

GPUs móviles

Daniel Flores Tafur
Javier Sanchez Martinez

Introducción

- La era post-PC: A parte de PCs y portátiles "clásicos" han aparecido:
 - Smartphones
 - Tablets
- En el futuro, los habrán más que ordenadores de sobremesa y los portátiles

Introducción

- La necesidad de la potencia gráfica es casi tan exigente como en ordenadores de sobremesa
- Interfaz complicada
- Juegos 3D
- Videos en alta definición
- Resoluciones de pantalla enormes

Samsung Galaxy S4: un "móvil"



Resolución de pantalla: 1920 por 1080

Apple iPad (4ta generación)



Resolución de pantalla: 2048 por 1536

Introducción

- A la vez de proporcionar todo esto, las GPUs móviles deben:
 - Gastar poca energía
 - Ocupar muy poco espacio
 - No calentarse mucho
- En resumen, una tarea imposible... Pero, lo intentan

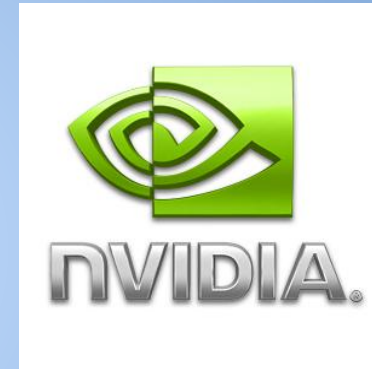
Introducción

- En el mundo de ordenadores, solo existen:



Introducción

- En el mundo de GPUs móviles existen:



- Y hay muchas más...

APIs móviles

- GPUs para ordenadores son una implementación de las APIs gráficas
- GPUs móviles también lo son
- La API principal: OpenGL ES (Embedded Systems)

APIs móviles

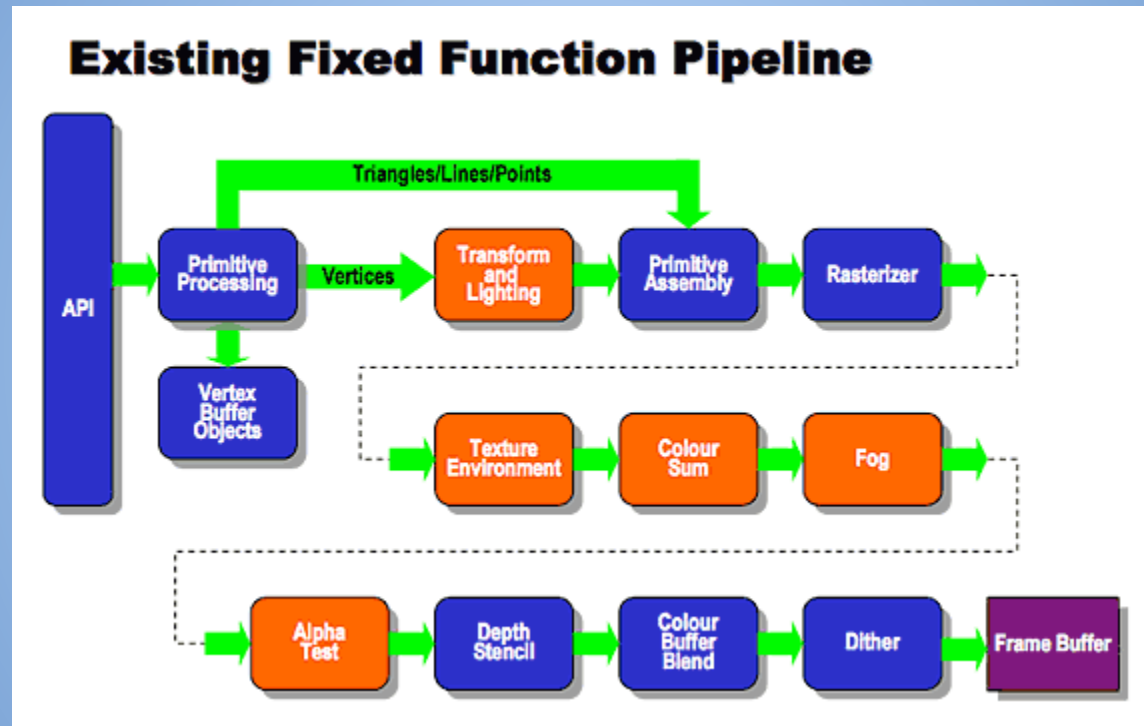
- OpenGL ES 1.0 = OpenGL 1.1 reducido
- Introducida antes de la era post-PC
- GPUs muy simples
- Procesamiento gráfico a través de CPUs sin FPU

APIs móviles

- OpenGL ES 1.1 = OpenGL 1.5 reducido
- GPUs más avanzadas
- Menos uso de CPU
- OpenGL ES 1.0 y OpenGL ES 1.1 utilizan el pipeline fijo

Pipeline fijo de OpenGL ES 1.0/1.

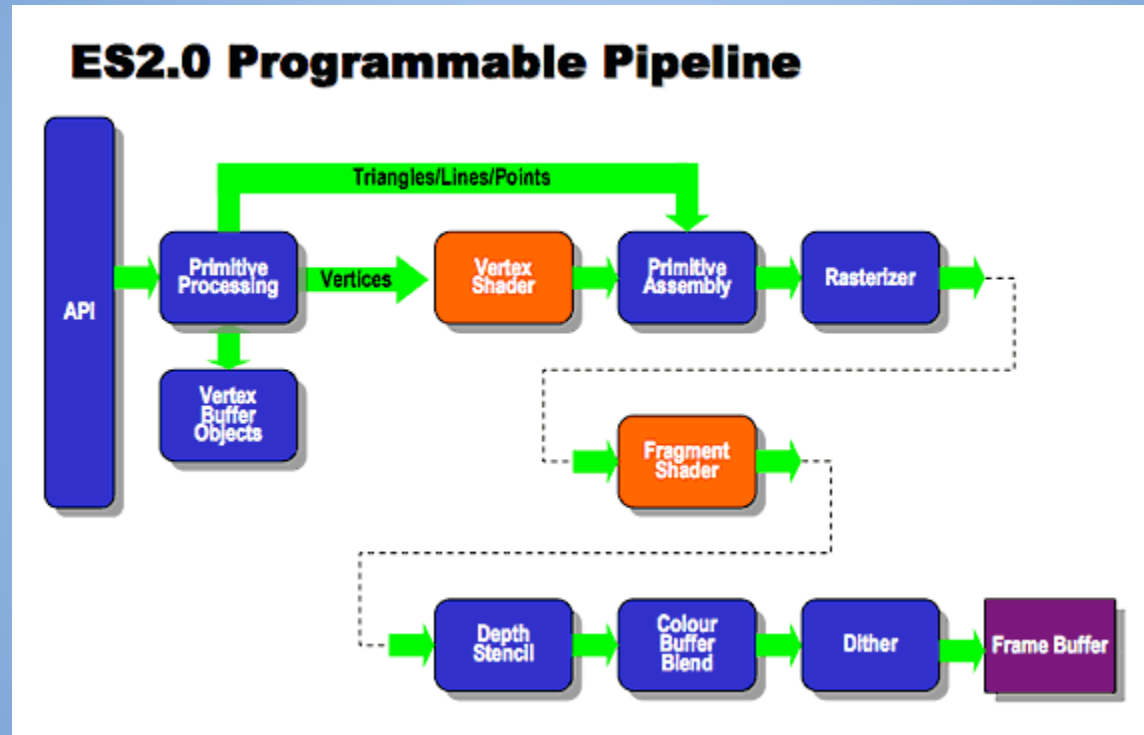
1



APIs móviles

- OpenGL ES 2.0
- Basado en OpenGL 2.0, pero con algunos rasgos de OpenGL 3.1
- Sale en marzo de 2007, entrando en la era post-PC
- Introduce el pipeline programmable

El pipeline programmable de Open GL ES 2.0



APIs móviles

- OpenGL ES 3.0 (agosto de 2012)
- Nuevo lenguaje de programación para shaders
- Efectos visuales y texturas en alta calidad
- GPUs casi tan complejas como de ordenadores

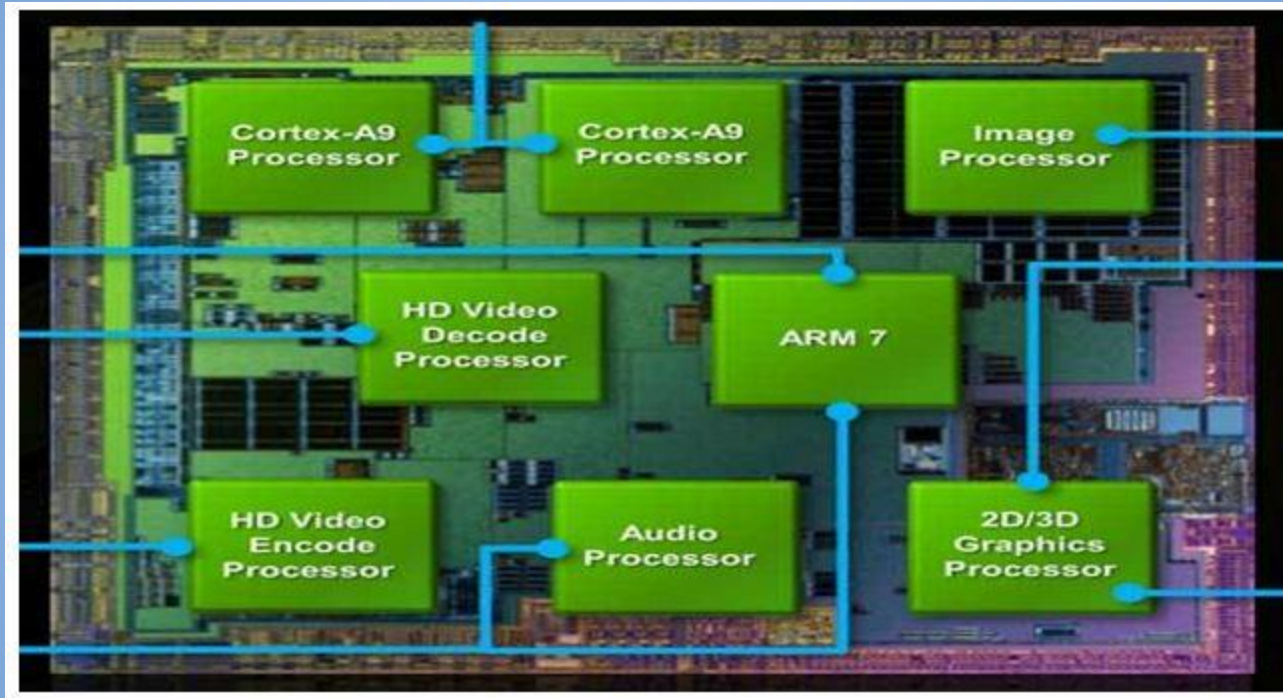
APIs móviles

- OpenGL ES es muy similar a OpenGL normal y corriente
- Tiene algunas restricciones respecto a OpenGL
- Si sabes programar en OpenGL, sabes programar en OpenGL ES

Arquitectura de GPUs móviles

- CPU + GPU están muy integradas
 - Se venden como una plataforma
- CPU, GPU y otros sistemas están en el mismo chipset
- Ahorra espacio, energía y reduce la salida de calor

Arquitectura de GPUs móviles



Arquitectura de GPUs móviles

- Renderizado: dos extremos
- IMR vs TBDR
- IMR: Immediate-Mode-Rendering
- Intenta renderizar toda la imagen de golpe
- Solía utilizarse en las primeras GPUs

Arquitectura de GPUs móviles

- TBDR: Tile-Based-Deferred-Rendering
- Se utiliza un buffer y la pantalla se separa en partes iguales
- El renderizado se hace parte por parte
- Solo se accede a los píxeles que aparecen en la pantalla

Arquitectura de GPUs móviles

- TBIMR: Tile Based Immediate Mode Rendering
- El renderizado se hace por partes...
 - Pero cada parte se renderiza como toda la pantalla en IMR
- Se utiliza en la mayoría de GPUs

Arquitectura de GPUs móviles

- Shaders especializados vs shaders unificados
- El mismo trend que en las GPUs de ordenadores
- Los shaders unificados habilitan el uso de cálculos en paralelo con la GPU
 - Con OpenCL

Arquitectura de GPUs móviles

	Adreno	GeForce ULP	Mali-400	Mali T604-	PowerVR
Unified Shader	yes	no	no	yes	yes
Render Mode	TBIMR	TBIMR	TBIMR	TBIMR	TBDR

- Casi en todas las GPUs modernas se utilizan shaders unificados y TBIMR

Adreno de Qualcomm

- Parte de la plataforma Snapdragon
- Implementación propia de Qualcomm
- Basada en GPUs AMD Radeon

Adreno de Qualcomm

- Serie 200, misma arquitectura que en GPUs de AMD
- Mejoras propias de Qualcomm
- Shaders unificados
- Soportaba OpenGL ES1.1/2.0, Direct3D, entre otros

Adreno de Qualcomm



- Diferencias en número de ALUs, cantidad de memoria...

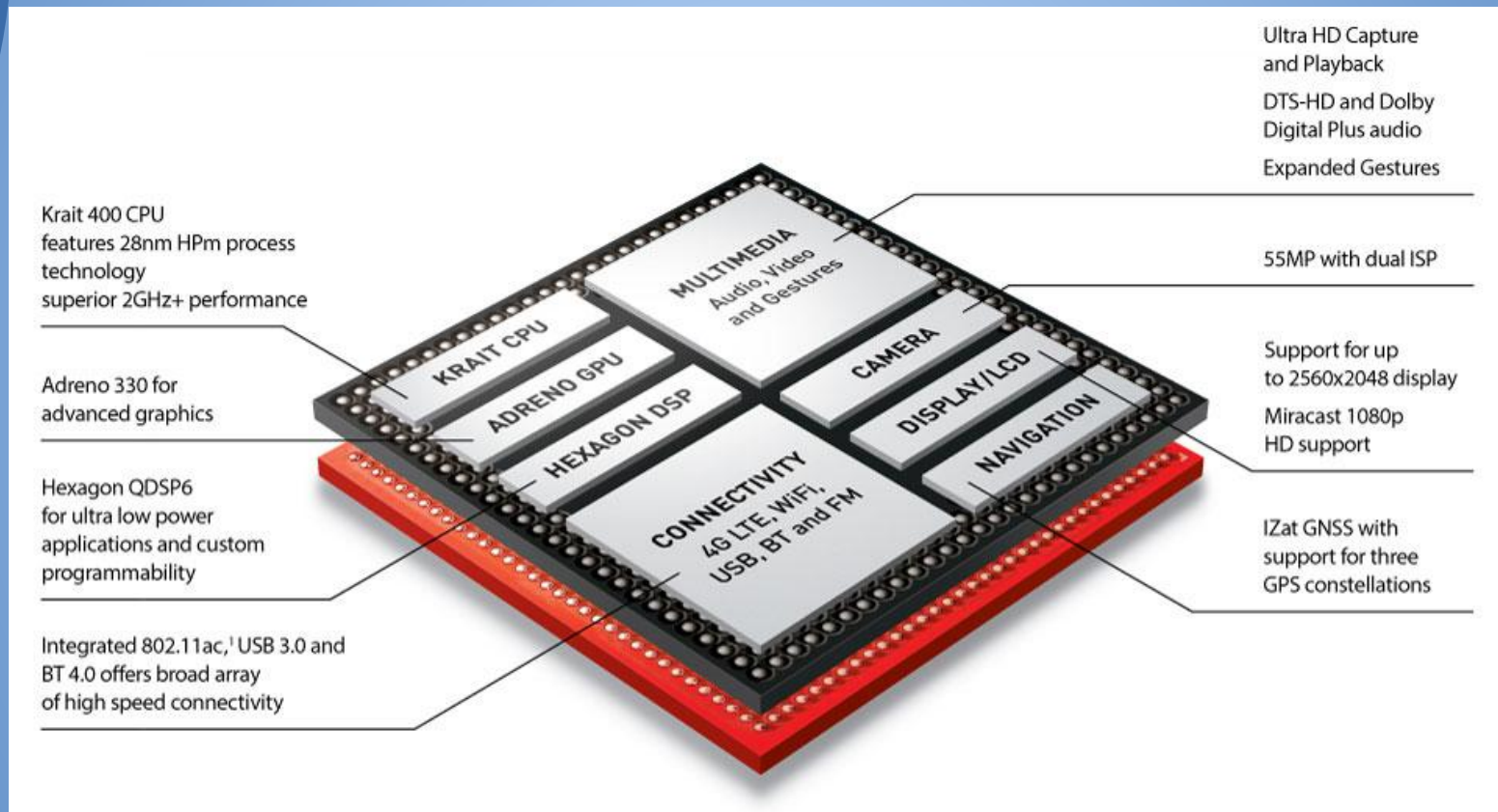
Adreno de Qualcomm

- Serie 300, la nueva arquitectura
- Shaders unificados
- Soporte para OpenGL ES 3.0 y OpenCL 1.2
- Modelos 305 y 320

Adreno de Qualcomm

- Próxima versión: Adreno 330, presentada este año
- Parte de la plataforma Snapdragon 800
- Según Qualcomm ofrece el rendimiento 50% más alto que Adreno 320

Adreno de Qualcomm



NVidia Tegra

- Primera versión, en 2008
- "NVidia Tegra" es el nombre de una plataforma
 - CPU ARM y GPU GeForce ULV
- Soporta OpenGL ES 2.0 y Direct3D

NVidia Tegra

- Segunda versión en el año 2010
- Ciertos cambios en la arquitectura (de GeForce ULV a GeForce ULP)
- Tegra 3 en 2011
- Misma arquitectura que Tegra 2, cambios cuantitativos

NVidia Tegra

NVIDIA Tegra 2 GPU

Vertex Unit
(Vec4)

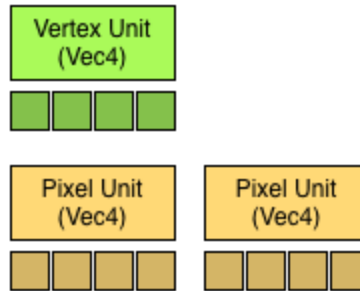


Pixel Unit
(Vec4)



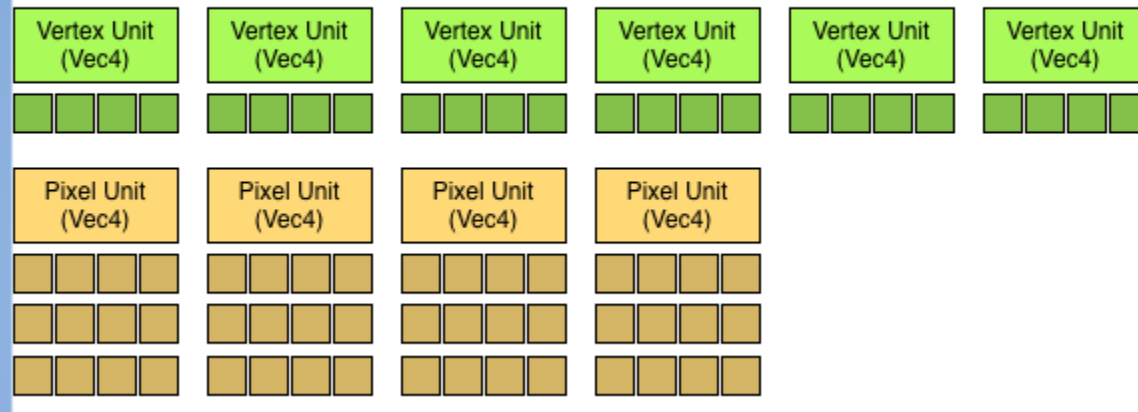
NVidia Tegra

NVIDIA Tegra 3 GPU



NVidia Tegra

NVIDIA Tegra 4 GPU



- NVidia Tegra 4, introducida este año
- Misma arquitectura que Tegra 3 y Tegra 2

NVidia Tegra 4

- Arquitectura basada en NV40 para GPUs de ordenadores
 - de año 2004...
- Es una de las pocas GPUs modernas que no tiene shaders unificados
- No soporta OpenGL ES 3.0
 - Ni, obviamente, OpenCL

NVidia Tegra 4

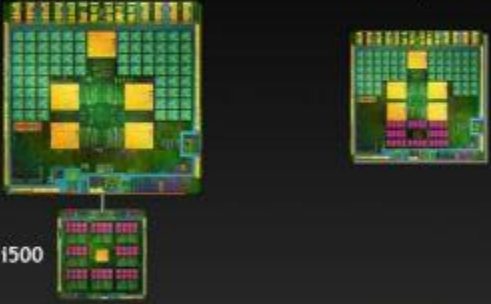
- Pero tener una arquitectura antigua no implica ser una mala GPU
 - La implementación es más importante
- NVidia dispone de sus propias herramientas para desarrolladores
 - Que permiten añadir efectos visuales avanzados
 - Compensa la falta de OpenGL ES 3.0

NVidia Tegra 4

Tegra 4 Family

	Tegra 4 <i>World's Fastest Mobile Processor</i> Superphone / Tablet	Tegra 4i <i>1st Integrated Tegra 4 LTE Processor</i> Smartphone
Quad CPU	ARM A15, 4+1	R4 ARM A9, 4+1
NVIDIA GPU	72 Core	60 Core
1500	✓	✓
Chimera	✓	✓

NVIDIA 1500



NVidia Tegra 4



NVidia Tegra 4

Tegra 4 vs Tegra 3 GPU stats

Tegra 4/ Tegra 3	
Vertex Shader	8x
Fragment ALU	8x
Pixel Rate	2.6x
Texture Rate	2.6x
Memory Rate	2.3x
Z-Kill Rate	1.3x
Triangle Rate	1.3x

Tegra 4 - 72 Core GPU @ 672 MHz
4 pixel pipes * 3 ALUs/pipe * 4 MADS/ALU +
6 VPEs * 4 MADS/VPE

Tegra 3 - 12 Core GPU @ 520 MHz
2 pixel pipes * 1 ALU/pipe * 4 MADS/ALU +
1 VPE * 4 MADS/VPE

Vivante

- Una empresa fundada en 2004 por un grupo de ingenieros de NVidia
- Es una de la veteranas del mercado
- Se sabe muy poco de la arquitectura que tienen sus GPUs

Vivante

- No es muy conocida para un consumidor europeo, pero esta presente en muchos dispositivos
 - Sobre todo, en mercado de Asia
- Esta en plataformas de Freescale, Marvell y Rockchip
 - Y, también, en Google TVs

Vivante

- Los productos más modernos son GC2000, GC4000 y GC6000
- Soportan OpenGL ES 3.0, OpenGL 3.0 y OpenCL 1.1/1.2
- GC2000 esta en plataforma Freescale i.MX6 y, por lo tanto, en algunos tablets de bajo coste

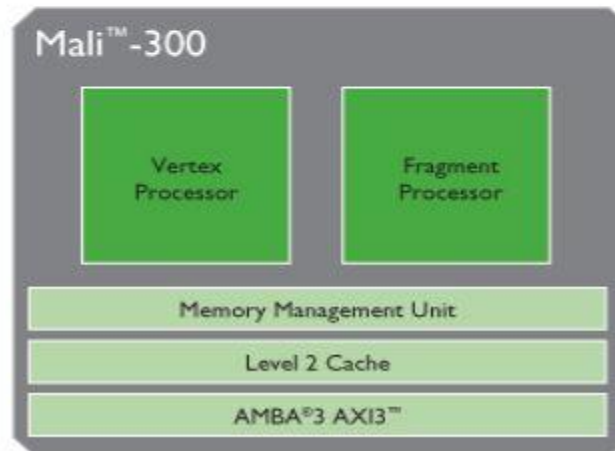
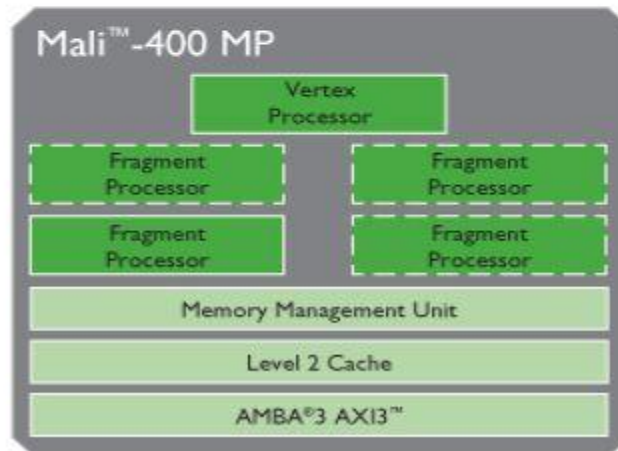
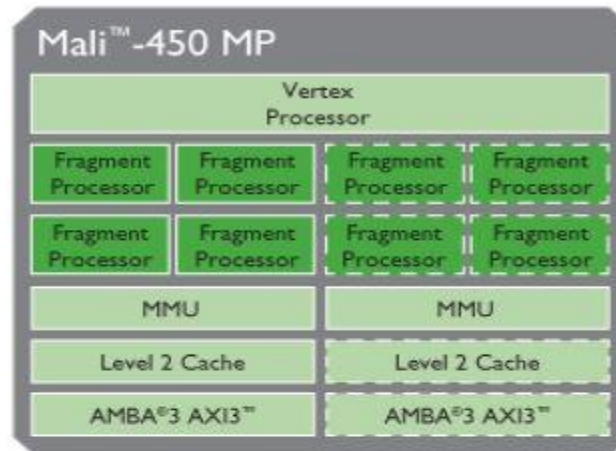
Mali

- GPUs diseñadas por ARM Holdings
- ARM no las produce, sino que vende licencias para su producción
 - Exactamente como con las CPUs
- Está presente en multitud de plataformas
 - La más conocida - Samsung Exynos

Mali

- Tiene dos series modernas: 400 y 600
- Serie 400 esta enfocada a aplicaciones gráficas
- No tiene shaders unificados
- Soporta OpenGL ES 2.0
- La GPU más reciente es Mali-450 MP, presentada en 2012

Mali

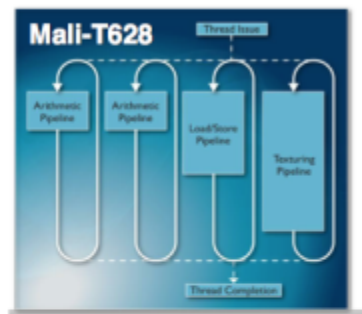


Mali

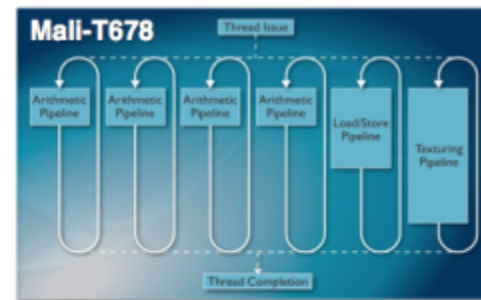
- Mali series 600: orientada a aplicaciones gráficas y a cálculos en paralelo
- La nueva arquitectura basada en shaders unificados
- Soporta OpenGL 3.0 y OpenCL 1.1

Mali

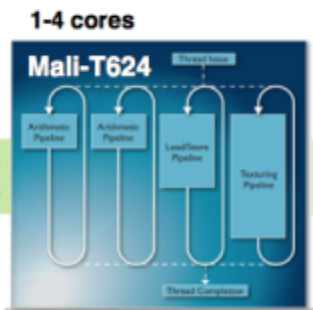
Mali-T600 – GPU Compute Scalability



4-8 cores



1-8 cores

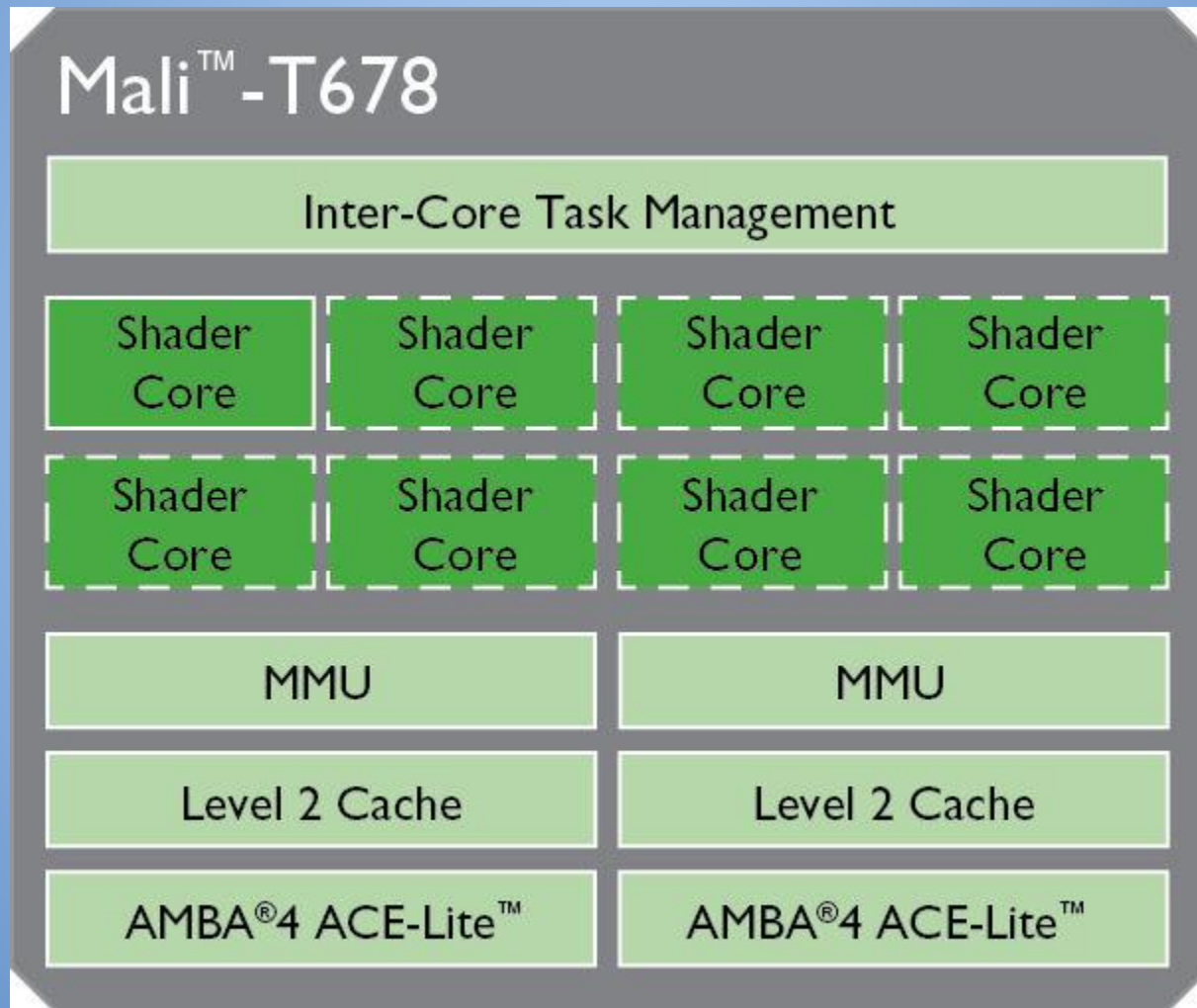


1-4 cores

2x Graphics
Performance

4x Compute
Performance

Mali



StemCell

- Arquitectura de la empresa ZiiLabs fundada en 1994 en Singapur
- Comparada en 2009 por Creative
- En 2012 Intel compró los patentes para poder utilizar sus tecnologías

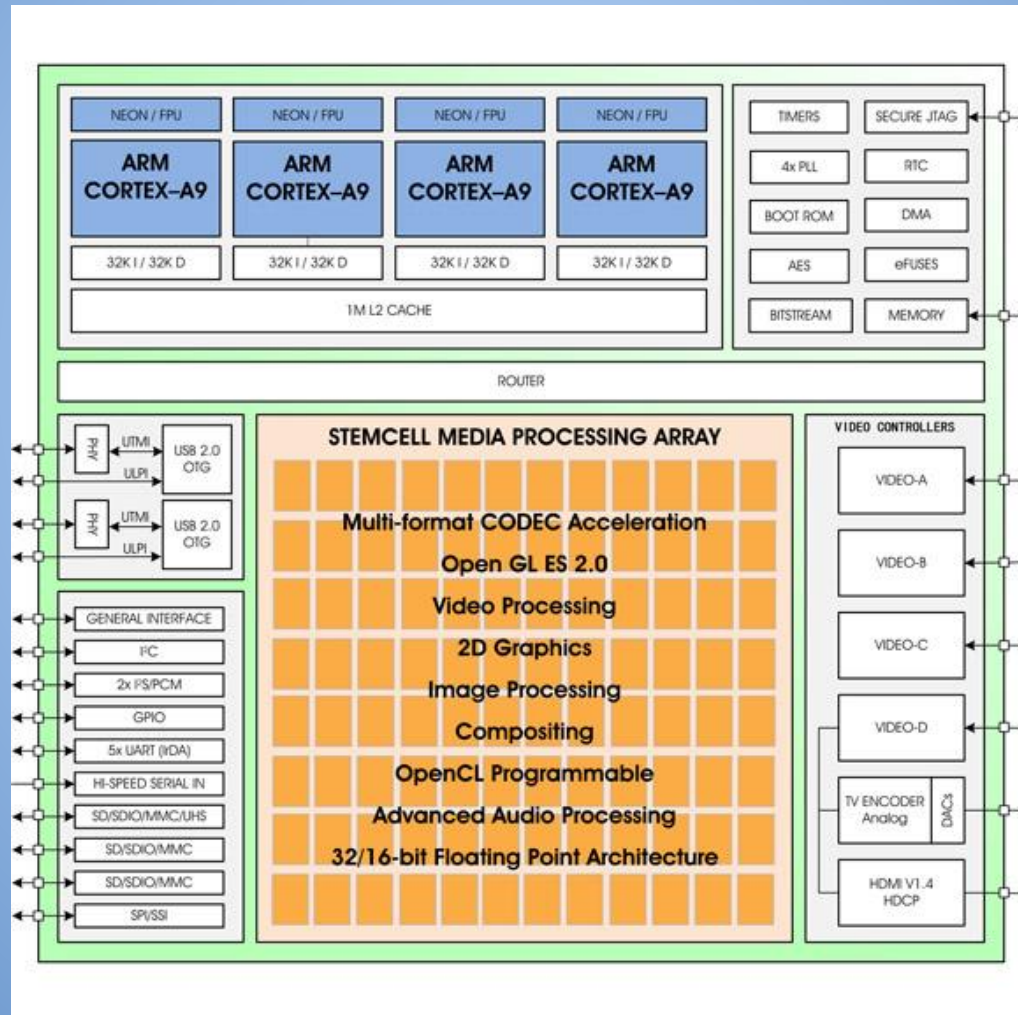
StemCell

- Tiene dos generaciones...
- La primera, formada por plataformas ZMS-05 y ZMS-08
- La segunda, formada por plataformas ZMS-20 y ZMS-40

StemCell

- Arquitectura muy similar a las arquitecturas de GPUs para ordenadores
 - Como NVidia Cuda, AMD Stream Processors
- Basada en núcleos completamente programables
- Se utilizaban tanto para gráficos, como decodificación de videos

StemCell ZMS-40



StemCell

- La GPU más avanzada, ZMS-40, dispone de 96 núcleos en el StemCell Array
- Se utilizan para gráficos, decodificación de video, audio
- Soporta OpenGL ES 2.0 y OpenCL 1.1

StemCell

- A pesar de tener una arquitectura innovadora, han sido peor que la competencia
 - Rendimiento de ZMS-40 inferior a NVidia Tegra 3
- No obstante, en decodificación de videos han superado a la competencia

StemCell

- Solo han sido utilizados en los productos de Creative y pocos más...
- No ha salido la nueva generación para seguir en competición
- No obstante, por la compra de Intel se puede pensar que volverán a aparecer

PowerVR

- División de la empresa británica Imagination Technologies
- Al principio intentaron competir en el mercado de GPUs para ordenadores
 - Pero han sido eliminados, como muchos otros
- Pero se han quedado en el mercado de GPUs móviles utilizando la misma arquitectura

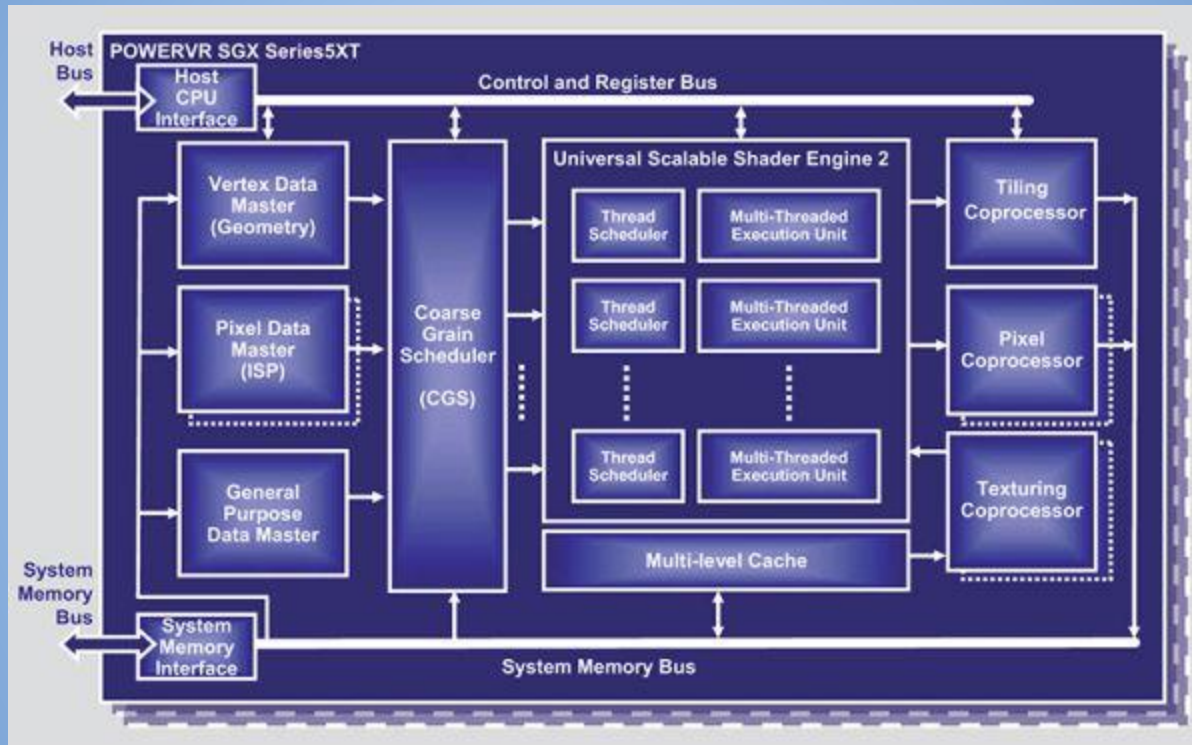
PowerVR

- Son conocidas sobretodo por ser las GPUs en productos de Apple
 - Estan en iPad y iPhone
- Durante mucho tiempo han sido referencia en cuanto al rendimiento
 - A pesar de que sus GPUs, normalmente son más antiguas que la competencia

PowerVR

- La serie actual es la serie 5
- GPUs SGX545, SGX543, SGX544, SGX554
 - Todos de los años 2009 y 2010
- Soportan OpenGL ES 2.0 y, algunos, OpenCL 1.1

PowerVR



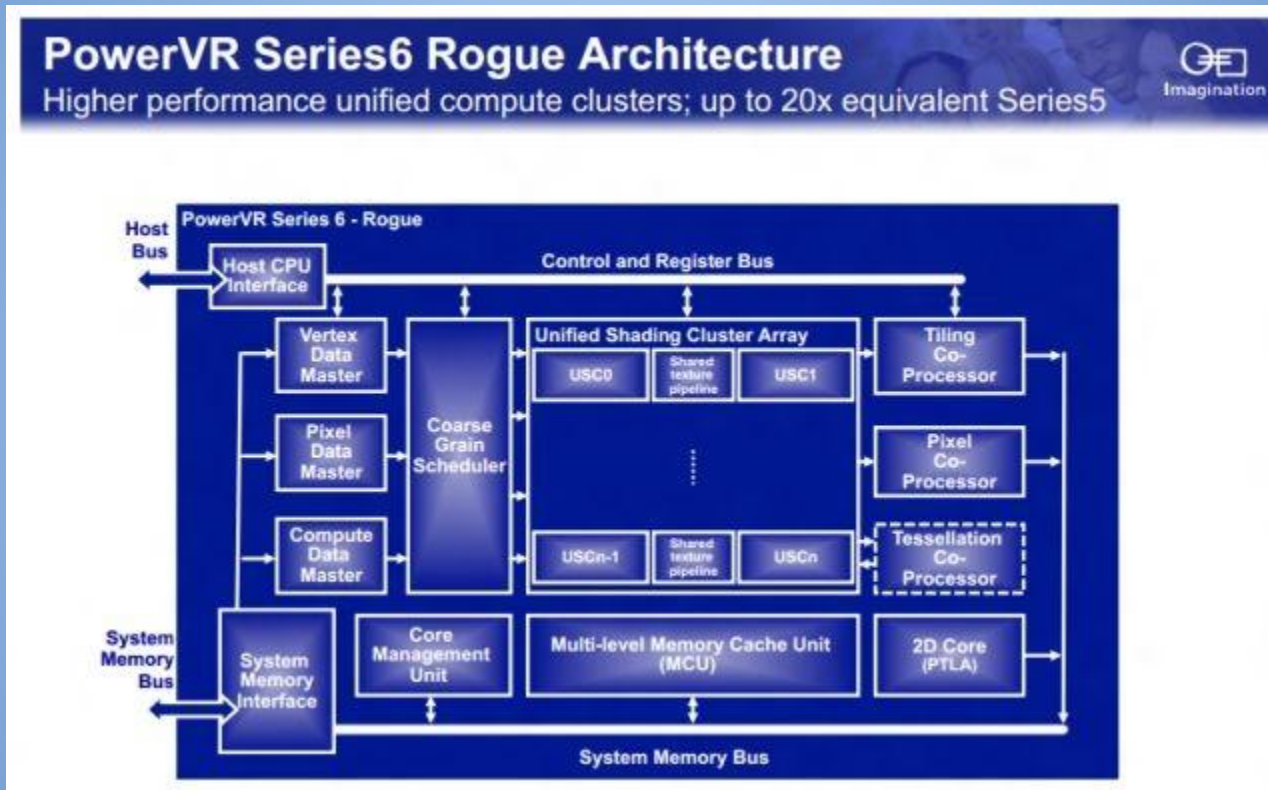
PowerVR

- Shaders unificados
- Las únicas GPUs que utilizan el método de renderizado TBDR
 - Especialmente eficiente en GPUs móviles
 - Reduce el bandwidth, elimina los píxeles invisibles

PowerVR

- La serie 6, presentada en el 2012, la más moderna
- Sigue con TBDR
- Soporta OpenGL ES 3.0 y OpenCL
- Evolución de la arquitectura
 - Se parece mucho a la arquitectura de GPUs de AMD

PowerVR



- Unified Shading Cluster Array

PowerVR

- El modelo de negocio de Imagination Technologies es el mismo que de ARM
 - No producen GPUs, solo venden licencias
- Todavía en el mercado no hay productos con GPUs de la serie 6
 - Pero han habido compradores de licencias, como Texas Instruments

SmartGPU

- Finalmente, como una pequeña curiosidad...
- Un acelerador gráfico para microcontroladores
- Facilita mucho el desarrollo de interfaces gráficas

SmartGPU



SmartGPU

- Permite dibujar fácilmente simples objetos geométricos
 - Como líneas, rectángulos
 - Con llamadas al estilo `drawLine()`
- Permite escribir el texto en la pantalla
 - Con fuentes y tamaño de letra diferentes
- Resolución de pantalla hasta 320 por 240 píxeles

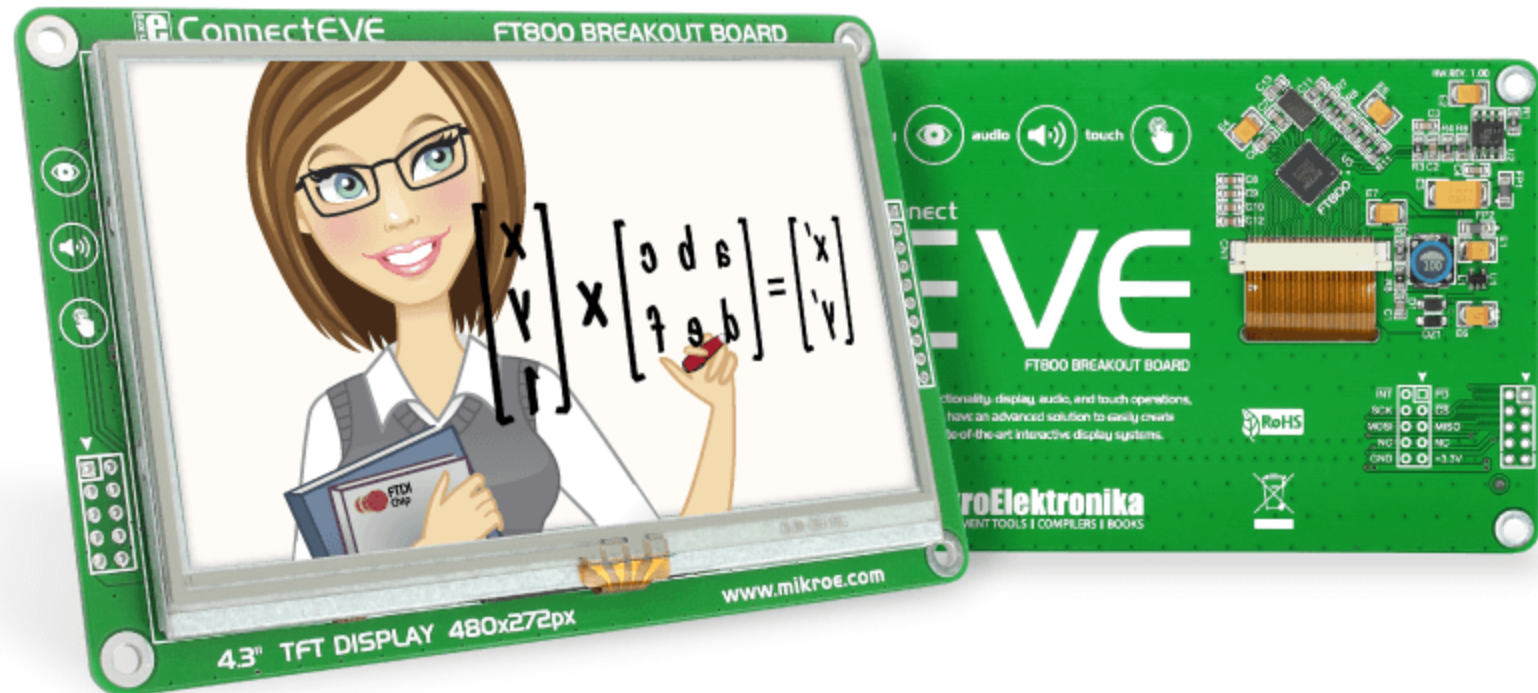
SmartGPU



SmartGPU

- Existen varias soluciones
- Algunos permiten visualizar hasta simples objetos 3D
- Se pueden conectar a microcontroladores de tan solo 8bits
- También son aceleradores





ConnectEVE