

UNIVERSIDAD DE OVIEDO

**MÁSTER EN TECNOLOGÍAS DE LA INFORMACIÓN Y
COMUNICACIONES EN REDES MÓVILES - TICRM**

TESIS DE MÁSTER

**DISEÑO, IMPLEMENTACIÓN Y VALIDACIÓN DE UNA
ESTACIÓN DE TRABAJO PARA LA RESOLUCIÓN
EFICIENTE DE PROBLEMAS DE SCATTERING**

YASSIN ALHARIRI NÚÑEZ

JULIO 2012

**TUTOR: JESÚS ALBERTO LÓPEZ FERNÁNDEZ
COTUTOR: MIGUEL LÓPEZ PORTUGUÉS**

RESUMEN

La presente Tesis aborda un problema crítico de la actualidad como es la eficiencia energética de los recursos computacionales. Esto resulta de vital importancia, ya que para mantener los centros de datos operativos, se requiere gran cantidad de consumo energético. El objetivo fundamental de esta Tesis se centra en el diseño, desarrollo e implementación de una estación de trabajo que permita obtener un alto rendimiento computacional al tiempo que un reducido consumo energético para la resolución de problemas de scattering, tanto acústico como electromagnético, mediante herramientas software previamente diseñadas por miembros del Área de conocimiento a la que pertenecen los directores de la Tesis.

El trabajo se ha dividido en varias etapas:

Primero, se ha realizado un estudio de las tecnologías que se encuentran en el mercado del hardware, como son los microprocesadores y aceleradores gráficos, para conocer el funcionamiento de su arquitectura, las prestaciones que aportan y la evolución en el segmento tecnológico, centrandó el interés mayoritariamente en la eficiencia energética de cada componente.

En segundo lugar, se ha realizado un estudio teórico para poder seleccionar los componentes que permitan construir un prototipo eficiente desde el punto de vista energético y computacional. De manera que sea capaz de ejecutar software de propósito general consumiendo la mínima cantidad de energía.

Finalmente, se ha realizado el ensamblaje y pruebas de funcionamiento del prototipo diseñado, comparando su rendimiento computacional y eficiencia energética con una estación de trabajo previamente disponible a la que se ha denominado PC de referencia.

PALABRAS CLAVE

Rendimiento – Eficiencia energética - Hardware – Sistemas heterogéneos – HPC.

ÍNDICE GENERAL

1. INTRODUCCIÓN.....	1
1.1. MOTIVACIÓN.....	2
1.2. OBJETIVOS	2
1.3. REVISIÓN DEL ESTADO DEL ARTE	3
1.4. ORGANIZACIÓN DE LA TESIS	4
2. COMPUTACIÓN DE ALTO RENDIMIENTO.....	6
2.1. LEY DE AMDAHL.....	6
2.2. TAXONOMÍA DE FLYNN	7
2.2.1. SISD.....	7
2.2.2. SIMD	7
2.2.3. MISD	8
2.2.4. MIMD.....	9
3. ESTUDIO DE LAS DIFERENTES ARQUITECTURAS DE PROCESADORES.....	11
3.1. CPUS.....	11
3.1.1. X86-64.....	11
3.1.1.1. Fabricante Intel	11
3.1.1.1.1. Arquitectura Nehalem	11
3.1.1.1.1.1. Características de la arquitectura.....	11
3.1.1.1.1.2. Eficiencia energética.....	13
3.1.1.1.1.3. Mercado basado en la arquitectura	14
3.1.1.1.2. Arquitectura Sandy Bridge.....	14
3.1.1.1.2.1. Características de la arquitectura.....	15
3.1.1.1.2.2. Eficiencia energética.....	16
3.1.1.1.2.3. Mercado basado en la arquitectura	18
3.1.1.2. Fabricante Amd	18
3.1.1.2.1. Arquitectura K10.....	18
3.1.1.2.1.1. Phenom ii.....	18
3.1.1.2.1.1.1. Características de la arquitectura.....	18
3.1.1.2.1.1.2. Eficiencia energética.....	19
3.1.1.2.1.1.3. Mercado basado en la arquitectura	20
3.1.1.2.1.2. Athlon ii.....	20
3.1.1.2.1.2.1. Características de la arquitectura.....	20
3.1.1.2.1.2.2. Eficiencia energética.....	21
3.1.1.2.1.2.3. Mercado basado en la arquitectura	22
3.1.1.2.2. Arquitectura Bulldozer.....	22
3.1.1.2.2.1. Características de la arquitectura.....	22
3.1.1.2.2.2. Eficiencia energética.....	23
3.1.1.2.2.3. Mercado basado en la arquitectura	23
3.1.2. Arm.....	24
3.1.2.1. Procesadores clásicos.....	24
3.1.2.2. Procesadores cortex.....	26
3.1.3. Sparc.....	28
3.1.3.1. Características generales	29
3.1.3.2. Sparc viiifx	29

3.1.4.	<i>Mips</i>	30
3.1.4.1.	<i>Características Mips64</i>	30
3.1.5.	<i>PowerPC</i>	31
3.1.5.1.	<i>PowerPC 970</i>	31
3.2.	ACCELERADORES HARDWARE	32
3.2.1.	<i>Fabricante Nvidia</i>	32
3.2.1.1.	<i>Tecnología Fermi</i>	32
3.2.1.1.1.	<i>Especificaciones generales</i>	32
3.2.1.1.2.	<i>Arquitectura Fermi</i>	33
3.2.1.1.3.	<i>Arquitectura de la caché</i>	35
3.2.1.1.4.	<i>Mercado basado en la arquitectura</i>	35
3.2.1.1.4.1.	<i>Tesla</i>	35
3.2.1.1.4.2.	<i>Geforce</i>	36
3.2.1.2.	<i>Tecnología Kepler</i>	37
3.2.1.2.1.	<i>Especificaciones generales</i>	38
3.2.1.2.2.	<i>Eficiencia energética</i>	38
3.2.1.2.2.1.	<i>GPU Boost</i>	38
3.2.1.2.3.	<i>Mercado basado en la arquitectura</i>	39
3.2.1.3.	<i>Tecnología TEGRA</i>	39
3.2.2.	<i>Fabricante Ati</i>	40
3.2.2.1.	<i>Serie HD 6000 (Northern Islands)</i>	40
3.2.2.1.1.	<i>Mercado basado en la arquitectura</i>	41
3.2.2.2.	<i>Serie HD 7000 (Southern Islands)</i>	41
3.2.2.2.1.	<i>Eficiencia energética</i>	42
3.2.2.2.2.	<i>Mercado basado en la arquitectura</i>	42
3.2.2.3.	<i>Apus Amd (Amd Fusion)</i>	43
3.2.2.3.1.	<i>Plataformas actuales</i>	43
3.2.2.3.2.	<i>Mercado basado en la arquitectura</i>	43
4.	DISEÑO Y ELECCIÓN	45
4.1.	ESTUDIO TEÓRICO Y SELECCIÓN DEL PROTOTIPO	45
4.1.1.	<i>Estudio de la CPU</i>	45
4.1.1.1.	<i>Arquitectura x86-64</i>	45
4.1.1.2.	<i>Otras arquitecturas</i>	50
4.1.2.	<i>Estudio de la GPU</i>	50
5.	PRUEBAS DE RENDIMIENTO	52
5.1.	MÉTRICAS UTILIZADAS	52
5.2.	BENCHMARKS UTILIZADOS	52
5.2.1.	<i>Whetstone</i>	53
5.2.2.	<i>Linpack</i>	54
5.2.3.	<i>Phoronix Test Suite</i>	54
5.2.4.	<i>Códigos del Área de TSC de la Universidad de Oviedo</i>	55
5.3.	RESULTADOS	56
5.3.1.	<i>Whetstone</i>	56
5.3.2.	<i>MatLab</i>	57
5.3.3.	<i>Linpack</i>	62
5.3.4.	<i>Phoronix Test Suite</i>	64

5.3.5. Códigos del Área de TSC de la Universidad de Oviedo	68
6. CONCLUSIONES	73
BIBLIOGRAFÍA.....	75
ANEXOS.....	80

ÍNDICE DE FIGURAS

Figura 1. Gasto energético ^[1] .	1
Figura 2. Gráfico CPU vs GPU por Nvidia ^[4] .	2
Figura 3. Ilustración de la Ley de Amdahl.	6
Figura 4. Arquitectura SISD.	7
Figura 5. Arquitectura SIMD.	8
Figura 6. Arquitectura MISD.	8
Figura 7. Memoria compartida en la arquitectura MIMD.	9
Figura 8. Memoria distribuida en la arquitectura MIMD.	9
Figura 9. Arquitectura MIMD.	10
Figura 10. Intel QuickPath Interconnect.	12
Figura 11. Tecnología Hyper-Threading.	12
Figura 12. Fase 1 (Turbo Boost desactivado).	13
Figura 13. Fase 2 (Turbo Boost activado).	13
Figura 14. Arquitectura Sandy Bridge ^[22] .	15
Figura 15. Turbo Boost 2.0 desactivado.	17
Figura 16. Turbo Boost 2.0 activado.	17
Figura 17. GPC Fermi ^[57] .	33
Figura 18. Composición SM.	34
Figura 19. GPU Boost.	39
Figura 20. Variación del rendimiento normalizado utilizando el R.T. y SPECfp2006.	48
Figura 21. Variación del rendimiento normalizado utilizando el R.T. y SPECfp_rate2006.	49
Figura 22. Multímetro Velleman DVM890.	52
Figura 23. Eficiencia energética proporcionada por Whetstone para 500000 repeticiones.	56
Figura 24. Resultados del tiempo (en centésimas de segundo) y energía requerida al realizar el benchmark interno de MatLab.	57
Figura 25. Comparación del consumo energético al realizar el producto de matrices.	58
Figura 26. Vista en detalle del consumo energético al realizar el producto de matrices.	59
Figura 27. Comparación de la eficiencia (EDP) al realizar el producto de matrices.	60
Figura 28. Vista en detalle de la eficiencia (EDP) al realizar el producto de matrices.	60
Figura 29. Comparación del consumo energético al aplicar la FFT.	61
Figura 30. Comparación de la eficiencia (EDP) al aplicar la FFT.	62
Figura 31. Eficiencia energética medida mediante Linpack para un problema de tamaño 200x200.	63
Figura 32. Eficiencia energética medida mediante Linpack para un problema de tamaño 2048x2048.	63
Figura 33. Eficiencia energética medida mediante los benchmarks de Phoronix.	64
Figura 34. Comparación Phoronix tiempo (Segundos) parte 1.	65
Figura 35. Comparación Phoronix tiempo (Segundos) parte 2.	65
Figura 36. Comparación Phoronix energía parte 1.	66
Figura 37. Comparación Phoronix energía parte 2.	66
Figura 38. Comparación Phoronix eficiencia energética parte 1.	67
Figura 39. Comparación Phoronix eficiencia energética parte 2.	67
Figura 40. Tiempo requerido para el código MST.	68
Figura 41. Consumo de energía para el código MST.	69
Figura 42. Eficiencia energética para el código MST.	69
Figura 43. Tiempo requerido para el código FMM.	70
Figura 44. Consumo de energía para el código FMM.	70
Figura 45. Eficiencia energética para el código FMM.	71

Figura 46. Tiempo requerido para el código FMM con FFT.	71
Figura 47. Consumo de energía para el código FMM con FFT.	72
Figura 48. Eficiencia energética para el código FMM con FFT.	72
Figura 49. SMX GeForce Serie GTX 680 ^[64]	95

ÍNDICE DE TABLAS

Tabla 1. Familia Athlon II ^[33]	22
Tabla 2. Familia AMD FX Bulldozer ^[36]	24
Tabla 3. Microprocesadores Cortex-A ^[37]	27
Tabla 4. Microprocesadores Cortex-R ^[37]	27
Tabla 5. Microprocesadores Cortex-M ^[37]	28
Tabla 6. PowerPC Serie 970.	31
Tabla 7. Comparación modelos Nvidia.....	33
Tabla 8. Arquitectura caché.	35
Tabla 9. Comparativa Tesla ^[59]	36
Tabla 10. Nvidia GeForce Serie GTX 500 ^{[61] [62]}	37
Tabla 11. Nvidia GeForce Serie GT 500 ^{[61] [62]}	37
Tabla 12. Comparación modelos Nvidia 2 ^[64]	38
Tabla 13. Nvidia GeForce Serie GTX 600 ^[62]	39
Tabla 14. AMD Radeon Serie 6000 ^[70]	41
Tabla 15. AMD Radeon Serie 7000 ^{[73] [74]}	42
Tabla 16. Familia AMD APu ^[77]	44
Tabla 17. Datos Top500. Sparc Viiiifx 2.0 GHz.....	50
Tabla 18. Datos Top500. PowerPC BQC 1.60 GHz.	50
Tabla 19. Planificación temporal.	80
Tabla 20. Generación Core i3.	86
Tabla 21. Generación Core i5.	86
Tabla 22. Generación Core i7.	87
Tabla 23. 2º Generación Core i3.	90
Tabla 24. 2º Generación Core i5.	90
Tabla 25. 2º Generación Core i7.	91
Tabla 26. Micro Phenom II.	94
Tabla 27. Resultados de Whetstone en el PC de referencia para 500000 repeticiones.	96
Tabla 28. Resultados de Whetstone en el prototipo para 500000 repeticiones.	96
Tabla 29. Resultados del benchmark interno de Matlab en el PC de referencia.	97
Tabla 30. Resultados del producto de matrices en el PC de referencia (doble precisión).	97
Tabla 31. Resultados de aplicar la FFT en el PC de referencia (doble precisión).	97
Tabla 32. Resultados del producto de matrices en el PC de referencia (simple precisión).....	97
Tabla 33. Resultados de aplicar la FFT en el PC de referencia (simple precisión).....	98
Tabla 34. Resultados del benchmark interno de Matlab en el prototipo.	98
Tabla 35. Resultados del producto de matrices en el prototipo (doble precisión).	98
Tabla 36. Resultados de aplicar la FFT en el prototipo (doble precisión).	98
Tabla 37. Resultados del producto de matrices en el prototipo (simple precisión).....	99
Tabla 38. Resultados de aplicar la FFT en el prototipo (simple precisión).....	99
Tabla 39. Resultados del producto de matrices en la GPU del prototipo (doble precisión).....	99
Tabla 40. Resultados de aplicar la FFT en la GPU del prototipo (doble precisión).	99
Tabla 41. Resultados del producto de matrices en la GPU del prototipo (simple precisión).....	99
Tabla 42. Resultados de aplicar la FFT en la GPU del prototipo (simple precisión).....	100
Tabla 43. Resultados para una matriz de 200x200 en 1 core 1º hilo del PC de referencia.	101
Tabla 44. Resultados para una matriz de 200x200 en 2 cores 1º hilo del PC de referencia.....	101
Tabla 45. Resultados para una matriz de 200x200 en 2 cores 2º hilo del PC de referencia.....	101
Tabla 46. Resultados para una matriz de 200x200 en 3 cores 1º hilo del PC de referencia.....	101

Tabla 47. Resultados para una matriz de 200x200 en 3 cores 2º hilo del PC de referencia.....	101
Tabla 48. Resultados para una matriz de 200x200 en 3 cores 3º hilo del PC de referencia.....	102
Tabla 49. Resultados para una matriz de 200x200 en 4 cores 1º hilo del PC de referencia.....	102
Tabla 50. Resultados para una matriz de 200x200 en 4 cores 2º hilo del PC de referencia.....	102
Tabla 51. Resultados para una matriz de 200x200 en 4 cores 3º hilo del PC de referencia.....	102
Tabla 52. Resultados para una matriz de 200x200 en 4 cores 4º hilo del PC de referencia.....	102
Tabla 53. Resultados para una matriz de 2048x2048 en 1 core 1º hilo del PC de referencia.	102
Tabla 54. Resultados para una matriz de 2048x2048 en 2 cores 1º hilo del PC de referencia.....	103
Tabla 55. Resultados para una matriz de 2048x2048 en 2 cores 2º hilo del PC de referencia.....	103
Tabla 56. Resultados para una matriz de 2048x2048 en 3 cores 1º hilo del PC de referencia.....	103
Tabla 57. Resultados para una matriz de 2048x2048 en 3 cores 2º hilo del PC de referencia.....	103
Tabla 58. Resultados para una matriz de 2048x2048 en 3 cores 3º hilo del PC de referencia.....	103
Tabla 59. Resultados para una matriz de 2048x2048 en 4 cores 1º hilo del PC de referencia.....	103
Tabla 60. Resultados para una matriz de 2048x2048 en 4 cores 2º hilo del PC de referencia.....	104
Tabla 61. Resultados para una matriz de 2048x2048 en 4 cores 3º hilo del PC de referencia.....	104
Tabla 62. Resultados para una matriz de 2048x2048 en 4 cores 4º hilo del PC de referencia.....	104
Tabla 63. Resultados para una matriz de 200x200 en 1 core 1º hilo del prototipo.	104
Tabla 64. Resultados para una matriz de 200x200 en 2 cores 1º hilo del prototipo.	104
Tabla 65. Resultados para una matriz de 200x200 en 2 cores 2º hilo del prototipo.	105
Tabla 66. Resultados para una matriz de 200x200 en 3 cores 1º hilo del prototipo.	105
Tabla 67. Resultados para una matriz de 200x200 en 3 cores 2º hilo del prototipo.	105
Tabla 68. Resultados para una matriz de 200x200 en 3 cores 3º hilo del prototipo.	105
Tabla 69. Resultados para una matriz de 200x200 en 4 cores 1º hilo del prototipo.	105
Tabla 70. Resultados para una matriz de 200x200 en 4 cores 2º hilo del prototipo.	106
Tabla 71. Resultados para una matriz de 200x200 en 4 cores 3º hilo del prototipo.	106
Tabla 72. Resultados para una matriz de 200x200 en 4 cores 4º hilo del prototipo.	106
Tabla 73. Resultados para una matriz de 2048x2048 en 1 core 1º hilo del prototipo.	106
Tabla 74. Resultados para una matriz de 2048x2048 en 2 cores 1º hilo del prototipo.	106
Tabla 75. Resultados para una matriz de 2048x2048 en 2 cores 2º hilo del prototipo.	107
Tabla 76. Resultados para una matriz de 2048x2048 en 3 cores 1º hilo del prototipo.	107
Tabla 77. Resultados para una matriz de 2048x2048 en 3 cores 2º hilo del prototipo.	107
Tabla 78. Resultados para una matriz de 2048x2048 en 3 cores 3º hilo del prototipo.	107
Tabla 79. Resultados para una matriz de 2048x2048 en 4 cores 1º hilo del prototipo.	107
Tabla 80. Resultados para una matriz de 2048x2048 en 4 cores 2º hilo del prototipo.	108
Tabla 81. Resultados para una matriz de 2048x2048 en 4 cores 3º hilo del prototipo.	108
Tabla 82. Resultados para una matriz de 2048x2048 en 4 cores 4º hilo del prototipo.	108
Tabla 83. Resultados del test Phoronix. Primera parte en el PC de referencia.	109
Tabla 84. Resultados del test Phoronix. Segunda parte en el PC de referencia.	109
Tabla 85. Resultados del test Phoronix. Primera parte en el prototipo.	109
Tabla 86. Resultados del test Phoronix. Segunda parte en el prototipo.	110
Tabla 87. Resultados al ejecutar FMM en la CPU del PC de referencia sin uso de la FFT.....	111
Tabla 88. Resultados al ejecutar FMM en la CPU del PC de referencia con uso de la FFT.....	111
Tabla 89. Resultados al ejecutar MST en la CPU del PC de referencia.....	111
Tabla 90. Resultados al ejecutar FMM en la CPU del prototipo sin uso de la FFT.....	111
Tabla 91. Resultados al ejecutar FMM en la CPU del prototipo con uso de la FFT.....	111
Tabla 92. Resultados al ejecutar MST en la CPU del prototipo.....	112
Tabla 93. Resultados al ejecutar los códigos del área de TSC en la GPU del prototipo.....	112

1. INTRODUCCIÓN

La escasez de fuentes de energía y el impacto que el consumo energético tiene sobre el medioambiente está motivando en los últimos tiempos una toma de conciencia, tanto de las instituciones privadas como públicas, sobre la importancia del consumo eficiente de energía. Esta concienciación ha llegado también a las organizaciones relacionadas con las tecnologías de la información (TI), ya que los centros de datos y de supercomputación requieren un gran consumo energético.

El consumo energético por parte de las TI ha ido incrementándose de manera muy notable en los últimos años. En la figura 1, se muestra la evolución del gasto energético relacionado con el coste de mantenimiento y adquisición de servidores desde el año 1996 hasta el 2011.

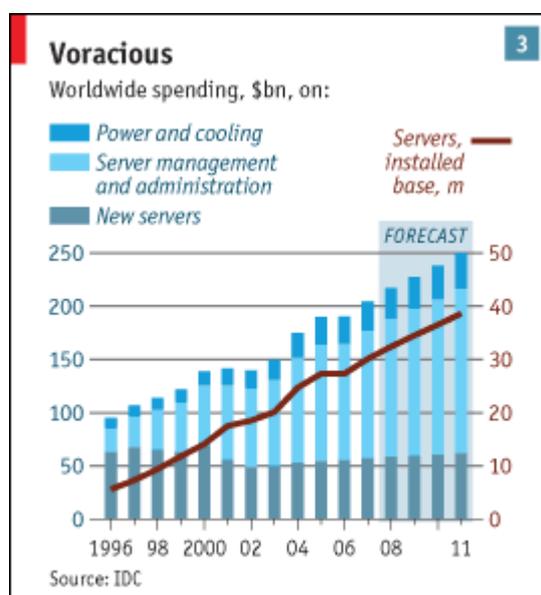


Figura 1. Gasto energético ^[1] .

Un punto significativo del gráfico anterior, es que el número de servidores que se encuentran en los centros de proceso de datos se incrementa al tiempo que crece el gasto energético utilizado para la gestión y refrigeración de esos mismos servidores. Esto implica que algunos centros de datos y de computación de alto rendimiento (High Performance Computing o HPC) estén llegando a sus límites en cuanto a consumo energético, capacidad de refrigeración y espacio físico.

En la actualidad, se pueden encontrar algunas organizaciones que están sobrepasando estos límites, con el problema asociado del incremento de los costes de energía. Además, se están planteando medidas provisionales en lugar de enfrentarse y resolver de forma eficaz sus problemas de refrigeración y consumo energético ^[1] . De hecho, bajo la tendencia actual, el coste de la energía usada para alimentar el hardware de TI por año excederá en poco tiempo el coste de adquisición del hardware ^[2] .

El problema de consumo energético que se ha mencionado, introduce lo que se va a plantear y desarrollar en la presente Tesis de Máster. A lo largo de esta memoria, se ha realizado un estudio de la tecnología disponible para reducir el consumo energético en el campo de la computación de alto rendimiento.

1.1. Motivación

Hasta mediados de la primera década de este siglo, y debido a las mejoras en el diseño y la fabricación de los procesadores (CPU), se consiguió elevar la frecuencia de reloj para realizar un número mayor de operaciones. Dicho aumento de frecuencia conllevó además un aumento en el consumo eléctrico para mantener esa potencia de cálculo de la CPU. Como consecuencia de ello, la frecuencia de reloj finalmente se estancó en torno a 3 GHz alrededor del año 2005. En su lugar, para seguir desarrollando mejoras en la computación, se comenzó a duplicar el número de núcleos ^[3]. Desde el año 2005 en adelante, las mejoras de rendimiento en la computación no provienen directamente de incrementar la frecuencia del reloj, sino que vienen de saber aprovechar el trabajo en paralelo de varios núcleos.

Además de utilizar los recursos propios de la CPU, en la última década se han empezado a utilizar las unidades de procesamiento gráfico (o Graphical Processing Unit, GPU) como herramientas de cálculo (ver figura 2), ya que disponen de múltiples núcleos de cálculo. Aunque originalmente estos chips se desarrollaron para dedicarse al procesamiento de gráficos, en la actualidad, empresas como Nvidia ofrecen un lenguaje propio para utilizarlas en el campo de la computación de alto rendimiento.

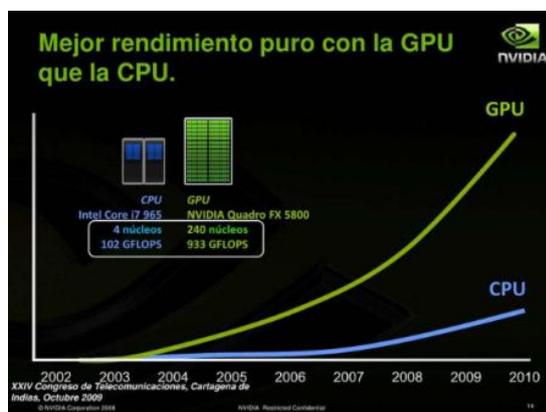


Figura 2. Gráfico CPU vs GPU por Nvidia ^[4].

La motivación de la presente Tesis radica en combinar adecuadamente las tecnologías CPU y GPU de manera que se integren en un prototipo de estación de trabajo que lleve asociada una alta eficiencia energética y computacional.

1.2. Objetivos

El objetivo principal de la presente Tesis de Máster consiste en el desarrollo de una estación de trabajo que combine tecnologías CPU y GPU disponibles en el mercado, de manera que se consiga un prototipo con una alta eficiencia energética y computacional. A tal fin, hay una serie de objetivos intermedios que se deben lograr.

Por una parte, en la fase de estudio teórico es necesario seleccionar una métrica adecuada que permita seleccionar los componentes más adecuados sin la necesidad de realizar medidas a priori. Asimismo, es de gran importancia la elección de un conjunto de pruebas de rendimiento (o benchmarks) que, basándose en la métrica escogida, permitan confirmar que el prototipo diseñado produce los resultados de eficiencia esperados.

Resulta también de importancia la difusión de los resultados a los que ha dado lugar este estudio, para que usuarios potenciales de este tipo de sistemas tengan una documentación de referencia.

1.3. Revisión del estado del arte

Como se ha reflejado en la motivación de la Tesis, el consumo energético es un verdadero problema a la hora de conseguir computadores más rápidos. Para evitar este sobreconsumo energético, los fabricantes continúan desarrollando nuevas y mejores soluciones.

Por un lado, está el fabricante Intel, el cual ha desarrollado en su arquitectura Sandy Bridge ^[20] la tecnología Turbo Boost 2.0 ^[24]. También ha mejorado la tecnología SpeedStep ^[18], además de producir un chip con un nuevo proceso de fabricación reducido, lo que implica menos consumo.

Dentro de la misma arquitectura de computadores (x86-64) se encuentra el fabricante AMD. La última tecnología en CPU que proporciona a día de hoy este fabricante se denomina Bulldozer ^[34]. Dicha arquitectura, obtiene una mejora en el consumo energético gracias a varios factores: reducción del voltaje, incorporación de un nuevo estado de alimentación para reducir el consumo, ajustes en los controles de encendido y apagado de la memoria, y uso de las tecnologías Cool'n'Quiet y PoweNow!^[30].

Finalmente, y dentro de las arquitecturas CPU, estarían los procesadores basados en la arquitectura RISC ^[37] ^[52] (Reduced Instruction Set Computer), que suelen presentar una reducida frecuencia de trabajo y una simplicidad en su juego de instrucciones, lo que les permite lograr un consumo muy reducido.

Además de las arquitecturas de CPU, actualmente se encuentran en desarrollo y continua mejora lenguajes como CUDA ^[57] y OPENCL. Dentro del ámbito de la computación técnica aplicada a resolver problemas de la ciencia y de la ingeniería, se ha podido contrastar que el uso de las tarjetas aceleradoras ofrece una buena solución para la computación de propósito general, tanto por su bajo consumo ^[3] en comparación con los supercomputadores más importantes como por su bajo coste de adquisición.

Por último, merece la pena destacar la existencia de multitud de iniciativas y actividades relacionadas con la eficiencia energética en sectores TI. En la actualidad, las actividades relacionadas con la Green TI están relacionadas con las empresas y también patrocinadas desde las administraciones públicas. A continuación, se destacan algunas de las actividades más importantes de este sector ^[2].

- Climate Savers: es un grupo sin ánimo de lucro de consumidores y fabricantes con conciencia ecológica. Se creó en 2007 por Google e Intel. La asociación está compuesta por fabricantes y clientes, unos se comprometen a crear productos que cumplan los objetivos de eficiencia energética y los otros se comprometen a adquirir dichos productos.
- The Green Grid: se basa en un consorcio global dedicado a la eficiencia energética centrada en los centros de procesamiento de datos y en la computación de negocio. Sus principales objetivos son: definir modelos y métricas significativas centradas en el cliente final, desarrollar nuevos estándares, métodos de medida y nuevas tecnologías para mejorar el

rendimiento en los centros de datos. Además de promover la adopción de estándares, medidas y tecnologías energéticamente eficientes.

- Energy Star: en 1992, la Agencia de Protección Medioambiental de EEUU lanzó un programa de carácter voluntario cuya finalidad era promover en los centros de datos la idea de eficiencia energética. Dicho programa tuvo una gran aceptación, y en la actualidad es de vital importancia, ya que sirve como fuente de información creíble y objetiva para que los consumidores y empresarios puedan tomar decisiones a la hora de renovar equipos de bajo consumo, implicando una mejora significativa en el medio ambiente.
- The Green500: es una base de datos que se actualiza cada seis meses y muestra información y clasificación de los supercomputadores más eficientes del mundo.

1.4. Organización de la Tesis

La presente Tesis se organiza en seis capítulos cuyo contenido se detalla a continuación.

En el segundo capítulo, denominado “Computación de alto rendimiento”, aparece definida la ley de Amdahl y la taxonomía de Flynn. Estas definiciones sirven para ayudar al lector a comprender los conceptos básicos sobre rendimiento y organización de los computadores dependiendo de su arquitectura.

Posteriormente en el tercer capítulo, denominado “Estudio de las diferentes arquitecturas de procesadores”, se recoge, en su primera parte, las principales tecnologías que utilizan los microprocesadores actuales de diversos fabricantes. Se centra en las tecnologías que han desarrollado los fabricantes para conseguir reducir el consumo energético de sus chips. En el mismo capítulo y cerrando el bloque de microprocesadores, se han explicado otras alternativas basadas en la arquitectura RISC, como por ejemplo procesadores ARM, procesadores PPC, etc. Estos últimos procesadores, son muy eficientes y su evolución parece situarlos como una opción muy interesante para el futuro.

La última parte del tercer capítulo, se centra en el tema de los aceleradores gráficos. En esta parte, se ha sintetizado la información sobre las arquitecturas GPU más influyentes para la ejecución de software de propósito general, englobando las soluciones aportadas por Nvidia y AMD Ati.

En el cuarto capítulo, denominado “Diseño y elección”, se explican los métodos, expresiones y decisiones que se han aplicado en el proceso de selección de la CPU y de la GPU con mayor eficiencia.

En el quinto capítulo, denominado “Pruebas de rendimiento”, se describen las métricas y benchmarks utilizados. Además, en la última parte del capítulo se muestran los resultados obtenidos de los diferentes experimentos realizados.

Por último, en el capítulo de conclusiones aparecen recogidas las conclusiones que se han obtenido a lo largo del estudio.

Asimismo, conviene mencionar la presencia de múltiples anexos recogidos en la parte final del documento. Es especialmente destacable el contenido del ANEXO I, en el que se encuentra la planificación temporal que se ha llevado a cabo para realizar la presente Tesis.

2. COMPUTACIÓN DE ALTO RENDIMIENTO

En los últimos años, se ha incrementado la necesidad de resolver nuevos problemas planteados por la ciencia y la ingeniería. Los requisitos de dichos problemas cada vez crecen más en cuanto a necesidades de cómputo, debido a que cada vez son más complejos y necesitan trabajar con grandes volúmenes de datos sin perder capacidad de respuesta en un tiempo limitado ^{[5][6]}.

El procesamiento paralelo de datos ha ganado terreno y se ha establecido como una herramienta esencial para conseguir dar solución a estos problemas. La computación de alto rendimiento se apoya en tecnologías como los clusters, supercomputadores y en el uso de los paradigmas de programación paralela.

Uno de los principales problemas de los sistemas paralelos es que los algoritmos son más complejos de diseñar y desarrollar, ya que la concurrencia introduce nuevos tipos de errores software, como las condiciones de carrera (este problema surge cuando el resultado depende del orden en que se ejecuten las instrucciones). Otro problema que aparece en la computación paralela es la comunicación y la sincronización entre las diferentes sub tareas. Para definir la ganancia en tiempo que se consigue en un programa al paralelizarlo, se puede emplear la ley de Amdahl ^{[7][8][9]}.

2.1. Ley de Amdahl

En el año 1967, Amdahl apreció que el hecho de agregar más procesadores, a partir de cierto valor, a un sistema de procesamiento paralelo no producía una mejora significativa de velocidad de procesamiento. La Ley de Amdahl sintetiza esta observación mediante la siguiente expresión:

$$A = \frac{1}{(1 - F_m) + \frac{F_m}{A_m}} \quad (1.1)$$

Donde A es la ganancia en velocidad conseguida en el sistema completo debido a la mejora de uno de los subsistemas, Am es el factor de mejora que se ha introducido en el subsistema mejorado, y Fm es la fracción de tiempo que el sistema utiliza el subsistema mejorado.

Por ejemplo, si se tiene una tarea que tiene dos partes independientes, A y B. Optamos porque B ocupa el 25% del tiempo total de computación en resolverse. Mejorando este tiempo en cinco veces, se puede observar en la figura 3 como el tiempo total de cómputo no mejora en gran medida. Por el contrario, si mejoramos en dos veces el tiempo de cómputo de A, se observa en la figura 3 que el tiempo total de cómputo tiene una gran mejora ^[10].

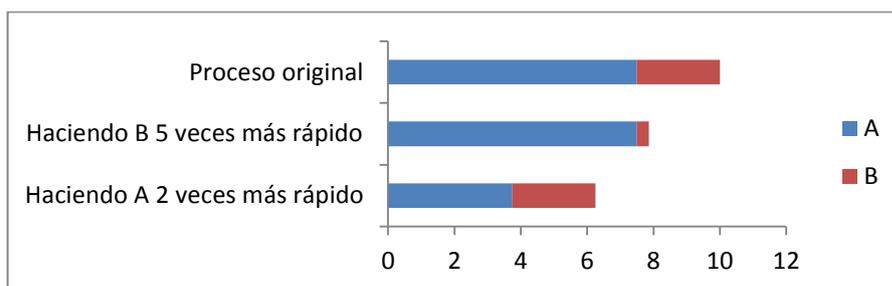


Figura 3. Ilustración de la Ley de Amdahl.

2.2. Taxonomía de Flynn

Una forma de clasificar las arquitecturas de computadores puede ser mediante la taxonomía de Michael Flynn. Flynn clasifica las arquitecturas teniendo en cuenta el número de procesadores, el número de programas que pueden ser ejecutados en los procesadores y la estructura de memoria ^[11]. La clasificación de Flynn se compone de 4 categorías: SISD, SIMD, MISD, MIMD.

2.2.1. SISD

Los computadores SISD (Single Instruction stream, Single Data stream) tienen una CPU la cual ejecuta una instrucción a la vez (single, instruction stream) y genera un ítem de datos por cada instrucción (single data stream). En la figura 4 se muestra la arquitectura SISD.



Figura 4. Arquitectura SISD.

En la arquitectura SISD existe un registro llamado contador de programa, el cual apunta a la instrucción que se está ejecutando e incrementa su valor. La instrucción a ejecutar es leída desde la memoria mientras que el registro contador se incrementa para apuntar a la siguiente instrucción.

2.2.2. SIMD

Los computadores SIMD (Single Instruction stream, Multiple Data stream) tienen una unidad de control que se encarga de ejecutar un flujo de instrucción simple, pero tienen más de un elemento procesando. La unidad de control genera las señales de control para todos los datos que se están procesando, es decir, muchos datos son llamados mediante una simple unidad de control. En la figura 5, se observa la arquitectura general.

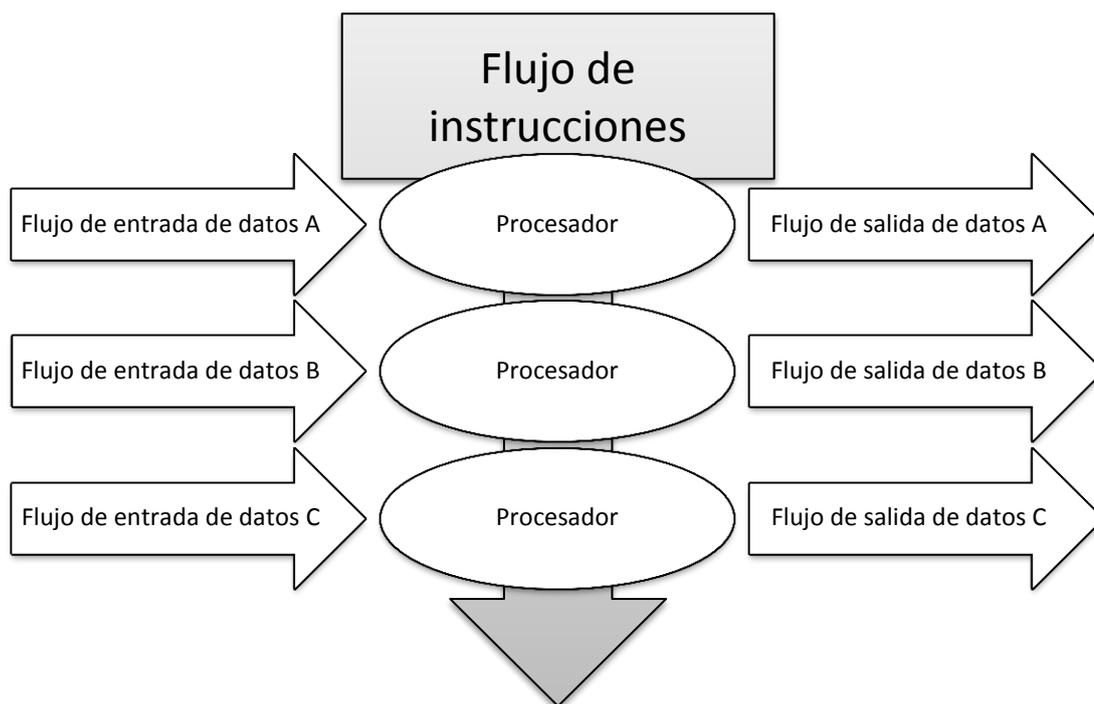


Figura 5. Arquitectura SIMD.

2.2.3. MISD

Los computadores MISD (Multiple Instruction stream, Single Data stream) pueden ejecutar varios programas distintos con el mismo ítem (vector, matriz,...) de datos. Esto implica que algunas instrucciones sólo requieren una parte de todos los datos para poder ejecutarse. Un ejemplo de esta arquitectura en el modelo pipeline.

Las arquitecturas en pipeline consisten en ir transformando un flujo de datos en un proceso comprendido por varias fases secuenciales, siendo la entrada de cada una la salida de la anterior ^[12]. Dichas arquitecturas entran dentro de este grupo de máquinas MISD. La razón por la que estas arquitecturas son agrupadas dentro de las máquinas MISD es que los elementos de un vector pueden tener el mismo grupo de datos, y todas las etapas del pipeline representan múltiples instrucciones que están siendo aplicadas al vector. En la figura 6, se observa la arquitectura general.

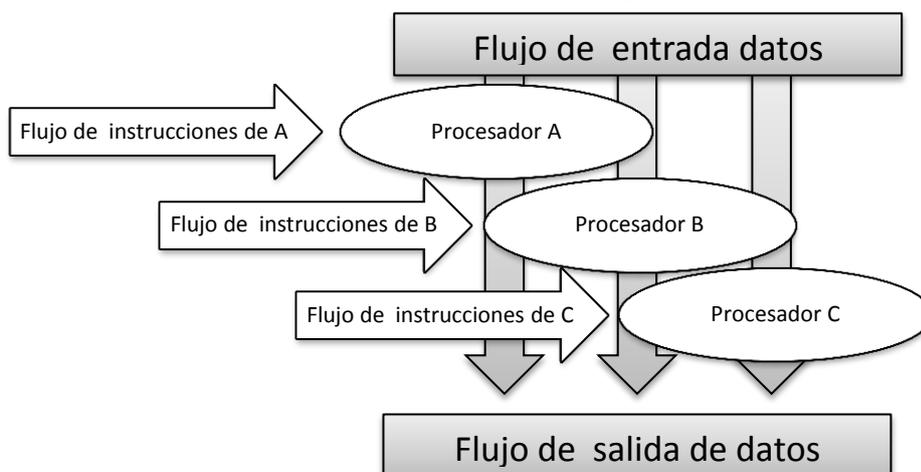


Figura 6. Arquitectura MISD.

2.2.4. MIMD

Los computadores MIMD (Multiple Instruction stream, Multiple Data stream) son también llamadas arquitecturas multiprocesadores. Estas máquinas tienen más de un procesador y cada procesador puede ejecutar un programa diferente con múltiples flujos de datos. En muchos sistemas MIMD cada procesador tiene acceso a una memoria global, la cual puede reducir el tiempo de comunicación de los procesadores, como se indica en la figura 7.

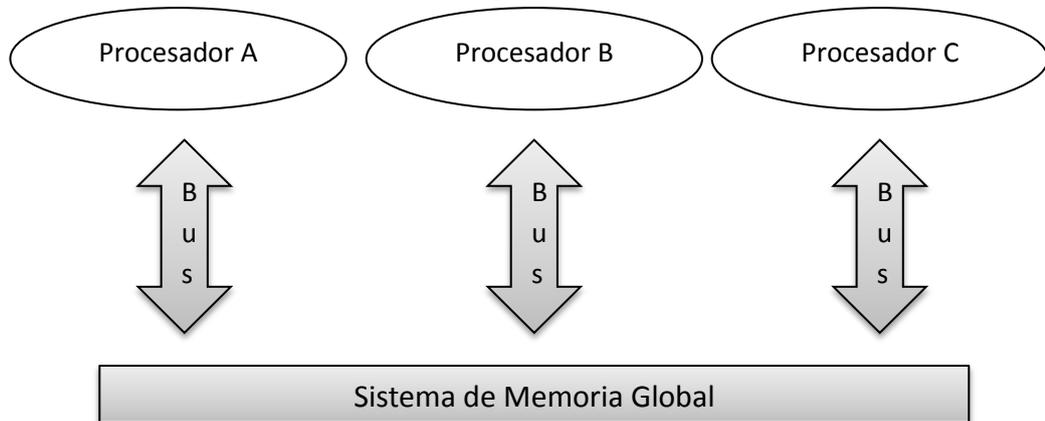


Figura 7. Memoria compartida en la arquitectura MIMD.

Además cada procesador tiene una memoria privada, como se ilustra en la figura 8. Muchas de las arquitecturas MIMD son utilizadas para la resolución de problemas concurrentes.

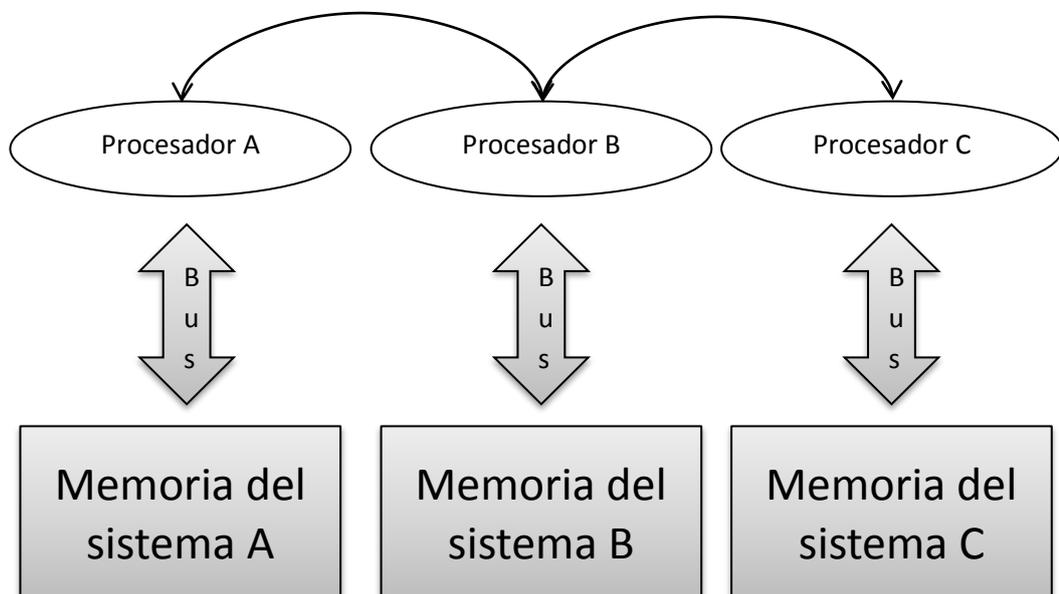


Figura 8. Memoria distribuida en la arquitectura MIMD.

Los computadores basados en la arquitectura MIMD son los más complejos, pero ofrecen grandes promesas para obtener solución a los problemas de forma concurrente. En la figura 9, se ilustra la arquitectura general de la categoría MIMD.

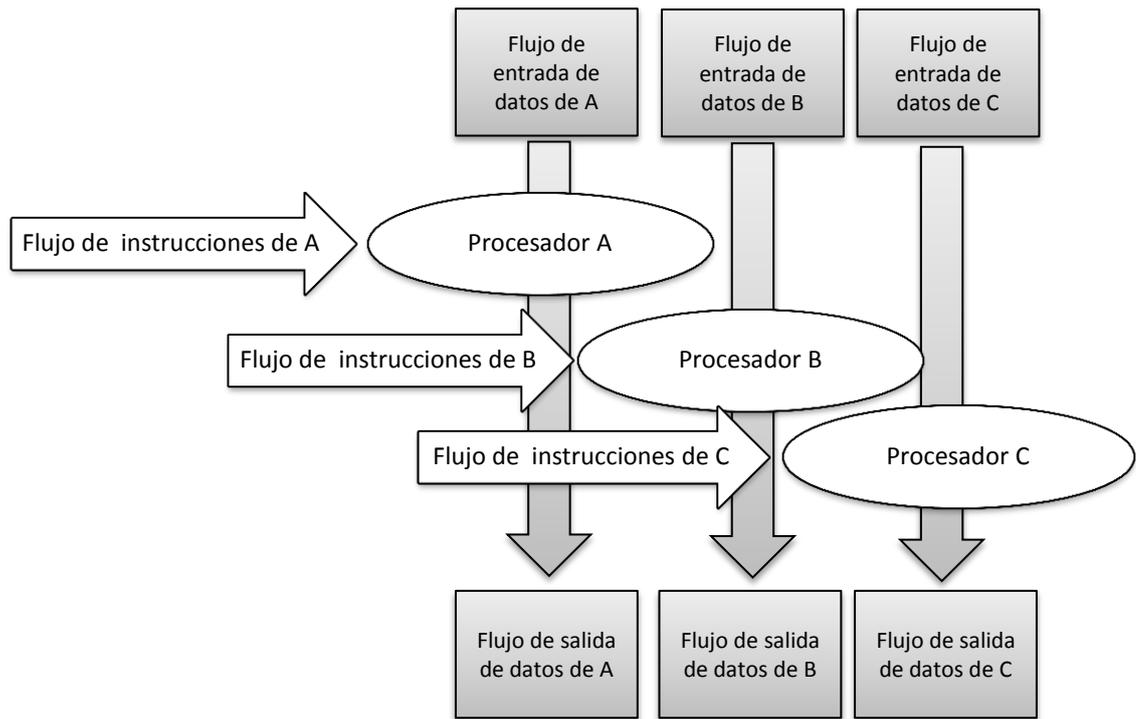


Figura 9. Arquitectura MIMD.

3. ESTUDIO DE LAS DIFERENTES ARQUITECTURAS DE PROCESADORES

La tecnología hardware se puede definir como la organización e interconexión de todos los componentes que forman el computador. Un computador está formado principalmente por memorias, unidades aritméticas, elementos de procesamiento y buses de datos. El sistema de memoria realiza funciones de almacenamiento, la unidad aritmético lógica y la unidad de punto flotante son las encargadas de realizar las operaciones, los buses sirven para comunicar diferentes componentes de datos, además de las interfaces de salida y entrada.

En este capítulo de la Tesis, se mostrarán diferentes componentes que se pueden encontrar en el mercado en la actualidad, y se detallarán sus características y las novedades incorporadas en lo relativo a eficiencia energética.

3.1. CPUs

La CPU o unidad central de procesamiento es la parte de un computador que se encarga de interpretar las instrucciones que contienen los programas y el procesamiento de datos ^[13]. Este bloque recogerá las arquitecturas de microprocesadores más eficientes de la actualidad.

3.1.1. X86-64

En este apartado se tratarán los fabricantes Intel y AMD (Advanced Micro Devices) cuyos microprocesadores (CPU) están basados en el juego de instrucciones x86 de 64 bits.

3.1.1.1. FABRICANTE INTEL

Las arquitecturas de computadores que el fabricante Intel está comercializando en la actualidad son: Nehalem y su evolución Sandy Bridge que se describen a continuación.

3.1.1.1.1. ARQUITECTURA NEHALEM

La arquitectura que ha sucedido a la anterior denominada Intel Core 2 se ha empezado a comercializar en 2008 con el nombre de Nehalem. El primer microprocesador lanzado con esta arquitectura ha sido el procesador de sobremesa Intel Core i7, mientras que para ordenadores portátiles esta arquitectura se empezó a ver a partir del año 2010.

3.1.1.1.1.1. CARACTERÍSTICAS DE LA ARQUITECTURA

A continuación se detallan las mejoras en las especificaciones de los procesadores de la familia Nehalem ^{[14] [15] [16] [17]}.

Proceso de fabricación: Los micros están formados por 731 millones de transistores para la variante de cuatro núcleos y 1170 millones de transistores para la variante de seis núcleos. La fabricación consta de procesadores de 45 a 32 nm. Además todos los procesadores son monolíticos, es decir, todos los procesadores están en un encapsulado.

IMC (Control de Memoria Integrado de Triple Canal DDR3): A diferencia de las generaciones anteriores de procesadores Intel, en Nehalem el controlador de memoria está fuera del chipset y se estableció dentro de la CPU, reduciendo con ello la latencia y aumentando el ancho de banda del que dispone la memoria. Este nuevo controlador permite manejar hasta 3 canales de memoria DDR3 con un ancho de banda máximo de 192 bits.

Gráfica integrada: El microprocesador dispone de un procesador de gráficos integrado (IGP) localizado en off-die, pero en el mismo paquete de CPU.

Nuevo modelo de interconexión: El nuevo modelo de interconexión punto-a-punto se denomina Intel QuickPath Interconnect, esta nueva tecnología reemplaza al bus frontal (FSB) que viene acompañando a los procesadores Intel desde antes incluso que los primeros Pentium. En la figura 10 se puede observar su funcionamiento.

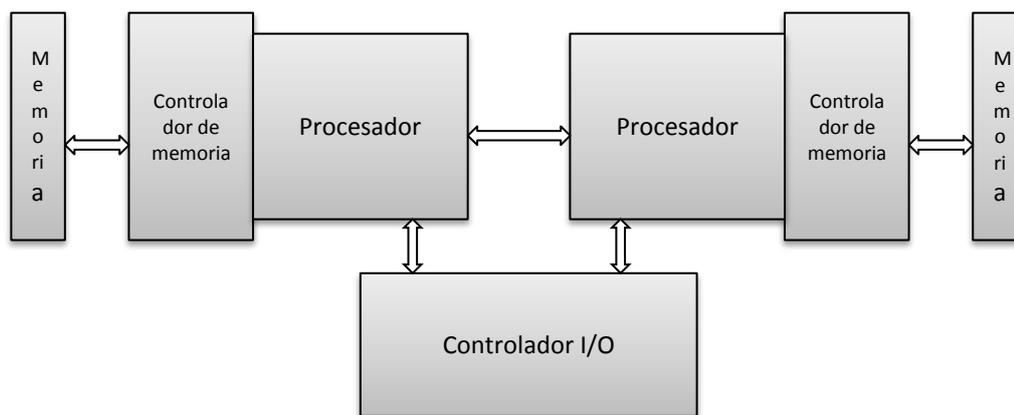


Figura 10. Intel QuickPath Interconnect.

HT (Hyper-Threading): Esta tecnología ya se había incorporado en algunos procesadores de la arquitectura Netburst (Pentium 4) y ahora se ha retomado. El Hyper-Threading consiste en que cada núcleo físico tiene capacidad para trabajar con 2 hilos de software aprovechando los tiempos muertos inevitables que hay al trabajar con un solo hilo. Obviamente no es lo mismo que tener dos núcleos físicos, pero en determinados contextos logra escalar el rendimiento decentemente. Por ejemplo un software altamente paralelizado puede aprovechar esta tecnología de muy buena manera y ocuparía casi totalmente los ciclos de cómputo disponibles. En la figura 11 se puede observar el funcionamiento.

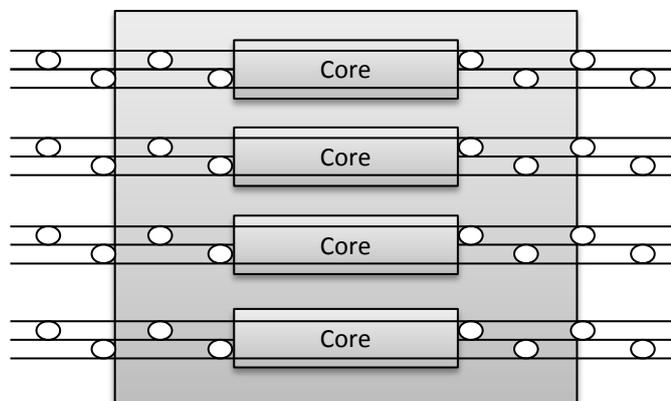


Figura 11. Tecnología Hyper-Threading.

Caché L3 (Memoria Caché L3 Compartida): En esta arquitectura Intel integró la memoria caché de tercer nivel compartida para todos los núcleos. La memoria L1 se mantiene en 64 KB por núcleo como pasaba en la anterior arquitectura, mientras que la memoria L2 bajó de 6MB compartidos a 256 KB dedicados por núcleo. Este cambio se compensó con la adicción de la memoria ya mencionada caché L3 de 8 MB.

3.1.1.1.2. EFICIENCIA ENERGÉTICA

En este apartado se detallarán dos importantes mejoras que se incluyen dentro de la arquitectura Nehalem. La primera se denomina SpeedStep y la segunda se denomina Turbo Boost ^{[16] [17] [18]}.

SpeedStep: La micro arquitectura Nehalem incorpora nuevas mejoras en la técnica SpeedStep que ya se había incorporado en los procesadores Pentium. Es una técnica de escalado dinámica de frecuencia que permite modificar la velocidad de reloj mediante software. Con esta técnica se pretende cambiar la frecuencia del reloj para minimizar el consumo y el calor disipado mientras el procesador está en reposo o con poca actividad.

Turbo (Intel Turbo Boost): En algunas ocasiones no se aprovechan todos los núcleos que tiene el micro procesador, además hay ocasiones que no requieren todo el potencial de la CPU. Con el fin de obtener más velocidad en esos momentos Intel incluye la tecnología Turbo Boost en su arquitectura. Esta tecnología consiste en darle prioridad al núcleo utilizado apagando los otros núcleos para no consumir más energía de forma innecesaria. Para “compensar” el reposo de los otros núcleos esta tecnología aumenta la frecuencia (vía aumento del multiplicador) del núcleo utilizado de forma dinámica de acuerdo al nivel de carga solicitada, lo que significa que en aplicaciones que no trabajen con todos los núcleos se puede acceder a una mayor velocidad de procesamiento de manera estable y sin riesgo de que la temperatura ni el consumo eléctrico se disparen. El funcionamiento de dicha tecnología queda representado en las siguientes figuras.

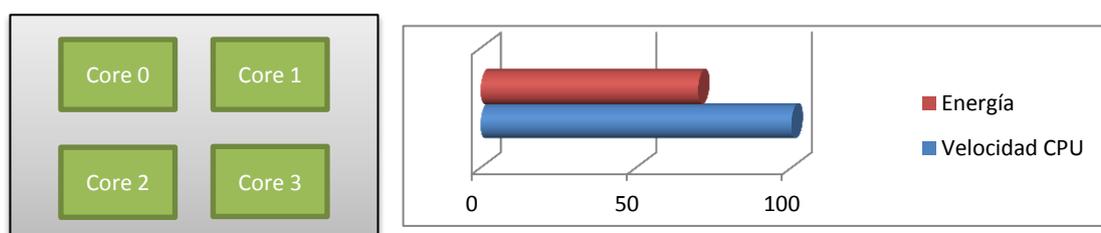


Figura 12. Fase 1 (Turbo Boost desactivado).

La primera fase se muestra en la figura 12. En ella, se puede observar un ejemplo donde el microprocesador está trabajando con los 4 cores de los que dispone a un ritmo moderado, lo que produce un consumo medio de los recursos eléctricos. En la siguiente imagen se puede ver teóricamente qué sucede cuando se activa la tecnología Turbo Boost.

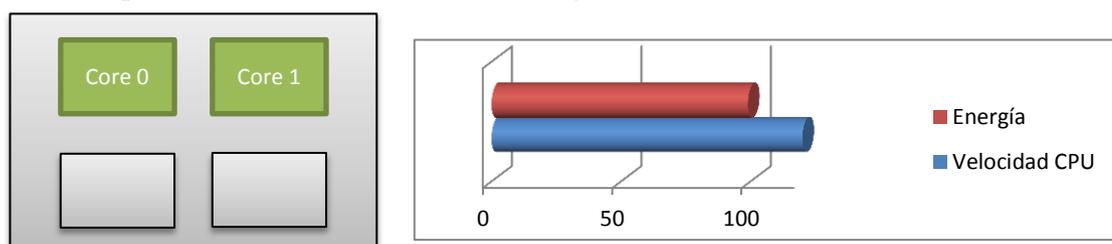


Figura 13. Fase 2 (Turbo Boost activado).

Al activar la tecnología Turbo Boost se produce un aumento de energía, que se entrega a la CPU para utilizar la capacidad de potencia adicional. La figura 13 retrata un ejemplo en el que sólo dos núcleos están activos para ofrecer una velocidad de cómputo más elevada gracias al aumento de la frecuencia producido por el incremento de energía.

3.1.1.1.3. MERCADO BASADO EN LA ARQUITECTURA

El fabricante Intel aprovechando el cambio de arquitectura hacia Nehalem ha comercializado tres variantes de productos que son Core i3, Core i5 y Core i7, refiriéndose a los niveles de prestaciones: bajo nivel (i3), nivel medio (i5) y gamma alta (i7). Acorde con la arquitectura Nehalem todos comparten el uso de la tecnología Intel QuickPath Interconnect. Debido a la nueva interconexión de puertos I/O, estos procesadores no son compatibles con las placas de generaciones pasadas de Intel, lo que no permite el simple remplazo de los anteriores ^[19].

Core i3: Los procesadores i3 son la primera etapa dentro de los Intel Core que supusieron una auténtica revolución en compromiso entre rendimiento y consumo. Basándose en los modelos dual-core, apostaron por la introducción del sistema Intel Hyper-Threading para dar soporte a ejecución multi-hilo. Las nuevas prestaciones introducidas por el Core i3 permitieron el uso de portátiles para actividades antes impensables como la edición y codificación de video, creación de contenidos en HD o el uso de videojuegos.

Core i5: La familia de procesadores Core i5 sigue la estela de los Core i3 pero dando un paso más en el nivel de prestaciones. Aunque con un rendimiento más modesto que los posteriores Core i7, los Core i5 buscaban ser procesadores asequibles y útiles para usuarios exigentes.

Core i7: En 2008 fue lanzado el primer microprocesador de 4 núcleos de la familia Intel Core, hasta que en marzo de 2010 se lanzó el primer microprocesador de la familia Intel Core de 6 núcleos creándose la gama Core i7 Extreme Edition.

Para ver el mercado actual de los microprocesadores que se están comercializando se debe acudir al ANEXO II.

3.1.1.1.2. ARQUITECTURA SANDY BRIDGE

Sandy Bridge es el nombre de la arquitectura de microprocesadores sucesora de Nehalem, se ha presentado a finales de 2010 y se empezó a comercializar a principios de 2011. Esta arquitectura se denomina la arquitectura Intel Core de segunda generación. Los microprocesadores que incorporan su tecnología vienen definidos como ix-2xxx ^{[20] [21]}.

Intel se ha caracterizado por crear una arquitectura y en la siguiente generación mejorarla. Sandy Bridge pertenece a la primera etapa, es decir, Sandy Bridge es una nueva micro arquitectura. Hay que destacar que a mediados o finales de 2012 se espera que se comercialicen los nuevos procesadores (mejora de los Sandy Bridge) denominados Ivy Bridge.

La nueva arquitectura se muestra en la figura 14, en ella se observa que el chip gráfico es un claro protagonista. La información que ha proporcionado el fabricante en cuanto a los transistores es: El chip gráfico está formado por unos 114 millones de transistores y cada núcleo tiene 55 millones de

transistores, mientras que el número total de transistores de todo el microprocesador asciende a 955 millones ^[22].

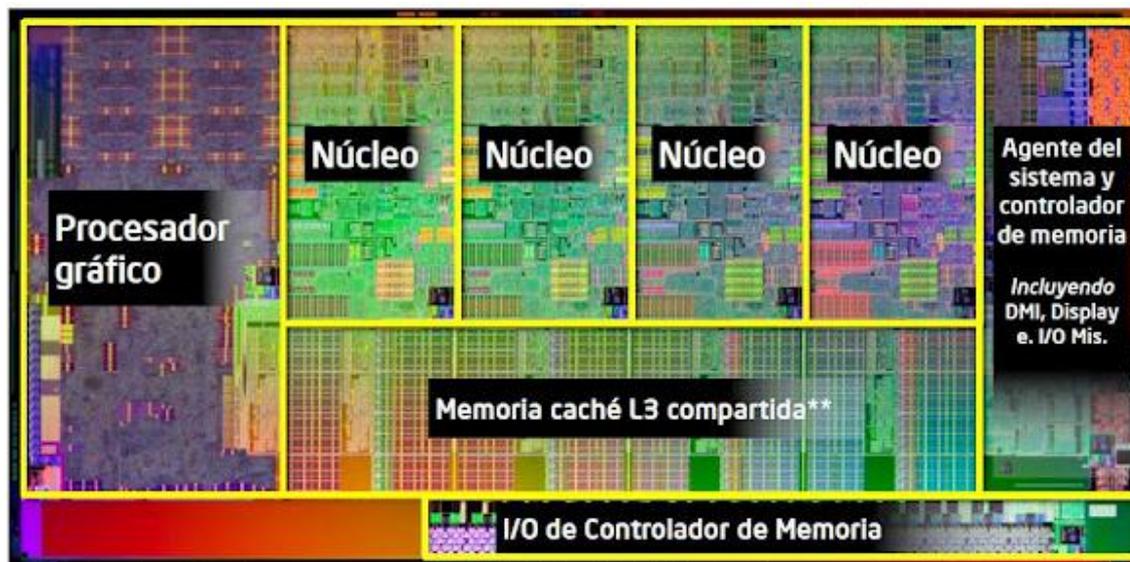


Figura 14. Arquitectura Sandy Bridge ^[22].

3.1.1.1.2.1. CARACTERÍSTICAS DE LA ARQUITECTURA

A continuación se muestran las principales características de la nueva arquitectura ^[23].

Integración de GPU: Intel sigue con su empeño de integrar el chip CPU y el chip gráfico GPU dentro de un mismo compuesto, y ahora con la arquitectura Sandy Bridge parece que lo ha conseguido de una forma más eficiente que en Nehalem. En Nehalem el encapsulado estaba compuesto de un procesador con VGA integrado, sin embargo no dejaba de ser dos componentes separados en el mismo chip, el procesador se fabricaba a 32nm y el chip gráfico se fabricaba a 45nm que se interconectaban entre ellos de forma directa. Por otro lado, Sandy Bridge sale al mercado a principios de 2011 con la idea de “todo junto”, en la nueva arquitectura nos encontramos que el clip gráfico y el procesador son totalmente indivisibles y están diseñados en un mismo chip fabricado con un proceso de 32nm. Una clara ventaja de unificar todo en un mismo chip es la alta velocidad de interconexión que se consigue, además también se consigue una reducción de consumo y un tamaño muy reducido que implica una disminución del costo de fabricación.

Mejoras en los Cores: En el apartado de núcleos, Sandy Bridge no supone una verdadera revolución comparado con Nehalem, ya que los núcleos no han sufrido aparentemente modificaciones revolucionarias. Sin embargo se han realizado mejoras en varios aspectos de la arquitectura para hacerla más eficientes. Es decir, el rendimiento máximo teórico de una arquitectura pasa por que las unidades de ejecución se aprovechen al máximo y la pipeline esté ocupada en todo momento. Para llenarla, se necesita predecir en la línea de ejecución de un programa qué instrucciones se van a ejecutar en el futuro. En la práctica, no siempre hay instrucciones ejecutándose o la pipeline se llena con algunas instrucciones que no hacen falta procesar.

La solución que se propone en Sandy Bridge es la de mejorar la unidad de predicción (Branch Predictor), así como la gestión de las instrucciones descodificadas o micro ops (μops).

El mejorar la unidad de predicción implica mejorar la arquitectura ya que supone aumentar el rendimiento del procesador sin modificar la frecuencia. Si se sabe qué instrucciones vendrán después, la CPU puede estar trabajando al máximo de su rendimiento.

En cuanto a la gestión de la μ ops (Instrucciones descodificadas para que sean procesables por la electrónica y la lógica en el procesador), Sandy Bridge trabaja con una caché de μ ops, de modo que las instrucciones descodificadas estarán accesibles sin volver a descodificarlas. Si una μ op se localiza en la caché, no hace falta usar toda la lógica dedicada a la descodificación, que además estaría apagada para ahorrar más energía.

Hyper-Threading y Turbo Boost: Los nuevos microprocesadores siguen soportando las tecnologías que se incorporaron en Nehalem. Se han producido nuevas mejoras para el HyperThreading y para la tecnología Turbo Boost, aunque como pasaba en la anterior arquitectura, algunas características están capadas o desactivadas para los diferentes modelos que se encuentran en el mercado. El HyperThreading puede llegar a ejecutar hasta 4 hilos de forma simultánea, mientras que con Turbo Boost activado se puede llegar hasta los 3.8 GHz de velocidad, también gracias a la tecnología nueva de Turbo Boost la GPU que inicialmente cuenta con frecuencias de 650 MHz a 850 MHz puede alcanzar una velocidad de 1.35GHz.

Memoria caché: La nueva arquitectura está formada por 64KB de caché de L1 por núcleo (32 para datos y otros 32 para instrucciones) y 256 KB de caché L2 por núcleo. Por otro lado, la caché de L3 puede llegar hasta 8MB compartida con un bus en anillo para poder compartirse con el núcleo gráfico.

Anchos de banda: Cuenta con un ancho de banda del bus de anillo de 256 bits por ciclo, este anillo permite que los núcleos accedan a los datos en la caché o en la memoria del sistema cuando se necesitan. La ventaja de esta estructura es su escalabilidad, de manera que se puede hablar con dos, cuatro, cinco, o los cores que sean necesarios. Además, no necesariamente tienen que ser los mismos tipos de núcleos. Todos los procesadores cuentan con un ancho de línea con caché de 64 Bytes, mientras que el controlador de memoria se ha mejorado con un ancho de banda máximo de 25,6 GB/s y que además incluye soporte para DDR3 a 1600 MHz.

Instrucciones AVX: Como bien es conocido las instrucciones SSE han sido un clásico dentro de los procesadores de Intel. En Sandy Bridge se alcanza un ancho de 256 bits para las operaciones de coma flotante usando las extensiones AVX (Advanced Vector Extensions). Estas son de especial interés para el procesamiento de datos en aplicaciones como la encriptación, donde se trabaja con cadenas de bits de 128, 256 o más bits. Si se puede trabajar con 256 bits se evita tener que invertir ciclos de reloj en partir las cadenas para procesarlas en bloques. Así, una operación que tardaría varios ciclos en procesarse se asume que con Sandy Bridge se realizaría en uno.

3.1.1.1.2.2. EFICIENCIA ENERGÉTICA

Intel mantiene la filosofía de que todo aquello dentro del silicio que no esté haciendo algo útil se puede apagar. En la nueva micro arquitectura esta idea se ha llevado al extremo. Tanto los núcleos CPU como la GPU y el agente del sistema pueden modificar sus voltajes y frecuencias dinámicamente, de modo que se aprovecha cualquier oportunidad para aumentar el rendimiento.

Uno de los efectos que se derivan de esta optimización es la mejora de Turbo Boost que ya se incluía en la arquitectura Nehalem y ahora la mejora de esta tecnología se denomina Turbo Boost 2.0.

En Turbo Boost 2.0 la estrategia consiste en aprovechar que el procesador no alcanza los límites máximos de temperatura y limitaciones térmicas hasta que no ha pasado un tiempo desde que se le exige el máximo rendimiento. De esta forma, se le puede forzar durante unos milisegundos hasta que se alcanza el nivel máximo de disipación especificado por su TDP (Thermal Design Power).

En funcionamiento de Turbo Boost 2.0 se detalla en los gráficos de las figuras 15 y 16.

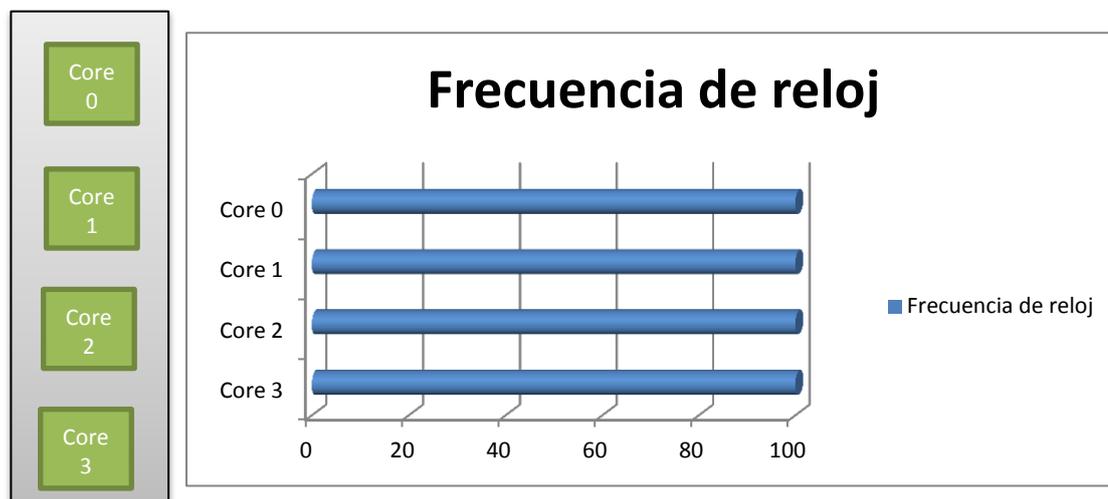


Figura 15. Turbo Boost 2.0 desactivado.

Cuando el procesador funciona por debajo de los límites de temperatura, consumo estimado de corriente, consumo estimado de energía y la carga de trabajo del usuario exige mayor desempeño, la frecuencia del procesador aumentará de forma dinámica hasta alcanzar su límite superior como se puede ver en la figura 16. La tecnología Intel Turbo Boost 2.0 posee varios algoritmos que funcionan en paralelo para administrar la corriente, energía y temperatura, a fin de maximizar el desempeño y la eficiencia energética ^[24].

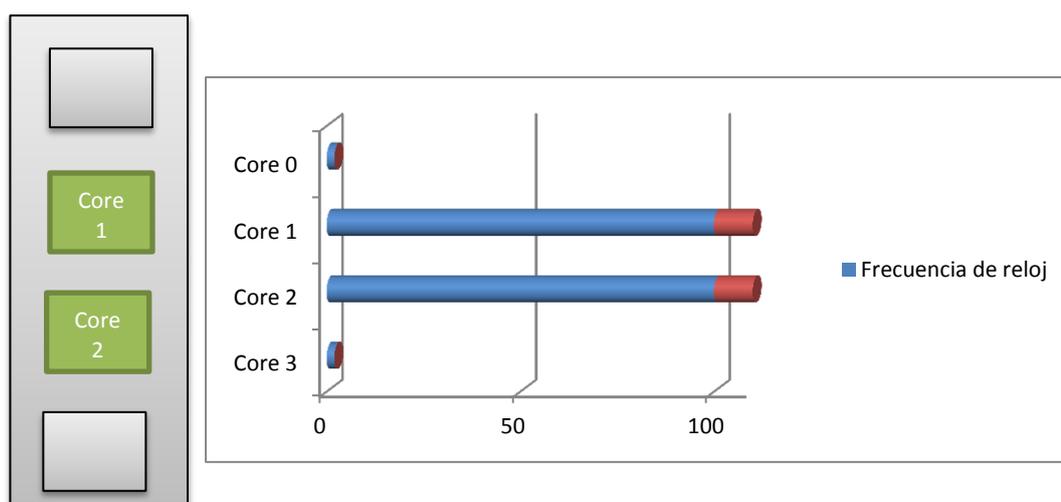


Figura 16. Turbo Boost 2.0 activado.

3.1.1.1.2.3. MERCADO BASADO EN LA ARQUITECTURA

El mercado de Sandy Bridge es similar al de la arquitectura Nehalem. El fabricante Intel proporciona su arquitectura en las tres familias que corresponden a gama baja (i3 de 2 generación), gama media (i5 de segunda generación) y gama alta (i7 de segunda generación).

Para ver el mercado actual de los microprocesadores que se están comercializando se debe acudir al ANEXO III.

3.1.1.2. FABRICANTE AMD

En este apartado de la memoria se tratarán las alternativas comerciales que proporciona el fabricante AMD.

AMD (Advanced Micro Devices) es una compañía americana que desarrolla procesadores de cómputo y productos tecnológicos para el mercado de la informática. AMD es el segundo proveedor de microprocesadores basados en la arquitectura x86. Además dentro de su gama de productos se encuentran placas base, circuitos integrados, sistemas embebidos y procesadores gráficos.

En el mercado actual se pueden encontrar las siguientes micro arquitecturas de procesadores:

- Arquitectura K10.
- Arquitectura Bulldozer.

3.1.1.2.1. ARQUITECTURA K10

La arquitectura K10 es una arquitectura que se empezó a comercializar a finales de 2007 como sucesora de la familia K8 desarrollada por AMD. Los primeros microprocesadores en utilizar esta arquitectura fueron los procesadores denominados Opteron, que eran microprocesadores que se utilizan como servidores de trabajo. La arquitectura K10 también fue utilizada en los microprocesadores Phenom, Athlon 64 y Sempron 64. Los microprocesadores que utilizaban esta micro arquitectura estaban basados en la tecnología de fabricación de 65 nm.

En 2008 apareció una evolución en la arquitectura, que AMD denominó arquitectura Shanghai. Esta arquitectura se fabrica con la tecnología de 45nm y es la que se puede encontrar en el mercado actualmente. Los microprocesadores que incorporan esta arquitectura son designados con el nombre Phenom II xn y Athlon II xn series. AMD tiene pensado dejar de producirla a lo largo de 2012 ^[25] ^[26] .

3.1.1.2.1.1. PHENOM II

El Phenom II es una subfamilia de micro procesadores fabricados con la tecnología de 45 nm la cual sucede al Phenom original. El micro procesador fue lanzado a finales de 2008.

3.1.1.2.1.1.1. CARACTERÍSTICAS DE LA ARQUITECTURA

Los procesadores de tres y cuatro núcleos en la subfamilia Phenom II se diseñaron para una mejor comunicación entre núcleos. Al tener una conexión directa en todos los núcleos repercute en una mejoría de la eficiencia del sistema. El microprocesador es compatible con instrucciones de 32 y 64

bits además tiene el controlador de memoria integrado. A continuación se enumeran las principales ventajas de la arquitectura ^{[27] [28]}.

Caché inteligente equilibrada de AMD

- Caché L3 compartida entre todos los núcleos de 6MB o de 4MB.
- La caché L2 es de 512KB por núcleo.

Acelerador ancho de coma flotante AMD

- La unidad de coma flotante (FPU) es de 128 bits.
- Unidad de coma flotante (ruta de datos internos de 128 bits) de alto rendimiento por núcleo, lo que proporciona unas rutas de datos más largas para cálculos de coma flotante más rápidos.

Tecnología HyperTransport

- Existe un enlace de 16 bits con capacidad de 4000MT/s.
- Hasta 8.0 GB/s de ancho de banda de entrada/salida en HyperTransport, dispone de un modo de trabajo basado en la tecnología HyperTransport Generation 3.0 donde se tendría hasta 16 GB/s.
- Los tiempos de acceso de entrada y salida son mejores gracias al ancho de banda disponible entre el procesador y el sistema llegando a los 37GB/s (bus de HyperTransport + bus de memoria).

Controlador DRAM integrado con tecnología de optimización de memoria de AMD

- Gracias a que se dispone de un controlador de memoria integrado, la arquitectura proporciona una baja latencia y un ancho de banda elevado.
- Admite hasta 17.1 GB/s de ancho de banda de memoria para DDR2 y hasta 21 GB/s de ancho de banda de memoria para DDR3.

AMD Virtualization (AMD-V) con indexación de virtualización rápida

- Con esta nueva tecnología se consigue contribuir a una ejecución más segura y eficaz del software de virtualización al permitir una mejor experiencia de los sistemas virtuales. Para conseguir la mejora final, se han desarrollado una serie de características de silicio diseñadas para aumentar el rendimiento, la fiabilidad y la seguridad de los entornos de virtualización.

3.1.1.2.1.1.2. EFICIENCIA ENERGÉTICA

A continuación se detalla las principales tecnologías que ha desarrollado el fabricante AMD en cuanto a temas de ahorro energético ^{[28] [29] [30]}.

Tecnología Cool'n'Quiet 3.0: Con esta tecnología se pretende que la plataforma sea más silenciosa y esté mejor refrigerada mientras se proporciona a la vez un uso de la energía y un rendimiento eficiente. La velocidad del reloj de la CPU y el voltaje es automáticamente reducido cuando el computador está en bajo uso o en espera, con esto se consigue que los núcleos operen de forma dinámica a una menor frecuencia y consumo, dependiendo del uso y de la carga de trabajo.

Tecnología AMD CoolCore: La tecnología se encarga de reducir el consumo del procesador desactivando las partes del procesador que no se están usando. Por ejemplo, el controlador de memoria puede desactivar la lógica de escritura cuando se está leyendo información de la memoria, lo que contribuye a reducir el consumo de energía del sistema. El sistema CoolCore funciona automáticamente sin necesidad de controladores ni activación de la BIOS, la alimentación se puede activar o desactivar dentro de un solo ciclo de reloj, ahorrando energía sin comprometer el rendimiento.

Dual Dynamic Power Management: Gracias a esta mejora se consigue mejorar la eficiencia de la plataforma proporcionando un rendimiento de memoria según la demanda, mientras permite un consumo energético reducido del sistema. La tecnología se encarga de habilitar funciones de gestión de alimentación más granular para conseguir reducir el consumo energético del procesador. Incluye placas de alimentación independientes para los núcleos y el controlador de memoria, para un consumo energético y un rendimiento óptimo.

3.1.1.2.1.3. MERCADO BASADO EN LA ARQUITECTURA

En la actualidad el fabricante proporciona varios modelos de la subfamilia AMD Phenom II. Se pueden agrupar mediante el número de núcleos: Phenom II X2 que constarían de 2 núcleos, Phenom II X4 que constarían de hasta 4 núcleos operativos y por último están los Phenom II X6 que contendrían 6 núcleos dentro del mismo chip. Para poder ver la lista de microprocesadores que actualmente se encuentran en el mercado se debe consultar el ANEXO IV.

3.1.1.2.1.2. ATHLON II

Athlon II es una subfamilia de micro procesadores de AMD que fue lanzada al mercado en 2009. Es una línea de procesadores basada en la arquitectura K10 que es complementaria a la subfamilia Phenom II comentada anteriormente. Sin embargo una característica destacable a diferencia de la familia de Phenom II, es que no posee caché de L3. En estos procesadores, se ha intentado cubrir esa diferencia, aumentando el nivel de la cache de segundo nivel de 512KB a 1MB por cada núcleo ^[31].

3.1.1.2.1.2.1. CARACTERÍSTICAS DE LA ARQUITECTURA

La fabricación del microprocesador se realiza con un proceso de 45 nm. El microprocesador cuenta con la tecnología AMD Digital Media XPress 2.0 que se encarga de ofrecer compatibilidad con las instrucciones SSE, SSE2, SSE3, SSE4a y MMX ^[32].

A continuación se muestran las características principales de la subfamilia Athlon II.

Procesamiento multinúcleo

AMD se ha encargado de mejorar y optimizar las características de la arquitectura AMD64 con el fin de conseguir una integración perfecta de los múltiples núcleos en el mismo procesador, en el que cada núcleo tiene su propia caché L1 y su caché L2.

Multicaché dedicada de AMD

Cada núcleo tiene su propia caché L2 dedicada como se ha comentado anteriormente, lo que le permite al núcleo un acceso independiente y simultáneo a su caché L2 y elimina la necesidad de que los núcleos tengan que decidir quién accede a la caché. Con esto, se consigue reducir la latencia en los accesos a la caché L2.

Tecnología HyperTransport 3.0

Como se ha explicado en la anterior arquitectura Phenom II, mediante la tecnología de HyperTransport que ha desarrollado AMD se consigue una mejora en el rendimiento. En la arquitectura Athlon se pueden conseguir hasta velocidades de transferencia de 4.4 GT/s (Gigas por segundo) o 2.2 GHz.

Capacidad simultánea para 32 y 64 bits

La tecnología AMD 64 realiza un enfoque totalmente nuevo en los programas de 64 bits que duplica el número de registros en el procesador y permite a los usuarios utilizar aplicaciones de 32 y 64 bits.

3.1.1.2.1.2.2. EFICIENCIA ENERGÉTICA

En cuanto al tema de la gestión de energía se ha conseguido una mejora mediante la tecnología AMD PowerNow! 3.0 en donde se destacan las siguientes características.

Tecnología Cool'n'Quiet 3.0

Esta tecnología también estaba presente en la anterior arquitectura Phenom II. En Athlon II se tiene hasta ocho estados de rendimiento diferentes que ayudan a mejorar la eficiencia energética. Las transiciones de estado simplificadas ayudan a reducir la latencia y la sobrecarga de software en los cambios de estado de rendimiento.

AMD Dynamic Power Management

En la micro arquitectura Athlon II cada núcleo del procesador, el controlador de memoria integrado y el controlador de HyperTransport reciben la alimentación de una placa de voltaje dedicada, proporcionando una mejora de la eficiencia mientras se reduce el consumo energético.

Controlador de memoria integrado de dos canales

El procesador está conectado directamente a la memoria para optimizar el rendimiento, reducir la latencia y elevar la producción.

Control de temperatura multipunto

Esta arquitectura dispone de varios sensores térmicos en la placa base con interfaz digital. La reducción de energía se produce de forma automática, cuando la temperatura supera el límite predefinido. Además existe un interfaz adicional que se encarga de la gestión de temperatura de memoria.

Tecnología AMD CoolCore

La tecnología AMD CoolCore ya se había introducido en la anterior arquitectura Phenom II. En Athlon II también existe un controlador que tiene la capacidad de reducir automáticamente el consumo de energía del procesador desactivando las partes que no se estén usando.

3.1.1.2.1.2.3. MERCADO BASADO EN LA ARQUITECTURA

La subfamilia Athlon II se encuentra en la última fase de producción, ya que los microprocesadores serán sustituidos por la nueva arquitectura Bulldozer. En la tabla 1 se muestra el estado del mercado que existe en la actualidad.

Tabla 1. Familia Athlon II ^[33] .

Procesador	Velocidad de reloj	Memoria Caché L2	TDP Max	Núcleos
Athlon II X3 455	3.3GHz	1536KB	95W	3
Athlon II X4 631	2.6GHz	4096KB	100W	4
Athlon II X3 460	3.4GHz	512KB	95W	3

3.1.1.2.2. ARQUITECTURA BULLDOZER

Bulldozer es la micro arquitectura diseñada por la empresa AMD que sucederá a la micro arquitectura descrita anteriormente AMD K10. El microprocesador está diseñado para disipar entre 10 y 125W. Los núcleos en Bulldozer soportan todos los juegos de instrucciones actualmente implementados en procesadores Intel y además añaden otros sets de instrucciones propuestos por AMD como son XOP y FMA4 ^[34] .

3.1.1.2.2.1. CARACTERÍSTICAS DE LA ARQUITECTURA

La nueva arquitectura mejora la temperatura y aumenta la velocidad del reloj con un proceso de fabricación basado en la tecnología de 32 nm. Se detallan a continuación las principales mejoras que incorpora la arquitectura.

Tecnología AMD Turbo Core

Los procesadores basados en la arquitectura Bulldozer aparecen dentro de la serie de procesadores denominados AMD FX y también dentro de la serie denominada Opteron que está por aparecer en el mercado. Los AMD FX Bulldozer vienen equipados con la tecnología AMD Turbo CORE, dicha tecnología permite aumentar la frecuencia de ejecución de las instrucciones bajo ciertas condiciones. El funcionamiento es similar al Turbo Boost tratado anteriormente en las arquitecturas de Intel.

Nuevas instrucciones que mejoran el rendimiento

- AVX: Las extensiones vectoriales avanzadas se han diseñado para aumentar el paralelismo que se utiliza en aplicaciones científicas y de 3D las cuales requieren cálculos complejos.
- FMA4 y XOP: Este juego de instrucciones tiene la capacidad de mejorar el rendimiento de numerosas funciones vectoriales (coma flotante y enteros). Además mediante la nueva mejora basada en el acelerador de AMD para operaciones de coma flotante, se dispone de chips capaces de realizar operaciones duales de 128 bits que consiguen realizar instrucciones AVX de 256 bits o de operar de forma separada por cada núcleo.

Caché AMD Balanced Smart

En la arquitectura Bulldozer la caché L3 es compartida entre los núcleos y tiene hasta 8 MB de capacidad, en donde:

- Incluye nuevas mejoras en la planificación y pre-lectura de datos.
- Se produce un aumento en el tamaño de las colas de datos.
- Coherencia para 8 núcleos.

Tecnología HyperTransport

- La nueva arquitectura dispone de un enlace de 16 bits que dé hasta 5600 MT/s.
- Existe un ancho de banda de Entrada/Salida HyperTransport de hasta 8.0 GB/s, si se trabaja en el modo HyperTransport Generation 3.0 se consigue una velocidad de hasta 16 GB/s.
- El ancho de banda formado por el bus HyperTransport + bus de memoria es como máximo 37 GB/s.

Controlador DRAM integrado con la tecnología AMD Memory Optimizer

- Se basa en un controlador de memoria integrada que dispone de un ancho de banda alto y baja latencia.
- Es compatible con DDR3-1866.
- Compatible con las nuevas memorias de bajo voltaje de 1.35V y 1.2V
- Hasta 29.9 GB/s de ancho de banda de memoria para DDR3.
- Incluye mejoras en la pre-lectura de datos ^[35].

3.1.1.2.2.2. EFICIENCIA ENERGÉTICA

En cuanto al rendimiento de la energía, los procesadores AMD FX basados en la arquitectura Bulldozer no presentan una verdadera revolución como sucedía en Intel de Nehalem a Sandy Bridge. En este apartado se comentará las principales características que ADM utiliza en su arquitectura eficiente.

Reducción de voltaje: Se ha reducido el voltaje de forma individual en cada núcleo.

Estado de alimentación C6: Se encarga de reducir el consumo energético del procesador cuando está en un estado de inactividad, produciéndose una reducción de hasta un 46%. El estado C6 permite un voltaje aún menor en todos los núcleos para una operación más silenciosa y a la vez reducir el gasto energético.

Separación de controles de encendido: El control de encendido está separado del controlador de la memoria.

Uso de la tecnología Cool'n'Quiet: La tecnología Cool'n'Quiet como se ha visto anteriormente funciona de forma automática reduciendo el consumo cuando el computador está en bajo uso.

Uso de la tecnología CoolCore: En cuando a la tecnología CoolCore como sucedía en la arquitectura Phenom II la alimentación puede encenderse o apagarse dentro de un único ciclo de reloj ahorrando energía sin impactar en el rendimiento ^[35].

3.1.1.2.2.3. MERCADO BASADO EN LA ARQUITECTURA

En la tabla 2 se muestran los productos AMD FX Bulldozer que ofrece el fabricante.

Tabla 2. Familia AMD FX Bulldozer ^[36].

Procesador	Velocidad de reloj	Memoria Caché L2	Memoria Caché L3	TDP Max	Núcleos
FX 4-Core Black Edition FX-4100	3.6 GHz	1024KB	8192KB	95W	4
FX8-Core Black Edition FX-8120	3.1GHz	1024KB	8192KB	125W	8
FX6-Core Black Edition FX-6100	3.3GHz	1024KB	8192KB	95W	6
FX8-Core Black Edition FX-8150	3.6GHz	1024KB	8192KB	125W	8

3.1.2. ARM

ARM es una arquitectura perteneciente a la familia RISC (reduced instruction set computer) de 32 bits desarrollada por ARM Holdings. ARM es el principal proveedor de microprocesadores de 32 bits integrados, ofrece una gama de procesadores basados en una arquitectura común que proporciona un alto rendimiento y además es el líder en la industria de los microprocesadores en cuanto a temas de eficiencia energética. Originalmente fue concebida para su uso en computadores personales y los primeros productos basados en ARM eran los Acorn Achimedes, lanzados en 1987. En la actualidad están disponibles dos grandes grupos de procesadores, los procesadores clásicos compuestos por ARM7, ARM9 y ARM11 y los procesadores denominados Cortex. En el siguiente apartado se describirán los diferentes procesadores centrándose en los más actuales ^{[37][38]}.

3.1.2.1. PROCESADORES CLÁSICOS

Los procesadores ARM clásicos incluyen las familias de procesadores ARM11, ARM9 y ARM7 como se ha visto anteriormente. Estos procesadores están proporcionando soluciones rentables para cierto abanico de aplicaciones.

ARM7

Fue introducido en 1994. La familia ARM7 ha tenido un enorme éxito y ha ayudado a establecer la arquitectura ARM en el mundo digital. Cuenta con más de 10 millones de procesadores a lo largo de estos años, que han impulsado una amplia variedad de aplicaciones con un consumo de energía aceptable. Todavía la familia de procesadores ARM7 se sigue utilizando en simples dispositivos de 32 bits. En la actualidad el fabricante propone actualizar la arquitectura hacia ARM Cortex-M0 y ARM Cortex-M3 los cuales ofrecen un mayor rendimiento además de mejoras técnicas con respecto a ARM7 ^[39].

ARM9

Los procesadores de la familia ARM9 ofrecen soluciones para micro controladores, DSP, y aplicaciones en JAVA, ofreciendo simplicidad, un diseño flexible y reducción de consumo.

Según los datos del fabricante se han producido más de cinco mil millones de procesadores ARM9. Además ARM9 es la familia de procesadores ARM más popular con más de 250 licencias que pertenecen a distintos titulares. En la actualidad todavía se siguen utilizando los microprocesadores en diversos productos y aplicaciones. La popularidad de la familia ARM9 es debido a su pequeño tamaño, estabilidad y facilidad de incorporación en diseños complejos. Dentro de la familia se encuentran tres procesadores (ARM926EJ-S, ARM946E-S, ARM968-S).

La familia ARM9 aparece en varios productos que se enumeran a continuación: de cara al consumidor director (PDA, Set top box, juguetes electrónicos, cámaras electrónicas...), en productos relacionados con las redes (Wireless LAN, 802.11, Bluetooth, Firewire, SCSI, 2.5G/3G, etc.) , en productos relacionados con el sector del automóvil (navegador, ABS, ...), y en sistemas embebidos como son los controladores USB, escáner médicos, etc.^[40]

Características técnicas

- Basado en la arquitectura ARMv5TE.
- Sistema eficiente basado en 5 etapas pipeline que mejoran el rendimiento del sistema: Extraer, decodificar, ejecutar, Memoria, rescritura.
- Soporta el conjunto de instrucciones ARM y Thumb.
- Basado en la arquitectura Harvard, que separa la interfaz de memoria de instrucciones y de datos, con ello se consigue un aumento del ancho de banda de la memoria disponible, también se consigue un acceso simultáneo a la memoria de instrucciones y de datos, lo que conlleva a una mejora en el rendimiento.
- Registros de 31x32 bits.
- Alu de 32 bits.

Controlador de memoria

- Las operaciones de memoria son controladas por la MMU o MPU.
- La MMU proporciona: Soporte para la memoria virtual y FCSE (Fast Context Switching Extensions).
- La MPU se encarga de la activación de la protección de la memoria y de delimitar el espacio entre aplicaciones.
- En cuanto a los buffers: Se encargan de desacoplar el procesador interno de la memoria externa. Además se puede almacenar hasta 16 palabras en 4 direcciones independientes.

Diseño de caché flexible

- Arquitectura de caché Harvard.
- Tamaño variable desde 4KB hasta los 128KB.
- Las cachés de instrucciones y de datos pueden tener tamaños diferentes.
- Longitud de la línea compuesta por 8 palabras.
- No se bloquea.
- Direcciones virtuales.

Diseño flexible TCM (Tightly-Coupled Memory)

- Sistema de organización basado en Harvard
- Los tamaños pueden ser de 0KB o 4 KB hasta 1MB.
- Pueden tener tamaños independientes.
- Puede ser RAM o ROM.
- Se permiten los estados de espera.
- Direcciones físicas ^[40] .

ARM11

Los procesadores de la familia ARM11 pueden encontrarse en la actualidad en múltiples teléfonos inteligentes. También se pueden encontrar en aplicaciones de consumo, en el hogar y en sistemas embebidos. Estos chips trabajan en un rango de frecuencias aproximado de 350 MHz hasta los 1 GHz con un consumo de potencia mínimo. La tecnología de fabricación se basa en 45 y 65nm. El software de ARM11 es compatible con todas las generaciones anteriores de procesadores ARM. En ARM11 aparecen 32bits para SIMD que se usan para el procesamiento de medios y cachés para mejorar el rendimiento del operativo.

A continuación se describen las características que proporciona el fabricante del chip ARM11.

- Potente conjunto de instrucciones procedente de la arquitectura ARMv6
- Conjunto de instrucciones Thumb que reduce el ancho de banda de memoria y los requisitos de tamaño hasta en un 35%.
- Incluye ARM Jazelle, que es una tecnología enfocada a la eficiencia de ejecución de Java en sistemas embebidos.
- Extensiones ARM DSP.
- Las extensiones de procesamiento SIMD mejoran hasta dos veces el rendimiento del procesamiento de video.
- Incluye tecnología para la seguridad denominada ARM TrustZone.
- Se mejora la eficiencia energética y reducción de código gracias a la tecnología Thumb-2.
- Bajo consumo energético: 0.21 mW/MHz incluyendo los controladores caché.
- Mejoras en el procesador de números enteros: 8 etapas de pipeline ofrecen una alta frecuencia de reloj. Además incluye una unidad de predicción y una pila de retorno.
- Mejoras en el sistema de diseño de memoria: Soporta 4-64K tamaños de caché. Opcionalmente se pueden incluir memorias que están acopladas con DMA para mejorar las tareas multimedia ^[41] .

3.1.2.2. PROCESADORES CORTEX

En esta sección se expone la información que proporciona el fabricante acerca de los procesadores Cortex -A series, Cortex-R series, y Cortex-M series.

Cortex-A series

Los procesadores ARM Cortex-A son una familia de procesadores que proporcionan un conjunto de soluciones para dispositivos que van desde portátiles de bajo coste, teléfonos inteligentes, plataformas móviles, sistemas de televisión digital, descodificadores, creación de redes, impresoras, hasta otras

posibles soluciones en productos. Existen varios modelos en el mercado, que serían, Cortex-A15, Cortex - A9, Cortex – A7 y Cortex –A5, todos comparten la misma arquitectura y por lo tanto existe una compatibilidad en las aplicaciones. Estos procesadores soportan los conjuntos de instrucciones Thumb y Thumb 2 ^[42] .

Los procesadores de la familia Cortex-A15 tendrán un proceso de fabricación de 20 nm y saldrán a finales del 2012 ^[43] . En la tabla 3 se muestra información de cada microprocesador.

Tabla 3. Microprocesadores Cortex-A ^[37] .

Núcleo	Versión de la arquitectura	Características	Caché (Instrucciones/Datos)/MMU	MIPS efectivos / MHz
Cortex-A5	ARMv7-A	VFP, NEON, Jazelle RCT y DBX, Thumb-2, Segmentación de 8 fase, 1-4 núcleos SMP.	Variable (L1), MMU+TrustZone.	Más de 1500 (1.5 DMIPS/MHz).
Cortex-A8	ARMv7-A	VFP, NEON, Jazelle RCT, Thumb-2, Segmentación Superescalar de 13 etapas.	Variable (L1+L2), MMU+TrustZone.	Más de 2000 (2.0 DMIPS/MHz de reloj desde 600 MHz hasta más de 1 GHz).
Cortex-A9	ARMv7-A	Perfil de aplicaciones, (VFP), (NEON), Jazelle RCT y DBX, Thumb-2.	MMU+TrustZone	2.5 DMIPS/MHz

Cortex-R series

La familia ARM Cortex-R ha sido pensada para ofrecer una solución en tiempo real en los procesadores incrustados con exigentes restricciones de tiempo de respuesta. Las aplicaciones pensadas para utilizar un microprocesador de la serie R son:

- Aplicaciones de procesamiento de telefonía móvil: Los teléfonos inteligentes y en módems.
- Aplicaciones de uso en sistemas empresariales: Discos duros, las interconexiones de una empresa, impresoras, etc.
- Aplicaciones que se desarrollan para la electrónica de consumo: Set top box, televisión digital, reproductores multimedia, cámaras, etc.
- Aplicaciones que se utilizan en sistemas integrados, sistemas industriales y mercado de la automoción ^[44] .

En la tabla 4 se muestra información del microprocesador R4.

Tabla 4. Microprocesadores Cortex-R ^[37] .

Núcleo	Versión de la arquitectura	Características	Caché (Instrucciones/Datos)/MMU	MIPS efectivos / MHz
Cortex-R4	ARMv7-R	Perfil Embebido, Thumb-2.(FPU).	Variable caché, MPU opcional.	600 DMIPS/475MHz.

Cortex-M series

La familia Cortex –M es una familia de procesadores pensada para la eficiencia energética, son procesadores fáciles de utilizar, diseñados para ayudar a los desarrolladores y satisfacer las necesidades en productos integrados. Cubren un gran abanico de demanda, la micro arquitectura fue pensada para un menor costo del producto final, una mayor conectividad, mejor reutilización del código y sobre todo una mayor eficiencia energética ^[45] .

En la tabla 4 se muestra información de cada microprocesador.

Tabla 5. Microprocesadores Cortex-M ^[37] .

Núcleo	Versión de la arquitectura	Características	Caché (Instrucciones/Datos)/MMU	MIPS efectivos / MHz
Cortex-M0	ARMV6-M	Perfil microcontrolador, Thumb-2. (Instrucciones 16-bit Thumb y BL, MRS, MSR, , ISB, DSB y DMB).	No tiene.	0.9 DMIPS/MHz
Cortex-M1	ARMv6-M	FPGA targeted, Perfil microcontrolador, Thumb-2, (instrucciones 16 bit Thumb & BL, MRS, MSR, ISB, DSB, y DMB).	No tiene.	Más de 136 DMIPS / 170 MHz.
Cortex-M3	ARMv7-M	Perfil microcontrolador, Thumb-2 únicamente. Reparto de instrucciones por Hardware.	Sin caché, MPU opcional.	125 DMIPS/100MHz
Cortex-M4	ARMV7-ME	Perfil microcontrolador, Thumb y Thumb-2, FPU. MAC, SIMD e instrucciones divididas.	MPU opcional.	1.25 DMIPS/MHz

3.1.3. SPARC

SPARC (Scalable Processor ARChitecture) es una arquitectura RISC big-endian. Fue originalmente diseñada por Sun Microsystems y dirigido por el ingeniero Kaa en 1985, se basa en los diseños RISC I y II de la Universidad de California en Berkeley que fueron definidos entre los años 1980 y 1982.

La empresa Sun Microsystems diseñó esta arquitectura y la licenció a otros fabricantes como Texas Instruments, Cypress Semiconductor, Fujitsu, LSI Logic entre otros ^[46] .

3.1.3.1. CARACTERÍSTICAS GENERALES

Hay que destacar que SPARC es la primera arquitectura RISC abierta, por tanto las especificaciones de diseño están publicadas, así otros fabricantes de microprocesadores pueden desarrollar su propio diseño.

Una de las ideas innovadoras de la arquitectura es la ventana de registros que permite hacer fácilmente compiladores de alto rendimiento y una significativa reducción de memoria en las instrucciones load/store en relación con otras arquitecturas RISC. Las ventajas se aprecian sobre todo en programas grandes.

En cuanto a la CPU de SPARC está compuesta de una unidad entera, UI (Integer Unit) que procesa la ejecución básica y una FPU (Floating-Point Unit) que ejecuta las operaciones y cálculos de reales. La IU y la FPU pueden o no estar integradas en el mismo chip. También, aunque no es una parte formal de la arquitectura, las computadoras basadas en SPARC de Sun Microsystems tienen una unidad de manejo de memoria (MMU) y una gran caché de direcciones virtuales (para instrucciones y para datos) que suelen utilizar un bus de datos de direcciones de 32 bits.

Características

- Uso de ventanas de registros.
- 32 registros de enteros de 32 bits.
- 16 registros de punto flotante de 64 bits (para el caso de la doble precisión) que se pueden utilizar como 32 registros de 32 bits (para precisión simple).
- Modos de direccionamiento:
 - Inmediato, (consta de 13 bits).
 - Directo, (offset de 13 bits).
 - Indirecto, (registro + offset de 13 bits o registro + registro).
- Utiliza instrucciones retardadas (saltos, load y store).
- Manejo de memoria:
 - Espacio virtual de 4 Gigabytes.
 - Unidad de manejo de memoria (MMU) que trabaja con páginas de tamaño configurable.

3.1.3.2. SPARC VIII FX

En este bloque se detallará el funcionamiento de la revisión SPARC64 VIII fx. Este procesador es el utilizado en el supercomputador “K Supercomputer” que consiguió aparecer en junio de 2011 en el top 10 de los más eficientes del mundo (lista green500) ^[47].

SPARC64 VIII fx, cuyo nombre en clave es Venus, es una versión de ocho núcleos basada en SPARC64 VII. La nueva versión incluye un controlador de memoria y 760 millones de transistores. El procesador es capaz de operar hasta 128 GFLOPS y está fabricado con la tecnología de 45 nm por Fujitsu ^[48].

Las especificaciones que se han dado a conocer de la revisión son las siguientes:

- **Caché:** Posee una cache L1 de 32KiB para datos y 32 KiB para instrucciones. En cuanto a la cache L2 tiene una capacidad de 5 MiB.

- **Registros:** Tiene un registro de 192 para números enteros y 256 para operaciones de FP (float point). Incorpora un nuevo juego de instrucciones denominado FMA (Floating-point Multiply and ADD) en el que es necesario instrucciones de 32 bits compuestas por un campo de 8 bits x 4 (3 de lectura + 1 de escritura).
- **Tamaño de página:** Puede paginar los siguientes tamaños 8KiB, 64KiB, 512KiB, 4MiB, 32MiB, 256MiB, 2GiB.
- **SIMD:** La unidad SIMD FP es capaz de ejecutar hasta dos hilos paralelos, en donde existe un registro de doble precisión que puede ser utilizado para el cálculo de un solo hilo. Los registros de operación pueden estar definidos por valores simples o dobles.
- **Potencia:** Consumo de 58W (TYP, 30°C), también dispone de refrigeración líquida para el control de temperaturas.

3.1.4. MIPS

El procesador denominado MIPS (Microprocessor without Interlocked Pipeline Stages) pertenece a la familia de microprocesadores RISC como sucedía con ARM o SPARC. Este procesador se puede encontrar en automóviles, sistemas de redes como pueden ser routers, cámaras digitales, sistemas de reproducción y video, GPS, microcontroladores, impresoras, sistemas de VoIP y videoconsolas.

Originalmente MIPS fue una arquitectura de 32 bits, pero en la actualidad se pueden encontrar versiones de 64 bits. Existen diferentes revisiones denominadas MIPS I, MIPS II, MIPS III, MIPS IV MIPS V y la última MIPS 32/64. Esta última revisión dispone de registros de control, y nuevas extensiones que se detallan a continuación ^[49]:

- **MIPS-3D:** Son un conjunto de instrucciones SIMD enfocadas a tareas 3D comunes.
- **MDMX:** Conjunto de instrucciones más extenso para SIMD que utilizan los registros de coma flotante de 64 bits.
- **MIPS16:** Se utiliza para comprimir los flujos de instrucciones para conseguir reducir el tamaño de los programas, es similar a la tecnología Thumb de ARM tratada en esta memoria anteriormente.
- **MIPS MT:** Esta extensión sirve para añadir funciones multi-hilo.

3.1.4.1. CARACTERÍSTICAS MIPS64

MIPS64 fue la primera arquitectura del mundo de 64 bits introducida en 1991. La arquitectura actualmente cubre diferentes segmentos del mercado como se ha comentado en el párrafo anterior. En este bloque se enumerarán las especificaciones que proporciona el fabricante ^{[50] [51]}.

- El microprocesador permite decodificar instrucciones de 32 bits.
- Dispone de 32 registros de propósito general, con un tamaño de 64 bits que pueden ser utilizados para operaciones de coma flotante.
- Dispone de hasta 64 bits de espacio de direcciones virtuales, puede tener hasta 59 bits de espacio de direcciones físicas.
- Modelo de direccionamiento simplificado.
- Soporte para 8 bits, 16 bits, 32 bits y variables de 64 bits.
- Software para manejo de la tabla de paginado y también para operaciones de datos de la pila.
- Soporte para multiplicación y división de enteros.

- Puede soportar precisión doble para operaciones de coma flotante.
- Compatible con los sistemas big-endian y Little-endian.
- Compatible con MIPS I, MIPS II, MIPS IV y MIPS V
- MMU opcional con:
 - Mecanismos de traducción de direcciones como TLB o BAT.
 - Tamaño de página programable.
 - Software de gestión para la tabla de paginado.
- Se puede incluir una caché opcional.

3.1.5. POWERPC

Es una arquitectura de computadores tipo RISC que fue desarrollada por IBM, Motorola y Apple (AIM). Los microprocesadores de la familia PowerPC (también denominada PPC) son producidos por IBM y Freescale Semiconductor que fue la división de semiconductores y microprocesadores de Motorola. El principal mercado de estos microprocesadores eran los computadores de Macintosh de Apple computer hasta el año 2006 y también se usaban en varios modelos de IBM ^[52].

En el siguiente bloque se expondrá la información acerca del microprocesador PPC 970, este procesador es utilizado en varios centros de supercomputación en España como son Marenstrum (BSC-CNS), Magerit (UPM), ITC Atlante (Gobierno de canarias), La Palma, Picasso, Altamira, Caesaraugusta, Tirant ^[53].

3.1.5.1. POWERPC 970

El micro procesador PPC 970 (también se denomina PowerPC G5, haciendo referencia a la 5ª generación de procesadores PowerPC) fue diseñado por IBM en el año 2002, es un procesador de alto rendimiento basado en la arquitectura RISC de 64 bits. La fabricación del microprocesador se basa en la tecnología de 130 nm mientras que sus revisiones 970FX y 970 MP utilizan la tecnología de 90 nm. Estos microprocesadores pueden ejecutar instrucciones de 32 bits de forma nativa y contienen más de 58 millones de transistores ^{[54][55]}. En la tabla 6 se muestran las especificaciones.

Tabla 6. PowerPC Serie 970.

	970	970FX	970MP
Cores	1	1	2
Velocidad del reloj	1.6 - 2.0 GHz	1.4 - 2.0 GHz	1.4 – 2.5 GHz
Caché L1	64KB para instrucciones y 32KB para datos	64KB para instrucciones y 32KB para datos	64KB para instrucciones y 32KB para datos
Caché L2	512KB	512KB	1MB por cada núcleo
Año	2002	2004	2005

3.2. Aceleradores hardware

En los últimos años, ha nacido un nuevo concepto informático que se denomina GPGPU, este concepto no es más que utilizar la tecnología GPU para un propósito más general. Originalmente las GPUs eran arquitecturas pensadas para jugar y para conseguir gráficos 3D interactivos. El punto de partida ha sido 2003 cuando se empezaron a popularizar. Se crearon los primeros códigos acelerados para GPU donde se observó su gran capacidad de rendimiento para el coste de la tecnología. Por el contrario, era muy complicado implementar los algoritmos y se requería una gran cantidad de conocimiento, sobre todo había que conocer a la perfección el sistema de renderización del chip gráfico (se requería un gran tiempo de desarrollo y un perfil de usuario especializado). En 2004 emerge con fuerza el concepto de GPGPU donde se mostraban los primeros esfuerzos académicos para intentar construir algoritmos de propósito general adaptados en tarjetas GPU, como pueden ser simulaciones físicas, procesamiento de señal, biología computacional, etc. En 2005 se demuestra que gracias a las GPU el tiempo de cómputo es menor que en las CPU actuales. A partir de ese momento no se discute sobre el rendimiento ni la polivalencia del hardware gráfico, pero todavía a la GPU le quedaban aun dos asignaturas pendientes: la precisión de los resultados ya que no existían variables “floats” y la facilidad de programación porque todavía se necesitaba mucho tiempo de desarrollo. Ambos aspectos se resuelven con la llegada de CUDA en la serie 8 del fabricante de tarjetas gráficas Nvidia (Noviembre de 2006). Desde entonces y hasta la actualidad se están desarrollando múltiples aplicaciones de cálculo intensivo para GPU ^[56].

En este bloque se recogerá la información que proporcionan los fabricantes de las diferentes GPU que están disponibles en el mercado.

3.2.1. FABRICANTE NVIDIA

En este apartado se explicará en qué consiste la tecnología “Fermi”, que es el modelo de arquitectura que incluyen los productos del fabricante Nvidia en la actualidad.

3.2.1.1. TECNOLOGÍA FERMI

La arquitectura Fermi (GF100) ha sido lanzada a principios de 2010. Es la primera arquitectura que da soporte nativo para DirectX 11 por parte del fabricante Nvidia. En este apartado se expondrán los puntos más importantes por los cuales Nvidia ha conseguido dar un salto tanto cualitativo como cuantitativo respecto a la anterior arquitectura. En cuanto al aspecto cuantitativo, se refiere al incremento de los componentes que se han incluido en el nuevo chip, y en cuando al aspecto cualitativo, se refiere a que se han agregado nuevas tecnologías como el soporte DirectX 11, mejoras en el tratamiento geométrico, también en el apartado visual relacionado con la calidad de imagen, y el rendimiento en computo de propósito general también ha experimentado un gran avance que es el tema que se afronta en esta tesis ^[57].

3.2.1.1.1. ESPECIFICACIONES GENERALES

Para tener un punto de comparación con las arquitecturas anteriores de Nvidia se puede observar la tabla 7, la cual muestra las principales diferencias entre la arquitectura Fermi y las arquitecturas anteriores.

Tabla 7. Comparación modelos Nvidia.

	NVIDIA G92	NVIDIA GT200	NVIDIA FERMI (GF100)
Unidades Shader	128	240	512
ROPs	16	32	48
Unidades de textura	64	80	64
DirectX	10	10	11
Transistores	754M	1400M	3200M
Tipo de memoria	GDDR3	GDDR3	GDDR5
Ancho de banda del bus	256 bit	512 bit	384 bit
Fabricación	65nm	55nm	40nm

Se observa de la tabla anterior el número de procesadores Shader o también denominados “CUDA Cores” por Nvidia ha incrementado respecto a las arquitecturas anteriores, esta mejora tiene una gran aceptación en el software de propósito general.

Otra nueva evolución es el aumento de transistores respecto a las anteriores arquitecturas. EL chip consta de unos 3200 millones de transistores. Además se puede ver en la tabla que también las unidades ROPs se han incrementado a 48. Este aumento tiene que ver con los cambios que se han realizado en el motor de operaciones geométricas que se utilizan para trabajar en el desarrollo de gráficos.

Con la arquitectura Fermi, Nvidia ha dado el paso hacia el tipo de memoria GDDR5, las generaciones anteriores se habían mantenido con chips GDDR3. Un aspecto que puede llamar la atención de la tabla anterior es que Nvidia utiliza un sistema de interfaz de memoria de 384 bits en lugar de los 512 de la generación anterior. Este suceso se da porque Fermi incorpora sólo seis controladores de memoria de 64 bits, en lugar de los 8 del GT200. Por lo que $6 \times 64 = 384$ bits, la reducción de 512 a 384 queda compensada con el uso de las altas frecuencias de la memoria GDDR5.

3.2.1.1.2. ARQUITECTURA FERMI

La arquitectura Fermi se ha organizado de la siguiente manera: en primer lugar está el GF100 que se compone como ya se ha mencionado anteriormente de 512 CUDA Cores, cada uno de estos 512 procesadores está organizado en 4 bloques denominados GPC (Graphics Processing Cluster), cada GPC está formado por 4 bloques llamados Streaming Multiprocessors (SM), que cada SM contiene 32 Cores. En la figura 17 se muestra un GPC.

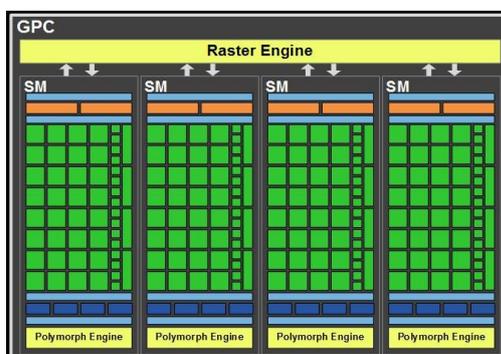


Figura 17. GPC Fermi [57].

Se puede demostrar que los 512 Cuda Cores salen de: 32 Cores * 4 Streaming Multiprocessors * 4 Graphics Processing Clusters.

Streaming Multiprocessors (SM): Para comprender el funcionamiento de la arquitectura se hará un aumento a estos bloques o matrices los cuales forman los GPC. Cada SM está formado por 32 Cores como se ha comentado anteriormente. El resultado de ampliar la figura 17 se puede observar en la figura 18.



Figura 18. Composición SM.

De la figura 18 se observa que cada uno de los 32 Cuda Core posee una propia unidad de cálculos de punto flotante (FP Unit) y una unidad de cálculo de enteros (INT Unit), estas unidades son aprovechadas por los algoritmos que se utilizan para cómputo de propósito general.

Para diseñar la arquitectura, Nvidia dividió el chip en 4 grandes grupos denominados GPC, en lugar de uno solo como sucedía en la generación anterior, metafóricamente hablando como cada GPC está formado por 4 SM la arquitectura Fermi se podría denominar GPU Quad-Core.

3.2.1.1.3. ARQUITECTURA DE LA CACHÉ

Una de las principales características de las tarjetas, es el nivel de paralelismo (capacidad de ejecutar múltiples instrucciones al mismo tiempo) además sus capacidades gráficas, requieren una memoria caché que ayuda a administrar el tráfico y las instrucciones de los datos. En la tabla 8, se muestra un resumen de la caché en Fermi comparándola con la anterior arquitectura.

Tabla 8. Arquitectura caché.

	GT200	Fermi (GF100)
L1 cache de textura	12KB	12KB
L1 caché	-	16 o 48KB
Memoria Shared	16KB	16 o 48KB
L2 caché	256 KB (sólo lectura)	768 KB (lectura y escritura)

En la arquitectura Fermi se dispone de una memoria caché de tamaño 16/48KB compartida. Para comprender mejor este dato, se debe saber que cada SM tiene 64KB de memoria caché compartida y programable, la cual puede ser configurada de dos formas. La primera forma de configuración es dejar 48 KB como memoria caché compartida y 16 KB restantes como caché L1. La segunda forma de configuración es dejar 16 KB como memoria caché compartida y los otros 48KB restantes como caché L1.

Por otro lado se encuentra la caché L2 la cual se incrementa desde los 256KB a los 768KB, pero a diferencia de la anterior arquitectura GT200 donde la caché sólo permitía la lectura, en Fermi (GF100) la caché opera en lectura/escritura. Esto permite mejorar el rendimiento para programación de texturas y de software de cómputo general. Con este sistema lo que se consigue es, por ejemplo, aliviar los cuellos de botella de la memoria, ya que en ciertos cálculos en vez de mover datos a la memoria se pueden dejar en la caché y así se evitan las múltiples operaciones de lectura/escritura hacia la memoria (DRAM).

3.2.1.1.4. MERCADO BASADO EN LA ARQUITECTURA

Los productos de Nvidia que se dedican a entornos de trabajo HPC y actualmente están en el mercado son la serie Geforce y la serie Tesla.

3.2.1.1.4.1. TESLA

En la actualidad se está comercializando la Serie GPU NVIDIA Tesla 20, este chip está basado en la arquitectura de cálculo FERMI. Fermi como ya se había mencionado anteriormente está pensada para optimizar y acelerar las aplicaciones científicas en las GPU. La GPU NVIDIA Tesla Serie 20 pertenece a la tercera generación de la arquitectura CUDA que ha sido diseñada para el cálculo paralelo.

La principal ventaja de la serie Tesla es el aumento de redimiendo en potencia de cálculo de precisión doble producido en gran parte por el aumento en el número de transistores que han creado un chip más eficiente. También se destaca una mejora para el tratamiento de grandes cantidades de datos y comunicación PCIe gracias a la tecnología Nvidia GPUDIRECT que se detalla a continuación ^[58].

La tecnología Nvidia GPUDirect facilita la comunicación entre la GPU y otros dispositivos del bus PCIe evitando que la CPU se sobrecargue de procesamiento innecesario. Con la versión GPUDirect v1.0 se ha permitido que otros controladores de dispositivos de otras marcas participen en la comunicación, por ejemplo, los adaptadores de infiniband. Mediante este software dichos adaptadores se pueden comunicar directamente con los controladores de CUDA, lo que elimina la necesidad de que los datos pasen por la CPU. Mientras que la versión GPUDirect v2.0 ha proporcionado una comunicación de igual a igual basada en el protocolo p2p entre dos GPUs del mismo sistema, con lo que también se evita carga de trabajo adicional a la CPU.

Algunas herramientas que el fabricante Nvidia proporciona para facilitar el desarrollo de software en la serie Tesla son:

- Controlador de TCC para WINDOWS: Es un controlador de rendimiento que permite utilizar el escritorio remoto, servicios de Windows y reduce la carga de trabajo de la ejecución del Kernel CUDA en Windows.
- Monitorización de GPU: Mediante el software nvsmi se dispone de una herramienta para monitorizar la GPU. Esta herramienta proporciona la temperatura de la GPU y la velocidad.

La información que ofrece el fabricante acerca de los productos que se pueden encontrar en el mercado de la Serie 20 de Tesla es la siguiente.

Tabla 9. Comparativa Tesla ^[59] .

	Tesla C2075	Tesla C2070	Tesla C2050
Pico de rendimiento de operaciones con coma flotante de doble precisión	515 Gigaflops	515 Gigaflops	515 Gigaflops
Pico de rendimiento de operaciones en coma flotante con precisión simple	1030 Gigaflops	1030 Gigaflops	1030 Gigaflops
Núcleos CUDA	448	448	448
Cantidad de memoria (GDDR5)	6 GigaBytes	6 GigaBytes	3 GigaBytes
Ancho de banda de memoria	144 GBytes/s	144 GBytes/s	144 GBytes/s
TDP	225W	238W	238W

3.2.1.1.4.2. GEFORCE

La denominación GeForce hace referencia a la serie de tarjetas gráficas que desarrolla el fabricante estadounidense NVIDIA. En la actualidad la serie GeForce ha conocido trece generaciones a lo largo de su existencia. Además existen versiones profesionales de estas unidades, conocidas con el nombre de Quadro ^[60] .

Varios fabricantes también utilizan los procesadores de NVIDIA para crear tarjetas gráficas, en las cuales se puede destacar a Gigabyte, Asus, Msi, Bfg, XFX.

La arquitectura Fermi se ha introducido en la serie Geforce 400. La última serie en salir al mercado ha sido la GeForce 500, dicha serie de tarjetas gráficas ha incluido algunos cambios significativos con respecto a la serie GeForce 400 sobre todo en términos de rendimiento y gestión de energía. Al igual que la serie GeForce 400, las NVIDIA GeForce 500 son compatibles con DirectX 11, OpenGL 4.1 y OpenCL 1.0. En las tablas 10 y 11, se puede ver los diferentes modelos de la serie 500.

Tabla 10. Nvidia GeForce Serie GTX 500 ^{[61][62]} .

	Núcleos CUDA	Reloj del procesador (MHz)	Cantidad de memoria estándar	Ancho de banda de memoria (GB/s)	Consumo máximo de la GPU (W)
GeForce GTX 590	1024	1215	3072 MB (1536 MB por GPU) (GDDR5)	327.7	365
GeForce GTX 580	512	1544	1536 MB (GDDR5)	192.4	244
GeForce GTX 570	480	1464	1280 MB (GDDR5)	152.0	219
GeForce GTX 560 Ti	384	1645	1024 MB (GDDR5)	128.0	170
GeForce GTX 560	336	1620	1024 MB (GDDR5)	128.0	150
GeForce GTX 555	288	776	1024 MB (GDDR5)	91.9	150
GeForce GTX 550 Ti	192	1800	1024 MB (GDDR5)	98.4	116

Tabla 11. Nvidia GeForce Serie GT 500 ^{[61][62]} .

	Núcleos CUDA	Reloj del procesador (MHz)	Cantidad de memoria estándar	Ancho de banda de memoria (GB/s)	Consumo máximo de la GPU (W)
GeForce GT 545 GDDR5	144	1740	1024 MB (GDDR5)	64	105
GeForce GT 545 DDR3	144	1440	1536 MB (DDR3)	43	70
GeForce GT 530	96	1400	1024 MB (DDR3)	28.8	50
GeForce GT 520	48	1620	1024 MB (DDR3)	14.4	29
GeForce GT 510	48	1046	1024 MB (DDR3)	14.4	25

3.2.1.2. TECNOLOGÍA KEPLER

La arquitectura Kepler ha sido presentada con el producto GK104 que dispone de 1536 procesadores y 3.5 mil millones de transistores, fabricados en un proceso de 28nm. A finales del mes de marzo de 2012 se presentó el primer chip de escritorio denominado Nvidia Geforce 680 GTX. En este apartado se recogerá la información de la arquitectura que sucede a Fermi por parte del fabricante Nvidia ^[60] .

3.2.1.2.1. ESPECIFICACIONES GENERALES

En la nueva arquitectura, Nvidia ha rediseñado sus propios Stream Processors o Cuda Cores, en Kepler habrá muchos más, pero serán algo más simples. La finalidad de este rediseño es conseguir compartir elementos entre cada SP y por lo tanto ahorrar espacio. Nvidia no ha especificado una equivalencia directa entre los Cuda Cores de Kepler y los de Fermi, la única información que han proporcionado es que la nueva arquitectura Kepler será el doble de rápida por vatio consumido que la arquitectura de la anterior generación ^[63] ^[64].

La primera tarjeta comercial basada en la arquitectura Kepler se denomina Nvidia GTX 680, este chip gráfico está compuesto de 8 bloques SMX como el mostrado en el ANEXO V. Cada SMX cuenta con 192 Stream Processors, por tanto, el total de este modelo es de 1536 SP.

En la tabla 12 se puede ver una comparación de la evolución introducida por Kepler.

Tabla 12. Comparación modelos Nvidia 2 ^[64].

GPU	GT200 (Tesla)	GF110 (Fermi)	GK104 (Kepler)
Transistores	$1.4 \cdot 10^9$	$3 \cdot 10^9$	$3.5 \cdot 10^9$
Cuda Cores	240	512	1536
Frecuencia de reloj	648 MHz	722 MHz	1006 MHz
GFLOPs	1063	1581	3090
Unidades de Textura	80	64	128
Frecuencia de reloj de memoria	2484 MHz	4008 MHz	6008 MHz
Ancho de banda	159 GB/s	192.4 GB/s	192.26 GB/s
TDP	183 W	244 W	195W

De la tabla anterior se puede observar que el nuevo chip basado en Kepler ha mejorado en ciertos aspectos, como por ejemplo el aumento de frecuencia, el incremento de los Cuda Cores, el aumento de las unidades de Textura. Pero también tiene otros números que ha reducido para bien, como es el caso del TDP.

3.2.1.2.2. EFICIENCIA ENERGÉTICA

Nvidia ha trabajado en el consumo de su primera tarjeta de 28 nm, ha conseguido que la tarjeta 680 GTX tenga una frecuencia elevada de trabajo (1006 MHz) con un consumo reducido de 195W, cuanto menos consumo también se genera menos calor y se ahorra en ventiladores para expulsar el calor.

Además de la mejora en cuanto a tecnología reducida de fabricación, Nvidia ha incorporado al chip gráfico GTX 680 un sistema de control de consumo denominado GPU BOOST que se describe a continuación.

3.2.1.2.2.1. GPU BOOST

Cada aplicación hace un uso distinto de la GPU y por lo tanto la tarjeta consume más o menos dependiendo de la situación. La nueva Geforce GTX 680 tiene un límite máximo de consumo establecido y GPU Boost se encarga de controlar que no se supere dicha cifra.

El funcionamiento de GPU Boost es similar al Turbo Boost de Intel. Cuando no se esté ejecutando aplicaciones que utilizan toda la carga de la GPU no se alcanzará el límite de la tarjeta, por lo que en ese momento se puede aumentar la velocidad del núcleo gráfico para mejorar el rendimiento sin sobrepasar el límite establecido de consumo o de calor. En la figura 19 se muestra un ejemplo del funcionamiento.

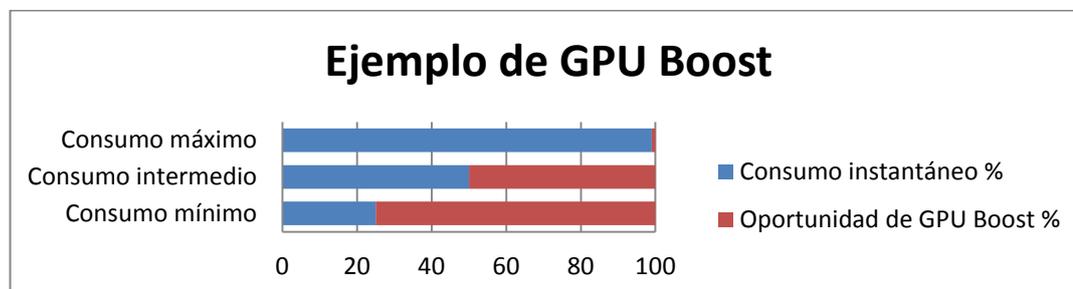


Figura 19. GPU Boost.

En el caso de la tarjeta gráfica 680 GTX, cuenta con una velocidad de reloj de 1006 MHz, pero si la aplicación no está utilizando todos los recursos de la tarjeta, el consumo no sobrepasará el TDP estipulado. Por lo tanto podrá subir su frecuencia hasta los 1058 MHz para acelerar el rendimiento de la aplicación.

3.2.1.2.3. MERCADO BASADO EN LA ARQUITECTURA

A continuación se muestran las especificaciones técnicas de la primera tarjeta gráfica basada en la nueva arquitectura que existe hasta la fecha.

Tabla 13. Nvidia GeForce Serie GTX 600 ^[62].

	Núcleos CUDA	Reloj del procesador (MHz)	Cantidad de memoria estándar	Ancho de banda de memoria (GB/s)	Consumo máximo de la GPU (W)
GeForce GTX 680	1536	1006/1058	2048 MB (GDDR5)	192.2	195

3.2.1.3. TECNOLOGÍA TEGRA

El chip Tegra que desarrolla el fabricante NVIDIA, es un dispositivo compuesto por un microprocesador que corresponde con la arquitectura ARM y un componente gráfico (GPU) de muy bajo consumo. En la actualidad estos chips se producen para dispositivos móviles o sistemas embebidos orientados a temas multimedia como son la reproducción de audio y de video. En cuanto a la implementación de la tecnología Tegra en temas de HPC, a finales de 2011 se ha dado pública una noticia relacionada con el centro de supercomputación de Barcelona. En dicho centro se pretende investigar y construir el primer supercomputador híbrido basado en CPU ARM y GPU. Por esta razón se recoge en este capítulo las principales características de la tecnología Tegra ^[65].

Los procesadores Tegra que actualmente produce el fabricante son Nvidia Tegra 2 y Nvidia Tegra 3 que se detallan a continuación.

Nvidia Tegra 2

El chip pertenece a la segunda generación de dispositivos Tegra. Cuyas características más destacables son ^[66] :

- **CPU de doble núcleo ARM Cortex A9:** Cortex A9 es el primer chip de doble núcleo para terminales móviles que ya se ha detallado en el apartado de ARM de esta tesis.
- **GPU GeForce con ULP:** Se trata de la tecnología de gráficos NVIDIA, diseñada para aplicaciones que demanda poca energía. Permite que aplicaciones exigentes tengan un bajo consumo energético.
- **Sistema global de administración de energía:** Mediante esta herramienta y los sensores que dispone la arquitectura Tegra, se puede recopilar todos los datos de actividad, frecuencia de trabajo, temperatura y operaciones. Dichos datos se utilizan para determinar la condición de funcionamiento óptimo y garantizar que los procesadores siempre se ejecuten con el mínimo nivel posible de energía para proporcionar el máximo rendimiento cuando se necesite.

Nvidia Tegra 3

Este chip pertenece a la tercera generación de dispositivos Tegra. El chip destaca por su GPU de bajo consumo y por su CPU compuesta por hasta 5 núcleos:

- **Super 4-PLUS-1:** El chip Tegra tiene cuatro núcleos más el quinto que se encarga de conservar la energía. Está compuesto por una arquitectura variante que permite el uso de los cuatro núcleos para ofrecer el máximo rendimiento, pero cada uno de los núcleos se activa y desactiva de forma independiente y automática dependiendo de la carga de trabajo. El núcleo para el ahorro de batería se encarga de las tareas de bajo consumo por ejemplo, en los terminales móviles son el consumo en modo espera, las aplicaciones secundarias del sistemas operativo, etc. ^{[67] [68]} .

3.2.2. FABRICANTE ATI

El fabricante AMD en la actualidad propone dos modelos de tarjetas gráficas. Por un lado ofrece la Serie HD 6000 (Northern Islands) y por otro lado a finales de 2011 anunció la nueva serie HD 7000 cuyos primeros modelos se están empezando a comercializar en el primer cuatrimestre del año 2012.

3.2.2.1. SERIE HD 6000 (NORTHERN ISLANDS)

La serie Radeon HD 6000 se lanzó al mercado en octubre de 2010 con el objetivo de conseguir un rendimiento similar a su serie anterior (Serie 5000), pero con un menor consumo y un menor precio final. Las primeras tarjetas gráficas que se comercializaron fueron la gama media Radeon HD 6870/6850 el 22 de octubre de 2010, el 15 de diciembre de 2010 salieron las de gama alta. Por último el 8 de marzo de 2011 se presentó el modelo de más alta gama denominado Radeon HD 6990. A continuación se detallan las principales características de la serie ^[69] .

La principal novedad de esta serie relacionada con la HPC es la nueva denominación EyeSpeed que hace referencia al tema de GPU para cálculos de propósito general (GPGPU), antes denominado ATI Stream. Las características de la nueva denominación son las siguientes:

- **Tecnología de Procesamiento Paralelo Acelerado (APP):** La GPU es compatible con nuevos estándares, como son el Open CL 1.1, Directcompute 11. Además la tecnología APP también se encarga de maximizar el rendimiento de las operaciones de coma flotante con doble precisión y de la codificación y conversión de video.
- **Acelerador de reproducción de video dedicado UVD3:** La serie de chips incluye una nueva mejora en el sistema de decodificación de video unificado, ahora denominado de tercera generación, dando soporte a MPEG-4 AVC/H.264, VC-1, MPEG-2 (SD y HD), Codificación multivista (MVC) y Adobe Flash.
- **Nueva versión del kit de desarrollo:** Para que los programadores puedan aprovechar las nuevas posibilidades de paralelismo han publicado una versión nueva de su herramienta (AMD APP SDK v2.2) con certificación OpenCL 1.1. Para poder trabajar con la GPU.

3.2.2.1.1. MERCADO BASADO EN LA ARQUITECTURA

En la tabla 14 se muestra el mercado de la serie 6000 de AMD.

Tabla 14. AMD Radeon Serie 6000 ^[70] .

	Líneas de procesamiento	Reloj del procesador (MHz)	Cantidad de memoria estándar	Ancho de banda de memoria (GB/s)	Consumo máximo de la GPU (W)
AMD Radeon HD 6990	3072	1250	4 GB (GDDR5)	320	375
AMD Radeon HD 6970	1536	1375	2 GB (GDDR5)	176	250
AMD Radeon HD 6950	1408	1250	2 GB (GDDR5)	160	225
AMD Radeon HD 6870	1120	1050	1 GB (GDDR5)	134.4	151
AMD Radeon HD 6850	960	1000	1GB (GDDR5)	128	127
AMD Radeon HD 6790	800	850	1GB (GDDR5)	134.4	150
AMD Radeon HD 6770	720	850	1 GB (GDDR5)	73.6	108
AMD Radeon HD 6750	720	700	1 GB (GDDR5)	73.6	86
AMD Radeon HD 6670	480	800	1 GB (GDDR5)	64	66
AMD Radeon HD 6570	480	650	512MB DDR3 / 1 GB (GDDR5)	28.8 / 64	44
AMD Radeon HD 6450	160	625	512MB DDR3 / 1 GB (GDDR5)	12.8 / 28.8	27

3.2.2.2. SERIE HD 7000 (SOUTHERN ISLANDS)

La serie 7000 del fabricante ATi apareció en el mercado en el primer trimestre del 2012, estas nuevas GPUs se fabrican bajo una nueva arquitectura denominada Graphics Core Next. La información técnica que ha proporcionado el fabricante es la siguiente ^[71] :

- Proceso de fabricación con litografía de 28 nm.
- Direccionamiento de 64 bits.
- Cuando se produce un fallo en el paginado, la GPU se encarga de enviar una interrupción a la CPU avisando del error de memoria.
- Soporta lenguajes de alto nivel, como puede ser C++ para programar algoritmos de propósito general.
- La arquitectura del chip está basado en el modelo MIMD.
- Nueva mejora en la tecnología Eyefinity 2.0, con esta tecnología se permite dar salida de video hasta seis monitores, además permite diferente resolución en cada pantalla.
- Se ha mejorado el UVD3, la nueva versión del sistema de aceleración de decodificación de vídeo da soporte para MVC (Multi View Codec) que permite la decodificación de contenidos 3D estereoscópicos.
- VCE (Video Codec Engine), mediante esta función se permite que la GPU ayude a la CPU en trabajos de codificación de video para permitir una compresión más ajustada.

3.2.2.2.1. EFICIENCIA ENERGÉTICA

Las mejoras que incluye la nueva serie de tarjetas gráficas del fabricante ATI son ^[72] :

- **PowerTune:** Esta tecnología permite al chip gráfico regular dinámicamente el TDP en función del rendimiento, para conseguir ajustar el consumo.
- **Nuevo método de refrigeración:** Se utiliza un método de refrigeración de cámara líquida en vez de cámara de vapor. La diferencia entre ambos métodos reside en que la cámara líquida tiene mayor cantidad de refrigerante que la cámara de vapor, además se ha mejorado el contenedor de líquido para evitar pérdidas por porosidad del material.
- **AMD ZeroCore:** Mediante esta tecnología los chips gráficos reducen el consumo estando en reposo, llegando incluso a parar el sistema de refrigeración o utilizando el mínimo para llegar a un consumo energético muy bajo.

3.2.2.2.2. MERCADO BASADO EN LA ARQUITECTURA

En la tabla 15 se muestra el mercado de la serie 7000 de AMD.

Tabla 15. AMD Radeon Serie 7000 ^[73] ^[74] .

	Líneas de procesamiento	Reloj del procesador (MHz)	Cantidad de memoria estándar	Ancho de banda de memoria (GB/s)	Consumo máximo de la GPU (W)
AMD Radeon HD 7970	2048	1375	3 GB (GDDR5)	265	300
AMD Radeon HD 7950	1792	1250	3 GB (GDDR5)	240	300
AMD Radeon HD 7870	1280	1200	2 GB (GDDR5)	153.6	225
AMD Radeon HD 7850	1024	1200	2GB (GDDR5)	153.6	225
AMD Radeon HD 7770	1125	640	1GB (GDDR5)	72	125
AMD Radeon HD 7750	1125	512	1GB (GDDR5)	72	75

3.2.2.3. APUS AMD (AMD FUSION)

AMD Fusion es el nombre comercial que se le dio a una serie de Unidades de Procesamiento Acelerado (APUs), que están siendo desarrolladas desde el 2006. El chip final es una fusión entre AMD y ATI combinando poder de procesamiento y funciones de GPUs actuales en un mismo encapsulado. La tecnología se mostró al público en el CES del 2011, en enero de 2012 se rebautizó el nombre comercial y pasó a llamarse AMD Heterogeneous Systems Architecture (HSA) ^{[75][76]}.

3.2.2.3.1. PLATAFORMAS ACTUALES

En la actualidad existen dos plataformas: **Brazos** compuesta por Ontario y Zacate y la plataforma **Llano** que es la plataforma que se encuentra actualmente en el mercado. Llano se presentó a mediados de 2011 y está basada en la arquitectura K10 de AMD, mientras que la plataforma Brazos (Ontario y Zacate) está basada en la nueva arquitectura Bobcat. A continuación se exponen las principales características de cada plataforma.

Brazos

Fue presentada el 5 de enero de 2011 como una plataforma orientada principalmente al mercado de subportátiles y netbooks. Destacan los procesadores de 40 nm Ontario (Apu con TDP de 9W) y Zacate (APU de 18 W). Ambos procesadores cuentan con uno o dos núcleos que están basados en la arquitectura Bobcat x86, soporte para DirectX 11 y OpenCL. Las características principales de **Ontario y Zacate** se muestran a continuación.

- Proceso de fabricación de Ontario y Zacate basado en la tecnología de 40 nm.
- Todos los modelos soportan: SSE, SSE2, SSE3, SSSE3, SSE4a, Nx bit, AMD64, PowerNow!, AMD-V.
- Soporte de memoria: DDR3 SDRAM, DDR3L SDRAM (Un canal, hasta 1066 MHz).
- Soporte GPU: Shaders Unificados (Vertex shader/Geometry shader/Pixel shader)

Llano

A continuación se muestran las principales características de la plataforma.

- Procesador basado en la tecnología de fabricación de 32nm.
- Se basa en una CPU actualizada que se denomina Stars y una GPU denominada Redwood.
- Procesadores de dos, tres y cuatro núcleos.
- Reemplazará a los Athlon II.
- Incluye una controladora para PCIe 2.0 y soporta DDR3-1600 de doble canal.
- No dispone de caché L3 y tiene 1 MB de caché L2 ^[75].

3.2.2.3.2. MERCADO BASADO EN LA ARQUITECTURA

Como se ha comentado anteriormente en la actualidad el fabricante AMD comercializa los productos basados en la plataforma Llano. En la tabla 16 se muestran dichos productos.

Tabla 16. Familia AMD APu ^[77] .

Procesador	Velocidad de reloj	Memoria Caché L2	TDP Max	Gráficos HD	Núcleos
A5-3650 APU	2.6GHz	1024KB	100W	Radeon HD 6530D	4
A6-3500 APU	2.1GHz	1024KB	65W	Radeon HD 6530D	3
A4-3400 APU	2.7GHz	1024KB	65W	Radeon HD 6410D	2
A8-3850 APU	2.9GHz	4096KB	100W	Radeon HD 6500D	4
A4-3300 APU	2.5GHz	1024KB	65W	Radeon HD 6410D	2

4. DISEÑO Y ELECCIÓN

En este capítulo, está recogida toda la información relacionada con el estudio teórico llevado a cabo con el fin de seleccionar los componentes del prototipo.

4.1. Estudio teórico y selección del prototipo

A continuación está explicado el método utilizado para seleccionar teóricamente los componentes CPU y GPU del prototipo eficiente.

4.1.1. ESTUDIO DE LA CPU

El presente estudio se divide en dos partes. Por un lado, se ha seleccionado el mejor microprocesador de la gama de escritorio actual de la familia x86-64. Por otra parte, se compara dicho procesador con los mejores microprocesadores de otras arquitecturas menos extendidas. El uso de la arquitectura x86-64 es un requisito impuesto por la necesidad de mantener la compatibilidad con el software más usado en el Área (sistemas operativos, paquetes de software de simulación, etc.).

4.1.1.1. ARQUITECTURA X86-64

Para conseguir seleccionar la CPU más eficiente de la actualidad se ha realizado un estudio por fabricantes. Los microprocesadores que se han incluido son los más actuales que se encuentran en el mercado de escritorio, ya que los microprocesadores denominados móviles no se pueden adquirir individualmente, aunque por norma general los microprocesadores móviles tienen un menor consumo.

Los fabricantes analizados en este estudio han sido Intel y AMD. Dentro de Intel se encuentran los microprocesadores de la arquitectura Sandy Bridge descrita anteriormente en el apartado “Estudio de las diferentes arquitecturas de procesadores”, también por parte de Intel se ha incluido la anterior arquitectura denominada Nehalem. En cuanto al fabricante AMD se han incluido los microprocesadores basados en la arquitectura K10, como son los Phenom II y los Athlon II, además de los microprocesadores basados en su última arquitectura Bulldozer.

Después de explicar las arquitecturas que se han analizado, se procede a describir el proceso de selección del microprocesador más eficiente. Para ello, lo primero que se ha de comprender es la expresión utilizada para calcular los GFLOPS/W de forma teórica. Antes de continuar con la explicación sobre cómo calcularlos, se define en qué consiste esta medida.

En el mundo de la informática los FLOPS hacen referencia a una medida del rendimiento de un computador para cálculos científicos que requieren un gran uso de operaciones de coma flotante, el término FLOPS deriva del inglés (floating point operations per second). Los computadores tienen un amplio rango de rendimientos en coma flotante, por lo que a menudo se usan unidades mayores que los FLOPS, en el caso de este estudio GFLOPS equivale a 10^9 FLOPS^[78].

Una vez comprendido la definición de lo que representa la unidad FLOPS, para poder seleccionar el microprocesador más eficiente, se ha realizado primero un cálculo del rendimiento teórico utilizando la expresión (4.1).

$$R.T.(GFLOPS) = Frecuencia\ de\ reloj\ (GHz) * Núcleos * Instrucciones\ por\ núcleo \quad (4.1)$$

Donde: la frecuencia del reloj viene dada en GHz por el fabricante, el número de núcleos físicos, también proporcionado por el fabricante y las instrucciones por núcleos dependen de la arquitectura. Las arquitecturas analizadas disponen de dos juegos de instrucciones, el set de instrucciones más antiguo es el SSE4 que en la expresión anterior para calcular los GFLOPS tiene un valor de 4 instrucciones/ (core*ciclo) para el fabricante Intel y AMD que soporten el set ^{[79][80]}. El segundo set de instrucciones denominado AVX ha sido desarrollado por el fabricante Intel que está presente en la arquitectura Sandy Bridge. Tiene, por parte del fabricante Intel, un peso de 8 instrucciones/(core*ciclo) ^{[81][82]}, mientras que en el caso de AMD el set de instrucciones AVX se incorpora dentro de la arquitectura Bulldozer, pero, al tener una unidad cálculo de coma flotante compartida entre dos cores, el peso asignado a esta variable es de 4 instrucciones/(core*ciclo). Para finalizar, se muestra el siguiente ejemplo del cálculo teórico de rendimiento pico de un microprocesador Intel i5-2500S Sandy Bridge que soporta el set de instrucciones AVX y cuyos datos son:

- Frecuencia de reloj: 2,7 GHz
- Núcleos físicos : 4
- Set de instrucciones : AVX

Por lo tanto, aplicando la expresión (4.1) se tiene:

$$\frac{2700 \cdot 10^6 \text{ ciclos}}{\text{segundo}} * 4 \text{ cores} * \frac{8 \text{ instrucciones}}{\text{core} * \text{ciclo}} = 8,64 \cdot 10^{10} \text{ FLOPS} \quad (4.2)$$

Convirtiendo los FLOPS a GFLOPS se tiene como resultado final 86,4 GFLOPS teóricos pico para el microprocesador Intel i5-2500S.

Para continuar con el estudio teórico se introduce el concepto de TDP, cuyo dato es proporcionado por el fabricante. La potencia de diseño térmico o TDP representa la máxima cantidad de calor que necesita disipar el sistema de refrigeración de un computador. Por ejemplo, si el TDP en un microprocesador es de 30 W significa que puede disipar 30 W de calor sin exceder la máxima temperatura de unión de los transistores en el circuito integrado, a partir de la cual el transistor deja de funcionar. Por lo tanto el TDP en ningún caso indica la máxima potencia que en algún momento el circuito integrado podría extraer, indica más bien la máxima potencia que podría extraer cuando se ejecutan aplicaciones reales ^[83]. En el prototipo de computador eficiente que se pretende diseñar para la presente Tesis de Máster, al estar trabajando con grandes cantidades de datos, el microprocesador se va a acercar al límite del TDP, por esta razón se puede seleccionar el TDP como una referencia de consumo máximo. En el caso de que el prototipo tuviera una función diferente, por ejemplo de servidor web, el TDP no sería un dato a tener en cuenta, ya que un servidor web puede atender o no peticiones. Por tanto, el microprocesador estaría consumiendo en función del número de peticiones/tareas y no estaría a máximo rendimiento durante todo el tiempo.

Siguiendo con el estudio teórico, se calcula con el TDP la relación GFLOPS/W. Para poder filtrar correctamente y no dejar atrás microprocesadores que en la práctica son eficientes pero teóricamente no destacan, lo que se ha realizado ha sido recurrir a la base de datos de SPEC. SPEC es un consorcio sin fines de lucro que incluye a vendedores de computadores, universidades, grupos de investigación, publicadores y consultores de todo el mundo. Su principal objetivo es crear un benchmark estándar para medir el rendimiento de los computadores y publicar los resultados de estos test. La base de datos

de SPEC se ha utilizado de la siguiente manera: se han tomado los datos de referencia denominados SPECfp2006 y SPECfp_rate2006:

- SPECfp2006: Este dato corresponde a la velocidad de ejecución de una sola tarea.
- SPECfp_rate2006: Este dato corresponde a la capacidad de la máquina de llevar a cabo un cierto número de tareas similares. Tradicionalmente usado para medir el rendimiento de multiprocesadores.

Posteriormente se han dividido los datos de SPEC entre el TDP para calcular los parámetros SPECfp2006/W y SPECfp_rate2006/W. Una vez calculados los parámetros GFLOPS/W (rendimiento teórico), además de los últimos mencionados (SPECfp2006/W y SPECfp_rate2006/W), se ha realizado un filtro para seleccionar aquellos candidatos que optan a ser el microprocesador más eficiente.

Continuando con el estudio, se ha realizado una suma ponderada de los datos teóricos (calculados mediante la expresión 4.2) y los datos prácticos (proporcionados por SPEC), variando los diferentes pesos que se le pueden dar a los diferentes datos, es decir, primero se otorga un porcentaje del 1% al dato teórico, mientras que el dato práctico recibe el porcentaje del 99%. En la siguiente suma ponderada, se incrementa el valor a 2% en el peso teórico, mientras que el valor otorgado al dato práctico decrece al 98% y así sucesivamente hasta que el peso teórico alcanza el 99%, mientras que el dato práctico estaría con el 1%. Una vez que se obtiene la gráfica recorriendo todos los pesos posibles, se realiza el promedio y se obtiene así el microprocesador más eficiente. La explicación de este cálculo se detalla a continuación.

Se parte, de que el dato de SPEC es una medida de rendimiento. Dicho rendimiento viene dado por la siguiente expresión.

$$SPEC = \frac{\mu_{CPU}}{\mu_{CPUref}} \quad (4.3)$$

De la expresión (4.3), el numerador (μ_{CPU}) representa las tareas/s que ofrece la CPU que se quiere evaluar, mientras que el denominador (μ_{CPUref}) representa las tareas/s que ofrece la CPU de la máquina de referencia contra la que se compara. Por lo tanto, al normalizar la expresión (4.3) se obtiene:

$$SPECn = \frac{\mu_{CPU}}{\mu_{CPUref}} \cdot \frac{\mu_{CPUref}}{\mu_{CPUMAX}} \quad (4.4)$$

Simplificando la expresión (4.4) se obtiene:

$$SPECn = \frac{\mu_{CPU}}{\mu_{CPUMAX}} \quad (4.5)$$

La expresión (4.5), refleja por tanto, un cociente entre las tareas/s que es capaz de ofrecer cada computador.

Por otro lado, al normalizar la expresión de rendimiento teórico quedaría:

$$R.T.n = \frac{R.T.CPU}{R.T.CPUMAX} \quad (4.6)$$

La expresión (4.6), representa un cociente de GFLOPS entre dos CPUs.

Ya por último, dividiendo cada rendimiento de la CPU entre su consumo. De las expresiones (4.5) y (4.6) se obtiene las expresiones (4.7) y (4.8), que sirven para obtener los datos que se han utilizado para seleccionar a priori la CPU más eficiente.

$$SPECn = \frac{\mu_{CPU}/W_{CPU}}{\mu_{CPUMAX}/W_{CPUMAX}} \quad (4.7)$$

$$R.T._n = \frac{R.T._{CPU}/W_{CPU}}{R.T._{CPUMAX}/W_{CPUMAX}} \quad (4.8)$$

En la figura 20, se muestra la gráfica comparativa, utilizando los datos SPECfp2006/W y GFLOPS/W. En dicha figura, el eje de ordenadas, representa la suma de los datos anteriores normalizados, es decir, $(SPECfp2006/W)_n$ más $(R.T./W)_n$. Mientras que el eje de abscisas representa el peso otorgado al valor práctico (valor de SPEC).

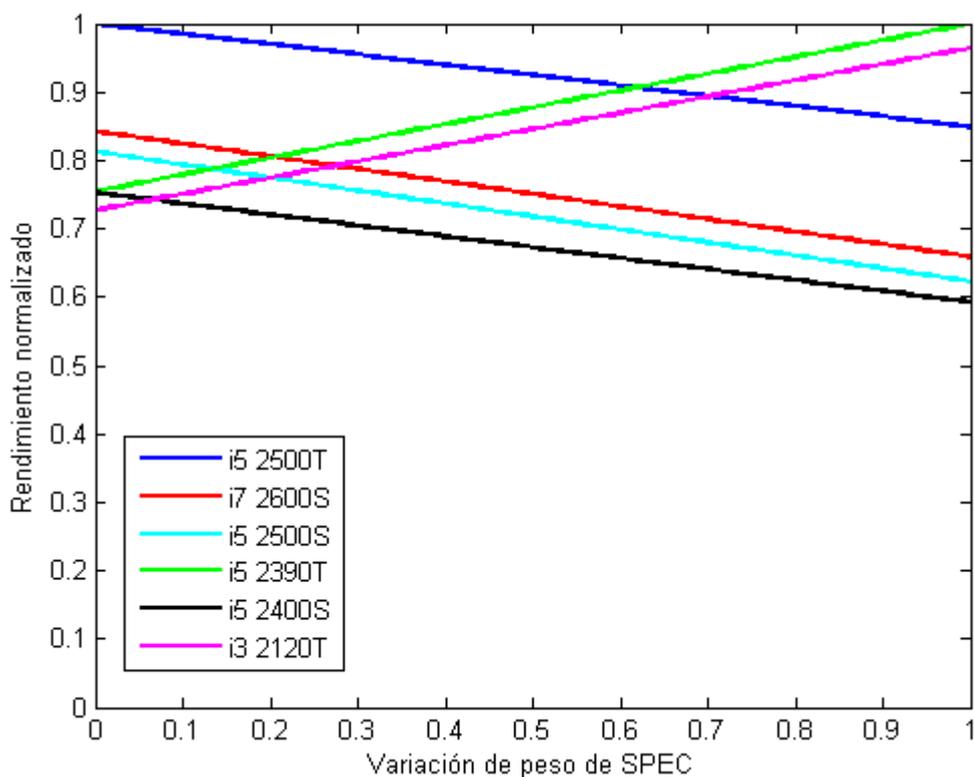


Figura 20. Variación del rendimiento normalizado utilizando el R.T. y SPECfp2006.

A continuación se muestran los promedios de la gráfica de la figura 20.

- i5_2500T = 0.9153
- i7_2600S = 0.7430
- i5_2500S = 0.7104
- i5_2390T = 0.8685
- i5_2400S = 0.6658
- i3_2120T = 0.8372

En la figura 21, se muestra la gráfica comparativa, utilizando los datos SPECfp_rate2006/W y GFLOPS/W. En dicha figura, el eje de ordenadas, representa la suma de los datos anteriores normalizados, es decir, $(\text{SPECfp_rate2006/W})_n$ más $(\text{R.T. /W})_n$. Mientras que el eje de abscisas representa el peso otorgado al valor práctico (valor de SPEC).

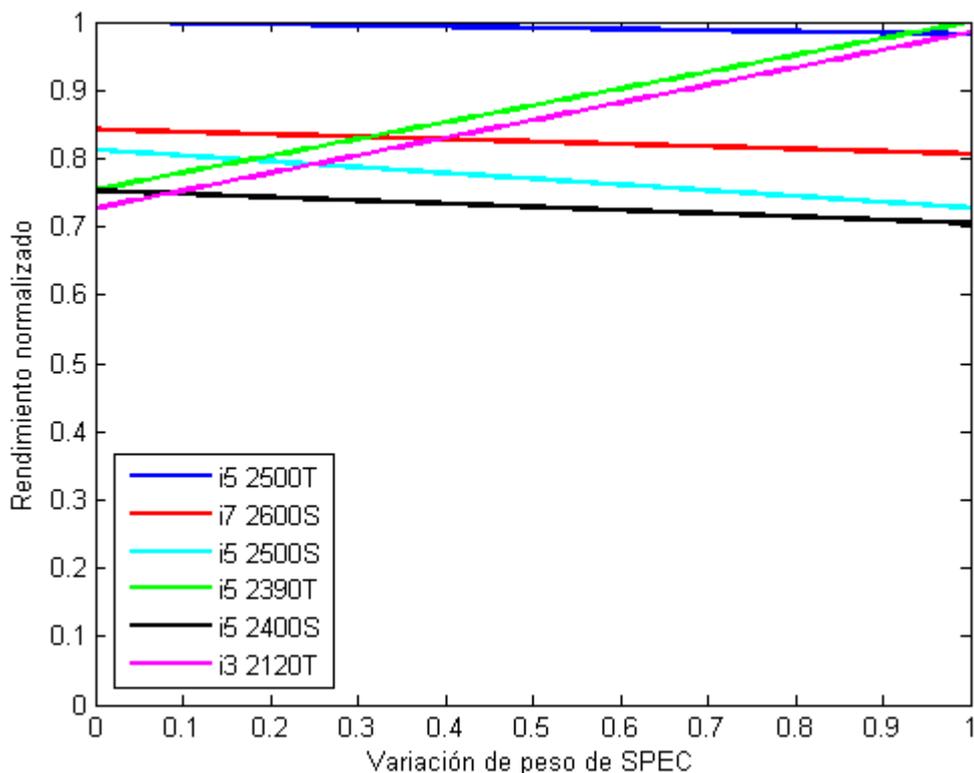


Figura 21. Variación del rendimiento normalizado utilizando el R.T. y SPECfp_rate2006.

A continuación se muestran los promedios de la gráfica de la figura 21.

- i5_2500T = 0.9812
- i7_2600S = 0.8163
- i5_2500S = 0.7623
- i5_2390T = 0.8685
- i5_2400S = 0.7216
- i3_2120T = 0.8471

En la actualidad, los algoritmos que se diseñan, suelen aprovechar la capacidad multitarea de la máquina. Por lo tanto, para la elección de la CPU más eficiente se debería tener más en cuenta a la figura 20, que otorga más peso a la productividad, es decir, a la capacidad de procesamiento multihilo.

Finalmente, para obtener la CPU más eficiente de forma teórica, se ha seleccionado aquella CPU con mejor promedio en el rendimiento. Obteniendo como resultado, que el microprocesador más eficiente es el Core i5 2500T.

4.1.1.2. OTRAS ARQUITECTURAS

En este apartado se tratan las arquitecturas SPARC y PowerPC, que son las arquitecturas que aparecen en la última lista (noviembre 2011) de “Top500.org” y “green500.org”.

- **SPARC Viiiifx 2.0 GHz**

El modelo Sparc Viiiifx 2.0 GHz, está colocado en la lista top500 en la primera posición, mientras que en la lista green500 está en la posición 32, cuyos datos se muestran a continuación:

Tabla 17. Datos Top500. Sparc Viiiifx 2.0 GHz.

Sistema	Vendedor	Cores	Rpeak (TFlops)	Power (kW)
K computer, SPARC64 Viiiifx 2.0 GHz	Fujitsu	705024	11280.38	12659.9

El rendimiento teórico en esta CPU sería **0,891 GFlops/W**.

- **PowerPC BQC 1.60 GHz**

El modelo Power BQC, está colocado en la lista top500 en la posición 64, mientras que en la lista green500 está en la primera posición, cuyos datos se muestran a continuación:

Tabla 18. Datos Top500. PowerPC BQC 1.60 GHz.

Sistema	Vendedor	Cores	Rpeak (TFlops)	Power (kW)
BlueGene/Q, Power BQC 16C 1.60 GHz	IBM	16384	209.72	85.12

El rendimiento teórico en esta CPU es de **2.464 GFlops/W**.

Por último, cabe destacar que el microprocesador que se ha seleccionado (i5-2500T), tiene un rendimiento teórico de **1,636 GFlops/W**. Por lo tanto, el microprocesador i5-2500T sería más eficiente que la arquitectura SPARC Viiiifx 2.0 GHz pero bastante peor que la arquitectura PowerPC BQC 1.60 GHz.

4.1.2. ESTUDIO DE LA GPU

Otro punto importante que se ha tratado en la presente Tesis de Master, ha sido buscar un chip GPU eficiente. Para conseguir este objetivo se ha realizado un estudio teórico por fabricantes en el cual se han incluido los chips gráficos más actuales del mercado de escritorio. Como sucedía con la CPU, los chips gráficos móviles no se pueden adquirir individualmente por lo tanto no se han incluido en este estudio.

Los fabricantes que se han analizado en este estudio han sido Nvidia y AMD. Por parte del fabricante Nvidia, se han analizado los chips gráficos basados en la arquitectura Fermi además del último chip gráfico que se ha incorporado al mercado basado en la nueva arquitectura Kepler. En cuanto al fabricante AMD se han estudiado las gráficas de la Serie HD 6000 (Northern Islands) y la Serie HD 7000 (Southern Islands).

Para realizar este estudio se han utilizado dos métricas. Por un lado se ha evaluado la capacidad de cálculo teórico por vatio y por otro lado se ha tenido en cuenta el ancho de banda de memoria de la GPU, ya que es de vital importancia para enviar la información a los núcleos.

Para seleccionar el chip GPU más eficiente se ha calculado los GFLOPS que es capaz de ofrecer una tarjeta gráfica con las siguientes expresiones:

$$\text{Para Nvidia} \rightarrow R.T. (FLOPS) = \text{Cores CUDA} * \text{Reloj} \left(\frac{\text{Ciclos}}{s} \right) * 2 \left(\frac{\text{Instrucciones}}{\text{Ciclo} * \text{core}} \right) \quad (4.9)$$

$$\text{Para AMD} \rightarrow R.T. (FLOPS) = \text{Cores AMD} * \text{Reloj} \left(\frac{\text{Ciclos}}{s} \right) * 2 \left(\frac{\text{Instrucciones}}{\text{Ciclo} * \text{core}} \right) \quad (4.10)$$

Donde: los cores CUDA, los cores de AMD (o líneas de procesamiento AMD) y el dato de Reloj en las anteriores expresiones, son datos proporcionados por los fabricantes.

En el siguiente ejemplo se muestra el cálculo para el modelo de tarjeta gráfica GeForce GTX 555, cuyos datos son:

- Frecuencia de reloj: 776 MHz
- Cores CUDA: 288

Por lo tanto, aplicando la expresión (4.9) por parte del fabricante Nvidia, se obtiene que el rendimiento teórico pico en precisión simple es de 446,976 GFLOPS.

Como sucedía al realizar el estudio de la CPU eficiente, una vez calculados los GFLOPS teóricos se utiliza el TDP para calcular el término GFLOPS/W.

El segundo parámetro a tener en cuenta ha sido el ancho de banda de memoria. El ancho de banda de memoria en una GPU es un dato importante. Por muchos núcleos que tenga la GPU, sino hay un buen canal de comunicación que suministre datos a los cores, el funcionamiento/rendimiento de la tarjeta gráfica no sería óptimo. En las arquitecturas analizadas anteriormente de CPU este dato no es de vital importancia porque los anchos de banda en las CPU son muy similares y no sirven para marcar una diferencia. A la hora de evaluar el ancho de banda para esta Tesis, se ha seguido el modelo anterior de expresiones, en donde, se ha utilizado el TDP para calcular la relación ancho de banda de memoria/W.

Una vez obtenidos los dos términos, se realiza una suma normalizada dando igual peso tanto al parámetro de ancho de banda de memoria/W como al parámetro GFLOPS/W.

En el estudio se observa que las tarjetas gráficas más eficientes son las proporcionadas por el fabricante AMD. El Área de TSC de la Universidad de Oviedo tiene códigos programados en lenguaje CUDA y debido a que CUDA no está implementado en las tarjetas de AMD se han descartado.

Por último, se ha de concluir que aunque el ganador oficial sea el modelo GeForce 680 GTX, esta tarjeta ha quedado descartada para el prototipo, ya que en la actualidad no está bien desarrollado el driver que se utiliza en el ámbito de la supercomputación y por lo tanto no está bien implementado CUDA. Por lo tanto, como selección de GPU más eficiente se ha optado por la tarjeta GeForce GTX 560.

5. PRUEBAS DE RENDIMIENTO

En el siguiente capítulo está recogida la información acerca de los propios resultados obtenidos después de realizar las mediciones y las herramientas necesarias para obtenerlos.

5.1. Métricas utilizadas

Para la obtención de los resultados, se ha optado por utilizar:

- Un multímetro “Velleman DVM890” como el que se observa en la figura 22.
- Un multímetro “Hameg HM8112-3”.



Figura 22. Multímetro Velleman DVM890.

La principal tarea del multímetro Velleman DVM890, conectado al cable de alimentación que se observa en la figura 22, ha sido mostrar la intensidad de corriente eléctrica mientras el computador ejecuta el banco de pruebas de rendimiento (o abreviadamente, del inglés benchmark). Para el estudio realizado, se ha seleccionado siempre el valor pico de corriente eléctrica (valor máximo) durante la ejecución del benchmark analizado. Una vez obtenida la intensidad de corriente medida en amperios, además de la tensión que alimenta al computador (previamente medida al inicio de cada sesión en el laboratorio) se puede obtener la potencia pico consumida por el computador en Watios. Para finalizar, con esa potencia obtenida y el tiempo de ejecución del benchmark, se puede calcular los Julios y el EDP (Energy delay product), que es una medida que se suele aplicar en estos entornos de trabajo ^[84]
_[85].

Otros sistemas de medidas que se han utilizado, han sido por un lado el tiempo en segundos que el computador tardar en ejecutar el benchmark y en algunos casos también se ha tenido en cuenta la información proporcionada por el benchmark. Como ha sucedido en el caso del benchmark Whetstone en donde después de ejecutarlo devuelve los MIPS (millones de instrucciones por segundo) que puede ofrecer la CPU.

5.2. Benchmarks utilizados

En este bloque se describen los benchmarks que se han utilizado para medir en el PC de referencia y en el prototipo eficiente.

5.2.1. WHETSTONE

Whetstone es un benchmark científico diseñado en el Laboratorio Nacional de Física de Inglaterra. Es considerado uno de los padres de los benchmarks sintéticos por ser el primero específicamente con este fin ^[86].

El benchmark recibe el nombre de Whetstone ya que fue en dicha localidad donde se desarrolló en la década de 1960 por Brian Wichmann. Dicho benchmark se había diseñado con el fin de crear un test para el compilador denominado ALGOL 60.

En 1972 aparece la primera implementación práctica de Whetstone escrita en lenguaje Fortran por Curnow ^[87], más adelante el benchmark se portó a lenguajes como Pascal, C, Fortran 77, etc. El benchmark fue diseñado para medir la velocidad de ejecución de una variedad de instrucciones de coma flotante. Lo que pretende es intentar estimar la velocidad de la CPU con la FPU.

El benchmark demostró ser una buena medida de rendimiento hasta el comienzo de la década de 1990, en donde se reconoció que Whetstone no funcionaría adecuadamente para medir el rendimiento de computadores con procesadores paralelos. Sin embargo, para el estudio, se ha creado un script denominado Whetstone_4.sh que lanza cuatro copias del mismo benchmark, con ello se consigue evaluar el rendimiento de un microprocesador en paralelo.

Los resultados que proporciona el benchmark se denominan MWIPS (Millones de Instrucciones Whetstone Por Segundo), en donde una instrucción Whetstone se puede definir como una instrucción de punto flotante promedio.

El benchmark está compuesto de diferentes bucles. Cada bucle tiene un tiempo de ejecución y un tipo de instrucciones diferentes. Al comenzar el código a ejecutarse se establece la cantidad de veces que se repetirá en cada bucle, es decir, se asigna un peso a cada tipo de instrucción o bucle. Para terminar se llama a cada uno de estos bucles para imprimir los resultados y calcular el rendimiento en MWIPS.

El código que ejecuta el benchmark es un código muy pequeño, cabiendo completamente en la caché interna del procesadores modernos, por lo tanto es capaz de mantener el pipeline de la FPU completa además de mantener a la FPU permanentemente ocupada.

El benchmark no es tan perfecto como se piensa, consume entre 40% y 50% del tiempo de ejecución en subrutinas matemáticas, por lo que las bibliotecas usadas para las pruebas pueden alterar los resultados de manera significativa. Además existe otro problema que afecta a los computadores basados en arquitecturas RISC. El problema está relacionado con la declaración de variables globales del benchmark, ya que este benchmark no mostraría la eficiencia de verdad en una arquitectura RISC. Los computadores basados en la arquitectura RISC están pensados para instrucciones pequeñas, también cuentan con un gran número de registros de procesador, por lo tanto para optimizar el rendimiento habría que utilizar variables locales.

5.2.2. LINPACK

El benchmark Linpack se desarrollo en Argone National Laboratory por Jack Dongarra en 1976, tiene fama mundial ya que es muy utilizado en entornos científicos y de ingeniería ^[88]. Para la presente Tesis de Master se ha optado por utilizar el código en C escrito por Bonnie Toy.

Linpack es un paquete de subrutinas de FORTRAN para resolver varios sistemas de ecuaciones lineales. El software de Linpack se basa en una descomposición utilizando el algebra lineal numérico. La idea general es la siguiente, dado un problema relacionado con una matriz A, se trata de descomponer A en un producto de matrices más sencillo, bien estructurado para que sea más fácilmente manipulable con el fin de resolver el problema original ^[89].

Linpack se basa en realizar operaciones de coma flotante de forma intensiva, por lo tanto, los resultados dependen de las capacidades de la FPU de que se disponga. El programa basa gran parte de su tiempo ejecutando las rutinas denominadas BLAS (Subrutinas de Álgebra Básica). El mayor tiempo en ejecución se consume en la rutina DGEFA, que se encarga de realizar la descomposición LU mediante pivoteo parcial. La siguiente rutina, denominada DGESL, aprovecha la descomposición realizada anteriormente para resolver el sistema de ecuaciones lineales.

Linpack se puede utilizar para medir la eficiencia en los sistemas multiprocesador, ya que realiza esencialmente los cálculos con matrices en un test fácilmente paralelizable. Por ejemplo, la página web Top 500 clasifica los computadores basándose en Linpack.

Los resultados que muestra vienen dados en MFlops (millones de operaciones de coma flotante por segundo), estos datos son referidos a operaciones de suma y multiplicación de doble precisión.

5.2.3. PHORONIX TEST SUITE

Phoronix Test Suite (PTS) es una herramienta de código abierto para Linux y otros sistemas operativos. El benchmark PTS ha sido respaldado por sitios como Linux.com, LinuxPlanet y ha sido denominado como “la mejor plataforma de referencia” por la Softpedia ^[90].

La herramienta nació a partir de los primeros desarrollos internos en Phoronix Media durante los años de funcionamiento de Phoronix.com, que era una de las primeras bases de datos de análisis de hardware para Linux. Phoronix se puso en marcha en junio de 2004, es a partir de finales de 2007 cuando se empezó a desarrollar el benchmark. El desarrollador principal del software es Michael Larabel, fundador de Phoronix. Otra persona a destacar del paquete software es el socio Matthew Tippet, encargado de la dirección global. Además más de dos docenas de diferentes corporaciones e individuos han contribuido directamente al desarrollo del paquete.

La herramienta de Phoronix viene con más de 130 perfiles de primera y 60 tipos de test de prueba. El rango de pruebas de la suite se compone de: monitorización del consumo de la batería, en dispositivos portátiles basado en benchmarks multi-hilos de trazado de rayos, medidas de la CPU, análisis de gráficos, análisis de la memoria del sistema, análisis del almacenamiento de disco y análisis de componentes de la placa base.

El benchmark Phoronix test suite está en continuo desarrollo, su código base es de aproximadamente 55000 líneas. El núcleo la suite está escrito en PHP5 para su compatibilidad con diversas plataformas y diferentes conjuntos de características, también hace uso de XML ^[91].

5.2.4. CÓDIGOS DEL ÁREA DE TSC DE LA UNIVERSIDAD DE OVIEDO

Para realizar mediciones de la eficiencia computacional del prototipo construido se han empleado varios códigos desarrollados por miembros del ÁREA de TSC (Teoría de la Señal y Comunicaciones) de la Universidad de Oviedo. A continuación se explica de forma muy somera en qué consiste cada código:

- MST: El código MST (del inglés Memory Saving Technique ^[61]) se emplea en el ámbito de diagnóstico de antenas para la reconstrucción de fuentes electromagnéticas. Permite la obtención de corrientes eléctricas y/o magnéticas equivalentes sobre la superficie de una antena a partir del procesado de las medidas de campo radiado por la misma. Partiendo de las corrientes reconstruidas es posible, por ejemplo, detectar defectos de diseño o de fabricación de la antena. Asimismo, el método permite obtener el campo en cualquier punto a partir de las corrientes reconstruidas. A fin de reconstruir las corrientes equivalentes, el código MST debe resolver un sistema de ecuaciones (en general no cuadrado) mediante el método CGNR ^[92] que comporta varios productos matriz vector por iteración y varias decenas de iteraciones. La resolución del sistema mencionado resulta de una alta exigencia computacional. El código está pensado para ejecutarse en una máquina, ya que utiliza paralización en memoria compartida.
- FMM: El código FMM (del inglés Fast Multipole Method ^[93]) posibilita la obtención eficiente del campo de presiones acústico dispersado por un objeto tridimensional iluminado por una fuente acústica. A tal fin, se plantea una ecuación integro-diferencial que relaciona la presión incidente sobre el objeto (conocida) con la presión total y su derivada respecto de la dirección normal (desconocida). Mediante el método de los elementos de contorno (BEM, Boundary Elements Method ^[94]) la ecuación mencionada anteriormente se planteará numéricamente, resultando en un sistema de N ecuaciones con N incógnitas cuya resolución puede resultar muy costosa. En el código FMM se emplea un método iterativo eficiente para la resolución de sistemas no simétricos (GMRES, Generalized Minimum Residual ^[95]) que requiere un producto matriz vector por iteración. Asimismo, el código mencionado agiliza el cálculo del producto matriz vector mediante un esquema acelerador denominado Fast Multipole Method. Para ejecutarse el código, se requiere la instalación de las librerías MPI (Interfaz de paso de mensajes). Además el código está pensado para paralización en memoria compartida, en memoria distribuida y también se ha programado para GPU.
- FMM-FFT: Este código implementa el algoritmo FMM-FFT ^[96] que es muy similar al del párrafo de arriba (FMM), pero utiliza la Fast Fourier Transform (FFT) que permite disminuir ligeramente el coste computacional del FMM.

5.3. Resultados

Este apartado contiene la comparación de los resultados obtenidos después de realizar las pruebas entre el PC de referencia y el prototipo.

5.3.1. WHETSTONE

Los resultados obtenidos se han producido compilando el benchmark con las siguientes opciones de compilación: `gcc -o2 whetstone.c -o Whetstone -lm`.

Para realizar las pruebas con el benchmark, se ha creado un script que se encarga de lanzar el programa utilizando diferentes hilos de la CPU. El script, tiene la función de evaluar el rendimiento de la CPU, utilizando Whetstone con 500000 repeticiones. Cabe destacar que, los resultados en el prototipo, para realizar una comparación más justa contra el PC de referencia, se han realizado sin la GPU conectada, ya que la GPU inactiva consume 14.5W.

Los resultados se pueden observar en el ANEXO VI. Mientras que a continuación se puede ver la comparación en gráficos.

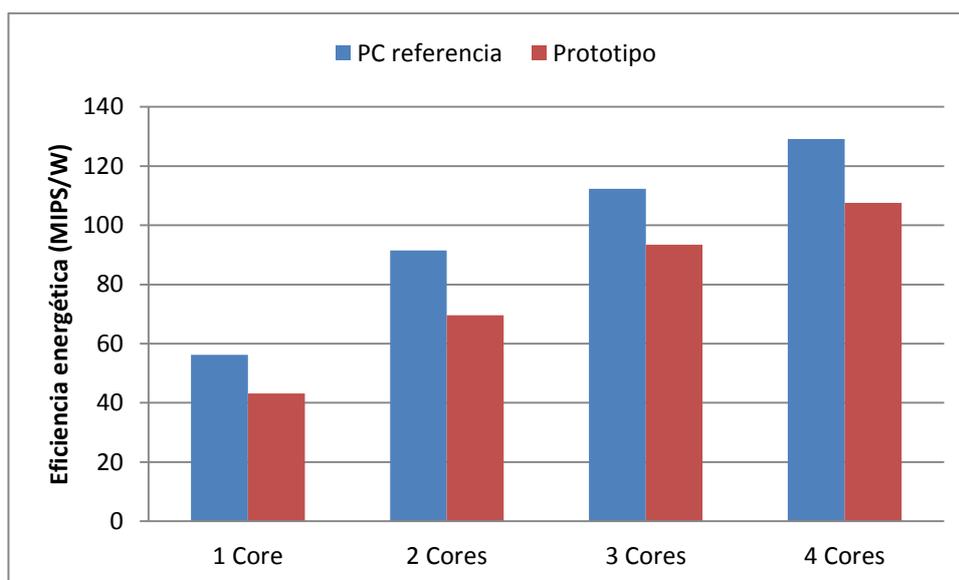


Figura 23. Eficiencia energética proporcionada por Whetstone para 500000 repeticiones.

Las figuras 23, muestran la eficiencia energética que ofrece la CPU del PC de referencia contra la CPU del prototipo. En ambas figuras, el eje de abscisas muestra el número de hilos que se han utilizado para realizar la prueba, mientras que el eje de ordenadas informa sobre la eficiencia energética medida en MIPS/W.

De la figura 23, se puede observar que ambos computadores han sido más eficientes cuando trabajan con 4 hilos en paralelo. Esto se debe, a que el aumento del consumo energético de utilizar un hilo a utilizar los cuatro hilos, es mínimo, mientras que el rendimiento aproximadamente, se va multiplicando por tantos cores como estén ejecutando el problema.

Los resultados en este experimento, han mostrado que el PC de referencia ha sido más eficiente que el prototipo. Estos resultados, se deben al aumento del consumo energético en el prototipo, debido al funcionamiento de la fuente de alimentación (de 750w). El fabricante de la fuente de alimentación, asegura que la fuente es eficiente cuando se trabaja por encima del 20% de su capacidad ^[97], es decir, cuando se trabaja por encima de 150W y en las pruebas realizadas, no se ha alcanzado el 12% del rendimiento total. Por lo tanto, la fuente en el prototipo, no ha sido eficiente. Mientras que por el contrario, la fuente de alimentación en el PC de referencia está bien dimensionada alcanzando un rendimiento óptimo.

5.3.2. MATLAB

Los resultados obtenidos de las diferentes pruebas realizadas con MatLab, se muestran en el ANEXO VII, mientras que a continuación se podrán ver los gráficos comparativos entre ambos computadores.

La primera prueba que se ha realizado con la herramienta MatLab, ha sido mediante el benchmark propio de MatLab. Para esta prueba, se debe tener en cuenta que algunos tests del benchmark interno de MatLab, requieren de aceleración gráfica. Por lo tanto, para realizar la prueba se ha conectado al prototipo de estación eficiente la GPU.

En la figura 24, se observan los resultados obtenidos al ejecutar el benchmark interno de MatLab en ambos computadores. El tiempo en el primer gráfico se representa en centésimas de segundo, mientras que la energía requerida para realizar la prueba se expresa en julios. Se debe destacar del gráfico, que para este tipo de medidas interesa que el dato sea menor, es decir, el computador más eficiente será aquel que tarde menos segundos en resolver el problema y que consuma la mínima cantidad de energía.

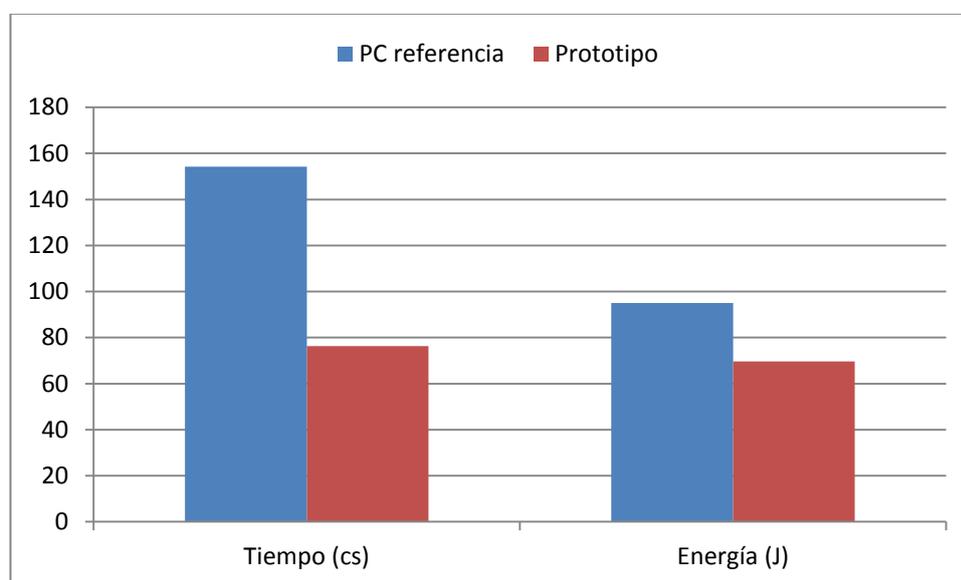


Figura 24. Resultados del tiempo (en centésimas de segundo) y energía requerida al realizar el benchmark interno de MatLab.

En el experimento realizado con el benchmark interno, cuyos resultados se muestran en la figura 24, se puede observar que el prototipo ha realizado la prueba más rápido, además de que ha consumido menos energía. Por lo tanto, se puede concluir diciendo que para esta prueba, el prototipo ha sido más eficiente energéticamente hablando que el PC de referencia.

Después de ejecutar el benchmark interno de MatLab, se han realizado dos pruebas más específicas dentro del mismo entorno de trabajo.

En la siguiente prueba se ha creado un programa, que al ejecutarlo, se encarga de generar dos matrices aleatorias del mismo tamaño. El programa tiene como objetivo, obtener el tiempo que se tarda en realizar el producto entre ambas matrices. A diferencia de lo que sucedía con el benchmark de Whetstone, el uso de este script implica que el computador requiera almacenar los datos de las matrices en la memoria RAM disponible. Por lo tanto, se ha jugado con este factor, creando matrices de diferentes tamaños, que abarcan desde una parte de los recursos hasta ocupar toda la memoria RAM disponible. Cabe destacar, que al realizar este experimento utilizando la GPU del prototipo, sólo se ha podido ocupar hasta 1GB, que es el tamaño máximo de memoria que tiene disponible el acelerador gráfico.

Los experimentos se han realizado por un lado, con datos de entrada de doble precisión (8 bytes por dato), y por el otro lado, se han utilizado datos de entrada de simple precisión (4 bytes por datos). Además, se debe saber del estudio realizado, que para aquellas mediciones, en donde sólo se requiere el uso de la CPU en el prototipo, se han realizado con la GPU desconectada.

En los gráficos que muestran los resultados del experimento del producto de matrices (figura 25, 26, 27 y 28), el eje de abscisas representa siempre el número de elementos de la matriz de entrada, mientras que el eje de ordenadas, o bien representa la energía consumida para realizar el producto de las matrices o el EDP requerido. Los resultados obtenidos se muestran en los siguientes gráficos comparativos.

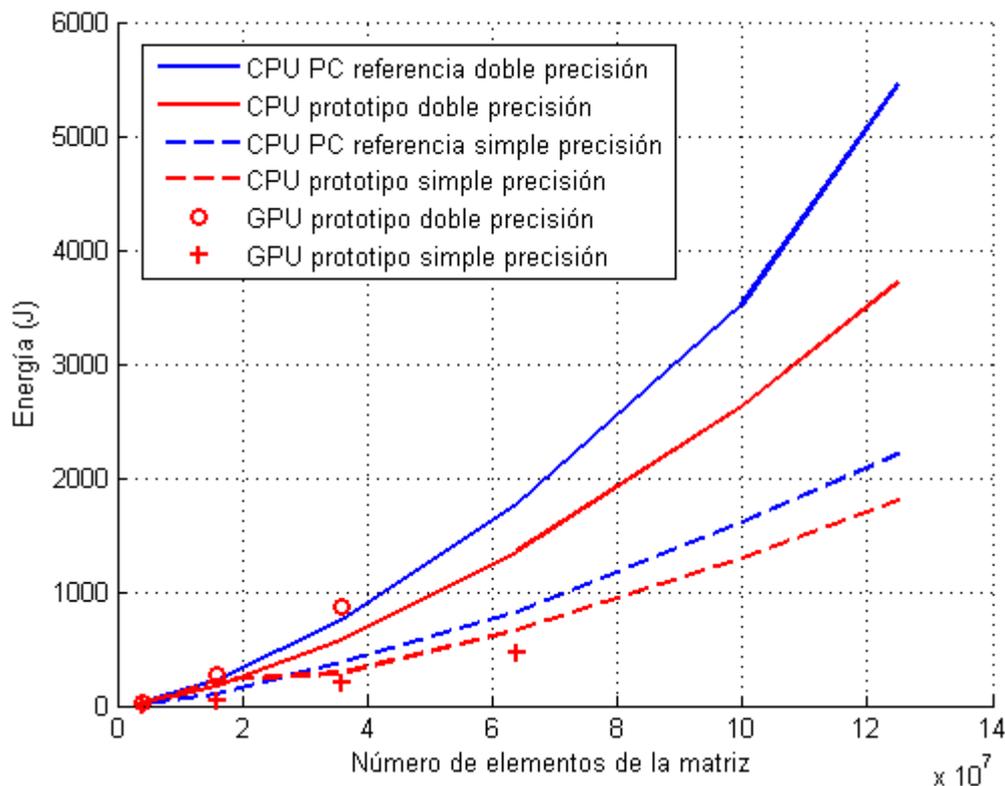


Figura 25. Comparación del consumo energético al realizar el producto de matrices.

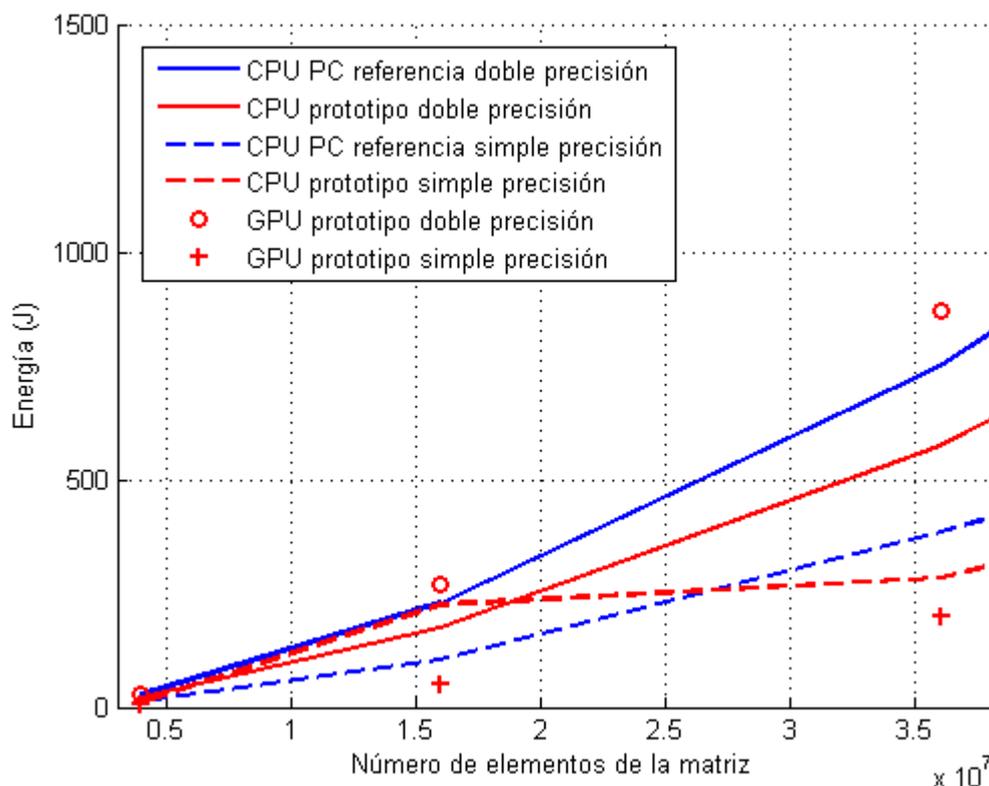


Figura 26. Vista en detalle del consumo energético al realizar el producto de matrices.

En los gráficos de las figuras 25 y 26, se observa la comparación de los resultados obtenidos en cuanto a consumo energético, después de realizar el producto de matrices mediante el MatLab, tanto en GPU como en CPU.

Lo primero que se puede destacar del gráfico de la figura 25, es el mayor consumo cuando se ha trabajado con la GPU del prototipo y datos de doble precisión. Sin embargo, al trabajar con la GPU, pero declarando los datos en simple precisión, se observa que se ha tenido el menor consumo energético. Además cabe destacar, que para esta prueba, el prototipo cuando trabaja con la CPU, tanto con datos simples, como con dobles, en general ha consumido menos que cuando se trabaja con el mismo tipo de datos en el PC de referencia.

En los gráficos de las figuras 27 y 28, se puede observar la comparación del EDP requerido para realizar la operación de producto de matrices en MatLab, para los diferentes tamaños de matrices.

Lo primero que se observa de los gráficos de las figuras 27 y 28, es que la CPU más eficiente ha sido la del prototipo, ya que es la CPU que menos EDP ha requerido para realizar el producto. Además, se puede observar que lo más eficiente energéticamente hablando en este experimento ha sido trabajar con la GPU y datos de simple precisión.

Por último, hay que destacar que a medida que crece el tamaño de la matriz, el prototipo ha sido todavía más eficiente. Esto es debido, a que el prototipo tiene a su disposición dos módulos de memoria RAM, por lo que se puede aprovechar del DUAL CHANNEL, realizando el envío de información más rápidamente que en el PC de referencia.

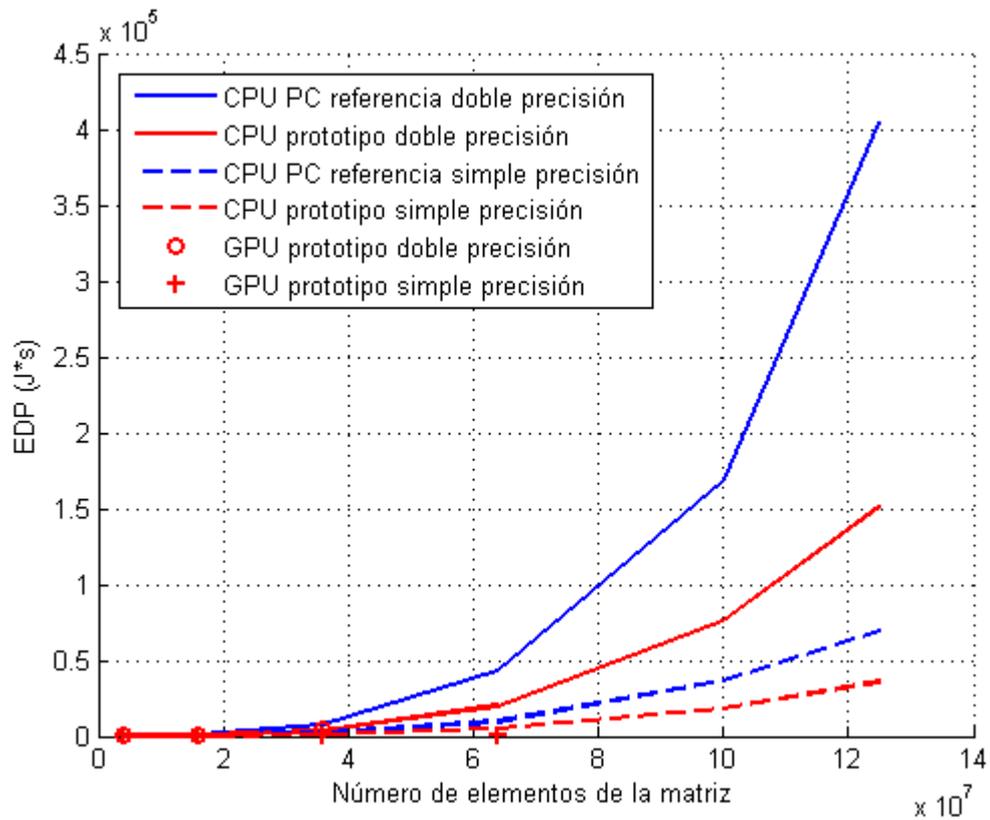


Figura 27. Comparación de la eficiencia (EDP) al realizar el producto de matrices.

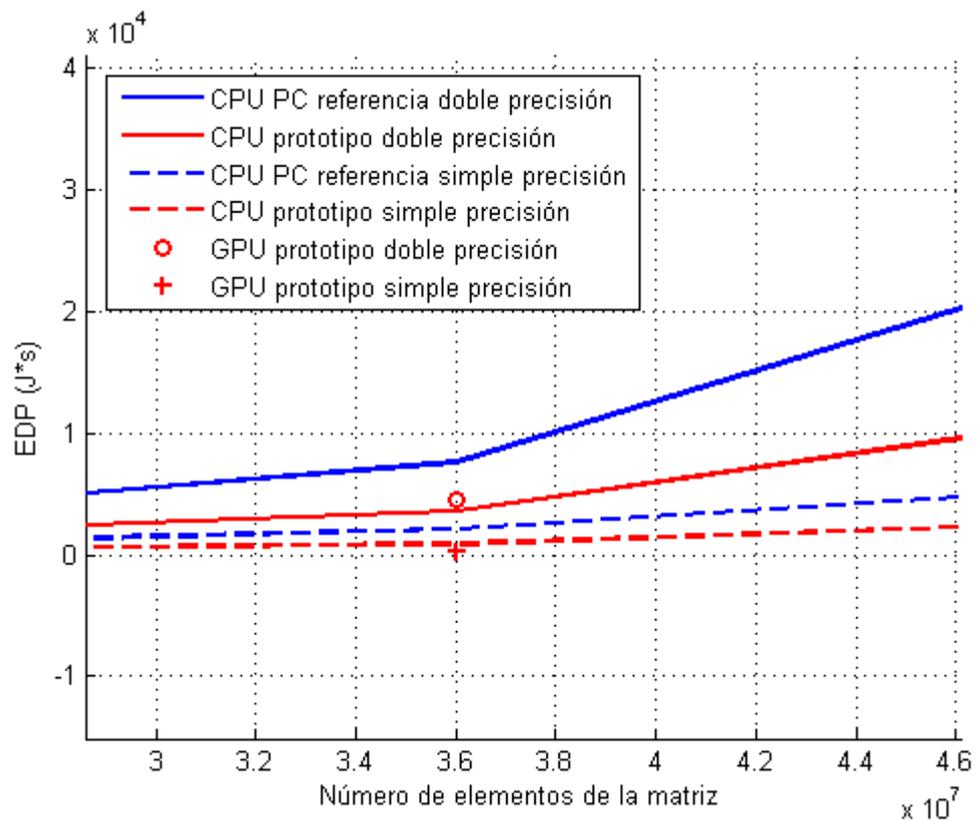


Figura 28. Vista en detalle de la eficiencia (EDP) al realizar el producto de matrices.

La tercera prueba que se ha realizado con MatLab, ha sido mediante el uso de la FFT (Transformada rápida de Fourier). Para esta prueba, se ha creado un script que aplica la FFT a un vector generado de forma aleatoria, en donde, se han capturado los amperios y el tiempo que se invierte en obtener la FFT de dicho vector.

Como sucedía con la segunda prueba, los experimentos se han realizado para datos de simple y doble precisión. Además para aquellas pruebas, que sólo requieran del uso de la CPU en el prototipo, se han realizado con la GPU desconectada.

Los gráficos que muestran los resultados del experimento de utilizar la FFT en MatLab se muestran en las figuras 29 y 30. En estos gráficos, el eje de abscisas representa siempre el número de elementos del vector de entrada, mientras que el eje de ordenadas, o bien representa la energía consumida para realizar la FFT o el EDP requerido.

La figura 29, muestra la energía requerida al aplicar la FFT a vectores con distintos tamaños. Lo primero que se puede observar, es que el prototipo trabajando con la CPU, tanto para datos de simple como de doble precisión, ha tenido el menor consumo energético que la CPU del PC de referencia. Por otro lado, en la misma figura, se observa que la GPU en este experimento ha tenido el mayor consumo energético.

La figura 30, informa sobre la eficiencia energética al aplicar la FFT sobre vectores de diferentes tamaños. En ella, se puede observar que lo más eficiente energéticamente, ha sido trabajar con la CPU del prototipo y datos de entrada de simple precisión.

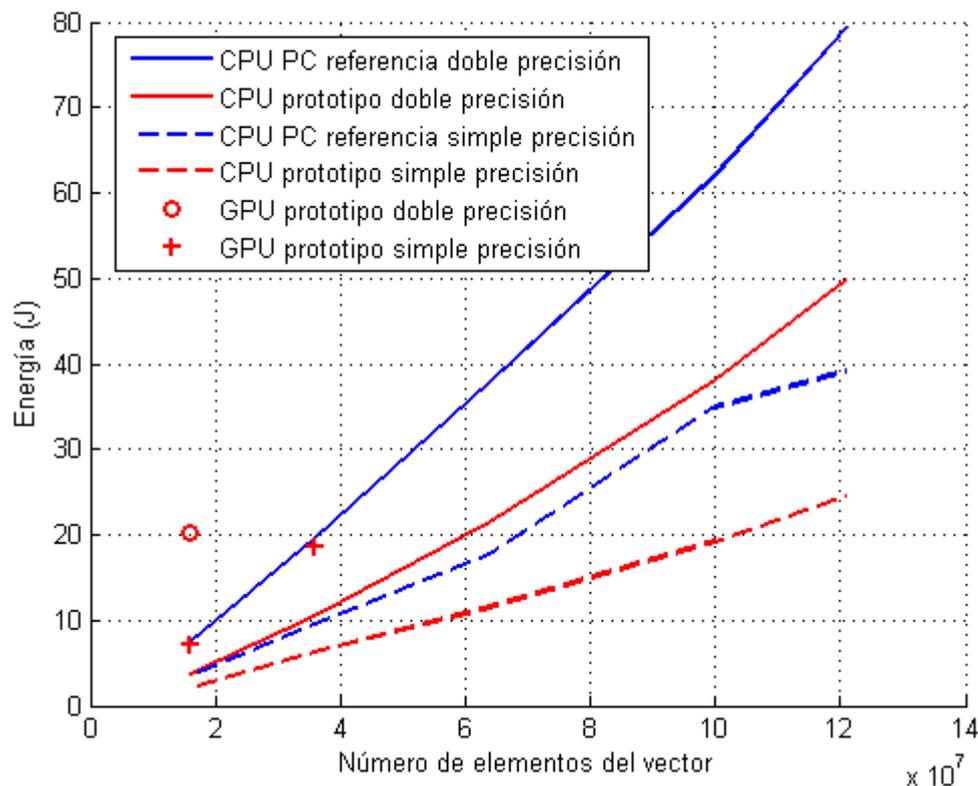


Figura 29. Comparación del consumo energético al aplicar la FFT.

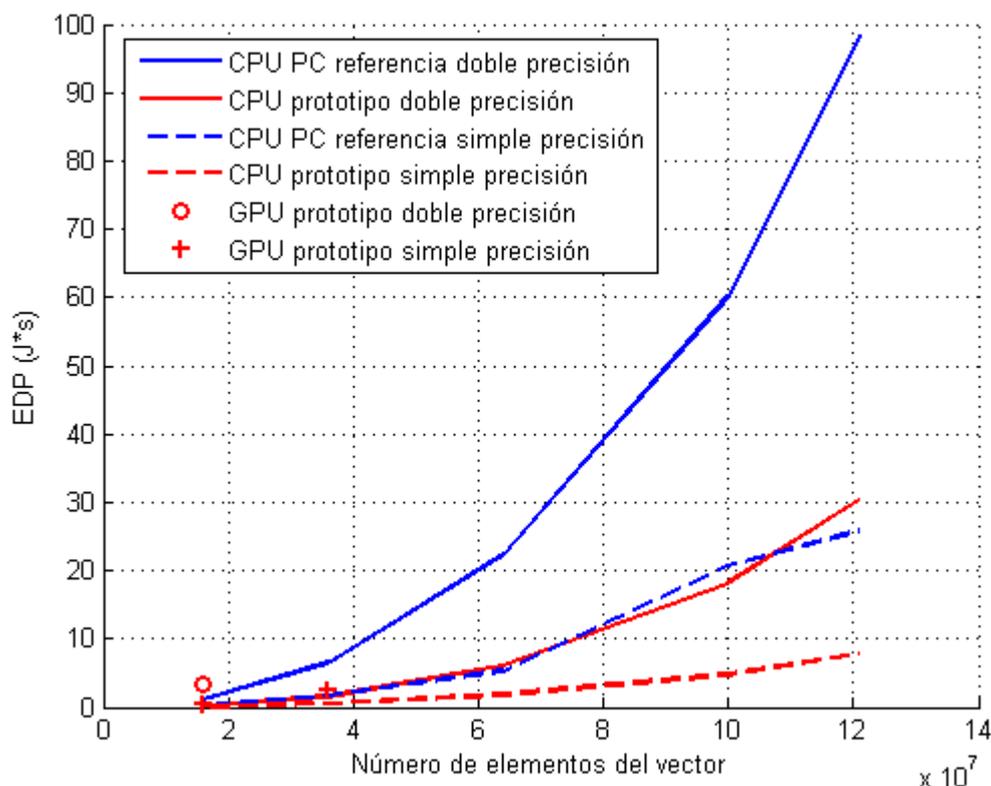


Figura 30. Comparación de la eficiencia (EDP) al aplicar la FFT.

5.3.3. LINPACK

Para obtener los resultados mediante el benchmark Linpack, lo primero que se ha realizado ha sido modificar el código original del benchmark, ya que este sólo se ejecutaba en un solo hilo de la CPU.

Una vez modificado el código se ha creado un script para poder lanzar el programa en diferentes hilos a la vez, hasta ocupar los cuatro hilos disponibles de ejecución en el prototipo y en el PC de referencia.

Para compilar el código se ha realizado con las siguientes opciones de compilación: `icc -O2 -xavx -xhost linpack.c -o linpack -lm`. En donde: `-xavx` sirve para poner el flag de instrucciones `avx` activo, y `-xhost` sirve para comprobar si la máquina de verdad tiene `avx` u otras opciones mejores de compilación. Además, los resultados que se han obtenido en el prototipo, para realizar una comparación más justa contra el PC de referencia, se han realizado sin la GPU conectada.

En los siguientes gráficos, se muestra un promedio de todos los resultados capturados que se pueden ver en el ANEXO VIII.

Las figuras 31 y 32, muestran los resultados obtenidos al ejecutar el benchmark de Linpack para una matriz de entrada de diferente tamaño. Primero se ha realizado para un tamaño de 200x200 (figura 31) y por último se ha realizado para una matriz de tamaño 2048x2048 (figura 32). En ambos gráficos, el eje de abscisas representa siempre el número de hilos/cores que se han utilizado para la prueba, mientras que el eje de ordenadas, refleja la eficiencia energética obtenida tras la prueba realizada.

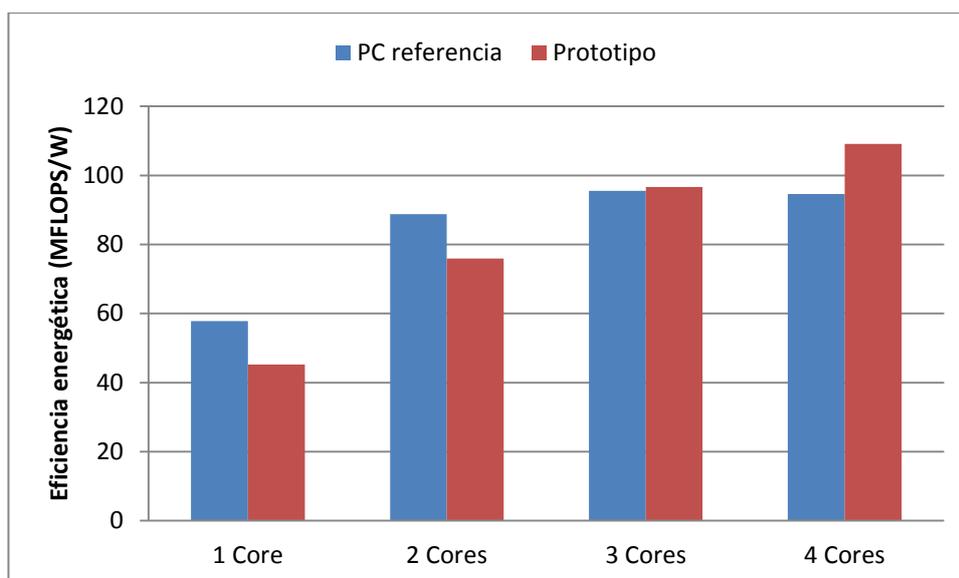


Figura 31. Eficiencia energética medida mediante Linpack para un problema de tamaño 200x200.

De la figura 31, se puede observar que el PC de referencia, ha sido más eficiente cuando ha utilizado 1 y 2 cores para la prueba. Mientras que cuando la misma prueba, se ha realizado utilizando 3 y 4 cores, el PC que ha resultado ser más eficiente ha sido el prototipo. Cuando se ejecuta la prueba para 3 y 4 cores en el PC de referencia, no se aprecia una mejoría en la eficiencia. Esto se debe a que Linpack logra acercarse al pico teórico de rendimiento y apenas se beneficia de la tecnología HyperThreading de Intel. Por último, cabe destacar que la configuración más eficiente ha sido utilizar los 4 cores disponibles en el prototipo.

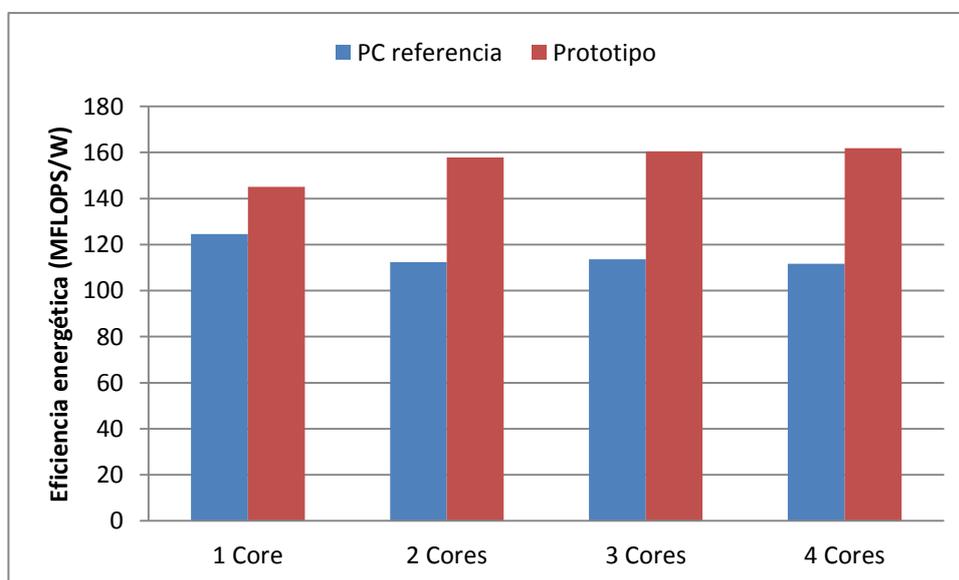


Figura 32. Eficiencia energética medida mediante Linpack para un problema de tamaño 2048x2048.

En la figura 32, se observa que el prototipo en todos los experimentos ha sido más eficiente que el PC de referencia. La configuración más eficiente energéticamente hablando, para la prueba de Linpack con un tamaño de matriz de 2048x2048 ha sido utilizar el prototipo con 4 hilos en ejecución.

5.3.4. PHORONIX TEST SUITE

En los siguientes gráficos (figuras 33, 34, 35, 36, 37, 38 y 39) aparecen recogidos los resultados obtenidos una vez que se han ejecutado los diferentes benchmarks dentro del paquete Phoronix. Para realizar estas pruebas, se ha dejado conectada la GPU al prototipo, ya que existen tests que requieren de aceleración gráfica. Las tablas con los resultados capturados, se puede ver en el ANEXO IX.

La figura 33, muestra los resultados obtenidos al ejecutar los benchmarks “FFTE”, “Himeno Benchmark”, “Java SciMark” y “SciMark”. En dicha figura, el eje de abscisas muestra el benchmark que se ha ejecutado, mientras que el eje de ordenadas refleja la eficiencia energética que ofrecen los computadores.

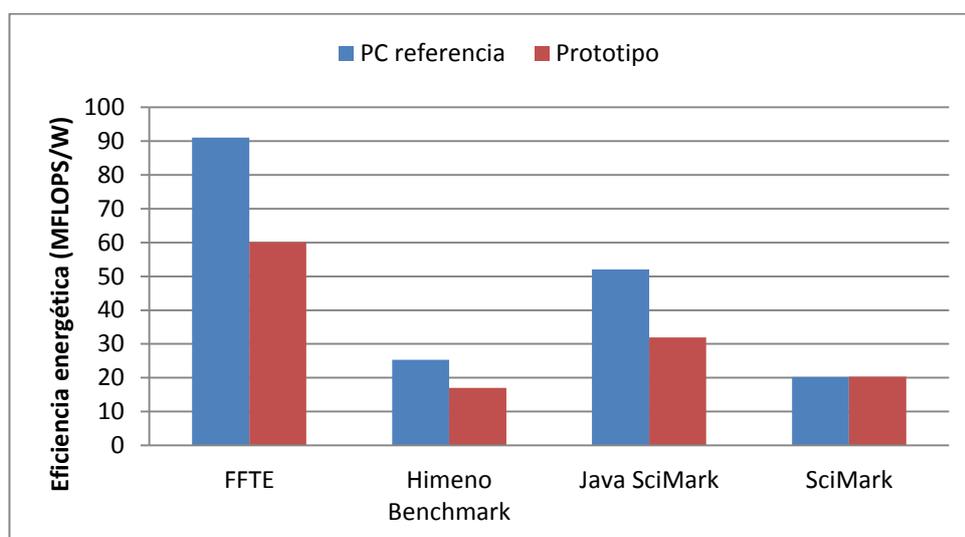


Figura 33. Eficiencia energética medida mediante los benchmarks de Phoronix.

De la figura 33, se puede observar, que todos los benchmarks probados, a excepción de “SciMark”, han sido más eficientes energéticamente en el PC de referencia que en el prototipo.

Las figuras 34, 35, 36, 37, 38 y 39, muestran los resultados obtenidos de aquellos benchmarks dentro del paquete Phoronix, que no devuelven una unidad tipo MFLOPS al ejecutarlos, como sucedía en los benchmarks de la figura 33. En estas figuras, se representan, primero el tiempo de ejecución requerido para los diferentes tests (figuras 34 y 35), la energía requerida por cada benchmark para realizar las pruebas (figuras 36 y 37) y por último, se muestra la comparación de la eficiencia energética en las figuras 38 y 39.

En la figura 34, se muestran los tiempos requeridos para los benchmarks “C-Ray”, “Crafty”, “Minion”, “Timed MrBayes Analysis”. En dicha figura, se muestra que para esos cuatro benchmarks, el prototipo ha tardado menos tiempo que el PC de referencia. Además, los benchmarks que también se muestran en la figura 35, han requerido menos tiempo en el prototipo que en el PC de referencia.

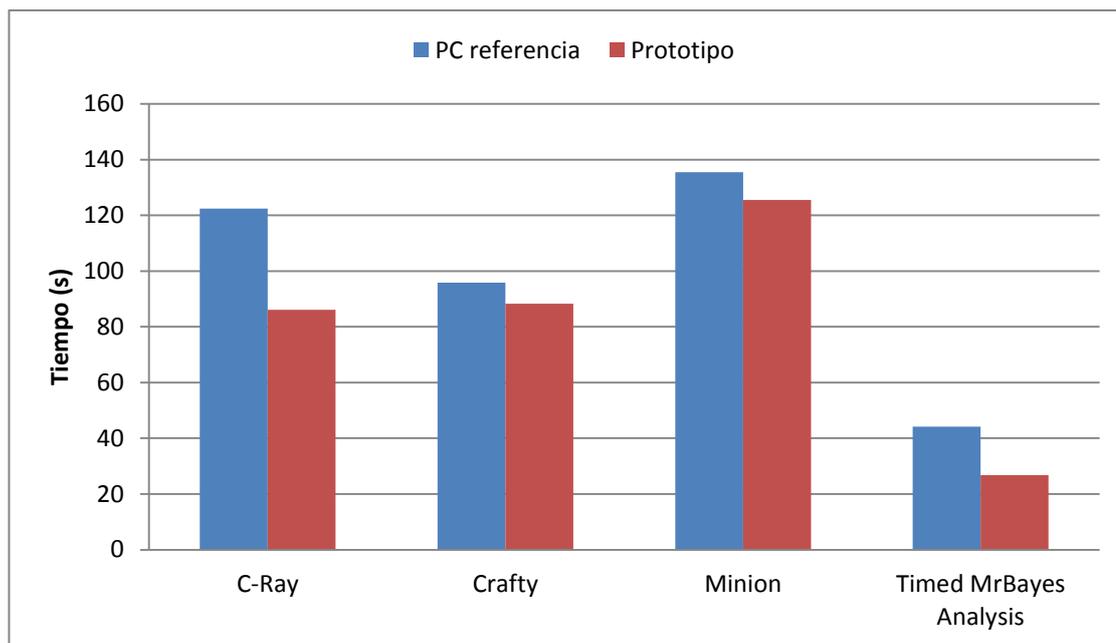


Figura 34. Comparación Phoronix tiempo (Segundos) parte 1.

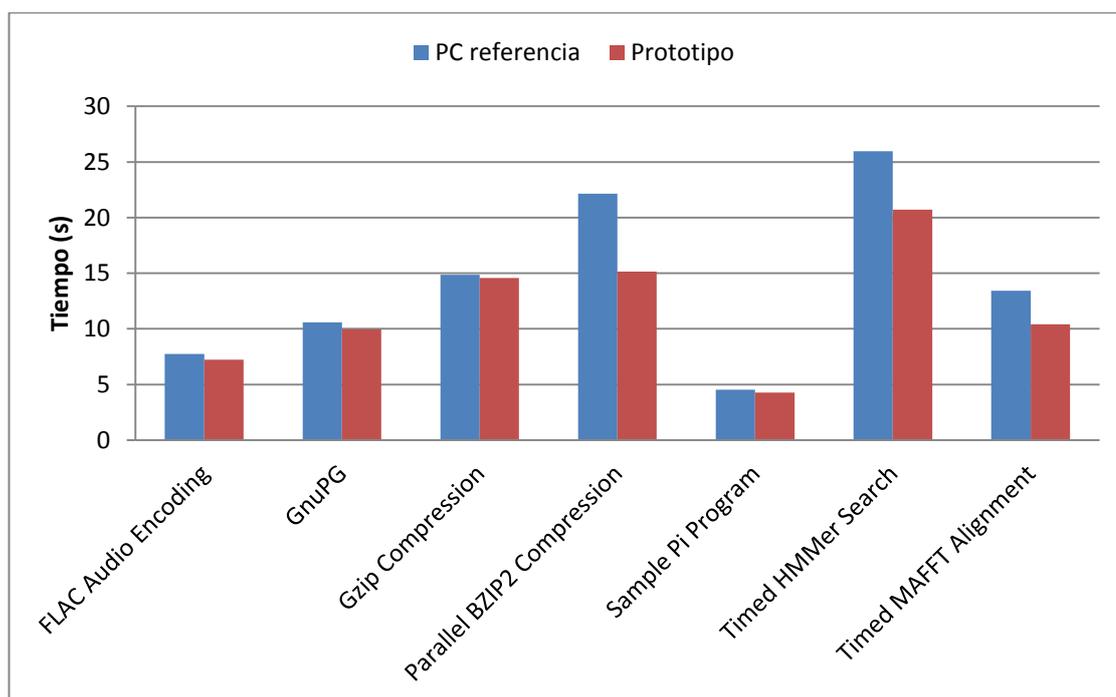


Figura 35. Comparación Phoronix tiempo (Segundos) parte 2.

De las figuras 36 y 37, se observa, que todos los benchmark que aparecen en el gráfico a excepción de “Parallel BZIP2 Compression” y de “Timed MrBayes Analysis” han tenido un menor consumo de energía en el PC de referencia que en el prototipo.

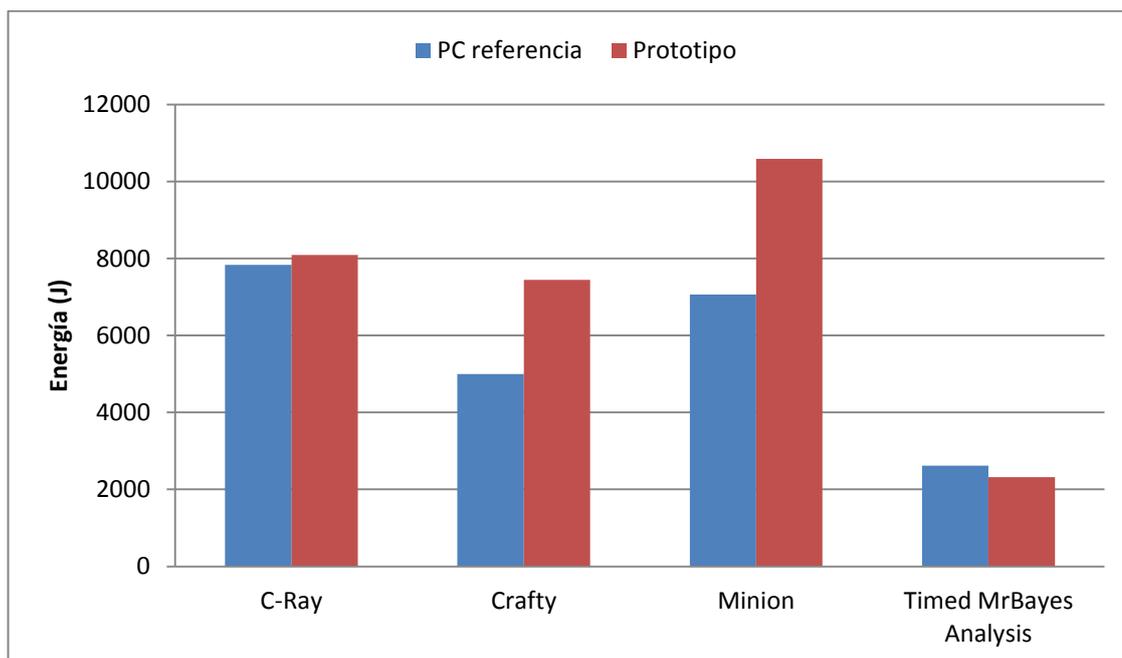


Figura 36. Comparación Phoronix energía parte 1.

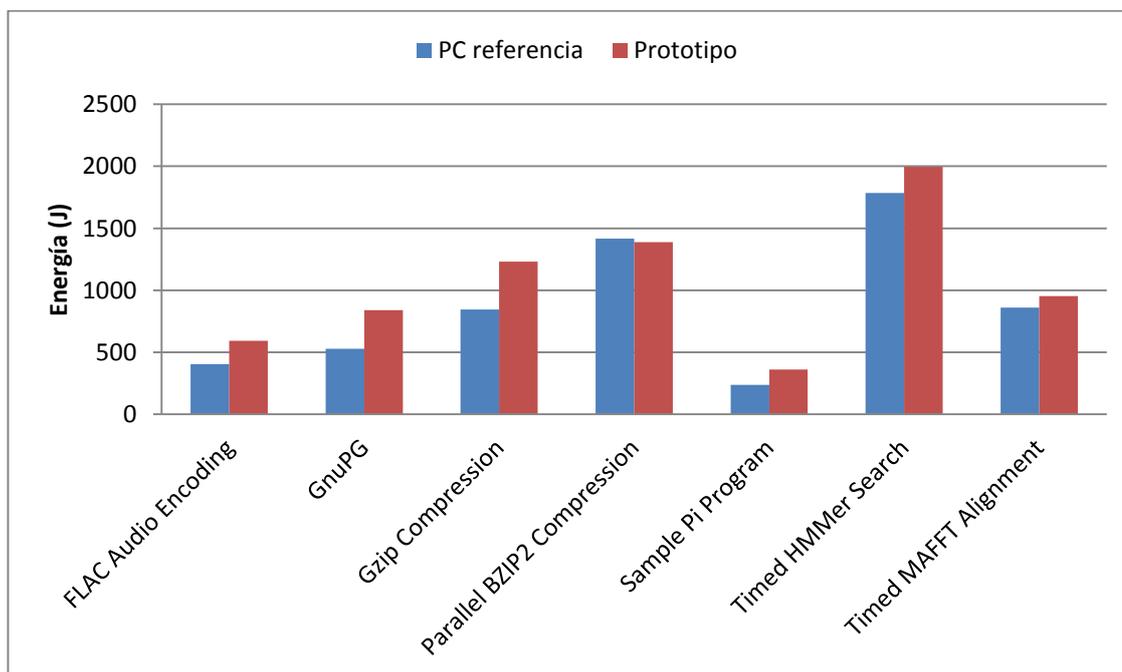


Figura 37. Comparación Phoronix energía parte 2.

Por último, de las figuras 38 y 39, se puede observar que en pruebas como, “C-ray”, “Timed MrBayes Analysis”, “Parallel BZIP2 Compression”, “Timed HMMer Search” y “Timed MAFFT Alignment”, han sido más eficiente energéticamente hablando el prototipo que el PC de referencia.

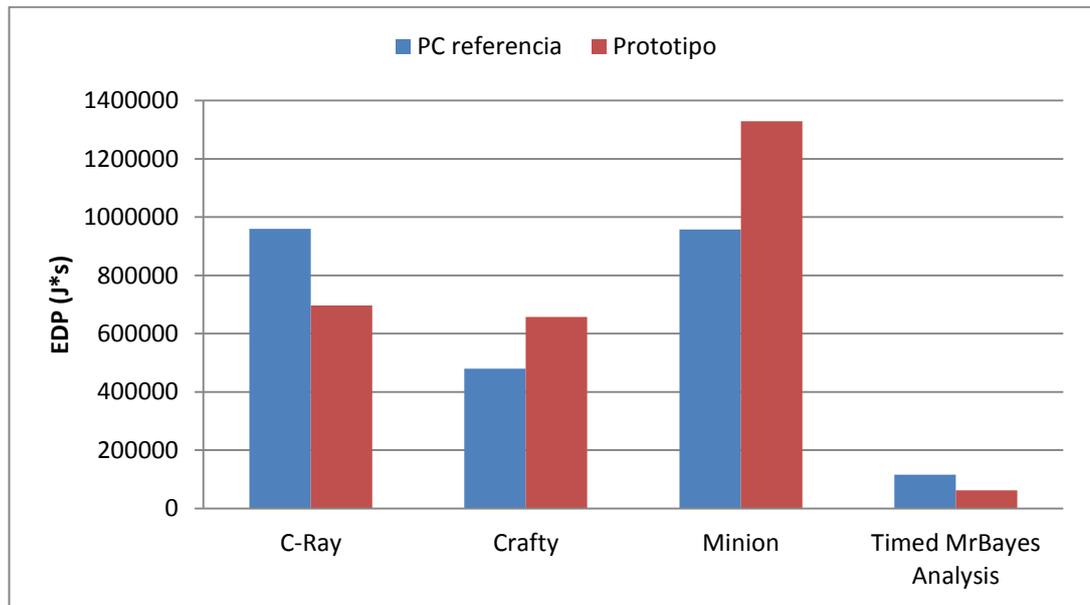


Figura 38. Comparación Phoronix eficiencia energética parte 1.

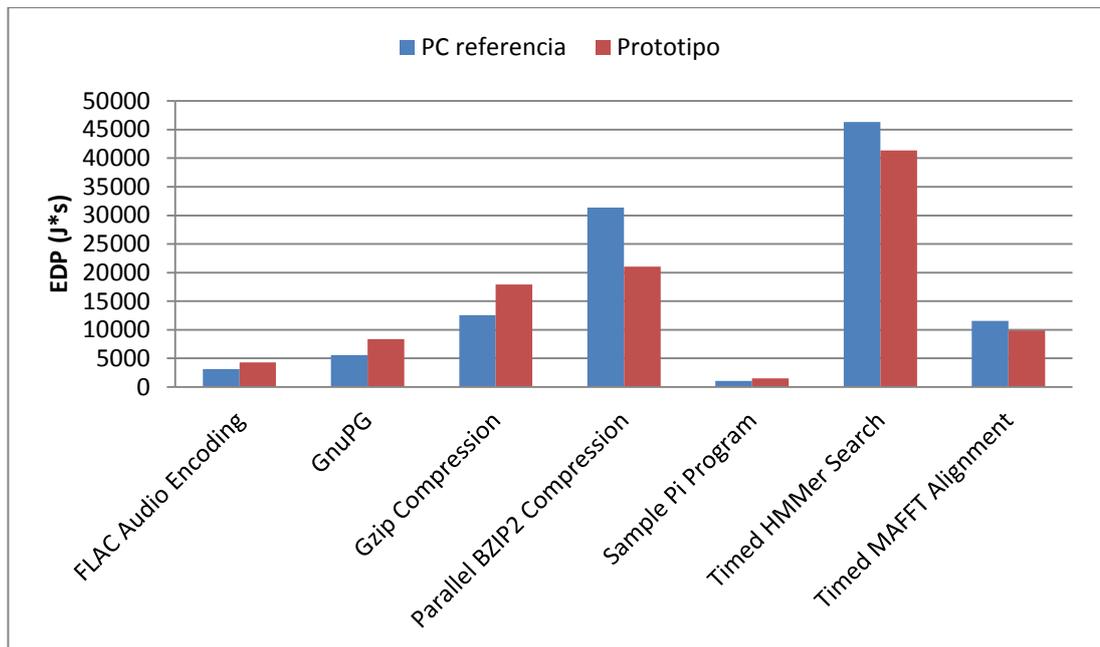


Figura 39. Comparación Phoronix eficiencia energética parte 2.

5.3.5. CÓDIGOS DEL ÁREA DE TSC DE LA UNIVERSIDAD DE OVIEDO

Para obtener los resultados, lo primero que se ha realizado ha sido instalar las librerías MPI necesarias para el funcionamiento correcto de algunos códigos. Además, para el caso del código FMM que precisa de la FFT se requiere de la instalación de la librería FFTW 3.3.1. Una vez preparado el entorno de trabajo, tanto en el PC de referencia como en el prototipo, se han ejecutado los códigos obteniendo los resultados que se pueden ver en el ANEXO X. Se debe saber del estudio realizado, que para aquellas mediciones, en donde sólo se requiera el uso de la CPU en el prototipo, se han realizado con la GPU desconectada.

Las gráficas están agrupadas por las diferentes pruebas realizadas, es decir, las figuras 40, 41 y 42 recogen el tiempo (en segundos), la energía requerida (en Kilo Julios) y la eficiencia energética (en Mega Julios * segundos) para el código MST. Mientras que las figuras 43, 44, y 45 recogen la misma información para el código FMM sin FFT. Por último, las figuras 46, 47 y 48 también muestran dicha información, pero para el código FMM que utiliza la FFT para agilizar el proceso.

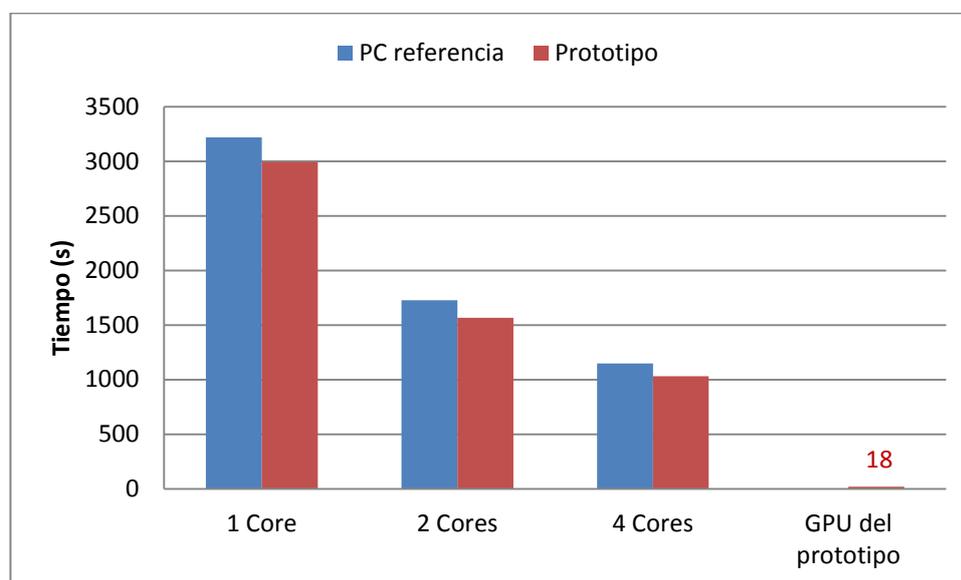


Figura 40. Tiempo requerido para el código MST.

De la gráfica 40, se observa, que para resolver el código MST cuando se ejecuta en su versión de GPU en el prototipo, apenas ha tardado unos pocos segundos. Por otro lado, en la misma gráfica se puede observar que siempre el prototipo ha sido más rápido en su versión de CPU que el PC de referencia.

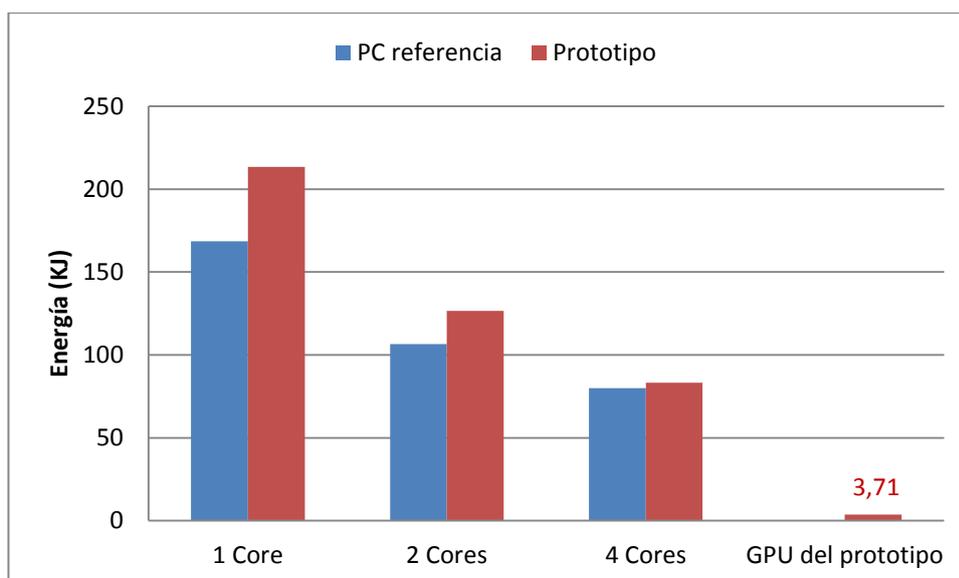


Figura 41. Consumo de energía para el código MST.

En la figura 41, se puede observar la energía que requieren ambos computadores para realizar las simulaciones basadas en el problema de MST. Se puede observar que el PC de referencia ha consumido menos que el prototipo al resolver el código en su versión de CPU. Por último, se observa, que el menor consumo para resolver el código de MST viene dado cuando se ha utilizado la GPU del prototipo.

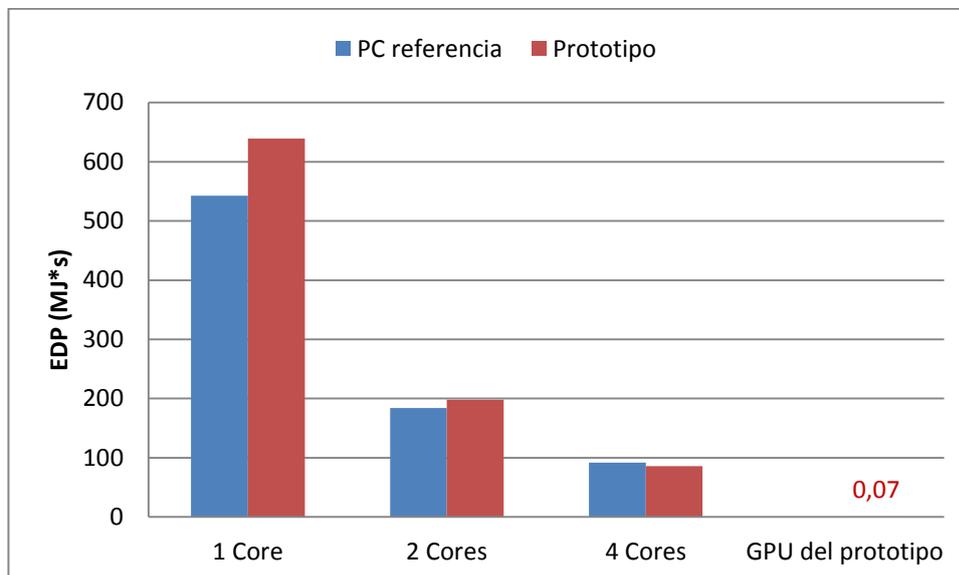


Figura 42. Eficiencia energética para el código MST.

En la figura 42, se muestra la eficiencia energética para el código de MST. En ella, se observa que al resolver el código utilizando la GPU se ha tenido el mejor resultado energéticamente hablando. Mientras que si se resuelve el código con la CPU, el prototipo utilizando 4 cores ha sido la opción más eficiente.

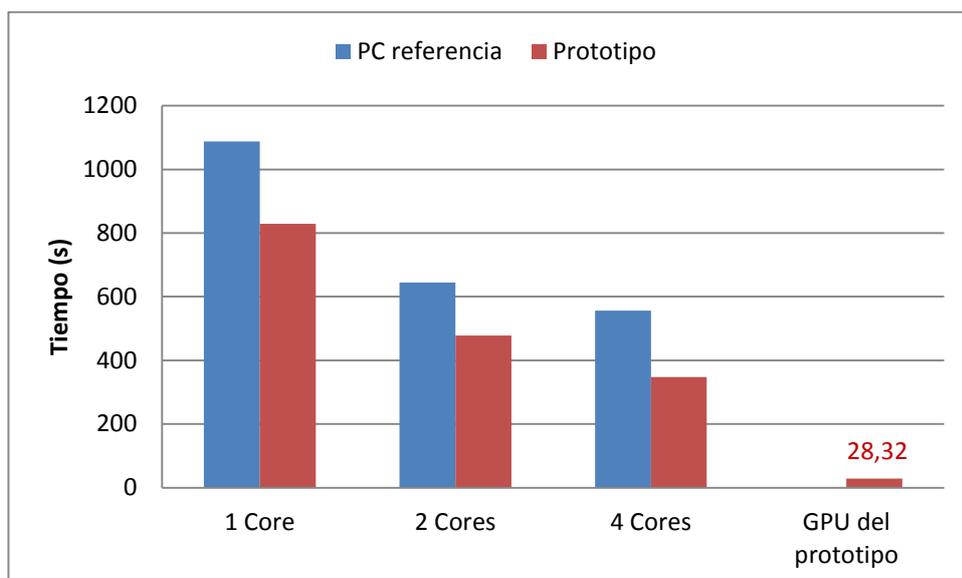


Figura 43. Tiempo requerido para el código FMM.

La figura 43 muestra el tiempo que se requiere para resolver el código FMM sin el uso de la FFT. En ella, se puede observar, que la GPU ha sido la opción más rápida con bastante diferencia respecto a resolver el código con la CPU. Además, también, el prototipo en su versión de CPU siempre ha resuelto el código más rápido que el PC de referencia para los distintos escenarios.

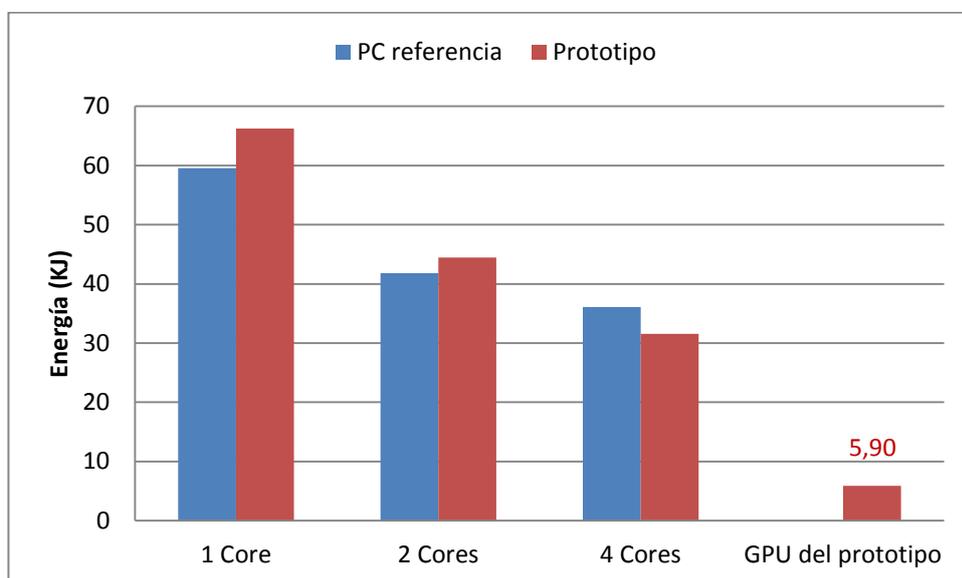


Figura 44. Consumo de energía para el código FMM.

La figura 44 muestra el consumo de energía que requiere el código FMM sin el uso de la FFT en los dos computadores. La gráfica muestra que para 1 y 2 cores, el PC de referencia ha consumido menos que el prototipo, sin embargo, cuando se han utilizado los 4 cores el prototipo ha consumido menos energía. Por último, se sigue observando que la opción que menos energía ha requerido para resolver este código se ha producido por el uso de la GPU.

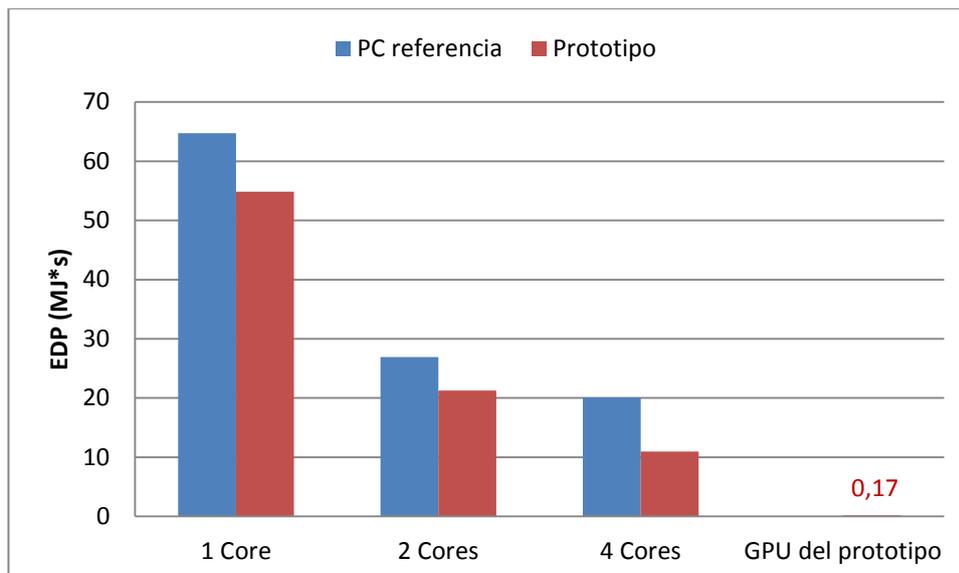


Figura 45. Eficiencia energética para el código FMM.

La figura 45 muestra la eficiencia energética para el código FMM sin el uso de la FFT. En ella, se observa que la GPU ha sido la opción más eficiente, como se podía haber previsto. Ya que ha sido la que menos ha consumido y además ha sido la que menos tiempo ha necesitado para resolver el problema. Por otro lado, la gráfica muestra la eficiencia también de la CPU al resolver el problema de FMM. Se observa, que la opción en cuanto a CPU más eficiente ha sido resolver el problema utilizando los 4 cores del prototipo.

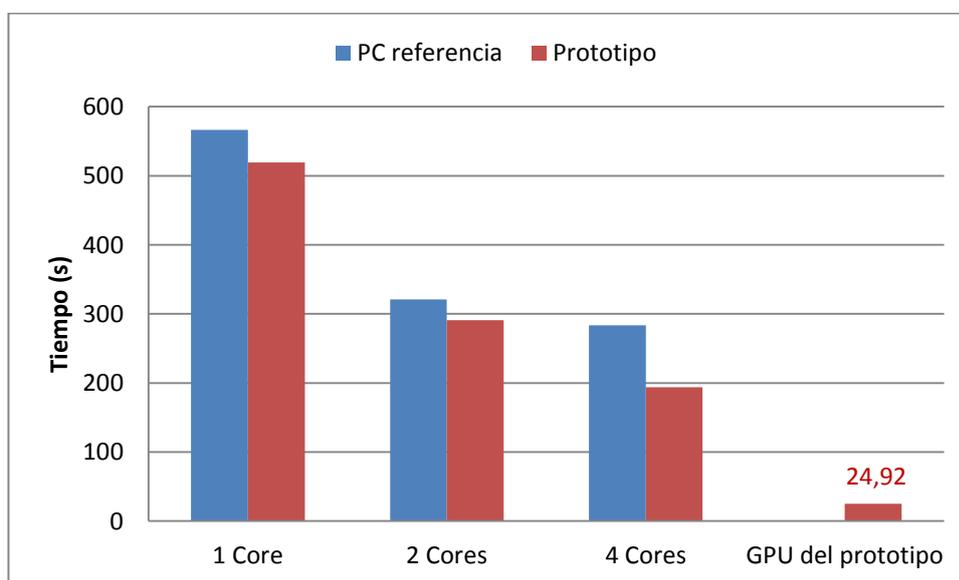


Figura 46. Tiempo requerido para el código FMM con FFT.

La figura 46, muestra el tiempo que se ha necesitado para ejecutar en ambos computadores el código FMM en su versión con FFT. En dicha gráfica, se observa que la GPU es la opción que requiere menos tiempo para resolver el problema. Además, se observa que al utilizar la CPU para resolver este código, se requiere menos tiempo con el prototipo que con el PC de referencia.

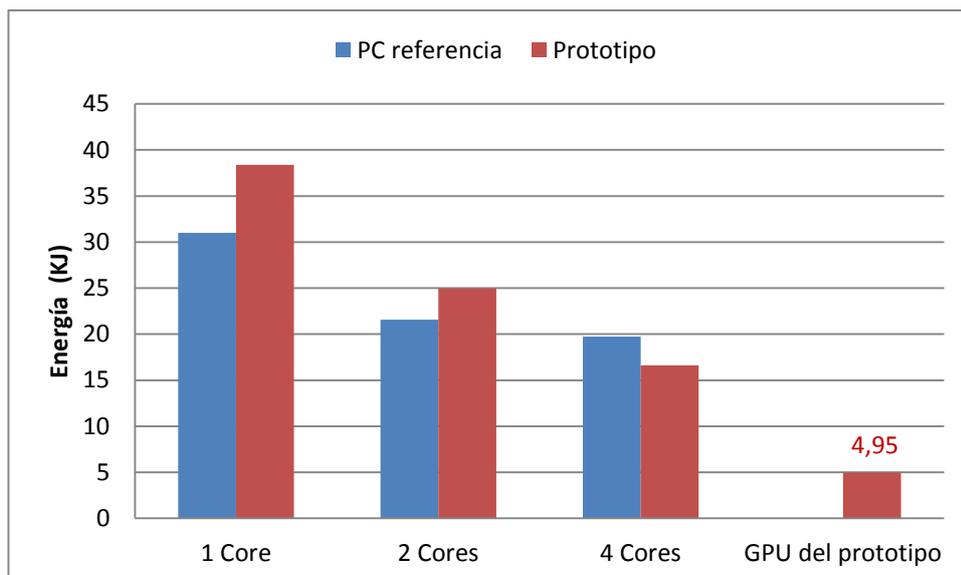


Figura 47. Consumo de energía para el código FMM con FFT.

La gráfica 47, muestra la energía requerida para resolver el problema FMM con FFT en ambos computadores. En dicha gráfica, la GPU sigue siendo la opción que menos energía ha requerido como se venía viendo hasta ahora. En el caso de resolver el problema utilizando la CPU, el resultado que menos consumo ha requerido viene dado de utilizar el prototipo con sus 4 cores.

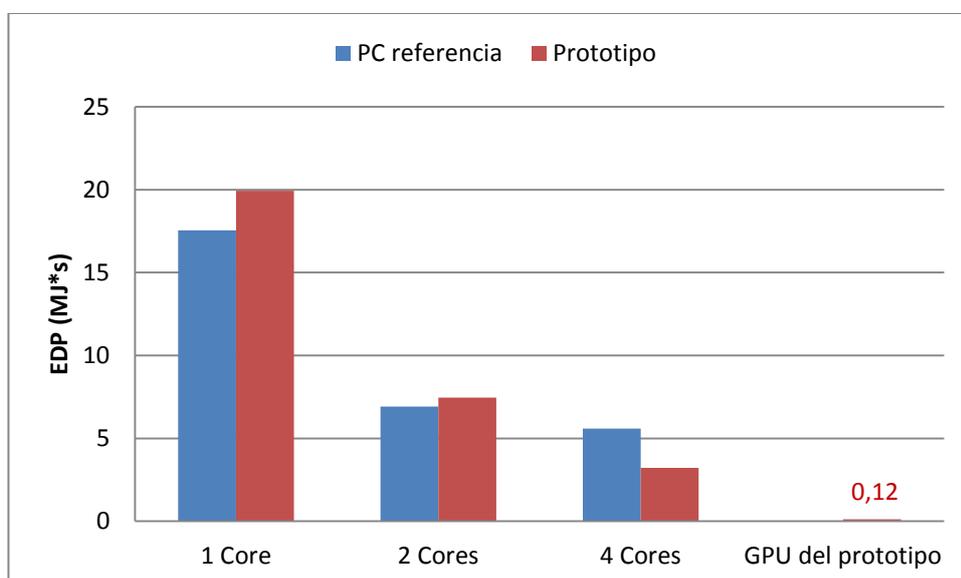


Figura 48. Eficiencia energética para el código FMM con FFT.

En la figura 48, se muestra la eficiencia energética que produce el código FMM con FFT. En ella se observa que la GPU vuelve a ser la herramienta más eficiente a la hora de resolver el problema. Además, el prototipo utilizando 4 cores, ha sido la opción más eficiente a la hora de resolver el código FMM con FFT utilizando la CPU.

6. CONCLUSIONES

El principal objetivo de esta Tesis de Máster era diseñar y probar el rendimiento de un prototipo de computador eficiente. A tal fin, se han llevado a cabo una serie de pasos que se detallan a continuación.

En la primera fase de la Tesis, se ha creado una base de conocimiento. Para ello, se realizó un proceso de recopilación de información sobre el hardware que se puede encontrar en la actualidad. Se continuó estableciendo una métrica teórica para seleccionar los componentes que, a priori, fueran los más eficientes.

Una vez seleccionados los componentes, se ha montado el equipo y se han realizado las diferentes pruebas, cuyas conclusiones se resumen a continuación.

Antes de centrarse en las conclusiones obtenidas, se debe tener en cuenta que las medidas que se han capturado al realizar cada prueba son medidas pico. Es decir, siempre se ha capturado el valor máximo de corriente que ha requerido el computador para realizar el experimento, debido a que no se disponía de una herramienta capaz de capturar una muestra de medidas de los consumos instantáneos para luego realizar el promedio.

La primera conclusión que se ha obtenido ha sido que, si se quiere trabajar de forma eficiente con un computador y no se pretende hacer uso de la GPU, es muy conveniente que la GPU esté físicamente desconectada. Una GPU que está conectada a una estación de trabajo eficiente pero que permanece inactiva consume energía innecesariamente, lo que puede afectar de manera importante a su eficiencia energética.

Al evaluar la CPU con el benchmark de Whetstone, se ha observado que el PC de referencia era más eficiente que el prototipo. Este resultado es debido a que, tal y como se comentó anteriormente, el prototipo tiene una fuente dimensionada para soportar un consumo mucho mayor del que se requiere en la prueba, lo que provoca que la fuente de alimentación trabaje en una zona en la que el fabricante no garantiza una alta eficiencia energética. Sin embargo, el PC de referencia dispone de una fuente de alimentación de mucha menor potencia, por lo que trabaja en una zona en la que su rendimiento, garantizado por el fabricante, está próximo al máximo. Por ello, parece que es de suma importancia conocer la carga de la fuente para poder dimensionarla correctamente con respecto al hardware adquirido. Asimismo, resulta de interés que la carga de la estación de trabajo no resulte excesivamente baja (lo que se conoce como estado de reposo o “idle”), ya que el prototipo que se ha diseñado para la Tesis obtiene unos resultados peores cuando menor es la carga computacional que se ejecuta en él.

Por su parte, las pruebas realizadas con MatLab y con los códigos del Área de TSC se ajustan mejor a la carga de trabajo esperada en un computador que ha sido especialmente diseñado para ejecutar software de cálculo científico. Estas pruebas suponen una mayor carga de trabajo, lo que permite alcanzar el punto de mínima eficiencia garantizado por la fuente, especialmente cuando se utiliza la GPU. Al comparar los resultados entre MatLab y los códigos desarrollados por varios miembros del Área de TSC, se puede afirmar que el computador diseñado para trabajar conjuntamente con una CPU más una GPU es la opción más eficiente. Otra conclusión especialmente relevante, es que se ha podido comprobar que los códigos programados en el lenguaje nativo de la tarjeta aceleradora (códigos del

Área de TSC) son notablemente más eficientes que aquellos códigos implementados en un lenguaje de muy alto nivel como MatLab. Es decir, en términos de eficiencia computacional y energética, el tiempo dedicado a programar los códigos en un lenguaje especializado para tarjetas aceleradoras se ve recompensado.

Mediante las pruebas realizadas con Linpack, se pudo observar que si el código que se quiere ejecutar está muy optimizado, resulta más interesante que la CPU tenga núcleos físicos en vez de núcleos virtuales. Al realizar las diferentes pruebas con este benchmark, se vio que la tecnología Hyper-Threading de Intel apenas consiguió mejorar la eficiencia en el PC de referencia.

Por último, estarían los resultados obtenidos con Phoronix, en donde se observó variedad en el rendimiento. Esto es debido a que esta suite, contiene múltiples pruebas de software. En algunos casos, cuando la prueba requería de aceleración 3D, se pudo ver un rendimiento mucho mayor con el prototipo. Nuevamente, se confirma que la estación de trabajo es especialmente eficientemente cuando se hace uso de la combinación de CPU más GPU.

Para finalizar, se puede concluir que gracias al estudio realizado se han descubierto diversos factores a tener en cuenta a la hora de diseñar una estación de trabajo eficiente, que inicialmente eran desconocidos para el autor y sus directores. Por lo tanto, el estudio ha cumplido con las expectativas, logrando una estación de trabajo que se ha validado como una herramienta eficiente al combinar adecuadamente el uso de la CPU más la GPU.

La continuación natural y futura línea de esta Tesis consistiría en utilizar los conocimientos adquiridos a lo largo del presente estudio para diseñar los servidores o estaciones GPGPU (del inglés, General-Purpose Computing on Graphics Processing Units) a emplear en el Área de TSC. Así, se pretende mejorar la eficiencia (computacional y energética) de los recursos informáticos empleados en el ámbito de la resolución de problemas de dispersión.

BIBLIOGRAFÍA

- [1] Consumo eléctrico. Recurso web disponible. <http://www.economist.com/node/12411920>. [Febrero 2012].
- [2] J. Garbajosa, E. Huedo and M. López. “Tecnologías para la eficiencia energética en los sistemas TI”. Universidad Politécnica de Madrid. 2008.
- [3] H. Lin and W. Fend. “A first look at integrated GPUs for Green High-Performance Computing”, T.R.W. Scogland, Dept. of Computer Science Virginia Tech.
- [4] Recurso web disponible. http://www.slideshare.net/daniel_b4e/la-evolucion-del-sistema-que-est-cambiando-el-mundo-hpc-high-performance-computing. [Diciembre 2011].
- [5] M. G. Araújo, J. M. Taboada, F. Obelleiro, J. M. Bértolo, L. Landesa, J. Rivero, and J. L. Rodríguez, “Supercomputer aware approach for the solution of challenging electromagnetic problems”, *Progress In Electromagnetics Research*, Vol. 101, 241– 256, 2010.
- [6] Jesús A. López-Fernández, Miguel López-Portugués, Yuri Álvarez, Cebrián García, David Martínez-Álvarez and Fernando Las-Heras, “Fast Antenna Characterization using the Sources Reconstruction Method on Graphics Processors”, *Progress in Electromagnetic Research (PIER)*, vol. 126, pp. 185–201, 2012.
- [7] G.S. Almasi and A. Gottlieb. “Highly Parallel Computing”. Benjamin-Cummings publishers, Redwood city, CA, 1989.
- [8] Recurso web disponible. http://es.wikipedia.org/wiki/Condici%C3%B3n_de_carrera. [Diciembre 2011].
- [9] S. Iturriaga, P. Maya, D. Puntos and S. Nesmachnow. “Proyecto Fetón - Cluster de computadores de Alto rendimiento. Desempeño con Acceso Remoto”, 2008.
- [10] Amdahl, G.M. Validity of the single-processor approach to achieving large scale computing capabilities. In *AFIPS Conference Proceedings* vol. 30 (Atlantic City, N.J., Apr. 18-20). AFIPS Press, Reston, Va., 1967, pp. 483-485.
- [11] M. Rossainz. “Apuntes del curso programación concurrente y paralela”, Universidad Autónoma de Puebla, 2011.
- [12] Definición de arquitectura pipeline. Recurso web disponible. <http://fermat.usach.cl/~msanchez/lcc/Arquitectura-Pipeline.pdf>. [Enero 2012].
- [13] Definición de CPU. Recurso web disponible. http://es.wikipedia.org/wiki/Unidad_central_de_procesamiento. [Enero 2012].
- [14] Noticia: Core i7 to make leap to Xeon in early 2009. Recurso web disponible. <http://legacy.macnn.com/articles/08/11/13/xeon.i7.in.early.2009/>. [Febrero 2012].
- [15] Intel, “Intel Core i7 Processor Series and Extreme edition Series Datasheet”, vol.2, 2009.
- [16] Arquitectura Nehalem. Recurso web disponible. <http://www.madboxpc.com/review-roundup-intel-nehalem-core-i7-920-940-965-extreme>. [Febrero 2012].
- [17] L. Marín, “Arquitectura Nehalem: La arquitectura de próxima generación”, 2008.
- [18] Explicación de la tecnología SpeedStep. Recurso web disponible. <http://www.intel.com/support/sp/processors/sb/cs-028855.htm>. [Febrero 2012].
- [19] A. Díaz. “Intel Para Laptops”, Universidad de Las Palmas de Gran Canaria, 2011.
- [20] Sandy Bridge. Recurso web disponible. http://es.wikipedia.org/wiki/Sandy_Bridge. [Febrero 2012].
- [21] About computadoras. Recurso web disponible. <http://computadoras.about.com/od/conocer-mi-computadora/a/que-Es-Sandy-Bridge.htm>. [Febrero 2012].
- [22] Noticias Hardware. Recurso web disponible. <http://noticias-hardware.blogspot.com/2011/01/intel-sandy-bridge-introduccion-y.html>. [Febrero 2012].

- [23] Sandy Bridge. Recurso web disponible.
http://www.pactual.com/articulo/laboratorio/especiales/8231/conoce_arquitectura_sandy_bridge_inte1_fondo.html. [Febrero 2012].
- [24] Turbo Boost 2.0. Recurso web disponible. <http://www.intel.com/espanol/technology/turboboost/>. [Febrero 2012].
- [25] AMD. “Guide to AMD Desktop Processor for Personal Computers”, 2011.
- [26] Noticia sobre los procesadores AMD. Recurso web disponible.
<http://www.muycomputer.com/2011/07/21/amd-matara-sus-cpus-athlon-ii-semprom-y-phenom-ii-en-2012>. [Febrero 2012].
- [27] AMD. “AMD Phenom II Shanghai. Perspectivas – ProfessionalSAT”. AMD Developers Blogs, 2008.
- [28] Documento sobre la arquitectura AMD. Recurso web disponible.
<http://www.amd.com/es/products/desktop/processors/phenom-ii/Pages/phenom-ii-key-architectural-features.aspx>. [Febrero 2012].
- [29] Tecnologías de AMD eficientes con la energía. Recurso web disponible.
<http://sites.amd.com/es/business/it-solutions/power-efficiency/Pages/power-efficiency.aspx>. [Febrero 2012].
- [30] PowerNow!. Recurso web disponible. <http://www.amd.com/us/products/technologies/amd-powernow-technology/pages/amd-powernow-technology.aspx>. [Febrero 2012].
- [31] Athlon II. Recurso web disponible. <http://www.amd.com/es/products/desktop/processors/athlon-ii-x2/Pages/amd-athlon-ii-x2-dual-core-processors-product-brief.aspx>. [Febrero 2012].
- [32] Características de Athlon II. Recurso web disponible.
<http://www.amd.com/es/products/desktop/processors/athlon-ii-x2/Pages/athlon-ii-key-features.aspx>. [Febrero 2012].
- [33] Mercado de la familia AMD. Recurso web disponible.
<http://shop.amd.com/es/Home/Search?SearchTerms=athlon>. [Febrero 2012].
- [34] Bulldozer micro arquitectura. Recurso web disponible.
http://es.wikipedia.org/wiki/Bulldozer_%28microarquitectura%29. [Febrero 2012].
- [35] Características de AMD FX. Recurso web disponible.
<http://www.amd.com/es/products/desktop/processors/amdfx/Pages/amdfx-key-architectural-features.aspx>. [Febrero 2012].
- [36] Familia AMD FX Bulldozer. Recurso web disponible.
<http://shop.amd.com/es/All/Search?SearchTerms=AMD%20FX>. [Febrero 2012].
- [37] ARM información. Recurso web disponible. http://es.wikipedia.org/wiki/Arquitectura_ARM. [Febrero 2012].
- [38] Procesadores ARM. Recurso web disponible.
<http://www.arm.com/products/processors/index.php>. [Febrero 2012].
- [39] ARM7. Recurso web disponible.
<http://www.arm.com/products/processors/classic/arm7/index.php>. [Febrero 2012].
- [40] ARM9. Recurso web disponible.
<http://www.arm.com/products/processors/classic/arm9/index.php>. [Febrero 2012].
- [41] ARM11. Recurso web disponible.
<http://www.arm.com/products/processors/classic/arm11/index.php>. [Febrero 2012].
- [42] Cortex-A. Recurso web disponible. <http://www.arm.com/products/processors/cortex-a/index.php>. [Febrero 2012].
- [43] Noticia Cortex-A15. Recurso web disponible. <http://www.xataka.com/componentes-de-pc/arm-cortex-a15-en-20-nanometros-para-finales-de-2012>. [Febrero 2012].

- [44] Cortex-R. Recurso web disponible. <http://www.arm.com/products/processors/cortex-r/index.php>. [Febrero 2012].
- [45] Cortex-M. Recurso web disponible. <http://www.arm.com/products/processors/cortex-m/cortex-m4-processor.php>. [Febrero 2012].
- [46] Sun SPARC. Recurso web disponible. http://es.wikipedia.org/wiki/Sun_SPARC. [Febrero 2012].
- [47] Lista de Green500 junio 2011. Recurso web disponible. <http://www.green500.org/lists/2011/06/top/list.php>. [Febrero 2012].
- [48] T. Maruyama. “SPARC64 VIIIfx: Fujitsu’s New Generation Octo Core Processor for PETA Scale Computing”, LSI Development Division Next Generation Technical Computing Unit Fujitsu Limited, 2009.
- [49] Mips technologies. Recurso web disponible. <http://mips.com/products/overview/>. [Febrero 2012].
- [50] Especificaciones MIPS64. Recurso web disponible. <http://www.mips.com/products/architectures/mips64/#specifications>. [Febrero 2012].
- [51] D. Sweetman “See MIPS Run: The Morgan Kaufmann Series in Computer Architecture and Design”, MK: second edition, 2006.
- [52] PPC. Recurso web disponible. <http://www.freescale.com/webapp/sps/site/homepage.jsp?code=PCPPCP>. [Febrero 2012].
- [53] M. González. “Acceso a los recursos de RED & PRACE”. Santander, 2010.
- [54] C. Shamieh. “Understanding 64-bit PowerPC architecture. Critical considerations in 64-bit microprocessor design”, 2004.
- [55] N. Rohrer. “The IBM PowerPC 970FX power envelope and power management. New processor shows promise with advanced power management techniques”, 2004.
- [56] M. Ujaldón Martínez. “Supercomputación y procesadores gráficos”. Universidad de Málaga: Departamento de arquitectura de Computadores.
- [57] NVIDIA Corporation. “NVIDIA’s Next Generation CUDA Compute Architecture”, V1, 2009.
- [58] Información del fabricante Nvidia. Recurso web disponible. http://www.nvidia.es/page/tesla_computing_solutions.html. [Marzo 2012].
- [59] Productos de la familia Tesla Serie 20. Recurso web disponible. http://www.nvidia.es/page/personal_computing.html. [Marzo 2012].
- [60] Información general de GeForce. Recurso web disponible. <http://es.wikipedia.org/wiki/GeForce>. [Marzo 2012].
- [61] Diferencia entre los distintos tipos de relojes y rendimiento. Recurso web disponible. http://www.nvidia.es/object/graphics_cards_buy_now_es.html. [Marzo 2012].
- [62] Productos de la familia Geforce serie 500. Recurso web disponible. http://www.nvidia.es/object/geforce_family_es.html. [Marzo 2012].
- [63] Noticia 680 GTX. Recurso web disponible. <http://www.noticias3d.com/articulo.asp?idarticulo=1603>. [Marzo 2012].
- [64] NVIDIA Corporation. “NVIDIA GeForce GTX 680”, V1, 2012.
- [65] Noticia BSC. Recurso web disponible. <http://www.nvidia.es/object/bsc-deploys-cpu-gpu-supercomputer-20111114-es.html>. [Marzo 2012].
- [66] Nvidia Tegra 2. Recurso web disponible. <http://www.nvidia.com/object/tegra-2.html>. [Marzo 2012].
- [67] NVIDIA Corporation. “Variable SMP (4-PLUS-1™) – A Multi – Core CPU Architecture for Low Power and High Performance”, 2011.
- [68] Artículo. The Benefits of Quad Core CPUs in Mobile Devices. Revision 1.1. 2011.
- [69] Serie HD 6000. Recurso web disponible. http://es.wikipedia.org/wiki/Radeon_HD. [Marzo 2012].

- [70] Productos AMD serie 6000. Recurso web disponible.
<http://www.amd.com/ES/PRODUCTS/DESKTOP/GRAPHICS/AMD-RADEON-HD-6000/Pages/amd-radeon-hd-6000.aspx>. [Marzo 2012].
- [71] Análisis de la presentación de la serie 7000. Recurso web disponible.
<http://www.legitreviews.com/article/1805/1/>. [Marzo 2012].
- [72] Consumo de la serie 7000. Recurso web disponible.
<http://www.legitreviews.com/article/1805/13/>. [Marzo 2012].
- [73] Productos AMD serie 7000. Recurso web disponible.
<http://www.amd.com/us/products/desktop/graphics/7000/Pages/amd-radeon-7000-series.aspx>. [Marzo 2012].
- [74] Productos AMD. Recurso web disponible. <http://products.amd.com/en-us/GraphicCardResult.aspx>. [Marzo 2012].
- [75] AMD Fusion. Recurso web disponible.
<http://www.amd.com/us/products/technologies/fusion/Pages/fusion.aspx>. [Marzo 2012].
- [76] AMD HSA. Recurso web disponible. <http://www.muycorner.com/2012/01/20/amd-dejara-de-utilizar-fusion-a-partir-de-ahora-sera-heterogeneous>. [Marzo 2012].
- [77] Familia AMD APu. Recurso web disponible.
<http://shop.amd.com/es/All/Search?SearchTerms=apu>. [Marzo 2012].
- [78] Definición FLOPS. Recurso web disponible.
http://es.wikipedia.org/wiki/Operaciones_de_coma_flotante_por_segundo. [Marzo 2012].
- [79] K. Kuah. “Motion Estimation with Intel Streaming SIMD Extensions 4 (Intel SSE4)”, 2008.
- [80] Y. Le. “Schema Validation with Intel Streaming SIMD Extensions 4 (Intel SSE4)”, 2008.
- [81] G. Ben, I. Neoran and I. Tubi. “Practical Intel AVX Optimization on 2nd generation Intel Core Processors”, 2011.
- [82] C. Lomont. “Introduction to Intel Advanced Vector Extensions”, 2011.
- [83] Definición TDP. Recurso web disponible.
http://es.wikipedia.org/wiki/Potencia_de_dise%C3%B1o_t%C3%A9rmico. [Marzo 2012].
- [84] EDP. S. Huang, S. Xiao, W. Feng, “On the energy efficiency of graphics processing units for scientific computing”, Proceedings of the 2009 IEEE International Symposium on Parallel & Distributed Processing, IEEE Computer Society, (2009), 1-8.
- [85] EDP. T.R.W. Scogland, H. Lin and W. Feng, “A first look at integrated GPUs for green high-performance computing”, Computer Science - Research and Development, 25(3-4), (2010), 125-134.
- [86] Explicación del benchmark Whetstone. Recurso web disponible.
<http://es.wikipedia.org/wiki/Whetstone>. [Abril 2012].
- [87] H.J. Curnow and B. A. Wichmann. “A synthetic benchmark. The computer Journal”. V. 19 nº 1, 1976.
- [88] Explicación del benchmark Linpack. Recurso web disponible.
<http://es.wikipedia.org/wiki/Linpack>. [Abril 2012].
- [89] Jack J. Dongarra, Piotr Luszczek, and Antoine Petit. “The Linpack Benchmark: Past, Present, and Future”, 2001.
- [90] Explicación PTS. Recurso web disponible. http://en.wikipedia.org/wiki/Phoronix_Test_Suite. [Abril 2012].
- [91] Phoronix test suite. Recurso web disponible. <http://www.phoronix-test-suite.com/?k=home>. [Abril 2012].
- [92] H. C. Wang and K. Hwang, “Multicoloring of grid-structured PDE solvers on shared-memory multiprocessors”, IEEE Transactions on Parallel and Distributed Systems, Vol. 6, No. 11, 1195– 1205, 1995.

- [93] J. A. López-Fernández, M. L. Portugués, J. M. Taboada, H. J. Rice, and F. Obelleiro, "HP- FASS: a hybrid parallel fast acoustic scattering solver", *International Journal of Computer Mathematics* Vol. 88, No. 9, 1960–1968, 2011.
- [94] T. W. Wu, *Boundary Element Acoustics: Fundamentals and Computer Codes (Advances in Boundary Elements)*, WIT Press, 2000.
- [95] Y. Saad and M. H. Schultz, "GMRES: a generalized minimal residual algorithm for solving nonsymmetric linear systems", *SIAM Journal on Scientific and Statistical Computing* Vol. 7, 856–869, 1986.
- [96] J. M. Taboada, L. Landesa, F. Obelleiro, J. L. Rodriguez, J. M. Bértolo, M. G. Araújo, J. C. Mouriño, and A. Gómez, "High scalability FMM- FFT electromagnetic solver for supercomputer systems," *IEEE Antennas Propagation Magazine*, Vol. 51, No. 6, 20–28, 2009.
- [97] Fuente de alimentación test. Recurso web disponible.
http://www.plugloadsolutions.com/psu_reports/XFX%20TECHNOLOGY_XPS-750W-3XB_ECOS%202831.1_750W_Report.pdf. [Junio 2012].

ANEXOS

ANEXO I

En este anexo se encuentra la planificación temporal, en donde se puede ver un esquema-tabla que recoge el tiempo que se ha invertido para realizar las diferentes tareas. Además, está detallado todo el trabajo que se ha realizado en el laboratorio.

Planificación temporal

Las tareas realizadas se han agrupado en cuatro etapas. Por un lado, está la etapa de formación, que abarca 104 días y comprende la búsqueda de la información, profundizar en el estado del arte y adquirir los conocimientos sobre los componentes hardware. En la segunda etapa, se hace referencia al desarrollo, en donde se recoge el trabajo que se ha realizado en el laboratorio. La tercera etapa, hace referencia a todo aquello relacionado con la toma de decisiones. Por último, se encuentra la cuarta etapa, que refleja como ha ido evolucionando la documentación de la presente memoria.

Tabla 19. Planificación temporal.

Tareas	Comienzo	Fin	Duración
Formación	10/11/2011	03/04/2012	104 Días
[Formación] Introducción de los conceptos básicos de HPC	10/11/2011	12/12/2011	23 Días
[Formación] Organización de los procesadores	02/01/2012	26/01/2012	19 Días
[Formación] Arquitectura x86	30/01/2012	06/02/2012	6 Días
[Formación] Arquitectura AMD	06/02/2012	13/02/2012	6 Días
[Formación] ARM	13/02/2012	17/02/2012	5 Días
[Formación] SPARC	17/02/2012	27/02/2012	7 Días
[Formación] MIPS/PowerPC	27/02/2012	29/02/2012	3 Días
[Formación] Nvidia (GPU)	29/02/2012	06/03/2012	5 Días
[Formación] ATI (GPU)	06/03/2012	12/03/2012	5 Días
[Formación] Búsqueda de Benchmarks	14/03/2012	19/03/2012	4 Días
[Formación] Nvidia Kepler	23/03/2012	27/03/2012	3 Días
[Formación] Estudio de componentes hardware eficientes	12/03/2012	03/04/2012	17 Días
Desarrollo	12/03/2012	14/06/2012	69 Días
[Desarrollo] Instalación del entorno de trabajo en el PC de referencia	12/03/2012	14/03/2012	3 Días
[Desarrollo] Instalación y configuración del paquete MatLab en PC de referencia	15/03/2012	15/03/2012	1 Día
[Desarrollo] Instalación del compilador de Intel ICC en el PC de referencia	02/04/2012	02/04/2012	1 Día
[Desarrollo] Pruebas con Whetstone y Dhrystone PC de referencia	10/04/2012	11/04/2012	2 Días
[Desarrollo] Pruebas con MatLab PC de referencia	12/04/2012	13/04/2012	2 Días
[Desarrollo] Pruebas Linpack PC de referencia	18/04/2012	20/04/2012	3 Días
[Desarrollo] Pruebas Phoronix PC de referencia	23/04/2012	25/04/2012	3 Días
[Desarrollo] Pruebas Códigos TSC PC de referencia	26/04/2012	27/04/2012	2 Días

[Desarrollo] Instalación y configuración del prototipo	31/05/2012	31/05/2012	1 Día
[Desarrollo] Instalación ICC, librerías, MatLab en el prototipo	1/06/2012	1/06/2012	1 Día
[Desarrollo] Pruebas Códigos TSC en el prototipo	04/06/2012	06/06/2012	3 Días
[Desarrollo] Pruebas con Whetstone, MatLab CPU, Linpack en prototipo	13/06/2012	13/06/2012	1 Día
[Desarrollo] Pruebas Phoronix prototipo	14/06/2012	14/06/2012	1 Día
Toma de Decisiones	13/03/2012	20/04/2012	29 Días
[Toma de decisiones] Instalación del sistema operativo Centos 6.2	13/03/2012	14/03/2012	2 Días
[Toma de decisiones] Selección final de los componentes del prototipo	27/03/2012	03/04/2012	6 Días
[Toma de decisiones] Selección de Benchmarks a medir en PC de referencia	10/04/2012	20/04/2012	9 Días
Documentación	09/12/2011	25/06/2012	142 Días
[Documentación] Conceptos básicos de HPC	09/12/2011	12/12/2011	2 Días
[Documentación] Organización de los Procesadores	16/01/2012	30/01/2012	11 Días
[Documentación] Arquitecturas x86	03/01/2012	06/01/2012	4 Días
[Documentación] Arquitectura AMD	09/02/2012	13/02/2012	3 Días
[Documentación] ARM	15/02/2012	17/02/2012	3 Días
[Documentación] Reorganización de la estructura de la Tesis	22/02/2012	24/02/2012	3 Días
[Documentación] SPARC	24/02/2012	27/02/2012	2 Días
[Documentación] MIPS/PowerPC	27/02/2012	29/02/2012	3 Días
[Documentación] Nvidia (GPU)	02/03/2012	06/03/2012	3 Días
[Documentación] ATI (GPU)	08/03/2012	12/03/2012	3 Días
[Documentación] Formato específico en la memoria	12/03/2012	15/03/2012	4 Días
[Documentación] Corrección de la memoria	15/03/2012	19/03/2012	3 Días
[Documentación] Estudio teórico de primeros componentes eficientes	14/03/2012	19/03/2012	4 Días
[Documentación] Corrección del estudio teórico de los componentes eficientes	20/03/2012	26/03/2012	5 Días
[Documentación] Nvidia Kepler	26/03/2012	26/03/2012	1 Día
[Documentación] Estudio teórico y selección del prototipo definitivo	27/03/2012	03/04/2012	6 Días
[Documentación] Diario de laboratorio del PC de referencia y resultados obtenidos	05/04/2012	27/04/2012	17 Días
[Documentación] Diario del PC eficiente y resultados obtenidos	04/06/2012	25/06/2012	16 Días

Desarrollo del trabajo en el laboratorio

El desarrollo del trabajo se divide en dos partes, por un lado está la información técnica del PC de referencia y del prototipo y por el otro lado, está descrito el trabajo que se ha realizado en el laboratorio para obtener los resultados.

Desarrollo en el PC de referencia

A continuación se muestran las características técnicas y el trabajo realizado en el laboratorio con el PC de referencia.

Información técnica

El pc que se ha utilizado de referencia cuenta con las siguientes características técnicas:

- **Microprocesador (CPU):** Core i3-2100 Sandy Bridge, el modelo tiene dos núcleos físicos con una frecuencia de trabajo por cada núcleo de 3,1 GHz. El modelo tiene a su disposición la tecnologías de Hyper-Threading con lo cual dispone de hasta 4 hilos virtualizados de trabajo.
- **Tarjeta gráfica (GPU):** Intel HD Graphics 3000. El modelo de tarjeta gráfica es el que viene integrado junto a los núcleos del microprocesador.
- **Memoria RAM:** Tiene disponibles 4 GB, SDRAM, DDR3 a 1333 MHz.
- **Disco duro (HDD):** Se ha utilizado un HDD a 7200 r.p.m. En el disco duro se ha instalado el sistema operativo CentOS 6.2.

Trabajo de laboratorio

A continuación se detalla todo el trabajo que se ha realizado en el laboratorio utilizando el PC de referencia.

En el primer día, se ha instalado el sistema operativo Windows 7 Professional de 64 bits en el PC utilizado para las pruebas. Después de instalar el operativo, se han instalado los drivers necesarios para que el sistema funcionase correctamente.

En el segundo día de laboratorio, se ha instalado el sistema operativo CentOS 6.2 (Linux) en el PC de referencia, también se ha configurado el acceso a internet y los permisos de usuario y root.

Una vez que el PC ha tenido acceso a internet, dentro del entorno de trabajo de Windows se ha realizado la primera toma de contacto para verificar que todo funcionaba correctamente. El software que se ha utilizado ha sido por un lado, el paquete CPUZ, que se encarga de medir los parámetros del computador tales como: la velocidad de trabajo del reloj, la carga en el bus de datos y la cantidad de memoria que se está utilizando en el sistema. El otro paquete software utilizado ha sido el programa "SuperPI". Este software se encarga de calcular una serie de decimales del número PI, ya que la tarea de calcular decimales implica poner a trabajar al microprocesador. Al ejecutar el software de "SuperPI" mediante la monitorización ofrecida por el paquete CPUZ, se ha podido observar un aumento de la frecuencia de trabajo del microprocesador hasta su límite, además del propio aumento de los parámetros de carga del bus de datos y cantidad de memoria utilizada por el sistema.

Al finalizar las primeras pruebas en el entorno de trabajo de Windows, el PC de referencia al tener acceso a la red, ha descargado las actualizaciones de Windows 7 y las ha instalado. Dichas actualizaciones no se debieron de instalar correctamente produciendo un error después de reiniciar el computador.

En el tercer día de laboratorio, se ha intentado solucionar el error que había en Windows de la sesión anterior. Se ha optado por formatear como solución para recuperarlo. Después de ver que no era fiable, se ha descartado el uso de Windows para las pruebas, ya que puede fallar en cualquier momento. Por lo tanto, se ha optado por utilizar sólo el sistema CentOS 6.2.

En el cuarto día de laboratorio, después de los inconvenientes sufridos en días anteriores, se ha conseguido instalar de nuevo el sistema operativo Centos 6.2. Una vez instalado el operativo se ha procedido a instalar el software matemático MatLab en su versión 2011b.

Ese mismo día de trabajo, se ha comprobado que el entorno funciona correctamente mediante dos pruebas. La primera prueba consistió en ejecutar el comando “bench;” dentro del entorno de MatLab. Como resultado se obtuvo las puntuaciones de los diferentes benchmarks que tiene el paquete MatLab predefinidos, también se pudo observar que para realizar los diferentes benchmarks el microprocesador ha necesitado poner a trabajar a todos sus cores al 100%. Por lo tanto, se ha conseguido demostrar que el entorno de trabajo para medir funciona correctamente.

La segunda y última prueba que se ha realizado en el cuarto día de laboratorio, ha sido realizar una factorización de una matriz LU de tamaño 8000x8000, en donde se ha confirmado que todo funciona correctamente.

En el quinto día de trabajo en el laboratorio, se ha instalado el compilador de Intel denominado ICC, además se han observado las primeras pruebas con el benchmark denominado Whetstone que ha sido compilado con “gcc”. A la hora de realizar la prueba, ha aparecido un problema, ya que el benchmark solamente funciona en un núcleo. Para solventar este problema se ha creado un script denominado “whetstone_4.sh”, encargado de lanzar 4 veces el programa. Además, el script se encarga de almacenar los resultados de las pruebas en ficheros para luego poder consultarlos. La sesión del laboratorio ha finalizado mediante la comprobación del funcionamiento del benchmark dry.sh, el cual viene configurado por defecto y se han tenido que modificar algunos parámetros para su ejecución.

En el sexto día, se ha realizado una revisión de todo el equipo de referencia, para ver que todo funcionaba correctamente y empezar a realizar las mediciones. Se ha visto que al lanzar diferentes programas intensivos ocupaban toda la CPU disponible, con ello se ha asegurado que todo funcionaba correctamente.

En el séptimo día de laboratorio, se han obtenido los primeros resultados utilizando el PC de referencia. Primero se ha compilado el benchmark Whetstone, y se ha modificado el script creado en el quinto día del laboratorio, con el objetivo de lanzar el programa para que se ejecute en uno, dos, tres y cuatro hilos del computador. Lanzándolo en diferentes hilos se pudo observar los diferentes consumos pico que requería la CPU para ejecutar el benchmark. Una vez terminadas las pruebas con Whetstone, se ha pasado a medir el consumo necesario para ejecutar el benchmark Dhrystone.

El benchmark Dhrystone funcionaba y compilaba correctamente, pero debido a que el código original tiene algunas estructuras obsoletas y el sistema de evaluación de la CPU es similar al de Whetstone, finalmente se ha descartado su uso para evaluar el PC de referencia.

En el octavo y noveno día de sesión, se han obtenido resultados del consumo energético utilizando el paquete software MatLab. Se han realizado dos tipos de medidas. El primer tipo de medida ha consistido en realizar productos de matrices de diferentes tamaños, pasando desde ocupar toda la memoria disponible en el PC de referencia, hasta un conjunto de matrices que apenas ocupaban junto con otros procesados del computador 1 GB. Para la segunda prueba, se ha utilizado la herramienta FFT de MatLab. Como sucedía con el producto de matrices, mediante un vector de cierto tamaño al que se le ha aplicado la FFT, se ha ocupado la memoria disponible desde prácticamente el 100% hasta menos de 1 GB para obtener distintos resultados y poder comparar.

En los siguientes tres días de laboratorio, se han instalado en el PC de referencia el benchmark Linpack y el paquete software Phoronix-test-suite.

En cuanto al benchmark Linpack, han aparecido algunas dificultades a la hora de ejecución, ya que el código original sólo se ejecuta en un solo hilo de la CPU. Por tanto, como sucedía con el benchmark Whetstone, se tuvo que crear un script para lanzar el paquete en varios hilos de la CPU.

En cuanto a la suite de Phoronix. Una vez que se instaló el paquete, se tuvo que actualizar las librerías de PHP a su versión 5.2 para cumplir con los requisitos de ejecución. Además se tuvo que realizar una selección de benchmarks útiles para la Tesis, ya que la suite dispone de una gran variedad de benchmarks que sirven para medir diferentes componentes, por ejemplo: pruebas en compiladores, pruebas en juegos en 3D, pruebas de video, pruebas de audio, pruebas de aceleración 3D, pruebas de tarjetas de red, etc. Por tanto, se han seleccionado aquellos benchmarks con más relación con la HPC.

En los últimos cuatro días de sesión de trabajo con el PC de referencia, se han instalado los diferentes benchmarks dentro de la suite de Phoronix para poder evaluar el consumo. Los benchmarks dentro de la suite no vienen instalados, por lo que hay que descargarlos y son paquetes de varios Gigabytes, además tienen que descargarse de su propio repositorio.

Después de instalar los diferentes paquetes software necesarios en la suite de Phoronix, se ha procedido a instalar las librerías MPI y el paquete de librerías FFTW 3.3.1 que requieren los códigos del Área de TSC para su correcto funcionamiento. Una vez cumplidos los requisitos para el correcto funcionamiento de los códigos de TSC, se ha procedido a realizar las medidas con dichos códigos.

Desarrollo en el prototipo

A continuación se muestran las características técnicas y el trabajo realizado en el laboratorio con el prototipo.

Información técnica

El prototipo cuenta con las siguientes características técnicas:

- **Microprocesador (CPU):** Core i5-2500T Sandy Bridge, el modelo tiene cuatro núcleos físicos con una frecuencia de trabajo por cada núcleo de 2,3 GHz. Con un TDP de 45W.
- **Tarjeta gráfica (GPU):** Modelo GeForce GTX 560 compatible con CUDA.
- **Memoria RAM:** Tiene disponibles 8 GB, DDR3 a 1333 MHz.
- **Disco duro (HDD):** Se ha utilizado un HDD a 7200 r.p.m. En el disco duro se ha instalado el sistema operativo CentOS 6.2.

Trabajo de laboratorio

A continuación se detalla todo el trabajo que se ha realizado en el laboratorio utilizando el prototipo eficiente.

En el primer día de laboratorio con el prototipo, una vez se ha montado el equipo, se ha comprobado que la EFI (antigua BIOS) estaba configurada de forma óptima. Sobre todo, hay que destacar que se

ha dejado activada la tecnología Turbo Boost 2.0, explicada anteriormente en esta memoria. También se ha desactivado el Bluetooth, las controladoras que no se estaban utilizando y el USB 3.0. Seleccionado la opción optimizada en el control de VRM. Por último, la placa venía por defecto con overclocking, por lo tanto se ha desactivado. Una vez que todo se ha configurado, se ha pasado a instalar el sistema operativo (CentOS 6.2) tal y como se había instalado en el PC de referencia, es decir, con las mismas opciones. Además se ha actualizado el software a sus últimas versiones.

Después del primer reinicio, han surgido los primeros problemas debido a los drivers por defecto de la tarjeta gráfica denominados “nouveau”, que se instalan en Linux cuando se detecta la gráfica. Para solucionar este problema, se ha creado una lista negra para evitar este driver en el arranque.

Ya para terminar la sesión, se han instalado los drivers correctos de CUDA en su versión 4.2, el SDK y el Toolkit correspondiente. Dejando el prototipo configurado y operativo para el resto de sesiones.

En el segundo día de laboratorio con el prototipo, se han instalado las herramientas necesarias para empezar a realizar las primeras mediciones. Lo primero que se realizó, ha sido comprobar que la temperatura de la CPU y de la placa base marcaban los valores correctos.

Una vez comprobado el estado de temperaturas, se ha pasado a copiar el directorio denominado “opt” del PC de referencia al prototipo. Con ello, se ha evitado instalar de nuevo el compilador de Intel, las librerías Mpich2 y las librerías FFTW, que se requieren para el funcionamiento de los códigos del Área de TSC. Lo segundo y último que se ha realizado ha sido volcar el directorio de MatLab (/usr/local), para evitar realizar una instalación de MatLab en el prototipo.

En el tercer día de trabajo con el prototipo, se han empezado a realizar las medidas con el computador. Las primeras medidas realizadas han sido utilizando el código “MST_CPU” en uno, dos y cuatro cores.

En el cuarto día de laboratorio, se han realizado las medidas del código FMM sin FFT (Fass_1.4.7) en su versión de CPU, en uno, dos y cuatro cores. Y después, del código FMM que requiere del uso de la FFT (Fass_1.5.0) también en uno, dos y cuatro cores.

En el quinto día de laboratorio, se han realizado las medidas de los códigos del departamento de TSC en su versión de CUDA. Para realizar estas medidas, hay que conectarse remotamente al prototipo, ya que pueden surgir complicaciones al compartir la gráfica para realizar los cálculos y a la vez sacar imagen hacia el monitor. Al realizar estas medidas, se ha visto el verdadero potencial del prototipo.

En el sexto día de laboratorio, se han realizado las medidas del código “Whetstone” para poder comparar los resultados con el PC de referencia. Además, se han realizado las medidas con los scripts en Matlab para CPU, tanto en simple como en doble precisión. Por último, se han capturado los datos del benchmark Linpack.

En el séptimo día de laboratorio, se ha instalado la herramienta Phoronix y se han realizado las medidas con dicha herramienta.

ANEXO II

Los datos que se proporcionan en el siguiente anexo, han sido extraídos de la página del fabricante Intel [www.intel.com] a fecha de febrero de 2012.

Tabla 20. Generación Core i3.

Procesador	Velocidad de reloj	Memoria Caché	Tipo de memoria	Intel HD Graphics	TDP Max	Núcleos/Subprocesos
Core i3-560	3.33GHz	4MB	DDR3-1066/1333	SI	73W	2/4
Core i3-550	3.20GHz	4MB	DDR3-1066/1333	SI	73W	2/4
Core i3-540	3.06GHz	4MB	DDR3-1066/1333	SI	73W	2/4
Core i3-390M	2.66GHz	3MB	DDR3-800/1066	SI	35W	2/4
Core i3-380UM	1.33GHz	3MB	DDR3-800	SI	18W	2/4
Core i3-380M	2.53GHz	3MB	DDR3-800/1066	SI	35W	2/4
Core i3-370M	2.40GHz	3MB	DDR3-800/1066	SI	35W	2/4
Core i3-350M	2.26GHz	3MB	DDR3-800/1066	SI	35W	2/4
Core i3-330UM	1.20GHz	3MB	DDR3-800	SI	18W	2/4
Core i3-330M	2.13 GHz	3MB	DDR3-800/1066	SI	35W	2/4

Tabla 21. Generación Core i5.

Procesador	Velocidad de reloj	Memoria Caché	Tipo de memoria	Intel HD Graphics	TDP Max	Núcleos/Subprocesos
Core i5-760	2.80GHz	8MB	DDR3-1066/1333	NO	95W	4/4
Core i5-750S	2.40GHz	8MB	DDR3-1066/1333	NO	82W	4/4
Core i5-750	2.66GHz	8MB	DDR3-1066/1333	NO	95W	4/4
Core i5-680	3.60GHz	4MB	DDR3-1066/1333	SI	73W	2/4
Core i5-670	3.46GHz	4MB	DDR3-1066/1333	SI	73W	2/4
Core i5-661	3.33GHz	4MB	DDR3-1066/1333	SI	87W	2/4
Core i5-660	3.33GHz	4MB	DDR3-1066/1333	SI	73W	2/4
Core i5-655k	3.20GHz	4MB	DDR3-1066/1333	SI	73W	2/4
Core i5-650	3.20GHz	4MB	DDR3-1066/1333	SI	73W	2/4
Core i5-580M	2.66GHz	3MB	DDR3-800/1066	SI	35W	2/4
Core i5-	1.33GHz	3MB	DDR3-800	SI	18W	2/4

560UM						
Core i5-560M	2.66GHz	3MB	DDR3-800/1066	SI	35W	2/4
Core i5-540UM	1.20GHz	3MB	DDR3-800	SI	18W	2/4
Core i5-540M	2.53GHz	3MB	DDR3-800/1066	SI	35W	2/4
Core i5-520UM	1.06GHz	3MB	DDR3-800	SI	18W	2/4
Core i5-520M	2.40GHz	3MB	DDR3-800/1066	SI	35W	2/4
Core i5-520E	2.40GHz	3MB	DDR3-800/1066	SI	35W	2/4
Core i5-480M	2.66GHz	3MB	DDR3-800/1066	SI	35W	2/4
Core i5-470UM	1.33GHz	3MB	DDR3-800	SI	18W	2/4
Core i5-460M	2.53GHz	3MB	DDR3-800/1066	SI	35W	2/4
Core i5-450M	2.40GHz	3MB	DDR3-800/1066	SI	35W	2/4
Core i5-430UM	1.20GHz	3MB	DDR3-800	SI	18W	2/4
Core i5-430M	2.26GHz	3MB	DDR3-800/1066	SI	35W	2/4

Tabla 22. Generación Core i7.

Procesador	Velocidad de reloj	Memoria Caché	Tipo de memoria	Intel HD Graphics	TDP Max	Núcleos/Subprocesos
Core i7-980	3.33GHz	12MB	DDR3-1066		130W	6/12
Core i7-970	3.20GHz	12MB	DDR3-800/1066		130W	6/12
Core i7-960	3.20GHz	8MB	DDR3-800/1066	NO	130W	4/8
Core i7-950	3.06GHz	8MB	DDR3-800/1066	NO	130W	4/8
Core i7-940	2.93GHz	8MB	DDR3-800/1066	NO	130W	4/8
Core i7-930	2.80GHz	8MB	DDR3-800/1066	NO	130W	4/8
Core i7-920	2.66GHz	8MB	DDR3-800/1066	NO	130W	4/8
Core i7-880	3.06GHz	8MB	DDR3-1066/1333	NO	95W	4/8
Core i7-875K	2.93GHz	8MB	DDR3-1066/1333	NO	95W	4/8
Core i7-870S	2.66GHz	8MB	DDR3-1066/1333	NO	82W	4/8
Core i7-870	2.93GHz	8MB	DDR3-1066/1333	NO	95W	4/8
Core i7-860S	2.53GHz	8MB	DDR3-1066/1333	NO	82W	4/8
Core i7-860	2.80GHz	8MB	DDR3-1066/1333	NO	95W	4/8

Core i7-840QM	1.86GHz	8MB	DDR3-1066/1333	NO	45W	4/8
Core i7-820QM	1.73GHz	8MB	DDR3-1066/1333	NO	45W	4/8
Core i7-740QM	1.73GHz	6MB	DDR3-1066/1333	NO	45W	4/8
Core i7-720QM	1.60GHz	6MB	DDR3-1066/1333	NO	45W	4/8
Core i7-680UM	1.46GHz	4MB	DDR3-800	SI	18W	2/4
Core i7-660UM	1.33GHz	4MB	DDR3-800	SI	18W	2/4
Core i7-660UE	1.33GHz	4MB	DDR3-800	SI	18W	2/4
Core i7-660LM	2.26GHz	4MB	DDR3-800/1066	SI	25W	2/4
Core i7-640UM	1.20GHz	4MB	DDR3-800	SI	18W	2/4
Core i7-640M	2.80GHz	4MB	DDR3-800/1066	SI	35W	2/4
Core i7-640LM	2.13GHz	4MB	DDR3-800/1066	SI	25W	2/4
Core i7-620UM	1.06GHz	4MB	DDR3-800	SI	18W	2/4
Core i7-620UE	1.06GHz	4MB	DDR3-800	SI	18W	2/4
Core i7-620M	2.66GHz	4MB	DDR3-800/1066	SI	35W	2/4
Core i7-620LM	2.00GHz	4MB	DDR3-800/1066	SI	25W	2/4
Core i7-620LE	2.00GHz	4MB	DDR3-800/1066	SI	25W	2/4
Core i7-610E	2.53GHz	4MB	DDR3-800/1066	SI	35W	2/4
Core i7-990X-Extreme	3.46GHz	12MB	DDR3-1066	NO	130W	6/12
Core i7-980X-Extreme	3.33GHz	12MB	DDR3-1066	NO	130W	6/12
Core i7-975-Extreme	3.33GHz	8MB	DDR3-800/1066	NO	130W	4/8
Core i7-965-Extreme	3.20GHz	8MB	DDR3-800/1066	NO	130W	4/8
Core i7-940XM-Extreme	2.13GHz	8MB	DDR3-1066/1333	NO	55W	4/8
Core i7-920XM	2.00GHz	8MB	DDR3-1066/1333	NO	55W	4/8

La leyenda de sufijos es la siguiente:

- K- Procesadores con el multiplicador desbloqueado.
- S- Procesadores más eficientes energéticamente.
- T- Procesadores muy eficientes energéticamente, con frecuencias de reloj más bajas que las de la serie.

- M- Procesadores portátiles.
- LM- Procesadores portátiles de baja tensión.
- UM- Procesadores portátiles de muy baja tensión.
- QM- Procesadores portátiles de cuádruple núcleo.
- XM- Procesadores portátiles de cuádruple núcleo con el multiplicador desbloqueado.
- E- Procesadores para sistemas embebidos.
- LE- Procesadores para sistemas embebidos de baja tensión.
- UE- Procesadores para sistemas embebidos de muy baja tensión.
- QE- Procesadores para sistemas embebidos de cuádruple núcleo.

ANEXO III

Los datos que se proporcionan en el siguiente anexo, han sido extraídos de la página del fabricante Intel [www.intel.com] a fecha de febrero de 2012.

Tabla 23. 2ª Generación Core i3.

Procesador	Velocidad de reloj	Memoria Caché	Tipo de memoria	Intel HD Graphics	TDP Max	Núcleos/Subprocesos
Core i3-2370M	2.40GHz	3MB	DDR3-1066/1333	SI	35W	2/4
Core i3-2367M	1.40GHz	3MB	DDR3-1066/1333	SI	17W	2/4
Core i3-2357M	1.30GHz	3MB	DDR3-1066/1333	SI	17W	2/4
Core i3-2350M	2.30GHz	3MB	DDR3-1066/1333	SI	35W	2/4
Core i3-2340UE	1.30GHz	3MB	DDR3-1066/1333	SI	17W	2/4
Core i3-2330M	2.20GHz	3MB	DDR3-1066/1333	SI	35W	2/4
Core i3-2330E	2.20GHz	3MB	DDR3-1066/1333	SI	35W	2/4
Core i3-2312M	2.10GHz	3MB	DDR3-1066/1333	SI	35W	2/4
Core i3-2310M	2.10GHz	3MB	DDR3-1066/1333	SI	35W	2/4
Core i3-2310E	2.10GHz	3MB	DDR3-1066/1333	SI	35W	2/4
Core i3-2130	3.40GHz	3MB	DDR3-1066/1333	SI	65W	2/4
Core i3-2125	3.30GHz	3MB	DDR3-1066/1333	SI	65W	2/4
Core i3-2120T	2.60GHz	3MB	DDR3-1066/1333	SI	35W	2/4
Core i3-2120	3.30GHz	3MB	DDR3-1066/1333	SI	65W	2/4
Core i3-2105	3.10GHz	3MB	DDR3-1066/1333	SI	65W	2/4
Core i3-2102	3.10GHz	3MB	DDR3-1066/1333	SI	65W	2/4
Core i3-2100T	2.50GHz	3MB	DDR3-1066/1333	SI	35W	2/4
Core i3-2100	3.10GHz	3MB	DDR3-1066/1333	SI	65W	2/4

Tabla 24. 2ª Generación Core i5.

Procesador	Velocidad de reloj	Memoria Caché	Tipo de memoria	Intel HD Graphics	TDP Max	Núcleos/Subprocesos
Core i5-2435M	2.40GHz	3MB	DDR3-1066/1333	SI	35W	2/4
Core i5-2557M	1.70GHz	3MB	DDR3-1066/1333	SI	17W	2/4
Core i5-	3.40GHz	6MB	DDR3-		95W	4/4

2550K			1066/1333			
Core i5-2540M	2.60GHz	3MB	DDR3-1066/1333	SI	35W	2/4
Core i5-2537M	1.40GHz	3MB	DDR3-1066/1333	SI	17W	2/4
Core i5-2520M	2.50GHz	3MB	DDR3-1066/1333	SI	35W	2/4
Core i5-2515E	2.50GHz	3MB	DDR3-1066/1333	SI	35W	2/4
Core i5-2510E	2.50GHz	3MB	DDR3-1066/1333	SI	35W	2/4
Core i5-2500T	2.30GHz	6MB	DDR3-1066/1333	SI	45W	4/4
Core i5-2500S	2.70GHz	6MB	DDR3-1066/1333	SI	65W	4/4
Core i5-2500K	3.30GHz	6MB	DDR3-1066/1333	SI	95W	4/4
Core i5-2500	3.30GHz	6MB	DDR3-1066/1333	SI	95W	4/4
Core i5-2467M	1.60GHz	3MB	DDR3-1066/1333	SI	17W	2/4
Core i5-2450P	3.20GHz	6MB	DDR3-1066/1333		95W	4/4
Core i5-2450M	2.50GHz	3MB	DDR3-1066/1333	SI	35W	2/4
Core i5-2430M	2.40GHz	3MB	DDR3-1066/1333	SI	35W	2/4
Core i5-2410M	2.30GHz	3MB	DDR3-1066/1333	SI	35W	2/4
Core i5-2405S	2.50GHz	6MB	DDR3-1066/1333	SI	65W	4/4
Core i5-2400S	2.50GHz	6MB	DDR3-1066/1333	SI	65W	4/4
Core i5-2400	3.10GHz	6MB	DDR3-1066/1333	SI	95W	4/4
Core i5-2390T	2.70GHz	3MB	DDR3-1066/1333	SI	35W	2/4
Core i5-2380P	3.10GHz	6MB	DDR3-1066/1333		95W	4/4
Core i5-2320	3.00GHz	6MB	DDR3-1066/1333	SI	95W	4/4
Core i5-2310	2.90GHz	6MB	DDR3-1066/1333	SI	95W	4/4
Core i5-2300	2.80GHz	6MB	DDR3-1066/1333	SI	95W	4/4

Tabla 25. 2ª Generación Core i7.

Procesador	Velocidad de reloj	Memoria Caché	Tipo de memoria	Intel HD Graphics	TDP Max	Núcleos/Subprocesos
Core i7-3930K	3.20GHz	12MB	DDR3-1600	NO	130W	6/12
Core i7-2860QM	2.50GHz	8MB	DDR3-1066/1333/1600	SI	45W	4/8
Core i7-	2.30GHz	8MB	DDR3-	SI	45W	4/8

2820QM			1066/1333/1600			
Core i7-2760QM	2.40GHz	6MB	DDR3-1066/1333/1600	SI	45W	4/8
Core i7-2720QM	2.20GHz	6MB	DDR3-1066/1333/1600	SI	45W	4/8
Core i7-2715QE	2.10GHz	6MB	DDR3-1066/1333/1600	SI	45W	4/8
Core i7-2710QE	2.10GHz	6MB	DDR3-1066/1333/1600	SI	45W	4/8
Core i7-2700K	3.5GHz	8MB	DDR3-1066/1333	SI	95W	4/8
Core i7-2677M	1.80GHz	4MB	DDR3-1066/1333	SI	17W	2/4
Core i7-2675QM	2.20GHz	6MB	DDR3-1066/1333	SI	45W	4/8
Core i7-2670QM	2.20GHz	6MB	DDR3-1066/1333	SI	45W	4/8
Core i7-2657M	1.60GHz	4MB	DDR3-1066/1333	SI	17W	2/4
Core i7-2655LW	2.20GHz	4MB	DDR3-1066/1333	SI	25W	2/4
Core i7-2649M	2.30GHz	4MB	DDR3-1066/1333	SI	25W	2/4
Core i7-2640M	2.80GHz	4MB	DDR3-1066/1333	SI	35W	2/4
Core i7-2637M	1.70GHz	4MB	DDR3-1066/1333	SI	17W	2/4
Core i7-2635QM	2.00GHz	6MB	DDR3-1066/1333	SI	45W	4/8
Core i7-2630QM	2.00GHz	6MB	DDR3-1066/1333	SI	45W	4/8
Core i7-2629M	2.10GHz	4MB	DDR3-1066/1333	SI	25W	2/4
Core i7-2620M	2.70GHz	4MB	DDR3-1066/1333	SI	35W	2/4
Core ii7-2617M	1.50GHz	4MB	DDR3-1066/1333	SI	17W	2/4
Core i7-2610UE	1.50GHz	4MB	DDR3-1066/1333	SI	17W	2/4
Core i7-2600S	2.80GHz	8MB	DDR3-1066/1333	SI	65W	4/8
Core i7-2600K	3.40GHz	8MB	DDR3-1066/1333	SI	95W	4/8
Core i7-2600	3.40GHz	8MB	DDR3-1066/1333	SI	95W	4/8

La leyenda de sufijos es la siguiente:

- K- Procesadores con el multiplicador desbloqueado.
- S- Procesadores más eficientes energéticamente.
- T- Procesadores muy eficientes energéticamente, con frecuencias de reloj más bajas que las de la serie.
- M- Procesadores portátiles.
- LM- Procesadores portátiles de baja tensión.
- UM- Procesadores portátiles de muy baja tensión.
- QM- Procesadores portátiles de cuádruple núcleo.

- XM- Procesadores portátiles de cuádruple núcleo con el multiplicador desbloqueado.
- E- Procesadores para sistemas embebidos.
- LE- Procesadores para sistemas embebidos de baja tensión.
- UE- Procesadores para sistemas embebidos de muy baja tensión.
- QE- Procesadores para sistemas embebidos de cuádruple núcleo.

ANEXO IV

Los datos que se proporcionan en el siguiente anexo, han sido extraídos de la página del fabricante AMD [shop.amd.com/es/] a fecha de febrero de 2012.

Tabla 26. Micro Phenom II.

Procesador	Velocidad de reloj	Memoria Caché L2	Memoria Caché L3	TDP Max	Núcleos
Phenom II X2 550	3.1GHz	512KB	6144KB	80W	2
Phenom II X2 545	3.0GHz	512KB	6144KB	80W	2
Phenom II X2 Black 560	3.3GHz	1024KB	6144KB	80W	2
Phenom II X2 565	3.4GHz	1024KB	6144KB	80W	2
Phenom II X4 850	3.3GHz	2048KB	6144KB	95W	4
Phenom II X4 955	3.2GHz	2048KB	6144KB	125W	4
Phenom II X4 960T	3.0GHz	2048KB	6144KB	95W	4
Phenom II X4 945	3.0GHz	2048KB	6144KB	95W	4
Phenom II X4 965	3.4GHz	2048KB	6144KB	125W	4
Phenom II X6 1045T	2,7GHz	3072KB	6144KB	95W	6
Phenom II X6 1055T	2.8GHz	3072KB	6144KB	125W	6
Phenom II X4 975	3.6GHz	2048KB	6144KB	125W	4
Phenom II X4 Black 970	3.5GHz	2048KB	6144KB	125W	4
Phenom II X6 1075T	3.0GHz	3072KB	6144KB	125W	6
Phenom II X4 980	3.7GHz	2048KB	6144KB	125W	4
Phenom II X4 910e	2.6GHz	2048KB	6144KB	65W	4
Phenom II X6 Black 1100T	3.3GHz	3072KB	6144KB	125W	6
Phenom II X6 Black 1090T	3.2GHz	3072KB	6144KB	125W	6

ANEXO V

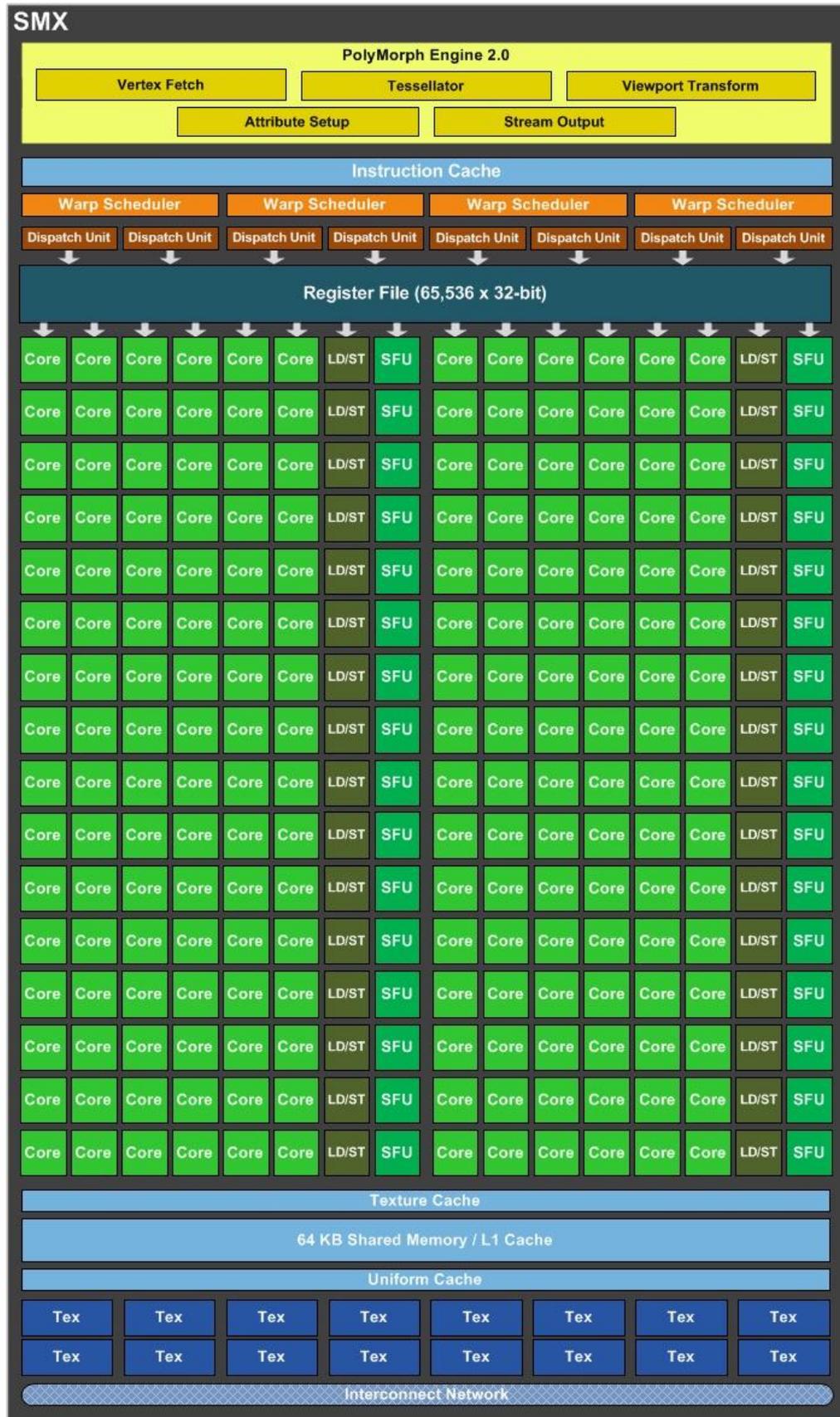


Figura 49. SMX GeForce Serie GTX 680 ^[64].

ANEXO VI

Primero se muestran los resultados capturados al ejecutar el benchmark Whetstone en el PC de referencia y después en el prototipo.

Tabla 27. Resultados de Whetstone en el PC de referencia para 500000 repeticiones.

	Amperios (Pico)	Tiempo (s)	MIPS	Tensión (V)	Potencia requerida PICO (W)	MIPS / W
1 Core	0,22	17	2941,2	238	52,36	56,17
2 Cores	0,27	17	2941,2	238	64,26	91,54
		17	2941,2			
3 Cores	0,28	22	2272,7	238	66,64	112,34
		17	2941,2			
		22	2272,7			
4 Cores	0,28	23	2173,9	238	66,64	129,13
		24	2083,3			
		23	2173,9			
		23	2173,9			

Tabla 28. Resultados de Whetstone en el prototipo para 500000 repeticiones.

	Amperios (Pico)	Tiempo (s)	MIPS	Tensión (V)	Potencia requerida PICO (W)	MIPS / W
1 Core	0,3	16	3125	241,5	72,45	43,13
2 Cores	0,35	17	2941,2	241,5	84,525	69,59
		17	2941,2			
3 Cores	0,35	19	2631,6	241,5	84,525	93,4
		19	2631,6			
		19	2631,6			
4 Cores	0,35	22	2272,7	241,5	84,525	107,55
		22	2272,7			
		22	2272,7			
		22	2272,7			

ANEXO VII

Las pruebas que han requerido del uso de la CPU del PC de referencia, se han capturado con una tensión de **237V**. A continuación se muestran los resultados obtenidos.

Tabla 29. Resultados del benchmark interno de Matlab en el PC de referencia.

Bench Interno de MatLab	LU (s)	FFT (s)	ODE (s)	Sparse (s)	2-D (s)	3-D (s)	Amperios PICO	Potencia (W)	Energía (W*s)	EDP (J * s)
1º Vez	0,1021	0,1447	0,1441	0,1473	0,2530	0,7169	0,26	61,62	92,93	140,15
2º Vez	0,1682	0,1406	0,1478	0,2080	0,2385	0,7082	0,26	61,62	99,29	159,98
3º Vez	0,1030	0,1625	0,1425	0,1491	0,2347	0,7174	0,26	61,62	93	140,35

Tabla 30. Resultados del producto de matrices en el PC de referencia (doble precisión).

Dimensiones de las matrices	Amperios (Pico)	Tiempo (s)	Tensión (V)	Potencia requerida PICO (W)	Energía (W*s)	EDP (J * s)
11180x11180	0,31	74,21	237	73,47	5452,21	404608,41
10000x10000	0,31	47,88	237	73,47	3517,74	168429,56
8000x8000	0,31	24,14	237	73,47	1773,57	42813,88
6000x6000	0,31	10,18	237	73,47	747,92	7613,87
4000x4000	0,31	3,11	237	73,47	228,49	710,61
2000x2000	0,29	0,39	237	68,73	26,8	10,45

Tabla 31. Resultados de aplicar la FFT en el PC de referencia (doble precisión).

Tamaño del vector	Amperios (Pico)	Tiempo (s)	Tensión (V)	Potencia requerida PICO (W)	Energía (W*s)	EDP (J * s)
121x10 ⁶	0,27	1,24	237	63,99	79,35	98,39
100x10 ⁶	0,27	0,97	237	63,99	62,07	60,21
64x10 ⁶	0,27	0,59	237	63,99	37,75	22,27
36x10 ⁶	0,25	0,33	237	59,25	19,55	6,45
16x10 ⁶	0,21	0,15	237	49,77	7,47	1,12

Tabla 32. Resultados del producto de matrices en el PC de referencia (simple precisión).

Dimensiones de las matrices	Amperios (Pico)	Tiempo (s)	Tensión (V)	Potencia requerida PICO (W)	Energía (W*s)	EDP (J * s)
11180x11180	0,30	31,17	237	71,1	2216,19	69078,55
10000x10000	0,30	22,62	237	71,1	1608,28	36379,34
8000x8000	0,30	11,55	237	71,1	821,21	9484,92
6000x6000	0,30	5,42	237	71,1	385,36	2088,66
4000x4000	0,30	1,47	237	71,1	104,52	153,64
2000x2000	0,25	0,19	237	59,25	11,26	2,14

Tabla 33. Resultados de aplicar la FFT en el PC de referencia (simple precisión).

Tamaño del vector	Amperios (Pico)	Tiempo (s)	Tensión (V)	Potencia requerida PICO (W)	Energía (W*s)	EDP (J * s)
121x10 ⁶	0,25	0,66	237	59,25	39,11	25,81
100x10 ⁶	0,25	0,59	237	59,25	34,96	20,62
64x10 ⁶	0,25	0,30	237	59,25	17,78	5,33
36x10 ⁶	0,24	0,17	237	56,88	9,67	1,64
16x10 ⁶	0,21	0,07	237	49,77	3,48	0,24

Las pruebas que han requerido del uso de la CPU del prototipo, se han capturado con una tensión de **240V**. A continuación se muestran los resultados obtenidos.

Tabla 34. Resultados del benchmark interno de Matlab en el prototipo.

Bench Interno de MatLab	LU (s)	FFT (s)	ODE (s)	Sparse (s)	2-D (s)	3-D (s)	Amperios PICO	Potencia (W)	Energía (W*s)	EDP (J * s)
1° Vez	0,0621	0,0744	0,1315	0,1542	0,2221	0,1104	0,38	91,2	68,83	51,94
2° Vez	0,0631	0,0769	0,1325	0,1524	0,2177	0,1320	0,38	91,2	70,64	54,72
3° Vez	0,0632	0,0989	0,1316	0,1520	0,2077	0,1079	0,38	91,2	69,43	52,86

Tabla 35. Resultados del producto de matrices en el prototipo (doble precisión).

Dimensiones de las matrices	Amperios (Pico)	Tiempo (s)	Tensión (V)	Potencia requerida PICO (W)	Energía (W*s)	EDP (J * s)
11180x11180	0,38	40,67	240	91,2	3709,1	150849,1
10000x10000	0,38	28,88	240	91,2	2633,85	76065,59
8000x8000	0,38	14,87	240	91,2	1356,14	20165,8
6000x6000	0,38	6,31	240	91,2	575,47	3631,22
4000x4000	0,38	1,91	240	91,2	174,19	332,7
2000x2000	0,33	0,25	240	79,2	19,8	4,95

Tabla 36. Resultados de aplicar la FFT en el prototipo (doble precisión).

Tamaño del vector	Amperios (Pico)	Tiempo (s)	Tensión (V)	Potencia requerida PICO (W)	Energía (W*s)	EDP (J * s)
121x10 ⁶	0,34	0,61	240	81,6	49,78	30,37
100x10 ⁶	0,33	0,48	240	79,2	38,02	18,25
64x10 ⁶	0,32	0,28	240	76,8	21,5	6,02
36x10 ⁶	0,29	0,15	240	69,6	10,44	1,57
16x10 ⁶	0,26	0,06	240	62,4	3,74	0,22

Tabla 37. Resultados del producto de matrices en el prototipo (simple precisión).

Dimensiones de las matrices	Amperios (Pico)	Tiempo (s)	Tensión (V)	Potencia requerida PICO (W)	Energía (W*s)	EDP (J * s)
11180x11180	0,38	19,78	240	91,2	1803,94	35681,93
10000x10000	0,38	14,15	240	91,2	1290,48	18260,29
8000x8000	0,38	7,28	240	91,2	663,94	4833,48
6000x6000	0,38	3,11	240	91,2	283,63	882,09
4000x4000	0,37	0,94	240	88,8	225,6	212,06
2000x2000	0,31	0,13	240	74,4	9,67	1,26

Tabla 38. Resultados de aplicar la FFT en el prototipo (simple precisión).

Tamaño del vector	Amperios (Pico)	Tiempo (s)	Tensión (V)	Potencia requerida PICO (W)	Energía (W*s)	EDP (J * s)
121x10 ⁶	0,32	0,32	240	76,8	24,58	7,87
100x10 ⁶	0,32	0,25	240	76,8	19,2	4,8
64x10 ⁶	0,3	0,16	240	72	11,52	1,84
36x10 ⁶	0,29	0,09	240	69,6	6,26	0,56
16x10 ⁶	0,28	0,03	240	67,2	2,02	0,06

Los resultados obtenidos, que han requerido del uso de la GPU del prototipo se muestran a continuación.

Tabla 39. Resultados del producto de matrices en la GPU del prototipo (doble precisión).

Dimensiones de las matrices	Amperios (Pico)	Tiempo (s)	Tensión (V)	Potencia requerida PICO (W)	Energía (W*s)	EDP (J * s)
6000x6000	0,71	5,12	240	170,4	872,45	4466,93
4000x4000	0,72	1,55	240	172,8	267,84	415,15
2000x2000	0,60	0,21	240	144	30,24	6,35

Tabla 40. Resultados de aplicar la FFT en la GPU del prototipo (doble precisión).

Tamaño del vector	Amperios (Pico)	Tiempo (s)	Tensión (V)	Potencia requerida PICO (W)	Energía (W*s)	EDP (J * s)
16x10 ⁶	0,53	0,16	240	127,2	20,35	3,26

Tabla 41. Resultados del producto de matrices en la GPU del prototipo (simple precisión).

Dimensiones de las matrices	Amperios (Pico)	Tiempo (s)	Tensión (V)	Potencia requerida PICO (W)	Energía (W*s)	EDP (J * s)
8000x8000	0,85	2,28	240	204	465,12	1060,47
6000x6000	0,83	1	240	199,2	199,2	199,2
4000x4000	0,68	0,32	240	163,2	52,22	16,71
2000x2000	0,50	0,05	240	120	6,00	0,3

Tabla 42. Resultados de aplicar la FFT en la GPU del prototipo (simple precisión).

Tamaño del vector	Amperios (Pico)	Tiempo (s)	Tensión (V)	Potencia requerida PICO (W)	Energía (W*s)	EDP (J * s)
36x10 ⁶	0,55	0,14	240	132	18,48	2,59
16x10 ⁶	0,49	0,06	240	117,6	7,06	0,42

ANEXO VIII

Las pruebas con el PC de referencia se han realizado con una tensión de **237V**. Primero se han realizado las pruebas para una matriz de 200x200 (como marcaba el código por defecto) y luego se han realizado las pruebas para una matriz de un tamaño 2048x2048. Los resultados obtenidos se muestran a continuación.

Tabla 43. Resultados para una matriz de 200x200 en 1 core 1º hilo del PC de referencia.

Repeticiones	Tiempo (s)	MFLOPS	Amperios (PICO)	Consumo (W)
1024	0,51	2929,777778	0,23	54,51
2048	1,03	3160,209738		
4096	2,06	3232,858238		
8192	4,12	3187,067044		
16384	8,24	3173,581570		
32768	16,47	3223,595033		

Tabla 44. Resultados para una matriz de 200x200 en 2 cores 1º hilo del PC de referencia.

Repeticiones	Tiempo (s)	MFLOPS	Amperios (PICO)	Consumo (W)
1024	0,52	3125,096296	0,28	66,36
2048	1,02	3090,754579		
4096	2,07	3125,096296		
8192	4,81	2743,986992		
16384	10,70	2409,067809		

Tabla 45. Resultados para una matriz de 200x200 en 2 cores 2º hilo del PC de referencia.

Repeticiones	Tiempo (s)	MFLOPS	Amperios (PICO)	Consumo (W)
1024	0,51	3125,096296	0,28	66,36
2048	1,03	3515,733333		
4096	2,07	3214,384762		
8192	4,81	2672,291370		
16384	10,67	2448,388828		

Tabla 46. Resultados para una matriz de 200x200 en 3 cores 1º hilo del PC de referencia.

Repeticiones	Tiempo (s)	MFLOPS	Amperios (PICO)	Consumo (W)
1024	0,93	1674,158730	0,29	68,73
2048	1,85	1736,164609		
4096	3,74	1649,610948		
8192	7,46	1676,653751		
16384	14,92	1677,904052		

Tabla 47. Resultados para una matriz de 200x200 en 3 cores 2º hilo del PC de referencia.

Repeticiones	Tiempo (s)	MFLOPS	Amperios (PICO)	Consumo (W)
1024	0,93	1714,991870	0,29	68,73
2048	1,86	1736,164609		
4096	3,73	1709,779129		
8192	7,45	1689,241241		
16384	14,93	1690,510393		

Tabla 48. Resultados para una matriz de 200x200 en 3 cores 3º hilo del PC de referencia.

Repeticiones	Tiempo (s)	MFLOPS	Amperios (PICO)	Consumo (W)
1024	0,52	3348,317460	0,29	68,73
2048	1,03	3125,096296		
4096	2,06	3160,209738		
8192	4,13	3116,439520		
16384	14,93	3133,801300		

Tabla 49. Resultados para una matriz de 200x200 en 4 cores 1º hilo del PC de referencia.

Repeticiones	Tiempo (s)	MFLOPS	Amperios (PICO)	Consumo (W)
1024	0,94	1694,329317	0,30	71,1
2048	1,87	1674,158730		
4096	3,75	1679,156219		
8192	7,50	1671,671124		
16384	14,92	1698,165535		

Tabla 50. Resultados para una matriz de 200x200 en 4 cores 2º hilo del PC de referencia.

Repeticiones	Tiempo (s)	MFLOPS	Amperios (PICO)	Consumo (W)
1024	0,93	1714,991870	0,30	71,1
2048	1,86	1625,772640		
4096	3,73	1684,183633		
8192	7,46	1659,343166		
16384	14,91	1659,343166		

Tabla 51. Resultados para una matriz de 200x200 en 4 cores 3º hilo del PC de referencia.

Repeticiones	Tiempo (s)	MFLOPS	Amperios (PICO)	Consumo (W)
1024	0,94	1736,164609	0,30	71,1
2048	1,87	1664,252465		
4096	3,75	1704,597980		
8192	7,49	1669,190900		
16384	14,93	1684,183633		

Tabla 52. Resultados para una matriz de 200x200 en 4 cores 4º hilo del PC de referencia.

Repeticiones	Tiempo (s)	MFLOPS	Amperios (PICO)	Consumo (W)
1024	0,93	1674,158730	0,30	71,1
2048	1,87	1644,787524		
4096	3,72	1736,164609		
8192	7,46	1661,794190		
16384	14,91	1680,410256		

Tabla 53. Resultados para una matriz de 2048x2048 en 1 core 1º hilo del PC de referencia.

Repeticiones	Tiempo (s)	MFLOPS	Amperios (PICO)	Consumo (W)
4	0,88	6756,941503	0,23	54,51
8	1,77	6837,381283		
16	3,54	6796,923405		
32	7,07	6766,892816		
64	14,15	6796,923405		

Tabla 54. Resultados para una matriz de 2048x2048 en 2 cores 1º hilo del PC de referencia.

Repeticiones	Tiempo (s)	MFLOPS	Amperios (PICO)	Consumo (W)
2	0,86	3418,690641	0,26	61,62
4	1,71	3459,879685		
8	3,44	3418,690641		
16	6,88	3434,021093		
32	13,41	3520,858408		

Tabla 55. Resultados para una matriz de 2048x2048 en 2 cores 2º hilo del PC de referencia.

Repeticiones	Tiempo (s)	MFLOPS	Amperios (PICO)	Consumo (W)
2	0,84	3418,690641	0,26	61,62
4	1,68	3502,073340		
8	3,36	3470,332494		
16	6,73	3486,130669		
32	13,46	3502,073340		

Tabla 56. Resultados para una matriz de 2048x2048 en 3 cores 1º hilo del PC de referencia.

Repeticiones	Tiempo (s)	MFLOPS	Amperios (PICO)	Consumo (W)
1	0,80	1817,531733	0,26	61,62
2	1,60	1840,833422		
4	3,15	1864,740350		
8	6,33	1858,705591		
16	12,26	1916,063479		

Tabla 57. Resultados para una matriz de 2048x2048 en 3 cores 2º hilo del PC de referencia.

Repeticiones	Tiempo (s)	MFLOPS	Amperios (PICO)	Consumo (W)
1	0,78	1864,740350	0,26	61,62
2	1,55	1889,276407		
4	3,07	1908,106404		
8	6,12	1917,662864		
16	12,26	1917,662864		

Tabla 58. Resultados para una matriz de 2048x2048 en 3 cores 3º hilo del PC de referencia.

Repeticiones	Tiempo (s)	MFLOPS	Amperios (PICO)	Consumo (W)
1	0,97	3022,842251	0,26	61,62
2	1,92	3071,337047		
4	3,86	3038,836126		
8	7,72	3046,896699		
16	11,68	4023,397742		

Tabla 59. Resultados para una matriz de 2048x2048 en 4 cores 1º hilo del PC de referencia.

Repeticiones	Tiempo (s)	MFLOPS	Amperios (PICO)	Consumo (W)
1	0,89	1650,402379	0,26	61,62
2	1,78	1650,402379		
4	3,55	1659,942277		
8	7,11	1652,777058		
16	13,16	1786,438655		

Tabla 60. Resultados para una matriz de 2048x2048 en 4 cores 2º hilo del PC de referencia.

Repeticiones	Tiempo (s)	MFLOPS	Amperios (PICO)	Consumo (W)
1	0,85	1709,345321	0,26	61,62
2	1,72	1729,939843		
4	3,43	1709,345321		
8	6,82	1722,159004		
16	13,46	1745,714370		

Tabla 61. Resultados para una matriz de 2048x2048 en 4 cores 3º hilo del PC de referencia.

Repeticiones	Tiempo (s)	MFLOPS	Amperios (PICO)	Consumo (W)
1	0,84	1751,036670	0,26	61,62
2	1,69	1729,939843		
4	3,35	1745,714370		
8	6,72	1751,036670		
16	13,41	1752,372320		

Tabla 62. Resultados para una matriz de 2048x2048 en 4 cores 4º hilo del PC de referencia.

Repeticiones	Tiempo (s)	MFLOPS	Amperios (PICO)	Consumo (W)
1	0,85	1729,939843	0,26	61,62
2	1,69	1719,580921		
4	3,39	1735,166247		
8	6,76	1740,424326		
16	13,47	1744,388847		

Las pruebas con el prototipo se han realizado con una tensión de **241.5V**. Como sucedía con el PC de referencia, primero se muestran los datos para una matriz de 200x200 y luego para un tamaño de matriz de 2048x2048.

Tabla 63. Resultados para una matriz de 200x200 en 1 core 1º hilo del prototipo.

Repeticiones	Tiempo (s)	MFLOPS	Amperios (PICO)	Consumo (W)
2048	0,97	3388,658635	0,31	74,87
4096	1,93	3368,367265		
8192	3,86	3358,312438		
16384	7,72	3429,983740		
32768	15,46	3383,562907		

Tabla 64. Resultados para una matriz de 200x200 en 2 cores 1º hilo del prototipo.

Repeticiones	Tiempo (s)	MFLOPS	Amperios (PICO)	Consumo (W)
1024	0.50	3348,317460	0,36	86,94
2048	1.00	3308,925490		
4096	2,00	3328,504931		
8192	3,99	3242,174832		
16384	7,99	3294,391410		
32768	15,98	3268,074558		

Tabla 65. Resultados para una matriz de 200x200 en 2 cores 2º hilo del prototipo.

Repeticiones	Tiempo (s)	MFLOPS	Amperios (PICO)	Consumo (W)
1024	0,50	3196,121212	0,36	86,94
2048	0,99	3384,317460		
4096	2,00	3409,195960		
8192	3,99	3270,449612		
16384	7,99	3279,984451		
32768	15,97	3291,989468		

Tabla 66. Resultados para una matriz de 200x200 en 3 cores 1º hilo del prototipo.

Repeticiones	Tiempo (s)	MFLOPS	Amperios (PICO)	Consumo (W)
1024	0,57	2929,777778	0,37	89,36
2048	1,14	3057,159420		
4096	2,29	2855,417936		
8192	4,56	2892,119966		
16384	9,16	2892,119966		
32768	18,29	2864,505835		

Tabla 67. Resultados para una matriz de 200x200 en 3 cores 2º hilo del prototipo.

Repeticiones	Tiempo (s)	MFLOPS	Amperios (PICO)	Consumo (W)
1024	0,57	2992,113475	0,37	89,36
2048	1,14	2992,113475		
4096	2,28	2899,573883		
8192	4,57	2914,597582		
16384	9,15	2899,473883		
32768	18,29	2871,817991		

Tabla 68. Resultados para una matriz de 200x200 en 3 cores 3º hilo del prototipo.

Repeticiones	Tiempo (s)	MFLOPS	Amperios (PICO)	Consumo (W)
1024	0,57	2604,246914	0,37	89,36
2048	1,14	2812,586667		
4096	2,28	2653,383648		
8192	4,57	2937,427328		
16384	9,14	2826,720268		
32768	18,29	2918,377864		

Tabla 69. Resultados para una matriz de 200x200 en 4 cores 1º hilo del prototipo.

Repeticiones	Tiempo (s)	MFLOPS	Amperios (PICO)	Consumo (W)
1024	0,66	2424,643678	0,37	89,36
2048	1,32	2343,822222		
4096	2,65	2363,518207		
8192	5,31	2435,139971		
16384	10,61	2464,479007		

Tabla 70. Resultados para una matriz de 200x200 en 4 cores 2º hilo del prototipo.

Repeticiones	Tiempo (s)	MFLOPS	Amperios (PICO)	Consumo (W)
1024	0,66	2163,528205	0,37	89,36
2048	1,33	2445,727536		
4096	2,65	2592,245776		
8192	5,31	2373,490858		
16384	10,62	2414,237482		

Tabla 71. Resultados para una matriz de 200x200 en 4 cores 3º hilo del prototipo.

Repeticiones	Tiempo (s)	MFLOPS	Amperios (PICO)	Consumo (W)
1024	0,66	2704,410256	0,37	89,36
2048	1,32	2533,861862		
4096	2,65	2435,139971		
8192	5,31	2445,727536		
16384	10,60	2456,407569		

Tabla 72. Resultados para una matriz de 200x200 en 4 cores 4º hilo del prototipo.

Repeticiones	Tiempo (s)	MFLOPS	Amperios (PICO)	Consumo (W)
1024	0,66	2383,548023	0,37	89,36
2048	1,32	2383,548023		
4096	2,66	2489,014749		
8192	5,31	2419,429391		
16384	10,60	2480,782065		

Tabla 73. Resultados para una matriz de 2048x2048 en 1 core 1º hilo del prototipo.

Repeticiones	Tiempo (s)	MFLOPS	Amperios (PICO)	Consumo (W)
4	0,52	11486,800555	0,33	79,7
8	1,05	11373,069856		
16	2,09	11544,523171		
32	4,19	11691,400056		
64	8,38	11646,946063		
128	16,75	11624,845596		

Tabla 74. Resultados para una matriz de 2048x2048 en 2 cores 1º hilo del prototipo.

Repeticiones	Tiempo (s)	MFLOPS	Amperios (PICO)	Consumo (W)
4	0,88	6837,812832	0,36	86,94
8	1,75	6796,923405		
16	3,48	6857,791376		
32	6,94	6930,196413		
64	13,87	6898,979312		

Tabla 75. Resultados para una matriz de 2048x2048 en 2 cores 2º hilo del prototipo.

Repeticiones	Tiempo (s)	MFLOPS	Amperios (PICO)	Consumo (W)
4	0,88	6756,941503	0,36	86,94
8	1,75	6878,323685		
16	3,47	6857,791376		
32	6,94	6888,636015		
64	13,88	6893,803784		

Tabla 76. Resultados para una matriz de 2048x2048 en 3 cores 1º hilo del prototipo.

Repeticiones	Tiempo (s)	MFLOPS	Amperios (PICO)	Consumo (W)
2	0,64	4631,774417	0,36	86,94
4	1,27	4631,774417		
8	2,54	4594,720222		
16	5,06	4727,078418		
32	10,13	4727,078418		

Tabla 77. Resultados para una matriz de 2048x2048 en 3 cores 2º hilo del prototipo.

Repeticiones	Tiempo (s)	MFLOPS	Amperios (PICO)	Consumo (W)
2	0,64	4631,774417	0,36	86,94
4	1,27	4594,720222		
8	2,53	4727,078418		
16	5,06	4698,077936		
32	10,13	4674,181304		

Tabla 78. Resultados para una matriz de 2048x2048 en 3 cores 3º hilo del prototipo.

Repeticiones	Tiempo (s)	MFLOPS	Amperios (PICO)	Consumo (W)
2	0,64	4558,254188	0,36	86,94
4	1,27	4594,720222		
8	2,53	4650,526540		
16	5,07	4659,959657		
32	10,15	4674,181304		

Tabla 79. Resultados para una matriz de 2048x2048 en 4 cores 1º hilo del prototipo.

Repeticiones	Tiempo (s)	MFLOPS	Amperios (PICO)	Consumo (W)
2	0,85	3545,308813	0,36	86,94
4	1,69	3502,073340		
8	3,39	3491,428740		
16	6,77	3512,783044		
32	13,44	3534,400171		

Tabla 80. Resultados para una matriz de 2048x2048 en 4 cores 2º hilo del prototipo.

Repeticiones	Tiempo (s)	MFLOPS	Amperios (PICO)	Consumo (W)
2	0,85	3545,308813	0,36	86,94
4	1,71	3502,073340		
8	3,42	3491,428740		
16	6,85	3465,098207		
32	13,33	3567,329365		

Tabla 81. Resultados para una matriz de 2048x2048 en 4 cores 3º hilo del prototipo.

Repeticiones	Tiempo (s)	MFLOPS	Amperios (PICO)	Consumo (W)
2	0,84	3589,625173	0,36	86,94
4	1,67	3502,073340		
8	3,36	3534,400171		
16	6,69	3539,846088		
32	13,39	3559,039676		

Tabla 82. Resultados para una matriz de 2048x2048 en 4 cores 4º hilo del prototipo.

Repeticiones	Tiempo (s)	MFLOPS	Amperios (PICO)	Consumo (W)
2	0,85	3459,879685	0,36	86,94
4	1,69	3502,073340		
8	3,39	3502,073340		
16	6,77	3512,783044		
32	13,45	3531,683491		

ANEXO IX

La captura de resultados mediante la herramienta Phoronix 3.8.0, se ha producido con una tensión de **237V** que alimentaba al PC de referencia. Los resultados obtenidos se muestran a continuación.

Tabla 83. Resultados del test Phoronix. Primera parte en el PC de referencia.

Benchmark	Amperios (PICO)	MFLOPS	Tensión (V)	Consumo (W)	MFLOPS/W
FFTE	0,22	4747,04	237	52,14	91,04
Himeno Benchmark	0,22	1320,64	237	52,14	25,33
Java SciMark	0,22	2712,55	237	52,14	52,02
SciMark	0,24	1151,94	237	56,88	20,25

Tabla 84. Resultados del test Phoronix. Segunda parte en el PC de referencia.

Benchmark	Amperios (PICO)	Tiempo (s)	Tensión (V)	Consumo (W)	Energía (W*s)	EDP (J * s)
C-Ray	0,27	122,45	237	63,99	7835,58	959466,22
Crafty	0,22	95,88	237	52,14	4999,18	479321,69
FLAC Audio Encoding	0,22	7,75	237	52,14	404,09	3131,66
GnuPG	0,21	10,59	237	49,77	527,06	5581,61
Gzip Compression	0,24	14,86	237	56,88	845,24	12560,22
Minion	0,22	135,55	237	52,14	7067,58	958010,06
Parallel BZIP2 Compression	0,27	22,15	237	63,99	1417,38	31394,93
Sample Pi Program	0,22	4,53	237	52,14	236,19	1069,96
Timed HMMer Search	0,29	25,96	237	68,73	1784,23	46318,63
Timed MAFFT Alignment	0,27	13,43	237	63,99	859,39	11541,55
Timed MrBayes Analysis	0,25	44,17	237	59,25	2617,07	115596,09

A continuación se muestran los datos capturados al utilizar el benchmark de Phoronix test suite 3.8.0 en el prototipo, con una tensión de **241V**.

Tabla 85. Resultados del test Phoronix. Primera parte en el prototipo.

Benchmark	Amperios (PICO)	MFLOPS	Tensión (V)	Consumo (W)	MFLOPS/W
FFTE	0,35	5066,63	241	84,35	60,07
Himeno Benchmark	0,36	1478,40	241	86,76	17,04
Java SciMark	0,35	2689,17	241	84,35	31,88
SciMark	0,36	1764,64	241	86,76	20,34

Tabla 86. Resultados del test Phoronix. Segunda parte en el prototipo.

Benchmark	Amperios (PICO)	Tiempo (s)	Tensión (V)	Consumo (W)	Energía (W*s)	EDP (J * s)
C-Ray	0,39	86,11	241	93,99	8093,48	696929,47
Crafty	0,35	88,27	241	84,35	7445,57	657220,86
FLAC Audio Encoding	0,34	7,23	241	81,94	592,43	4283,24
GnuPG	0,35	9,95	241	84,35	839,28	8350,86
Gzip Compression	0,35	14,58	241	84,35	1229,82	17930,82
Minion	0,35	125,54	241	84,35	10589,3	1329380,6
Parallel BZIP2 Compression	0,38	15,15	241	91,58	1387,44	21019,67
Sample Pi Program	0,35	4,26	241	84,35	359,33	1530,75
Timed HMMer Search	0,40	20,71	241	96,4	1996,44	41346,36
Timed MAFFT Alignment	0,38	10,40	241	91,58	952,43	9905,29
Timed MrBayes Analysis	0,36	26,76	241	86,76	2321,70	62128,63

ANEXO X

A continuación se muestran los resultados obtenidos al utilizar los códigos del Área de TSC en el PC de referencia.

Tabla 87. Resultados al ejecutar FMM en la CPU del PC de referencia sin uso de la FFT.

Hilos en ejecución	Amperios (PICO)	Tensión (V)	Tiempo (s)	Consumo (W)	Energía (W*s)	EDP (J * s)
1	0,23	238	1087,37	54,74	59522,63	64723126,32
2	0,27	240	644,783	64,8	41781,94	26940283,59
4	0,27	240	556,57	64,8	36065,74	20073106,69

Tabla 88. Resultados al ejecutar FMM en la CPU del PC de referencia con uso de la FFT.

Hilos en ejecución	Amperios (PICO)	Tensión (V)	Tiempo (s)	Consumo (W)	Energía (W*s)	EDP (J * s)
1	0,23	238	566,223	54,74	30995,05	17550108,51
2	0,28	240	321,09	67,2	21577,25	6928238,56
4	0,29	240	283,581	69,6	19737,24	5597105,58

Tabla 89. Resultados al ejecutar MST en la CPU del PC de referencia.

Hilos en ejecución	Amperios (PICO)	Tensión (V)	Tiempo (s)	Consumo (W)	Energía (W*s)	EDP (J * s)
1	0,22	238	3220	52,36	168599,2	542889424
2	0,26	237	1729	61,62	106540,98	184209354,42
4	0,29	240	1148	69,6	79900,8	91726118,4

Los resultados obtenidos, utilizando los códigos del Área de TSC en el prototipo se muestran a continuación.

Tabla 90. Resultados al ejecutar FMM en la CPU del prototipo sin uso de la FFT.

Hilos en ejecución	Amperios (PICO)	Tensión (V)	Tiempo (s)	Consumo (W)	Energía (W*s)	EDP (J * s)
1	0,335	238,5	828,717	79,9	66214,49	54873073,51
2	0,39	238,5	477,878	93,02	44452,21	21242733,21
4	0,38	238,5	347,987	90,63	31538,06	10974834,89

Tabla 91. Resultados al ejecutar FMM en la CPU del prototipo con uso de la FFT.

Hilos en ejecución	Amperios (PICO)	Tensión (V)	Tiempo (s)	Consumo (W)	Energía (W*s)	EDP (J * s)
1	0,31	238,5	519,396	73,94	38401,54	19945606,27
2	0,36	238,5	290,936	85,86	24979,76	7467511,46
4	0,36	238,5	193,62	85,86	16624,21	3218779,54

Tabla 92. Resultados al ejecutar MST en la CPU del prototipo.

Hilos en ejecución	Amperios (PICO)	Tensión (V)	Tiempo (s)	Consumo (W)	Energía (W*s)	EDP (J * s)
1	0,3	237,5	2995	71,25	213393,75	639114281,3
2	0,34	237,5	1567	80,75	126535,25	198280736,8
4	0,34	237,5	1032	80,75	83334	86000688

Los resultados obtenidos, utilizando los códigos del Área de TSC en la GPU del prototipo se muestran a continuación.

Tabla 93. Resultados al ejecutar los códigos del área de TSC en la GPU del prototipo.

Código TSC	Amperios (PICO)	Tensión (V)	Tiempo (s)	Consumo (W)	Energía (W*s)	EDP (J * s)
MST GPU	0,86	239,5	18	205,97	3707,46	66734,28
FMM sin FFT	0,87	239,5	28,321	208,37	5901,11	167125,20
FMM con FFT	0,83	239,5	24,92	198,79	4953,72	123446,76