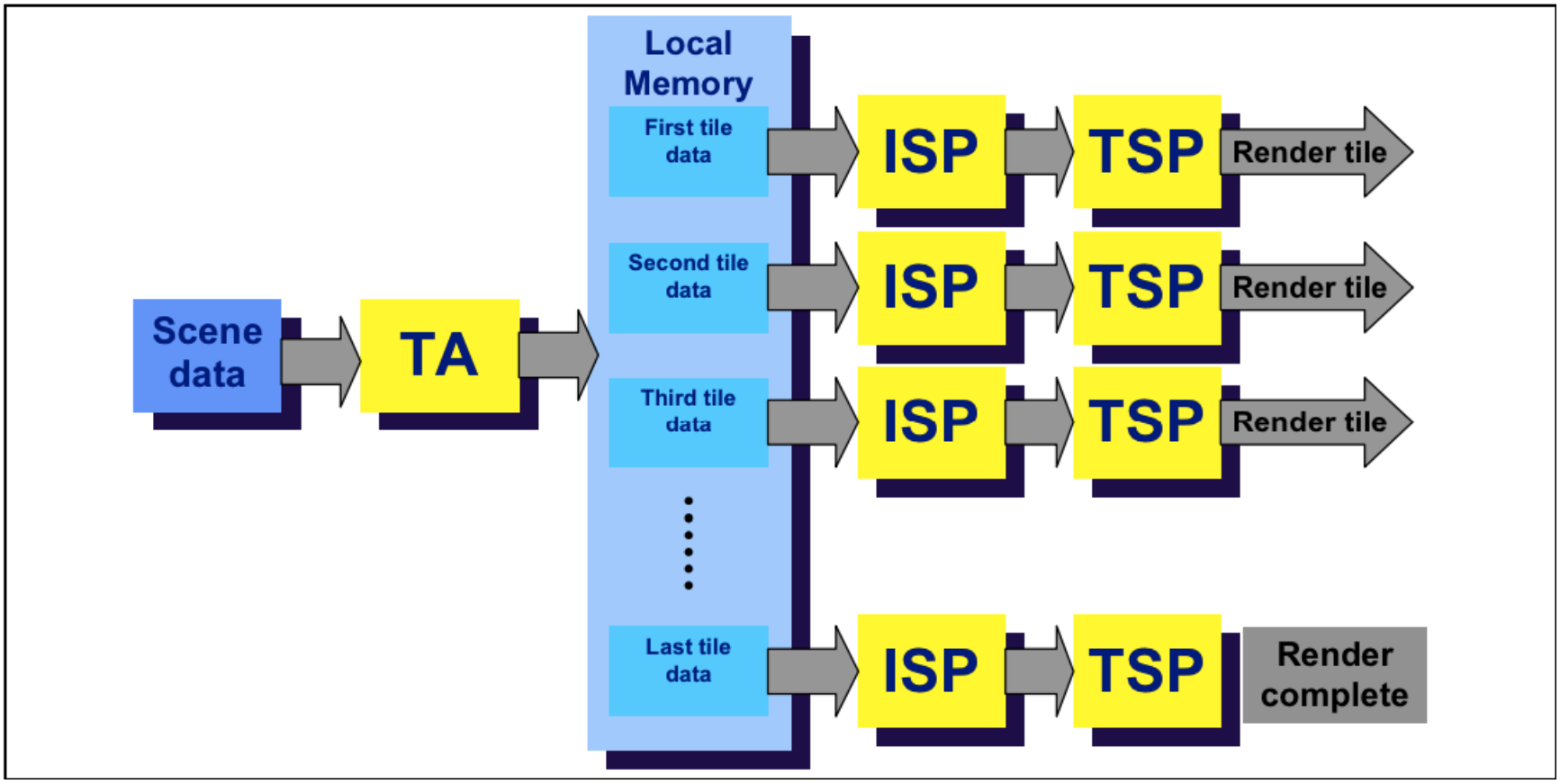


Image from AnandTech

PowerVR SGX



- TA (Tile Accelerator) – store scene data and split up screen into tiles
- ISP (Image Synthesis Processor) – perform Hidden Surface Removal with z-testing
- TSP (Texture and Shading Processor) – run pixel shader

Image from ImgTec

PowerVR SGX Series5XT

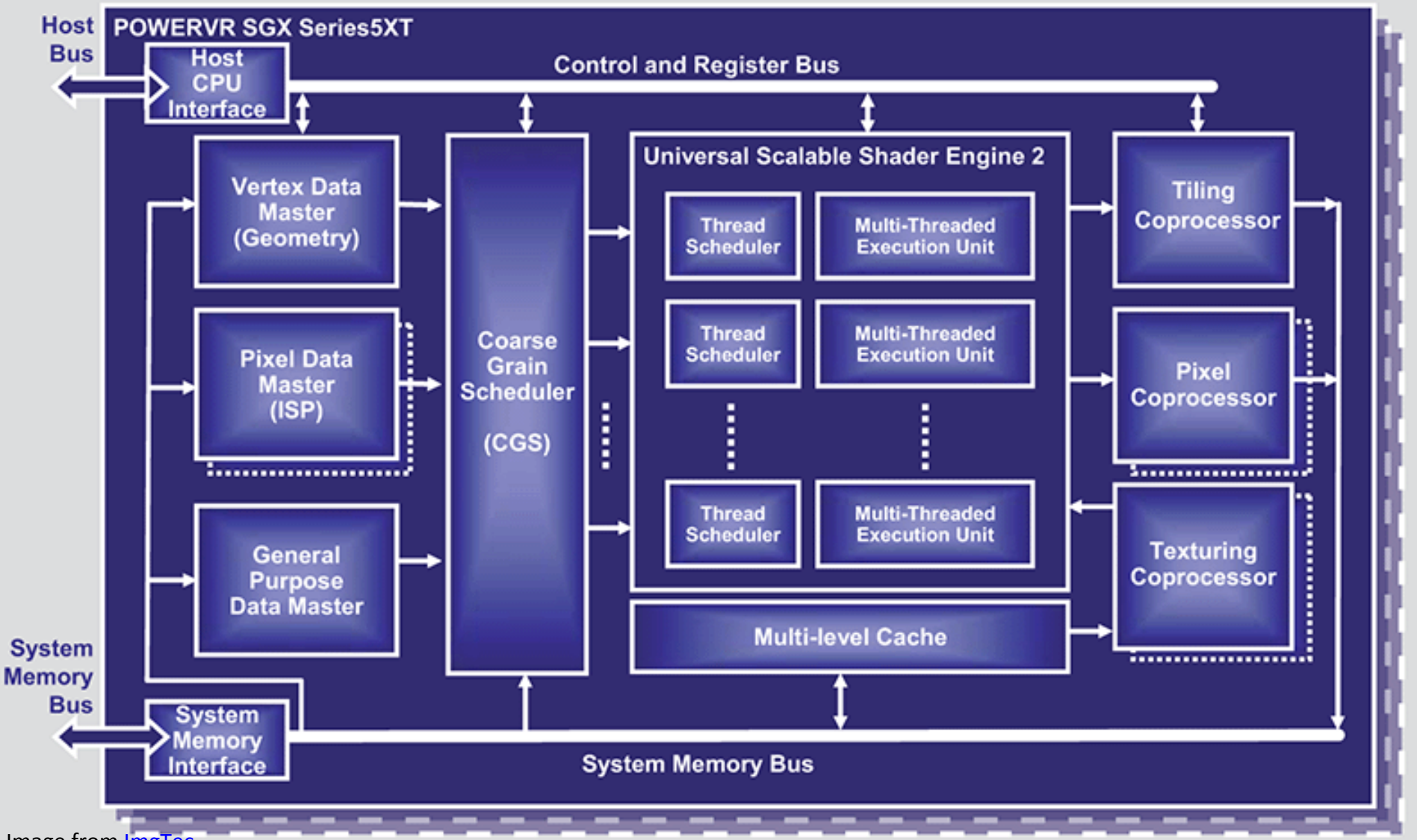


Image from ImgTec

Summarizing PowerVR SGX Series5XT

- Used in Apple A5, A5X
- Unified shader architecture (called USSE2)
- Tile based deferred rendering (TBDR)
  - Will cover in more detail next week
- Multi-core architecture

Mobile GPU Families

- Qualcomm Adreno
  - Unified shaders, 4-wide SIMD
  - immediate mode with early-z
- Imagination Technologies' PowerVR SGX Series5XT
  - Unified shaders, 4-wide SIMD
  - Tile based deferred rendering
- NVIDIA Mobile GeForce
  - Separate vertex (4) & pixel (8/12) shaders , scalar
  - immediate mode with early-z
- ARM Mali
  - Separate vertex (1) & pixel (4) shaders , 4/2-wide SIMD
  - immediate mode with early-z

Analysis by AnandTech

Demands for Mobile

- Higher screen resolutions
  - Requires more memory bandwidth
  - Pixel count growing higher than geometry?
- Longer battery life
- Higher quality mobile gaming

Case Study: the new iPad

- Screen resolution of 2048x1536
  - Quadruple the pixels of previous 1024x768 version
  - Higher than nearly all desktop and laptop displays
- Battery life approximately equal to previous version
- Gaming performance?

iPad Gaming Performance

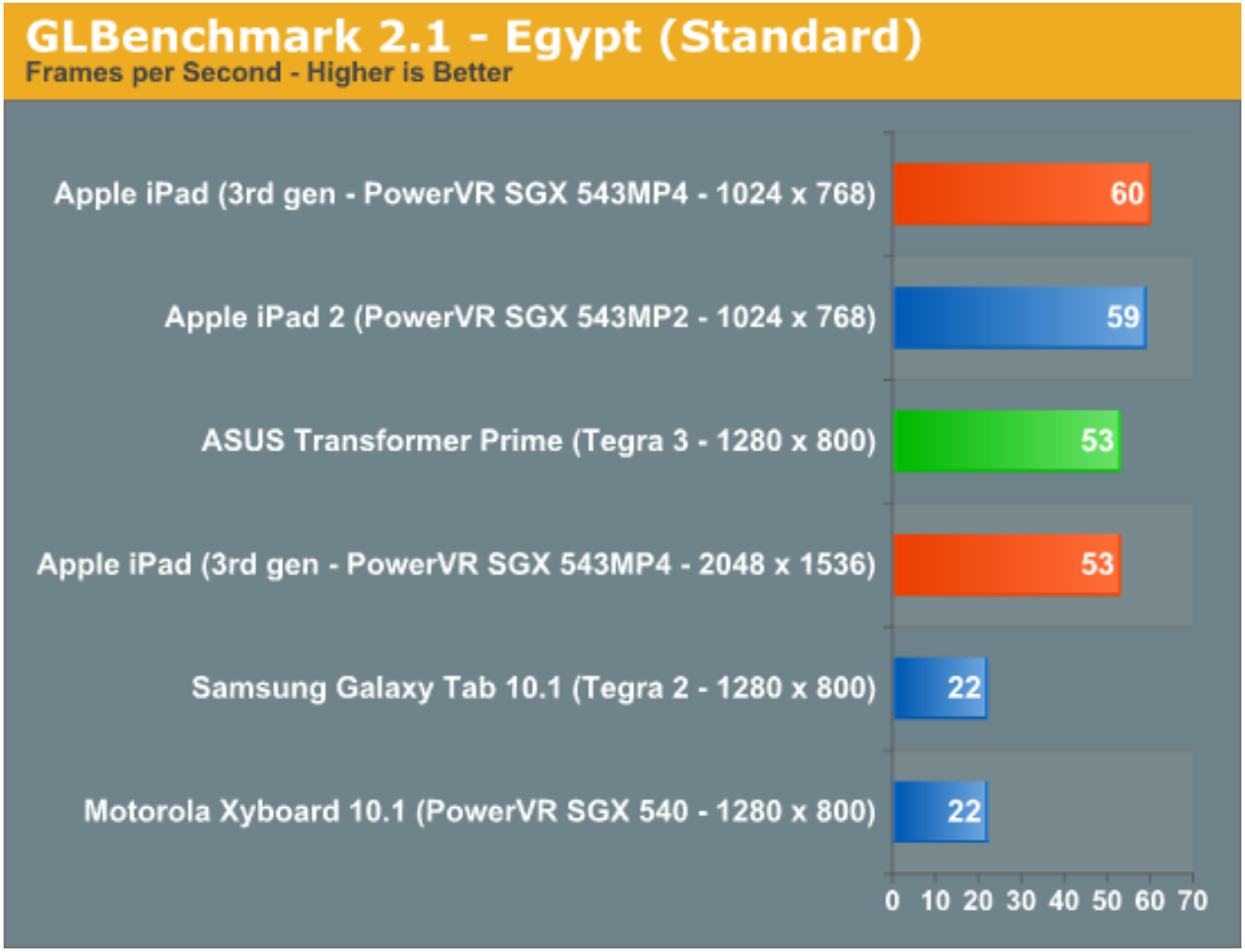
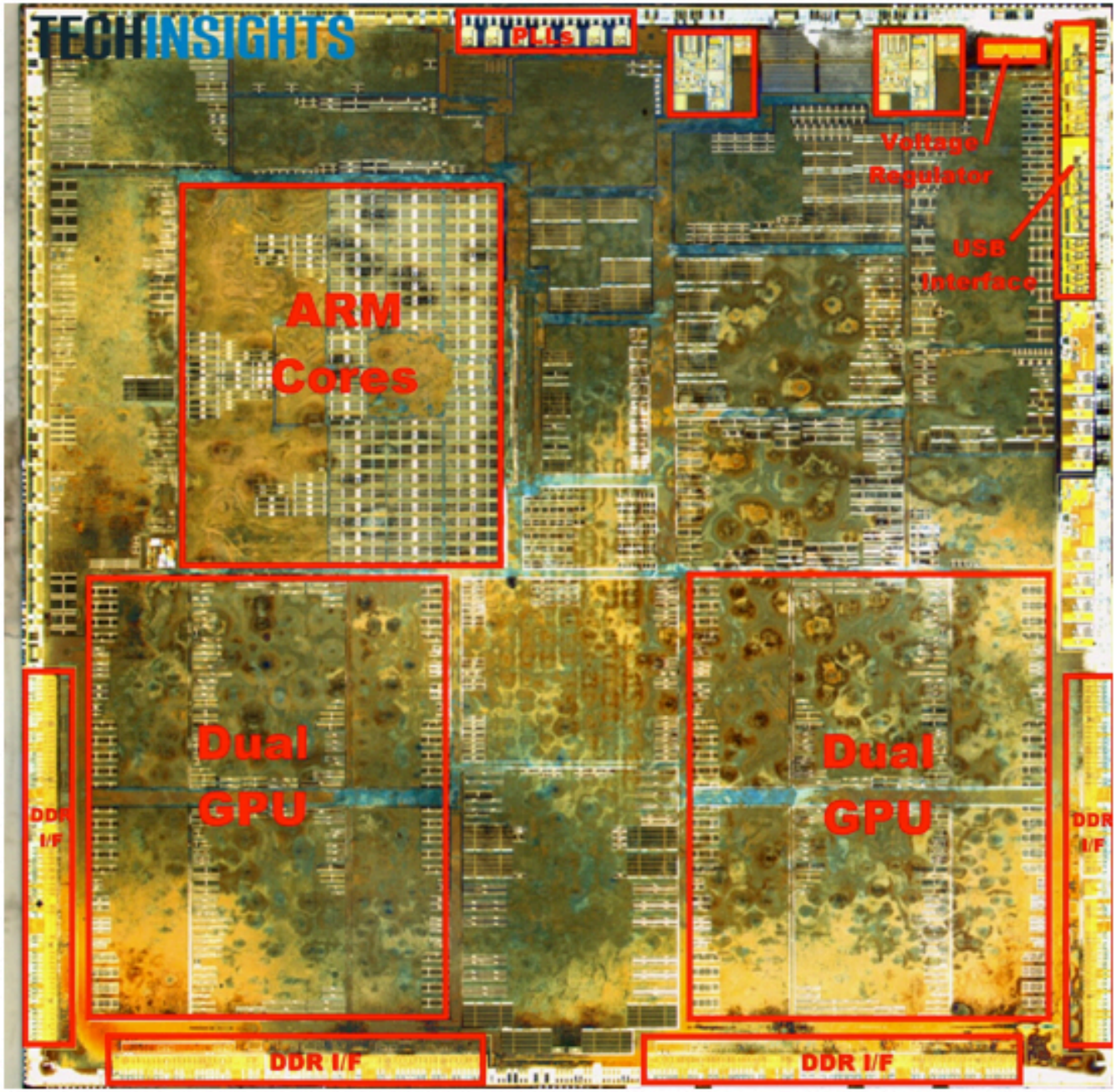


Image from AnandTech



Apple A5X Die Shot

Image from [UBMTechInsights]

Apple iPad Statistics

	Apple iPad 2	Apple iPad (2012)	11” Apple MBA
CPU	A5 @ 1GHz	A5X @ 1GHz	Sandy Bridge @ 1.8GHz
GPU	POWERVR SGX543MP2 @ 250MHz	POWERVR SGX543MP4 @ 250MHz	Sandy Bridge IGP @ 350MHz/ 1.2GHz
Memory Interface	64-bit @ 800MHz = 6.4GB/s	128-bit (for GPU)	128-bit @ 1.3GHz = 20.8GB/s
Die Size	122mm^2	**163mm^2**	149mm^2
Battery Size	25Wh	**42.5Wh**	35Wh

Data and Image from AnandTech

What will the future bring?

- GPU Compute
  - PowerVR SGX Series5XT OpenCL capable, but no drivers
  - Could do compute the old-fashioned way with GLSL
  - Direct3D 11 means Compute Shader support
- PowerVR Series6 press release suggests 100-1000 GFLOPS
- Kepler-based GPU coming to a super phone near you?