

UNIVERSIDAD DE CANTABRIA

DPTO. DE ELECTRÓNICA Y COMPUTADORES.



**IMPACTO DEL SUBSISTEMA DE  
COMUNICACIÓN EN EL RENDIMIENTO DE  
LOS COMPUTADORES PARALELOS:  
DESDE EL HARDWARE HASTA LAS  
APLICACIONES.**

Presentada por:

**Valentin Puente Varona**

Dirigida por:

**Ramón Bevide Palacio.**

**SANTANDER, OCTUBRE DE 1999**

UNIVERSIDAD DE CANTABRIA

DPTO. DE ELECTRÓNICA Y COMPUTADORES.



**IMPACTO DEL SUBSISTEMA DE  
COMUNICACIÓN EN EL RENDIMIENTO DE  
LOS COMPUTADORES PARALELOS:  
DESDE EL HARDWARE HASTA LAS  
APLICACIONES.**

Memoria presentada para optar al  
grado de Doctor en Ciencias Físicas  
por:

**Valentin Puente Varona**

El Director:

**Dr. D. Ramón Beivide Palacio**

Declaro: Que el presente trabajo ha sido  
realizado en el Departamento de Elec-  
trónica y Computadores bajo mi dirección,  
reuniendo las condiciones exigidas a los  
trabajos de Doctorado

UNIVERSIDAD DE CANTABRIA

DPTO. DE ELECTRÓNICA Y COMPUTADORES.

**IMPACTO DEL SUBSISTEMA DE  
COMUNICACIÓN EN EL RENDIMIENTO DE  
LOS COMPUTADORES PARALELOS:  
DESDE EL HARDWARE HASTA LAS  
APLICACIONES.**

---

*A **Valentin** y **Esther**. Sin su constante animo y apoyo frente a todas las dificultades, este trabajo no habría llegado ni tan siquiera a iniciarse.*

## **Agradecido a:**

Los *jefes* y amigos, **José Angel** y **Món**, por su constante ayuda y por haberme enseñado los secretos del *vudú*. A **Cruz**, por las constructivas discusiones por *emilio* y su ayuda con la lengua del *Shakespeare*.

Toda la gente de ATC, en especial: a SysOp **Borja**, el maestro del Perl y el terror de la *quota*, a virtual **Prelle**, el monstruo de la POO, a Jota **Vinuesa**, el dominador de las *cnumas*, y a **Carmen**, la experta del transistor. Gracias también por las cervezas y las conversaciones pseudo-filosóficas.

Las pobres máquinas, por haber soportado duramente las horas y horas de simulación sin liarla demasiadas veces.



# Tabla de Contenidos

<b>Prólogo</b> .....	<i>I</i>
----------------------	----------

## **Capítulo 1**

<b>Introducción</b> .....	1
---------------------------	---

1.1	Introducción.....	2
1.2	Taxonomía de los Computadores Paralelos.....	5
1.2.1	Computadores de Paso de Mensajes.....	5
1.2.2	Computadores de Memoria Compartida .....	7
1.3	Arquitecturas S-SMP.....	11
1.4	Importancia de la Red de Interconexión.....	13
1.4.1	Características Básicas de Algunos de los Encaminadores/redes Actuales más Significativos.....	17
1.4.2	Cray T3E.....	18
1.4.3	SGI Origin 2000: SGI Spider.....	19
1.4.4	Myrinet.....	20
1.5	Motivación y Objetivos.....	22
1.6	Estructura.....	25

## **Capítulo 2**

<b>Tecnologías, Métricas y Evaluación de Rendimiento</b> .....	27
--	----

2.1	Introducción.....	28
2.2	Análisis de los costes Hardware.....	28
2.2.1	La Importancia de Los Costes Hardware Asociados.....	28
2.2.2	Tecnologías Síncrona y Asíncrona.....	29
2.2.3	Tecnologías de Implementación Hardware.....	35
2.2.4	Metodología de Diseño.....	36
2.2.5	Estimación de los Costes Hardware.....	40
2.3	Evaluación del Rendimiento.....	44
2.3.1	Cargas de Trabajo Sintéticas.....	44
2.3.2	Métricas y Normalización de los Resultados.....	47
2.3.3	Cargas Reales: Impacto del Subsistema de Comunicaciones en una Arquitectura ccNUMA.....	53
2.3.4	Sensibilidad al Throughput y a la Latencia. Análisis Inicial.....	69
2.3.5	Efectos Combinados: Correlación Latencia y Throughput y su Impacto en las Aplicaciones.....	76
2.4	Conclusiones .....	78

## **Capítulo 3**

<b>Encaminamiento: una Nueva Propuesta Para Balancear Latencia y Throughput</b> .....	81
---	----

3.1	Introducción.....	82
3.1.1	Encaminamiento Determinista.....	82
3.1.2	Encaminamiento Adaptativo.....	83

3.2	Algoritmos de Encaminamiento Adaptativos para Redes k-ary n-cube. Anomalías asociadas. . . .	87
3.3	Nueva Teoría de Evitación de Interbloqueos con Encaminamiento Completamente Adaptativo.	89
3.3.1	Condiciones Básicas . . . . .	90
3.3.2	Conceptos Generales. . . . .	91
3.3.3	Condiciones Necesarias Para una Función de Encaminamiento Libre de Interbloqueos. . . . .	95
3.3.4	Encaminamiento Adaptativo Para Redes k-ary n-cube. . . . .	98
3.4	Implicaciones Tecnológicas. . . . .	105
3.4.1	Implementación Hardware del Bubble Router. . . . .	105
3.4.2	Implementaciones Alternativas. . . . .	112
3.4.3	Sumario. . . . .	119
3.5	Análisis de Rendimiento. . . . .	119
3.5.1	Cargas Sintéticas. . . . .	119
3.5.2	Cargas Reales. . . . .	125
3.6	Conclusiones . . . . .	136

## Capítulo 4

### **Efectos de la Implementación: un Mecanismo Simple de Arbitraje Distribuido.** 139

4.1	Introducción. . . . .	140
4.2	Estructuras Consideradas. . . . .	142
4.2.1	Router Basado en la Estructura OAC. . . . .	142
4.2.2	Router basado en la Estructura SIC. . . . .	144
4.2.3	Router Basado en la Estructura SAC. . . . .	146
4.3	Diseños Hardware: Aplicación a Encaminadores de Redes k-ary n-cubes. . . . .	150
4.4	Análisis Comparativo Bajo Cargas de Trabajo Sintéticas. . . . .	154
4.4.1	Encaminador Aislado. . . . .	154
4.4.2	Red 8-ary 2-cube y 4-ary 3-cube. . . . .	159
4.5	Análisis Comparativo Bajo Cargas de Trabajo Reales. . . . .	164
4.5.1	Características de la Arquitectura. . . . .	164
4.5.2	Análisis Comparativo para FFT. . . . .	165
4.5.3	Análisis Comparativo para Radix. . . . .	169
4.5.4	Análisis Comparativo para LU. . . . .	172
4.6	Conclusiones. . . . .	175

## Capítulo 5

### **Dependencias Topológicas . . . . .**177

5.1	Introducción. . . . .	178
5.2	Redes evaluadas. Redes Directas de Bajo Grado. . . . .	179
5.2.1	Malla . . . . .	179
5.2.2	Toro . . . . .	180
5.2.3	Midimew . . . . .	181
5.2.4	Análisis Comparativo de Características Topológicas. . . . .	184
5.3	Limitaciones de las Redes Midimew y Soluciones Propuestas. . . . .	186
5.3.1	Evitación de Interbloqueos. . . . .	186
5.3.2	Complejidad del Encaminamiento. . . . .	187

5.4 Implementaciones Hardware. ....	189
5.5 Análisis de Rendimiento. ....	192
5.5.1 Cargas de Trabajo Sintéticas. ....	192
5.5.2 Análisis de Rendimiento con Cargas Reales. ....	198
5.5.3 Distribución de Datos. Localidad Espacial en la Midimew. ....	198
5.5.4 Evaluación de Prestaciones. ....	203
5.6 Conclusiones . ....	208

## Capítulo 6

<b>Alternativas Estructurales. ....</b>	<b>209</b>
6.1 Introducción. ....	210
6.2 Propuestas Arquitectónicas para redes k-ary n-cube. ....	212
6.2.1 Encaminadores con Buffer en la Entrada. ....	212
6.2.2 Encaminadores con Buffer en la Salida. ....	214
6.3 Implementación Hardware. ....	218
6.3.1 Implementación de las Memorias Multipuerto. Memorias FIFO Segmentadas. ....	219
6.3.2 Resultados de la Implementación Hardware. ....	223
6.4 Otras alternativas a la evitación del HLB. ....	226
6.4.1 Encaminador Completamente Multiplexado. ....	227
6.4.2 Encaminador Parcialmente Multiplexado. ....	231
6.4.3 Implementaciones Hardware. ....	233
6.5 Evaluación de Rendimiento. ....	236
6.5.1 Cargas Sintéticas. ....	236
6.5.2 Cargas Reales. ....	240
6.6 Conclusiones. ....	249

## Capítulo 7

<b>Conclusiones y Perspectivas Futuras . ....</b>	<b>251</b>
---	------------

7.1 Conclusiones y Principales Aportaciones del Trabajo. ....	251
7.2 Perspectivas y Líneas de Trabajo Abiertas. ....	257

## Apéndice A

<b>Herramientas e Infraestructura de Simulación. ....</b>	<b>259</b>
---	------------

A.1 Introducción. ....	260
A.2 Simulador de Redes de Interconexión. ....	260
A.2.1 Características Básicas de SICOSYS. ....	262
A.2.2 Análisis de Rendimiento y Comparación con otras Alternativas de Simulación. ....	266
A.3 Cargas de Trabajo Reales: Simulador Conducido por Ejecución. ....	270
A.3.1 Características Básicas de RSIM. ....	271
A.3.2 NETSIM frente a SICOSYS. ....	274
A.3.3 Implementación ED-SICOSYS. ....	276
A.4 Cargas Reales. Simulador Conducido por Trazas. ....	280
A.4.1 La Simulación Conducida por Traza en el Análisis de la Red de Interconexión. ....	281

A.4.2 Simulación Conducida por Trazas con Realimentación desde la Red (TDS-NF). .....	283
A.4.3 Validación de la Metodología Propuesta. ....	287
<b>Bibliografía</b> .....	<b>293</b>

# Lista de Figuras

## Capítulo 1

<b>Introducción</b> .....	1
Figura 1-1. Requerimientos de potencia de cálculo para algunas aplicaciones típicas [37]. .....	3
Figura 1-2. Evolución del rendimiento máximo alcanzado por los computadores paralelos [128]. .....	3
Figura 1-3. Roadmap del proyecto ASCI [83]. .....	4
Figura 1-4. Estructura de una arquitectura de paso de mensajes. Caso del IBM SP-2. ....	6
Figura 1-5. Arquitecturas UMA. ....	8
Figura 1-6. Estructura de una arquitectura de memoria compartida UMA. Caso del Sun E10000 Enterprise Server [29]. ....	9
Figura 1-7. Arquitectura de un sistema ccNUMA. ....	10
Figura 1-8. Estructura de una arquitectura de memoria compartida NUMA. Caso del SGI Origin 2000.10	
Figura 1-9. Arquitectura del encaminador del T3E [113]. ....	18
Figura 1-10. Arquitectura del encaminador SGI SPIDER [55]. ....	20
Figura 1-11. Ejemplo de una red de interconexión basada en Myrinet. ....	21
Figura 1-12. Estructura de una interface de red Myrinet. Procesador LANai. ....	22

## Capítulo 2

<b>Tecnologías, Métricas y Evaluación de Rendimiento.</b> .....	27
Figura 2-1. Distribución del reloj con lógica completamente síncrona. ....	30
Figura 2-2. Distribución del reloj con lógica síncrona self-timed. ....	31
Figura 2-3. Protocolo de comunicación entre encaminadores. ....	32
Figura 2-4. Ciclo de diseño completo. ....	39
Figura 2-5. Esquemático del circuito bajo estudio. ....	41
Figura 2-6. Layout del circuito bajo análisis. ....	42
Figura 2-7. Representación BNF de la latencia y el throughput. ....	51
Figura 2-8. Representación de la latencia y el throughput en formato CNF. ....	51
Figura 2-9. (a) Representación CNF de la latencia al incorporar el tiempo de inyección, (b) Detalle. ....	51
Figura 2-10. Análisis comparativo de dos redes con la representación CNF. ....	52
Figura 2-11. Análisis comparativo de dos redes con la representación T-CNF. ....	53
Figura 2-12. Desglose del tráfico originado por FFT en bytes por instrucción para el tamaño por defecto (a.- tamaño L2 =1MB, asociatividad=4 y línea de cache=64 bytes, b.- tamaño L2 =8-KB, asociatividad=4 y línea de cache=64 bytes, c.- Variando el tamaño de la línea de cache en L2 para 1 MB y 32 procesadores). ....	56
Figura 2-13. Desglose del tráfico originado por Radix en bytes por instrucción para el tamaño por defecto(a.- tamaño L2 =1MB, asociatividad=4 y línea de cache=64 bytes, b.- tamaño L2 =8-KB, asociatividad=4 y línea de cache=64 bytes, c.- Variando el tamaño de la línea de cache en L2 para 1 MB y 32 procesadores). ....	58
Figura 2-14. Patrón de comunicación del algoritmo LU. ....	59
Figura 2-15. Desglose del tráfico originado por LU en bytes por instrucción para el problema por defecto(a.- tamaño L2 =1MB, asociatividad=4 y línea de cache=64 bytes, b.- Variando el tamaño de la línea de cache en L2 para 1 MB y 32 procesadores). ....	60
Figura 2-16. Desglose del tráfico originado por Water-Nsq en bytes por instrucción en punto flotante para el problema por defecto(a.- tamaño L2 =1MB, asociatividad=4 y línea de cache=64 bytes, b.- Variando el tamaño de la línea de cache en L2 para 1 MB y 32 procesadores). ....	61
Figura 2-17. Curvas de speedup para cada una de las aplicaciones. ....	65
Figura 2-18. Descomposición del tiempo de ejecución para cada una de las aplicaciones con 16 procesadores. ....	66
Figura 2-19. Descomposición del tiempo de ejecución para cada una de las aplicaciones con 64 procesadores. ....	66
Figura 2-20. Porcentaje de mensajes vs. distancia recorrida en un toro 4x4 y 8x8. ....	68

Figura 2-21.	Porcentaje de mensajes vs. distancia recorrida en una malla 4x4 y 8x8. ....	68
Figura 2-22.	Tiempo de ejecución para un toro 4x4, normalizado al tiempo de ejecución de un Xbar 16x16. ....	71
Figura 2-23.	Tiempo de ejecución para una malla 4x4, normalizado al tiempo de ejecución de un Xbar 16x16. ....	72
Figura 2-24.	Tiempo de ejecución para un toro 8x8, normalizado al tiempo de ejecución de un Xbar 64x64. ....	72
Figura 2-25.	Tiempo de ejecución para una malla 8x8, normalizado al tiempo de ejecución de un Xbar 64x64. ....	72
Figura 2-26.	Speedup para redes toro (a) malla (b). Análisis comparativo frente al Xbar. ....	73
Figura 2-27.	Variación en el tiempo de ejecución debido a el retraso de los encaminadores para 16 procesadores (Toro 4x4). ....	73
Figura 2-28.	Variación en el tiempo de ejecución debido a el retraso de los encaminadores para 64 procesadores (Toro 8x8). ....	74
Figura 2-29.	Variación en el tiempo de ejecución debido al ancho de los enlaces para 16 procesadores (Toro 4x4). ....	75
Figura 2-30.	Variación en el tiempo de ejecución debido al ancho de los enlaces para 64 procesadores (Toro 8x8). ....	76
Figura 2-31.	Variación del tiempo de ejecución al variar de forma correlacionada la latencia de los encaminadores y su ancho de banda con un Toro 4x4. ....	78
Figura 2-32.	Variación del tiempo de ejecución al variar de forma correlacionada la latencia de los encaminadores y su ancho de banda con un Toro 4x4. ....	78

### Capítulo 3

#### **Encaminamiento: una Nueva Propuesta Para Balancear Latencia y Throughput. .... 81**

Figura 3-1.	Infrutilización de los recursos con mecanismos de encaminamiento deterministas. ....	83
Figura 3-2.	(a) Función de selección en Zig-Zag. (b) Función de selección X/Y dinámica. ....	86
Figura 3-3.	Dependencia cíclica de recursos en un anillo. ....	88
Figura 3-4.	Producción de deadlock con encaminamiento determinista (a) y con adaptativo (b) en un Toro 2-D. ....	89
Figura 3-5.	Control de flujo burbuja con información local. ....	100
Figura 3-6.	Control de flujo burbuja con información remota. ....	101
Figura 3-7.	Estado del anillo unidireccional cuando todos las colas adaptativas están bloqueados. ....	103
Figura 3-8.	Representación equivalente para la situación mostrada en la Figura 3-7. ....	103
Figura 3-9.	Ejemplo de producción de inanición en el caso de un solo canal Determinista con control de flujo basado en la Burbuja. ....	104
Figura 3-10.	Bubble Router: estructura del encaminador adaptativo con mecanismo de evitación de interbloqueo basado en la burbuja (caso de un toro bidimensional). ....	107
Figura 3-11.	Estructura de la Unidad de Decisión de Encaminamiento (RDU). ....	108
Figura 3-12.	Unidad de Control de la RDU. ....	108
Figura 3-13.	Modelo CPN de la máquina de estados de la RDU [129]. ....	110
Figura 3-14.	Estructura del pipeline del Bubble Router. ....	112
Figura 3-15.	Encaminador determinista con mecanismo de evitación de interbloqueo basado en la burbuja. ....	113
Figura 3-16.	Estructura del pipeline del encaminador burbuja-determinista. ....	113
Figura 3-17.	Encaminador determinista con mecanismo de evitación de interbloqueo basado en los canales virtuales. ....	114
Figura 3-18.	Estructura del pipeline del encaminador canales virtuales-determinista. ....	115
Figura 3-19.	Encaminador adaptativo con mecanismo de evitación de interbloqueo basado en canales virtuales. ....	116
Figura 3-20.	Estructura del pipeline del encaminador canales virtuales-adaptativo. ....	117
Figura 3-21.	Latencia y Throughput para un Toro 8x8 bajo tráfico uniforme ( $L=20$ phits). ....	123
Figura 3-22.	Latencia y Throughput para un Toro 8x8 bajo tráfico bimodal ( $L_{corto}=20$ phits, $L_{largo}=200$ phits con probabilidad 0.1). ....	123

Figura 3-23.	Latencia y Throughput para un Toro 8x8 bajo tráfico bimodal corto ( $L_{\text{corto}}=4$ phits, $L_{\text{largo}}=20$ phits con probabilidad 0.8).....	123
Figura 3-24.	Latencia y Throughput para un Toro 8x8 bajo tráfico bit-reversal ( $L=20$ phits). ....	124
Figura 3-25.	Latencia y Throughput para un Toro 8x8 bajo tráfico matriz transpuesta ( $L=20$ phits). ...	124
Figura 3-26.	Latencia y Throughput para un Toro 8x8 bajo tráfico perfect-shuffle ( $L=20$ phits). ....	124
Figura 3-27.	(a) Tiempo de ejecución normalizado para FFT con 64 procesadores, (b) Detalle.....	127
Figura 3-28.	Traza de ejecución de la red de respuestas para FFT con 64 procesadores (en términos estructurales), (a) Carga Promedio, (b) latencia promedio. ....	128
Figura 3-29.	Traza de ejecución de la red de respuestas para FFT con 64 procesadores (en términos reales), (a) Carga Promedio, (b) latencia promedio. ....	129
Figura 3-30.	(a)Tiempo de ejecución normalizado para Radix con 64 procesadores, (b) Detalle.....	130
Figura 3-31.	Traza de ejecución de la red de respuestas para Radix con 64 procesadores (en términos estructurales), (a) Carga Promedio, (b) latencia promedio. ....	132
Figura 3-32.	Traza de ejecución de la red de respuestas para Radix con 64 procesadores (en términos reales), (a) Carga Promedio, (b) latencia promedio. ....	133
Figura 3-33.	(a)Tiempo de ejecución normalizado para Lu con 64 procesadores, (b) Detalle. ....	134
Figura 3-34.	Traza de ejecución de la red de respuestas para LU con 64 procesadores, (a) Carga Promedio, (b) latencia promedio. ....	135
Figura 3-35.	Detalle de la traza de ejecución para LU con 64 procesadores, (a)latencia promedio, (b) Carga Promedio.....	136

## Capítulo 4

### Efectos de la Implementación: un Mecanismo Simple de Arbitraje Distribuido. .... 139

Figura 4-1.	Estructura parcial del router para la propuesta OAC.....	142
Figura 4-2.	Estructura interna del crossbar OAC.....	143
Figura 4-3.	Estructura parcial del router para la propuesta SIC.....	145
Figura 4-4.	Estructura interna del arbitrio en SIC.....	146
Figura 4-5.	Wave Front Arbiter 4x4 y celda de arbitrio.....	147
Figura 4-6.	Mecanismo de evitación de un posible starvation en WFA. ....	148
Figura 4-7.	Incorporación de la función de selección en la matriz del árbitro WFA. ....	150
Figura 4-8.	Multiplexación a nivel de paquete con el árbitro SAC.....	151
Figura 4-9.	Segmentación del router con arbitrio OAC.....	152
Figura 4-10.	Segmentación del router con arbitrio SIC y SAC. ....	153
Figura 4-11.	Configuración del encaminador aislado para una red 2D.....	155
Figura 4-12.	Throughput máximo con la longitud del paquete en 2-D.....	157
Figura 4-13.	Throughput máximo con la longitud del paquete en 3-D.....	157
Figura 4-14.	Latencia y Throughput para un toro 2-D de 64 nodos bajo tráfico uniforme ( $L=20$ phits). (a) Rendimiento estructural, (b) Rendimiento Tecnológico. ....	160
Figura 4-15.	Latencia y Throughput para un toro 2-D de 64 nodos bajo tráfico bimodal corto ( $L_{\text{largo}}=20$ phits, $L_{\text{corto}}=4$ phits con $P_{\text{corto}}=0.8$ ) (a) Rendimiento estructural, (b) Rendimiento Tecnológico.....	161
Figura 4-16.	Latencia y Throughput para un toro 2-D de 64 nodos bajo tráfico Bit-Reversal ( $L=20$ phits). ....	161
Figura 4-17.	Latencia y Throughput para un toro 3-D de 64 nodos bajo tráfico uniforme ( $L=20$ phits).....	162
Figura 4-18.	Latencia y Throughput para un toro 3-D de 64 nodos bajo tráfico Perfect-Shuffle ( $L=20$ phits).....	162
Figura 4-19.	Latencia y Throughput para un toro 3-D de 64 nodos bajo tráfico Bit-Reversal ( $L=20$ phits).....	163
Figura 4-20.	(a)Tiempos de ejecución de FFT para 64 procesadores en un Toro 8x8 y 4x4x4, (b) Detalle.....	166
Figura 4-21.	Evolución de la carga promedio (en terminos estructurales) sobre la red de peticiones en la ejecución de FFT: (a) redes 2-D, (b) redes 3-D. ....	168
Figura 4-22.	Evolución de la carga promedio (en terminos reales) sobre la red de respuestas en la ejecución de FFT: (a) redes 2-D, (b) redes 3-D. ....	168

Figura 4-23.	Evolución de la latencia promedio (en terminos reales) sobre la red de respuestas en la ejecución de FFT: (a) redes 2-D, (b) redes 3-D.....	168
Figura 4-24.	(a) Tiempos de ejecución de Radix para 64 procesadores en un Toro 8x8 y 4x4x4, (b) Detalle.....	169
Figura 4-25.	Evolución de la carga promedio (en términos estructurales) sobre la red de peticiones en la ejecución de Radix: (a) redes 2-D, (b) redes 3-D.....	171
Figura 4-26.	Evolución de la carga promedio (en términos reales) sobre la red de respuestas en la ejecución de Radix: (a) redes 2-D, (b) redes 3-D.....	171
Figura 4-27.	Evolución de la latencia promedio (en términos reales) sobre la red de respuestas en la ejecución de Radix: (a) redes 2-D, (b) redes 3-D.....	171
Figura 4-28.	(a)Tiempos de ejecución de Radix para 64 procesadores en un Toro 8x8 y 4x4x4, (b)Detalle.....	172
Figura 4-29.	Evolución de la carga promedio (en términos estructurales) sobre la red de peticiones en la ejecución de LU:(a) redes 2-D, (b) redes 3-D.....	173
Figura 4-30.	Evolución de la carga promedio (en términos reales) sobre la red de respuestas en la ejecución de LU: (a) redes 2-D, (b) redes 3-D.....	174
Figura 4-31.	Evolución de la latencia promedio (en términos reales) sobre la red de respuestas en la ejecución de LU: (a) redes 2-D, (b) redes 3-D.....	175

## Capítulo 5

### Dependencias Topológicas ..... 177

Figura 5-1.	(a) Grafo circulante con $N=24$ y $b=4$ and (b) su transformación en una malla con enlaces periféricos.....	182
Figura 5-2.	Construcción de la red Midimew a partir del grafo circulante isomorfo.....	183
Figura 5-3.	Redes Midimew (a) horizontal y (b) vertical para $N=14$ .....	184
Figura 5-4.	Evolución de las características topológicas de cada red, al variar el número de nodos.....	185
Figura 5-5.	Cálculo de Routing-Record para la malla, el toro y la midimew.....	188
Figura 5-6.	Router empleado en la comparativa de las 3 topologías evaluadas.....	190
Figura 5-7.	Estructura de la RDU con encaminamiento basado en tabla.....	191
Figura 5-8.	Estructura del pipeline de los encaminadores empleados.....	191
Figura 5-9.	Latencia base en nanosegundos (con una carga de 0.05% normalizada con respecto a la bisección).....	193
Figura 5-10.	Throughput máximo aceptado expresado en phits consumidos por nanosegundo.....	194
Figura 5-11.	Comparativa de latencia base y carga máxima aceptada entre las tres topologías (Redes de 16 Nodos).....	195
Figura 5-12.	Comparativa de latencia base y carga máxima aceptada entre las tres topologías (Redes de 64 Nodos).....	195
Figura 5-13.	Comparativa de latencia base y carga máxima aceptada entre las tres topologías (Redes de 256 Nodos).....	195
Figura 5-14.	Histograma de distancias recorridas para tráfico Random.....	196
Figura 5-15.	Histograma de distancias recorridas para tráfico Matriz Traspuesta.....	196
Figura 5-16.	Histograma de distancias recorridas para tráfico Perfect Shuffle.....	196
Figura 5-17.	Histograma de distancias recorridas para tráfico Bit Reversal.....	197
Figura 5-18.	Mapas de ocupación de los buffers adaptativos para el Toro 16x16 y la Midimew de 256 nodos con tráfico Matriz Traspuesta (expresada en phits).....	197
Figura 5-19.	Asignación de una matriz compartida por columnas en Toro y Malla. Ejemplo de 9 procesos sobre un Toro 3x3 o Malla 3x3.....	199
Figura 5-20.	Asignación de una matriz compartida por columnas en Midimew. Ejemplo 9 procesos.....	200
Figura 5-21.	Relación entra la numeración de nodos en el nivel de red y en el nivel de aplicación.....	201
Figura 5-22.	Histograma de distancias recorridas para cada aplicación en un Toro 8x8 (64 Procesadores).....	202
Figura 5-23.	Histograma de distancias recorridas para cada aplicación en una Malla 8x8 (64 Procesadores).....	202
Figura 5-24.	Histograma de distancias recorridas para cada aplicación en una Midimew 64 (64 Procesadores).....	202

Figura 5-25.	Tiempo de ejecución normalizado a la malla. ....	204
Figura 5-26.	Detalle del tiempo de ejecución. ....	204
Figura 5-27.	Evolución del tiempo de ejecución de FFT al variar el número de nodos de la red. ....	205
Figura 5-28.	Detalle del tiempo de ejecución. ....	205
Figura 5-29.	Doble reasignación de nodos para FFT .....	206
Figura 5-30.	Resultados de ejecución con la nueva etiquetación de nodos. ....	207
Figura 5-31.	Resultados de ejecución con la nueva etiquetación de nodos. Detalle. ....	207

## Capítulo 6

<b>Alternativas Estructurales. ....</b>	<b>209</b>	
Figura 6-1.	Diferentes Arquitecturas de Almacenamiento. ....	211
Figura 6-2.	Esquema del router determinista, basado en el algoritmo de la Burbuja, con almacenamiento en la entrada. ....	213
Figura 6-3.	Esquema del router adaptativo, basado en el algoritmo de la Burbuja, con almacenamiento en la entrada. ....	213
Figura 6-4.	Encaminador determinista con método de evitación de deadlock basado en la Burbuja y con buffers en la salida. ....	215
Figura 6-5.	Encaminador adaptativo con método de evitación de deadlock basado en la Burbuja y con buffers en la salida. ....	216
Figura 6-6.	Mapas de ocupación promedio para las colas adaptativas y deterministas en un toro 16x16 con tráfico Matriz Transpuesta. ....	217
Figura 6-7.	Encaminador adaptativo con método de evitación de deadlock basado en la Burbuja y con esquema de almacenamiento híbrido. ....	217
Figura 6-8.	FIFO Multipuerto segmentada. ....	221
Figura 6-9.	FIFO Multipuerto segmentada con multiplexación de dos canales virtuales. ....	222
Figura 6-10.	Estructura que permite reducir el número de etapas en el buffer multipuerto para el caso del encaminador adaptativo. ....	223
Figura 6-11.	Estructura del pipeline para los encaminadores con almacenamiento en la entrada (determinista y adaptativo). ....	223
Figura 6-12.	Estructura del pipeline para los encaminador determinista con almacenamiento en la salida y los caminos de salida adaptativos del encaminador híbrido. ....	224
Figura 6-13.	Estructura del pipeline para los canales de entrada deterministas en el encaminador adaptativo con almacenamiento híbrido. ....	224
Figura 6-14.	Router multiplexado con Burbuja adaptativa y almacenamiento en la entrada. ....	228
Figura 6-15.	Incremento en el número de canales virtuales para la evitación del HLB en el encaminador multiplexado. ....	229
Figura 6-16.	Estructura de la unidad de decisión de encaminamiento multiplexada. ....	231
Figura 6-17.	Encaminador adaptativo parcialmente multiplexado con buffer en la entrada. ....	233
Figura 6-18.	Ocupación en phits de los canales deterministas en el encaminador completamente multiplexado (a) y en el parcialmente multiplexado (b), para un patrón de destinos uniforme en un toro 8x8. ....	233
Figura 6-19.	Estructura del pipeline para el encaminador adaptativo multiplexado y los canales adaptativos del parcialmente multiplexado. ....	234
Figura 6-20.	Evolución de la latencia y el throughput ofrecido en toro 8x8 con tráfico uniforme (L=10 phits). ....	237
Figura 6-21.	Evolución de la latencia y el throughput ofrecido en toro 8x8 con tráfico Bit Reversal (L=10 phits). ....	237
Figura 6-22.	Evolución de la latencia y el throughput ofrecido en toro 8x8 con tráfico Matriz transpuesta (L=10 phits). ....	238
Figura 6-23.	(a)Tiempo de ejecución normalizado para FFT con 64 procesadores, (b) Detalle. ....	242
Figura 6-24.	Traza de ejecución de la red de respuestas (en términos estructurales) para FFT con 64 procesadores, (a) Carga Promedio, (b) Latencia Promedio. ....	243
Figura 6-25.	Traza de ejecución de la red de respuestas (en términos reales) para FFT con 64 procesadores, (a) Carga Promedio, (b) Latencia Promedio. ....	243
Figura 6-26.	(a)Tiempo de ejecución normalizado para Radix con 64 procesadores, (b) Detalle. ....	244

Figura 6-27.	Traza de ejecución de la red de respuestas (en términos estructurales) para Radix con 64 procesadores, (a) Carga Promedio, (b) Latencia Promedio. ....	245
Figura 6-28.	Traza de ejecución de la red de respuestas (en términos reales) para Radix con 64 procesadores, (a) Carga Promedio, (b) Latencia Promedio. ....	246
Figura 6-29.	(a)Tiempo de ejecución normalizado para LU con 64 procesadores, (b) Detalle.....	247
Figura 6-30.	Traza de ejecución en la red de respuestas (en términos estructurales) para LU con 64 procesadores, (a) Carga promedio, (b) latencia promedio.....	248
Figura 6-31.	Detalle de la traza de ejecución para LU con 64 procesadores, (a) Carga promedio, (b) latencia promedio. ....	248

## Capítulo 7

<b>Conclusiones y Perspectivas Futuras</b> .....	251
--	-----

## Apéndice A

Herramientas e Infraestructura de Simulación. ....	259
<b>Figura A-1.</b> Parámetros de entrada del experimento. ....	265
<b>Figura A-2.</b> Representación de la Evolución Temporal de la Latencia Media. ....	266
<b>Figura A-3.</b> Encaminador Prueba Utilizado en la Comparación SICOSYS Frente a Simulación Vhdl. ....	267
<b>Figura A-4.</b> (a)Latencia Promedio de SICOSYS frente a Leapfrog y error relativo, (b) Throughput.. ....	268
<b>Figura A-5.</b> Rendimiento de SICOSYS frente a Leapfrog y ganancia relativa. ....	268
<b>Figura A-6.</b> Encaminador prueba utilizado en la comparación SICOSYS frente a simulador funcional. ....	269
<b>Figura A-7.</b> (a)Latencia promedio de SICOSYS frente a simulador funcional y error relativo, (b) Throughput. ....	269
<b>Figura A-8.</b> Rendimiento de SICOSYS frente a simulador funcional y ganancia relativa. ....	270
<b>Figura A-9.</b> Arquitectura del sistema de emulado por RSIM. ....	272
<b>Figura A-10.</b> Estructura del procesador emulado. Tomado de [95]. ....	273
<b>Figura A-11.</b> Comparativa entre NETSIM y SICOSYS con un encaminador determinista basado en burbuja sobre un toro 8x8 con tráfico uniforme. ....	275
<b>Figura A-12.</b> Comparativa entre NETSIM y SICOSYS con un encaminador determinista basado en canales virtuales sobre un toro 8x8 con tráfico uniforme. ....	276
<b>Figura A-13.</b> Arquitectura de ED-SICOSYS. ....	277
<b>Figura A-14.</b> Estructura del simulador EDD-SICOSYS (Execution Driven Distributed SICOSYS ). ....	279
<b>Figura A-15.</b> Relaciones de velocidad y precisión con cada uno de los planteamientos de la simulación. ....	283
<b>Figura A-16.</b> Temporalización absoluta (a) y relativa (b). ....	284
<b>Figura A-17.</b> Aproximación inicial. ....	285
<b>Figura A-18.</b> Fase 1: extracción de la traza y validación de resultados. ....	288
<b>Figura A-19.</b> Aplicación de la metodología. ....	288
<b>Figura A-20.</b> Error relativo en valor absoluto entre la simulación conducida por trazas y por ejecución y ganancia en velocidad para FFT. ....	291
<b>Figura A-21.</b> Error relativo en valor absoluto entre la simulación conducida por trazas y por ejecución y ganancia en velocidad para Radix. ....	291

# Lista de Tablas

## Capítulo 1

### Introducción ..... 1

Tabla 1-1. Características de la red de interconexión de varios sistema comerciales y prototipos académicos [37] .....	17
--	----

## Capítulo 2

### Tecnologías, Métricas y Evaluación de Rendimiento. .... 27

Tabla 2-1. Estimación de Costes Hardware: desde el nivel lógico al nivel físico.....	42
Tabla 2-2. Condiciones de operación de la librería tecnológica ECPD07. ....	43
Tabla 2-3. Desglose del número de instrucciones para FFT con el tamaño por defecto (64K complejos) para un caso de 32 procesadores [137]. ....	54
Tabla 2-4. Desglose del número de instrucciones para Radix con el tamaño por defecto para el caso de 32 procesadores. (*Las pausas son puntos de sincronización basados en flags) .....	57
Tabla 2-5. Desglose del número de instrucciones para LU con el tamaño por defecto para el caso de 32 procesadores. ....	59
Tabla 2-6. Desglose del número de instrucciones para Water con el tamaño por defecto para el caso de 32 procesadores. ....	61
Tabla 2-7. Principales Parámetros del Procesador. ....	62
Tabla 2-8. Principales Parámetros de la Jerarquía de Memoria. ....	63
Tabla 2-9. Número de referencias con un único procesador. ....	64
Tabla 2-10. Número de referencias con 16 procesadores en un Crossbar.....	64
Tabla 2-11. Número de referencias con 64 procesadores en un Crossbar.....	65
Tabla 2-12. Tiempo de ejecución sobre un Crossbar (en millones de ciclos de CPU). ....	65
Tabla 2-13. Número de paquetes manejados por cada una de las redes.....	67
Tabla 2-14. Distancia media recorrida por los mensajes en cada aplicación. ....	69
Tabla 2-15. Encaminadores base empleados en el análisis de tiempos de ejecución Malla vs Toro. ....	70
Tabla 2-16. Variación del Tiempo de Ejecución (Tanto por 1). ....	71
Tabla 2-17. Notación de los experimentos probados. ....	77

## Capítulo 3

### Encaminamiento: una Nueva Propuesta Para Balancear Latencia y Throughput. .... 81

Tabla 3-1. Características de tiempo y área de las colas FIFO en condiciones típicas. ....	111
Tabla 3-2. Características de todos los módulos del encaminador bajo condiciones típicas. ....	111
Tabla 3-3. Características de todos los módulos del encaminador burbuja-determinista en condiciones típicas de operación.....	113
Tabla 3-4. Características de todos los módulos del encaminador canales virtuales-determinista bajo condiciones típicas. ....	115
Tabla 3-5. Características de todos los módulos del encaminador canales virtuales-adaptativo bajo condiciones típicas. ....	117
Tabla 3-6. Características de los diseños para redes bidimensionales. ....	119
Tabla 3-7. Latencia base (ns.): Toro 8x8.....	121
Tabla 3-8. Throughput máximo alcanzado por cada propuesta para un toro 8x8. ....	122
Tabla 3-10. Velocidades relativas procesador/red en cada caso.....	126
Tabla 3-9. Resumen de los parámetros más significativos del sistema.....	126

## Capítulo 4

### Efectos de la Implementación: un Mecanismo Simple de Arbitraje Distribuido. .... 139

Tabla 4-1.	Resultados finales de tiempo y área para cada una de las implementaciones en 2-D.	153
Tabla 4-2.	Resultados finales de tiempo y área para cada una de las implementaciones en 3-D.	153
Tabla 4-3.	Rendimiento con patrón de destinos uniforme en configuración aislada.	158
Tabla 4-4.	Latencia base para los encaminadores 2-D (nanosegundos).	159
Tabla 4-5.	Latencia base para los encaminadores 3-D (nanosegundos)	159
Tabla 4-6.	Throughput máximo para encaminadores 2-D	159
Tabla 4-7.	Throughput máximo para encaminadores 3-D.	159
Tabla 4-8.	Velocidades relativas entre el procesador y la red en cada caso.	165

## Capítulo 5

### Dependencias Topológicas ..... 177

Tabla 5-1.	Comparación de las características topológicas de las tres topologías.	185
Tabla 5-2.	Características de área y tiempo para los encaminadores.	192
Tabla 5-3.	Distancias promedio recorridas por los mensajes para cada aplicación y topología.	203

## Capítulo 6

### Alternativas Estructurales. .... 209

Tabla 6-1.	Configuración de los espacios de almacenamiento para cada encaminador.	225
Tabla 6-2.	Características de área y tiempo para cada uno de los routers evaluados.	225
Tabla 6-3.	Características de área y tiempo para los routers multiplexados.	234
Tabla 6-4.	Latencia base en nanosegundos para una carga de 0.05% con respecto a la bisección.	236
Tabla 6-5.	Throughput máximo aceptado expresado en phits consumidos por nanosegundo (phits consumidos por ciclo).	237
Tabla 6-6.	Velocidades relativas red/procesador en cada caso.	241

## Capítulo 7

### Conclusiones y Perspectivas Futuras ..... 251

## Apéndice A

### Herramientas e Infraestructura de Simulación..... 259

Tabla A-1.	Ficheros SGML de simulación para encaminador determinista burbuja.	264
Tabla A-2.	Comparativa de rendimiento y precisión entre RSIM y ED-SICOSYS.	278
Tabla A-3.	Comparativa de rendimiento y precisión entre ED-SICOSYS y EDD-SICOSYS.	279
Tabla A-4.	Relaciones de disparo entre mensajes.	286
Tabla A-5.	Estructura de las entradas del fichero de traza.	287
Tabla A-6.	Número de mensajes generados por aplicación y número de procesadores.	290
Tabla A-7.	Resultados de Error y Speedup (promedios en valor absoluto).	291