

UNIVERSIDAD CARLOS III DE MADRID

DEPARTAMENTO DE INGENIERÍA TELEMÁTICA



TESIS DOCTORAL

**CONTRIBUCIONES AL DISEÑO DE REDES ÓPTICAS DE
NUEVA GENERACIÓN: PROTECCIÓN, CALIDAD DE
SERVICIO Y EFICIENCIA ENERGÉTICA**

Autor: Isaac Seoane Pujol

Ingeniero en Telecomunicación, Máster en Ingeniería Telemática

Directores: José Alberto Hernández Gutiérrez y David Larrabeiti López

Doctores en Ingeniería de Telecomunicaciones

Leganés, Abril de 2012

UNIVERSIDAD CARLOS III DE MADRID

DEPARTMENT OF TELEMATICS ENGINEERING



Ph.D. THESIS

**CONTRIBUTIONS TO NEXT GENERATION OPTICAL
NETWORKS DESIGN: RESILIENCE, QUALITY OF SERVICE
AND ENERGY EFFICIENCY**

Author: Isaac Seoane Pujol, Ms.C.

Supervisor: José Alberto Hernández Gutiérrez, Ph.D.

Supervisor: David Larrabeiti López, Ph.D.

Leganes, April 2012

**CONTRIBUCIONES AL DISEÑO DE REDES ÓPTICAS DE
NUEVA GENERACIÓN: PROTECCIÓN, CALIDAD DE
SERVICIO Y EFICIENCIA ENERGÉTICA**

**CONTRIBUTIONS TO NEW GENERATION OPTICAL NETWORKS DESIGN:
RESILIENCE, QUALITY OF SERVICE AND ENERGY EFFICIENCY**

Autor: Isaac Seoane Pujol

Directores: Dr. José Alberto Hernández Gutiérrez y Dr. David Larrabeiti López

Tribunal nombrado por el Mgfco. y Excmo. Sr. Rector de la Universidad Carlos III de Madrid, el
día ___ de Marzo de 2012. Firma del tribunal calificador:

Firma:

Presidente:

Vocal:

Vocal:

Vocal:

Secretario:

Calificación:

Leganés, ___ de Abril de 2012.

A Tere.

A mis padres.

A mi hermano.

A ti, Aroa,
que estás a punto de llegar

Y a todos aquellos a los que alguna vez les dije:

No puedo, tengo que hacer la tesis.

Y me esperaron con paciencia

Agradecimientos

Antes de comenzar esta parrafada busqué "Gracias" en el DRAE¹, para saber la forma más correcta de escribir cómo agradecer tanta ayuda que he recibido de todas las personas que me han acompañado en este camino y que me han tratado con gracia². A muchos les nombro a continuación; a los que no, espero me perdonen el despiste y me den una pequeña gracia³. A todos intento corresponder con la mía.

Comenzaré por agradecer la *gracia* que ofrecen⁴ y con la que actúan⁵ mis dos directores de tesis, José Alberto y David. Sin su perseverancia, tesón, eficacia, confianza y paciencia, nunca hubiese puesto punto y final a este documento. Gracias a los dos, por sacarme de mis vasos de agua y cortar mis frases pasivas interminables.

Este trabajo ha sido posible en parte gracias⁶ a la financiación recibida del proyecto T2C2 (Tecnologías Telemáticas para la Colaboración Ciudadana), financiado por el Ministerio de Economía y Competitividad, o como quiera que se haya llamado o se llame ahora, y del resto de proyectos de I+D+i en los que participo y he participado.

Quería agradecer a las personas que han formado y/o forman parte del grupo de investigación ADS-COM: Ángel, Gerson, Sergio, Raquel, Alfonso, Gonzalo, etc., su ayuda, colaboración y compañerismo en todo momento... excepto en esas reuniones infinitas... que compensamos con contadas pero excepcionales sobremesas. Ánimo en lo vuestro y aquí me tenéis para lo que necesitéis.

El día a día me ha llevado a compartir cosas de agradecer con muchas otras personas: Lisardo y David, con nuestros proyectos frikis. Roberto, con alguna caña sincera. Goyo y su sabiduría hedonista. Marco...deja de leer y ajusta el rumbo. Hablar de ciencias políticas con María. De equipos segundones con Alberto y Jaime. De la Fuerza con Paco. De cocina con Iria o de mil y una cosas variadas con Pablo (nombrado quedas) y Norberto. También agradecería de antemano al #tuistar Pablo que me honre con un RT. Y al resto de compañeros del Departamento de Ingeniería Telemática. Enumerar a toda la gente a la que tendría algo que agradecer, aunque sólo sea una sonrisa una mañana o una charla en un rato de café, es remitirse a la lista de personal del departamento. Gracias a todos.

Especialmente, existe un peligroso grupo de procrastinadores a cuyos pomos agradezco agarrarme, que están a ambos lados del límite de lo profesional y lo personal, y que se agrupan bajo su propia denominación de origen: los "glotones". Gracias a todos, por estricto orden alfabético: Antonio, Carlos, Ignacio, Isaías,

¹La cita de arriba es el resultado de lo que me sucedió cuando quise dar las gracias y me pregunté lo que significaba esta palabra, y del susto que me dio nuestro querido diccionario

²6. f. *Benevolencia y amistad de alguien*

³10. f. *Perdón o indulto de pena que concede el poder competente.*

⁴3. f. *Don o favor que se hace sin merecimiento particular; concesión gratuita.*

⁵4. f. *Afabilidad y buen modo en el trato con las personas.*

(las acepciones 1 y la 2 en realidad os las merecéis todos sólo por leer esto .. pero las buscáis en <http://www.rae.es/gracia>)

⁶*gracias a. loc. prepos.* Por causa de alguien o algo que produce un bien o evita un mal.

Iván, Julia, Manolo, Mar, Raquel, Richi, Rocío... Gracias por dejarnos formar parte de esta tropa. Será por casas rurales, restaurantes y bares por catar y aventuras por compartir juntos. ¡Qué hambre!

Habitantes de los pasillos en general, gracias. Con vosotros dan fruto hasta los pomos de las puertas. Seguid poniendo en práctica vuestra gracia⁷ que os caracteriza. Se me olvidará nombrar mucha gente ... a los que no hará esto ni pizca de gracia⁸ ... pero gracias a todos!!!

La lista de amigos a los que nombrar haría desistir de leer esto a los que aún no lo hayan hecho. En primer lugar es especial para mí agradecer a Nacho-y-Sara (con esa voz solemne de cura esquiviano), nuestras aventuras paralelas y su amistad paciente. Y a muchos más: Molo, Raquel, Edu, Kiko,... y todos aquellos con los que cogí este tren hace años sin saber que realmente pasaría por aquí.

Nunca podría haberse realizado este trabajo sin la existencia eterna e incondicional del trappulismo como forma de vida y religión. Gracias por vuestros infinitos correos infinitos. Gracias por mi nominación a muchas cosas y nombres, por vuestra preocupación y por confiar en que estaría toda la semana trabajando duro, cuando en realidad todos sabéis que en la universidad nos dedicamos a jugar todo el día con la Wii en red... y el premio es para: Perla, Brugue, Leonidas, Koseki, Zorro, Yago, Rubito, Quero, Bronson. Recogen este premio nuestras *cheerleaders* preferidas en compañía de la cantera: Bea, Silvia, Ainhoa, Chejo, Rata, Ali, Ana, Lidia y las conquistas internacionales de Quero que le apagan el despertador. El año que viene el premio a la actriz revelación del concurso de villancicos es para tí! (ssshhh, que viene, que viene...).

Ahí fuera de los pasillos, pero en lo más profundo de mi vida, están las personas sin las que ni siquiera sería yo mismo. Las que se merecen la gracia⁹ divina por soportar mis peores gracias¹⁰, y por ayudarme incondicionalmente a conseguir mis mejores gracias¹¹. Por supuesto me refiero a mi familia.

A mis abuelas y a mi tía, que me siguen recordando lo mucho que me gusta ser un niño. Mayor y "fuertecito" pero niño al fin y al cabo. Tío, ahora nos podremos llamar de tú, aunque el usted siempre nos dio para buenas risas. Otro doctor más para el clan.

A mi suegro, Juan, que siempre me ofrece jugosas soluciones por si hubiese que convencer al personal. Y al resto de su *troupe* que nos acompaña: mi suegra Teresa y el "cuñao" Juan Carlos.

Especialmente, gracias a mis padres, que día a día me seguís dando lecciones de superación y entrega para conseguir lidiar con los avatares de la vida y de paciencia infinita. Estas palabras no hacen justicia a lo mucho que tengo que agradecerlos. Pero sabéis lo mucho que os quiero. Y lo mucho que os va a tocar disfrutar aún en las etapas venideras.

A mi hermano Javi, en cuya vitalidad e imaginación y capacidad para estirar el tiempo, me fijo como modelo a seguir en cada nueva aventura en la que me embarco. Junto a él está su inseparable Bea. Gracias por estos últimos años de gran ayuda y soporte, y felicidades por vuestra(s) próxima(s) aventura(s).

Ojalá hubiese podido darle las gracias en persona hoy a mi abuelo Ramón, siempre preocupado por la cantidad de horas que pasaba en la universidad. Y enseñarle este libro como ejemplo tangible de aquello por lo que él siempre me preguntaba: «¿qué hacéis cuando no hay chavales? Si ahora están de vacaciones ... a qué os dedicáis en la Universidad»... pues no sé abuelo, pero ¡somos campeones del mundo!

Y gracias, por supuesto, a ti, Tere. Ya es la segunda vez que te dejo para el final en una parrafada de estas. La razón de ello es que en realidad siempre has sido y eres la pieza más importante de todo este puzzle, la fuente de la que parten y hacia donde van todos mis esfuerzos, y mi refugio de paz y felicidad. Se quedan muy cortas las palabras. Gracias por lo que consigues que sea a tu lado. Y gracias por confiar en mí para acompañarte en esta aventura en que nos hemos metido y que ha empezado en paralelo con la recta final de la tesis. Gracias por soportar el peso de esa tripita graciosa que se mueve. ¡Que viene, que viene!

Y como todo esta verborrea era para, en definitiva, que quede constancia que tengo mucho que agrade-

⁷5. f. *Habilidad y soltura en la ejecución de algo*

⁸9. f. *irón. Cosa que molesta e irrita.*

⁹13. f. Rel. En el cristianismo, favor sobrenatural y gratuito que Dios concede al hombre para ponerlo en el camino de la salvación.

¹⁰12. f. coloq. Acción o dicho de un niño que le sirve de lucimiento. Referido a personas adultas, u. t. en sent. irón.

¹¹14. f. Col. Proeza, hazaña, mérito.

ceros a todos, gracias. Os merecéis alguna compensación por haberme soportado en alguna parte del camino ... pero hoy no que es tarde. Mañana.

Empecé este rollo leyendo el diccionario, esto es lo que buscaba:

dar s, o las s. 1. *locs. verbs.*

Manifiestar el
agradecimiento por el
beneficio recibido.

gracias. 1. *expr. U.*

para expresar
nuestro agradecimiento por
cualquier beneficio, favor
o atención que se nos dispensa.

Abstract

Next Generation Optical Switching Networks development introduces several new challenges. This Ph D. Thesis makes contributions to some research topics that are still open in the state of the art:

- quality of service based on delay differentiation,
- network reliability modelling,
- and energy consumption reduction.

Regarding quality of service based on delay differentiation, two solutions have been proposed for different deployment scenarios. The first one, proposes low cost redesign of optical TT-FR ring nodes, in order to achieve an optimal relationship between traffic load and the delay suffered because of the node relative position in the topology. The second one proposes a variation of the dynamic bandwidth assignment (DBA) in passive optical networks (PON) to reduce delay jitter of the high-priority traffic.

The second topic raised, network reliability and failure propagation treatment, deals with the design of a model for the network performance based on the network failure and repairing rates. These rates depend on the topology and the failure treatment protocols chosen by the network operator. Design results, which have been validated using an optical ring topology, will enable the network operators and service providers to design the failure control and repair mechanisms deployed, in order to be able to offer the desired reliability.

Finally, there are two new contributions on energy savings. The first one is the optimization of the flow allocation problem based on energy constraints applied to Wide Area Networks (WAN); and the second one is focused on designing energy aware DBA mechanisms in passive optical networks, modifying the DBA algorithm at the OLT which changes the TDM transmission slots assignment in order to allow the ONUs to stay in a low consumption mode for a long time.

Keywords: *Optical Switching Networks, Passive Optical Networks, Network Reliability, Energy Efficiency, Quality of Service, Optical Ring topologies.*

Resumen

El desarrollo de la siguiente generación de redes ópticas conmutadas se enfrenta a múltiples retos tecnológicos. Este trabajo de tesis doctoral realiza contribuciones al modelado y análisis de diferentes aspectos abiertos en el estado del arte, en particular:

- la implementación de calidad de servicio basada en diferenciación de retardo,
- el modelado de mecanismos de protección en la red,
- y la reducción del consumo energético.

En relación con la calidad de servicio ofrecida en base al retardo que sufre el tráfico, se proponen dos soluciones con diferente ámbito de aplicación. La primera de ellas propone modificaciones de bajo coste en el diseño de los nodos TT-FR de un anillo óptico, con el objeto de poder balancear la relación entre la carga de tráfico en los nodos y el retardo que sufre el tráfico, debido a la posición relativa del nodo de entrada del tráfico en el anillo. La segunda propuesta está relacionada con el protocolo de asignación dinámica de ancho de banda (DBA) en redes ópticas pasivas. En este caso, se propone una solución de planificación entre los propios dispositivos ópticos que dan acceso a la red (ONUs), y que agregan el tráfico hacia la red troncal, para reducir la variación del retardo que sufre el tráfico de alta prioridad.

En el segundo de los aspectos investigados, la actuación en caso de fallos, se presenta un método novedoso de modelado del comportamiento de la red en base a sus tasas de fallo y a las tasas de recuperación. Estas tasas dependerán de la topología y del diseño del planteamiento de actuación elegido por el proveedor de servicio. Los resultados de este diseño, del cual se presenta un ejemplo para una topología en anillo, permitirán al operador de la red la posibilidad de ajustar los mecanismos implementados en ella para poder ofrecer los valores de fiabilidad deseados.

Finalmente, se aborda el estudio de soluciones que permitan la reducción de consumo energético. El primero de los estudios realizados se basa en la optimización del encaminamiento en una red de área extensa (WAN). La segunda propuesta se centra en redes ópticas pasivas (PON) alterando el comportamiento del algoritmo de DBA por parte de la OLT, y realizando una asignación de tiempos para la transmisión que permite a las ONUs permanecer en un estado de bajo consumo la mayor cantidad de tiempo posible.

Palabras clave: *Redes de conmutación Ópticas, Redes Ópticas Pasivas, Recuperación frente a fallos, Eficiencia energética, Calidad de Servicio, Anillo Óptico.*

Índice general

1. Presentación	1
1.1. Motivación del estudio	1
1.2. Objetivos del estudio	2
1.3. Estructura del documento	4
I Estado del arte	7
2. Redes ópticas: estado del arte y conceptos básicos	9
2.1. Introducción	9
2.2. Componentes básicos de una red de conmutación óptica	10
2.2.1. Medios de Transmisión Ópticos: Fibra óptica	12
2.2.2. Dispositivos de transmisión y recepción	18
2.2.2.1. LEDs (<i>Light Emitting Diode</i>)	18
2.2.2.2. Láseres	18
2.2.2.3. Filtros ópticos y fotodetectores	20
2.2.3. Componentes de sistemas de transmisión y conmutación en redes ópticas	20
2.2.3.1. Bloques básicos pasivos de redes ópticas	22

2.2.3.2.	Otros bloques de conmutación óptica	22
2.3.	Topología, conmutación y control en una red óptica	26
2.3.1.	Transparencia en redes ópticas	26
2.3.2.	Multiplexación en el dominio óptico	26
2.3.2.1.	TDM	27
2.3.2.2.	SDM	28
2.3.2.3.	WDM	28
2.3.3.	Granularidad en los sistemas de conmutación	29
2.3.3.1.	Conmutación de circuitos: OCS	30
2.3.3.2.	Conmutación de paquetes: OPS	31
2.3.3.3.	Conmutación basada en ráfagas: OBS	31
2.3.4.	Gestión y control de Redes Ópticas	32
2.3.5.	GMPLS - Generalized Multi-Protocol Label Switching	34
2.4.	Protección frente a fallos en redes ópticas	39
2.4.1.	Protección y restauración	40
2.4.2.	Clasificación de técnicas de protección y restauración	41
2.4.3.	Otras arquitecturas de protección en redes ópticas:	42
2.4.3.1.	Protección en SoNet/SDH	42
2.4.3.2.	RPR - <i>Resilient Packet Ring</i>	43
2.4.3.3.	Protección en WDM	45
2.4.3.4.	P-ciclos	45
2.5.	Despliegue de redes ópticas	46
2.5.1.	Redes de área extensa (WAN - Wide Area Networks)	47
2.5.2.	Redes ópticas Metropolitanas	49
2.5.3.	Red de acceso: FTTx	49

2.5.3.1.	Redes ópticas pasivas (PONs)	52
2.5.3.2.	Gestión de acceso en EPON	55
2.5.3.3.	Gestión del ancho de banda en PON	56
2.6.	Eficiencia Energética	57
2.6.1.	Eficiencia energética en redes ópticas	60
2.6.2.	Energy Efficient Ethernet (EEE)	61
2.7.	Resumen y conclusiones	62

II Soluciones de eficiencia propuestas para varios aspectos de las redes de conmutación ópticas 65

3. Propagación de fallos en Anillos Ópticos GMPLS: modelo y análisis de prestaciones 67

3.1.	Introducción	67
3.2.	Trabajos anteriores y planteamiento del problema	69
3.2.1.	Propagación de fallos en redes multicapa	69
3.2.2.	Fallos en redes basadas en GMPLS	71
3.2.3.	Propagación de fallos basados en modelos epidémicos	73
3.3.	El modelo SID: <i>Susceptible-Infected-Disabled</i>	74
3.4.	Análisis y diseño del modelo	76
3.4.1.	Modelo de Cadenas de Markov en Tiempo Continuo (CTMC)	76
3.4.2.	Caso de estudio: un anillo GMPLS de 8 nodos	78
3.4.3.	Consideraciones para el modelo general de un anillo GMPLS de N nodos	82
3.5.	Resultados Numéricos	85
3.5.1.	Métricas de prestaciones	85
3.5.2.	Resultados para las probabilidades de estado estacionario	88
3.5.3.	Tiempos de primera ocurrencia (<i>First-Passage Times</i>): MTTF y MTTR	90

3.5.4. Anillo GMPLS de 32 nodos	92
3.6. Resumen del capítulo y conclusiones	93
4. Análisis del retardo medio y su varianza en anillos WDM libres de colisión mediante el uso de segmentos de recirculación del tráfico bloqueado	95
4.1. Introducción y trabajos anteriores relacionados	95
4.2. Control de colisión mediante recirculación de tráfico bloqueado	98
4.3. Análisis de los sistemas de recirculación y del retardo extremo a extremo	100
4.3.1. Análisis de la probabilidad de bloqueo	100
4.3.2. Análisis del retardo extremo a extremo	103
4.3.2.1. Recirculación con almacenamiento-reenvío	103
4.3.2.2. Recirculación directa o <i>bypass</i>	106
4.4. Ejemplos numéricos	108
4.4.1. Análisis y validación mediante simulación	109
4.4.2. Análisis del retardo para diferentes valores de N	109
4.4.3. Análisis del retardo sufrido cuando un nodo presenta un pico de carga	111
4.5. Resumen y conclusiones	112
5. Algoritmo de encaminamiento para eficiencia energética en redes Ethernet	115
5.1. Introducción	115
5.2. Descripción del funcionamiento de EEE-FA	119
5.2.1. Ejemplo de funcionamiento de EEE-FA en una red de tres nodos	119
5.2.2. El algoritmo EEE-FA	121
5.3. Experimentos con el algoritmo EEE-FA	122
5.3.1. Escenario de simulación	123
5.3.2. Resultados de los experimentos	124

5.4. Resumen y conclusiones	126
6. Análisis y simulación de algoritmos basados en IPACT para diferenciación de servicios basada en retardo en PONs	129
6.1. Introducción y trabajos relacionados	129
6.1.1. Estudio de trabajos anteriores en diferenciación de servicios	130
6.2. Estudio analítico del modelo de asignación de intervalos TDM para DBA en PON	132
6.2.1. Suposiciones previas para el análisis del modelo	133
6.2.2. Análisis del tiempo medio de ciclo	134
6.2.3. Análisis del retardo máximo	138
6.3. Análisis del modelo de asignación de TDM para PON con diferenciación basada en retardo	139
6.3.1. Descripción del algoritmo de planificación con diferenciación basado en retardo	139
6.3.2. Análisis del tiempo de ciclo medio para cada clase de tráfico	141
6.4. Análisis con tiempo de ciclo de alta prioridad garantizado	145
6.5. Resultados experimentales	147
6.5.1. Validación de los tiempos de ciclo medios $E[T]$, $E[T^{(HP)}]$ y $E[T^{(LP)}]$.	147
6.5.2. Experimentos bajo condiciones de tráfico elevado	148
6.6. Resumen y conclusiones	149
7. Incremento del ahorro energético de una E-PON	153
7.1. Trabajos relacionados	153
7.2. Formulación del problema	156
7.3. Aumento del porcentaje de ahorro energético	160
7.4. Estimación del ahorro energético	163
7.5. Diseño del tiempo de espera óptimo	164

7.6. Resultados numéricos obtenidos mediante la simulación de una PON de 8 ONUs	168
7.7. Conclusiones y trabajos futuros	170
III Conclusiones y Trabajos Futuros	173
8. Conclusiones y Trabajos Futuros	175
8.1. Conclusiones Finales	175
8.2. Trabajos Futuros	178
A. Lista de Contribuciones	181
Referencias	185
Acrónimos	199

Índice de figuras

1.1. Esquema del ámbito de trabajo de la presente Tesis Doctoral	3
2.1. Espectro electromagnético	11
2.2. Ejemplos de fibras ópticas	12
2.3. Reflexión de la luz en el interior de la fibra (a) y tipos de fibras (b)	14
2.4. Atenuación en la fibra óptica en función de la longitud de onda y la distancia	15
2.5. Efecto de la dispersión modal sobre un pulso óptico	16
2.6. Anchos de banda normalizados de fuentes de luz comerciales tipo LED	18
2.7. Comparativa entre la forma de espectro de un laser y de un diodo LED sintonizados a la misma frecuencia	19
2.8. Frecuencias disponibles en de láseres comerciales	21
2.9. Esquemas funcionales de componentes básicos pasivos	23
2.10. Esquema de un Cross-Bar Switch en ambos modos de funcionamiento	24
2.11. Esquemas funcional de un OADM	25
2.12. Esquemas funcional de un OXC	25
2.13. Jerarquías de interfaces para su etiquetado en GMPLS	38
2.14. Despliegue actual de tecnologías en FTTH y previsiones a lo largo de 2012	51
2.15. Esquema básico de una red PON	53

3.1. Los planos de control y datos en la arquitectura de GMPLS	71
3.2. Diagrama de transición de estados para el modelo SID	75
3.3. Anillo basado en GMPLS de 8 nodos de ejemplo	78
3.4. Diagrama de tasas de transición y estados para el modelo SID en un anillo GMPLS de 8 nodos	79
3.5. Ejemplo de posibles configuraciones del anillo y su estado correspondiente: (a) (0:0:0), (b) (0:0:1), (c) (0:1:0), (d) (0:0:2), (e) (0:0:3) y (f) (0:1:1)	83
3.6. Grupos de estados estudiados	87
3.7. Impacto de t en las probabilidades de estado estacionario de la CTMC de un anillo GMPLS de 8 nodos	89
3.8. Efecto de δ en las probabilidades estacionarias de una CTMC para un anillo GMPLS de 8 nodos	90
3.9. Tiempo Medio para el primer fallo - Mean Time To Failure (MTTF)	91
3.10. Tiempo medio para la reparación -Mean Time To Repair (MTTR)	91
3.11. Efecto de δ/β en las probabilidades estacionarias de un anillo de 32 nodos	92
4.1. Ejemplo de anillo WDM de 4 nodos con recirculación de tráfico. Notación de las intensidades de tráfico.	99
4.2. "bypass de recirculación" (izquierda) y "almacenamiento y reenvío" (derecha)	99
4.3. Análisis de validación con Simulink de Matlab para una red en anillo de $N = 8$ nodos. Resultados analíticos (a la izquierda) y simulados (derecha) se ha obtenido suponiendo recirculación con "almacenamiento-reenvío" (arriba) y "bypass"(abajo).110	
4.4. Valor medio y varianza del retardo que experimentan los paquete usando ambas políticas para una cantidad de nodos $N + 1$ variable	111
4.5. Retardo medio y varianza en cada nodo suponiendo que el nodo con el pico de tráfico es el n_1 (arriba) y el nodo n_6 (abajo) en un anillo WDM de 8 nodos	112

5.1. Consumo energético frente a carga del enlace para BASE-TX, 1000BASE-T y 10GBASE-T EEE, según el estudio realizado en [96]	117
5.2. Ejemplo de encaminamiento de 3 demandas sobre una red de tres nodos utilizando dos algoritmos con objetivos de coste distinto: (a) Shortest-Path Routing (SPR), (b) Energy Efficient Ethernet Flow Allocation (EEE-FA)	120
5.3. Topología de red de prueba	123
5.4. Consumo de potencia instantáneo para EEE-FA y SPR para diferentes capacidades de enlace	125
6.1. Esquema básico de una red PON	130
6.2. Esquema básico del funcionamiento del protocolo de DBA para una ONU	131
6.3. Definiciones para una PON con $N = 3$ ONUs	134
6.4. Carga de tráfico en la red mínima para que el tiempo de ciclo T_{ciclo} no dependa de la influencia de T_{min}	138
6.5. Colas de espera para las clases de servicio HP y LP en cada ONU	139
6.6. Algoritmo con diferenciación basada en retardo. Caso $M = 1$	140
6.7. Algoritmo con diferenciación basada en retardo. Casos $M = 1, 2$ y 3 y $N = 3$. Instantes de tiempo referidos a la ONU_i con $i = 1$	142
6.8. Algoritmo con diferenciación de servicios dentro de los umbrales correspondientes. Caso $M = 1$ y $N = 3$	145
6.9. Tiempos de ciclos medios $E[T^{(HP)}]$ y $E[T^{(LP)}]$ para diferentes valores del cociente M/N : (a) $\frac{M}{N} = \frac{1}{32}$, (b) $\frac{M}{N} = \frac{4}{32}$, (c) $\frac{M}{N} = \frac{16}{32}$, (d) $\frac{M}{N} = \frac{32}{32}$	148
6.10. Tiempos medios de ciclo $E[T]$, $E[T^{(HP)}]$ y sus desviaciones típicas ($E[T] \pm 2Std[T]$) para diferentes relaciones M/N : (a) $\frac{M}{N} = \frac{1}{32}$, (b) $\frac{M}{N} = \frac{4}{32}$, (c) $\frac{M}{N} = \frac{16}{32}$, (d) $\frac{M}{N} = \frac{32}{32}$	149
6.11. Tiempos medios de ciclo y paquetes perdidos con el algoritmo DBA de diferenciación basada en retardo en condiciones de cantidades de tráfico altas: $T_{Limit} = 3,2ms$, tiempo medio de ciclo (a) , el número de paquetes perdidos (b); $T_{Limit} = 8ms$, tiempo medio de ciclo (c) y pérdida de paquetes (d).	150

7.1. Esquema básico de temporización de los cambios de estado de una ONU.	159
7.2. Esquema básico de la gestión del enlace ascendente en EPON con EEE. Nomenclatura del modelo.	160
7.3. Ejemplo de cálculo del tráfico estimado de una ONU con $\alpha = 0,75$	166
7.4. Ejemplo de cálculo del porcentaje de potencia consumida respecto del máximo en función de la carga para diferentes valores objetivo del retardo medio máximo . .	167
7.5. Ejemplo de cálculo del retardo medio esperable en función de la carga para varios objetivos de consumo de potencia	168
7.6. Distribución de probabilidad del tamaño de paquete, con tamaño medio de 1250 bytes, mínimo de 64 bytes y máximo de 1518 bytes.	169
7.7. P_{cons} en función de D_q para varios valores de ρ_T	170
7.8. P_{cons} en función de ρ_T para varios valores de D_q	171

Índice de tablas

2.1. Longitudes de onda más utilizadas en función del tipo de fibra	13
2.2. Resumen de características de tipos de conmutación óptica	32
2.3. Resumen de características de Redes Ópticas Pasivas (PONs)	55
3.1. Notación y parámetros correspondientes al modelo SID	76
3.2. Reglas de generación de transiciones para Q_N	82
3.3. Matriz generadora infinitesimal para un anillo óptico GMPLS de 8 nodos (Q_8)	84
5.1. Algoritmo EEE-FA	122
5.2. Parámetros de la red de simulación	124
5.3. Resumen de resultados de simulación	127
7.1. Valores de los parámetros de diseño de la simulación de una PON con EEE	156
7.2. Valores de T_{sync} y de potencia nominal consumida en una ONU en cada uno de las opciones de funcionamiento	160
7.3. Valores de los parámetros usados en la simulación de una red PON con EEE	169

Capítulo 1

Presentación

1.1. Motivación del estudio

Las redes de conmutación ópticas (*OSN-Optical Switching Networks*) han evolucionado en los últimos años para proporcionar soluciones de alta velocidad y alto ancho de banda en las capas bajas de todo tipo de redes de comunicaciones, desde las grandes redes del núcleo de Internet hasta las redes más cercanas a los usuarios finales.

En las últimas tres décadas, se han definido muchos y diversos tipos nuevos de redes ópticas, topologías de red, nuevos componentes y técnicas de conmutación, y muchas aplicaciones están aprovechando las ventajas que proporciona su utilización. Cada uno de estos nuevos avances presenta nuevas ventajas, pero también sus propios inconvenientes, de forma que hay soluciones que se ajustan mejor que otras, dependiendo del escenario final de utilización.

Tanto el mundo industrial como el académico están plenamente interesados en incorporar soluciones completamente ópticas para mejorar las prestaciones de las redes ya que permiten mejorar aspectos relativos a la reducción de costes. Sin embargo, no todos los escenarios o redes de conmutación pueden implementarse en la actualidad con redes completamente ópticas. Esto conlleva el hecho de que algunos flujos de tráfico en ciertos nodos intermedios necesiten conversión del medio óptico al electrónico, y viceversa (*OEO-Optical-Electrical-Optical Conversion*), lo cual introduce ciertas desventajas intrínsecas a este proceso.

1.2. Objetivos del estudio

El principal objetivo del trabajo de tesis doctoral que se presenta en este documento es el estudio de diversos aspectos relacionados con las redes de conmutación ópticas, con el fin de minimizar las desventajas inherentes a las soluciones actuales, y presentar las correspondientes contribuciones propuestas para llevar a cabo este objetivo. Los objetivos parciales de la tesis doctoral están divididos en función del área de alcance de la red y de su topología, y del tipo de prestaciones a mejorar: calidad de servicio, ahorro de energía y fiabilidad (Fig. 1.1).

Los resultados del primer trabajo que se presenta, proporcionan un mecanismo para redes en anillo que permite balancear el retardo y la probabilidad de bloqueo que sufre el tráfico ofrecido a la red en dichos nodos. La solución propuesta tiene en cuenta la posibilidad de usar la conversión OEO junto con un sistema de almacenamiento (*buffering*) electrónico de los datos, o elegir una solución completamente óptica que conecta los nodos con una fibra óptica adicional, de forma que el tráfico bloqueado se vuelva a recircular en dirección opuesta y tenga otra oportunidad de ser transmitido correctamente hacia su destino.

Las redes de conmutación ópticas se utilizan habitualmente para desplegar redes de área extensa (WAN - *Wide Area Networks*), no sólo debido a sus características de alta capacidad sino también por su baja tasa de fallos. La protección y recuperación frente a fallos en redes ópticas es otro campo que ha despertado gran interés para su investigación en los últimos años. En este trabajo de tesis se presenta también un estudio realizado dentro de este paradigma de protección y recuperación del funcionamiento de la red. Dicho estudio proporciona el diseño de un modelo que describe la tasa de fallos en anillos ópticos, con el objetivo de proporcionar al operador de la red de la información necesaria para adecuar el diseño y la implementación de sus estrategias y mecanismos de recuperación y tasas de intervención para conseguir ofrecer un servicio con un alto nivel de protección y fiabilidad (por ejemplo, durante el 99.999 % del tiempo).

Otro escenario en el cual las redes de conmutación ópticas están muy extendidas es el de las redes de acceso. Las Redes Ópticas Pasivas (PON-*Passive Optical Networks*) con uno de los tipos de red de acceso más desplegados en la actualidad por los operadores de red en sus nuevas instalaciones. Este tipo de red tiene características óptimas para topologías punto a multipunto, como las que requieren los Proveedores de Servicio de Internet (ISP - *Internet Service Providers*), en su red de acceso, como ocurre por ejemplo en las redes FTTH (*Fiber To The Home*).

En la red de acceso, los usuarios solicitan a los proveedores de servicio el cumplimiento de

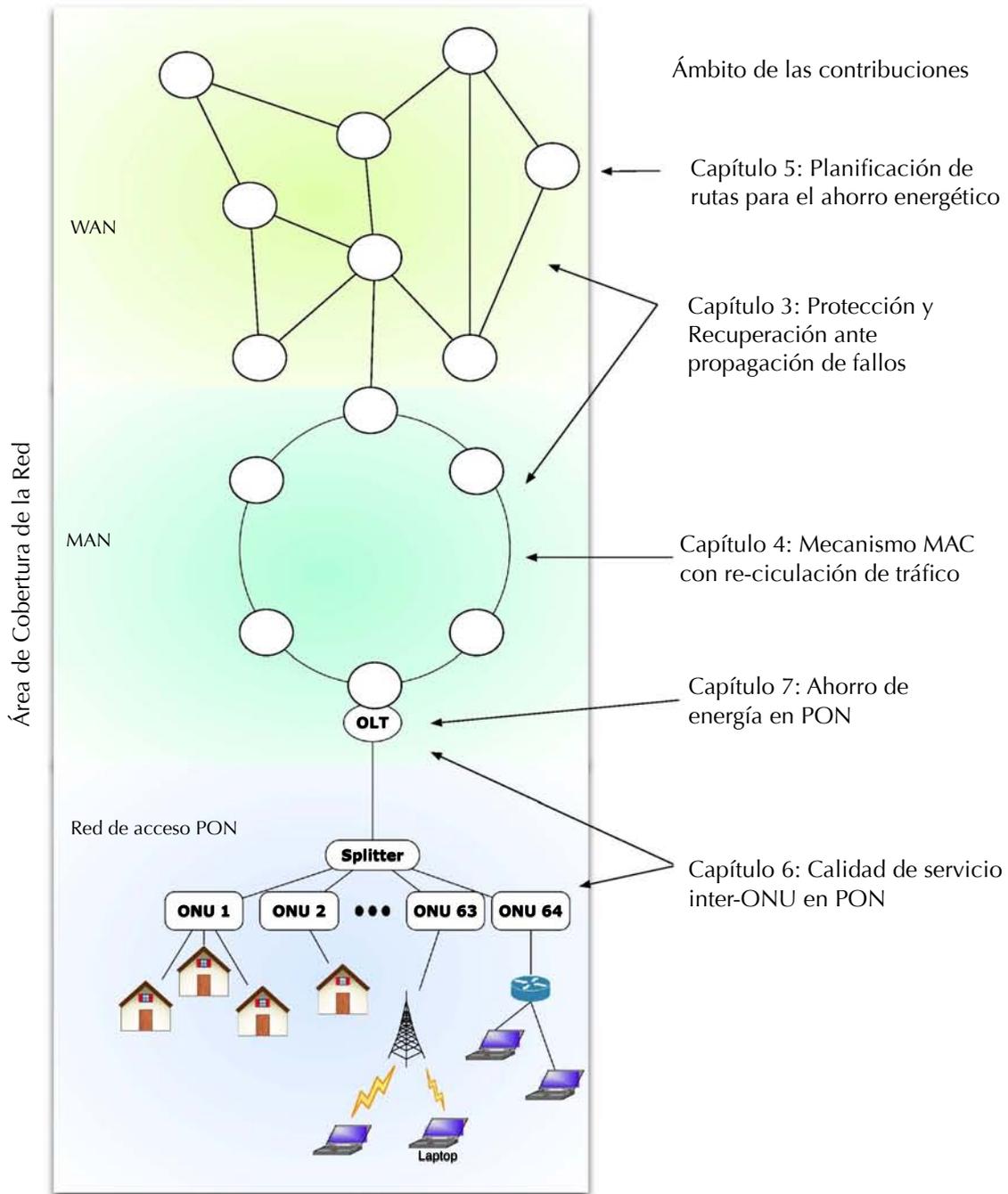


Figura 1.1: Esquema del ámbito de trabajo de la presente Tesis Doctoral

requisitos de garantía de calidad de servicio para su tráfico prioritario. Por lo tanto, la existencia de protocolos y mecanismos de calidad de servicio (QoS - *Quality of Service*) es un requisito indispensable que se debe cumplir en las redes PON. Los últimos trabajos realizados en el curso de esta tesis doctoral proponen mecanismos para conseguir garantías de baja variabilidad del retardo (*jitter*) para las clases más prioritarias en una red PON.

Finalmente, en las últimas décadas ha aumentado el interés hacia un tema de investigación común a todos los campos científicos: el ahorro de energía. La preocupación por reducir el consumo energético en cualquier desarrollo tecnológico está presente desde las primeras etapas del mismo. No sólo es una preocupación relacionada con la optimización del consumo de materias primas o de su impacto en el medio ambiente, sino que también está directamente relacionada con la siempre inherente preocupación económica de la reducción de costes. En esta línea se presenta también en esta tesis varios trabajos encaminados a conseguir la minimización del consumo de energía tanto en redes de acceso como en redes de mayor cobertura, proponiendo para ello sencillas modificaciones de los protocolos de planificación y encaminamiento del tráfico respectivamente.

1.3. Estructura del documento

El contenido del documento está estructurado en tres partes principales. Dado que el campo de las comunicaciones ópticas es muy extenso, la primera parte del documento, formada por el capítulo 2, presenta una breve introducción al paradigma de las redes de conmutación óptica y se centra en la base de aquellos aspectos más importantes relacionados con los trabajos de investigación que se presentan a continuación.

En la segunda parte, se describen detalladamente las contribuciones y trabajos de investigación, realizados durante la elaboración de esta tesis. Cada uno de ellos presenta un breve estudio de las soluciones existentes, el diseño y modelado de la contribución propuesta, así como los resultados de simulación y experimentación obtenidos que permiten validar la viabilidad y la posible mejora prevista tras su implementación. Tal y como puede verse en la Fig. 1.1, este trabajo de tesis aborda estudios relacionados con los principales escenarios y tipos de conmutación ópticas:

- El capítulo 3 presenta el estudio relacionado con la fiabilidad de la red y la propagación de errores en un anillo óptico, arquitectura típica en escenarios de red metropolitana (MAN - *Metropolitan Area Networks*).

- Para este mismo tipo de topología en anillo y extensión de la red (MAN), el capítulo 4 describe la contribución relacionada con la modificación del protocolo de control de acceso al medio (MAC - *Medium Access Control*) de los nodos del anillo y su impacto en las prestaciones de retardo y bloqueo del tráfico.
- El capítulo 5 presenta un estudio relacionado con el ahorro energético en redes basadas en Ethernet (EEE - *Energy Efficient Ethernet*), y como ejemplo se aplica su uso en el encaminamiento de tráfico sobre una red mallada extensa (WAN).
- En el capítulo 6 se describe la contribución propuesta para la mejora de la calidad de servicio basada en retardo en una red óptica pasiva (PON) en la red de acceso.
- Y finalmente, dentro del ámbito de la red de acceso, el capítulo 7 describe una propuesta de aumento del ahorro energético en redes PON.

La tercera parte (capítulo 8), presenta las conclusiones globales más importantes extraídas de la realización de este trabajo de tesis doctoral junto con un breve resumen de las líneas de investigación detectadas de mayor relevancia para la continuación de la misma.

Asimismo, una lista de las publicaciones científicas resultantes de los trabajos de investigación realizados en el marco de trabajo de esta tesis puede encontrarse en el apéndice A.

Parte I

Estado del arte

Ten muy presente que tu enfoque determina tu realidad. Muchas de las verdades que creemos dependen del punto de vista.

Maestro Yoda

Capítulo 2

Redes ópticas: estado del arte y conceptos básicos

Este capítulo trata de introducir brevemente al lector de este trabajo en los conceptos más relevantes del paradigma de las redes ópticas, de forma que sea de utilidad para entender los problemas abordados y las soluciones propuestas que han llevado a la conclusión de esta tesis doctoral. También se intenta mostrar una idea del estado del arte actual de las tecnologías y técnicas usadas en redes ópticas describiendo las soluciones existentes para poder identificar aquellos temas abiertos a la investigación.

2.1. Introducción

Las comunicaciones ópticas fueron evolucionado en el último cuarto del siglo XX desde los primeros enlaces ópticos direccionales punto a punto hasta las redes de conmutación completamente ópticas que están desplegándose en la actualidad.

De acuerdo a [1], la expresión *redes de conmutación ópticas* se usa para referirse a los distintos tipos existentes de redes flexibles y reconfigurables que hacen uso de las técnicas existentes de multiplexación óptica, sintonización de longitud de onda y conmutación a nivel óptico en sus nodos de interconexión.

Desde su aparición, las técnicas ópticas han sido reconocidas por proveer al paradigma telemático con las ventajas intrínsecas del uso de medios y dispositivos ópticos en la arquitectura

de la red. Sobre todo, el uso de fibra óptica como medio de transmisión, ha permitido aprovechar sus bien conocidas y extraordinarias características como medio de transmisión guiado: muy baja atenuación, gran ancho de banda e inmunidad frente a interferencias electromagnéticas.

Además del medio de transmisión, los dispositivos de transmisión y recepción ópticos han ido evolucionando para llegar a ser muy baratos en su fabricación así como cada vez más precisos y de mejores prestaciones. El uso del láser permite incrementar enormemente el producto distancia-ancho de banda gracias a su potencia de transmisión, sin olvidarnos de que pueden sintonizarse a las longitudes de onda deseadas con gran precisión. Estas características permitieron el desarrollo de las técnicas de multiplexación por división en longitudes de onda (WDM-Wavelength División Multiplexing), claves para el despliegue actual de las redes ópticas.

Mediante WDM, se incrementa la capacidad y la velocidad de transmisión que puede ofrecerse al usuario por fibra óptica, dado que cada canal puede funcionar sintonizado a una longitud de onda distinta que no interfiere con el resto. El concepto de *Bypass óptico - Optical bypassing* permite a los diseñadores de red evitar las conversiones de las señales ópticas a señales eléctricas (OEO - optical-electrical-optical) en los nodos de conmutación intermedios. Manteniendo el tráfico entrante en el dominio óptico, se evita el cuello de botella del procesamiento electrónico y se reduce la potencia consumida en los nodos intermedios y, por tanto, la huella de carbono final del equipo. Además, se reduce el retardo, el número de puertos de conversión a implementar y el procesamiento de información de las cabeceras de control.

La aparición de redes completamente ópticas ha sido rápidamente aceptada desde sus comienzos tanto por la industria del sector como por los grupos de investigación académicos. Desde entonces hasta ahora, un sinnúmero de trabajos de investigación, desarrollo e implementación se han centrado tanto en el despliegue como en la estandarización: algoritmos, técnicas de conmutación, arquitecturas optimizadas, etc., siendo muchos de estos mecanismos nuevos y muchos otros heredados del más extendido mundo de las redes electrónicas. En las siguientes secciones del capítulo se hace un repaso a los principales componentes y aspectos tecnológicos y científicos involucrados en la resolución del puzzle del paradigma de las redes ópticas.

2.2. Componentes básicos de una red de conmutación óptica

La luz, como el resto de fenómenos físicos que conforman el espectro electromagnético, puede caracterizarse por la frecuencia a la que se emite. Frecuencia y longitud de onda están relacionadas

entre sí en función de la velocidad a la que se propaga en el medio la señal: $v = \lambda f$, por lo que unas veces se usa la frecuencia y otras veces la longitud de onda indistintamente para referirnos a la señal utilizada (Fig. 2.1¹).

En el caso de señales de longitudes de onda bajas, como la luz visible, rayos ultravioleta (UV) o rayos-X, por ejemplo, es más frecuente utilizar la longitud de onda o λ (lambda) para caracterizarlas. Las señales luminosas visibles se sitúan en la parte del espectro electromagnético, a longitudes de onda mayores que los rayos ultravioletas, entre las longitudes de onda de 400 a 800nm (desde el azul/violeta hasta el rojo intenso), seguidas por el infrarrojo cercano (hasta los 2500nm).

Por lo tanto, la longitud de onda (en adelante nos referiremos a ella indistintamente como λ o lambda para abreviar), identifica una señal por su "color" o frecuencia de emisión y la caracteriza diferenciándola de otras señales transmitidas a frecuencias diferentes. Los dispositivos y medios utilizados en comunicaciones ópticas podrán por tanto emitir, recibir o propagar señales de uno o varios "colores" o longitudes de onda simultáneamente.

Dentro del paradigma de comunicaciones ópticas esta sección describe resumidamente los bloques más importantes utilizados para construir redes de conmutación ópticas, desde los dispositivos de transmisión y detección de luz hasta los componentes de conmutación necesarios para construir topologías avanzadas e implementar las técnicas de multiplexación que se describen en posteriores secciones de este capítulo.



Figura 2.1: Espectro electromagnético

¹http://rodcontre.files.wordpress.com/2011/08/electromagnetic_spectrum.png

2.2.1. Medios de Transmisión Ópticos: Fibra óptica

La idea general de la comunicación óptica aparece en paralelo con el desarrollo de las comunicaciones: el concepto de transmitir símbolos usando la luz es tan antiguo como la idea de hacerlo usando cualesquiera de los otros fenómenos físicos que se conocen en comunicaciones, como por ejemplo el sonido o el electromagnetismo. Sin embargo, el concepto que hace despegar la investigación en comunicaciones ópticas en la segunda mitad del siglo XX es la posibilidad de enviar la luz de forma guiada, más concretamente, la posibilidad de usar fibras ópticas para transmitir luz guiada.

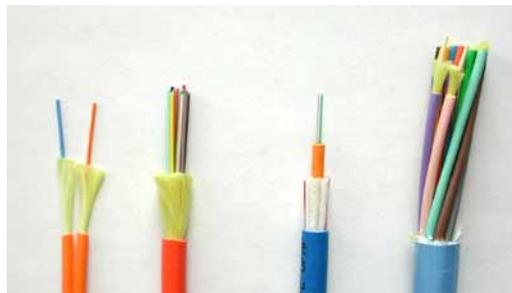


Figura 2.2: Ejemplos de fibras ópticas

La fibra óptica no sólo permite una manera muy efectiva en cuanto a coste de transmitir luz entre dos puntos con muy poca atenuación en largas distancias, sino que además es un medio de gran capacidad, ancho de banda y más protegido frente a interferencias que el resto de medios guiados que se utilizan, sobrepasando con creces las mejores características de los cables de cobre o coaxiales tradicionales.

Una fibra óptica consiste en un núcleo de cristal de silicio transparente recubierto por una capa con un índice de refracción mayor. La diferencia de los índices de refracción de ambas capas, según las leyes de Snell, permite que la luz rebote y vuelva al núcleo de la fibra avanzando así a lo largo del eje de la misma. El ángulo de incidencia de la luz a la entrada de la fibra y los ángulos de rebote ocasionados por las diferencias de índices de refracción, provocan que la velocidad de propagación de la luz guiada en la fibra sea menor respecto de la velocidad natural de la luz en el vacío. Los dispositivos ópticos actuales (Láseres y LEDs) generan luz en frecuencias concretas pero con un cierto grado de dispersión, lo que hace que, en general, varios modos de luz estén presentes a la entrada del medio de transmisión óptico.

Dependiendo de las dimensiones del núcleo, de la diferencia de índices de refracción y de la frecuencia de la luz que se transmite, uno o varios modos pueden propagarse a lo largo de la

Fibra Óptica Plástica (POF)	Fibra Óptica de Índice Gradual	Fibra Monomodo
650 nm 850 nm	850 nm 1300 nm	1310 nm 1490 - 1625 nm

Tabla 2.1: Longitudes de onda más utilizadas en función del tipo de fibra

fibra. Las fibras ópticas monomodo (Fig. 2.3(a)) se construyen usando diámetros de núcleo muy pequeños en relación a la longitud de onda de la luz que se desea transmitir. De esta manera, sólo un modo de luz va a puede propagarse a la frecuencia seleccionada.

Cuando las dimensiones del núcleo son grandes en relación a la longitud de onda transmitida, los pulsos de luz pueden propagarse en múltiples modos que seguirán caminos de propagación de diferente longitud. Este tipo de fibras se denominan fibras multimodo. Cada modo de propagación tendrá una dirección de propagación diferente dentro de la fibra, y tardarán tiempos distintos en recorrerla, lo que da lugar a la llamada dispersión modal y la interferencia intersimbólica (ISI). Para distancias cortas, las fibras multimodo pueden estar construidas con núcleos de polímeros de plástico (fibras POF -Polymer Optical Fibers y HCS), de forma que se consigan enlaces muy flexibles y baratos a costa de peores prestaciones en transmisión.

Como se ha comentado, el cristal de silicio del núcleo de la fibra ha de tener un índice de refracción menor que la capa que lo recubre. Cuando la relación entre este índice de refracción es constante en toda la fibra, ésta se denomina fibra multimodo de salto abrupto de índice (*Step-index multimode fiber*), mientras que si la relación entre ambos va decreciendo progresivamente, se denominan fibras multimodo de índice gradual (*grad-index multimode fiber*). Dado que el tipo de fibra además está relacionado con el tamaño de su núcleo, esto determina un rango más adecuado de longitudes de onda que suelen usarse en cada uno de ellos².

En comunicaciones digitales de datos, la señal transmitida consiste generalmente en pulsos de luz para transmitir los símbolos en los que se codifican los datos. Habitualmente se usan codificaciones de símbolo con constelaciones sencillas de modulación en amplitud de tipo ASK (*Amplitude Shift Keying*), también denominada OSK (*ON-OFF Shift Keying*), con generación de símbolos de pulso estrecho NRZ (Non-Return to Zero) [2], de forma que los bits transmitidos se identifiquen por la presencia o no de luz. Debido a las características físicas intrínsecas de las fibras y a la forma en la que la luz se propaga a través de los cristales de silicio, aparecen algunos inconvenientes que han de ser tenidos en cuenta a la hora de usar las fibras ópticas como medios

²<http://www.thefoa.org> - The Fiber Optics Association

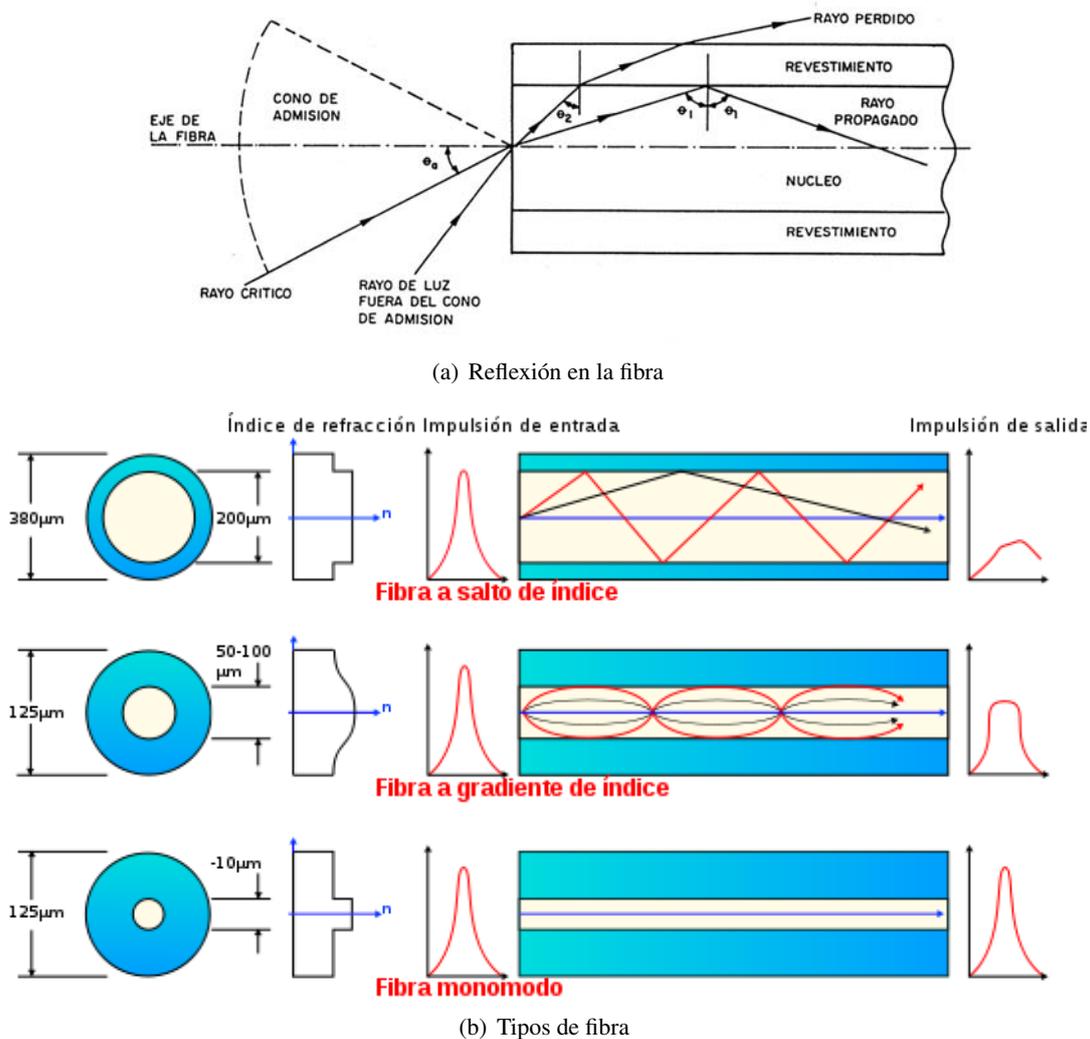


Figura 2.3: Reflexión de la luz en el interior de la fibra (a) y tipos de fibras (b)

de transmisión de alta velocidad en comunicaciones digitales debido a que afectan a la calidad el pulso de luz recibido:

Atenuación Aunque se ha comentado que una de las ventajas de la fibra es su baja atenuación frente a los cables coaxiales y de par trenzado, ésta no deja de existir. Sin embargo, hablamos de una atenuación perfectamente admisible en enlaces de varias decenas de kilómetros. (0.2 - 0.5 dB/km). Al igual que en otros medios de transmisión, la atenuación depende directamente de la pureza de los materiales de construcción de las fibras. En este caso la atenuación está relacionada con la pureza de los cristales de silicio y su transparencia. Estas características afectan a la luz transmitida de forma diferente en función de la frecuencia de la misma. Es por ello que, en general, se seleccionan tres rangos de longitudes de onda más aptos para la transmisión a través de la fibra óptica en función de la curva de atenuación. En

la Figura 2.4, se muestra la atenuación en función de la distancia en kilómetros y las tres ventanas más aptas para la sintonización de los dispositivos transmisores.

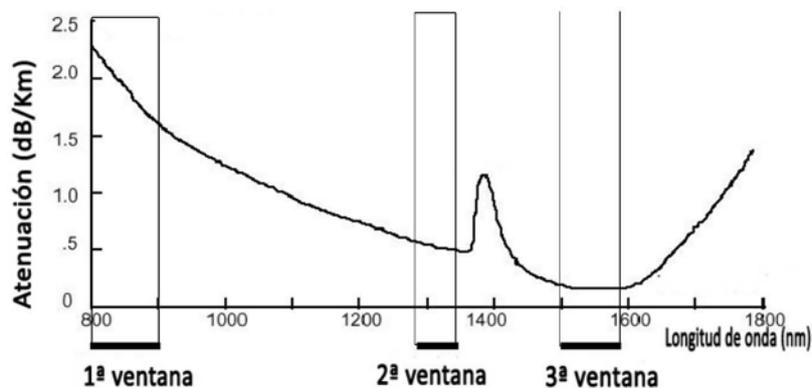


Figura 2.4: Atenuación en la fibra óptica en función de la longitud de onda y la distancia

Dispersión La dispersión es el efecto por el cual las diferentes componentes de la señal transmitida viajan por caminos distintos por el medio físico, lo cual se traduce en un retardo de propagación distinta extremo a extremo. Este efecto se va acumulando a lo largo de la fibra por lo que es dependiente de su longitud. En comunicaciones digitales, el resultado de este efecto es un ensanchamiento en el pulso transmitido, el cual interfiere en los adyacentes, provocando la denominada interferencia intersimbólica (ISI - *Inter Symbolic Interference*). Existen varias razones por las que el pulso puede verse afectado por efectos de dispersión:

- **Dispersión modal:** Cuando las dimensiones de la fibra permiten que varios modos de luz se propaguen a la longitud de onda seleccionada, lo que sólo ocurre en fibras multimodo, cada uno seguirá distintos caminos físicos dentro del núcleo de silicio debido a que cada uno de ellos sufrirá diferentes ángulos de refracción. Por lo tanto, cada uno de ellos llegará al extremo final de la fibra con un retardo distinto. En las fibras de salto de índice, las diferencias relativas en los retardos de cada modo serán mayores que en las de índice gradual. La graduación del salto de índices permite equalizar la dispersión modal, aumentando el producto ancho de banda-distancia, pero con un coste de producción más elevado.
- **Dispersión por guía de ondas:** este tipo de dispersión depende de cómo afectan las características de guiado de la fibra a las diferentes longitudes de onda. En función de los índices de transmisión, de la forma de la fibra y del núcleo y su revestimiento, parte de la energía transmitida se propaga también por el recubrimiento del núcleo ya que la

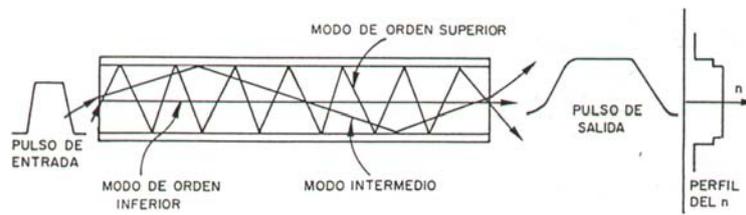


Figura 2.5: Efecto de la dispersión modal sobre un pulso óptico

reflexión no es completa, con lo que cada parte de la señal viajará por medios distintos y por tanto con velocidades distintas.

- **Dispersión Cromática:** la dispersión material o cromática es aquella relacionada con el hecho de que el índice de refracción de los materiales depende de la longitud de onda de la señal transmitida. Dado que el pulso de luz no es infinitamente estrecho, el índice que sufre no es constante y provoca un ensanchamiento del pulso en recepción. La dispersión cromática se mide en $\text{ps}/(\text{nm km})$, donde ps es el tiempo de ensanchamiento del pulso, en función de la longitud de onda central en nm y de la longitud de la fibra en km. Un valor típico para una fibra monomodo es de 17 ps/nm km a una longitud de onda de trabajo de 1550 nm.
- **Dispersión por modos de polarización:** este tipo de dispersión surge del hecho de que el núcleo de la fibra no es perfectamente circular. Estas imperfecciones de fabricación provocan que diferentes polarizaciones de la señal tengan diferentes velocidades de propagación. Se ha comprobado que este tipo de dispersión es particularmente dañina a velocidades superiores a 10Gbps.

No-linealidades: dado que en general, la potencia de transmisión en la fibra es pequeña, ésta se suele modelar como un medio lineal. Sin embargo, cuando es necesario introducir potencias más elevadas en la fibra, aparecen efectos no lineales que suponen limitaciones a sistemas de alta velocidad y a la multiplexación en longitud de onda (WDM). Existen no-linealidades debidas a los efectos que provoca el hecho de que el índice de refracción depende de la potencia de la señal, y otro grupo que aparece debido a fenómenos de dispersión por la interacción de las ondas de luz con partículas vibrantes en el núcleo de silicio, cambiando su dirección y su longitud de onda.

Interferencia cruzada (Crosstalk): La interferencia cruzada disminuye la relación señal-ruido (SNR - *Signal to Noise Ratio*), y por tanto la calidad, de la señal recibida aumentando la

probabilidad de error en bit (BER - *Bit Error Rate*) en el destino. Esta interferencia viene provocada por señales que comparten la misma fibra pero con diferentes longitudes de onda (interferencia entre canales) o por señales de la misma longitud de onda (interferencia cocanal) en fibras adyacentes debidas a las imperfecciones de los materiales de las mismas, o de los elementos de interconexión y conmutación. La primera de ellas limita el ancho del canal y puede eliminarse usando filtrados de banda, pero la segunda es más difícil de eliminar en caso de aparecer.

Ruido: Finalmente, la transmisión en la fibra puede verse degradada por diversas fuentes de ruido térmico. Las principales fuentes de ruido son los transmisores de luz y los amplificadores, pero también pueden existir fuentes de ruido en los fotodiodos detectores debido a sus componentes electrónicos.

Teniendo en cuenta los efectos de la atenuación y de la dispersión, los valores típicos de funcionamiento dentro de las fibras ópticas tradicionales son ventanas de 200nm con una atenuación por debajo de los 0.5dB/km centrada en 1300nm con un ancho de banda total de 25 THz en esta zona de funcionamiento (2ª ventana - Fig. 2.4). Para aplicaciones más avanzadas usando láseres de mayor longitud de onda, la 3ª ventana proporciona atenuaciones mucho más bajas (0.2 dB/km). Para distancias más cortas, y debido a su mayor atenuación, puede usarse la 1ª ventana (correspondiente a fibras de menos calidad o las que tienen núcleos de plástico cuyas atenuaciones pueden llegar a orden de 100dB/km [3]) que llega hasta frecuencias próximas al infrarrojo cercano (800nm) y a espectros visibles (fuentes de luz roja de 650nm e incluso azules a 460nm), lo que permite el uso de láseres más baratos y de dispositivos LED, a pesar de que a estas longitudes de onda, la dispersión modal es muy grande.

En resumen, con estas características, parece adecuado dedicar fibras monomodo de altas prestaciones a enlaces de larga distancia o redes de acceso ópticas o híbridas (por ejemplo redes de acceso FTTH - *Fiber To The Home* o mixtas junto con DSL -*Digital Subscriber Lines* o coaxial (HFC - *Hybrid Fiber Coax*) [3] donde los enlaces son del orden de decenas de kilómetros. Para enlaces más cortos, como podría ser en entornos con redes de área local (LAN - *Local Area Network*), oficinas o edificios, las fibras de silicio multimodo o incluso las construidas con polímeros plásticos cumplen su función a la perfección.

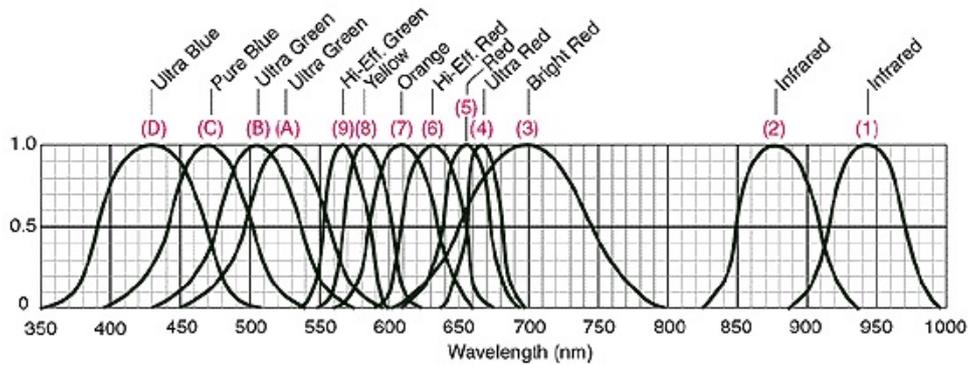


Figura 2.6: Anchos de banda normalizados de fuentes de luz comerciales tipo LED

2.2.2. Dispositivos de transmisión y recepción

Otros componentes importantes en los enlaces ópticos son tanto los transmisores como los receptores de señal óptica. Ambos dispositivos consisten no sólo en el generador y receptor de luz propiamente dichos sino en la electrónica asociada a su funcionamiento y que es necesaria para el filtrado, la modulación y la demodulación de la señal. Estos circuitos electrónicos además serán fuente de ruidos, y limitarán la velocidad final de funcionamiento del sistema.

Como dispositivos transmisores se utilizan fundamentalmente dos familias: láseres y LEDs, mientras que el dispositivo receptor es habitualmente un diodo fotodetector o fotodetector de avalancha, acompañado de los filtros y circuitos electrónicos necesarios.

2.2.2.1. LEDs (*Light Emitting Diode*)

Los diodos fotoemisores tipo LED son una fuente de luz de banda ancha barata y eficiente, y muy comunes en comunicaciones. Habitualmente son utilizados para enlaces de corto alcance y baja capacidad debido a que su producto ancho de banda por distancia es bajo. Sin embargo, debido a su bajo coste de fabricación y su eficiencia en cuanto a emisión de luz, siguen investigándose nuevos tipos superluminiscentes. Las longitudes de onda de trabajo habituales de un LED pueden verse en la Fig. 2.6³.

2.2.2.2. Láseres

El término *láser* es el acrónimo de "Light Amplification by Stimulated Emission of Radiation" (amplificación de luz mediante la radiación por emisión estimulada). La emisión estimulada

³http://www.oksolar.com/led/led_color_chart.htm

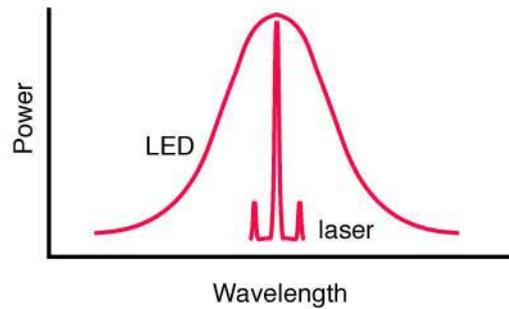


Figura 2.7: Comparativa entre la forma de espectro de un láser y de un diodo LED sintonizados a la misma frecuencia

permite generar haces de luz estrechos y direccionales de alta potencia y de forma coherente (con una o varias frecuencias concretas), generalmente en una sola longitud de onda, aunque el dispositivo puede reconfigurarse para poder emitir en más de una frecuencia, con una curva de intensidad espectral mucho más estrecha que la de un diodo LED alrededor de dicha frecuencia (Fig 2.7). En la actualidad los emisores de tipo láser pueden operar en todo el rango de frecuencias ópticas aptas para las fibras ópticas existentes en el mercado (Fig 2.8⁴).

Como se explicará más adelante, en WDM, es necesario disponer de varias frecuencias de trabajo, por lo que podemos tener una matriz de láseres sintonizados cada uno de ellos a longitudes de onda fija diferentes o un sólo dispositivo láser sintonizable a diferentes longitudes de onda. En el caso de transmisores y receptores sintonizables, es imprescindible que los transmisores ópticos cumplan con el requisito de permitir una sintonización, en una o varias longitudes de onda, lo más fina posible y a gran velocidad. Algunas de las características intrínsecas de los dispositivos de emisión láser que más influyen en las prestaciones de sistemas WDM son:

- el ancho espectral del láser, que afecta a la anchura efectiva de los canales y a la dispersión modal en la fibra,
- las inestabilidades en frecuencia, que son las pequeñas variaciones que puede sufrir un láser durante su funcionamiento y que afectan a la sintonización de la frecuencia deseada. Estas inestabilidades pueden deberse a inestabilidades en las corrientes eléctricas de la electrónica de amplificación del láser y/o a cambios de temperatura,
- el número de modos o longitudes de onda que el láser es capaz de amplificar, siendo lo más deseable que el láser amplifique sólo un modo para evitar un aumento de la dispersión. Esta característica depende directamente de las características físicas y de los materiales del láser

⁴Fuente de la Fig. 2.8 http://en.wikipedia.org/wiki/List_of_laser_types

con lo que también influye en su coste de fabricación,

- otras características deseables son el rango de longitudes de onda sintonizables, la velocidad de conmutación entre ellas y si el láser puede sintonizarse de forma continua o discreta (sólo ciertas longitudes de onda son sintonizables).

2.2.2.3. Filtros ópticos y fotodetectores

Los filtros ópticos se utilizan para seleccionar de una señal en banda ancha sólo el espectro del canal deseado o también para seleccionar una sola longitud de onda de todas las que se transmiten por el mismo medio. La señal filtrada es convertida de óptica a electrónica mediante un fotodetector. Tanto los filtros como los fotodetectores pueden ser sintonizables dinámicamente (en frecuencias continuas o discretas) o fijos en frecuencia, siendo una alternativa a los primeros una matriz de receptores sintonizados a varias frecuencias fijas. Sus características más importantes, al igual que en los transmisores, serán su rango de sintonización y la velocidad de selección de la frecuencia sintonizada.

2.2.3. Componentes de sistemas de transmisión y conmutación en redes ópticas

Para la construcción de sistemas de comunicaciones ópticas, existen diversos componentes fundamentales que son utilizados para lograr la conmutación de las diversas señales ópticas según la manera en las que se las quiera hacer llegar entre los nodos de la red. A continuación se describen brevemente los bloques básicos más relevantes y más utilizados en las topologías ópticas habituales, centrándonos sólo en sus funcionalidades como bloques de conmutación, pudiendo el lector referirse a literatura especializada en la materia para profundizar en los detalles de su construcción y funcionamiento a nivel físico [2], [4], [5].

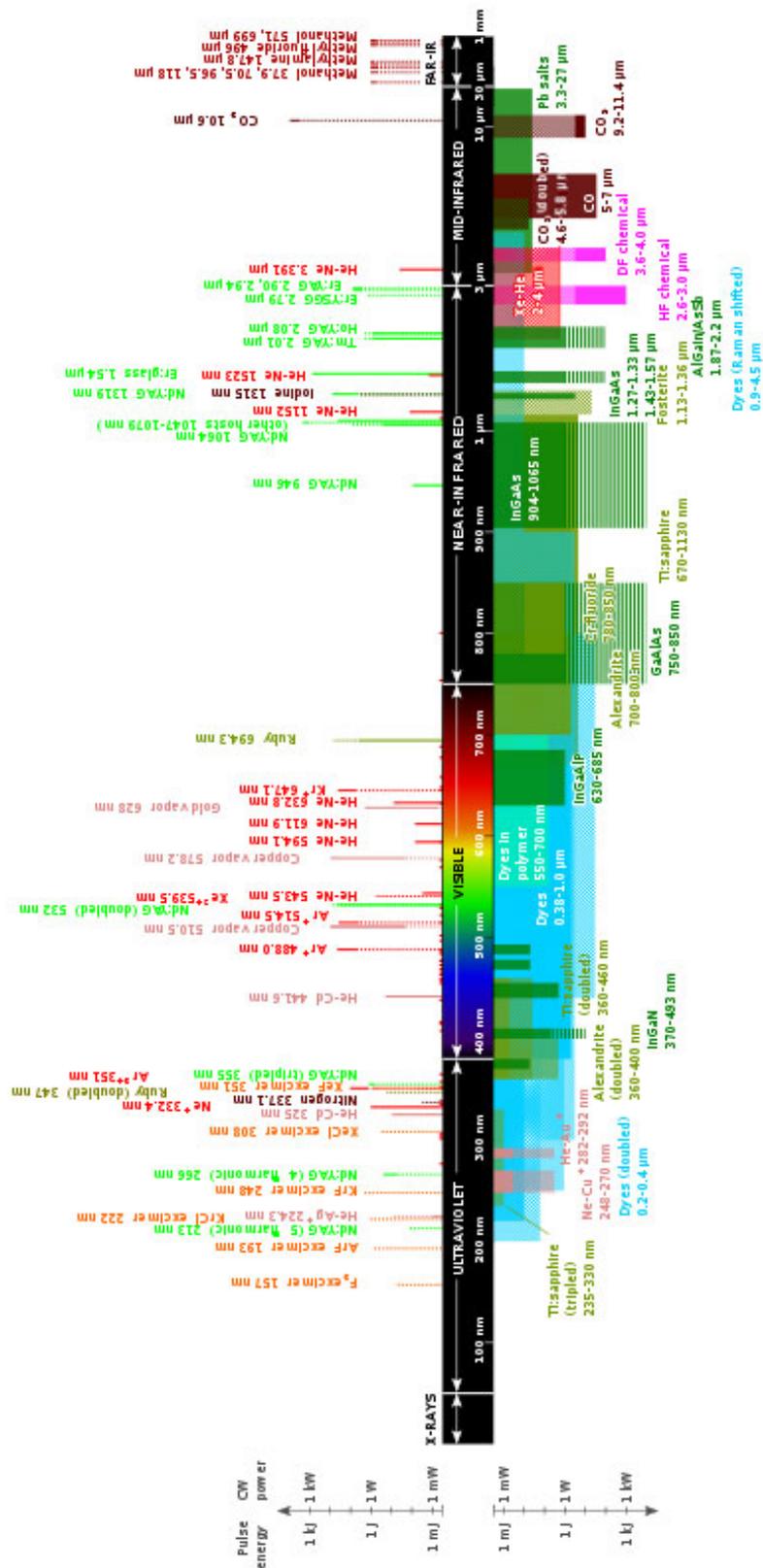


Figura 2.8: Frecuencias disponibles en de láseres comerciales

2.2.3.1. Bloques básicos pasivos de redes ópticas

Mezcladores (*combiners*) - $S \times 1$ Este bloque de conmutación permite combinar S entradas en una sola salida. Todas las longitudes de onda recibidas en cada una de sus S entradas son transmitidas mezcladas en el puerto de salida. En general, para evitar interferencias entre canales, aquellos que se van a combinar han de ser distintos o pertenecer a diferentes longitudes de onda.

Divisores (*splitters*) - $1 \times D$ Un *splitter* permite dividir la señal de su entrada entre los D puertos de salida, de forma que por cada uno de ellos se envíen todas y cada una de las longitudes de onda de entrada. La potencia de la señal de entrada queda dividida entre las D salidas.

Acopladores (*Couplers*) Un acoplador puede definirse como un mezclador seguido a continuación de un divisor.

Multiplexores (o mezcladores de bandas) Los multiplexores, escogen una longitud de onda proveniente de un puerto de entrada y la mezclan con la escogida en el resto de puertos, transmitiendo cada una de ellas en la salida. A diferencia de los mezcladores, en los puertos de entrada se selecciona una banda concreta, no se retransmite la señal completa que llega a la puerta de entrada, como puede verse en las Figuras 2.9(a) y 2.9(d).

Demultiplexores (o divisores de banda) La operación complementaria a la anterior la realiza el demultiplexor, bloque capaz de separar la longitud de onda que se desea transmitir por cada uno de sus puertos de la señal óptica recibida en el puerto de entrada.

Combinando estos bloques sencillos podemos construir conmutadores y encaminadores pasivos que configuran las rutas para los flujos de forma estática en los nodos de interconexión.

2.2.3.2. Otros bloques de conmutación óptica

Los bloques descritos en la sección 2.2.3.1 son todos ellos pasivos, esto es, no necesitan de alimentación externa para su funcionamiento. Combinando estos bloques se pueden construir topologías de conmutación sencillas para realizar comunicaciones punto a punto y sistemas de difusión *broadcast* y *multicast*, jugando con las posibles longitudes de onda que van a poder llegar a cada extremo. Pero para dotar a la red de flexibilidad en cuanto a número de flujos, nodos y/o usuarios, es necesario implementar bloques que permitan la reutilización de las mismas longitudes de onda

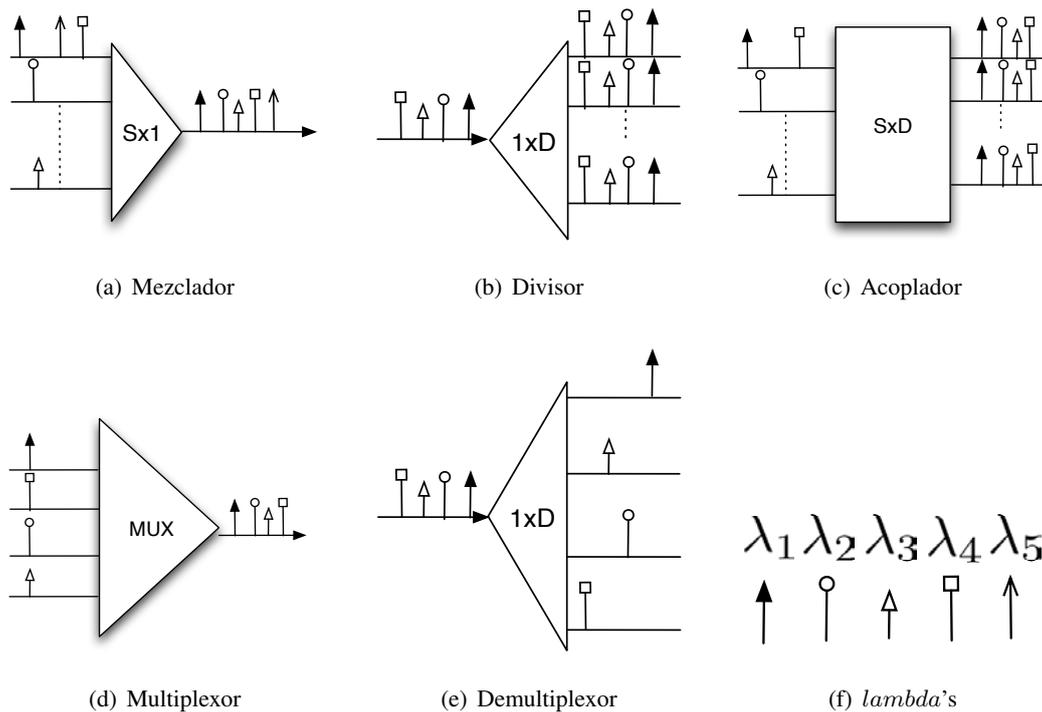


Figura 2.9: Esquemas funcionales de componentes básicos pasivos

para varios flujos de tráfico distintos en los nodos de interconexión y de esta forma flexibilizar el encaminamiento (reutilización espacial). Además, permitiendo la reconfiguración de los bloques mediante señales electrónicas, se puede controlar la configuración de conmutación del nodo de forma dinámica en función de las necesidades de encaminamiento.

Los bloques de conmutación electro-ópticos más importantes son:

2x2 Cross-Bar Switch Este elemento consiste básicamente en una matriz de conmutación de dos entradas y dos salidas. Mediante una señal de control se configura el bloque de forma que la señal del puerto de entrada salga por su correspondiente puerto de salida (modo puente o "bar o bypass"), o se cruzan las salidas de ambos puertos, quedando conectado la primera salida con la segunda entrada y la segunda con el primer puerto de entrada (modo cruce o "cross") tal y como puede verse en la figura 2.10).

OADM - Optical ADD-Drop Multiplexer El bloque básico de un OADM (Fig. 2.11) está formado por un demultiplexor en la entrada para separar las longitudes de onda. Cada uno de los puertos del multiplexor se conecta a uno o a varios bloques Cross-Bar controlados electrónicamente, de forma que se pueda seleccionar en qué conjunto de longitudes de onda se desean recibir y/o transmitir en cada momento ("add/drop") y cuales se quieren dejar pasar intactas,

cambiando el modo de funcionamiento del Cross-Bar. Aquellas longitudes de onda para las que el Cross-Bar esté en modo de cruce ("cross") son extraídas de la red y reemplazadas por nuevos datos modulados en la misma banda. Finalmente, un multiplexor vuelve a mezclar las diversas frecuencias en una sola fibra de salida. Por ejemplo en la figura 2.11, la tercera frecuencia que es extraída en el nodo, recibiendo así los datos enviados para dicho nodo y reemplazándolos por datos nuevos en la misma longitud de onda, mientras que el resto de frecuencias se dejan pasar a través del resto de cruces del OADM en modo *bypass*.

Los OADMs actuales admiten la configuración remota de cada una de sus entradas y salidas, permitiendo el desarrollo de redes ópticas malladas complejas mediante la conexión de los puertos y también la selección dinámica de las longitudes de onda de funcionamiento (ROADM -*Reconfigurable OADM*).

OXC - Optical Cross Connect El bloque OXC realiza una función equivalente al anterior pero a gran escala. En resumen, un OXC es un dispositivo con un gran número de puertos de entrada y salida, que es capaz de conmutar "lambdas" de un puerto a otro a demanda de la configuración de la red. La tecnología de red que se conecta en cada puerto puede ser distinta, lo que permite interconectar equipos WDM, SONET o conmutadores IP o ATM entre sí, lo que supone la posibilidad de crear topologías malladas complejas en redes ópticas. El OXC (Fig. 2.12 se compone de un sistema central que realiza la conmutación y la multiplexación de señales y de un conjunto de tarjetas de puertos que pueden o no incorporar la tecnología necesaria para realizar conversiones OE (Óptica-Electrónica) y/o OEO (Óptica-Electrónica-Óptica).

Tanto los OADM como los OXC pueden incorporar sistemas de conversión de longitudes de onda que permitirían aumentar aún más la flexibilidad de reutilización espacial del dispositivo de conmutación.

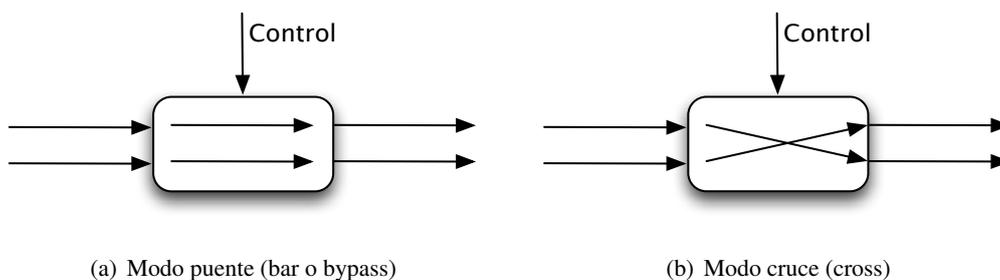


Figura 2.10: Esquema de un Cross-Bar Switch en ambos modos de funcionamiento

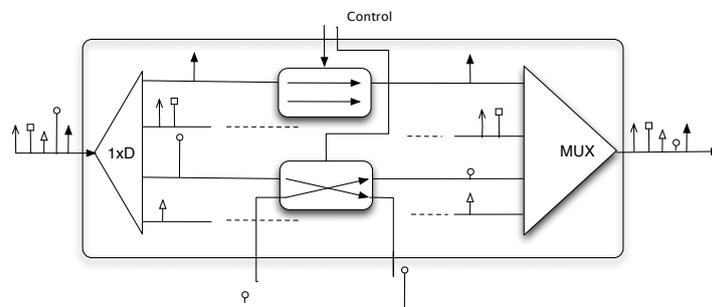


Figura 2.11: Esquemas funcional de un OADM

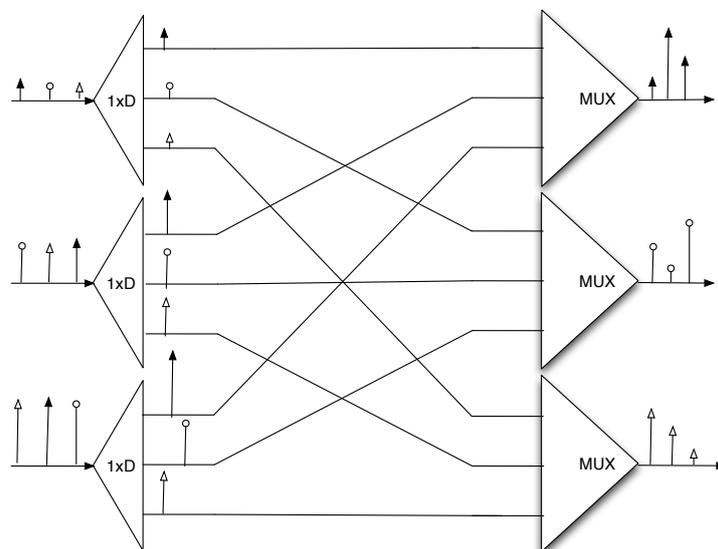


Figura 2.12: Esquemas funcional de un OXC

2.3. Topología, conmutación y control en una red óptica

2.3.1. Transparencia en redes ópticas

En los años 80, la tecnología óptica comenzó a desplegarse para enlaces punto a punto. La extensión de estos enlaces a topologías más complejas exigía que en cada punto de interconexión la señal óptica fuera convertida a señal eléctrica, procesada, y vuelta a convertir en señal óptica (conversión OEO). De esta forma, topologías sencillas como anillos y estrellas eran desplegadas usando nodos capaces de realizar la conversión OEO y enlaces ópticos entre los puntos de interconexión.

Mediante el uso de bloques de conmutación sencillos como los vistos en la sección 2.2.3.1, se comenzaron a construir topologías más complejas donde la señal óptica era replicada mediante acopladores en estrella (PSC - Passive Star Couplers) para conseguir topologías de distribución broadcast de un salto. Los anillos ópticos se podían construir interconectando en cascada enlaces ópticos con sus consecuentes multiplexores y demultiplexores para realizar la conversión OEO de cada una de las longitudes de onda (λ 's) que viajan por el anillo.

El término "*transparencia*" en redes ópticas se usa para designar aquellas redes en las que la comunicación extremo a extremo se hace siempre dentro del plano óptico, sin conversiones OEO en los nodos intermedios [6]. Por tanto, para los protocolos de orden superior, la red es completamente transparente, comportándose como un enlace directo punto a punto.

2.3.2. Multiplexación en el dominio óptico

En términos generales, la multiplexación permite a diferentes fuentes de tráfico compartir los recursos de que disponen para encaminar los datos al receptor. En el caso de la multiplexación en comunicaciones ópticas, esta permite compartir el medio de transmisión a diferentes fuentes de tráfico. Las formas en la que se organiza el medio compartido dividen las diferentes soluciones en tres grandes familias: multiplexación por división en tiempo (TDM - *Time Division Multiplexing*), en espacio (SDM - *Space Division Multiplexing*) y en frecuencia (WDM - *Wavelength Division Multiplexing*). A continuación se comentan las principales diferencias entre cada una de ellas.

2.3.2.1. TDM

La multiplexación por división en el tiempo es una técnica bien conocida que lleva usándose en muchos campos de las comunicaciones digitales desde la mitad del siglo XX. Esta solución asigna "ranuras" de tiempo en instantes concretos a las fuentes de tráfico para comenzar transmitir y hacer uso del medio compartido, de forma que no exista colisión de los datos ya que cada fuente tiene permiso para transmitir en la "ranura" o período de tiempo asignado. Su sencillez y la experiencia de su uso en otro tipo de tecnologías de comunicación permitieron su implantación en las redes punto a punto SONET a finales de los años 80.

En las redes ópticas actuales, TDM sigue utilizándose sobre todo en las redes de acceso ópticas pasivas (PON - *Passive Optical Networks*). Sin embargo, algunas de sus características intrínsecas se han convertido en inconvenientes para su uso en redes de muy alta velocidad: la principal de ellas es que los nodos de conmutación han de funcionar a la velocidad de la línea para gestionar la sincronización de todas las fuentes de tráfico, en lugar de a la velocidad de generación de datos de dichas fuentes (en general mucho menor que la de transmisión óptica). Esto condiciona la máxima velocidad de funcionamiento del sistema, limitado por la máxima velocidad de la electrónica disponible para la fabricación de estos equipos. La consecuencia de este inconveniente es que las redes TDM no puedan sacarle el máximo partido al enorme ancho de banda disponible en el escenario de la transmisión óptica.

Otro aspecto que limita su despliegue es la precisión de los dispositivos ópticos en cuanto a la dispersión óptica en el medio de transmisión. A velocidades superiores a 100Gbps, el coeficiente de dispersión óptica de la fibra no es despreciable en distancias de decenas de kilómetros. Un alto coeficiente de dispersión provoca un aumento de la interferencia entre los símbolos transmitidos (ISI) proporcional tanto a la velocidad de transmisión como a la distancia. Es por esto que las redes que utilizan TDM como técnica de multiplexado sean normalmente redes de baja capacidad o de corto alcance, para que el efecto de la dispersión modal sea mínimo. Sin embargo, la creciente evolución de los dispositivos ópticos de pulso estrecho ha aumentado el uso actual de redes TDM.

Otra característica a tener en cuenta es la precisión en la sincronización de los equipos de transmisión. Dado que cada fuente tiene asignado un instante de comienzo y un período de transmisión, cualquier error en la sincronización puede provocar colisiones con los datos transmitidos por otras fuentes. Además del uso de equipos más precisos, las redes que utilizan TDM suelen implementar protocolos que dejan márgenes de guarda entre las transmisiones de una y otra fuente

para evitar dichas colisiones. Pero estos tiempos de silencio actúan en detrimento de la utilización del medio compartido y por tanto disminuyen el ancho de banda efectivo de los sistemas.

Otra de las características de las redes ópticas que usan TDM como forma de acceso es que no proveen a priori de *transparencia* a cualquier protocolo de comunicaciones que utilice el usuario de la red. Dado que el tráfico de cada cliente ha de regirse por el subyacente criterio de la sincronización, la red no es completamente transparente ya que los clientes tendrán que adaptar sus protocolos y temporizaciones a la estructura de tramas del multiplexado TDM. Para evitar esto, los nodos intermedios han de ser dotados de complejidad extra de forma que en cualquier situación los datos del cliente permanezcan en el dominio óptico sin necesidad de conversiones electro-ópticas (OEO).

2.3.2.2. SDM

Una forma clásica de evitar el cuello de botella de las conversiones OEO en redes ópticas es usar multiplexado por división espacial, esto es, muchas fibras utilizadas en paralelo en lugar de una sola fibra para el enlace. Cada fibra óptica puede operar como un enlace independiente, permitiendo así flexibilidad en cuanto a ancho de banda o longitudes de onda a utilizar. Pero este sistema claramente adolece de falta de escalabilidad, por lo que habitualmente sólo se usa para enlaces de corto alcance, y con un número relativamente bajo de usuarios o flujos de tráfico a conmutar.

2.3.2.3. WDM

WDM aparece como la solución más flexible para sacar el máximo partido al gran ancho de banda disponible en la fibra óptica y para solventar los problemas de TDM y SDM. En WDM el tráfico de cada cliente se envía en una longitud de onda (λ) distinta, de la misma manera que tradicionalmente se hace en redes multiplexadas FDM (Frequency Division Multiplexing) con la frecuencia. En el extremo origen, un multiplexor se encarga de recoger la señal de cada portadora e inyectarla en la fibra óptica, siendo necesario un demultiplexor en el destino para separar cada una de las *lambda's* de los clientes para que sea enviado al receptor correspondiente. De esta forma, cada longitud de onda puede transportar un cierto número de canales que no se solapan con los de otras longitudes de onda.

A diferencia de TDM, cada longitud de onda se puede transmitir a diferente velocidad de

línea incluso por debajo de la frecuencia agregada que se usa en TDM. Al contrario que SDM, WDM explota todo el potencial de la fibra en cuanto a ancho de banda se refiere sin necesidad de instalar otras fibras en paralelo para realizar la conmutación, lo que garantiza un coste menor de despliegue. Además, una gran parte del despliegue de las redes WDM se consigue gracias a los componentes vistos en la sección 2.2 de este capítulo, todos ellos sencillos de construir e implementar en los nodos de interconexión.

El despliegue actual de redes de próxima generación WDM se está llevando a cabo en redes troncales que evolucionan hacia un paradigma de IP-sobre-WDM. De esta forma, los clientes finales de las operadoras disponen de una red troncal transparente para sus protocolos y topologías de niveles superiores, y con mejores prestaciones en cuanto a capacidad y tiempos de respuesta. Esto es posible gracias a bloques ópticos más avanzados que permiten la reconfigurabilidad de la red y la conversión de longitudes de onda y su encaminamiento de forma totalmente óptica [2].

2.3.3. Granularidad en los sistemas de conmutación

En los sistemas de conmutación óptica, la granularidad en la que trabajan los nodos intermedios permite dotar a la red de flexibilidad y escalabilidad. Al igual que en otros tipos de sistemas de comunicaciones, en redes ópticas existen muchos tipos de subdivisiones en la forma en la que se comparte el medio en relación a la cantidad de recursos que se reservan y a su flexibilidad para compartirlos:

- Conmutación a nivel de fibra: usando SDM con fibras dedicadas para cada enlace, proporcionando la mínima posibilidad de aumentar la granularidad del sistema.
- El siguiente nivel de granularidad sería el que puede conseguirse mediante la división en canales adyacentes de las longitudes de onda disponibles, de forma que los nodos intermedios puedan conmutar entre los distintos canales que viajan por la misma fibra.
- Cuando cada canal contiene una sola longitud de onda, estamos en un caso particular del anterior que proporciona una granularidad más fina. Puede asignarse a cada cliente una longitud de onda o usarse de forma mixta, por ejemplo con TDM, de forma que intervalos de tiempo distintos de la misma *lambda* pertenezcan a clientes diferentes. Esta forma de conmutación, a diferencia de las anteriores permite repartir por tanto la longitud de onda en varias *sub-lambdas*.

Las subdivisiones anteriores forman parte en mayor o menor medida de sistemas de conmutación de circuitos (OCS - *Optical Circuit Switching*).

- Cuando se permite la conmutación independiente de cada paquete de una longitud o sub-longitud de onda, se denomina OPS (*Optical Packet Switching*), siendo esta la forma de granularidad más fina que puede conseguirse.

- Finalmente, existe la posibilidad intermedia de realizar la conmutación entre ráfagas de paquetes de distintas λ 's lo que se denomina OBS - *Optical Burst Switching*

2.3.3.1. Conmutación de circuitos: OCS

OCS configura circuitos cerrados extremo a extremo entre los nodos de la red reservando una fibra, canal o longitud de onda para cada par de nodos, y una vez el camino está configurado, los datos permanecen en el dominio óptico hasta su destino, sin necesidad de realizar conversiones OEO intermedias.

Como todo sistema de conmutación de circuitos, OCS sólo es eficiente cuando existe un gran volumen de tráfico entre dos nodos para los que merezca la pena configurar un enlace dedicado. De otro modo, si la duración del tiempo de transmisión de datos es corta en relación con el tiempo de establecimiento de la conexión, estamos desaprovechando mucho ancho de banda. Además, y dado que el número de longitudes de onda disponibles es en general muy limitado, la utilización de OCS degenera en problemas de flexibilidad y escalabilidad: es imposible generar una malla de caminos ópticos estáticos entre todos los nodos. Para cada topología y matriz de tráfico concretas se debe resolver el problema del encaminamiento y disponibilidad de las longitudes de onda por separado, no existiendo algoritmos que lo resuelvan en un tiempo aceptable. Si además tenemos en cuenta que son redes estáticas, esto conlleva una baja adaptabilidad a un tráfico cambiante.

A cambio, OCS permite proporcionar circuitos con un alto nivel de calidad de servicio garantizado, asumiendo la latencia que suponga el establecimiento del circuito. Las operadoras actuales pueden ofrecer circuitos dedicados a través de fibra óptica de diferentes anchos de banda durante el tiempo que el cliente contrate el servicio.

2.3.3.2. Conmutación de paquetes: OPS

Para solventar el problema de la limitación en cuanto al número de *lambdas* de las redes WDM, los paquetes que los clientes transmiten entremezclados en las longitudes de onda disponibles, se conmutan de forma independiente en OPS. Esta forma de conmutación aprovecha de manera más eficiente el ancho de banda comparado con OCS. En este caso, el tráfico de usuario se transporta junto con la información de control, siendo los nodos intermedios los que se encargan de procesarla y realizar el encaminamiento hacia su destino. La conmutación de las diversas ráfagas de paquetes se realiza mediante multiplexación estadística en OPS, consiguiendo reducir el ancho de banda estimado para los enlaces de salida: dado que normalmente los flujos de datos mandan ráfagas de paquetes cuando están activos, se asume que la probabilidad de que todos estén activos a la vez es pequeña y se dimensiona el enlace con un ancho de banda menor que aquel que se usaría si todos los flujos estuviesen activos. Esto supone un aprovechamiento más eficiente del ancho de banda pero requiere implementar sistemas de colas de espera en los nodos para evitar pérdidas en momentos de picos de tráfico.

Esto supone que la tecnología de los conmutadores de los nodos intermedios, debe estar preparada para recibir la información de control y realizar la conversión OEO y su procesamiento a velocidad de línea, aunque los datos viajen de forma transparente en el dominio óptico. Los datos del paquete se retienen en memoria óptica (mediante el uso de FDLs - Fiber Optical Delays) hasta que se decide por dónde debe ser enviado, y llegado ese momento, se debe configurar la matriz de conmutación en ese nodo. Por lo tanto, el retardo total será la suma de la conversión óptico-eléctrica de la cabecera, más el tiempo de configuración de la matriz de conmutación, más el tiempo de reconstrucción de la cabecera del paquete en el dominio óptico.

Los FDLs se diseñan de forma que los paquetes más grandes puedan ser almacenados en ellos, lo que provoca grandes retardos innecesarios para los paquetes pequeños, por lo que OPS se utiliza habitualmente para casos donde se pueda fijar el tamaño de paquete.

2.3.3.3. Conmutación basada en ráfagas: OBS

Como una solución intermedia entre OCS y OPS para poder obtener los beneficios de ambas, aparece la técnica de conmutación a ráfagas OBS [7]. Esta técnica proporciona otra alternativa de conmutación en entornos WDM. En OBS, la conmutación se hace a nivel de ráfagas de paquetes y no de paquetes independientes. De esta forma se permite la multiplexación estadística de ráfagas

x	Usó BW	Latencia	Tolerancia fallos	Granularidad	Dificultad impl
OCS	Bajo	Alta	Baja	Baja	Bajo
OPS	Alto	Baja	Alta	Alta	Alto
OBS	Moderado	Baja	Media	Moderada	Moderado

Tabla 2.2: Resumen de características de tipos de conmutación óptica

de distintos flujos entre varias longitudes de onda, aunque la tecnología de conmutación óptica tenga tiempos de conmutación superiores al tiempo de transmisión del paquete. Dependiendo de las características del tráfico a conmutar, los resultados se acercan a las prestaciones de OCS cuando las ráfagas sean de tamaño grande y enviadas de forma continua, mientras que se diferenciará poco de OPS cuando se trate de enviar ráfagas muy pequeñas a intervalos de tiempo separados.

En el caso de OBS, existe un canal exclusivo para control por el que viaja la información de conmutación de las ráfagas para preconfigurar los nodos intermedios. Las ráfagas se ensamblan en el nodo origen con los paquetes almacenados en memoria con el mismo destino, de forma que puedan compartir cabecera. En paralelo, por el canal de control, se envía la información de encaminamiento de dicha ráfaga, de forma que cada conmutador tenga preconfigurado el puerto y longitud de onda de salida correspondientes y no sufra bloqueos. Estos bloqueos pueden llegar a producirse dado que el número de longitudes de onda de los nodos intermedios es limitado y pueden darse situaciones de contienda por su utilización. A la hora de diseñar la red OBS es necesario contar con el hecho de que el paquete de control necesita de procesado y conversión OEO, lo que marca también un límite para la capacidad de ráfagas que pueden enviarse.

Existen muchas familias de algoritmos de planificación de ráfagas para preconfigurar los conmutadores de una red OBS dependiendo de la manera en que se realiza la reserva de los recursos, pero la más habitual en las implementaciones que investigan esta técnica es JET (*Just Enough Time*) [8].

2.3.4. Gestión y control de Redes Ópticas

Como se ha comentado en las secciones anteriores, una de las principales características de las redes ópticas de nueva generación es el establecimiento dinámico de circuitos. El establecimiento, configuración, modificación y liberación de las conexiones ópticas en los conmutadores, es independiente de los datos que se están transmitiendo a través de la red óptica usando protocolos "fuera de banda".

La conmutación propiamente dicha, en cualquiera de sus granularidades vistas en la sección 2.3.3 se lleva a cabo en el plano de datos. El plano de datos se compone de las capas y dispositivos necesarios para asegurar la transmisión, la recepción y el reenvío de los datos entre origen y destino. Para que estas operaciones se realicen adecuadamente es necesario un plano de control que gestione las operaciones que se van a realizar. Este plano de control se encarga, entre otras, de las operaciones de reserva y liberación de recursos, configuración de caminos, encaminamiento y coordinación de las diferentes formas de conmutación y multiplexación que puedan coexistir en la red. El plano de control se puede implementar para alguna capa superior de la red, en la subcapa de control de acceso al medio (MAC - Medium Access Control- layer) para evitar así colisiones entre las tramas de datos a nivel óptico, o en la capa de red, para establecer las rutas y caminos ópticos entre los nodos. La incorporación de técnicas habitualmente desplegadas en redes IP al plano de control de redes ópticas permite la integración de las mismas de tres posibles formas:

- Modelo de pares o "peers": mediante un modelo formado por nodos de la misma categoría, donde los elementos IP y ópticos se interconectan entre sí creando un plano de control único y compartiendo toda la información de conmutación y encaminamiento.
- Modelo en capas u "overlay": un modelo de capas superpuestas donde cada capa gestiona los nodos de su propia tecnología de red de forma totalmente independiente. En este modelo, los nodos IP tienen sus propios algoritmos de encaminamiento distintos de los nodos de conmutación, y la información entre ellos se hace a través de interfaces especiales diseñadas para interconectar ambas capas y que están en proceso de estandarización [1].
- Modelo mixto: en el cual, ambas capas tienen sus propios algoritmos de encaminamiento, pero la información de acceso a los *routers* IP se pasa a través de la red óptica subyacente. Este modelo de conexión interdominio permite por tanto que ambos paradigmas, el óptico y el electrónico de los routers IP interactúen entre sí.

El Sector de Estandarización en Telecomunicaciones de la Unión Internacional de Telecomunicaciones (ITU-T) trabaja en un marco de estándares para lo que se denomina ASTN (*Automatic Switched Transport Network*) para estandarizar el plano de control. En concreto, para redes ópticas define una variante de ASTN denominada ASON (*Automatized Switched Optical Network*). ASON [2] no es una arquitectura implementable propiamente dicha sino un marco para la descripción de diversas funciones de la red: autodescubrimiento de topologías, encaminamiento, reserva

de recursos; y a su vez estandarizar las interfaces entre capas y planos, como por ejemplo en la Interfaz Nodo-Nodo (NNI - *Node-to-Node Interface*) o la Interfaz óptico Usuario-Red (O-UNI - *Optical User-Network Interface*). Esta última es la interfaz que permite el uso de protocolos GMPLS (*Generalized Multiprotocol Label Switching*) para establecer y liberar caminos ópticos a través de la red de conmutación.

Las redes con plano central de control IP que tienen como capa subyacente una red óptica WDM han venido siendo objeto de investigación por muchos autores en la última década [1, 2], ya que se consideran la opción más prometedora para construir redes ópticas conmutadas de gran capacidad a bajo coste.

En los conmutadores de redes ópticas transparentes (AONs - *All Optical Networks*), los multiplexores, OADMs y OXCs, pueden atravesarse de forma completamente óptica, sin acceder a la longitud de onda correspondiente, mientras que la señalización para la configuración de las conexiones entre los bloques se realiza en otro plano.

2.3.5. GMPLS - Generalized Multi-Protocol Label Switching

Las características de MPLS (*Multiprotocol Label Switching*) le hacen un gran candidato para ser implementado en el plano de control de redes multicapa dentro del paradigma de redes ópticas IP/WDM (redes automáticamente conmutadas - ASON), ya que permite separar las operaciones del plano de control de las operaciones del plano de datos funcionando en un canal aparte, permitiendo que coexistan conexiones dedicadas fijas con conexiones conmutadas dinámicamente.

La arquitectura de MPLS fue estandarizada por el IETF (*Internet Engineering Tasking Force*) en 1996 [9]. Esta arquitectura extiende y complementa la señalización y el encaminamiento IP de forma transparente al resto de protocolos implementados. GMPLS permite la introducción de los servicios orientados a conexión en un entorno no orientado a conexión como es IP. MPLS encapsula las tramas correspondientes a un camino o conexión virtual (LSP - *Label Switched Path*) usando etiquetas que dependen de la tecnología implementada en el nivel de enlace en el que se trabaje, incorporando las ventajas de los protocolos de red y evitando capas redundantes. Mediante el intercambio y el mapeo de etiquetas entre dominios y de forma interna al dominio MPLS, los routers MPLS pueden establecer, mantener, señalar y liberar conexiones virtuales entre los nodos de la red, y permite la reserva de recursos y la realización de encaminamiento basado en restricciones.

En el caso de las redes ópticas, al realizar tareas tanto de ingeniería de tráfico como de encaminamiento basado en restricciones, MPLS puede sustituir las tareas de ATM en gestión del tráfico y de SONET/SDH en recuperación rápida de encaminamiento (*Fast Reroute*). Además, la posibilidad de encapsular varios niveles de etiquetado permite a MPLS conmutar de forma transparente diferentes medios dentro de la misma capa de control.

El protocolo MPLS está orientado a dispositivos de conmutación que trabajan a nivel de paquete, pero como se describe en la sección 2.3.3, las redes ópticas permiten muchos grados de granularidad en conmutación. GMPLS (Generalized MPLS) [10] se diferencia de MPLS principalmente en que permite extender el etiquetado para redes de conmutación de circuitos: permite identificar etiquetas con instantes concretos de TDM, etiquetar longitudes de onda para su conmutación (MPLS - *Multi-Protocol Lambda Switching*) o conversión en bloque e incluso fibras completas (conmutación por puerto).

Cuando se calcula un camino a través de la red, los nodos intermedios han de ser informados mediante la señalización correspondiente. Para esta gestión se utilizan protocolos de reserva de recursos (por ejemplo RSVP-TE - *Resource Reservation Protocol with Traffic Engineering extension*), encaminamiento (OSPF - *Open Shortest Path First*) y de distribución de etiquetas (LDP - *Label Distribution Protocol*) dentro de la señalización básica de GMPLS, que a su vez son flexibles en cuanto a su nivel de granularidad y señalización y transparentes y compatibles con otros protocolos de su mismo cometido. Con esta información los nodos intermedios pueden realizar la reserva de recursos y la configuración de sus OADMs y OXCs adecuadamente, dependiendo del nivel de granularidad propuesta para la reserva del camino buscado. GMPLS persigue la posibilidad de interconectar nodos en redes con tecnologías y protocolos totalmente heterogéneos extremo a extremo.

Toda esta heterogeneidad se consigue definiendo puntos de interconexión (LSR - *Label switching routers*) que permitan la gestión de etiquetas en varios niveles de conmutación. Estos LSRs pueden clasificarse en función de su capacidad de conmutación de interfaces (ISC - *Interfaz Switching Capabilities*) en base a su granularidad:

- Interfaz PSC (*Packet Switching Capability*): interfaz capaz de llevar a cabo la conmutación a nivel de paquete, por ejemplo interfaces de encaminamiento IP o MPLS.
- Interfaz L2SC (*Layer 2 Switching Capability*): interfaz capaz de conmutar en función de las cabeceras a nivel de trama, por ejemplo interfaces de la capa de control de acceso al medio

(MAC).

- Interfaz TDMC (*Time Division Multiplexing Capable*): Interfaz capaz de llevar a cabo la conmutación en función de intervalos de tiempo, como pueden ser los sistemas síncronos de conexión y multiplexado DCS y ADM de los nodos intermedios en SONET/SDH.
- Interfaz LSC (*Lambda Switching Capability*): capaces de conmutar a nivel de canales o longitud de onda, como pueden ser los OXCs y OADMs, e incluso grupos de varias lambdas en bloques.
- Interfaz FSC (*Fiber Switching Capability*): permiten la conmutación a nivel físico dependiendo del puerto de conexión de la fibra. Por ejemplo un OXC puede trabajar a nivel de una sola fibra o de múltiples fibras de entrada/salida.

De acuerdo a esta jerarquía, las etiquetas de GMPLS se pueden anidar al igual que en MPLS (label stacking) permitiendo la conexión de túneles GMPLS que atraviesen extremo a extremo islas y/o dominios con diferentes tecnologías y protocolos implementados en las capas correspondientes al plano de datos, mediante una adecuada gestión en los interfaces de control. Dado que en GMPLS las etiquetas identifican aspectos físicos de los recursos disponibles en la red, podrían darse situaciones de bloqueo cuando estos recursos no estuviesen disponibles. Por ello, GMPLS introduce el concepto de conjunto de etiquetas, de forma que cada LSR puede generar conjuntos o rangos de etiquetas en función de su equipamiento y de si cumple o no los requisitos necesarios para establecer el LSP.

Las funciones básicas del plano de control y gestión de la red en GMPLS se pueden resumir en las siguientes:

1. Gestión y reserva de recursos: determina cómo se realiza el descubrimiento (por ejemplo mediante LMP - Link Management Protocol), la actualización y la distribución de los recursos de red, y el mantenimiento de las tablas de estado de enlaces y rutas, ya sea de forma distribuida o centralizada usando protocolos de anuncio o de encaminamiento. En GMPLS los recursos se controlan mediante LMP en un plano aparte de los datos, pudiéndose gestionar de forma transparente grupos de enlaces agregados entre nodos.
2. Cálculo de trayectos (*paths*): determina la manera en la que se realizan las peticiones de conexión, y la gestión del tráfico para optimizar el uso de recursos y evitar colisiones y bloqueos, y gestionando la asignación de los caminos obtenidos.

3. Señalización: determina cómo comunicar el establecimiento, modificación y liberación de las rutas. Ha de tener en cuenta la posible heterogeneidad de tecnologías implementadas en la red para su gestión óptima.
4. Respuesta ante fallos: el plano de control se encarga de gestionar las operaciones tanto de protección (establecimiento de los caminos redundantes, descubrimiento de rutas alternativas, etc.) como de la restauración posterior. Hay que tener en cuenta que normalmente se van a gestionar enlaces de gran capacidad, con lo que un fallo en uno de estos enlaces, afecta de forma crítica a un gran número de servicios y clientes de forma simultánea.

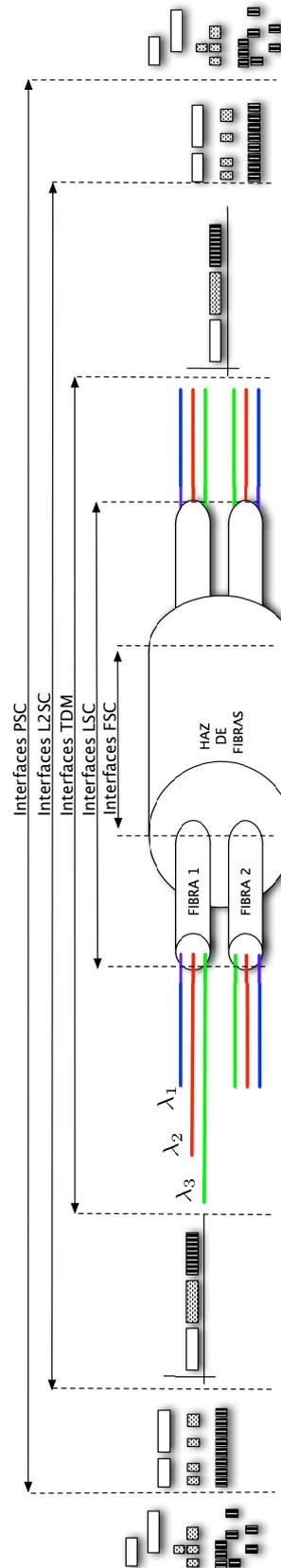


Figura 2.13: Jerarquías de interfaces para su etiquetado en GMPLS

2.4. Protección frente a fallos en redes ópticas

La protección de una red frente a posibles fallos es uno de los aspectos que más preocupa a los proveedores de servicio que operan a través de dicha red. Desde el inicio de las telecomunicaciones, la transmisión de datos era totalmente dependiente de la fiabilidad de los enlaces por los que se transmitía, provocando que el fallo de un enlace o nodo intermedio impidiese alcanzar el destino seleccionado.

Para el proveedor de servicios, el coste asociado a un fallo no sólo depende del coste de reparación del equipo o del enlace, sino también del coste de oportunidad del servicio que deja de prestar durante el tiempo que el enlace no está disponible. Esto supone que desde los inicios de las comunicaciones se implementen enlaces y dispositivos redundantes en el despliegue de la red, que puedan servir de alternativa cuando se produce el indeseado pero inevitable fallo en un elemento. Pero el problema relevante no se soluciona sólo mediante caminos alternativos, ya que surge el inconveniente del tiempo que se tarda en encontrar dichos caminos.

Actualmente, las redes de datos siguen filosofías de conmutación a nivel de paquete de forma predominante dada su flexibilidad y eficiencia en cuanto a número de usuarios y a aprovechamiento de los recursos, y disponen de mecanismos de encaminamiento que encuentran las rutas alternativas de forma autónoma (RIP, OSPF, LSR, AODV, etc.). Pero en entornos de grandes redes de comunicación, el tiempo de convergencia de estos algoritmos puede ser excesivamente alto para el servicio que se desea mantener operativo.

Desde el punto de vista del plano de control de la red, la atención a un fallo pasa por varias fases: su detección, su localización, notificación al sistema de control y su recuperación. Sin embargo, los tiempos de cada una de estas fases cambian en función de la capa donde se produce el error de acuerdo con a la influencia de dicha capa en el resto del sistema. Por ejemplo, en capas altas de la red, puede permitirse fallos del orden de décimas de segundo, mientras que en la capa física es deseable que los fallos no superen el orden de los microsegundos para minimizar la pérdida de datos.

La detección es recomendable que se realice lo más cerca posible del punto de fallo para evitar retardos innecesarios en la gestión del problema. En el caso de redes ópticas, la detección a nivel físico puede realizarse rápidamente en función de la ausencia o no de señal óptica. La localización del punto de fallo puede realizarse fácilmente gracias a la jerarquía de interfaces cuando se haya implementado una capa de control basada en GMPLS al igual que la notificación

a los protocolos de encaminamiento para que encuentren un camino alternativo. Finalmente, los LSRs involucrados gestionarán los recursos activos para lograr la recuperación del servicio usando el camino alternativo, mediante la multiplexación de los recursos de protección seleccionados.

2.4.1. Protección y restauración

La posibilidad de conocer de antemano los recursos disponibles y la existencia de caminos alternativos a los enlaces más susceptibles de fallar, proporciona la posibilidad de establecer líneas de protección que actúan como alternativa cuando se produce el fallo. Esencialmente existen dos posibles esquemas para gestionar la respuesta ante fallos de la red óptica: protección o restauración.

Si los recursos de protección alternativos (rutas, longitudes de onda, etc.) están calculados y preestablecidos de antemano, antes de que se haya producido cualquier fallo, podemos denominar a este mecanismo esquema de *protección*. En caso contrario, cuando un error ocurre, si la nueva ruta o longitud de onda alternativas han de ser descubierta en primer lugar, estaremos hablando de mecanismos de *restauración*.

La recuperación de un fallo usando mecanismos de protección se lleva a cabo mediante la simple conmutación de los recursos reservados para el camino que ha fallado a un nuevo camino alternativo evitando el punto de fallo en la red. En cambio, la recuperación mediante restauración ha de hacer uso de algoritmos de descubrimiento de nuevos recursos y de encaminamiento a través de ellos. De esta forma puede decirse que los esquemas de protección son la técnica proactiva de recuperarse ante un fallo mientras que los esquemas de restauración conforman técnicas reactivas ante los errores.

En general, los mecanismos de restauración consumirán más tiempo para restablecer la conexión, después de que los algoritmos implementados hayan descubierto y reservado los recursos alternativos. Pero los mecanismos de protección, a pesar de ser en esencia más rápidos, suponen la dedicación de recursos reservados sólo para casos de fallo, lo cual resta eficiencia al aprovechamiento de los mismos.

Por todo ello es necesario llegar a un equilibrio entre redundancia de recursos y tiempo de recuperación a la hora de seleccionar y diseñar los mecanismos de control para la recuperación de los servicios frente a errores. Ambos esquemas se pueden aplicar a varios niveles de la jerarquía de interfaces de control.

2.4.2. Clasificación de técnicas de protección y restauración

Podemos dividir los mecanismos de recuperación basados en protección en cinco subtipos, siendo posible implementar cada uno de ellos mediante conjuntos de etiquetas GMPLS. Cada uno de estos subtipos además puede referirse a dos conceptos diferentes: reserva de recursos o asignación de protecciones. Por ejemplo, en el tipo de protección 1:1 podemos referirnos a que los recursos reservados en el camino de respaldo o *backup* están reservados al tráfico de un solo camino principal o *normal* en caso de fallo o simplemente que un camino de respaldo sólo tiene asignado un camino principal a proteger. En las descripciones de los tipos de protección que realizaremos a continuación utilizaremos la terminología relativa a la asignación de protecciones, con la reserva de recursos sería igual [11].

1+1 Un camino de *backup* con recursos disjuntos es pre-configurado de antemano pero sólo protege un camino principal. Además, el tráfico del camino principal es enviado simultáneamente por el camino de respaldo. Es el PML (*Path Merge LSR*) el encargado de seleccionar en cada momento de cuál de los caminos se va a obtener el tráfico.

1:1 Existe un camino de respaldo para proteger un camino principal. El camino de *backup* es preconfigurado de antemano, pero en este caso sólo se envía el tráfico por el camino de respaldo si hay un fallo en el camino principal.

1:n Se preconfigura un camino de respaldo para proteger simultáneamente n caminos principales. Estos caminos principales pueden tener trayectos distintos siempre que tengan en común el PSL (*Path Switching LSR*) y/o el PML.

m:n m caminos de respaldo preestablecidos protegen n caminos principales. Estos caminos principales pueden tener trayectos distintos siempre que tengan en común el PSL y el PML.

Camino dividido Varios caminos de respaldo se utilizan para llevar el tráfico de un camino principal. Se puede configurar el ratio de tráfico a transportar por cada camino de respaldo. Este tipo de protección se utiliza cuando se quieren reservar recursos y no se pueden asegurar con un solo camino. El PML es el encargado de reordenar el tráfico que llega desde los diferentes caminos de respaldo.

De la misma manera, existen en GMPLS diferentes esquemas de recuperación por restauración:

Restauración con restablecimiento. Se calcula un nuevo camino que se establece después de que ocurra el fallo. El camino puede haberse calculado de antemano para acelerar el proceso de restauración, pero en ningún caso se habrá establecido la conmutación ni reservado recursos para ello.

Restauración con reserva de ancho de banda preseñalizada y sin preselección de etiquetas.

Se precalcula un camino alternativo y se señala a los puntos intermedios a través de él la posible reserva de recursos, pero no se le asigna etiqueta GMPLS. De esta manera, varios LSPs que no se prevea que vayan a fallar a la vez pueden compartir los recursos. Cuando se notifica un fallo, se señala el etiquetado de los LSPs y se usan los recursos preseñalizados.

Restauración con reserva de ancho de banda preseñalizada y con preselección de etiquetas.

El camino está precalculado y las etiquetas asignadas, pero los recursos de recuperación los pueden compartir varios LSPs que no se espera que vayan a fallar a la vez.

La selección de los mecanismos a implementar en el plano de control depende de muchos factores como la fiabilidad estadística deseada, el tiempo de recuperación y los recursos disponibles. Es necesario que las diferentes capas GMPLS a niveles de granularidad diferentes estén correctamente sincronizadas para que no se lleven a cabo acciones de recuperación redundantes y conseguir optimizar el tiempo de respuesta con el mínimo desperdicio de recursos.

2.4.3. Otras arquitecturas de protección en redes ópticas:

Además de los mencionados mecanismos de recuperación implementados gracias a GMPLS, existen otros esquemas de protección en las arquitecturas más comúnmente utilizadas en redes ópticas.

2.4.3.1. Protección en SoNet/SDH

Las redes SONET/SDH utilizan técnicas de protección que permiten una alta disponibilidad extremo a extremo para proporcionar calidad de servicio estable para los servicios ofrecidos a través de sus conexiones. Estos esquemas de protección son generalmente de doble camino (1+1 ó 1:1), donde uno de los caminos se preconfigura con antelación como alternativa (camino de respaldo o *backup*) en caso de fallo del otro, que será el que transporte los datos (camino primario).

Si se usa un esquema de protección 1+1 en enlaces unidireccionales donde se envía de forma simultánea el tráfico por ambos caminos, el receptor simplemente selecciona aquel que le llega con mejor calidad. Mientras que en esquemas de conexión bidireccional, el tráfico de ambas direcciones puede tenerse que conmutar por el camino de respaldo, lo cual requiere señalización, mediante un protocolo llamado APS (Automatic Protection-Switching).

Los anillos SONET/SDH están desplegados en gran parte de las infraestructuras de las grandes operadoras, donde muchos nodos están interconectados en una topología en anillo construida realmente a base de enlaces punto a punto entre ellos. Un anillo físico permite rápidamente restaurar un fallo en un enlace, reconfigurando la transmisión de tráfico en el sentido contrario para llegar al nodo destino por el otro posible camino que evita el enlace caído. Estos anillos se denominan SHRs (*Self Healing Rings*) ya que pueden restaurarse automáticamente haciendo uso de la otra mitad de la misma topología.

2.4.3.2. RPR - *Resilient Packet Ring*

RPR es una arquitectura de red en anillo bidireccional y conmutada a nivel de paquete. Está definida en el estándar IEEE 802.17 [12] y su cometido es la de combinar las funcionalidades de transporte por clases fiable (QoS), de alta disponibilidad y multiplexable en tiempo (TDM) de SONET/SDH con la simplicidad, bajo coste y aprovechamiento del ancho de banda de las redes Ethernet.

Un anillo RPR consiste en dos anillos de fibra con sentidos de giro opuestos, donde se permiten hasta 255 nodos. Los anillos RPR clásicos sólo disponen de un canal (esto es, cada fibra transporta sólo una longitud de onda). La disposición de los nodos en anillo permite que el nodo origen seleccione el sentido que mejor le convenga, típicamente el de menor distancia al destino. Los nodos intermedios, al reconocer en la cabecera MAC que no son destinatarios del paquete, lo reenvían hacia el siguiente nodo. Cuando el destino reconoce el paquete como suyo, extrae la información del anillo (esta técnica se denomina *destination stripping*), de forma que el resto de nodos en ese sentido de giro pueden reaprovechar el ancho de banda que quede disponible (reutilización espacial) y transmitir simultáneamente en segmentos del anillo que no pertenecen a la ruta de ese paquete. De esta manera, RPR permite un uso más justo de la red por todos los nodos, de forma opuesta a como ocurriría en Ethernet, y permite el aprovechamiento de todo el ancho de banda, a diferencia de SONET/SDH donde el 50 % de los recursos se dedican a la protección del anillo.

Gracias a su topología bidireccional con el anillo de doble fibra, RPR puede recuperarse de caídas de un enlace simple o de un solo nodo. Cuando sucede el fallo de dos o más elementos simultáneamente el anillo queda en general dividido en dos o más subredes disjuntas sin conexión entre ellas. Cuando se detecta un fallo, ya sea por falta de señal óptica o porque un nodo no recibe los mensajes de actualización de estado de sus vecinos, se difunde a través del anillo un paquete de redescubrimiento de la topología para informar al resto de nodos de la existencia del fallo (notificación).

Los nodos del anillo restantes pueden realizar una o las dos técnicas de protección implementadas en RPR, siendo obligatorio que todos ellos usen la misma simultáneamente para evitar colisiones en la recuperación:

steering la cual es de implementación obligatoria, es la técnica por la cual, el tráfico se recircula desde el nodo origen en el sentido contrario al que se estaba utilizando. Esta técnica permite reaprovechar el ancho de banda al máximo y restaura la conexión usando el camino más corto al destino, con lo que los paquetes sufren el mínimo retardo posible mientras se soluciona el problema. Sin embargo, es más lenta debido al tiempo de convergencia entre la notificación del fallo y el cambio de sentido en el envío de paquetes.

wrapping En esta técnica de restauración, el nodo origen sigue enviando los paquetes en el sentido elegido. Al llegar al nodo más cercano al fallo, éste reenvía todos los paquetes que le llegaron a través de la fibra correspondiente a un sentido del segmento anterior al fallo, por la fibra que se está usando para el sentido contrario, hasta llegar de nuevo al nodo más cercano al fallo pero desde el otro sentido. Este nodo, a su vez, vuelve a puentear ambos sentidos de giro del anillo y reenvía de nuevo el tráfico por la fibra de su sentido original. Al nodo de destino, el tráfico le llega por el sentido de giro original que seleccionó el origen antes del fallo. *Wrapping* permite una restauración más rápida y más eficiente en cuanto a probabilidad de pérdida de paquetes, pero supone un retardo añadido mayor a los paquetes enviados durante el tiempo en que esté caído el enlace.

En general, el tiempo de restauración de RPR, contando desde que se detecta el problema hasta que todos los nodos tienen actualizada la situación de la topología y han hecho las operaciones de cambio de sentido necesarias es del orden de 40-50 ms.

2.4.3.3. Protección en WDM

La mayoría de las redes WDM desplegadas se basan en anillos WDM interconectados entre sí. Cada uno de estos anillos dispone de las capacidades de recuperación típicas de anillos SHR.

En anillos WDM en general se utiliza una sólo fibra con múltiples longitudes de onda. Los nodos intermedios disponen de sistemas de conversión de longitud de onda y de transmisores y receptores sintonizables a varias frecuencias, de forma que se intenta que la red sea completamente óptica. En los OADM de los nodos intermedios se puede extraer la información de una longitud de onda y transmitir datos nuevos a cualquiera de las longitudes de onda arbitrariamente. Cada nodo por tanto puede recibir datos en la longitud de onda extraída y transmitir en la λ que añade. Cuando el número de nodos es igual al número de λ s disponibles, cada nodo dispone de su propia longitud de onda reservada para recepción.

Pero esto no ocurre en general, sino que lo normal es que el número de nodos sea mayor que el número de longitudes de onda disponibles. Es necesario que los nodos dispongan de dispositivos receptores y transmisores sintonizables para poder reutilizar las longitudes de onda. Los errores en este tipo de redes se deben no sólo a caídas de enlaces o nodos sino a bloqueos de las longitudes de onda disponibles por parte de sólo unos pocos nodos.

Además cuando los anillos WDM se interconectan entre sí en puntos intermedios, se desea que las características de restauración de los anillos sigan vigentes para la nueva topología. Para ello, la nueva estructura mallada se puede intentar cubrir mediante uno o varios anillos que engloben toda la malla de conexión WDM.

2.4.3.4. P-ciclos

Grover y Stamelakis propusieron en 1998 [13] un mecanismo de protección intradominio para la recuperación rápida de enlaces. Consiste en la preconfiguración de anillos de recuperación distribuidos en una red mallada de forma que puedan ser usados para recuperar cualquier enlace de la red. Este primer trabajo estaba planteado para la recuperación de los enlaces a nivel físico, con el objetivo de trasladar a una red mallada las propiedades de recuperación de un anillo en cuanto a velocidad de recuperación y simplicidad, aprovechando la eficiencia de las redes malladas en cuanto a uso de la capacidad de respaldo.

Las primeras ideas para el uso de los p-ciclos fueron para utilizarlos en redes tipo WDM y

SONET [13], aunque posteriormente se ampliaran a esquemas de redes MPLS/IP [14]. Originalmente los p-ciclos (*p-cycles - protection cycles*) se diseñaron como una forma rápida de protección frente a caídas de enlaces y posteriormente se estudió su uso en situaciones de caída de nodos, utilizando los "p-ciclos rodeando un nodo". Jugando con ambas posibilidades se extendió su uso a p-ciclos que permiten protegerse tanto de caídas de nodos como de enlaces. Además de extensiones a la protección de nodos también se han propuesto soluciones para la protección de segmentos de camino usando lo que denominan en [15] como "p-ciclos de flujo", puesto que realmente si un camino lleva un flujo, estos p-ciclos servirían para proteger flujos completos, y no enlaces, lo que se traduce en un concepto de protección completamente diferente.

Finalmente, al igual que ocurre en general con las técnicas de protección que suponen cierta redundancia o pre-configuración, ésta ha de hacerse de forma eficiente en cuanto a la relación entre el coste y la eficacia de la protección. Por tanto, determinar cuántos y qué posibles p-ciclos utilizar para proteger una red completa de la manera más eficiente es un tema de optimización relativamente complejo como puede verse en [16], [17] y [18], permitiéndose incluso un aumento de la complejidad de los p-ciclos de forma que puedan atravesar un nodo más de una vez y así proteger zonas de la red que no podrían ser protegidas con un p-ciclo simple.

2.5. Despliegue de redes ópticas

En función de la cobertura de la red y de la estimación de carga en la misma, es necesario recurrir a diferentes técnicas de conmutación, topologías y capas en la red óptica. Es posible clasificar la red dentro de una de las tres categorías mostradas en la figura 1.1. En función de su nivel en esta jerarquía tenemos:

- Red de acceso: que engloba la parte de la red más cercana al usuario final, por lo que soportará el tráfico agregado de varios niveles de usuarios en un radio de cobertura de decenas de kilómetros.
- Red metropolitana: que abarca desde la interconexión de las redes de acceso con la red del proveedor de servicio con el resto de redes metropolitanas que pudiesen coexistir en una zona del tamaño de un gran núcleo urbano (de decenas a cientos de kilómetros). Esta red ha de transportar todo el tráfico agregado que los proveedores de servicio ofrecen a los usuarios finales.

- Red de área extensa (*backbone* o núcleo): que interconecta las redes metropolitanas y los sistemas autónomos a niveles nacionales e incluso internacionales, por lo que su radio puede llegar a ser de miles de kilómetros y por ella se transmite tráfico heterogéneo agregado en múltiples redes metropolitanas.

2.5.1. Redes de área extensa (WAN - Wide Area Networks)

Las tecnologías de red predominantes en las redes troncales de área extensa actuales son IP, MPLS, ATM y SONET/SDH. En muchos casos SONET/SDH es a su vez el mecanismo de transporte subyacente a la red IP o ATM implementada. Las redes SONET/SDH son muy eficientes a la hora de transportar flujos TDM multiplexados de baja velocidad; en cambio IP y ATM están diseñadas para ser muy eficientes para flujos de paquetes conmutados mediante multiplexación estadística. A pesar de ello, debido a la gran demanda de tráfico y a la mejora en la velocidad de las interfaces de IP y ATM, están proliferando soluciones donde la capa de red IP o ATM trabaja directamente mapeada sobre la capa óptica.

SONET (*Synchronous Optical Network*) se desplegó primeramente en Estados Unidos y SDH (*Synchronous Digital Hierarchy*) en Europa y Japón y son actualmente los estándares de multiplexación más habituales para la interconexión de enlaces de gran distancia en el núcleo de la red, sobre todo para grandes operadoras y para enlaces submarinos internacionales, respectivamente. Surge a finales de los años 80 como solución a los problemas derivados de los tradicionales enlaces jerárquicos digitales asíncronos (PDH- *Plesiochronous Digital Hierarchy*) desplegadas en aquel entonces, gracias a las ventajas intrínsecas de ser un sistema síncrono:

- La multiplexación en SONET / SDH es muy sencilla dado que todas las entidades tienen sincronizados sus relojes. Esto evita la introducción de bits para la compensación de posibles errores en la sincronización de las frecuencias de reloj de entidades distintas, lo cual afecta sobre todo al tráfico de velocidades bajas que viaja agregado en las tramas de TDM. Tanto SONET como SDH utilizan velocidades binarias adecuadas a conseguir la mejor adaptación y compatibilidad con el resto de tecnologías coexistentes. Ambos permiten diferentes escalas de granularidad de velocidades.
- Ambos estándares, SONET y SDH introducen mucha información de gestión de red para realizar tareas de monitorización de prestaciones, identificación de conexiones y tipos de tráfico, identificación e información de errores y un canal de control independiente.

- La interoperabilidad entre entidades en SONET /SDH está asegurada ya que se definen interfaces que permiten la compatibilidad de funciones entre equipos de fabricantes distintos, todos ellos sincronizados con un reloj maestro.
- En SONET/SDH se han investigado numerosas técnicas, topologías y arquitecturas que permiten una restauración de servicio más rápida frente a PDH [4].

Para poder añadir o extraer canales físicos de distintas granularidades de velocidad binaria, es necesario un dispositivo que realice la función de multiplexado: el ADM (*Add-Drop Multiplexer*). Un equipo SONET suele configurarse en una de estas tres formas de funcionamiento: como un ADM para topología de anillo, como un ADM en una configuración lineal, o como un multiplexor terminal (TM - *Terminal Multiplexer*).

Las redes SONET desplegadas más básicas consisten en topologías punto a punto formadas por dos equipos TM en sus extremos. En estas topologías es posible insertar dispositivos que funcionen como ADM para conseguir una configuración lineal y permitir la interconexión con otros interfaces de baja velocidad. Usando únicamente dispositivos SONET/SDH configurados como ADM, se pueden interconectar entre sí formando anillos. Además, estos ADM han de tener implementados las rutinas de restauración correspondientes para activarlas en caso de detectar un fallo.

Otro componente importante de una red SONET es el DCS (Digital Crossconnect), que se utiliza en el nodo central para realizar la gestión, transmisión y la agregación de los flujos de tráfico para conmutar y conformar las tramas de datos que se transmitirán a través de la red SONET. La capacidad de tráfico que el DCS puede agregar depende de la granularidad de la interfaz sobre la que trabaja, por lo que será necesario establecer una jerarquía de DCS para poder agregar tráfico de todas las posibles velocidades binarias existentes.

A pesar de que el tráfico agregado en los enlaces de larga distancia es muy homogéneo, para ganar en capacidad, heterogeneidad de tráfico y reconfigurabilidad en los destinos, se comienza a desplegar WDM en las redes de larga distancia, para aumentar su capacidad y reconfigurabilidad, siendo el objetivo implementar redes IP sobre WDM.

2.5.2. Redes ópticas Metropolitanas

Dentro de las redes metropolitanas podemos considerar, por ejemplo, las redes que dan conexión a las sedes de una misma operadora o las redes de interconexión entre ellas. Debido a la extensión que han de cubrir y a la naturaleza del tráfico que en general atraviesa los nodos camino del destino sin necesidad de procesamiento, la configuración más habitual de estas redes *metro*, suele ser en topologías de anillos de decenas de kilómetros que interconectan sedes. El crecimiento de la oferta de ancho de banda y la posibilidad de solucionar el cuello de botella de las redes SONET/SDH, hace que WDM sea la tecnología más adecuada para dar el servicio esperado de la red *metro*.

Otra razón para el uso de WDM es su reconfigurabilidad: en la red metropolitana el tráfico es mucho más variable que en la red troncal, por lo que es necesario reorganizar la asignación de ancho de banda en los enlaces más rápida y eficientemente.

Otro uso muy extendido en la actualidad es dar conexión de gran capacidad a las empresas con sus centros de datos. Estos centros de datos, están habitualmente separados de las sedes centrales por distancias de decenas de kilómetros. Este escenario requiere de gran capacidad por el volumen de datos, tanto de transacciones como de respaldo, que se transmiten. Además, este tipo de conexiones utilizan gran variedad de protocolos y tasas, por lo que la reconfigurabilidad de WDM es muy atractiva para su uso en estas redes, también denominadas SANs (*Storage-Area Networks*) o redes para zonas de almacenamiento. Como los centros de datos suelen estar en las cercanías de las redes metropolitanas de fibra desplegadas es muy sencillo darles acceso a través de ellas.

Las topologías en anillo además permiten a la operadora un alto nivel de protección y una rápida respuesta en tiempo de restauración del servicio.

2.5.3. Red de acceso: FTTx

El despliegue de fibra óptica compite con la red tradicional de par de cobre o coaxial que lleva décadas desplegada en la llamada "última milla" para dar servicio de acceso a los abonados de las operadoras. En países donde la evolución de las redes de datos y el interés de los usuarios ha llegado más tarde, es una opción barata y de gran capacidad el despliegue directo de fibra hasta el usuario final. Dependiendo de la ubicación geográfica y del tipo de usuario final el acceso totalmente óptico a la red se denomina *Fiber-To-The-X* (Fibra-hasta-el-X) donde la X equivale por

ejemplo a un edificio (FTTB - *Fiber To The Building*), un hogar (FTTH - *Fiber To the Home*), un conjunto de ellos conectados mediante par trenzado, coaxial o tecnologías de red inalámbricas (FTTC - *Fiber to the Curb*), etc.

Para reducir los costes de despliegue y mantenimiento en la red de acceso, las arquitecturas FTTx suelen desarrollarse a base de componentes ópticos pasivos, por lo tanto su alcance es limitado en ausencia de amplificadores y otros componentes de regeneración de señal. Por esta razón, estas redes de acceso también se conocen con el nombre de Redes Ópticas Pasivas o PON. Las tecnologías de acceso y enlace que se pueden utilizar en redes PON son muy variadas: encapsulado ATM (APON, BPON), GPON y tramas Ethernet enviadas de forma transparente convertidas al plano óptico (EPON), lo que permite su fácil convergencia con el resto de redes con diferentes tecnologías coexistentes.

En la actualidad, GPON es la tecnología más desplegada a nivel global en redes PON (Fig. 2.14⁵). Pero dada la alta penetración que tiene Ethernet en entornos de operadoras de servicio de banda ancha, y el bajo coste de la fibra, es de esperar que las tecnologías FTTH basadas en EPON se encuentren a la cabeza de los posibles sucesores de los actuales servicios de banda ancha por DSL o GPON. Además, una futura migración a redes PON usando WDM permitiría atender la demanda de las nuevas aplicaciones emergentes de gran ancho de banda, permitiendo además agregar tráfico de otras posibles redes de acceso locales (redes WiMax o LTE por ejemplo) en su conexión con la red metropolitana.

La configuración de estas redes ópticas y la gran capacidad disponible en la fibra permite ofrecer servicios de baja latencia, gran capacidad y calidad de servicio garantizada con bajo coste de despliegue, siendo la solución más aceptada la de llevar fibra directamente hasta el punto de conexión del cliente. Existen varios tipos de topologías y de tecnologías de acceso dentro de FTTH, aunque habitualmente se suele usar este término para denominar al subconjunto formado por las redes ópticas pasivas. Las topologías típicas de FTTH suelen ser:

- conexiones punto a punto directas entre un conjunto de terminales de línea (OLT - *Optical Line Terminal*) y el equipo receptor del cliente mediante una fibra directa dedicada a esta conexión. Esta topología permite al operador ofrecer una calidad de servicio configurada a medida de cada cliente ya que no hay competición por repartirse el medio de transmisión entre ellos. Además, los transceptores Ethernet punto a punto son muy baratos lo que permi-

⁵<http://www.infonetics.com/pr/2011/Carrier-Next-Gen-FTTH-and-PON-Deployment-Strategies-Survey-Highlights.asp>

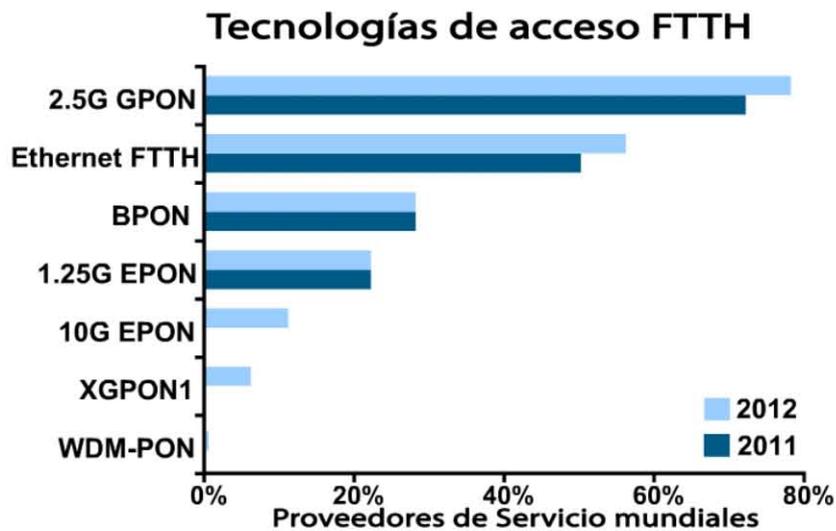


Figura 2.14: Despliegue actual de tecnologías en FTTH y previsiones a lo largo de 2012

te su rápida penetración en el mercado. Sin embargo, como es de esperar, esta configuración sufre de problemas de escalabilidad y de un alto coste de reparación de fallos puntuales por cliente.

- conexiones en árbol, entre el OLT de la operadora y diversas ONU (*Optical Network Units*), a las que se conectan los dispositivos de los clientes finales. De esta manera, se consigue simplificar el despliegue en el lado del operador que es compartido en la mayor parte del trayecto por todos los clientes. Las topologías punto-multipunto pueden conseguirse usando elementos de conmutación óptica para difundir todo el tráfico en el enlace descendente a todos los clientes mediante un separador/combinador (*splitter/combiner*) que separa o replica la información. En el sentido ascendente, este mismo dispositivo ha de agregar el tráfico que proviene de los clientes para entregarlo al OLT. El nodo de difusión puede ser activo (estrella activa) o pasivo (divisor o *splitter*).

En el primer caso, el nodo activo puede amplificar y regenerar la señal y, dependiendo de su estructura interna, realizar operaciones de conmutación con el tráfico para los clientes. Además, puede convertir la señal óptica a otras tecnologías de transmisión de forma que, a partir de este nodo, se pueden tener enlaces de par trenzado o coaxial con acceso DSL (soluciones Fibra-DSL), o enlaces inalámbricos [19](FWA - *Fiber Wireless Access*). En cualquiera de estos casos este nodo activo necesita de instalaciones propias cerca de los clientes con su consecuente coste de mantenimiento, aislamiento y alimentación eléctrica.

Los problemas derivados del consumo eléctrico de los equipos activos y su mantenimien-

to pueden evitarse utilizando elementos completamente pasivos. Las redes ópticas pasivas (PON) tienen las ventajas intrínsecas de su topología punto-multipunto y además añaden su sencillez de despliegue y mantenimiento así como su bajo coste. El elemento separador o divisor, es un divisor de potencia que permite enviar la señal que proviene del OLT a todas las ONUs de la red. Cada ONU recibirá todo el tráfico y seleccionará según el direccionamiento o la tecnología de acceso los paquetes o ráfagas de los que sea destinataria.

2.5.3.1. Redes ópticas pasivas (PONs)

Como se ha introducido en la sección anterior, una red PON es un tipo de red de acceso FTTx en la que todos sus elementos son pasivos. En sentido amplio se compone de los siguientes elementos 2.15:

- El OLT - Optical Line Terminal, en la parte de la red correspondiente al operador que será la conexión de la red de acceso con el siguiente nivel en la jerarquía (una red de acceso de orden superior o la red metropolitana correspondiente).
- Las ONUs - Optical Network Units (típicamente en grupos de potencias de 2 para facilitar su direccionamiento) que reciben todo el tráfico descendente desde el OLT, discriminan el suyo propio y envían en sentido ascendente sólo el tráfico que les corresponde. Este tráfico en el canal de subida se transmite sin coexistir con el de las otras ONUs por la parte del enlace entre la ONU y el *splitter*.
- Una fibra óptica común para la mayor parte de la red: entre el OLT y el *splitter*, y segmentos de fibra de menor tamaño entre el elemento divisor y las ONUs.
- Un divisor de potencia o *splitter* que interconecta la fibra principal hacia el OLT y los segmentos de enlace con cada una de las ONUs y que divide la señal que recibe del OLT en múltiples copias idénticas pero de menor potencia dirigidas a las ONUs. En el sentido ascendente, las señales que reciben los puertos del divisor se transmiten a la fibra común que llega al OLT.

El esquema básico de una red de estas características puede verse en la figura 2.15. En el canal descendente, el tráfico enviado es sincronizado y agregado por la OLT de forma que todas las ONUs tienen que discriminar mediante algún tipo de etiquetado o señalización sus paquetes

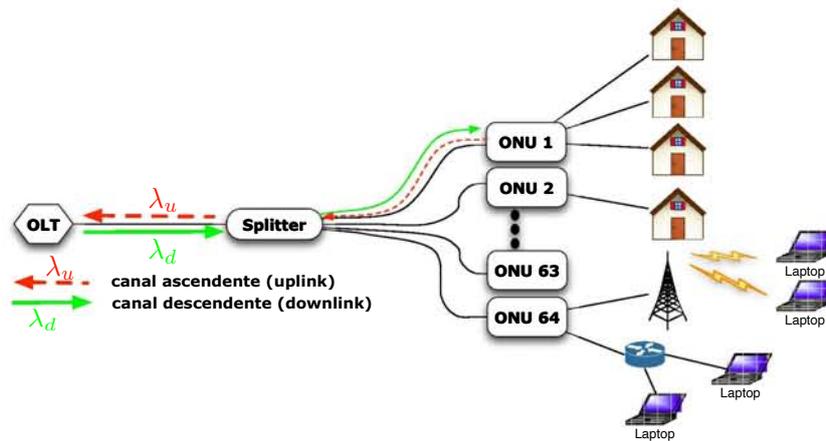


Figura 2.15: Esquema básico de una red PON

para los que son destinatarias. Sin embargo, en el canal ascendente, las ONUs han de competir por el recurso compartido que es la fibra común a partir del *splitter*. En el caso de que dos o más ONUs transmitan a la vez, las señales ópticas que llegan al divisor colisionan a la entrada de la fibra troncal y la señal se corrompe sin posibilidad de ser recibida correctamente en el OLT. Esta colisión no puede ser detectada por las ONUs debido a las características direccionales del divisor de potencia pasivo, esto es, una ONU no puede saber si otra en su misma PON está transmitiendo a la vez ni comunicarse directamente con ella sin pasar por el OLT. Por ello, es necesario utilizar mecanismos de control de acceso al medio (MAC) que arbitren la transmisión de las ONUs y eviten colisiones.

Estudios iniciales [20–22] determinaron que el clásico CSMA que se utiliza en otras redes de acceso no es aceptable debido al alto producto *ancho de banda x retardo* de este tipo de redes, que resultan en tiempos de acceso sin colisión asegurada muy poco eficientes. En lugar de ello, se ha aceptado ampliamente el uso de técnicas de "polling" con reservas de tipo "request-and-grant" como un método más eficiente para arbitrar el acceso al medio compartido del enlace ascendente. A la hora de gestionar la transmisión en el medio compartido, diversas estrategias se han estudiado para PONs:

- Por división en el tiempo (TDMA - *Time Division Multiple Access*)
- Por multiplexación en longitud de onda (WDMA - *Wavelength Division Multiple Access*)
- Por selección de subportadora (SCMA - *SubCarrier Multiple Access*).
- Mediante técnicas de división en código (OCDMA - *Optical Code Division Multiplexing Access*)

Sin embargo, sólo TDM y WDM se han considerado de momento para su despliegue por la madurez de la tecnología que precisan y por su coste, simplicidad y escalabilidad [23–25].

Para la capa de enlace, tanto ATM como Ethernet coexisten en los despliegues de PON, aunque las ventajas que ofrece Ethernet sobre ATM en cuanto a escalabilidad y eficiencia a la hora de ofrecer servicio de enlace a las redes IP, hacen de esta última la solución más aceptada. Las arquitecturas basadas en ATM y con multiplexación TDM pueden englobarse bajo el nombre de APON (*ATM PON*) o su sinónimo más comercial BPON (*Broadband PON*). Las redes PON basadas en Ethernet actualmente se despliegan usando uno de los dos estándares competidores que mejores prestaciones ofrecen: GPON (*Gigabit PON*) y EPON (*Ethernet PON*).

APON/BPON se desarrolló por el ITU-T dentro del estándar ITU-T G.983⁶, el cual define diferentes velocidades binarias múltiplos de la velocidad básica de SONET/SDH de 8kHz, mediante el envío para la sincronización del reloj de 8kHz desde la OLT a las ONUs.

El estándar GPON, también ha sido desarrollado y estandarizado por la ITU-T G.984⁷ y es un estándar que cubre desde el transporte de servicios TDM, datos y video usando tanto tramas ATM como Ethernet a velocidades que alcanzan 1.25 Gbps de subida y 2.5 Gbps de bajada. En general, GPON define ciclos de transmisión fijos para las ONUs cada 125 μs e implementa mecanismos que permitan fragmentar las tramas Ethernet tanto en las ONUs como en el OLT y ser transmitidas en las ranuras de tiempo asignadas por el OLT. Este funcionamiento en ciclos fijos requiere de la sincronización global de toda la PON.

En cambio, EPON transmite sobre la red Ethernet nativa servicios de datos, voz y video a los clientes, lo cual reduce los costes de despliegue al reaprovechar los dispositivos Ethernet y simplifica la arquitectura global de la red. Las tramas Ethernet no pueden fraccionarse en las ONUs o en el OLT pero a cambio, los ciclos de transmisión son variables, lo que permite adaptarse de forma flexible a los distintos tipos de modelos de tráfico de los servicios que ofrece la EPON. La descripción del estándar EPON ha sido realizada por el IEEE inicialmente en el IEEE802.ah [26] en 2004, donde se definen los parámetros para operar a 1 Gbps. El incremento de servicios que requieren de mayores velocidades binarias, como pueden ser los servicios de IPTV y el tráfico agregado de las redes de acceso inalámbrico (*wireless LAN, como puede ser Wi-Fi, o wireless MAN como WiMax*) hacia la red troncal, requieren de la revisión de este estándar. Esta labor se

⁶<http://www.itu.int/rec/T-REC-G/recommendation.asp?lang=es&parent=T-REC-G.983.1>

⁷<http://www.itu.int/rec/T-REC-G/recommendation.asp?lang=es&parent=T-REC-G.984.1>

	BPON	GPON	EPON
Estándar	ITU-T G.983	ITU-T G.984	IEEE 802.3ah (1 Gb/s) IEEE 802.3av (10Gb/s)
Velocidad binaria de bajada	155, 622 Mb/s, 1.2 Gb/s	155, 622 Mb/s, 1.2, 2.5 Gb/s	1.25 Gb/s, 10.3 Gb/s
Velocidad binaria de subida	155, 622 Mb/s	155, 622 Mb/s, 1.2, 2.5 Gb/s	1.25 Gb/s, 1.25 or 10.3 Gb/s
$\lambda_{Downstream}$	1490, 1550	1490	1490, 1550
$\lambda_{Upstream}$	1310	1310	1310
Protocolo (Nivel 2)	ATM	Ethernet over ATM/IP or TDM	Ethernet
Tecnología de Video	RF en 1550 or IP en 1490	RF en 1550 o IP en 1490	Video IP
Max PON Splits	32	64	16
Cobertura	< 20 km	< 60 km	< 20 km

Tabla 2.3: Resumen de características de Redes Ópticas Pasivas (PONs)

está llevando a cabo dentro del estándar IEEE802.3av [27] para desarrollar redes 10G-EPON que operaría a 10Gbps, teniendo en cuenta la posibilidad de compatibilizar la capa MAC de ambos estándares para facilitar la escalabilidad de las arquitecturas.

2.5.3.2. Gestión de acceso en EPON

Como se ha descrito en la sección anterior, la transmisión en las redes EPON tiene la peculiaridad de que las ONUs han de gestionarse de forma que no colisionen en el uso del canal ascendente hacia la OLT, ya que EPON no distribuye ciclos fijos para la transmisión de las ONUs, al contrario que GPON. Este hecho es particularmente importante en Ethernet PONs basadas en TDM donde las diferentes ONUs han de usar ciclos de transmisión disjuntos entre ellas para no colisionar en la fibra troncal entre el divisor de potencia y la OLT. Para solucionar el reparto de ciclos en EPON, dentro del estándar IEEE 802.3ah se define un protocolo de gestión llamado MPCP (*Multi Point Control Protocol*) para el arbitraje del canal ascendente, de forma que la OLT asigne ciclos de transmisión en función de las peticiones de las ONUs. Además de esta función, el protocolo MPCP realiza la función de descubrimiento, establecimiento de conexión con las ONUs y de sincronización cuando éstas entran en servicio (ranging”).

El mecanismo de reserva y asignación de ancho de banda se realiza mediante un sencillo mecanismo de "sondeo"("polling") que puede resumirse en los siguientes pasos:

1. Cada ONU en particular, estima la cantidad de tráfico que va a enviar en la siguiente asignación y solicita el correspondiente tiempo de acceso al canal compartido a la OLT. La estimación se realiza en función de los algoritmos de planificación y de gestión de ancho de banda implementados en la PON, generalmente teniendo en cuenta la ocupación de la cola de entrada a la PON, y en función de la calidad de servicio (QoS - Quality of Service) requerida para el tráfico en el sentido de la red hacia el OLT. En MPCP el paquete de reserva se

denomina mensaje de REPORT y permite señalar el tiempo solicitado de forma individual para las diferentes colas de entrada hasta un máximo de ocho, siendo estas colas utilizadas para priorizar el tráfico o para prestar servicios diferenciados en Internet, en función de los acuerdos de servicio establecidos (SLA - *Service Level Agreement*).

2. El OLT realiza la asignación y gestión del ancho de banda de acuerdo con un algoritmo de gestión dinámica (DBA - *Dynamic Bandwidth Allocation*) usando la información de reserva y QoS que recibe de todas las ONUs. Una vez calculados los instantes de transmisión correspondientes a cada una de ellas, les reenvía a cada ONU su correspondiente mensaje de asignación de ciclo. Este mensaje se denomina GATE. En él vienen especificados uno o varios ciclos de tiempo delimitados por sus instantes inicial y final en los que la ONU puede transmitir con garantía de no colisionar con el resto.

Las estructuras y tamaños de los mensajes pueden consultarse en el estándar IEEE802.3ah [26]. En general, los ciclos de transmisión serán variables y en ellos las ONUs envían una o varias tramas Ethernet. Esta variabilidad depende de la carga de tráfico instantánea y de las preferencias entre las diferentes clases de servicio cuando exista una planificación con QoS. Antes de finalizar su ciclo de transmisión cada ONU tiene la posibilidad de enviar un mensaje REPORT a la OLT para reservar el siguiente intervalo de tiempo en función de su nivel de carga, con la sobrecarga de tráfico que produce el envío periódico de estos mensajes de REPORT (56 bytes) ha de tenerse en cuenta a la hora del cálculo del tamaño del siguiente período a reservar.

Para garantizar la temporización de los ciclos de TDM asignados, la OLT gestiona y difunde un reloj global que las ONUs utilizan para encontrar los instantes de transmisión de sus intervalos asignados dentro de los paquetes de GATE. Además este reloj se usa también para calcular el retardo de ida y vuelta extremo a extremo (RTT - Round Trip Time). Ambas operaciones se realizan cada vez que una ONU se enciende en la red PON, mediante un proceso de registro (denominado '*ranging*'), en el que la OLT envía información del estado del contador de su reloj para que lo actualicen los dispositivos que descubra en el canal descendente.

2.5.3.3. Gestión del ancho de banda en PON

En general, MPCP no define ningún algoritmo particular para realizar la gestión DBA en la OLT, simplemente define el marco de trabajo para implementar los algoritmos DBA deseados en función de las necesidades de los servicios que se ofrezcan en la PON.

A la hora de diseñar un algoritmo de asignación dinámica de ancho de banda hay que tener en cuenta si se va a hacer de forma centralizada o distribuida. En el primer caso, el OLT será la encargada de calcular y señalar los intervalos de transmisión de las ONUs, mientras que en un algoritmo distribuido, las propias ONUs realizan cálculos que señalizan a la OLT mediante sus mensajes de REPORT. Otro parámetro importante es el momento de realizar la planificación, esto es, la OLT puede realizar la asignación de intervalo instantáneamente después de recibir el REPORT de una ONU o puede usar un algoritmo para el cual precise conocer información de todas las ONUs y se lo señala a todas de forma conjunta.

Existen en la literatura muchas referencias al diseño de algoritmos DBA para optimizar la gestión de los intervalos de transmisión TDM para garantizar que los servicios mantienen los requisitos de capacidad reservada, retardo, etc. en la red de acceso. Una primera clasificación de estos algoritmos de asignación dinámica de tiempos puede ser la siguiente:

- Planificación externa entre las ONUs (inter-ONU scheduling), en la que se gestiona de forma dinámica el acceso al canal de subida común entre todas las ONUs.
- Planificación interna a las ONUs (intra-ONU scheduling), en la que la ONU gestiona primero las asignadas a cada una de sus colas en función de los requisitos QoS de su tráfico entrante y divide el tiempo que tiene asignado de forma proporcional a las prioridades y al tráfico de sus colas internas.

2.6. Eficiencia Energética

El aumento exponencial del consumo de energía a nivel mundial desde finales del siglo XX está siendo uno de los temas más importantes a tratar en cualquier aspecto de la sociedad en estas primeras décadas del siglo XXI. En el aspecto industrial y más si cabe en el tecnológico, se abren numerosos frentes de actuación, desde la investigación en nuevas fuentes de energía como las renovables, la optimización del uso de las fuentes de energía tradicionales y el ahorro energético global. El objetivo de la investigación y de los acuerdos adoptados en todos los campos científicos, sociales, políticos y/o económicos a nivel internacional es conseguir un marco de trabajo para alcanzar una sociedad sostenible centrándose en la reducción de la huella de CO_2 . En definitiva es muy importante conseguir reducir el consumo de energía en el día a día tanto de empresas como de ciudadanos individuales, tanto con el uso racional de la misma, como con la

eficiencia de uso de los dispositivos tecnológicos.

Las tecnologías de la información juegan dos papeles indispensables⁸: por un lado está la reducción directa del consumo de energía utilizada en las tecnologías de información y comunicaciones (TICs) ya que ellas mismas son parte del problema del elevado consumo; y por otro lado, está la posibilidad del uso de las TICs para conseguir el ahorro de energía en otros sectores. Para lo concerniente a este trabajo de tesis doctoral, nos centraremos en el primero de estos dos paradigmas.

El consumo de energía en redes de comunicaciones puede clasificarse de la siguiente manera [28]:

- en primer lugar, se consume energía en la infraestructura de comunicaciones (enlaces, conmutadores, encaminadores, repetidores, amplificadores, etc.) para el transporte y regeneración de la señal de datos,
- en segundo lugar, existe un consumo de electricidad asociado a mantener operativos los equipos que dan servicio a la infraestructura de red,
- y finalmente, el porcentaje de consumo más importante se realiza en los equipos finales, ya sea en los servidores y equipos de control de los proveedores de servicio como en los equipos finales de los usuarios.

Uno de los papeles de la investigación en la eficiencia energética se centra en la definición de métricas que permitan cuantificar y modelar el consumo y sus variaciones. Por ejemplo, la estimación de cuánta energía se consume anualmente debido a Internet y la Web, es un valor complejo de obtener: en un estudio de la universidad de Berkeley de Noviembre de 2011 [29] se estima en un 2% aproximadamente el consumo de electricidad debido a Internet. Además, este estudio realiza el cálculo no sólo del gasto en electricidad de los equipos sino el coste en energía para su fabricación y renovación⁹. Otros estudios tratan de obtener modelos estadísticos que permitan analizar dónde están los puntos de máximo consumo dentro de la arquitectura de la red [30] para poder trabajar en optimizar estos.

En [31] se pueden encontrar un buen número de métricas para la cuantificación del consumo de energía en los tipos de infraestructura y tecnologías de red más habituales, y algunos aspectos

⁸<http://www.csee.usf.edu/~christen/energy/madridTalk10.pdf>

⁹Los cálculos de este estudio son públicos y en la fecha de elaboración de este trabajo de tesis doctoral están disponibles en <http://goo.gl/y4juZ>

clave para diseñar estrategias de reducción de consumo en general, como el apagado automático y remoto de los equipos que no se estén usando en cada momento, el balanceo inteligente de carga, etc. La tendencia tradicional era que la mayor parte de la energía se gastaba en la red de acceso [32], pero la reciente evolución en la velocidad de dicho acceso y en la demanda está sobrecargando el consumo en los routers (en cuyo diseño también hay esfuerzos orientados a nuevos diseños para optimizar sus operaciones y reducir su consumo [33]) en la red troncal. Pero a su vez, están apareciendo servicios y aplicaciones de nueva generación como son la televisión digital a través de Internet y los servicios en la nube (*Cloud Computing*) que cambian la balanza de gasto energético de nuevo hacia la red de acceso en el primer caso, y en el segundo caso, hacia los actuales grandes centros de computación y almacenamiento (*Data Centers*) [34] que han de soportar la demanda de miles de usuarios así como el coste de mantener los datos en la red incluso aquellos de acceso menos frecuente. El caso de los centros de datos es además un campo de estudio para otros sectores relacionados con la economía del sector energético ya que su ubicación y su alimentación eléctrica generan enormes cifras de gasto [35].

Otro aumento muy importante de la electricidad que se consume a nivel internacional viene originado por la proliferación de dispositivos de comunicaciones y de otros dispositivos electrónicos embebidos en todo tipo de objetos de consumo habitual, como por ejemplo electrodomésticos, vehículos, etc. La rápida penetración en el mercado de los dispositivos de comunicaciones a nivel de usuario ha disparado el consumo de energía del usuario medio, debido al auge de los dispositivos inalámbricos en mayor medida. Desde hace varios años se llevan a cabo proyectos de investigación para converger hacia un marco de objetivos y estándares comunes que conlleven un mayor ahorro energético en redes inalámbricas [36] y la literatura especializada está llena de publicaciones para optimizar el consumo de batería de los dispositivos móviles y de las actuales redes de sensores así como de la propia infraestructura de la red celular de telefonía móvil [37]. En el caso de los dispositivos móviles, alimentados a base de baterías, la optimización del consumo de las mismas no sólo garantiza el ahorro en cuanto a la energía empleada en la recarga de las baterías o por la fabricación, reciclaje y sustitución de las mismas sino que aumenta el tiempo de vida útil tanto de la batería como del dispositivo, y por consiguiente, las posibilidades de uso de estos dispositivos móviles en nuevas aplicaciones.

Las tareas de optimización en consumo energético son de lo más variadas y están llenas de relaciones de compromiso, desde la mejora de los terminales de transmisión y recepción, que permitirían reducir la potencia utilizada para conseguir las mismas prestaciones a las mismas dis-

tancias de alcance, así como los protocolos de comunicaciones, teniendo en cuenta las desventajas en cuanto a calidad de la señal o pérdidas de datos que pueda conllevar [38]. Esto es, conseguir el mismo servicio con la misma calidad, pero usando menos energía.

Existen numerosos trabajos en la literatura acerca de las posibles estrategias con las que abordar el problema de optimización del uso de la energía en las redes de comunicaciones ya que, como se puede ver, es un problema complejo dado que afecta a un número muy alto de parámetros, imposibles de gestionar simultáneamente. En [39] se puede consultar una extensa revisión de estas estrategias, clasificadas taxonómicamente en función del tipo de algoritmo o modificación a implementar para diversas redes objetivo.

La optimización de los algoritmos implementados en cualquiera de las capas de las torres de protocolos va encaminada a la gestión eficiente del uso que hacen dichos protocolos de los interfaces de comunicación, los procesadores y las memorias, etc. y por tanto del consumo de energía debido a las tareas de procesamiento de los protocolos de comunicación. Un extenso estudio que resume las tendencias actuales en todos los campos de la investigación en eficiencia energética para las redes de nueva generación junto a un listado de algunos proyectos internacionales que lo llevan a cabo puede encontrarse en [40].

2.6.1. Eficiencia energética en redes ópticas

La eficiencia en cuanto al consumo de energía también es un tema candente en la actualidad en la investigación en redes ópticas. No sólo desde el punto de vista de la optimización de los dispositivos de transmisión, recepción y conmutación ópticos, en los cuales se pretende minimizar la energía que consumen en sus operaciones dentro de las redes de comunicaciones, sino en los algoritmos y protocolos de comunicación implementados [41].

El uso eficiente por parte de los protocolos de los recursos dependerá de en qué tipo de red (truncal, metro, acceso) nos encontremos. Por ejemplo, en la red truncal [42], pueden implementarse configuraciones en las cuales los dispositivos físicos pueden estar apagados o consumiendo una potencia mínima durante el tiempo en que el dispositivo no esté transmitiendo información y encenderse cuando se requiera transmitir, en función de la gestión de los protocolos de reserva de recursos de niveles superiores (GMPLS), y seguir cumpliendo los requisitos de calidad de servicio previstos. También se puede recurrir a soluciones de planificación eficiente de las rutas IP y de la gestión de las longitudes de onda. En [41] puede encontrarse una comparativa de algoritmos de

eficiencia energética para la red troncal y sus prestaciones.

Las redes metropolitanas y de acceso ópticas en la actualidad tienden a usar Ethernet como protocolo de comunicación más habitual entre sus dispositivos, por lo que los trabajos en conseguir la eficiencia de Ethernet también aplican a estas infraestructuras de redes ópticas, sobre todo en aquellos enlaces que implementan tecnologías de máxima velocidad para dar servicio a las cada vez más demandadas aplicaciones multimedia de banda ancha [43].

Otra tendencia cada vez más extendida es el intentar desplegar la red sobre tecnologías completamente ópticas usando IP sobre WDM. En [44], los autores realizan un estudio del consumo de los servicios de acceso de banda ancha tradicionales ofrecidos sobre redes IP y la comparativa de su consumo en función del tiempo de infraestructura de red de acceso, y de las consecuencias de ahorro derivadas del cambio a tecnologías ópticas por parte de los proveedores. En [45] se puede ver la evolución de la escalabilidad de estas arquitecturas desde el punto de vista del ahorro energético, para dar cabida al incremento exhaustivo de la demanda de capacidad en la red, mediante el uso de redes ópticas pasivas (PON) como la medida de menor coste para ofrecer servicios de banda ancha, y de enlaces WDM en la red troncal por el consumo relativamente bajo al que se puede llegar en sus enlaces.

2.6.2. Energy Efficient Ethernet (EEE)

En la actualidad, el protocolo de nivel de enlace Ethernet (o su homólogo IEEE 802.3) es el más extendido sobre todo tipo de infraestructuras físicas de comunicaciones (incluidas las redes ópticas), para dar conectividad salto a salto entre los nodos de una LAN y cada vez más de una MAN. Es por ello que un esfuerzo considerable por parte de la comunidad investigadora se centre en la optimización del consumo de energía de dicho protocolo.

Desde 2006, el grupo de trabajo del IEEE (Institute of Electrical and Electronical Engineering) en este campo, el IEEE 802.3 Working Group, comenzó a desarrollar un marco de trabajo para lograr la eficiencia energética de las redes Ethernet. El estándar IEEE 802.3az aprobado en septiembre de 2010 recoge los cambios de funcionamiento a realizar para que los dispositivos Ethernet actuales puedan reducir el consumo de energía [46]. En resumen, este estándar utiliza el recurso de configurar el dispositivo en un modo de bajo consumo mientras no necesite transmitir o recibir datos, para reducir la potencia gastada en el enlace cuando no se envían paquetes. La reducción de energía que se conseguiría al introducir este modo de funcionamiento en todos los

equipos Ethernet a nivel mundial alcanzaría los mil millones de dólares según [47]. En dicho estudio se detallan los pormenores de la adopción del estándar y su influencia en las prestaciones de los equipos.

Sin embargo, el estándar de EEE (Energy Efficiency Ethernet) deja abierta aún la posibilidad de reducir el consumo de energía dado que la mayor parte del consumo se pierde en activar y desactivar el enlace [48]. De esta manera, se puede estudiar la posibilidad de aplicar la minimización del consumo de energía como métrica en la optimización de algoritmos de control de acceso o de encaminamiento para reducir el número de cambios de estado de los enlaces Ethernet.

2.7. Resumen y conclusiones

Como se ha descrito durante este capítulo, existen muy diversos campos dentro de la investigación en comunicaciones ópticas donde se puede aún encontrar formas de optimizar las tecnologías existentes. El presente trabajo de tesis doctoral engloba los trabajos realizados dentro de los temas descritos en las secciones anteriores.

Las dos primeras contribuciones tienen relación con el despliegue de topologías de red óptica en anillo. En primer lugar, el capítulo 3 presenta la contribución realizada dentro del marco de la protección y restauración de redes ópticas visto en la sección 2.4. El trabajo presentado permite obtener un modelo para el diseño de anillos ópticos con métricas útiles para el proveedor de servicio o el operador de red en cuanto al grado y frecuencia de las reparaciones que deberán llevarse a cabo en el anillo, en función de la tasa de fallos en cascada prevista, para poder ofrecer un servicio de operatividad completa en el 99,999 % del tiempo.

La siguiente contribución (capítulo 4) está relacionada con la optimización de la infraestructura interna de los nodos de un anillo óptico. el objetivo de esta contribución es conseguir reducir el bloqueo del tráfico que atraviesa nodos intermedios en su camino al nodo del anillo que lo conecta con la infraestructura de red troncal. Además permite balancear la prioridad de los nodos en cuanto a probabilidad de bloqueo del tráfico circulante y una segunda opción que permite reducir el porcentaje de pérdidas y el retardo por recirculación de los paquetes a costa de realizar la conversión OEO y posterior reenvío de parte del tráfico bloqueado.

El capítulo 5 se presenta el trabajo realizado en el campo de la eficiencia energética. Este capítulo presenta los estudios realizados en relación a la replanificación del encaminamiento de

los flujos de tráfico entrantes en una red con protocolos de reserva de recursos como GMPLS, de forma que se puedan reutilizar los enlaces que ya están consumiendo energía para encaminar dichos paquetes, sin necesidad de encender enlaces nuevos, pero a costa de sufrir una penalización en cuanto a la carga de los enlaces y al retardo que sufre el tráfico al no ser encaminado en general por la ruta más corta al destino. Este capítulo presenta los resultados obtenidos en redes fijas ópticas y de cobre tradicionales.

Finalmente, la contribución del capítulo 6 está relacionada con el estudio de un algoritmo de gestión del ancho de banda en una red óptica pasiva (PON) que resulte en una gestión más eficiente de la calidad de servicio en la misma. Para ello, se propone replanificar los intervalos de acceso de las ONUs basándose en el algoritmo IPACT, pero utilizando un algoritmo inter-ONU para garantizar la minimización del retardo de una clase de tráfico de acceso prioritaria frente al resto de clases de tráfico de todas las ONUs. De la misma manera, en el capítulo 7 se describen las modificaciones necesarias en la asignación de intervalos TDM para obtener un ahorro de energía sin degradar sustancialmente la QoS.

Parte II

Soluciones de eficiencia propuestas para varios aspectos de las redes de conmutación ópticas

Capítulo 3

Propagación de fallos en Anillos Ópticos GMPLS: modelo y análisis de prestaciones

3.1. Introducción

La fiabilidad y la capacidad de recuperación en caso de fallo se han convertido en una de las preocupaciones más importantes para los Proveedores de Servicio de Internet y Operadores de Red. En efecto, los operadores de red buscan formas de conseguir asegurar a sus clientes un nivel de fallo por debajo de 1:10000, lo que significa que su objetivo es conseguir tener su red operativa al menos en el 99,999 % del tiempo. Hay muchos métodos y técnicas para tratar los fallos de forma que la continuidad del servicio no se vea comprometida en un principio, o se vea rápidamente restablecida en caso de que ocurra el fallo [49, 50].

Las redes de datos actuales integran varias tecnologías de transporte de forma que el sistema completo se conforma mediante un modelo de capas apiladas. En este modelo, las capas superiores operan sobre topologías virtuales que se construyen sucesivamente sobre la estructura que conforman las capas inferiores [51]. Dentro de los protocolos que se utilizan habitualmente, GMPLS - Generalized Multi-Protocol Label Switching (GMPLS), descrito en 2.3.5, está ganando una gran aceptación como el protocolo más seleccionado para gestionar este tipo de redes tan heterogéneas [10].

GMPLS permite la interoperabilidad y convergencia de diferentes tecnologías de transporte utilizando un plano de control unificado, y aspira a facilitar retos tales como la provisión de servi-

cios, ingeniería de tráfico y recuperación de fallos [52]. Una arquitectura en capas puede mejorar la capacidad de recuperación de la red ya que proporciona flexibilidad a la gestión de los fallos y a su resolución [53,54]. A pesar de ello, en estas arquitecturas aparece un efecto indeseado como es la *propagación de fallos en cascada*, debido a que errores en las capas inferiores pueden afectar al correcto funcionamiento de los protocolos desplegados en las capas superiores. Otro efecto intrínseco a la estructura en capas es que un fallo en la capa inferior, puede manifestarse como un conjunto de fallos concurrentes de las capas superiores.

La mayor parte de los estudios en recuperación de una red asumen que el número de fallos cuya reparación está pendiente en un cierto instante es pequeño (por ejemplo, no más de uno o dos de forma simultánea), y que suceden de forma independiente unos de otros. Existe un amplio abanico de publicaciones que estudian exhaustivamente el tratamiento de este tipo de fallos, y los aspectos relativos a la operación y diseño de la red están cubiertos en ellos. Sin embargo, no abundan investigaciones tan extensas para escenarios arbitrarios de gran escala y/o para fallos múltiples. Este tipo de fallos son habitualmente causados por desastres naturales o ataques intencionados, y por lo tanto es más raro que sucedan que por ejemplo la rotura de un cable o fallos de funcionamiento de módulos hardware, pero sus consecuencias son a menudo mucho más dañinas. En esta contribución se pretende conseguir una guía de medida del impacto de este tipo específico en redes basadas en GMPLS teniendo en cuenta un tipo específico de fallos múltiples: aquellos fallos que pueden propagarse de un nodo a otro a través del plano de control y que pueden afectar a una parte considerable de la topología de la red.

Para ello se aprovecha una de las características clave de GMPLS: la separación del plano de control frente al plano de datos (véase sección 2.3.5), hasta el punto de que pueden desplegarse usando redes diferentes. Debido a esta disgregación, se puede separar la atención a los fallos ya que estos pueden producirse en cualquiera de los dos planos, incluso simultáneamente [55]. Un fallo en el plano de control conlleva la pérdida de funcionalidades de control y conmutación (por ejemplo, un nodo de conmutación que no puede reconfigurarse), mientras que un fallo en el plano de datos afecta a los servicios de reenvío de paquetes [56].

Además de los beneficios que tienen las arquitecturas multicapa por esta separación de planos, también aparecen nuevos requisitos como es la necesidad de protección frente a fallos del plano de datos independientemente [55,57]. Sin embargo, al menos en el alcance de la literatura encontrada en el momento de la realización de esta contribución, no hay estudios publicados para esta clase de escenarios, en el cual un fallo en un nodo en el plano de control puede terminar por provocar un

fallo en otro nodo o incluso el plano de datos, en otras palabras, situaciones donde los fallos de GMPLS se propaguen.

El objetivo concreto de este estudio es por lo tanto caracterizar el comportamiento transitorio y los posibles estados en los que se encuentre una de las arquitecturas ópticas más habituales en las redes de operador: un anillo óptico basado en GMPLS sujeto a un escenario de fallo múltiple. En esta contribución se considera que la propagación de fallos acontece de forma simultánea en dos sentidos: una propagación a nivel horizontal en el plano de control de nodo a nodo, y una propagación vertical desde el plano de control al plano de datos. Para poder calcular la fiabilidad de esta clase de topologías en anillos, se propone el diseño de un modelo mediante el uso de una cadena continua de Markov (CTMC - Continuous Time Markov Chain) .

3.2. Trabajos anteriores y planteamiento del problema

3.2.1. Propagación de fallos en redes multicapa

Los efectos negativos de la propagación de errores pueden ser evitados o limitados si las capas inferiores de la red se mantienen buscando de forma automática rutas alternativas para usar como respaldo en caso de producirse un fallo, gracias a que cuentan con sus propios mecanismos para ello. De esta forma, los procedimientos de autorrecuperación pueden activarse autónomamente en caso de error, de forma transparente a la capa superior [58].

Un punto de vista diferente es diseñar el nivel superior de la topología de red teniendo en cuenta las capacidades y restricciones de las capas inferiores. Por ejemplo, cuando se diseñan redes IP-sobre-WDM, las demandas de tráfico de la capa IP han de ser encaminadas a través de la infraestructura WDM. De esta forma, si ocurre un fallo en un nodo o en una fibra, la topología IP estará protegida y puede seguir conectada si hay capacidad libre suficiente para poder recuperar la conexión a nivel IP. Está demostrado que mantener estos respaldos se trata de un problema NP-completo [58].

Diversos algoritmos heurísticos se han propuesto en la literatura para encontrar las topologías virtuales de encaminamiento y protección, o para modificar las topologías hasta que la asignación sea la apropiada, como por ejemplo en [59–61], o en otros estudios más amplios [51, 62–66].

La mayoría de los trabajos de investigación en protección de redes óptica, incluidos aquellos relativos a redes multicapa, asumen que los fallos ocurren de forma independiente uno del otro.

Por tanto, la ocurrencia de fallos del tipo de cortes de fibras y fallos de funcionamiento de nodos, ocurren de forma aislada y son eventos que no tienen relación entre sí. Más aún, dado que se considera que los fallos múltiples son posibles aunque poco frecuentes [67], la atención se centra en fallos aislados, y en fallos de enlaces individuales en concreto, existiendo muy pocos estudios acerca de redes que puedan reponerse de un fallo de enlace doble.

Dentro de estos últimos, una forma específica de fallo múltiple que atrae mucha atención es la resultante de daños en estructuras físicas, como por ejemplo los conductos que son compartidos por un grupo de fibras no relacionadas entre sí. El concepto de "Grupos de Enlaces de Riesgo Compartido" ("*Shared-Risk Link Groups*" [68,69] y su generalización a "Grupos de Recursos Riesgo Compartido" ("*Shared Risk Resource Groups*" (SRRG) [70]), tienen en cuenta esta situación y se utilizan habitualmente en el diseño de la protección de la red. Sin embargo, en la contribución que forma parte de este trabajo de tesis doctoral, lo que se aborda es el caso de que los fallos se propaguen a través de los nodos en la red, lo que la diferencia de los casos de estudio encontrados que analizan la fiabilidad de la red en cuanto a sólo fallos aislado e incorrelados.

En cualquier caso, en el mundo real podemos encontrarnos con muchos más tipos de fallos, incluyendo aquellos en los cuales la disfunción puede propagarse a través de la red o cubrir un área geográfica amplia, esto es, afectando a varios elementos de red no relacionados entre sí simultáneamente. Las causas originarias de este tipo de fallos a gran escala suelen ser desastres naturales [71, 72], pero también pueden ser virus y gusanos informáticos, así como ataques intencionados [73]. Aunque la literatura relacionada con fallos a gran escala es extensa en el contexto del estudio de redes complejas (por ejemplo [74–76] y los trabajos referenciados en ellos), hay menos publicaciones acerca del modelado y análisis de este tipo de fallos en redes de ordenadores. Algunos fallos ocasionados por catástrofes se analizan en [77] y [78] pero no son aplicables directamente a redes de transporte, así como su impacto en la capa IP de Internet.

En [79] y [80] pueden encontrarse estudios acerca de fallos de red geográficamente correlados y que afectan a localizaciones específicas. En estos artículos se proponen modelos para evaluar la fiabilidad en escenarios de fallo dados de forma que se determinen las áreas más vulnerables de la red física. Además, se centran en las propiedades estructurales de las topologías dadas y su habilidad para responder ante errores localizados causados por errores no propagables.

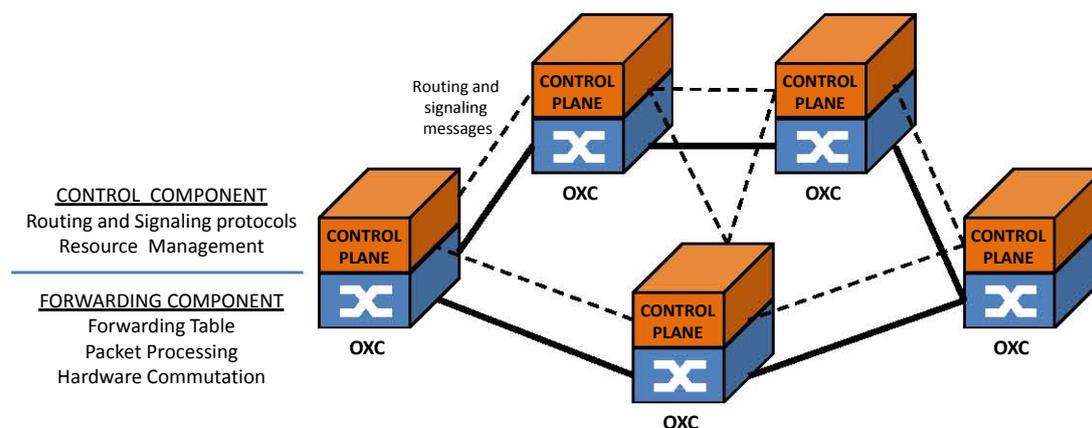


Figura 3.1: Los planos de control y datos en la arquitectura de GMPLS

3.2.2. Fallos en redes basadas en GMPLS

Habitualmente, cuando se tienen en consideración las redes GMPLS (redes ópticas), es posible distinguir dos partes diferentes en todos los nodos (Fig. 3.1). En primer lugar, hay un componente de transmisión donde el hardware diseñado específicamente para ello se encarga de encaminar, tan rápido como sea posible, los paquetes a su correspondiente puerto de salida de acuerdo a la tabla de encaminamiento. Por encima de este componente, existe otro hardware de control genérico que, mediante un sistema operativo de red específico, ejecuta los protocolos de encaminamiento y señalización, y configura la tabla de reenvío (cuando las conexiones son establecidas o liberadas). Aunque ambos componentes suelen estar ubicados en el mismo dispositivo, entre ellos existe algún grado de aislamiento. Incluso podrían estar situados en diferentes dispositivos y podrían enviar los mensajes de control a través de redes distintas [55].

En este tipo de escenarios sería posible que un ataque o fallo afectase sólo al plano de control o sólo al componente de reenvío, durante un corto espacio de tiempo. Sería incluso posible que, debido a un virus informático, un ataque premeditado o una configuración de software errónea, el fallo afectara sólo a un mecanismo de control independiente (por ejemplo, al protocolo de señalización o al de encaminamiento). En caso de que el protocolo de señalización falle y el módulo de encaminamiento siga funcionando, las conexiones ya no pueden establecerse o liberarse a través de ese nodo. En ese caso, no se puede usar el protocolo de encaminamiento para anunciar a los nodos vecinos que no hay suficiente capacidad disponible para que no intenten establecer nuevas conexiones a través del nodo con este fallo parcial. Por otro lado, si el módulo de señalización sigue operativo pero es el de encaminamiento el que falla, los cambios de estado del nodo local

no pueden ser señalizados a los vecinos, los cuales estarán trabajando con información obsoleta. Sin embargo, el nodo con el fallo aún podría procesar peticiones de conexión nuevas o liberar conexiones existentes.

Para este análisis, se asume en adelante que un fallo en el plano de control provoca fallos en ambos protocolos: señalización y encaminamiento. En este caso no es posible establecer ni liberar conexiones, y los nodos vecinos no recibirán señalización de actualización de la información del nodo mientras ocurra un fallo en el plano de control. Sin embargo, es posible que, por algún tiempo, el componente de reenvío continúe funcionando correctamente con la tabla de encaminamiento configurada apropiadamente para las conexiones establecidas previamente al fallo. También podrá darse el caso de que, un cierto tiempo después del fallo en el plano de control, el plano de datos termine por dejar de funcionar correctamente, causando un fallo completo del nodo y la interrupción de las conexiones establecidas a través de dicho nodo.

Es de la mayor importancia establecer algún mecanismo que sea capaz de restaurar la funcionalidad del componente de control con fallo lo más pronto posible y sincronizar de nuevo los componentes de control y re-envío. Esto puede conseguirse mediante nodos que implementan mecanismos de resincronización como "*Non-Stop Forwarding*" o "*Graceful Restart*" [81, 82]. Aun así, esto no es fácil de llevar a cabo y requiere de un tiempo de restauración debido a una primera fase de reinstalación y carga del componente de control y los procedimientos necesarios y los mensajes del protocolo para aquella resincronización [57].

El paradigma de la fiabilidad en el plano de control en redes GMPLS atrae la atención de la comunidad investigadora. En [83], los parámetros de señalización relativos al tiempo se estudian para optimizar la detección de fallos y la sobrecarga de control en anillos ópticos. El impacto de la pérdida de mensajes RSVP-TE y OSPF-TE (mensajes de señalización y encaminamiento) se estudian en [84]. Asimismo, el modelo analítico presentado en [85] puede usarse para cuantificar el número de enlaces requeridos en la topología del plano de control para que la probabilidad de perder el establecimiento o interrupción de una conexión se mantenga por debajo de determinado valor.

A diferencia de los trabajos referenciados, en esta contribución se considera un escenario en el cual los fallos pueden propagarse a través del plano de control GMPLS de un nodo a sus vecinos, y además considera que, una vez que la funcionalidad del plano de control se ve comprometida, es posible que el plano de datos se vea afectado, creándose un escenario de propagación de fallos bidireccional.

Nuestra contribución se centra en el estudio de topologías de red en anillo por dos razones:

1. Están ampliamente desplegados como parte de diversas tecnologías de transporte y se pueden encontrar de forma común en redes metropolitanas [62, 86].
2. Su estructura es muy adecuada para el estudio de la evolución de los fallos y sus efectos sobre la fiabilidad a través de Cadenas Continuas de Markov.

3.2.3. Propagación de fallos basados en modelos epidémicos

Teniendo en cuenta que los fallos en los que se centra este trabajo son aquellos que pueden propagarse, los modelos de propagación epidémicos que se aplican a otros campos de la ciencia pueden usarse para caracterizar la dinámica de la expansión de los errores por la red. El término "*red epidémica*" se utiliza para describir y estudiar cómo evoluciona una epidemia en un conjunto de individuos durante una cierta cantidad de tiempo, tanto en redes de contacto biológico como en redes de ordenadores (véase como ejemplo [87, 88] y sus referencias).

El auge y mitigación de una epidemia se puede caracterizar de forma probabilística, dependiendo de la tasa de infección y del grado de conexión de los nodos o individuos [89]. Las investigaciones llevadas a cabo en este campo estudian diferentes aspectos, incluyendo cómo evoluciona la epidemia a lo largo del tiempo o cómo inmunizar a una parte de la población para minimizar y controlar la propagación de la infección y de sus efectos. Ejemplos de aplicaciones en red donde se pueden aplicar estos modelos epidémicos pueden ser redes de alimentación eléctrica, redes sociales, redes neuronales o redes de ordenadores.

Se han propuesto un gran número de modelos epidémicos para caracterizar la propagación de virus en sistemas complejos, principalmente en el campo biológico. Una buena revisión de este campo puede encontrarse en el artículo [90]. Para esta contribución asumiremos que los fallos se propagan usando el modelo SID (Susceptible-Infection-Disabled) publicado recientemente en [91], y descrito en 3.3. Este estudio presenta un análisis basado en simulación de la robustez de topologías malladas bajo el riesgo de fallos propagables según el modelo SID.

A continuación se presenta el estudio hecho con el objetivo de complementar el trabajo de [91] proporcionando un estudio del diseño de un modelo analítico de SID sobre topologías en anillo, y se proporcionan unas reglas de diseño en base a las tasas de reparación con el objetivo de conseguir la fiabilidad y disponibilidad aseguradas para el servicio de la red.

3.3. El modelo SID: *Susceptible-Infected-Disabled*

Esta sección describe los estados del modelo asociados a un nodo y cual es el efecto de estar en cada estado desde el punto de vista funcional. Adicionalmente, se describen las suposiciones hechas acerca de la generación y propagación espontánea de un fallo, necesarias para poder desarrollar finalmente el modelo de propagación para el anillo mediante CTMC.

Los posibles estados en que puede encontrarse un nodo en el modelo SID son:

- El estado S (*Susceptible*), significa "estado susceptible de ser infectado". En este estado, ambos planos, el de datos y el de control funcionan correctamente, por lo que el nodo es susceptible de tener un fallo (o de "ser infectado") si al menos uno de sus vecinos ya ha sido infectado antes. De forma adicional, el modelo contempla la posibilidad de que un nodo en este estado pueda fallar de forma espontánea, lo que significa que dicho nodo será el causante de la infección.
- Con el estado I (*Infected*) se denominada al "estado de *infección* o fallo en el plano de control". En este caso, el plano de control GMPLS falla, pero el plano de reenvío de datos sigue funcionando. El nodo no puede participar en el establecimiento o eliminación de nuevos LSPs ni tampoco será capaz de modificar la configuración de los sus LPSs propios. Sin embargo, el tráfico de los LSPs que ya estuviesen establecidos se sigue enviando. Un nodo que se encuentra en este estado puede propagar el error a sus vecinos.
- Finalmente, existe el estado D (*Disabled*) o "estado de error o disfunción total", en el cual tanto el plano de control como el de datos fallan. La propagación de este tipo de fallo a nodos adyacentes no puede suceder debido a que la conectividad con dicho nodo ha quedado interrumpida.

La Fig. 3.2 muestra el diagrama de transición entre estados para el modelo SID, donde los valores de las flechas significan la tasa de transición entre los distintos estados (el número de eventos de transición ocurridos por unidad de tiempo). En definitiva, Fig. 3.2 denota las siguientes funcionalidades del modelo:

- Un nodo que se encuentra operativo dentro de la normalidad (se encuentra en estado S), deja de funcionar con una tasa β (en el caso de que exista al menos un nodo vecino infectado).

- Un nodo que presenta fallos en el plano de control (estado I, *infectado*) se repara y vuelve a estar de nuevo operativo (estado S) o puede desembocar en un fallo total del nodo (falla también el plano de datos y pasa al estado D, *deshabilitado*). Lo primero ocurre con una tasa δ , que es la tasa con la que el operador de red corrige los problemas en el plano de control de forma manual o automática, mientras que el segundo caso, ocurre a una tasa de c . El operador de red puede también reparar un nodo que está deshabilitado con una frecuencia t .
- Finalmente, β_F indica la frecuencia de aparición de fallos espontánea en un nodo de la red cuyos vecinos están funcionando correctamente.

El valor de la tasa β_F es en general mucho más pequeño que el de β , por lo tanto puede despreciarse para el cálculo de la tasa de propagación de fallos final por simplicidad. Esta relación permite también tener en cuenta la siguiente hipótesis de funcionamiento: cuando un nodo sufre un fallo espontáneo del plano de control, al ser la tasa de fallos espontánea mucho menor que la de propagación, no se permite en el modelo exista otro nodo aislado que sufra fallo; a partir del primer error, el resto de fallos se producen sólo por propagación entre vecinos. La tabla 3.1 resume los parámetros comentados del modelo SID.

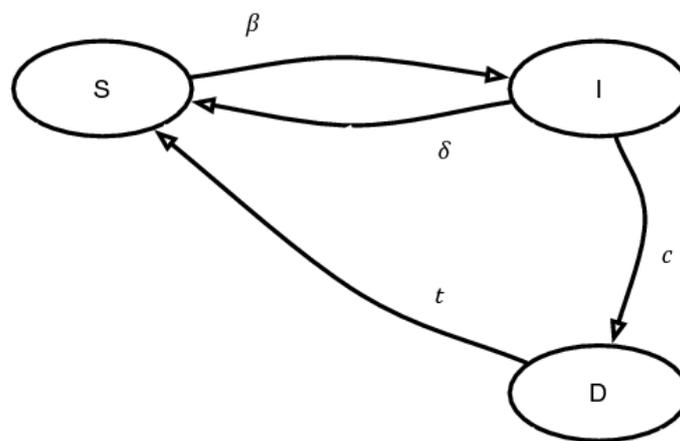


Figura 3.2: Diagrama de transición de estados para el modelo SID

Es importante remarcar que, en este escenario, la propagación epidémica de los fallos ocurre sólo entre entidades pertenecientes a los planos de control. Esto es, la propagación entre planos (el fallo del plano de datos una vez se ha dado el fallo en el de control) no es epidémica, sino que es la consecuencia de asumir que un conjunto de nodos en el estado I no pueden ser reparados y puestos de nuevo en el estado de operación normal. En ese momento, los nodos dejan de ser operativos

Parámetro	Descripción
β_F	Tasa de fallo espontáneo
β	Tasa de propagación de fallos
δ	Tasa de reparación del Plano de Control
c	Tasa de Fallo del Plano de Datos
t	Tasa de reparación de nodos Deshabilitados

Tabla 3.1: Notación y parámetros correspondientes al modelo SID

completamente (fallo del "forwarding" o reenvío del plano de datos) y entran en el estado D, en el que también el plano de datos funciona incorrectamente.

3.4. Análisis y diseño del modelo

3.4.1. Modelo de Cadenas de Markov en Tiempo Continuo (CTMC)

Como se ha comentado anteriormente, para el diseño del modelo de predicción de este trabajo, se considera que los fallos de cada nodo ocurren de forma independiente entre ellos y constan de la propiedad de no tener memoria, lo que significa que los tiempos entre fallos están distribuidos de forma aleatoria siguiendo una distribución exponencial.

Se asume este comportamiento para todos los posibles eventos que puedan darse en este escenario: fallos en el plano de control ("*propagación de la infección*"), fallos en el plano de datos ("*nodo deshabilitado*"), reparaciones del plano de control y reparaciones completas (cuando un nodo apagado, estropeado o que falla su plano de datos es vuelto a poner en funcionamiento). Con estas características, se hace apropiado el uso de cadenas de Markov de tiempo continuo (CTMC - *Continuous Time Markov Chain*) para el diseño del modelo. Estas cadenas de Markov se caracterizan por el denominado *diagrama de tasas de transición de estados (state transition-rate diagram)* (Fig. 3.2), el cual es un grafo que muestra los posibles estados en los que puede encontrarse el sistema, junto con arcos dirigidos que conectan los estados y que representan las tasas de transición (en eventos de fallo o reparación por unidad de tiempo) entre estados. Esta sección describe el uso de una CTMC cuyas estados representan las posibles situaciones de fallo que podrían ocurrir en una red GMPLS con topología en anillo. Para la descripción del modelo se utiliza un anillo GMPLS de ejemplo formado por 8 nodos. Las tasas de transición del modelo están calculadas de acuerdo a las tasas de propagación de errores del modelo SID de la Fig. 3.2.

El diagrama de tasas de transición proporciona la denominada *matriz generadora infinitesimal*

(infinitesimal generation matrix) Q , la cual caracteriza el comportamiento transitorio de la cadena de Markov. Para el modelo diseñado, la matriz Q se usará para calcular:

1. Las probabilidades en estado estacionario (*steady-state probabilities*) del sistema: el porcentaje de tiempo que el anillo está en una configuración dada a largo plazo.
2. Los tiempos de primera ocurrencia (*first-passage times*) de cada estado: la cantidad de tiempo que en media se tarda en alcanzar un estado desde otro cualquiera.

En resumen, usando la matriz Q , las soluciones para el estado estacionario de la cadena de Markov, requieren resolver el valor del parámetro p_i en el siguiente sistema de ecuaciones:

$$p_i Q_i = \sum_{j \neq i} Q_{ij} p_j, \quad i \in C \quad (3.1)$$

$$\text{donde} \quad \sum_{j \in C} p_j = 1 \quad (3.2)$$

$$\text{y} \quad Q_i = - \sum_{j \neq i} Q_{ij} \quad (3.3)$$

donde p_i es la probabilidad estacionaria del estado i , C es el espacio de estados y Q_{ij} es la frecuencia de transición entre desde el estado i al estado j , según se especifique en la matriz Q . Los valores de p_i denotan la cantidad de tiempo que, a largo plazo, la cadena de Markov se va a encontrar en cada estado.

La matriz Q permite además calcular los tiempos de transición esperados entre una pareja de nodos cualesquiera de la CTMC: tiempo de primera aparición u ocurrencia. El tiempo de primera ocurrencia desde el estado i al estado k , que denotaremos a partir de ahora como m_{ik} , es el tiempo medio en alcanzar el estado k por primera vez suponiendo que el proceso empezó en el estado i , y se calcula mediante el siguiente conjunto de ecuaciones:

$$m_{ik} = \frac{1}{Q_i} + \sum_{j \neq k} \frac{Q_{ij}}{Q_i} m_{jk}, \quad i, j, k \in C \quad (3.4)$$

donde $Q_i = \sum_{j \neq i} Q_{ij}$.

3.4.2. Caso de estudio: un anillo GMPLS de 8 nodos

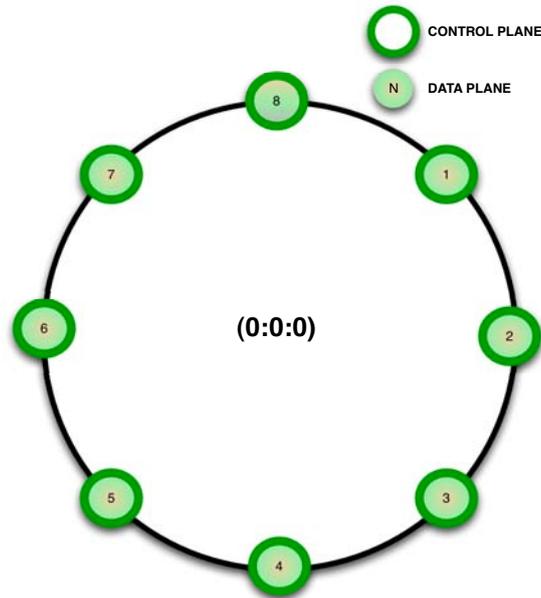


Figura 3.3: Anillo basado en GMPLS de 8 nodos de ejemplo

La Fig. 3.3 muestra un anillo GMPLS de 8 nodos que se usará para la descripción del diseño de un modelo con cadenas de Markov en tiempo continuo. El modelo CTMC utiliza las siguientes suposiciones:

1. Los nodos que ya tienen fallos en su plano de control ("*nodo infectado*"), sólo pueden propagar los errores a sus nodos vecinos. Básicamente, un nodo sólo puede sufrir errores en el plano de control si tiene al menos un vecino que ya los esté sufriendo. Excepto en el caso del primer nodo erróneo del anillo, que fallará de forma espontánea. La tasa espontánea de aparición del primer fallo en el plano de control de GMPLS es β_F .
2. Sólo los nodos con problemas en su plano de control pueden fallar definitivamente y quedar deshabilitados. Los nodos deshabilitados no pueden propagar los fallos a sus vecinos a ningún nivel, ni en el plano de control ni en el plano de datos.
3. Ambos tipos de nodos, aquellos con fallos sólo en el plano de control o aquellos deshabilitados completamente, pueden ser reparados por el operador de red, pero sólo si se encuentran adyacentes a un nodo que funciona correctamente. En otras palabras, los eventos de reparación tanto de nodos "*infectados*" o "*deshabilitados*" sólo pueden ocurrir en los extremos de la cadena de nodos con errores.

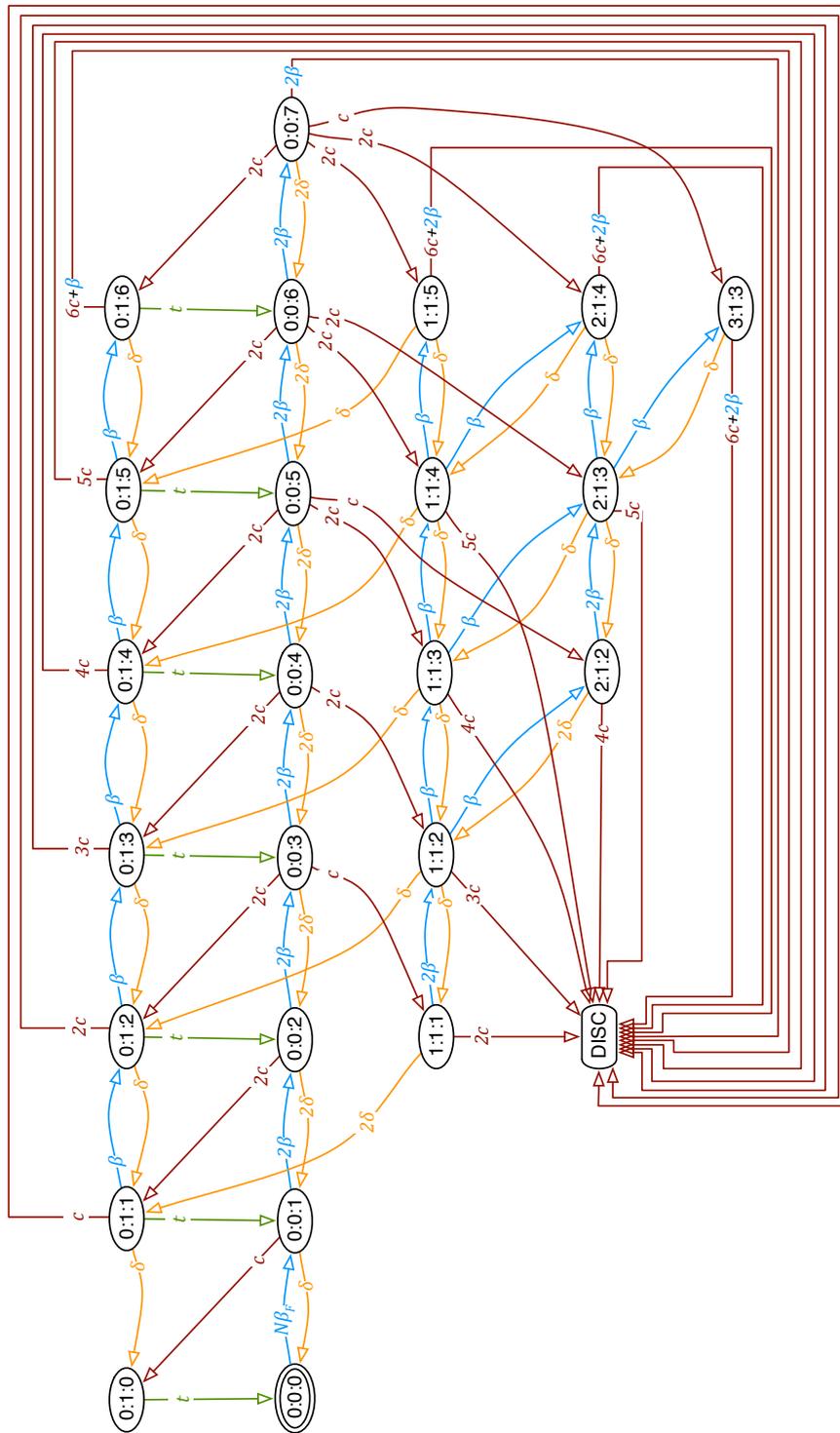


Figura 3.4: Diagrama de tasas de transición y estados para el modelo SID en un anillo GMPLS de 8 nodos

Para ilustrar esta descripción, la Fig. 3.4 muestra el diagrama completo de estados y transiciones del modelo para la topología en anillo de la Fig. 3.3. Cada estado está etiquetado mediante la siguiente tupla $(N_{II} : N_D : N_{Ir})$, donde N_D indica el número de nodos deshabilitados (nodos en el estado D), y N_{II} y N_{Ir} en los extremos indican el número de nodos con fallos en el plano de control (nodos en el estado I) a ambos lados del nodo deshabilitado respectivamente. Cuando $N_D = 0$ (esto es, no existe ningún nodo deshabilitado), entonces la notación puede ser simplificada a $(0 : 0 : N_{Ir})$. De la misma manera, el estado $(0 : 1 : 1)$ designa el caso de que exista un nodo deshabilitado y uno *infectado* adyacente a él, independientemente de su posición absoluta en el anillo (por convención siempre denotaremos el estado de forma que $N_{II} \leq N_{Ir}$).

Como ejemplo, consideraremos el anillo de la Fig. 3.3. Inicialmente, el estado en el que se encuentra el anillo es el $(0 : 0 : 0)$ ya que todos los nodos funcionan correctamente y no hay nodos en los estados I ni D. (véase Fig. 3.5(a)). En un determinado instante de tiempo, un nodo comienza a sufrir problemas de funcionamiento en el plano de control como se ilustra en la Fig. 3.5(b). Esto ocurre con una frecuencia $8\beta_F$, ya que cualquiera de los 8 nodos del anillo podría sufrir un error espontáneo, y lleva al anillo a estado $(0 : 0 : 1)$. A partir de este punto, el nodo con fallos en el plano de control puede ocasionar por propagación las siguientes transiciones:

- al estado $(0 : 1 : 0)$ si él mismo resulta deshabilitado debido al fallo de su plano de datos, lo cual puede ocurrir con una tasa c . Este caso se puede ver en la Fig. 3.5(c).
- al estado $(0 : 0 : 2)$ si el fallo en el plano de control se propaga a un vecino. Esto ocurre con una frecuencia 2β (véase Fig. 3.5(d)) debido a que puede causar la propagación del fallo hacia cualquiera de sus dos nodos adyacentes.
- al estado $(0 : 0 : 0)$ (Fig. 3.5(a)), si el operador de red repara el nodo y lo devuelve a su estado de operatividad normal, volviendo al estado en el que es susceptible de que un nodo adyacente le propagase errores. La tasa de esta transición es δ , que coincide con la tasa de reparación.

A continuación se describe en detalle lo que ocurre en el estado $(0 : 0 : 2)$, estado desde el cual las posibles transiciones serían:

- al estado $(0 : 0 : 3)$ cuando el error se propaga a un nodo adyacente (Fig. 3.5(e)), lo que de nuevo puede ocurrir con una tasa 2β .

- al estado $(0 : 1 : 1)$ si uno de los dos nodos sufre errores en el plano de datos y queda desconectado, lo que ocurriría con frecuencia $2c$ (Fig. 3.5(f)).
- al estado $(0 : 0 : 1)$ si se repara uno de los dos nodos, lo cual tiene una frecuencia de ocurrencia de 2δ (Fig. 3.5(b)).

En contraste con este estado, puede observarse el estado $(0 : 1 : 0)$. Desde este estado no pueden ocurrir la propagación de errores en el anillo puesto que los nodos desconectados o deshabilitados no propagan los fallos. Por tanto, la única posibilidad en este caso, es la transición al estado $(0 : 0 : 0)$ cuando las acciones de reparación del nodo por parte del operador sean llevadas a cabo.

El resto de estados y transiciones son fácilmente deducibles, y llevan al modelo mostrado en la Fig. 3.4. En esta figura, existe un estado especial denominado como "DISC". Este estado especial agrupa un conjunto de varios estados: todas aquellas configuraciones en las cuales haya más de un nodo desconectado, (esto es, $x : D : y$ con $D > 1$ y para cualquier valor de x e y), y todos los estados en los que no queden nodos operativos (por ejemplo, $0 : 0 : 8$ y $0 : 1 : 7$). Este estado "DISC", indica la "desconexión total" de la red, y representa los casos en los que el operador de red ha de tomar medidas urgentes para la reparación de los errores, porque el anillo ha quedado desconectado o totalmente inoperativo.

Se ha considerado que el operador es capaz de restaurar el funcionamiento de la red en este estado de desconexión con una frecuencia de t_R . El valor de t_R se considera de varios órdenes de magnitud menor que t para tener en cuenta que este estado es mucho más problemático y difícil de reparar que la simple desconexión de un nodo aislado.

Finalmente, el objetivo del trabajo enmarcado dentro de esta tesis doctoral se centra en el estudio del diseño del modelo para la elección de las tasas de reparación, δ y t , con el fin de conseguir tener el anillo operativo al menos el 99,999% del tiempo, lo que equivale a que la probabilidad de dicho estado especial "DISC" sea inferior a 10^{-5} .

La Tabla 3.3 muestra la matriz generadora infinitesimal Q_8 para este caso particular de CTMC, tal cual se obtiene del diagrama de estados y tasas de transición entre estados de la Fig. 3.4. Los espacios vacíos equivalen a tasas iguales a 0, excepto los valores diagonales que son:

$$Q_{ii} = -\sum_{j \neq i} Q_{ij} \quad (3.5)$$

Estado desde	a	Q_{ij}	Condición
(0 : 0 : 0)	(0 : 0 : 1)	$N\beta_F$	
(0 : 0 : 1)	(0:1:1)	0	
(0 : 1 : x)	(0 : 1 : x + 1)	β	$x > 0$
(y : 1 : x)	(y : 1 : x + 1)	β	$x \geq y, y \geq 0$
(y : 1 : x)	(y + 1 : 1 : x)	β	$x \geq y, y \geq 0$
(0 : 0 : x)	(0 : 0 : x + 1)	2β	$x \geq 1$
(x : 1 : x)	(x : 1 : x + 1)	2β	$x \geq 0$
(0 : 0 : N - 1)	DISC	2β	
(0 : 0 : x)	(0 : 1 : 0)	c	$x \geq 1$
(0 : 0 : x)	$(\frac{x-1}{2} : 1 : \frac{x-1}{2})$	c	x odd and $x \geq 2$
(0 : 0 : x)	(0 : 1 : x - 1)	2c	$x \geq 1$
(0 : 0 : x)	(y : 1 : x - y - 1)	2c	$y = 1, 2, 3, \dots$ si $y \leq x$
(y : 0 : x)	DISC	$(x + y)c$	
(0 : 1 : N - 2)	DISC	$(N - 2)c + \beta$	
(y : 1 : N - y - 2)	DISC	$(N - 2)c + 2\beta$	
(0 : 1 : x)	(0 : 0 : x)	t	
(0 : 0 : 1)	(0 : 0 : 0)	2δ	
(x : 1 : x)	(x - 1 : 1 : x)	2δ	$x \geq 1$
(0 : 0 : x)	(0 : 0 : x - 1)	δ	$x \geq 2$
(y : 1 : x)	(y : 1 : x - 1)	δ	$x > y$
(y : 1 : x)	(y - 1 : 1 : x)	δ	$x > y > 0$

Tabla 3.2: Reglas de generación de transiciones para Q_N

3.4.3. Consideraciones para el modelo general de un anillo GMPLS de N nodos

Partiendo del caso explicado anteriormente de 8 nodos, es sencillo inferir las reglas de construcción de la matriz generadora Q_N partiendo de la mostrada en la Tabla 3.3, para un anillo genérico de un número arbitrario de nodos N . Estas reglas están resumidas en la Tabla 3.2.

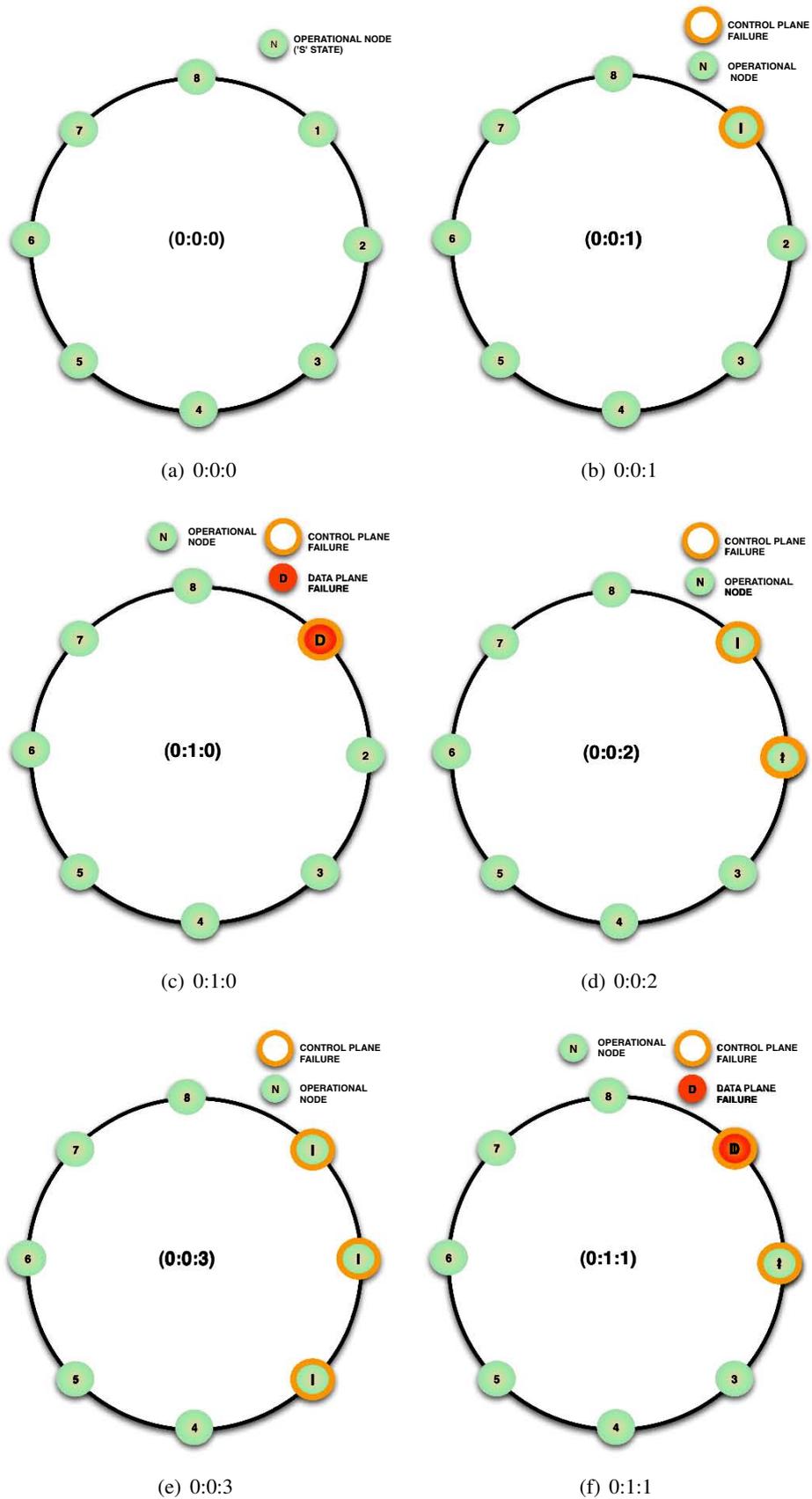


Figura 3.5: Ejemplo de posibles configuraciones del anillo y su estado correspondiente: (a) (0:0:0), (b) (0:0:1), (c) (0:1:0), (d) (0:0:2), (e) (0:0:3) y (f) (0:1:1)

Q_8	0:0:0	0:0:1	0:0:2	0:0:3	0:0:4	0:0:5	0:0:6	0:0:7	0:1:0	0:1:1	0:1:2	0:1:3	0:1:4	0:1:5	0:1:6	1:1:1	1:1:2	1:1:3	1:1:4	1:1:5	2:1:2	2:1:3	2:1:4	3:1:3	DISC
0:0:0	δ	$N\beta_F$																							
0:0:1		2δ	2β						c																
0:0:2			2δ	2β						$2c$															
0:0:3				2δ	2β						$2c$														
0:0:4					2δ	2β						$2c$													
0:0:5						2δ	2β						$2c$												
0:0:6							2δ							$2c$											
0:0:7								2β																	
0:1:0	t								δ																
0:1:1		t								δ															
0:1:2			t								β														
0:1:3				t								β													
0:1:4					t								β												
0:1:5						t								δ											
0:1:6							t								δ										
1:1:1								2δ								δ									
1:1:2										δ							2β								
1:1:3											δ							β							
1:1:4												δ							β						
1:1:5													δ							β					
2:1:2														δ							2β				
2:1:3																						δ			
2:1:4																							δ		
3:1:3																								δ	
DISC																									

Tabla 3.3: Matriz generadora infinitesimal para un anillo óptico GMPLS de 8 nodos (Q_8)

3.5. Resultados Numéricos

3.5.1. Métricas de prestaciones

A continuación se muestra el estudio de la probabilidad estacionaria de un conjunto de estados, claves para el diseño del sistema, los cuales representan los siguientes tipos de problemas en la red (ver Fig. 3.6). Dichos conjuntos de estados serían:

- Estado de operatividad completa o *Fully Operational* ($0 : 0 : 0$): En este estado, en el cual los nodos en el anillo están todos funcionando correctamente, pueden surgir errores sólo de forma espontánea. La probabilidad que caracteriza a este estado es $P_{(0:0:0)}$, esto es, el porcentaje de tiempo en el cual el sistema tiene a todos sus nodos operativos.
- Infección moderada (*Moderate Infection*), estados denotados por $(0 : \{0, 1\} : < N_{I,max})$: Este conjunto de estados contiene todas aquellas configuraciones del anillo en las cuales el número de nodos con problemas en su plano de control está por debajo de un determinado umbral $N_{I,max}$, y el número de nodos desconectados es 0 ó 1. Este caso se caracteriza por $P_{low,I}$ la cual se calcula como la suma de las probabilidades estacionarias de todos los estados que cumplen estas condiciones. Si el anillo está en alguno de los estados englobados en este grupo, indica una situación de propagación moderada de fallos a lo largo de los nodos del anillo, y requiere una atención no crítica por parte del operador para evitar que la propagación se convierta en un problema más severo.
- Infección Severa (*Severe Infection*), estados con $(0 : \{0, 1\} : > N_{I,max})$: este conjunto contiene todos los estados del anillo en los cuales el número de nodos con problemas excede el umbral $N_{I,max}$, y el número de nodos desconectados sigue siendo 0 ó 1. A este grupo de estados se le caracteriza con el valor $P_{high,I}$, que al igual que en el caso anterior es la suma de la probabilidad de encontrarse en cualquiera de todos los estados que lo componen. Este grupo de estados describe una situación de propagación severa en el anillo y de difícil control y restauración por parte del operador, y debe considerarse como el paso previo y detonante de la situación de desconexión completa.
- Desconexión, estado (*DISC*): este estado es aquel en el cual el anillo ha quedado desconectado porque tiene más de un nodo deshabilitado, o el fallo propagado ha llegado a la completa totalidad de nodos del anillo. En este caso, existen al menos dos nodos que no pueden comunicarse entre ellos, lo cual es inaceptable para la mayor parte de los operadores

de red. Este caso se caracteriza por la probabilidad P_{DISC} , esto es, el porcentaje de tiempo que el anillo está en un estado de desconexión debido al fallo total de más de un nodo o del fallo parcial de todos ellos.

El valor del umbral $N_{l,max}$ se elige entre 0 y N . En los resultados que se muestran a continuación se utiliza $N_{l,max} = N/2$, esto es, consideraremos que el problema ha pasado a un estado de propagación de fallos *severo* cuando el 50 % de los nodos tengan algún tipo de problema.

Los grupos descritos se muestran en la Fig. 3.6. El estado de operatividad completa está marcado usando un pentágono en la figura, los estados pertenecientes a una situación "moderada" se muestran dentro de rectángulos, mientras que los correspondientes a una situación "severa" están enmarcados dentro de figuras ovaladas.

Además, es interesante estudiar los tiempos medios de primera ocurrencia de los tres grupos de estados que designan situaciones de fallo para tener una medida de lo a menudo que podrían darse estas situaciones partiendo del estado inicial donde todos los nodos están operativos (estado (0 : 0 : 0)). Los siguientes ejemplos describen estas métricas de prestaciones en detalle.

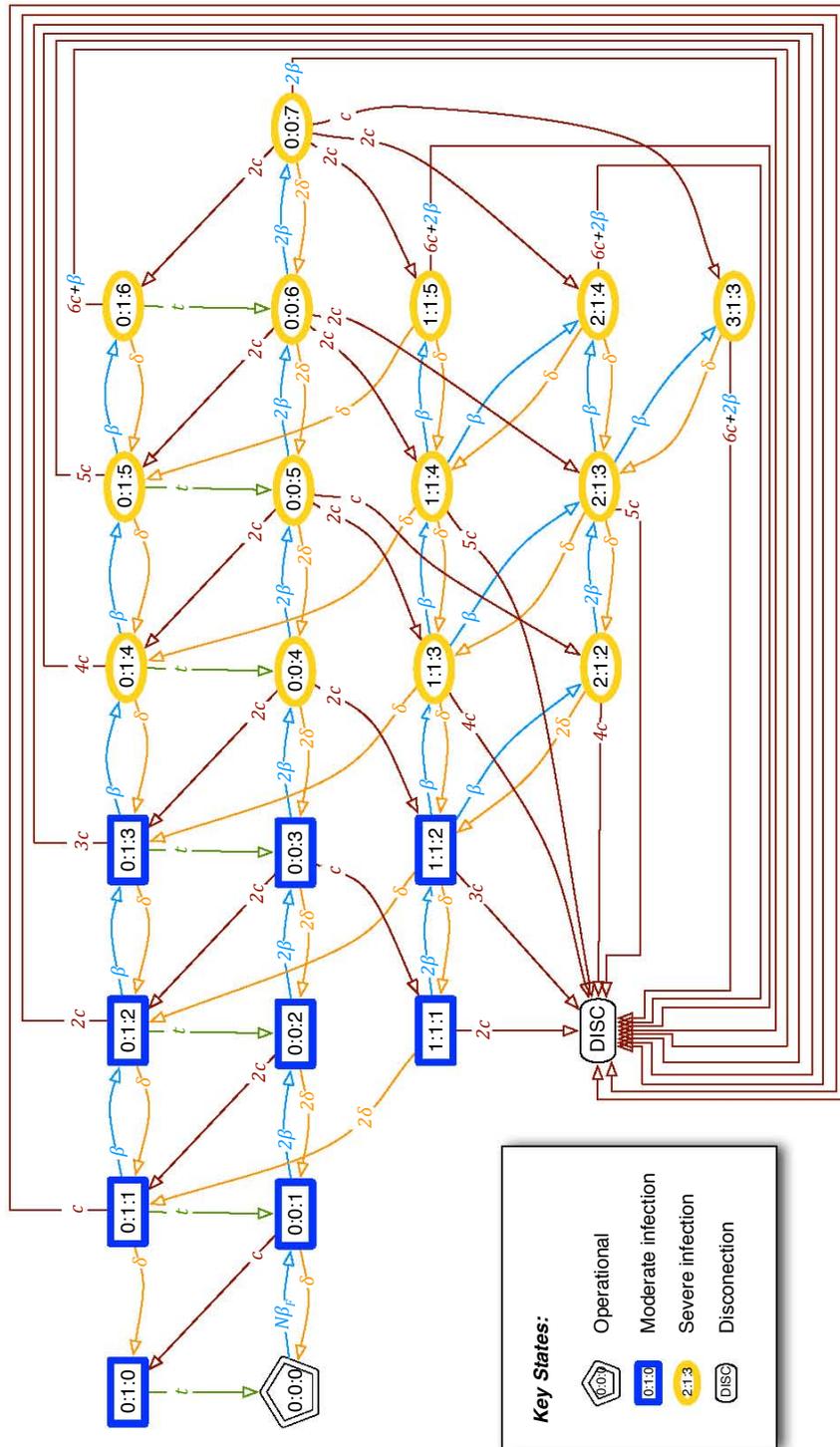


Figura 3.6: Grupos de estados estudiados

3.5.2. Resultados para las probabilidades de estado estacionario

Una vez resueltas las probabilidades correspondientes al estado estacionario para cada uno de los estados de la cadena de Markov del modelo, es sencillo mostrar el porcentaje de tiempo que el anillo permanece en cada estado como una función de las dos tasas de restauración o reparación: la tasa de reparación del plano de control del nodo, δ (esto es, la tasa de transición entre el estado S y el estado I del modelo SID) y la tasa t a la cual los nodos son reparados por completo tras un fallo global que los desconecte (transición entre los estados D y S). Las unidades de todas las tasas se presentan normalizadas a número de eventos de transición (reparación o fallo) que pueden ocurrir en un intervalo infinitesimal de tiempo según el modelo de la CTMC. Los resultados numéricos se presentan para diferentes valores de δ , β , t y c . Las figuras usan la notación acorde a los grupos de estado sujetos a estudio: el subíndice "000" se utiliza para denotar el estado de "operatividad completa" (o "Fully Operational"), el subíndice "High" representa el caso de "infección severa", "Low" se usa para "infección moderada" finalmente "DISC" hace referencia al estado de desconexión generalizada.

La Fig. 3.7 muestra las probabilidades en estado estacionario de la CTMC para el anillo GMPLS de 8 nodos para diferentes valores de δ , β , c y t . La tasa de fallo del plano de datos una vez ha dejado de funcionar el de control se ha fijado a un valor normalizado de $c = 1$ reparaciones por unidad de tiempo (por ejemplo, 1 hora) para todas las figuras. Los dos gráficos superiores utilizan un valor fijo de $\delta = 5$ (tasa de reparación del plano de control) mientras que los dos inferiores consideran $\delta = 100$. En las dos figuras de la izquierda se considera una tasa de propagación del errores en el plano de control por unidad de tiempo de $\beta = 1$ (tasa de infección), mientras que en los dos gráficos del lado derecho, la tasa de propagación está fijada en $\beta = 20$.

A la vista de la figura un operador de red puede sacar conclusiones importantes para la toma de decisiones del diseño de la protección y restauración de su red:

1. En primer lugar, se observa que para los cuatro escenarios, la probabilidad estacionaria es casi constante, independientemente del valor que tome t .
2. En segundo lugar, las figuras de las esquinas superior izquierda e inferior derecha, tienen la misma relación $\beta/\delta = 1/5$, y muestran unas probabilidades muy parecidas.
3. Sólo el gráfico inferior izquierdo de la Fig. 3.7 muestra un valor de P_{DISC} razonablemente bajo, lo que ocurre con una relación $\beta/\delta = 1/100$ baja.

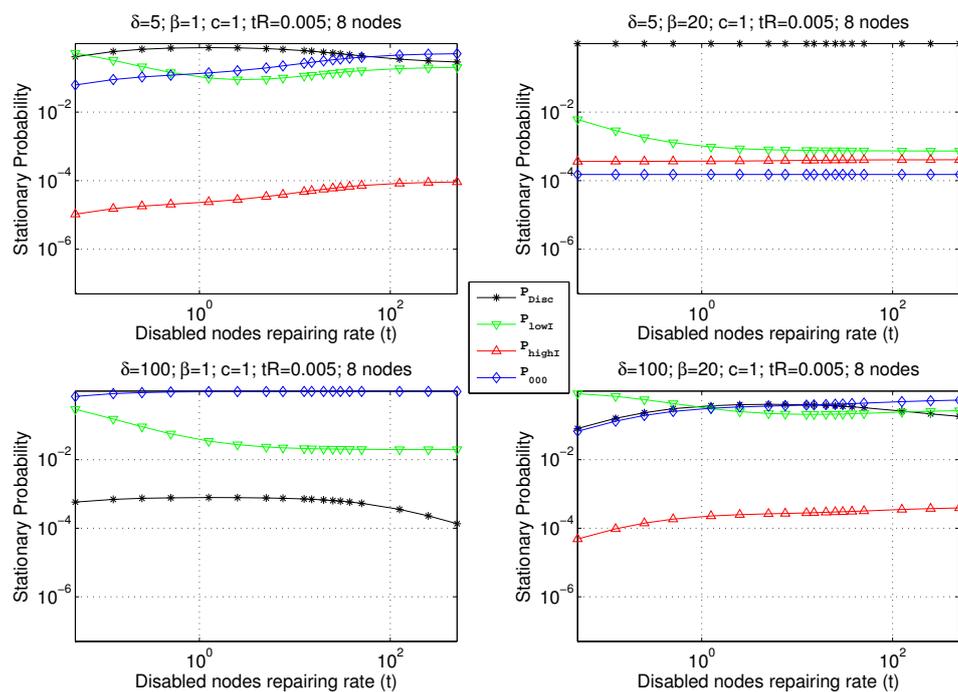


Figura 3.7: Impacto de t en las probabilidades de estado estacionario de la CTMC de un anillo GMPLS de 8 nodos

En resumen, la relación entre β/δ es la clave para conseguir valores de suficientemente bajos P_{DISC} tal y como se puede observar en la Fig. 3.8.

En la Fig. 3.8 se utilizan valores de $t = 0,5$ (gráficas superiores) y $t = 50$ (gráficas inferiores) para diferentes tasas de infección $\beta = 1$ (gráficas de la columna izquierda) y $\beta = 20$ (gráficas en la columna de la derecha). Las curvas demuestran un comportamiento similar en el hecho de que baja la probabilidad de desconexión total P_{DISC} para valores altos de δ y diferentes valores de β .

En definitiva, para poder conseguir una probabilidad de desconexión total por debajo de 10^{-5} , es deseable que $\delta > 4 \times 10^2 \beta$ cuando $\beta = 1$ (Fig. 3.8 columna de la derecha) y $\delta > 4 \times 10^3 \beta$ si $\beta = 20$ (Fig. 3.8 columna de la izquierda).

Por lo tanto, es seguro tener una tasa de reparación δ alrededor de tres órdenes de magnitud mayor que la tasas de propagación del fallo β . Dicho de otro modo, es necesario que $\delta > 10^3 \beta$ para garantizar la disponibilidad de la red en el 99,999 % del tiempo.

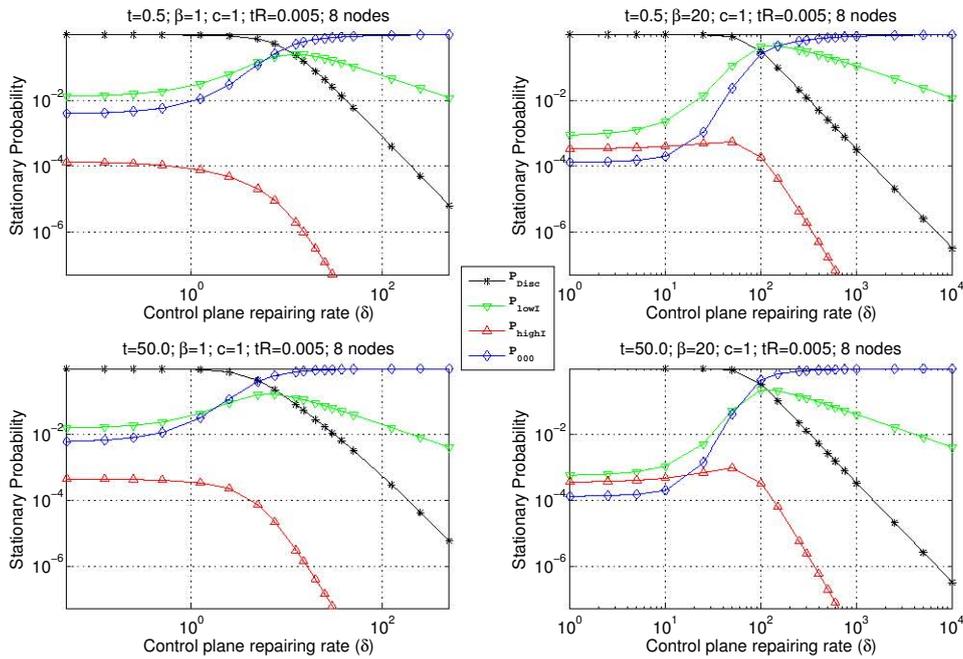


Figura 3.8: Efecto de δ en las probabilidades estacionarias de una CTMC para un anillo GMPLS de 8 nodos

3.5.3. Tiempos de primera ocurrencia (*First-Passage Times*): MTTF y MTTR

En esta sección se presenta el estudio hecho acerca de los tiempos medios de primera ocurrencia (FPT - *First-Passage Times*) para los tres subconjuntos de estados con errores de algún tipo: Fallo Moderado, Fallo Severo y Desconexión; comenzando siempre desde el estado inicial sin errores (0 : 0 : 0).

De esta manera, podemos calcular para ello el MTTF (*Mean Time To Failure*, o tiempo hasta el primer error), el cual se refiere al tiempo medio en alcanzar alguno de los estados con errores, desde el estado de operatividad completa (0 : 0 : 0), usando la expresión de la ecuación 3.4. Tal y como se muestra en la Fig. 3.9, el comportamiento es de nuevo independiente de t (ya que las figuras superior e inferior tienen la misma apariencia).

Finalmente, el análisis del tiempo medio para la reparación, MTTR (*Mean Time To Repair*), se muestra en la Fig. 3.10. Esta figura describe los tiempos medios que se requieren para volver a dejar el anillo en su estado original con todos los nodos funcionando (0 : 0 : 0), desde cualquiera de los grupos de estados de la cadena. Con el objetivo de conseguir tiempos de MTTR bajos (del orden de 10^{-2} unidades de tiempo), los valores de δ es necesario que sean de varios órdenes de magnitud superiores a β .

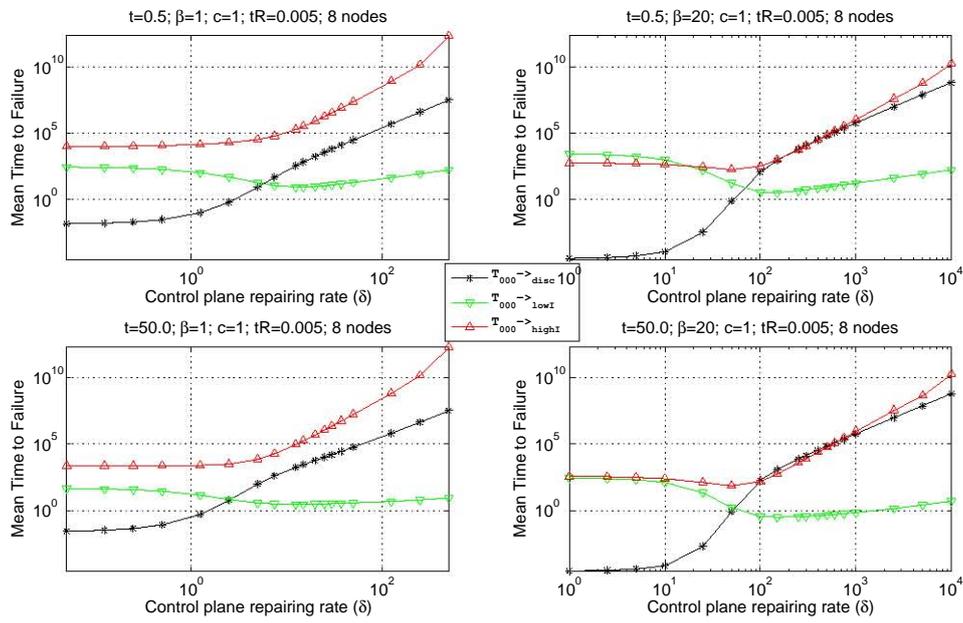


Figura 3.9: Tiempo Medio para el primer fallo - Mean Time To Failure (MTTF)

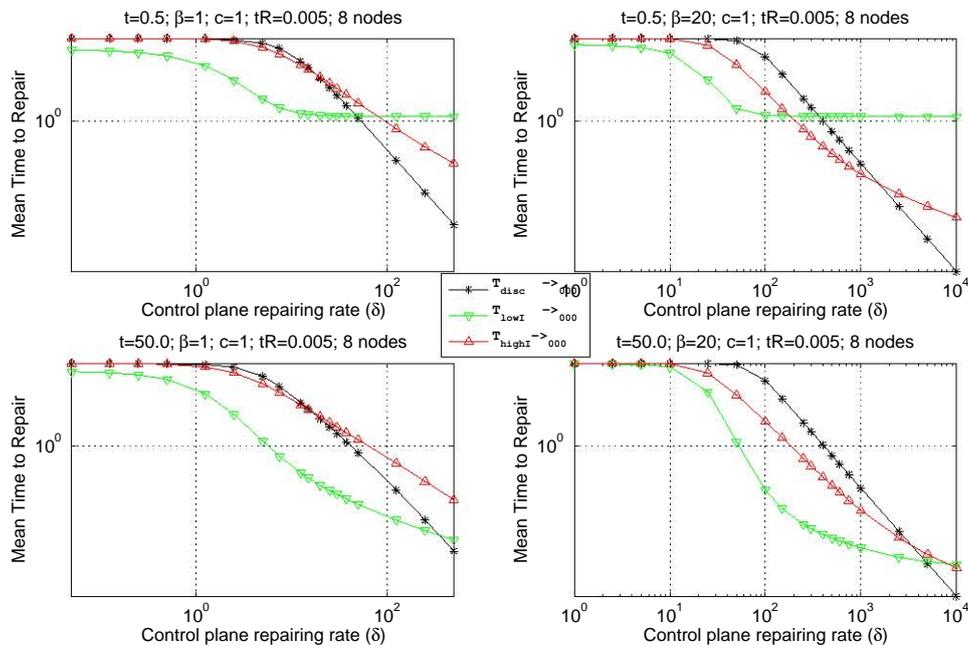


Figura 3.10: Tiempo medio para la reparación - Mean Time To Repair (MTTR)

3.5.4. Anillo GMPLS de 32 nodos

Esta sección, finalmente, presenta los resultados numéricos obtenidos para un anillo más grande: 32 nodos. Usando las mismas reglas que se describen en la subsección 3.4.3, se calcula la matriz generadora Q_{32} y con ella las probabilidades de estado estacionario se calculan resolviendo la cadena de Markov para el anillo de 32 nodos.

La Fig. 3.11 muestra la relación δ/β (eje de ordenadas) que se requiere para conseguir un cierto nivel de fallo del servicio (eje de abscisas) asumiendo diferentes valores de β y t . Tal y como puede observarse, para un nivel de fallo deseado por debajo de 10^{-5} el valor de δ debería de estar entre 10^2 y 10^3 veces el valor de β .

Esta conclusión es la misma a la que se llega en el caso del anillo de 8 nodos, por lo que una generalización de esta regla proporciona al operador una regla útil en el diseño del valor estratégico de su propia δ .

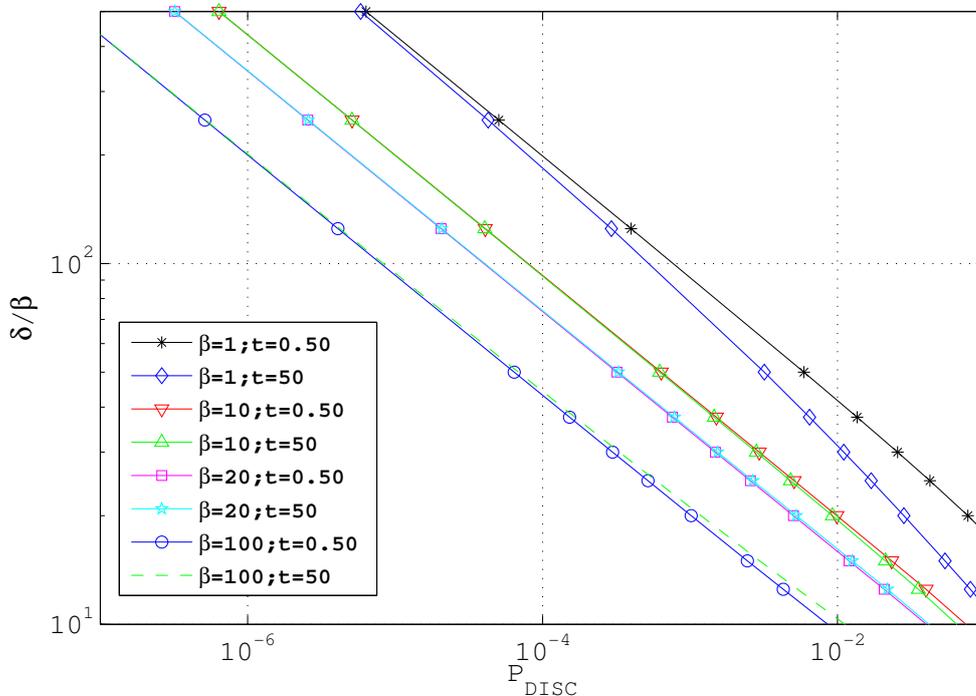


Figura 3.11: Efecto de δ/β en las probabilidades estacionarias de un anillo de 32 nodos

3.6. Resumen del capítulo y conclusiones

Esta contribución describe el uso de un modelo CTMC para caracterizar el comportamiento transitorio y los posibles estados de redes basadas en GMPLS con topología en anillo cuyos nodos pueden desarrollar fallos o desconectarse siguiendo un modelo de propagación epidémico de los fallos denominado SID. Se presentan un conjunto completo de resultados numéricos, centrado en el análisis de las probabilidades estacionarias resultantes junto con los parámetros relacionados con la gestión de fallos (MTTF y MTTR) de un número discreto de valores de las tasas de reparación δ y t .

El modelo descrito puede ser de utilidad a los operadores de red para encontrar las tasas de reparación de los planos de control y de datos respectivamente δ y t para mantener un cierto nivel de disponibilidad, digamos del 99.999 % del tiempo. Además, el modelo puede usarse también para estudiar la sensibilidad de la red a distintas combinaciones de las tasas de reparación del modelo SID, en términos de la esperanza del número de nodos e cada estado.

Como conclusión, de los resultados calculados para anillos de diferentes tamaños, una buena regla de diseño sería el conseguir que la tasa de reparación del plano de control de nodos "infectados" sea mucho más alta (de unos tres órdenes de magnitud) que la tasa de propagación de los errores β . Básicamente, cuando δ es muy grande con respecto a β conseguiremos que los nodos con problemas sean reparados muy rápidamente, reduciendo así al mínimo la probabilidad de infectar a sus vecinos. Esto puede verse claramente en el siguiente supuesto: cuando la CTMC transita desde el estado $(0 : 0 : 0)$ al estado $(0 : 0 : 1)$, entonces el siguiente movimiento es el tránsito a $(0 : 0 : 2)$, $(0 : 1 : 0)$ or de vuelta a $(0 : 0 : 0)$, cuyas tasas son 2β , c y δ respectivamente. Si se mantiene $\delta \gg (2\beta + c)$, el operador de red se asegura que la probabilidad de que los errores se propaguen a los nodos adyacentes sea muy pequeña:

$$\frac{2\beta}{2\beta + \delta + c} \tag{3.6}$$

Por tanto, el operador de red debe asegurar que $\delta \gg \beta$ con el objetivo de frenar la propagación de fallos antes de tener que pensar en su reparación.

Es de importancia para quien desee
alcanzar una certeza en su
investigación,
el saber dudar a tiempo.

Aristóteles

Capítulo 4

Análisis del retardo medio y su varianza en anillos WDM libres de colisión mediante el uso de segmentos de recirculación del tráfico bloqueado

4.1. Introducción y trabajos anteriores relacionados

La siempre creciente demanda de ancho de banda por parte de usuarios y aplicaciones ha hecho converger el acuerdo de la comunidad investigadora en la conclusión de que la fibra óptica y la multiplexación por división en longitud de onda (WDM) podrían ser la solución para satisfacer dicha demanda en la Internet de nueva generación. Es un hecho que las arquitecturas basadas en WDM, como se ha venido comentando en este documento, se han propuesto tanto para la red de acceso, como para las redes de área metropolitana y troncal. En lo que concierne a la redes metropolitanas, la arquitectura más desplegada en la actualidad sigue siendo la topología en anillo con tecnología SONET/SDH, la cual combina transmisión óptica con la conmutación basada en multiplexación por división en tiempo (TDM) en cada nodo del anillo.

Gracias a la aparición de los OADM reconfigurables (ROADMs), en los últimos años se han propuesto topologías en anillo con WDM más eficientes en coste para reemplazar los nodos SONET/SDH intermedios, donde el tráfico no debería sufrir la conversión OEO actual y conmutarse a

continuación hacia el nodo destino, sino que debería atravesarlos directamente de forma completamente óptica. Este tipo de arquitecturas necesitan la definición de nuevos mecanismos de control de acceso al medio (MAC) para arbitrar el uso del canal ya que, en la mayoría de los casos se dispone de una única longitud de onda común compartida por muchos nodos origen y se pueden ocasionar colisiones. Este es el caso de los anillos TT-FR (*Tunable-Transmitter Fixed-Receiver*), donde los nodos tendrán una sola longitud de onda configurada para recepción local de tráfico, y al mismo tiempo, disponen de un láser sintonizable que les permite transmitir tráfico en las longitudes de onda dedicadas para el resto de nodos del anillo [92,93].

En las topologías TT-FR, las colisiones en el receptor no pueden ocurrir ya que cada nodo destino está siempre escuchando en su propio canal dedicado, pero sí que pueden originarse colisiones en transmisión, ya que varios nodos fuente han de competir por acceder al canal dedicado para alcanzar sus destinos comunes. En este caso el uso de un mecanismo MAC es indispensable para evitar estas colisiones.

Las topologías TT-FR iniciales consideraban el ancho de banda repartido en intervalos temporales (con tiempos de envío máximos ajustados al tamaño máximo de paquete) con una prioridad superior para el tráfico en tránsito que para el tráfico generado localmente. Adicionalmente, todos los nodos disponían de mecanismos capaces de inspeccionar el canal, de forma que fuesen capaces de detectar la presencia de señal, y por consiguiente, conocer el estado del mismo (ocupado o disponible) para los siguientes intervalos, y decidir de alguna manera si transmitir o no transmitir en ellos.

En la solución de "múltiples anillos WDM" propuesta en [94], se utiliza un protocolo MAC libre de colisiones donde cada nodo mantiene tantas colas lógicas independientes como nodos destino haya (y por tanto, longitudes de onda en las que transmitir), de forma que un protocolo de Round Robin síncrono (SRR) la vaya escaneando secuencialmente para la transmisión de los paquetes. Esta solución reduce el llamado "bloqueo de cabecera de línea" o "*HOL (Head of Line) blocking*" el cual aparece en los nodos que sólo disponen de una cola FIFO (*First Input First Output*) común para todos los destinos, y en los cuales los paquetes en la cabecera de la cola encuentran bloqueado el canal de su destino, provocando el consiguiente bloqueo de todas las longitudes de onda, ya que aquella que deben usar los primeros está siendo ocupada. Sin embargo, los autores identifican algunas situaciones de retardo y colapso que pueden ocurrir y que originan la necesidad de definir un protocolo de equilibrio que gestione por encima del SRR la garantía de igualdad de oportunidades de todas las colas en el acceso al anillo. Este mecanismo se extiende

con la contribución en [95] para permitir la diferenciación de servicios para diferentes clases de QoS.

En paralelo, otra propuesta de anillo TT-FR llamada HorNet, es probada por los autores de [96]. HorNet se concibe originariamente para usar CSMA/CA (*Carrier Sense Multiple Access/Collision Avoidance*) como protocolo MAC para evitar colisiones [97], pero fue rápidamente reemplazado por un sistema MAC con canal de control dedicado en el cual un paquete testigo (*token*) advierte de la disponibilidad de los siguientes intervalos de tiempo futuros. El equilibrio entre nodos se resuelve aquí limitando el número máximo de reservas que un nodo dado puede realizar sobre el testigo que atraviesa el canal de control fuera de banda.

Sin embargo, ambas propuestas, RingO y HorNet necesitan del reparto en tiempo del ancho de banda disponible. Recientemente un nuevo protocolo MAC se ha propuesto para el Anillo de Red Óptica en Bus Dual (*Dual Bus Optical Ring Network-DBORN* [98]), y se ha demostrado que es posible emplear CSMA con tamaños variables de paquetes usando componentes ópticos pasivos baratos, esto es, simplificando la arquitectura óptica en anillo a un coste moderado [99]. Este protocolo MAC se basa en la idea de que cada nodo dispone de una ventana de detección creada mediante el retardo del tráfico en tránsito a través del nodo durante el tiempo máximo de trama usando Retardos de Línea de Fibra (FDL). Esto le proporciona al nodo el tiempo suficiente para decidir si hay o no suficiente espacio entre dos ráfagas de tráfico como para permitirse planificar el envío del siguiente paquete en espera en su cola local. A pesar de que esta técnica es muy eficiente, el hecho de que la prioridad le sea otorgada al tráfico en tránsito sobre el local puede originar serios problemas de bloqueo HOL en los nodos del anillo cercanos al destino, ya que sólo pueden aprovechar los espacios libres entre las ráfagas de paquetes en tránsito de todos los nodos que les precedan.

En la contribución que se presenta en este capítulo, se propone y analiza un protocolo MAC que resuelve la colisión en el acceso al canal basándose en recircular en sentido opuesto, el tráfico bloqueado perteneciente a segmentos del anillo precedentes en el sentido de la transmisión. En este anillo, la prioridad se le otorga al tráfico local sobre el tráfico en tránsito, y cuando ocurre la colisión, el tráfico entrante es reflejado por una fibra independiente en sentido contrario en el anillo al de la transmisión de los datos. De esta manera, los segmentos de anillo se usan en cierto modo como sistemas de almacenamiento en los cuales se le otorga al tráfico bloqueado una segunda oportunidad de atravesar el nodo que encontró bloqueado en el intento anterior. En esencia, el cambiar la prioridad sobre el tráfico local en lugar de sobre los paquetes en tránsito

provoca resultados en el acceso inferiores a la media para los nodos más distantes al destino, lo que se traduciría en un aumento del retardo experimentado. Sin embargo, el caso descrito anterior genera el efecto contrario: se le niega el acceso a los nodos del anillo situados más cerca del destino, ya que al llegar a ellos cada vez quedan menos intervalos libres para el acceso al medio entre medias de los paquetes en tránsito, una vez se van multiplexando el tráfico de las múltiples fuentes, y no es suficiente para mandar los paquetes grandes.

4.2. Control de colisión mediante recirculación de tráfico bloqueado

En esta sección se describe de forma resumida el modo de funcionamiento habitual de un anillo WDM bidireccional en el cual se recircula el tráfico bloqueado. La Fig. 4.1 muestra un ejemplo de anillo de 4 nodos bidireccional de este tipo, donde se asume que el tráfico viaja en el sentido de las agujas del reloj (sobre la línea sólida). De forma general, se elige el nodo n_0 como el nodo destino de los flujos de tráfico local, cuyas cargas ρ_i , $i = 1, 2, 3$ son generadas por los tres nodos precedentes n_1 , n_2 and n_3 . Siguiendo la arquitectura TT-FR, dicho tráfico es inyectado por los nodo transmisores en una longitud de onda λ_0 , y extraído del anillo en el nodo receptor n_0 , el cual sólo es capaz de recibir en dicha λ_0 . Además, se denominan como enlaces 1^+ , 2^+ y 3^+ (líneas sólidas en la figura) a las fibras conectando los pares de nodos 0-1, 1-2 y 2-3 respectivamente, en el sentido de las agujas del reloj (sentido normal de la transmisión), y se denota como 1^- , 2^- y 3^- (líneas discontinuas de la figura) a las fibras ópticas que conectan las mismas parejas de nodos anteriores, 0-1, 1-2 y 2-3 en el sentido contrario, opuesto a las agujas del reloj. Finalmente, el tráfico originado por el nodo n_3 debe atravesar los nodos n_2 y n_1 respectivamente para alcanzar el nodo destino n_0 ; el tráfico inyectado por el nodo n_2 debe viajar a través del nodo n_1 antes de recibirse por el destino, y que el tráfico generado en n_1 no atraviesa ningún nodo y es recibido directamente en el nodo n_0 a través del enlace 1^+ .

Básicamente, cuando un nodo tiene que transmitir cualquier paquete al nodo n_0 , antes de hacerlo tiene que comprobar la disponibilidad de la longitud de onda dedicada a este nodo y, en caso de estar disponible, bloquear el canal el acceso al canal λ_0 para prevenir colisiones con el tráfico en tránsito de nodos que le preceden en el anillos. El bloqueo en este caso consiste en que todo el tráfico que llegue de estos nodos precedentes a partir de ese momento se redirecciona en sentido contrario hacia el nodo anterior a través de la fibra que transporta el tráfico que viaja en sentido opuesto a las agujas del reloj (línea discontinua en la figura) Este tráfico bloqueado y que

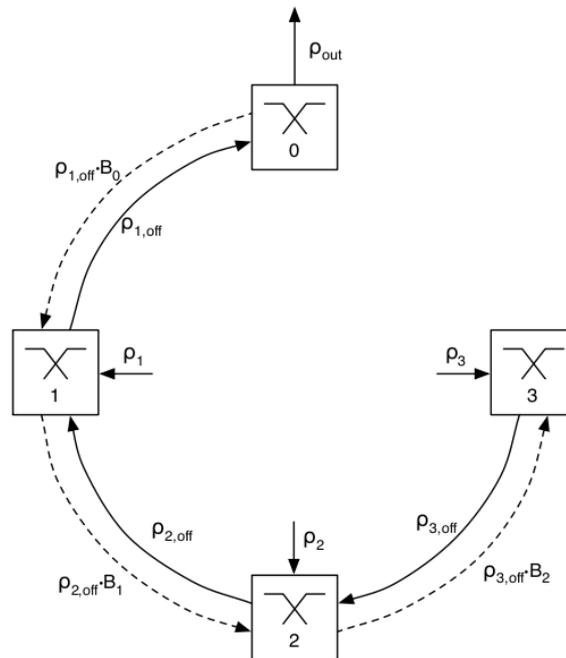


Figura 4.1: Ejemplo de anillo WDM de 4 nodos con recirculación de tráfico. Notación de las intensidades de tráfico.

viene redirigido de vuelta, se puede tratar en el nodo anterior usando dos políticas diferentes, tal y como se ilustra en las Fig. 4.2(a) and 4.2(b):

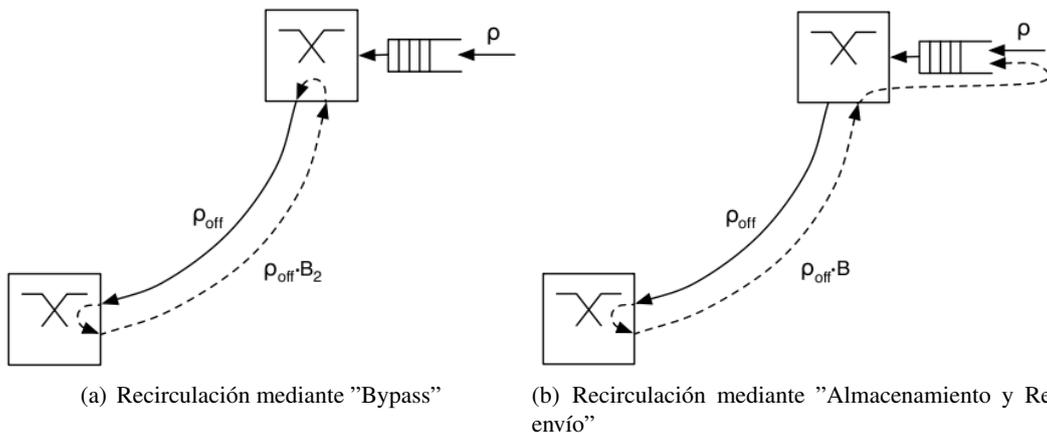


Figura 4.2: "bypass de recirculación" (izquierda) y "almacenamiento y reenvío" (derecha)

- Recirculación con *bypass* (Fig. 4.2(a)): En este caso, el tráfico recirculado tiene prioridad sobre el tráfico local y se deja pasar directamente (*bypass*) de nuevo hacia el nodo bloqueado sin realizar ninguna conversión OEO. En caso de que haya tráfico en el nodo que debe ser enviado, deberá esperar hasta que se libere el canal.
- Recirculación con almacenamiento y reenvío (Fig. 4.2(b)): En este caso, el tráfico que viene

devuelto desde el nodo bloqueado se convierte al dominio electrónico (conversión OE) y se almacena en la misma cola de espera que el tráfico que tiene el nodo para ser enviado. A este tráfico no se le asigna ninguna prioridad diferente que al tráfico local.

Evidentemente, es de esperar que los paquetes experimenten un retardo extremo a extremo inferior en el primer caso ya que no sufren conversión OEO para llegar de nuevo al nodo bloqueado. Sin embargo, si la recirculación del tráfico ocurre en demasiadas ocasiones, la calidad de la señal en su relación señal-ruido (SNR) puede verse degradada en exceso, significando la pérdida de los datos. El sistema de "bypass" para el tráfico bloqueado será preferible en situaciones de baja carga (la recirculación se realizará en pocas ocasiones), mientras que a cargas altas, cuando la esperanza de sufrir bloqueo sea alta, la conversión OEO puede ser imprescindible. Adicionalmente, cuando la longitud de los segmentos de fibra fuese demasiado larga, siendo entonces la atenuación considerable, entonces el sistema con almacenamiento y reenvío en el nodo anterior será preferible con el objeto de regenerar la señal. Sin embargo, en ambos casos, los segmentos del anillo actuarán como si de sistemas de almacenamiento óptico en serie ("buffers") se tratase.

4.3. Análisis de los sistemas de recirculación y del retardo extremo a extremo

Siendo $N + 1$ el número de nodos en el anillo de la figura Fig. 4.1, la notación $n_i, i = 0, \dots, N$ permite etiquetar cada uno de los nodos del anillo óptico. Puede inferirse el caso general del estudio particular del retardo experimentado por los paquetes inyectados por cualquiera de los nodos fuente $i = 1, \dots, N$ hacia el nodo de destino seleccionado n_0 , bajo ambas políticas de gestión de tráfico bloqueado: con reenvío directo o con almacenamiento y reenvío. El estudio del bloqueo descrito a continuación serviría para ambas situaciones.

4.3.1. Análisis de la probabilidad de bloqueo

Como se ha descrito con anterioridad, cada nodo dispone de un canal local para recepción. En este caso, el nodo de destino n_0 está recibiendo permanentemente la longitud de onda λ_0 . De esta forma, los nodos que le preceden en el anillo compiten por acceder a al recurso común que es el canal λ_0 .

ρ_i denota la carga de tráfico ofrecido por el nodo n_i en la longitud de onda λ_0 (esto es, el tráfico ofrecido con destino a n_0). Además, $\rho_{i,off}$ denota el tráfico ofrecido total en el segmento i^+ , el cual básicamente comprende la suma de las intensidades de los flujos de tráfico que lo atraviesan. De acuerdo a la ley de conservación de flujos en cada nodo, la suma de la intensidades de los flujos de entrada debe ser igual a la suma de los de salida. Esto se traduce en el siguiente sistema de $N + 1$ ecuaciones:

$$\begin{aligned}
\rho_{out} + \rho_{1,off}B_0 &= \rho_{1,off} \Rightarrow \rho_{out} = \rho_{1,off}(1 - B_0) \\
\rho_{1,off} + \rho_{2,off}B_1 &= \rho_1 + \rho_{2,off} + \rho_{1,off}B_0 \Rightarrow \rho_{1,off}(1 - B_0) = \rho_1 + \rho_{2,off}(1 - B_1) \\
\rho_{2,off} + \rho_{3,off}B_2 &= \rho_2 + \rho_{3,off} + \rho_{2,off}B_1 \Rightarrow \rho_{2,off}(1 - B_1) = \rho_2 + \rho_{3,off}(1 - B_2) \\
&\vdots \\
\rho_{N,off} &= \rho_N + \rho_{N,off}B_{N-1} \Rightarrow \rho_{N,off}(1 - B_{N-1}) = \rho_N
\end{aligned}$$

donde $\rho_{i+1,off}B_i$ indica la cantidad de tráfico ofrecido por al i y bloqueado en dicho nodo (con probabilidad B_i).

Se puede demostrar con las ecuaciones anteriores que:

$$\begin{aligned}
\rho_{out} &= \rho_{1,off}(1 - B_0) = \rho_1 + \rho_{2,off}(1 - B_1) \\
&= \rho_1 + \rho_2 + \rho_{3,off}(1 - B_2) = \dots = \\
&= \rho_1 + \rho_2 + \dots + \rho_{N,off}(1 - B_{N-1}) = \\
&= \sum_{i=1}^N \rho_i
\end{aligned} \tag{4.1}$$

lo cual confirma que no se pierden datos (ley de conservación de tráfico), esto es, todo el tráfico inyectado en la red es recibido en algún momento en n_0 .

Dado que $\rho_{out} = \sum_{i=1}^N \rho_i$, los valores del tráfico ofrecido $\rho_{i,off}$ en cada nodo pueden derivarse de:

$$\begin{aligned}
 \rho_{1,off} &= \frac{\rho_{out}}{1 - B_0} = \frac{1}{1 - B_0} \sum_{i=1}^N \rho_i \\
 \rho_{2,off} &= \frac{\rho_{1,off}(1 - B_0) - \rho_1}{1 - B_1} = \frac{1}{1 - B_1} \sum_{i=2}^N \rho_i \\
 \rho_{3,off} &= \frac{\rho_{2,off}(1 - B_1) - \rho_2}{1 - B_2} = \frac{1}{1 - B_2} \sum_{i=3}^N \rho_i \\
 &\vdots \\
 \rho_{N,off} &= \frac{\rho_{N-1,off}(1 - B_{N-2}) - \rho_{N-1}}{1 - B_{N-1}} = \frac{1}{1 - B_{N-1}} \rho_N
 \end{aligned} \tag{4.2}$$

Por lo tanto, en general para el nodo n_i :

$$\rho_{i,off} = \frac{1}{1 - B_{i-1}} \sum_{k=i}^N \rho_k, \quad i = 1, \dots, N \tag{4.3}$$

Esta información es muy útil a la hora de calcular las probabilidades de bloqueo B_i experimentadas por el tráfico en tránsito en cada nodo n_i , usando las conocidas ecuaciones de perdidas Erlang. Esencialmente, la probabilidad de bloqueo que experimenta el tráfico de todos los nodos sigue la expresión:

$$B_i = E_B(\rho_i + \rho_{i,off} B_{i-1}, 1), \quad i = 1, \dots, N \tag{4.4}$$

donde $E_B(\rho, M)$ se refiere a la ecuación Erlang-B para ρ unidades de tráfico servidas por M circuitos.

Cabe remarcar que:

$$E_B(\rho, M) = \frac{\frac{\rho^M}{M!}}{\sum_{j=0}^M \frac{\rho^j}{j!}} \tag{4.5}$$

Por lo tanto, usando Eq. 4.3 en el caso de $\rho_{i,off}$, la ecuación 4.4 de arriba resulta en:

$$B_i = \frac{\rho_i + \frac{B_{i-1}}{1 - B_{i-1}} \sum_{k=i}^N \rho_k}{1 + \rho_i + \frac{B_{i-1}}{1 - B_{i-1}} \sum_{k=i}^N \rho_k}, \quad i = 1, \dots, N \tag{4.6}$$

con $B_0 = 0$ dado que el nodo de destino nunca sufre bloqueo debido a que él nunca inyecta tráfico en su λ_0 .

De esta manera, dado un conjunto de intensidades de tráfico ρ_i destinado al nodo n_0 , se puede calcular las probabilidades de bloqueo B_i usando la ecuación 4.6 recursivamente: esto es, empezando en $B_0 = 0$, después $B_1 = \rho_1 + \frac{B_0}{1-B_0} \sum_{i=1}^N \rho_i = \rho_1$, y así sucesivamente. El siguiente paso es derivar la cantidad de tiempo que necesita cada paquete para llegar correctamente al destino siguiendo las dos políticas de recirculación: (1) almacenamiento-reenvío y (2) *bypass*.

4.3.2. Análisis del retardo extremo a extremo

4.3.2.1. Recirculación con almacenamiento-reenvío

En este caso, las tramas que llegan al nodo bloqueado necesitan viajar de nuevo en sentido contrario hacia el nodo anterior, donde sufren la conversión OE y esperan en la cola de reenvío junto con el tráfico local del nodo en que se encuentran. Si usamos la suposición de que el tamaño de dicha cola de espera es infinito, no se sobrepasará la capacidad de almacenamiento y todas las tramas de datos llegarán tarde o temprano al destino. Sin embargo, las tramas de datos necesitan atravesar de vuelta numerosas veces ciertos segmentos si encuentran el nodo final de los mismos bloqueado habitualmente, lo cual hace crecer ilimitadamente el retardo. Por lo tanto, es importante el estudio tanto de la esperanza media de dicho retardo como de su varianza para dar una idea más acertada no solo del retardo medio que se espera que sufran los paquetes sino también del *jitter* o variabilidad del mismo.

Bajo la ya comentada suposición de que las colas de almacenamiento electrónicas no tienen una limitación de tamaño, se puede aplicar el modelo de teoría de colas M/M/1 para calcular el retardo medio experimentado en la cola de cada nodo, que se denota como $E(Q_i)$, $i = 1, \dots, N$. Este valor se obtiene de:

$$E(Q_i) = \frac{E(S)}{1 - \rho_i} \quad (4.7)$$

donde $E(S)$ es el tiempo medio de servicio de una trama de datos de tamaño $E(B)$ bits transmitida sobre un enlace de capacidad C bits/s. En otras palabras: $E(S) = E(B)/C$. Para una trama de datos típica de 1500 bytes sobre un enlace de 10 Gbps, tiene un valor de $E(S) = 1500 \cdot 8/10^{10} =$

1,2 μs .

A continuación, el nombre de bucle j se usa para referirse a los segmentos j^+ and j^- . Por lo tanto, un paquete dado que se origina en el nodo n_l y con destino el nodo n_0 , experimenta el siguiente retardo:

$$D_l = Q_l + R_{eo} + \sum_{j=1}^l L_j + R_{oe} \quad (4.8)$$

el cual tiene en cuenta tanto el retardo de espera en cola experimentado hasta que se inyecta el paquete en el anillo como la cantidad de tiempo L_j transcurrido en cada bucle intermedio j , con $j = 0, \dots, l - 1$. Además, también hay que tener en consideración los retardos adicionales por las conversiones EO y OE (R_{eo} y R_{oe} en la ec. 4.8) sufridos por el paquete al entrar y salir del anillo respectivamente según el modelo de funcionamiento descrito.

A continuación, considérese sólo el bucle j : si el paquete encuentra el nodo siguiente n_j libre (lo cual ocurrirá con una probabilidad de $1 - B_j$), entonces, el paquete sólo sufre el retardo debido a la propagación proporcional a la longitud del segmento R_j . En caso contrario, el paquete encuentra el nodo bloqueado y es reflejado en sentido contrario hacia el nodo n_{j+1} , convertido de nuevo al dominio electrónico y almacenado en la cola de espera del nodo $j + 1$ (lo que ocurrirá con probabilidad B_j). En el siguiente intento, si el paquete encuentra esta vez el nodo n_j libre, entonces el paquete habrá experimentado un retardo total de $2R_j + Q_{j+1} + R_j$ (esto es, con probabilidad $B_j(1 - B_j)$). Si el paquete se transmite con éxito en el tercer intento (con probabilidad $B_j^2(1 - B_j)$), entonces habrá sufrido un retardo total de $4R_j + 2Q_{j+1} + R_j$, y así sucesivamente. Siguiendo este razonamiento, se puede inferir que el retardo total de un paquete para atravesar el bucle j viene dado por la siguiente ecuación:

$$L_j = R_j + \sum_{k=0}^m [2R_j + R_{oe} + Q_{j+1} + R_{eo}] \quad (4.9)$$

donde m es una variable aleatoria con una distribución geométrica de parámetro B_j . En definitiva, m indica el número de intentos extra que ha invertido un paquete dado en atravesar con éxito el bucle j , caracterizado por la probabilidad B_j de encontrar bloqueado el nodo siguiente n_j en su

camino extremo a extremo. Dicho de otro modo:

$$P(m = k) = (1 - B_j)B_j^k, \quad k = 0, 1, \dots \quad (4.10)$$

Es importante destacar de la ec. 4.9 que los paquetes deben sufrir el retardo de conversión OE and EO siempre que encuentren bloqueado el siguiente nodo, ya que deben ser almacenados en la memoria electrónica del nodo previo en el sentido de transmisión del anillo.

El retardo medio para el bucle L_j viene dado por:

$$\begin{aligned} E(L_j) &= E(R_j) + E(m)E(2R_j + R_{oe} + Q_{j+1} + R_{eo}) \\ &= E(R_j) + \frac{B_j}{1 - B_j}(2E(R_j) + E(R_{oe}) + E(Q_{j+1}) + E(R_{eo})) \\ &= R_j \frac{1 + B_j}{1 - B_j} + (R_{oe} + E(Q_{j+1}) + R_{eo}) \frac{B_j}{1 - B_j} \end{aligned} \quad (4.11)$$

donde $E(R_j) = R_j$ ya que el retardo de propagación será un valor fijo determinado por la longitud del segmento de fibra y la velocidad de propagación de la luz en dicha fibra (típicamente $2 \cdot 10^8 m/s$). Lo mismo se aplica a los retardos de conversión OE y EO los cuales son también cantidades de tiempo fijas dependientes de la tecnología del nodo, esto es, $E(R_{oe}) = R_{oe}$ y $E(R_{eo}) = R_{eo}$. La varianza del retardo en el bucle $Var(L_j)$ viene dada por¹:

$$\begin{aligned} Var(L_j) &= Var(R_j) + E(m)Var(2R_j + R_{oe} + Q_{j+1} + R_{eo}) \\ &\quad + (E(2R_j + R_{oe} + Q_{j+1} + R_{eo}))^2 Var(m) \\ &= \frac{B_j}{1 - B_j} Var(Q_{j+1}) + \frac{B_j}{(1 - B_j)^2} (2R_j + R_{oe} + E(Q_{j+1}) + R_{eo})^2 \end{aligned} \quad (4.12)$$

donde $Var(R_j) = 0$ debido a que el retardo de propagación será un valor fijo. El mismo razonamiento puede aplicarse a $Var(R_{oe}) = Var(R_{eo}) = 0$. Además, $Var(Q_{j+1}) = \left(\frac{E(S)}{1 - \rho_j}\right)^2$.

¹Es importante destacar que la varianza de una suma aleatoria de N variables aleatorias independientes entre sí X_i , $i = 1, 2, \dots$ se puede calcular como: $Var(\sum_{i=1}^N X_i) = E(N)Var(X_1) + E^2(X_1)Var(N)$

Finalmente, para el retardo extremo a extremo que sufre un paquete originado en el nodo n_l se obtienen las siguientes expresiones para su media y varianza respectivamente:

$$E(D_l) = R_{eo} + \sum_{j=0}^{l-1} \left[R_j \frac{1+B_j}{1-B_j} + (R_{eo} + E(Q_{j+1}) + R_{oe}) \frac{B_j}{1-B_j} \right] + R_{oe} + \frac{E(S)}{1-\rho_l} \quad (4.13)$$

$$Var(D_l) = \sum_{j=0}^{l-1} \left[\frac{B_j}{1-B_j} Var(Q_{j+1}) + \frac{B_j}{(1-B_j)^2} (2R_j + R_{eo} + E(Q_{j+1}) + R_{oe})^2 \right] + \left(\frac{E(S)}{1-\rho_l} \right)^2 \quad (4.14)$$

bajo la suposición de que los nodos actúan de forma independiente entre sí.

Como se ha descrito, a niveles de carga bajos, un paquete dado originado en el nodo l debería de experimentar idealmente un retardo medio de $E(Q_l) + \sum_{j=0}^{l-1} R_j$.

4.3.2.2. Recirculación directa o *bypass*

En este caso, la misma ecuación 4.8 descrita anteriormente puede aplicarse:

$$D_l = Q_l + R_{eo} + \sum_{j=1}^l L_j + R_{oe} \quad (4.15)$$

Sin embargo, el cálculo de Q_l y L_j es diferente. Lo primero de todo, el retardo que experimenta el paquete en el bucle j -ésimo, L_j , no incluye el retardo en cola, dado que en este caso, el tráfico que viene en sentido contrario no se almacena en la cola de espera y no sufre retardos de conversión OEO. Por lo tanto:

$$L_j = R_j + \sum_{k=0}^m 2R_j \quad (4.16)$$

Adicionalmente, el tiempo medio de espera en cola de los paquetes antes de entrar en el anillo

es también diferente que en el caso anterior:

$$E(Q_i) = \frac{E(S)}{(1 - \rho_i)(1 - B_{i-1})} \quad (4.17)$$

dado que la capacidad observada por la cola en el nodo n_i es sólo el remanente que deja libre el tráfico que viene recirculado desde el siguiente nodo, esto es, $C(1 - B_{i-1})$.

Por lo tanto, el retardo medio y su varianza, que sufren los paquetes en el bucle L_j viene dado por las expresiones:

$$\begin{aligned} E(L_j) &= E(R_j) + E(m)E(2R_j) = E(R_j) + \frac{B_j}{1 - B_j} 2E(R_j) \\ &= R_j \frac{1 + B_j}{1 - B_j} \end{aligned} \quad (4.18)$$

y

$$\begin{aligned} Var(L_j) &= Var(R_j) + E(m)Var(2R_j) + (E(2R_j))^2 Var(m) \\ &= \frac{B_j}{(1 - B_j)^2} (2R_j)^2 \end{aligned} \quad (4.19)$$

donde $Var(R_j) = 0$ ya que el retardo de propagación sigue siendo un valor fijo y se considera que $Var(Q_{j+1}) = \left(\frac{E(S)}{1 - \rho_j}\right)^2$.

Finalmente, los paquetes sufren un retardo extremo a extremo con la estrategia de "bypass" cuya media y varianza vienen definidas por las siguientes expresiones:

$$E(D_l) = \frac{E(S)}{(1 - \rho_l)(1 - B_{l-1})} + R_{eo} + \sum_{j=0}^{l-1} \left[R_j \frac{1 + B_j}{1 - B_j} \right] + R_{oe} \quad (4.20)$$

$$Var(D_l) = \left(\frac{E(S)}{(1 - \rho_l)(1 - B_{l-1})} \right)^2 + \sum_{j=0}^{l-1} \left[4 \frac{B_j}{(1 - B_j)^2} R_j^2 \right] \quad (4.21)$$

de nuevo bajo a suposición de independencia entre los nodos.

4.4. Ejemplos numéricos

Esta sección trata de validar las ecuaciones analíticas presentadas en la sección anterior mediante los resultados obtenidos en escenarios simulados para determinar qué aspectos del modelo tienen mayor impacto sobre el retardo y el *jitter* en el anillo. El escenario se compone de un anillo de 60 km de longitud con $N + 1$ etiquetados desde n_0 a n_N con el nodo n_0 actuando como destino del tráfico ofrecido al anillo por el resto de nodos. El tamaño de los paquetes inyectados sigue una distribución exponencial de media $E(B) = 1250$ bytes, lo que genera un tiempo de servicio $E(S) = \frac{1250 \times 8}{10 \cdot 10^9} = 1 \mu s$ en un interfaz óptico de salida de $C = 10$ Gbps interfaz. Adicionalmente, se asume que los nodos agregan tráfico que proviene de múltiples usuarios, lo que permite modelar que el perfil de tráfico inyectado al anillo pueda modelarse como un proceso de Poisson con una tasa $\lambda = \rho/E(S)$, siendo ρ la carga del tráfico de entrada. Finalmente, ambos retardos de conversión OE y EO se asume que son de $R_{oe} = R_{eo} = 25 \mu s$.

Para dar un ejemplo numérico en la estrategia de acceso con almacenamiento y reenvío utilizaremos un anillo de $N + 1 = 8$ nodos equiespaciados en los 60 km de longitud del anillo. Además, todos los nodos operan con el mismo patrón de tráfico de entrada con una carga de $\rho = 0,05$. De esta forma, el retardo en cola experimentado por un paquete antes de entrar al anillo será en media de:

$$E(Q) = \frac{1 \mu s}{1 - 0,05} = 1,0526 \mu s$$

y sufre un retardo de propagación por segmento de:

$$R = \frac{60/8 \text{ km}}{2 \cdot 10^5 \text{ km/s}} = 37,5 \mu s$$

El retardo total depende del número de segmentos que el paquete en cuestión debe atravesar para llegar al nodo destino (dependiendo de la posición relativa del nodo fuente en el anillo) y del número de veces que el paquete ha circulado de vuelta en cada segmento (lo cual depende de las probabilidades de bloqueo) antes de llegar a su destino. También es importante notar que el retardo de conversión R_{oe} y R_{eo} pueden constituir una parte importante del mismo.

4.4.1. Análisis y validación mediante simulación

En esta sección se explora el valor del retardo medio y su variabilidad en diferentes topologías bajo distribuciones de carga uniformes y no uniformes entre los nodos origen.

El primero de los experimentos permite validar la ecuación del retardo analítica derivada en la sección 4.3.2 con el modelo de simulación. Dicha simulación se ha realizado mediante un modelo de bloques de la plataforma Simulink del entorno de programación y análisis Matlab, que simula un anillo de 60 km de longitud con $N + 1 = 8$ nodos con diferentes cargas de tráfico de entrada, y asumiendo los valores de capacidad $C = 10$ Gbps, tamaño medio de paquete $E(B) = 1250$ bytes y $R_{oe} = R_{eo} = 25 \mu s$, igual que los descritos en la sección 6.5. Es importante señalar que la máxima carga permitida en cada nodo es de $\rho_{max} = 1/(N + 1)$, para que de ese modo la carga máxima ofrecida en el último segmento no supere la capacidad del enlace y no haya pérdidas por saturación del nodo receptor.

Fig. 4.3 muestra tanto el resultado analítico como los obtenidos por simulación en el anillo descrito, usando ambas políticas para arbitrar el tráfico bloqueado en los nodos de tránsito: recirculación con almacenamiento y reenvío (figuras de la parte superior) y con reenvío directo o "bypass" (gráficas de la parte inferior de la figura). En la figura se muestran el retardo experimentado por el primero y el último nodo en el anillo junto con el que ocupa la posición central en cuanto a distancia al nodo de destino. Puede verse que tanto los valores analíticos como los obtenidos mediante simulación se ajustan entre sí en ambas métricas: el retardo medio (puntos) como en la desviación estándar (barras de error), a diferentes cargas de tráfico de entrada. El lector de este trabajo ha de tener en cuenta que las barras de error representan desviación típica del retardo (no intervalos de confianza) en esta figura, en unidades de tiempo. En este caso, la máxima carga de entrada por nodo es de $\rho_{max} = 1/(N + 1) = 0,125$.

4.4.2. Análisis del retardo para diferentes valores de N

Los siguientes experimentos tienen en cuenta sólo los valores analíticos para mostrar los beneficios y defectos de las dos políticas de control de acceso al medio (con "bypass" y con almacenamiento-reenvío) en diferentes escenarios. Estos experimentos proporcionan una comparativa entre las dos políticas cuando varía el número de nodos del anillo. En todo los casos se mantiene la longitud del anillo fija en 60 Km, de forma que el retardo total mínimo por propagación es el mismo en todos los casos. También se mantienen constantes los valores del tamaño

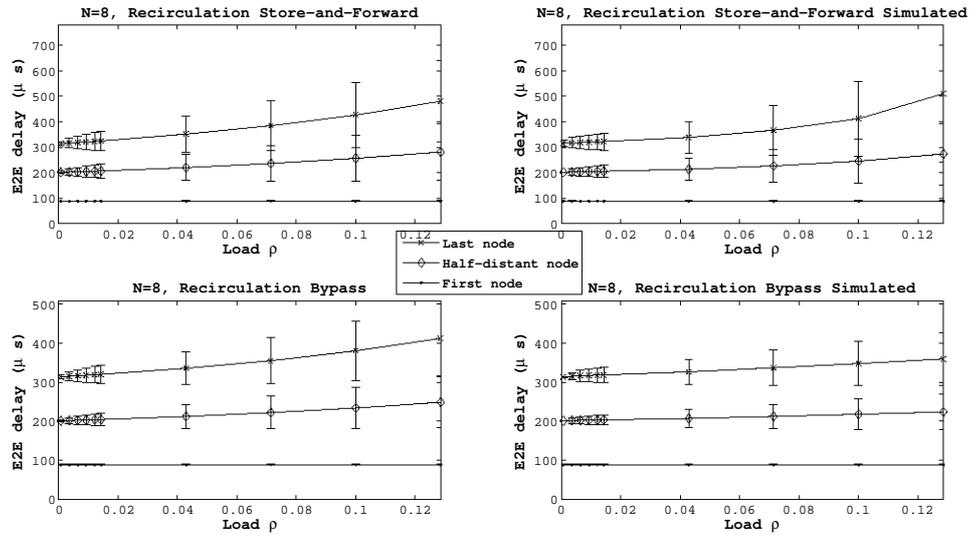


Figura 4.3: Análisis de validación con Simulink de Matlab para una red en anillo de $N = 8$ nodos. Resultados analíticos (a la izquierda) y simulados (derecha) se ha obtenido suponiendo recirculación con "almacenamiento-reenvío" (arriba) y "bypass"(abajo).

medio de trama de $E(B) = 1250$ bytes, la capacidad del enlace $C = 10$ Gbps y los retardos de conversión OEO de $R_{oe} = R_{eo} = 25 \mu s$, como en los experimentos de secciones anteriores.

Se han considerado dos casos de estudio: (a) $N + 1 = 8$ y (b) $N + 1 = 32$. En el primer caso, el retardo de propagación por segmento es $R = \frac{60 \cdot 10^3 / 8}{2 \cdot 10^8} = 37,5 \mu s$ mientras que en el segundo caso es de $R = \frac{60 \cdot 10^3 / 32}{2 \cdot 10^8} = 9,37 \mu s$. El retardo medio y la varianza se muestran en la Fig. 4.4 para los dos casos a niveles de carga distintos. Por claridad, de nuevo se muestran sólo los valores correspondientes al primero y el último nodos y los del que ocupa la posición central.

Tal y como se muestra en la figura 4.4, existe una gran diferencia de comportamiento (en términos relativos a la media y desviación típica del retardo) entre las dos políticas de gestión del tráfico bloqueado, especialmente cuando el sistema trabaja a altas cargas. En todos lo casos, la recirculación directa (mediante "bypass" muestra valores de retardo medio y varianza menores debido a que los paquetes no tiene que sufrir conversión al dominio electrónico y espera en cola en los nodos intermedios.

De estos resultados podemos extraer además las siguientes conclusiones:

- El primer nodo no experimentará grandes diferencias de comportamiento en términos de retardo medio y varianza en ningún caso. Eso es debido a que el primer nodo percibe un tratamiento de "alta prioridad" con respecto a los nodos más alejados, por lo que sólo sufre

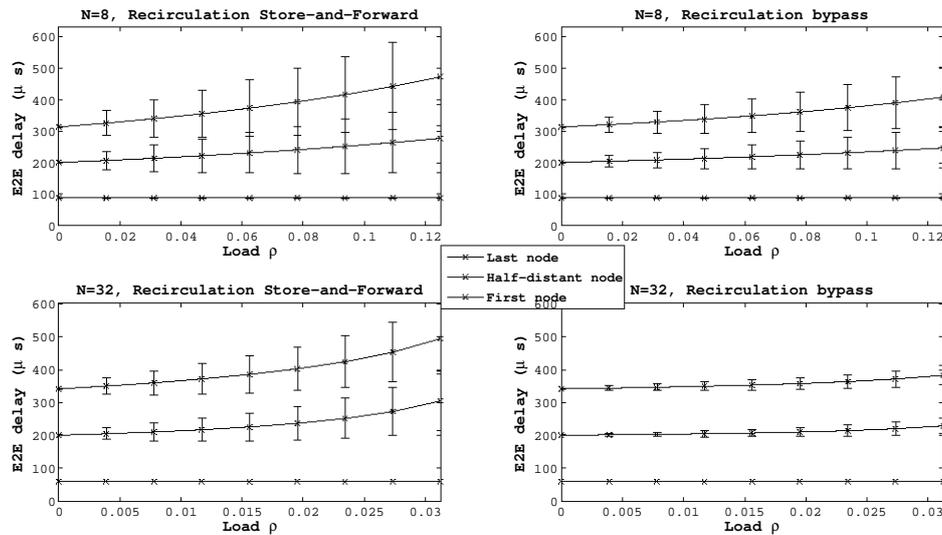


Figura 4.4: Valor medio y varianza del retardo que experimentan los paquete usando ambas políticas para una cantidad de nodos $N + 1$ variable

un retardo de $R + R_{oe} + R_{eo}$ en la mayor parte de los casos.

- A cargas bajas, no hay grandes diferencias entre el retardo medio experimentado por los paquetes en ambas políticas pero sí en la varianza ("jitter").
- Cuando se usa la técnica de recirculación mediante "bypass", el número de nodos (8 ó 32 nodos) no marca grandes diferencias en el retardo medio de atravesar el anillo de 60 Km de longitud fija, pero sí que hay grandes diferencias en la variabilidad de dicho retardo. Esta conclusión es muy importante ya que habrá de tenerse en consideración a la hora de incrementar el numero de nodos intermedios añadidos a un anillo donde la longitud del mismo no cambie.
- La situación opuesta ocurre si se utiliza la técnica de "almacenamiento y reenvío" ya que muestra un comportamiento prácticamente constante en términos de la variabilidad del retardo cuando aumenta el número de nodos en el anillo de 8 a 32, y sí que aparecen diferencias en relación con el retardo medio.

4.4.3. Análisis del retardo sufrido cuando un nodo presenta un pico de carga

La Fig. 4.5 muestra el retardo medio y la varianza observados en cada nodo de un anillo suponiendo que un determinado nodo del anillo presenta un pico de carga y por tanto consume un exceso de recursos de red respecto del resto de nodos en la topología. Para este experimento, todos

los nodos se presuponen ligeramente cargados ($\rho = 0,0312$) excepto un nodo en el cual el tráfico transmitido tiene una carga relativa de $\rho_{\text{peak}} = 0,75$. Puede verse cómo, cuando esto ocurre, el efecto que se observa es que el nodo más cargado bloquea el acceso de los nodos precedentes en el anillo, lo que incrementa de forma importante su respectivo retardo medio y especialmente su varianza.

Esto genera una situación particularmente dañina en cuanto al retardo medio sufrido si se está utilizando la política de "almacenamiento y reenvío", donde los nodos previos al nodo más cargado han de realizar múltiples operaciones de conversión OEO y de almacenamiento en sus colas de espera. En cambio, si se utiliza el sistema de "bypass" el efecto negativo afecta considerablemente a la variabilidad del retardo, en mayor medida que al valor absoluto de retardo medio.

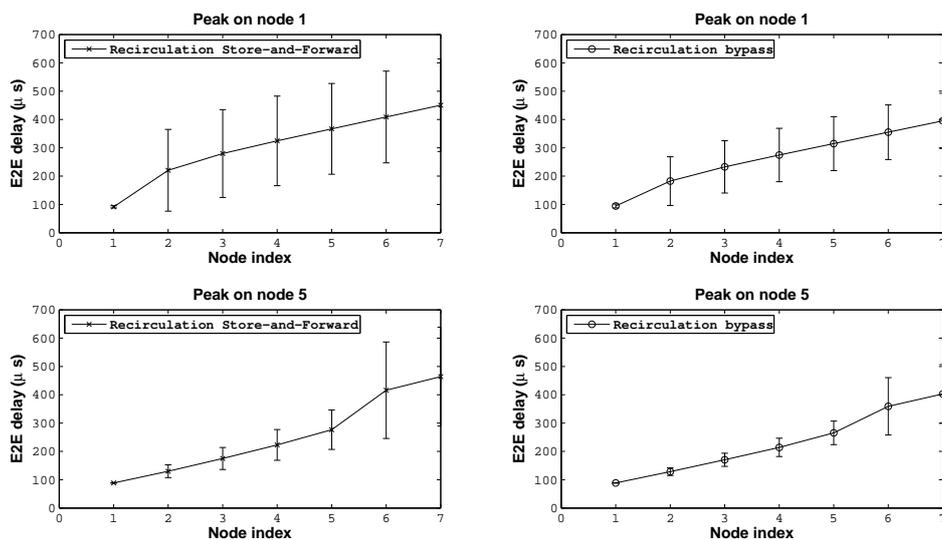


Figura 4.5: Retardo medio y varianza en cada nodo suponiendo que el nodo con el pico de tráfico es el n_1 (arriba) y el nodo n_6 (abajo) en un anillo WDM de 8 nodos

4.5. Resumen y conclusiones

Este capítulo presenta la contribución realizada dentro del marco de este trabajo de tesis doctoral, en la propuesta de un protocolo de control de acceso al medio (MAC) para evitar colisiones en el acceso al canal compartido en anillos ópticos WDM basados en TT-FR. La novedad introducida con esta propuesta de protocolo MAC es que el tráfico bloqueado es devuelto en el sentido opuesto al de transmisión en el anillo hacia los nodos precedentes, donde puede volverse a enviar

directamente (recirculación con reenvío directo o *"bypass"*) o almacenarse junto con los datos a transmitir por el anillo en una cola de espera para su posterior retransmisión (recirculación con *"almacenamiento y reenvío"*).

En ambos casos, se presenta un estudio analítico de los valores de retardo medio y varianza del mismo que sufren los paquetes transmitidos hacia un destino común por el resto de los nodos del anillo. Los experimentos realizados revelan un cambio en el equilibrio de prioridades de los nodos, siendo en ambos casos los más beneficiados los nodos más cercanos al destino y no al contrario como sucede en las soluciones habituales de anillos WDM, lo que permitiría balancear las probabilidades de bloqueo según sea conveniente.

En cuanto a ambas políticas de reenvío de tráfico bloqueado presentadas, los resultados revelan que la técnica de *"almacenamiento y reenvío"* introduce mayor retardo medio en los paquetes que la técnica de *"bypass"*, especialmente a cargas elevadas, y, lo que es más significativo, una tremenda variabilidad de dicho retardo, lo cual afecta negativamente al *"jitter"* que sufren los servicios, que utilizan el anillo rebajando los niveles de calidad de servicio (QoS) garantizables, lo cual es especialmente dañino para servicios de tiempo real. Mejoras futuras sobre el modelo y las políticas de acceso permitirían estudiar su efecto concreto sobre requerimientos estrictos de calidad de servicio.

Vivimos en una sociedad
profundamente dependiente de la
ciencia y la tecnología en la que
nadie sabe nada de estos temas.
Ello constituye una fórmula segura
para el desastre.

Carl Sagan

Capítulo 5

Algoritmo de encaminamiento para eficiencia energética en redes Ethernet

5.1. Introducción

La reducción del consumo de energía en los equipos de comunicaciones es un tema de investigación candente que está relacionado con la posibilidad de conseguir beneficios tanto económicos como medioambientales. Tal y como se desprende de estudios previos [100], la cantidad de energía consumida por las tecnologías de la información (IT) ronda los 800 Twh por año), de los cuales se podría conseguir un ahorro importante mediante el diseño de software y hardware más eficiente. Un ejemplo concreto sería el ahorro que supondría el seleccionar la velocidad más adecuada para un enlace de comunicaciones, como la solución propuesta en [43, 101], o conseguir el ahorro energético mediante el apagado del equipo innecesario en determinadas ocasiones [102–104].

En Ethernet existe la posibilidad de conseguir mejorar la eficiencia energética debido a que por un lado, la circuitería electrónica de las interfaces Ethernet siguen consumiendo cerca del 100 % de su potencia, mientras están desocupadas, y por otro lado, Ethernet está ampliamente desplegado y es comúnmente utilizado en todo tipo de entornos de red. De hecho, Ethernet es la tecnología presente en la mayoría de las redes de área local tanto corporativas como de uso personal o residencial, y está expandiéndose su uso a las redes metropolitanas e incluso a algunas soluciones de red troncal gracias a los últimos avances en estándares que ofrecen 40/100 Gbps en Ethernet, como se ha visto en algunas de las soluciones comentadas en secciones anteriores de este trabajo de tesis doctoral. Por lo tanto, cualquier mejora de diseño relacionada con la optimización

del gasto energético en Ethernet se traduciría en un gran ahorro a nivel global dada su actual ubicuidad de despliegue.

En relación con esto, el grupo de trabajo de IEEE802.3 [46] aprobó recientemente el estándar 802.3 de Eficiencia Energética en Ethernet (EEE) que, en resumen, intenta conseguir el ahorro de una gran parte del consumo de energía mediante la definición de dos modos de funcionamiento: activo o dormido. La idea básica es la de poner la capa física de Ethernet en un estado de bajo consumo (o *sleep*) cuando no haya datos pendientes de ser transmitidos. Este modo de bajo consumo gasta del orden del 10 % de la potencia total que se consume en el modo activo habitual. De este modo, los enlaces menos cargados deberían beneficiarse de este modo especial de funcionamiento ya que estarán largos períodos de tiempo en dicho modo de bajo consumo.

Aunque se esperaban a priori grandes ahorros de consumo gracias a la definición de estos dos modos de funcionamiento, estudios previos demuestran que esto no ocurre debido a los largos tiempos de conmutación de la tecnología entre el modo activo y el de bajo consumo. De hecho, la curva de consumo en EEE se desvía significativamente del modelo de consumo lineal proporcional a la carga, que estaría considerado como el perfil de consumo de potencia ideal [105]. La razón de estos resultados operacionales subóptimos radica en el hecho de que en EEE la transición entre el modo activo y el de bajo consumo es muy lenta y desperdicia a su vez una cantidad considerable de energía sólo en realizar esta conmutación entre modos. Se puede describir este hecho con valores concretos típicos. Por ejemplo, la transmisión de una trama de 1500 bytes en una red Ethernet 10G Base-T (10Gbps) tiene el siguiente patrón de eficiencia energética η :

$$\eta = \frac{T_{\text{frame}}}{T_w + T_{\text{frame}} + T_s} = \frac{1,2\mu s}{8,56\mu s} = 0,14 \quad (5.1)$$

ya que los intervalos de $4,48\mu s$ y $2,88\mu s$ del total de $8,56\mu s$ se dedican a la transición de activar el enlace o ponerlo en bajo consumo, respectivamente, mientras que sólo $1,2\mu s$ se dedican a la transmisión de la trama de datos. Esto es, el 86 % de la potencia se ha consumido sólo en la conmutación entre los dos modos de funcionamiento de la capa física. Esto hace que las gráficas de consumo de energía en función de la carga del enlace tengan la apariencia que se muestra en la Fig. 5.1 para el caso de interfaces Ethernet 10G BASE-T según [105]. Cabe resaltar que 100BASE-TX, 1000BASE-T and 10GBASE-T son estándares definidos para redes de área local con longitudes máximas de enlace del orden de 100 metros.

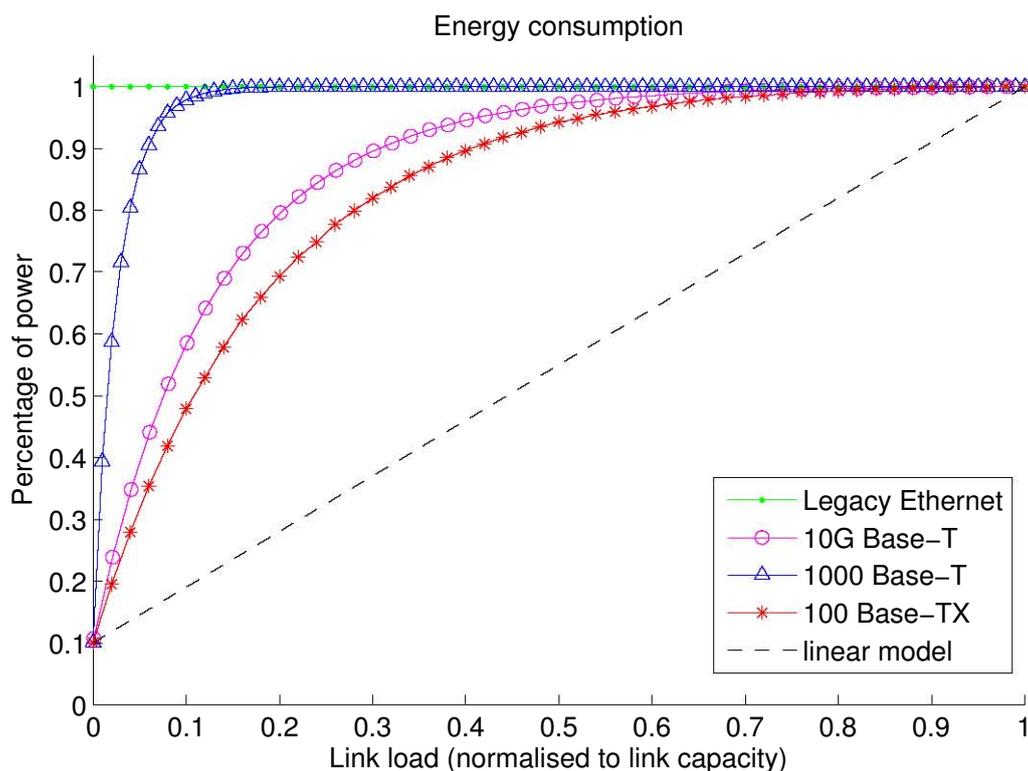


Figura 5.1: Consumo energético frente a carga del enlace para BASE-TX, 1000BASE-T y 10GBASE-T EEE, según el estudio realizado en [96]

Tal y como se muestra en la figura, la energía que se consume en los enlaces EEE está cercana al 100 % para cargas de tráfico en el enlace del orden del 10 % de su capacidad en el caso de 1000 BASE-T, del 55 % en 10G Base-T y del 45 % en 100 Base-TX. Sin embargo, se puede sacar una conclusión constructiva de estos resultados mostrados en la Fig. 5.1: existen dos rangos de porcentaje de carga del enlace en los que hay poca diferencia entre la curva de consumo frente a carga en EEE y la curva lineal proporcional: en situaciones de carga de tráfico muy bajas (por debajo del 5 %) y a cargas muy altas (del 70 % en adelante).

Para sacar partido de estos resultados a priori, se propone en este trabajo de tesis doctoral la definición de un algoritmo de encaminamiento de flujos basado en ahorro energético que se denominará EEE-FA (Energy Efficient Ethernet - Flow Allocation) en adelante. El objetivo de este algoritmo es calcular la mejor ruta posible en términos de aumento en el consumo de potencia, para cada nueva demanda de tráfico que llega a una red dada. El algoritmo EEE-FA funciona de la siguiente manera:

1. En primer lugar calcula los K posibles caminos más cortos (K-SPR - K-Shortest Path Routing) extremo a extremo para un determinado flujo en cuanto a número de saltos.

2. Una vez hecho esto, evalúa cual sería el impacto en la energía total consumida por la red para cada una de las posibilidades, teniendo en cuenta la carga de los enlaces debida a las demandas que ya han sido encaminadas previamente.
3. A continuación, selecciona aquella ruta que suponga el mínimo aumento de la métrica elegida: consumo de potencia al global de la red.
4. La nueva demanda se encamina por esta ruta y se actualizan los datos de carga y consumo de cada enlace para las próximas iteraciones del algoritmo.
5. En el caso de que llegue un nuevo flujo de tráfico a un nodo de la red se repiten los pasos del algoritmo.

Las prestaciones del algoritmo han sido comparadas con las que ofrece un algoritmo de encaminamiento clásico basado en el camino más corto en cuanto a número de saltos (SPR - Shortest Path Routing). Este algoritmo equivale a la situación de calcular el algoritmo EEE-FA usando K-SPR para un valor de $K=1$, esto es, tener en cuenta sólo el camino de menor número de saltos sin ninguna posible alternativa. Es importante recalcar, que el valor de K influye en la rapidez del algoritmo y en las probabilidades de encontrar la solución óptima: un valor de K muy elevado es el caso de un problema NP-completo de búsqueda de la solución óptima para la nueva demanda a costa de lentitud y uso de más potencia de cálculo, mientras que un valor de K pequeño supone una menor probabilidad de obtener la solución óptima pero una mayor velocidad de convergencia del algoritmo.

Los resultados de simulación para EEE-FA con K-SPR presentan beneficios en el ahorro de energía frente a SPR, sin embargo, el ahorro se consigue a costa de un aumento en el número medio de saltos por demanda, por no usarse siempre el camino de menor distancia, y un ligero aumento de la carga global de la red, como consecuencia de usarse caminos subóptimos (en términos de distancia medida en número de saltos) para encaminar las demandas que llegan.

Una idea similar fue propuesta en [106], pero con importantes diferencias:

- En primer lugar, el algoritmo EEE-FA calcula la mejor ruta posible en términos de su impacto en el consumo energético global basándose en llegadas independientes de nuevas demandas de capacidad al sistema y en instantes independientes entre sí. Sin embargo, el trabajo publicado en [106] formula un problema de optimización del encaminamiento, donde la función de consumo energético objetivo, necesita ser minimizada de acuerdo a un conjunto

de restricciones globales. Un algoritmo centralizado de este tipo, se resuelve para una matriz de tráfico dada que especifica todas las demandas de capacidad que va a tener la red, y por tanto, ha de conocerse de antemano, lo que no es posible en un gran número de casos.

- Y en segundo lugar, el algoritmo EEE-FA que se ha estudiado como parte de este trabajo de tesis doctoral, utiliza las curvas de consumo de energía frente a carga del enlace reales, en función de la tecnología de red utilizada, resultantes de aplicar EEE a distribuciones de tráfico de Poisson. En cambio, el trabajo de [106] utiliza un conjunto de curvas matemáticas teóricas para representar las curvas de consumo de potencia frente a carga, que no se corresponden con la realidad de EEE.

En las siguientes secciones de este capítulo se presenta la definición del algoritmo EEE-FA y se demuestra mediante un conjunto de redes de ejemplo su funcionamiento y sus prestaciones, así como los resultados obtenidos mediante la simulación de las mismas y las conclusiones que pueden sacarse de dichos resultados.

5.2. Descripción del funcionamiento de EEE-FA

5.2.1. Ejemplo de funcionamiento de EEE-FA en una red de tres nodos

Consideremos una red de tres nodos como la de la Fig. 5.2, interconectados entre sí mediante una topología en forma de triángulo, para describir un ejemplo de funcionamiento de EEE-FA, aplicando la métrica de EEE a la asignación de rutas para las demandas de capacidad que llegan a la red. Los tres enlaces, que interconectan los nodos etiquetados como N_1 , N_2 y N_3 , son todos de capacidad normalizada a C . Para el ejemplo, se asume que llegan tres demandas al sistema que han de ser encaminadas a través de la red triangular, en el orden siguiente:

- Primero, una demanda $d_{13}^{(1)} = 0,5C$ del nodo 1 al nodo 3 que requiere la reserva del 50 % de la capacidad del enlace.
- Una segunda demanda $d_{32}^{(2)} = 0,5C$, del nodo 3 al 2, también del 50 % de la capacidad del enlace.
- Y finalmente, una tercer flujo $d_{12}^{(3)} = 0,2C$, del nodo 1 al 2, que demanda el 20 % de la capacidad del enlace.

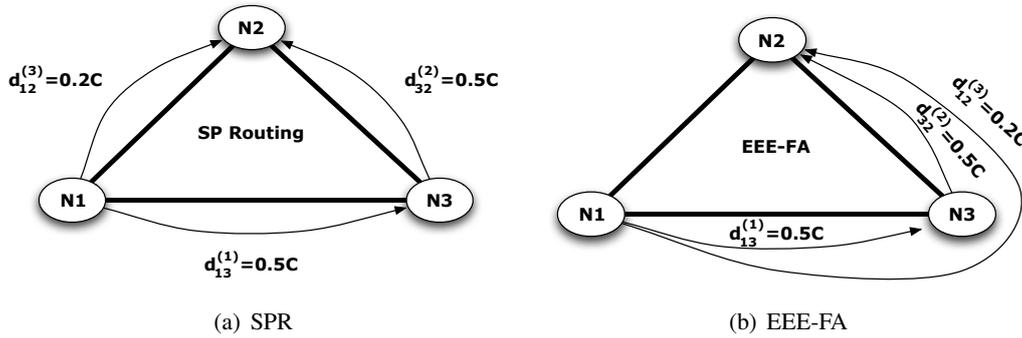


Figura 5.2: Ejemplo de encaminamiento de 3 demandas sobre una red de tres nodos utilizando dos algoritmos con objetivos de coste distinto: (a) Shortest-Path Routing (SPR), (b) Energy Efficient Ethernet Flow Allocation (EEE-FA)

Si se usa como métrica el camino más corto (Shortest Path Routing) para encaminar las demandas, el resultado es tal y como se ve en la Fig. 5.2(a) (líneas punteadas), haciendo uso de los tres enlaces directos entre los nodos. En este caso, la carga media por enlace en la red sería:

$$E[\rho_l] = \frac{0,5C + 0,5C + 0,2C}{3} = 0,4C \quad (5.2)$$

mientras que el consumo medio de energía por enlace (normalizado), asumiendo que la red utiliza por ejemplo una infraestructura de enlace 100BASE-TX (Fig. 5.1) sería:

$$E[P] = \frac{0,95 + 0,95 + 0,7}{3} = 0,86 \quad (5.3)$$

y el número medio de saltos por demanda quedaría como:

$$E[H] = \frac{1 + 1 + 1}{3} = 1 \quad (5.4)$$

Sin embargo, la tercera demanda podría ser encaminada de forma alternativa, utilizando alguno de los dos enlaces que ya están activos, puesto que han sido asignados para los flujos que llegaron anteriormente, tal y como se ve en la Fig. 5.2(b), incrementando la carga media por enlace y el número medio de saltos por demanda a:

$$E[\rho_l] = \frac{0,7C + 0,7C + 0}{3} = 0,47C \quad \text{y} \quad E[H] = \frac{1 + 1 + 2}{3} = 1,33 \quad (5.5)$$

respectivamente, pero reduciendo el consumo de energía a:

$$E[P] = \frac{0,98 + 0,98 + 0,1}{3} = 0,69 \quad (5.6)$$

Encaminar la tercera demanda de tráfico a través de los dos enlaces que ya están activos permite ahorrar energía en general, ya que esta estrategia permite mantener un enlace sin tráfico, y por tanto su consumo será el 10 % mínimo que requiere la tecnología. Además, los dos enlaces ya activos sufren un incremento muy ligero en la potencia que están consumiendo, debido al incremento de carga que conlleva utilizarlos para encaminar la tercera demanda (de 0.95 de consumo normalizado a 0,5C de carga pasan a 0.98 de consumo para 0,7C de carga) como puede comprobarse en la figura de consumo frente a carga (Fig. 5.1).

En la sección siguiente se define más en detalle el algoritmo EEE-FA, el cual tiene por objetivo reducir el consumo de energía en la red siguiendo las mismas premisas que en el anterior ejemplo.

5.2.2. El algoritmo EEE-FA

La descripción del algoritmo EEE-FA, resumido en la Tabla 5.1 utiliza la siguiente notación:

N el conjunto de nodos en la red.

L el conjunto de enlaces de la red.

D el número total de demandas que van a ser encaminadas.

$d_i(m, n)$ la i -ésima demanda de tráfico con origen en el nodo $m \in N$ y destino el nodo $n \in N$.

p_{ik} la lista ordenada de enlaces que atraviesa el k -ésimo camino más corto para la demanda $d_i(m, n)$ (siendo $k = 1, \dots, K$).

Ahora, para cada demanda $d_i(m, n)$, con $i = 1, \dots, D$, EEE-FA primero calcula K caminos más cortos p_{ik} , $k = 1, \dots, K$ entre los nodos m y n . Para todos los posibles caminos p_{ik} , EEE-FA evalúa el coste en términos de consumo energético de encaminar la demanda d_i por cada camino

p_{ik} con la siguiente expresión:

$$C_{ik} = \sum_{l \in p_{ik}} [f_j(\rho(l) + d_i(m, n)) - f_j(\rho(l))] \quad (5.7)$$

la cual básicamente suma el consumo de energía después de encaminar d_i sobre los enlaces especificados por el k -ésimo camino más corto p_{ik} . La expresión $f_j(\cdot)$ se refiere a las funciones de coste frente a carga mostradas en la Fig. 5.1, para 100BASE-TX, 1000BASE-T y 10GBASE-T. Esto es, $f_{100BASE-TX}(\rho(l))$ da el consumo de energía en el enlace $l \in L$ con carga $\rho(l)$.

Finalmente, EEE-FA selecciona el enlace ganador p_i^* que minimiza el impacto en consumo de potencia C_i^* para la demanda $d_i(m, n)$:

$$\text{Selecciona } p^* \text{ que minimiza } C_i = \min_k C_{ik} \quad (5.8)$$

En resumen, EEE-FA calcula los K caminos más cortos para las demandas de tráfico entrantes. Después de ello, calcula el impacto energético de encaminar las demandas de tráfico por cada uno de los K caminos calculados, y finalmente escoge aquel camino que supone un aumento menor de la potencia global que consume la red para las demandas de tráfico ya reservadas. La descripción de EEE-FA puede verse resumida en la Tabla 5.1.

1	Para cada nueva demanda entrante $d_i(m, n)$
2	Calcula los K caminos más cortos cuyo origen es m y su destino es n .
3	Para cada camino p_{ik} , con $k = 1, \dots, K$ para la demanda d_i
4	Calcula el impacto energético $C_{ik} = \sum_{l \in p_{ik}} [f_j(\rho(l) + d_i(m, n)) - f_j(\rho(l))]$ para todos los caminos p_{ik}
5	Escoge p^* que minimiza $C_i = \min_k C_{ik}$
6	Encamina la demanda d_i a través del camino p^* si hay suficiente capacidad libre en sus enlaces para ello. En caso contrario repite el paso 3 sin este camino
7	Actualiza el valor de $\rho(l)$ para cada enlace del camino p^*

Tabla 5.1: Algoritmo EEE-FA

5.3. Experimentos con el algoritmo EEE-FA

Esta sección ilustra el potencial del algoritmo EEE-FA en términos de ahorro energético. El algoritmo EEE-FA se compara con las prestaciones del algoritmo de encaminamiento por camino

más corto en términos de la energía media consumida por enlace, la carga media por enlace y el número medio de saltos por demanda de tráfico. Merece la pena notar que el método de encaminamiento por camino más corto (SPR en adelante), sería el mismo que EEE-FA con parámetro $K = 1$, esto es, el camino más corto al destino siempre se tiene en cuenta y se calcula su impacto en el consumo de energía.

5.3.1. Escenario de simulación

En este experimento se ha usado una topología de red de prueba consistente en 14 nodos interconectados entre sí usando 20 enlaces full-duplex, cuyos parámetros más importantes se resumen en la Tabla 5.2. Cabe destacar que el número de posibles parejas origen-destino es $N(N - 1)$, y que todos los enlaces se sume que son bidireccionales.

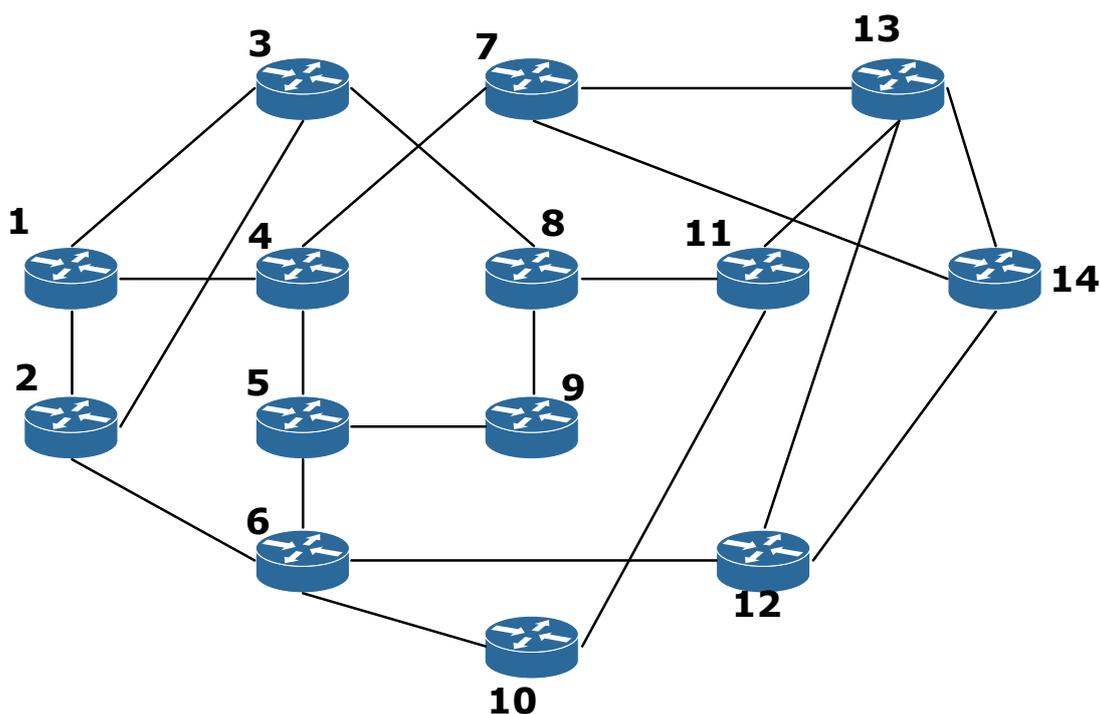


Figura 5.3: Topología de red de prueba

Como se observa en la Tabla. 5.2, se asume que las demandas de tráfico llegan a la red siguiendo un proceso de Poisson de tasa λ demandas por unidad de tiempo, y requieren una cantidad fija de $h_D = 0,1C$ unidades de capacidad por demanda y son encaminadas en un camino dado durante una cantidad de tiempo distribuida exponencialmente con un tiempo medio de $1/\mu$ unidades de tiempo. Los nodos origen y destino de cada demanda son elegidos aleatoriamente de entre el total

Parámetro	Valores de la red
Número de nodos N	14
Número de enlaces bidireccionales L	20
Carga media de la red	1 %, 5 %, 10 % 25 %, 50 % 70 %
Capacidad reservada por cada demanda de tráfico h_D	0,1C fija
Tamaño medio de la conexión $E(S)$	100 MBytes

Tabla 5.2: *Parámetros de la red de simulación*

del conjunto de nodos en la red, con la única restricción de que el mismo nodo no puede ser a la vez origen y destino de una demanda. Con esto se obtiene un número de posibles pares origen-destino de $14 \times 13 = 182$. Cada simulación tiene en consideración 1820 demandas extremo a extremo, con una media de 10 demandas o flujos por cada pareja de nodos origen-destino.

A continuación, la duración media de los flujos $1/\mu$ es calculada suponiendo un tamaño medio de la conexión de $E(S) = 100Mbytes$. En otras palabras:

$$\frac{1}{\mu} = \frac{8 \times E(S)}{h_D} \tag{5.9}$$

lo cual es $8s$ para 100BASE-TX, $0,8s$ para 1000BASE-T y $0,08s$ para 10GBASE-T.

Finalmente, con cada valor de μ , la tasa de demandas de tráfico λ se obtiene como:

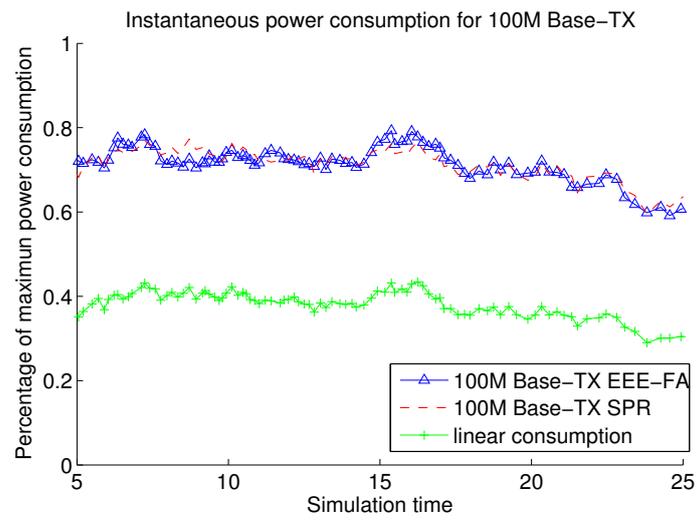
$$\lambda = \rho\mu \tag{5.10}$$

donde $\rho=0.01$ ó 0.05 (bajas cargas), 0.1 ó 0.25 (carga media) y 0.5 ó 0.7 (cargas altas).

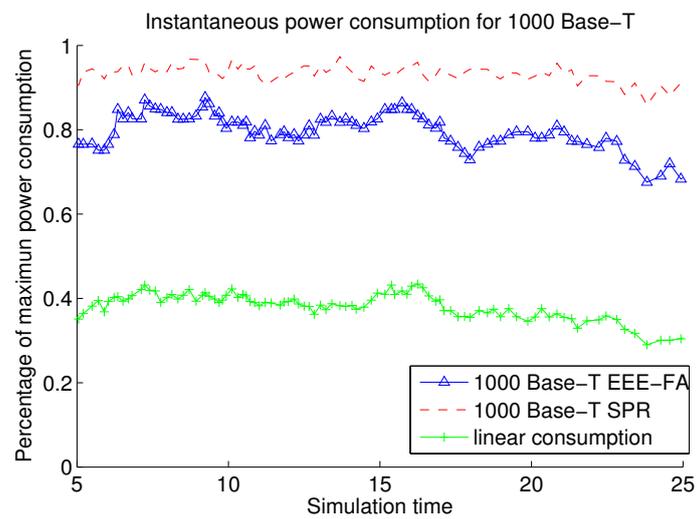
5.3.2. Resultados de los experimentos

La Fig. 5.4 muestra el consumo de potencia instantáneo en un escenario de simulación (con el mismo patrón de demandas) y para diferentes velocidades de enlace, 100BASE-TX, 1000BASE-T y 10GBASE-T, a carga de tráfico media ($\rho = 0,25$, carga media normalizada a la capacidad del enlace). Como puede observarse, se obtiene un ahorro muy considerable para 1000BASE-T, pero menor para 100BASE-TX y 10GBASE-T. Básicamente, 1000BASE-T ofrece más espacio a la mejora debido a sus altos tiempos de transición entre los modos activo y bajo consumo (EEE), lo cual resulta en unas peores prestaciones en la curva de consumo frente a carga (Fig. 5.1).

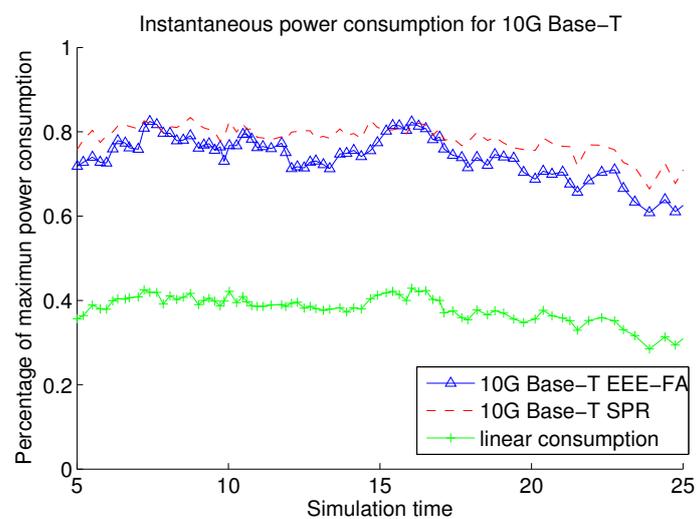
Finalmente, la Tabla 5.3 muestra un resumen de los resultados medios obtenidos para 20 simulaciones de 1820 demandas de tráfico extremo a extremo con 18 configuraciones experimentales



(a) 100BASE-TX



(b) 1000BASE-T



(c) 10GBASE-T

Figura 5.4: Consumo de potencia instantáneo para EEE-FA y SPR para diferentes capacidades de enlace

distintas, de acuerdo a:

- Algoritmos: SPR y EEE-FA.
- Capacidad de enlace: 100BASE-TX, 1000BASE-T y 10GBASE-T (100M, 1G and 10G en la tabla).
- Carga de enlace: baja ($\rho=0.01$ y $\rho=0.05$), media ($\rho=0.1$ y $\rho=0.25$) y alta ($\rho=0.5$ y $\rho=0.7$).

Para cada simulación, se ha calculado el número medio de saltos, $E(H)$, carga media por enlace $E(\rho_l)$ y consumo medio por enlace $E(P)$ a lo largo del tiempo de simulación, y finalmente se muestra la media de todas las simulación. Los resultados pueden verse resumidos en la Tabla 5.3.

Tal y como se muestra en la tabla, usando el algoritmo EEE-FA se obtiene un importante ahorro de energía a cargas de tráfico medias (25 %) y ahorros moderados a cargas bajas y altas de tráfico. Además, estas cantidades de ahorro se obtienen a costa de incrementar el número medio de saltos por demanda de tráfico, lo cual supone una clara degradación de las prestaciones globales debido al incremento de la carga total de la red y a un aumento del retardo medio por demanda. Finalmente, el porcentaje de demandas bloqueadas, esto es, aquellas que no han podido cursarse por ninguno de los posibles caminos es mayor cuando usamos EEE-FA que con SPR porque la carga de tráfico media de los enlaces es mucho mayor, lo que resulta en que algunos enlaces críticos pueden estar llenos.

En conclusión, EEE-FA debe usarse en escenarios que admiten ahorros de energía grandes, mientras que al mismo tiempo, la degradación producida por elegir caminos subóptimos (caminos más largos) no incremente la carga de la red excesivamente.

5.4. Resumen y conclusiones

Este capítulo presenta la descripción y los resultados de simulación de EEE-FA, un algoritmo de asignación de que intenta mejorar la eficiencia energética de una red Ethernet en términos de consumo de energía como función de la carga de los enlaces. Este algoritmo tiene en cuenta el incremento de energía que supone encaminar las demandas entrantes sobre distintos caminos, y selecciona aquel que supone el menor incremento global. Esto se hace en base a la llegada de mandas de tráfico a la red, lo cual facilita su implementación en redes con llegadas de flujos aleatorias.

(a) Cargas bajas

		$\rho = 0,01$			$\rho = 0,05$		
		100M	1G	10G	100M	1G	10G
Núm. medio de saltos	SPR	2,20	2,20	2,20	2,20	2,20	2,20
	EEE-FA	2,21	2,30	2,25	2,21	2,30	2,25
Carga media por enlace	SPR	1,093 %	1,093 %	1,093 %	1,115 %	1,115 %	1,115 %
	EEE-FA	1,103 %	1,151 %	1,124 %	1,124 %	1,173 %	1,145 %
Consumo medio	SPR	12,66 %	17,28 %	14,20 %	12,76 %	17,46 %	14,32 %
	EEE-FA	12,64 %	17,13 %	14,19 %	12,74 %	17,31 %	14,30 %
Demandas bloqueadas	SPR	0	0	0	0	0	0
	EEE-FA	0	0	0	0	0	0

(b) Cargas medias

		$\rho = 0,10$			$\rho = 0,25$		
		100M	1G	10G	100M	1G	10G
Núm. medio de saltos	SPR	2,20	2,20	2,20	2,20	2,20	2,20
	EEE-FA	2,21	2,30	2,25	3,35	3,31	3,48
Carga media por enlace	SPR	1,123 %	1,123 %	1,123 %	34,67 %	34,67 %	34,67 %
	EEE-FA	1,132 %	1,183 %	1,153 %	53,76 %	52,71 %	54,96 %
Consumo medio	SPR	12,76 %	17,50 %	14,32 %	61,11 %	81,29 %	67,32 %
	EEE-FA	12,64 %	17,35 %	14,30 %	61,06 %	68,27 %	63,67 %
Demandas bloqueadas	SPR	0	0	0	0	0	0
	EEE-FA	0	0	0	0,073 %	0,037 %	0,055 %

(c) Cargas Altas

		$\rho = 0,50$			$\rho = 0,70$		
		100M	1G	10G	100M	1G	10G
Núm. medio de saltos	SPR	2,22	2,22	2,22	2,35	2,35	2,35
	EEE-FA	3,20	2,98	3,27	2,97	2,74	3,02
Carga media por enlace	SPR	56,22 %	56,22 %	56,22 %	56,69 %	55,80 %	56,61 %
	EEE-FA	82,42 %	73,18 %	82,41 %	66,85 %	61,73 %	67,33 %
Consumo medio	SPR	73,67 %	85,79 %	77,97 %	80,68 %	85,03 %	82,13 %
	EEE-FA	76,56 %	78,69 %	78,58 %	81,39 %	82,88 %	82,40 %
Demandas bloqueadas	SPR	0	0	0	1,172 %	1,172 %	1,172 %
	EEE-FA	2,454 %	0,512 %	2,931 %	8,882 %	3,864 %	8,260 %

Tabla 5.3: Resumen de resultados de simulación

Los resultados experimentales muestran que se pueden conseguir importantes ahorros con EEE-FA en cierto tipo de escenarios, principalmente a cargas medias y para infraestructuras 1000BASE-T. Sin embargo, la reducción del coste en términos de consumo se consigue a costa de degradar otros aspectos del funcionamiento y administración de la red: rutas más largas y cargas de tráfico mayores. Lo cual se traduce en que los flujos experimentan en media un retardo extremo a extremo mayor.

Es importante resaltar que la mayor parte de las redes de área local utilizan el protocolo STP (*Spanning Tree Protocol*) para configurar sus rutas extremo a extremo. Por lo tanto, para poder aplicar el heurístico EEE.-FA propuesto, se recomienda que el administrador de red reemplace STP por otro tipo de solución que habilite la posibilidad de configurar rutas extremo a extremo a medida para las demandas de tráfico. Por ejemplo en el estándar IEEE802.1aq [107](SPB - *Shortest Path Bridging*), permite definir rutas basándose en conexiones IS-IS para construir árboles de distribución eficientes [108, 109]. Estas soluciones se proponen como líneas de investigación futuras para el autor de este trabajo de Tesis Doctoral.

El pesimista se queja del viento;
el optimista espera que cambie;
el realista ajusta las velas.

William George Ward

Capítulo 6

Análisis y simulación de algoritmos basados en IPACT para diferenciación de servicios basada en retardo en PONs

6.1. Introducción y trabajos relacionados

Las redes ópticas pasivas (en adelante PONs) se han propuesto en la literatura como una solución efectiva en coste al techo de ancho de banda de las tecnologías xDSL en la red de acceso [110, 111]. Las redes PON ofrecen las ventajas conocidas de las redes de comunicaciones ópticas tradicionales (un gran ancho de banda, baja atenuación con la distancia e inmunidad frente a las interferencias electromagnéticas) junto con los beneficios de implementación inherentes a las arquitecturas pasivas punto-multipunto (por ejemplo, se reduce el número de transceptores y no se precisa de alimentación eléctrica al ser pasivas, lo que redundaría en la reducción del coste).

La arquitectura típica de una PON (como se ha detallado en la sección 2.5.3.1 y que puede verse en la Fig. 6.1) consiste en una topología en árbol, donde las hojas o nodos terminales (Optical Network Units - ONU en adelante) se conectan al nodo raíz o terminal de línea (Optical Line Terminal - OLT en adelante) mediante un combinador/divisor (o *splitter/combiner*) pasivo (Véase Fig. 6.1). En el enlace descendente hacia los nodos de la PON (el sentido de la comunicación desde el OLT hacia las ONUs), el combinador/divisor pasivo sirve para reenviar una copia de la señal entrante a todas las ONUs conectadas a sus puertos de salida (en este caso, hace las funciones

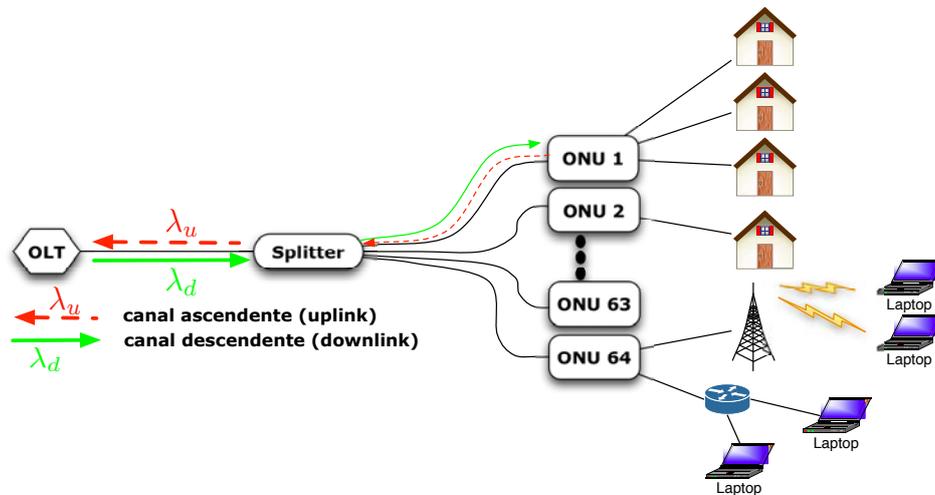


Figura 6.1: Esquema básico de una red PON

de divisor de potencia). Mientras que en el sentido ascendente (desde las ONUs hacia el OLT), el combinador/divisor, añade todas las señales que recibe generadas por las ONUs al canal que transporta el tráfico en dirección al OLT (en este caso, funcionando como un combinador pasivo). Ambos canales, ascendente y descendente, utilizan longitudes de onda distintas.

De este modo, la PON opera como una red de difusión-selección (*broadcast-and-select network*) en el canal de bajada, ya que los datos generados por el OLT se replican en el divisor/combinador y llegan a todas las ONUs, las cuales han de seleccionar aquellos datos destinados a cada una de ellas. Por otro lado, la longitud de onda del canal de subida se comparte por todas las ONUs, por lo que es necesario implementar algún mecanismo de arbitraje para evitar colisiones entre las señales de todas las ONUs, que comparten medio al mezclarse en el combinador/divisor pasivo. El Protocolo de Control MultiPunto definido para Ethernet (MPCP que como se ha descrito anteriormente en la sección 2.5.3.2 está definido en el estándar IEEE 802.3ah) especifica un mecanismo mediante el cual, el OLT primero estima los diferentes retardos entre él mismo y las ONUs (RTTs - Round Trip Times) y a continuación, planifica las ventanas temporales de TDM que asigna a cada ONU para la transmisión en el enlace ascendente en función de sus requerimientos de ancho de banda.

6.1.1. Estudio de trabajos anteriores en diferenciación de servicios

Existen propuestas en la literatura [112] de numerosos algoritmos de asignación dinámica de ancho de banda (DBA- Dynamic Bandwidth Allocation), de entre los cuales, el algoritmo IPACT

[113] (Interleaved Polling with Adaptive Cycle Time) parece el más extendido. En resumen, OLT y ONUs se comunican entre sí mediante dos mensajes de control: *Report* y *Gate*. Los mensajes *Report* los genera la ONU para informar a la OLT acerca del ancho de banda que necesita para planificar su próxima ventana de transmisión. Con esta información, el OLT decide el tamaño de ventana y el instante de comienzo adecuados para cada ONU. Finalmente, un mensaje *Gate* con esta información de planificación se envía a todas las ONUs (Fig. 6.2).

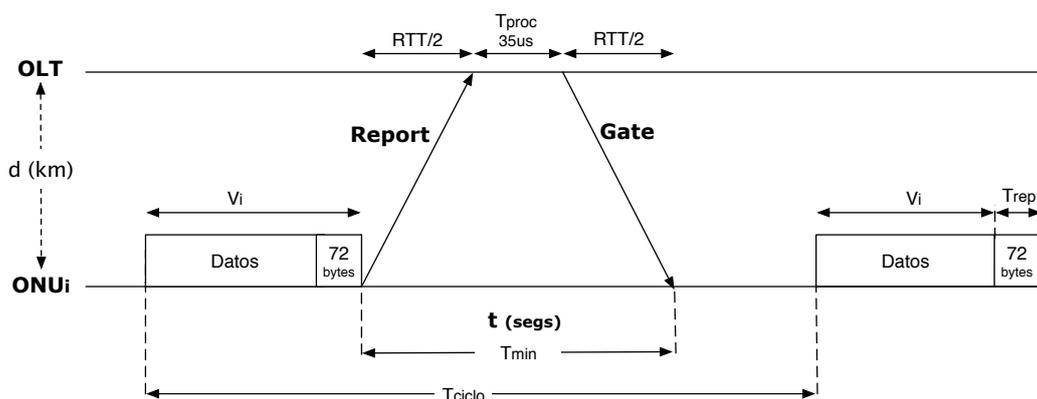


Figura 6.2: Esquema básico del funcionamiento del protocolo de DBA para una ONU

En relación al soporte para implementar calidad de servicio (QoS - Quality of Service) en PONs, la mayor parte de los estudios previos se centran en la diferenciación de los servicios en base al ancho de banda, en lugar de hacerlo en función del retardo. Una buena revisión de estas propuestas puede encontrarse en [112]. En este aspecto, también los autores de [114] proponen un algoritmo de reparto del ancho de banda que divide el ancho de banda total en unidades de tamaño fijo (por ejemplo de 10Mbps por unidad), de forma que, las ONU con QoS de alta prioridad reciban más unidades en el reparto que aquellas sin QoS (*best-effort*).

Otras propuestas de mejora se proponen en [115–117] donde las ONUs disponen de varias colas virtuales para separar diferentes clases de servicio, de forma que las ventanas de transmisión asignadas se reparten entre dichas colas de acuerdo a un algoritmo basado en pesos para la asignación de tiempo. De esta forma, se le asigna más tiempo de transmisión a las colas con tráfico de alta prioridad frente a las colas de tráfico no prioritario. En la mayor parte de las propuestas, es el OLT quien garantiza en definitiva un reparto justo del canal ascendente entre las ONUs, mientras que cada ONU se encarga de repartir el intervalo de transmisión que se le ha asignado en función del sistema de planificación de QoS que tenga implementado, que suele ser un algoritmo que asigna pesos en función de la jerarquía de las clases de servicio, como por ejemplo en [118].

Como conclusión, la mayor parte de los algoritmos de DBA para ofrecer QoS en PONs se centran en arquitecturas de servicios diferenciados basados en requisitos de reparto del ancho de banda entre las distintas clases de tráfico. Existen hasta la fecha pocos estudios que tengan en cuenta el impacto sobre el retardo medio y la variabilidad del retardo (*jitter*) desde el punto de vista de una ONU dada, como puede verse en los estudios de los autores de [119–122]. Aun así, estos estudios no proporcionan un algoritmo claro de diferenciación basado en retardo, que es una característica crítica para un gran número de aplicaciones actuales tales como la telefonía IP o la industria de los juegos en red interactivos.

De este modo, en este capítulo se propone un algoritmo de DBA para servicios diferenciados basado en retardo para redes PON con IPACT implementado. En las secciones siguientes se presenta su estudio analítico y los resultados de simulación. En resumen, este algoritmo consiste en transmitir más a menudo el tráfico de alta prioridad que el de baja prioridad, de forma que el tiempo entre ciclos de transmisión de los paquetes correspondientes al tráfico de alta prioridad sea mucho menor (de una forma proporcional y configurable) a los tiempos entre transmisiones de paquetes de baja prioridad. En la sección 6.2 se presenta una revisión de trabajos previos relativos al estudio del análisis de retardo de IPACT siguiendo el análisis de [123, 124], pero sin diferenciación de servicios. A continuación, la sección 6.4 analiza el caso de que exista un exceso de carga de tráfico en la red, y cómo el algoritmo garantiza la calidad de servicio y mantiene el retardo experimentado por los paquetes de alta prioridad por debajo de una cota máxima, a costa de sufrir un determinado porcentaje de pérdidas asumibles en el tráfico de baja prioridad. En la sección 6.5 se presentan los experimentos de simulación realizados con el fin de validar las ecuaciones derivadas del estudio analítico, de forma que se puedan comparar y evaluar las prestaciones y beneficios que ofrece este algoritmo frente al funcionamiento convencional de IPACT.

6.2. Estudio analítico del modelo de asignación de intervalos TDM para DBA en PON

Esta sección repasa el modelo analítico general de reparto de ancho de banda y retardo en una red PON que tenga implementado el algoritmo IPACT para la gestión dinámica de las reservas de canal ascendente por parte de las ONUs, tal y como proponen los autores de [124](Sección 3B). Se recomienda la lectura de dicho trabajo en caso de que el lector de esta tesis doctoral esté interesado en los detalles más concretos del modelo.

6.2.1. Suposiciones previas para el análisis del modelo

En el algoritmo IPACT, el OLT es el terminal encargado de asignar las ventanas de transmisión, mediante la información de los instantes garantizados y sus tamaños correspondientes, *grants*, en el mensaje *GATE*, de las ONUs, basándose en la información de reserva de ancho de banda que ha recibido de las ONUs, y ciertas reglas internas que garanticen el reparto justo de la capacidad total de la longitud de onda de subida.

Para simplificar el análisis, se asumen las siguientes condiciones:

- El análisis corresponde al comportamiento en estado estacionario del algoritmo IPACT, una vez sincronizadas las ONUs y medidos sus RTTs.
- Las ONUs se encuentran todas a la misma distancia, d km, del OLT.
- La carga de tráfico ofrecida por cada ONU es constante a lo largo del tiempo, ρ_i , $i = 1, \dots, N$. En otras palabras, la i -ésima ONU recibe de sus usuarios conectados un tráfico agregado que sigue una distribución de Poisson con una tasa media λ_i paquetes/segundo. Además, cada paquete requiere un tiempo de servicio medio fijo de $E[X] = \frac{1}{\mu}$ calculado como:

$$E[X] = \frac{1}{\mu} = \frac{8B}{C} \text{segundos} \quad (6.1)$$

donde B indica el tamaño medio de paquete en bytes y C denota la velocidad de transmisión del enlace. Para $B = 1518$ bytes y $C = 1$ Gbps, el tiempo de servicio que se invierte es de $E[X] = 12,14\mu s$ por paquete.

Por tanto, la i -ésima ONU dispone de una carga de tráfico ofrecido ρ_i , $i = 1, \dots, N$ de:

$$\rho_i = \frac{\lambda_i}{\mu} \quad (6.2)$$

Finalmente, se asume la condición de que la carga de tráfico ofrecida total de la red ρ_T , esto es, la suma de todas las cargas de tráfico ofrecidas independientemente por cada ONU, ρ_i , ha de ser menor que la unidad:

$$\rho_T = \sum_{i=1}^N \rho_i < 1 \quad (6.3)$$

6.2.2. Análisis del tiempo medio de ciclo

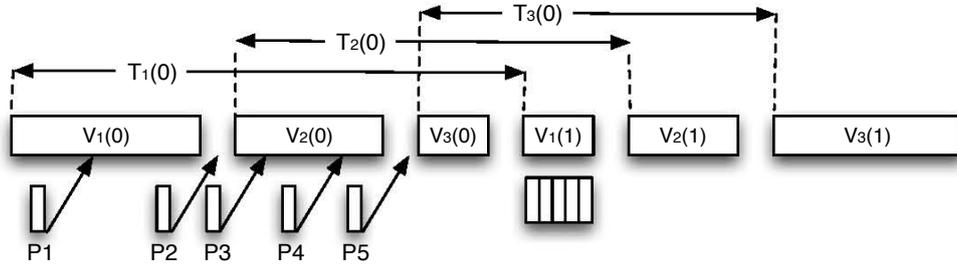


Figura 6.3: Definiciones para una PON con $N = 3$ ONUs

En primer lugar, se define el término tiempo de ciclo (en adelante T_{ciclo}) como el tiempo que transcurre entre el inicio de dos intervalos consecutivos asignados a la misma ONU para la transmisión de su tráfico de entrada en el canal ascendente de la PON. Para describir el análisis del tiempo medio de ciclo entre dos intervalos de transmisión de una ONU utilizaremos una red de ejemplo con $N = 3$ ONUs, cuyos intervalos asignados son como los que pueden verse en la Fig. 6.3. Con $T_i(k)$ se denota el k -ésimo tiempo de ciclo para la i -ésima ONU. $V_i(k)$, en cambio, designa el tamaño en tiempo de la ventana de transmisión de la ONU i en el instante discreto k , donde $k = 0, 1, \dots$ (recuadros de la figura). Es obvio notar que las ventanas de transmisión tendrán una longitud variable en función del número de paquetes que hayan llegado a la ONU durante la duración del ciclo anterior. Por ejemplo, en el instante $k = 1$, el tamaño de $V_1(1)$ en la figura depende de la cantidad de paquetes que hayan llegado a la ONU 1, $N_1(1)$, durante el tiempo de ciclo anterior, $T_1(0)$, que en este caso son cinco paquetes (P1 a P5 en la figura). Asumiendo la premisa de que la llegada de paquetes es un proceso de Poisson con tasa λ_i , la probabilidad de que hayan llegado exactamente n paquetes dentro de su tiempo de ciclo anterior correspondiente, de longitud $T_i(k)$ segundos, sigue una función de densidad de probabilidad de Poisson:

$$P(N_i(k + 1) = n) = \frac{(\lambda_i T_i(k))^n}{n!} e^{-\lambda_i T_i(k)}, \quad i = 0, 1, \dots \tag{6.4}$$

donde el número medio de llegadas de paquetes es:

$$E[N_i(k + 1)] = \lambda_i E[T_i(k)] \text{ paquetes} \tag{6.5}$$

donde $E[T_i(k)]$ es el tiempo de ciclo medio para la ONU i en el instante k . De este modo, bajo la suposición de que cada paquete requiere un tiempo medio de servicio de $S = \frac{1}{\mu}$ segundos, la el tamaño de ventana medio requerido por la i -ésima ONU para la próxima transmisión en el instante $k + 1$ es:

$$E[V_i(k + 1)] = \frac{1}{\mu} E[N_i(k + 1)] = \rho_i E[T_i(k)] \text{ segundos} \quad (6.6)$$

Finalmente, el tiempo de ciclo medio $E[T_i(k)]$ se calcula como la suma de las N ventanas de transmisión previas, más el tiempo de guarda entre dos ventanas de diferentes ONUs, más el tiempo de transmisión del mensaje de reserva o *Report*, los cuales se representan como los huecos vacíos entre ventanas de la Fig. 6.3:

$$E[T_i(k)] = \sum_{i=1}^N E[V_i(k)] + N(T_{guard} + T_{report}) \text{ segundos} \quad (6.7)$$

donde T_{guard} denota el tiempo de guarda entre ventanas TDM consecutivas ($T_{guard} = 1,5\mu s$ es el valor recomendado con velocidades de línea de 1Gbps) y $T_{report} = \frac{8 \cdot 64}{10^9} \approx 0,5\mu s$ es el tiempo de transmisión del mensaje *Report* que tiene un tamaño de 64 bytes [124].

En estado estacionario, se puede observar que el sistema presenta un tiempo medio de ciclo, $E[T]$, que no depende del instante de tiempo k . De la misma manera, las ventanas de transmisión $V_i(k)$ presentan un tamaño medio, $E[V_i]$, independiente del instante. Por tanto:

$$E[V_i] = \rho_i E[T], \quad i = 1, \dots, N \quad (6.8)$$

$$E[T] = \sum_{i=1}^N \rho_i E[T] + NT_0 \quad (6.9)$$

donde $T_0 = T_{guard} + T_{report}$ para abreviar las ecuaciones.

Resolviendo el sistema de ecuaciones anterior para $E[T]$ se obtiene:

$$E[T] = \frac{NT_0}{1 - \sum_i \rho_i} = \frac{NT_0}{1 - \rho_T} \quad (6.10)$$

de modo que $E[T]$ sólo depende de la carga de tráfico total que se agrega al canal ascendente ρ_T , del número de ONUs N y de la suma de los tiempos de guarda y de transmisión del mensaje de reserva T_0 ($T_0 = T_{guard} + T_{report}$).

Finalmente, vale la pena resaltar que el modelo representado por la Ec. 6.10 sólo es válido a cargas medias y altas (véase [124]), debido a las restricciones que impone el usar el sistema de "polling" de MPCP e IPCT: el mensaje con la información de la siguiente ventana de transmisión (mensaje *Gate*) no puede ser enviado por el OLT hasta que no haya recibido la información del mensaje de reserva (*Report*) y haya sido procesado en función del algoritmo de asignación. Para cumplir con esta condición, es necesario que el tiempo medio de ciclo, $E[T]$, sea mayor que el retardo de propagación de la red en ambos sentidos (RTT - Round-Trip Time) más el retardo de procesado del paquete *Report* por parte del OLT, T_{proc} y el tiempo de transmisión de su ráfaga de datos $V_i + T_{rep}$. En resumen, la ec. 5.8 es válida sólo si:

$$E[T] = \frac{NT_0}{1 - \rho_T} > RTT + T_{proc} + V_i + T_{rep} \quad (6.11)$$

La ec.6.11 indica que existe un valor mínimo tanto del tiempo de ciclo, $T_{ciclo_{min}}$, como de carga de la red, ρ_{min} , a partir de los cuales, la carga de tráfico influye más en la duración de un ciclo, que el tiempo que ha de esperar la ONU para recibir el nuevo instante de comienzo y la duración de su intervalo de transmisión garantizado. En estas condiciones, el modelo descrito sigue la ec. 6.10. El tiempo de ciclo mínimo ha de permitir que se espere a recibir la información y se transmita la ráfaga correspondiente a dicho ciclo, trabajando con una carga de red total ρ_{min} :

$$T_{c_{min}} = RTT + T_{proc} + V_i + T_{rep} = T_{min} + \rho_i * T_{c_{min}} \quad (6.12)$$

$$T_{c_{min}} = \frac{T_{min} + T_{rep}}{1 - \frac{\rho_{min}}{N}} \quad (6.13)$$

La ecuación 6.12 se puede simplificar si identificamos T_{min} con el tiempo mínimo que transcurre hasta que se recibe el paquete *Gate* con la información del siguiente intervalo de transmisión: $T_{min} = RTT + T_{proc}$ y que depende de la distancia entre el OLT y la ONU. Suponiendo que todas las ONUs están a la misma distancia y que tienen la misma distribución de carga $\rho_i = \frac{\rho_T}{N}$ y sustituyendo el valor de $T_{c_{min}}$ en la ec. 6.12 en la ec. 6.10 se obtiene la expresión:

$$\rho_{min} = \frac{T_{min} + T_{rep} - NT_0}{T_{min} + T_{rep} - T_0} = \frac{T_{min} - (N - 1)T_{rep}}{T_{min} - T_{guard}} \quad (6.14)$$

El denominador de la ec. 6.14 es un valor positivo dado que en general el tiempo T_{min} será mayor que el tiempo de guarda T_{guard} . La expresión del numerador sin embargo indica que para distancias cortas, en redes PON con un número suficiente de ONUs $N > 1 + \frac{T_{min}}{T_{rep}}$, la carga mínima se supera con el simple hecho de que todas las ONUs manden sus mensajes de *Report*.

Para una arquitectura típica de PON con velocidad de línea de 1Gbps y con $N = 32$ ONUs, y una distancia de $d = 10$ km entre el OLT y las ONUs, y suponiendo que el tiempo de procesamiento de los paquetes de reserva es de $T_{proc} = 35\mu s$, el requisito anterior se cumple para un valor de carga de la red mínimo $\rho_T > 0.47$.

En la Fig.6.4 pueden verse resumidos, para diferente número de ONUs en la red PON, los requisitos descritos del valor mínimo de carga $\rho_{T_{min}}$ en función de la distancia al OLT. En resumen, este valor de carga mínima es aquel a partir del cual, el tiempo mínimo para realizar la reserva no influye, gracias a que la PON está suficientemente cargada como para que el tiempo de ciclo dependa más del tiempo de servicio de los paquetes a enviar en el total del canal de subida que de la distancia al OLT.

En los casos en los que el requisito de carga media no se cumple, el tiempo medio de ciclo viene impuesto por los retardos de propagación RTT y de procesamiento del paquete de reserva. Con

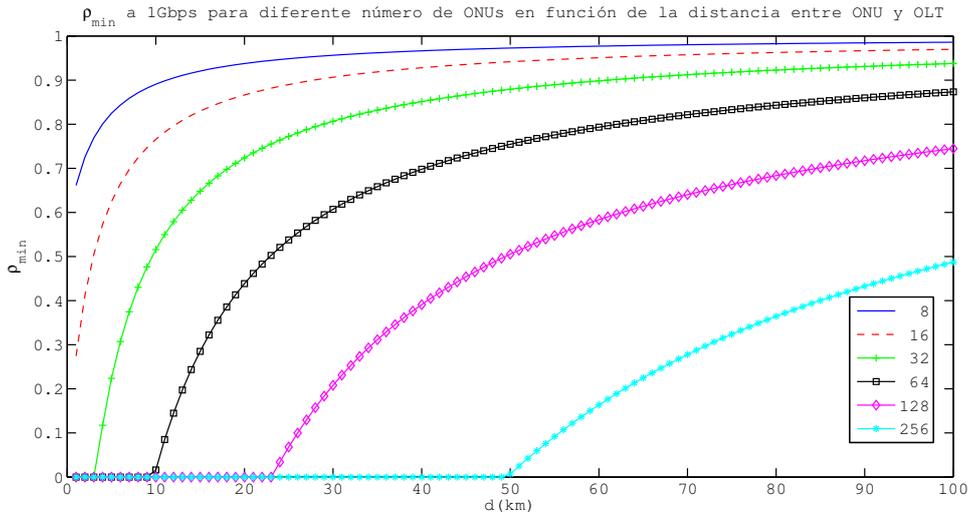


Figura 6.4: Carga de tráfico en la red mínima para que el tiempo de ciclo T_{ciclo} no dependa de la influencia de T_{min}

los datos del ejemplo anterior esto supone un valor mínimo para el tiempo medio de ciclo de:

$$T_{min} = RTT + T_{proc} = 2 \frac{10km}{2 \cdot 10^5 km/s} + 35\mu s = 135\mu s \quad (6.15)$$

6.2.3. Análisis del retardo máximo

El retardo que sufre un paquete aleatorio que llega a una ONU antes de ser transmitido por la PON se encuentra comprendido entre $E[T]$ y $2E[T]$. El valor mínimo será para los paquetes que lleguen justo antes de que la ONU realice las estimaciones de ancho de banda que necesita en su siguiente intervalo y genera el mensaje de *Report*, suponiendo que se encuentren la cola de espera vacía, mientras que el valor máximo ocurre cuando el paquete llega justo después del cálculo de la estimación de la siguiente reserva y del envío del mensaje *Report* al OLT, tal y como se describe en [124]. De este modo, el retardo medio experimentado por un paquete cualquiera será proporcional al $E[T]$. En adelante se utiliza $E[T]$ como la métrica elegida para evaluar las prestaciones del algoritmo de calidad de servicio en DBA propuesto en este trabajo de tesis doctoral, teniendo además en cuenta que, el retardo que experimentarán los paquetes antes de ser transmitidos por la ONU, será proporcional a dicho $E[T]$.

6.3. Análisis del modelo de asignación de TDM para PON con diferenciación basada en retardo

6.3.1. Descripción del algoritmo de planificación con diferenciación basado en retardo

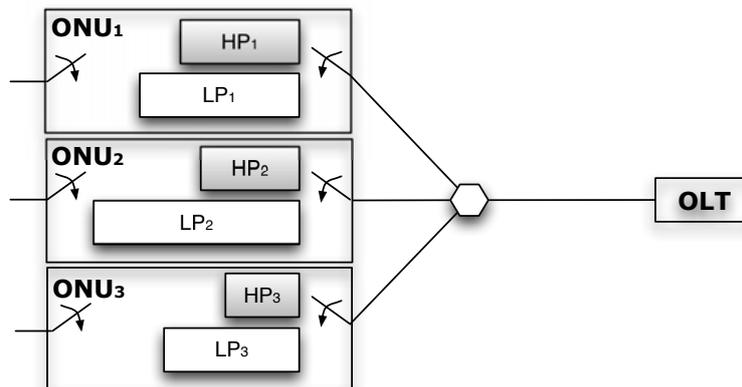


Figura 6.5: Colas de espera para las clases de servicio HP y LP en cada ONU

Para la descripción del algoritmo de planificación con diferenciación basada en retardo se considera que cada ONU dispone de dos colas de salida (Fig.6.5, una para cada clase de tráfico: tráfico de alta prioridad (*High Priority Class*, en adelante HP) y tráfico de baja prioridad (*Low Priority Traffic*, en adelante LP). La primera clase (HP) está destinada a aquellos paquetes correspondientes a aplicaciones sensibles al retardo, típicamente aplicaciones multimedia interactivas (aplicaciones de vídeo interactivo, voz sobre IP, juegos interactivos en red, etc.), mientras que la segunda (LP) corresponde al resto de aplicaciones que en este caso no tienen requisitos de calidad de servicio de este tipo (*best-effort*).

El objetivo principal de este trabajo consiste en el diseño de un algoritmo de planificación de los intervalos de transmisión que favorezca a la clase HP en detrimento de la clase LP, en base al retardo experimentado por el tráfico en la ONU correspondiente. La mayor parte de los estudios previos consultados proponen la planificación de la transmisión de los paquetes correspondientes al tráfico de HP en los primeros instantes de la ventana de transmisión asignada con el objetivo de priorizarlos y que tengan menos retardo. Este algoritmo difiere del resto en que propone separar los mensajes de *Report* para la reserva de intervalo TDM para la cola HP de los de la cola LP, de forma que se solicita al OLT la asignación de ancho de banda para que el tráfico correspondiente a HP se envíe más regularmente que el de LP. Esto supone a priori una reducción del tiempo de

espera de los paquetes que llegan a la cola de alta prioridad de una ONU dada, pero a costa de aumentar el retardo que sufre el tráfico de la cola LP. En la Fig. 6.6 se muestra un ejemplo de funcionamiento para una red PON de 3 ONUs.

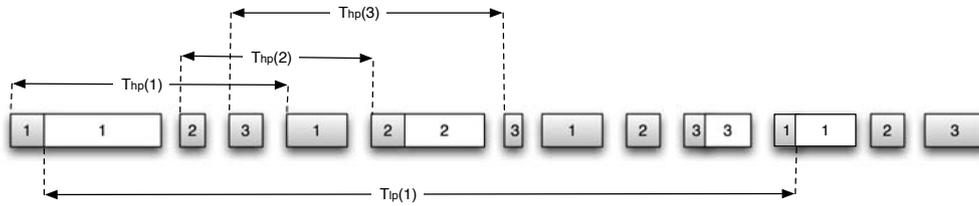


Figura 6.6: Algoritmo con diferenciación basada en retardo. Caso $M = 1$.

En este ejemplo, el tráfico HP (recuadros sombreados) de las tres ONUs se envía a la OLT más frecuentemente que el tráfico de la cola LP (recuadros en blanco). Se observan dos tipos de ciclos de transmisión distintos:

- unos tiempos de ciclo más cortos correspondientes a la espera entre dos asignaciones de canal consecutivas para reservas solicitadas para transmitir el tráfico de alta prioridad, $T^{(HP)}$, por parte de la misma ONU.
- y otros ciclos más grandes, $T^{(LP)}$, correspondientes al tiempo entre dos transmisiones de tráfico de baja prioridad de la misma ONU.

En general se cumple que $T^{(HP)} \leq T^{(LP)}$. Básicamente, el funcionamiento del algoritmo puede resumirse en los siguientes pasos:

- el OLT organiza los intervalos necesarios para los paquetes de HP de las tres ONUs y para los de LP de una de las ONUs (ONU 1 en el ejemplo) dentro de un primer ciclo de transmisión de HP.
- En una segunda iteración, el OLT planifica de nuevo intervalos para las tres colas de alta prioridad de las 3 ONUs respectivamente y para el tráfico de baja prioridad de la segunda ONU.
- Finalmente, en el tercer tiempo de ciclo de alta prioridad consecutivo, el OLT vuelve a asignar tiempo a las colas de alta prioridad y por fin asigna un intervalo para la cola de LP de la tercera ONU.

- En un cuarto ciclo, se vuelve a repetir el algoritmo, asignando ventanas de transmisión para el tráfico de alta prioridad de todas las ONUs y, de nuevo, para el tráfico de baja prioridad de la primera ONU como en la primera iteración.

Cabe mencionar que con este algoritmo, el tiempo de ciclo entre dos transmisiones consecutivas correspondientes al tráfico de baja prioridad de la misma ONU es la suma de tres tiempos de ciclo de alta prioridad, como puede verse en la Fig. 6.6.

En general, las ventanas de transmisión de HP son más pequeñas que las de LP por dos razones:

- Primero, las aplicaciones sensibles al retardo generan menos tráfico que las aplicaciones sin calidad de servicio o *best-effort*.
- Segundo, a la cola de salida de alta prioridad se le da servicio más a menudo que a la de baja prioridad, con lo que la cola LP agregará más tráfico durante su ciclo de espera y reservará un intervalo mayor para su ventana de transmisión siguiente.

Para generalizar el algoritmo, se usa M para designar el número de ONUs que reserva un intervalo de transmisión para su cola de baja prioridad en cada ciclo entre transmisiones de alta prioridad. La Fig. 6.7 muestra los casos posibles $M = 1, 2, 3$ en una PON de $N = 3$ ONUs. Como se ve en la figura, el caso de $M = 3$ produce los mismos tiempos de ciclo para ambas clases de tráfico, HP y LP, por lo que no habría diferenciación entre ambas clases de servicio.

La siguiente sección analiza el tiempo de ciclo medio para cada clase de tráfico.

6.3.2. Análisis del tiempo de ciclo medio para cada clase de tráfico

Siguiendo el mismo razonamiento descrito en la sección 6.2, la duración media de las ventanas de transmisión de HP reservadas por las ONUs depende de la cantidad media de tráfico agregada durante su tiempo de ciclo de espera previo entre dos intervalos de transmisión de tráfico de alta prioridad. La cantidad de tráfico medio agregado de HP será menor en general que la del tráfico de LP, dado que su tiempo de ciclo es menor.

A continuación se usa $E[V_i^{(HP)}]$ y $E[V_i^{(LP)}]$ para designar las ventanas de transmisión de tamaño medio requeridas para la i -ésima ONU para el tráfico de las colas HP y LP respectivamente.

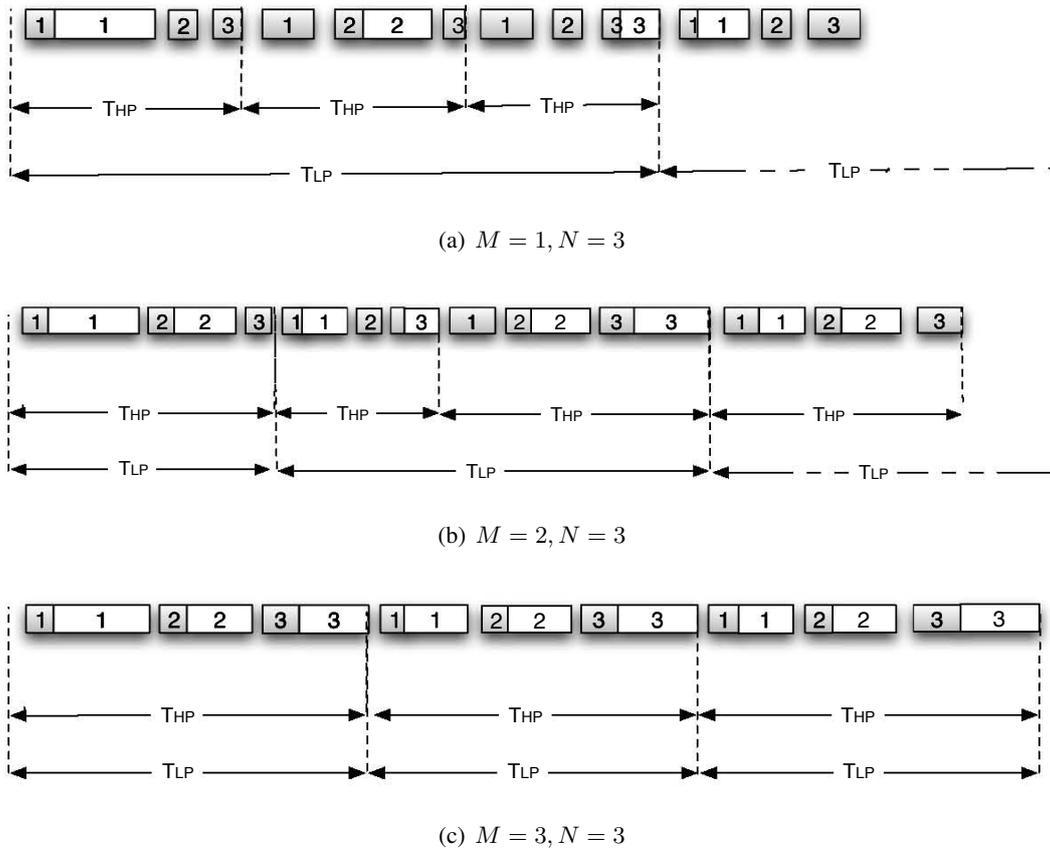


Figura 6.7: Algoritmo con diferenciación basada en retardo. Casos $M = 1, 2$ y 3 y $N = 3$. Instantes de tiempo referidos a la ONU_i con $i = 1$

Las siguientes ecuaciones, en estado estacionario, proporcionan el tamaño medio de las ventanas de transmisión para el tráfico HP y LP:

$$E[V_i^{(HP)}] = \rho_i^{(HP)} E[T^{(HP)}] \quad (6.16)$$

$$E[V_i^{(LP)}] = \rho_i^{(LP)} E[T^{(LP)}] \quad (6.17)$$

donde $\rho_i^{(HP)} = p_h \rho_i$ y $\rho_i^{(LP)} = (1 - p_h) \rho_i$ para una factor de proporcionalidad $p_h \in (0, 1)$.

Este factor p_h indica el porcentaje del tráfico total que corresponde a las aplicaciones sensibles al retardo y que van a ingresar en la ONU a través de la cola de alta prioridad HP. Además, ha de cumplirse de nuevo, que la cantidad de tráfico total ofrecido a la red no debe superar la unidad:

$$\rho_T = \sum_{i=1}^N \left(\rho_i^{(HP)} + \rho_i^{(LP)} \right) < 1 \quad (6.18)$$

En la siguiente ecuación se utilizan los términos $E[T^{(HP)}]$ y $E[T^{(LP)}]$ para denotar los tiempos de ciclo medios de HP y LP respectivamente. Es importante resaltar en la Fig. 6.7 cómo el tiempo medio de ciclo de HP incluye N ventanas de transmisión de tráfico de alta prioridad más otras M ventanas correspondientes a tráfico de colas de baja prioridad, junto con los N tiempos de guarda y de transmisión del mensaje de reserva, denotados por T_0 . Por lo tanto:

$$E[T^{(HP)}] = \sum_{i=1}^N E[V_i^{(HP)}] + \sum_{i=1}^M E[V_i^{(LP)}] + NT_0 \quad (6.19)$$

de la cual se puede obtener:

$$E[T^{(HP)}] = N\rho_i E[T^{(HP)}] + M\rho_i(1 - p_h)E[T^{(LP)}] + NT_0 \quad (6.20)$$

debido a que:

$$E[V_i^{(HP)}] = \rho_i p_{hp} E[T^{(HP)}] \quad (6.21)$$

$$E[V_i^{(LP)}] = \rho_i(1 - p_{hp})E[T^{(LP)}] \quad (6.22)$$

de la ecuaciones 6.16 y 6.17.

De la misma manera, para $E[T^{(LP)}]$, se tiene que:

$$M \cdot E[T^{(LP)}] = N \sum_{i=1}^N E[V_i^{(HP)}] + N \sum_{i=1}^M E[V_i^{(LP)}] + N^2 T_0 \quad (6.23)$$

gracias a que N tiempos de ciclo medios de HP corresponden a M tiempos de ciclo medios de LP, tal y como se ven la Fig. 6.7. De manera que:

$$E[T^{(LP)}] = \frac{N}{M} \sum_{i=1}^N E[V_i^{(HP)}] + \frac{N}{M} \sum_{i=1}^M E[V_i^{(LP)}] + \frac{N^2}{M} T_0 \quad (6.24)$$

Por lo tanto, es posible usar las ecuaciones 6.20 y 6.24 como un sistema de dos ecuaciones y dos incógnitas a resolver para obtener $E[T^{(HP)}]$ y $E[T^{(LP)}]$:

$$E[T^{(HP)}] = \frac{NT_0}{1 - \rho} \quad (6.25)$$

$$E[T^{(LP)}] = \frac{N}{M} \frac{NT_0}{1 - \rho} \quad (6.26)$$

Como conclusión, se puede observar que hay una relación de proporcionalidad, M/N , entre el tráfico HP y el LP como se observa en:

$$\frac{E[T^{(HP)}]}{E[T^{(LP)}]} = \frac{M}{N} \quad (6.27)$$

Esta relación de proporcionalidad entre ambos, permite al planificador definir el nivel de diferenciación basada en retardo entre el tráfico de alta prioridad (HP) y el de baja (LP), simplemente ajustando el valor apropiado de M en el planificador. En el caso de $M = 1$ se consigue la máxima diferenciación entre las dos clases de tráfico, mientras que en el caso $M = N$ no se obtiene diferencia entre ellos.

6.4. Análisis con tiempo de ciclo de alta prioridad garantizado

La sección anterior analizaba el mecanismo propuesto para la diferenciación de la clase de tráfico prioritario (HP) de la de baja prioridad (LP), de forma proporcional al retardo. Sin embargo, este mecanismo no garantiza una cota máxima al retardo que experimenta el tráfico de alta prioridad, esto es, el tráfico de la cola HP sufrirá un retardo añadido más pequeño que el que sufre el tráfico de la cola LP, pero aún así puede seguir siendo demasiado alto para los requisitos de las aplicaciones de tiempo real. Esta sección propone una extensión al algoritmo propuesto anteriormente mediante la definición de un límite superior estricto en la duración de los tiempos de ciclo de alta prioridad, incluso a costa de perder una cierta cantidad de tráfico de baja prioridad. Para la descripción del algoritmo, se usa $T_{Limit}^{(HP)}$ para referirse a dicha cota superior para el tiempo de ciclo del tráfico prioritario.

Básicamente, cuando el OLT recibe los *Report* de todas las ONUs, para la reserva y procesado de los siguientes tiempos de ciclo de HP (lo cual incluye las N ventanas de tráfico HP y las M ventanas de tráfico no prioritario, debe comprobar si la suma de todas las ventanas de transmisión pedidas por cada una de las ONUs, está por debajo de dicho umbral $T_{Limit}^{(HP)}$. En caso afirmativo, el OLT puede asignar todas las ventanas de transmisión requeridas por las ONUs. Sin embargo, si dicho total de tamaños de ventanas de transmisión, supera el umbral de $T_{Limit}^{(HP)}$, entonces la OLT sólo puede planificar intervalos más pequeños para el tráfico LP. Este exceso de tráfico de baja prioridad deberá ser retrasado hasta el siguiente ciclo LP o descartarse y contabilizarse como paquete perdido. La Fig. 6.8 muestra un ejemplo de la situación para valores de $N = 3$, $M = 1$.

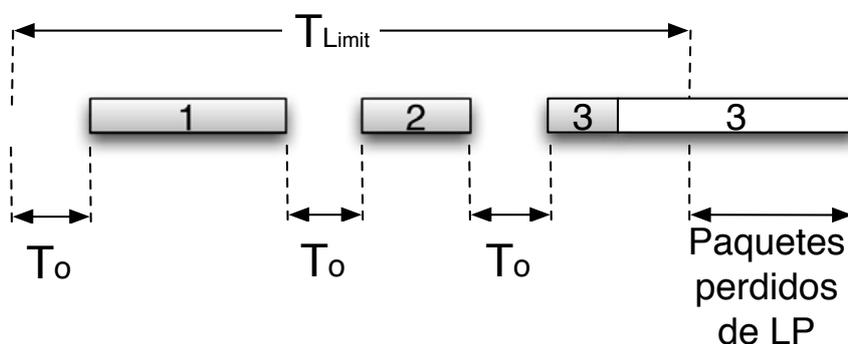


Figura 6.8: Algoritmo con diferenciación de servicios dentro de los umbrales correspondientes. Caso $M = 1$ y $N = 3$.

En este ejemplo, la cantidad total de las ventanas de transmisión excede el ciclo de tiempo de HP. Esto es:

$$T^{(HP)} = NT_0 + \sum_{i=1}^N V_i^{(HP)} + V_3^{(LP)} > T_{Limit}^{(HP)} \quad (6.28)$$

Por tanto, en un tiempo de ciclo de alta prioridad, siendo T' el número de paquetes ofrecidos por la ONU, cuya expresión es N_{Off} , es igual a:

$$\begin{aligned} E[N_{Off}] &= \left(\sum_{i=1}^N \lambda_i p_{hp} \right) T' \\ &+ M \frac{1}{N} \left(\sum_{i=1}^N \lambda_i (1 - p_{hp}) \right) \frac{N}{M} T' \\ &= \frac{\rho p_{hp} T'}{E[X]} + \frac{\rho (1 - p_{hp}) T'}{E[X]} \\ &= \frac{\rho T'}{E[X]} \text{ paquetes} \end{aligned} \quad (6.29)$$

el cual designa los paquetes ofrecidos de HP y de LP durante un tiempo de ciclo de duración T' .

El número total de paquetes N_{Limit} permitido en $T_{Limit}^{(HP)}$ necesita sustraer los N tiempos de guarda y los tiempos de procesamiento del mensaje *Report* y dividirlo por el tiempo medio de servicio por paquete $E[X]$. Dicho con otras palabras:

$$N_{Limit} = \frac{T_{Limit}^{(HP)} - NT_0}{E[X]} \text{ paquetes} \quad (6.30)$$

Finalmente, el porcentaje de paquetes perdidos de la clase de baja prioridad es:

$$\begin{aligned} P_{Loss} &= \frac{E[N_{Off}] - N_{Limit}}{E[N_{Off}]} \\ &= \frac{\rho \frac{T_{Limit}}{E[X]} - \frac{T_{Limit} - NT_0}{E[X]}}{\frac{T_{Limit} - NT_0}{E[X]}} \\ &= \frac{T_{Limit}(\rho - 1) + NT_0}{\rho T_{Limit}} \end{aligned} \quad (6.31)$$

6.5. Resultados experimentales

Esta sección muestra un cierto número de experimentos para demostrar las ecuaciones propuestas anteriormente y las conclusiones obtenidas para ejemplos de situaciones reales.

6.5.1. Validación de los tiempos de ciclo medios $E[T]$, $E[T^{(HP)}]$ y $E[T^{(LP)}]$

En el primero de los experimentos, el objetivo es validar las ecuaciones obtenidas anteriormente para el algoritmo de diferenciación de servicios basado en el retardo $E[T^{(HP)}]$ y $E[T^{(LP)}]$ (eqs. 6.20 and 6.24), comparándolo con el resultado que se obtendría en caso de no usar el algoritmo de diferenciación de clases en una red PON basada en IPACT. En ambos casos se consideran los siguientes parámetros de configuración:

- Velocidad de línea: $C = 1$ Gbps.
- Tiempo de Guarda $T_{guard} = 1,5\mu s$
- Tiempo de procesado: $T_{report} = 0,512\mu s$
- Porcentaje de tráfico de HP respecto del total: $p_h = 0,4$

Las Figs. 6.9(a)-6.9(d) muestran $E[T^{(HP)}]$ y $E[T^{(LP)}]$ para 32 ONUs y diferentes valores de M en el planificador. Como puede verse, para $M = 1$, la diferenciación por retardo es máxima (con proporción 1:32), mientras que para $M = 32$ no existe diferenciación de servicios. En las gráficas se presentan resultados para valores intermedios $M = 4$ y $M = 16$. Además, que las fórmulas analíticas cuadran de forma muy aproximada con los resultados de simulación, validando así el estudio analítico.

Las figs. 6.10(a) 6.10(d) muestran el tiempo de ciclo medio, más menos el doble de su desviación estándar para $E[T]$ y $E[T^{(HP)}]$. Como puede observarse, aunque los tiempos medios de ciclos son similares para ambos casos, la desviación estándar es mucho menor en el algoritmo propuesto en este trabajo, lo que se traduce en una menor dispersión de los tiempos de ciclo de HP, lo que supondría una variabilidad del retardo (*jitter*) pequeña, que es el requisito de muchas aplicaciones sensibles a retardo, especialmente para cargas elevadas y para valores bajos de M . El tiempo de ciclo medio y su varianza han sido estimados mediante simulación.

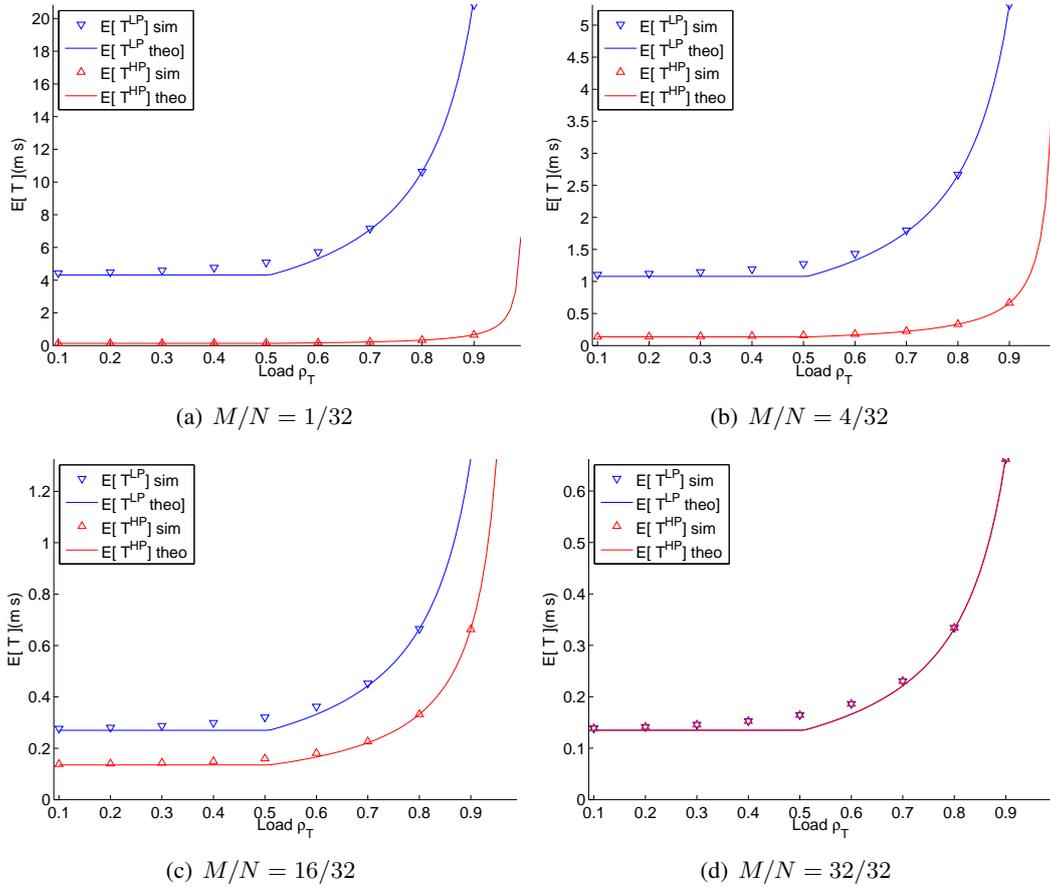


Figura 6.9: Tiempos de ciclos medios $E[T^{(HP)}]$ y $E[T^{(LP)}]$ para diferentes valores del cociente M/N : (a) $\frac{M}{N} = \frac{1}{32}$, (b) $\frac{M}{N} = \frac{4}{32}$, (c) $\frac{M}{N} = \frac{16}{32}$, (d) $\frac{M}{N} = \frac{32}{32}$

6.5.2. Experimentos bajo condiciones de tráfico elevado

Las figs. 6.11(a)-6.11(d) muestran el comportamiento del algoritmo propuesto para la diferenciación de servicios basada en retardo, en situaciones de una elevada carga de tráfico entrante, lo que significa que la carga total de la red es un valor dentro del rango $1 \leq \rho_T \leq 2$. Se consideran dos posibles umbrales límite en las simulaciones: $T_{Limit}^{(HP)} = 3,2ms$ (umbral superior) y $T_{Limit}^{(HP)} = 8ms$ (umbral inferior).

Como puede observarse, los valores de los tiempos de ciclo para el tráfico HP se mantienen por debajo del límite de $T_{Limit}^{(HP)}$ en ambos casos, mientras que los tiempos de ciclo de LP son mucho mayores. Además, para garantizar los umbrales de retardo, cierta cantidad del tráfico de baja prioridad es descartada, tal y como se ha explicado en la sección 6.4.

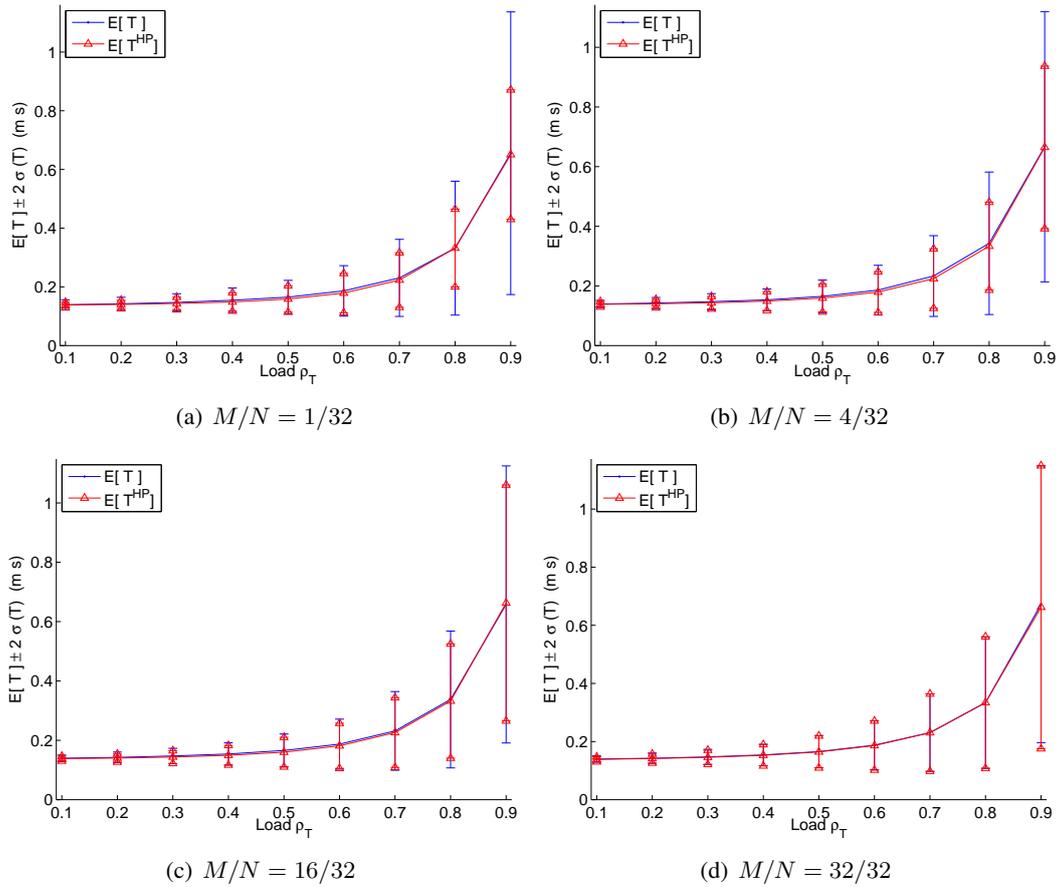


Figura 6.10: Tiempos medios de ciclo $E[T]$, $E[T^{(HP)}]$ y sus desviaciones típicas ($E[T] \pm 2Std[T]$) para diferentes relaciones M/N : (a) $\frac{M}{N} = \frac{1}{32}$, (b) $\frac{M}{N} = \frac{4}{32}$, (c) $\frac{M}{N} = \frac{16}{32}$, (d) $\frac{M}{N} = \frac{32}{32}$

6.6. Resumen y conclusiones

En este capítulo se ha descrito la propuesta de diseño de un nuevo algoritmo de asignación dinámica de ancho de banda, cuyo principal objetivo es el de proporcionar diferenciación de servicios basada en la variación de retardo, para redes ópticas pasivas (PON) basadas en IPACT. Básicamente, el algoritmo define dos tipos de ciclos de asignación de ventanas temporales de transmisión en un TDM dinámico por parte del OLT: un tipo de ciclo para el tráfico de alta prioridad (ciclos HP) y otro para el de baja (ciclos LP), de forma que se asignen ciclos más regulares a la clase de servicio prioritaria. De esta forma, en cada ciclo completo de asignación del OLT, todas las ONUs son consultadas para atender a su reserva de intervalos de transmisión para el tráfico de su cola de alta prioridad, mientras que sólo se atiende a la reserva de ventanas para transmitir tráfico de baja prioridad de un subconjunto de ONUs.

Esta estrategia de asignación pretende conseguir la creación de ciclos de espera para el tráfico

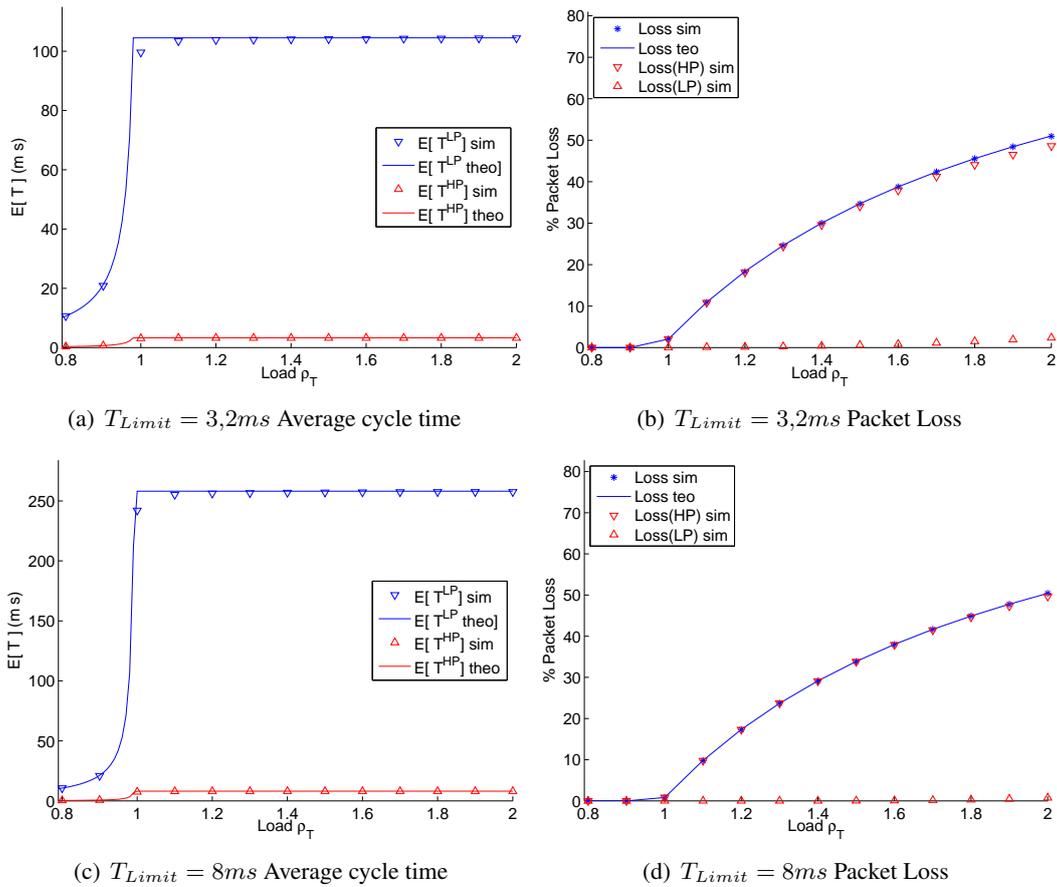


Figura 6.11: Tiempos medios de ciclo y paquetes perdidos con el algoritmo DBA de diferenciación basada en retardo en condiciones de cantidades de tráfico altas: $T_{Limit} = 3,2ms$, tiempo medio de ciclo (a), el número de paquetes perdidos (b); $T_{Limit} = 8ms$, tiempo medio de ciclo (c) y pérdida de paquetes (d).

de alta prioridad más cortos que los ciclos de espera que sufren los paquetes de baja prioridad, respetando una relación de proporcionalidad entre ellos. El trabajo descrito en este capítulo proporciona un modelo para caracterizar dicha relación de proporcionalidad, el cual ha sido validado mediante simulación.

Además, se propone la utilización de este modelo para conseguir una mejora en la diferenciación de servicios en función del retardo, proponiendo asignar una cota máxima a los tiempos de ciclo de la clase HP. Por supuesto, esto supone una degradación del servicio de la clase LP, que puede llegar a sufrir la pérdida de un porcentaje de paquetes, para mantener la garantía de no superar el umbral máximo de retardo fijado para la clase prioritaria. La implementación de este mecanismo es interesante cuando el tráfico se compone de paquetes que provienen de aplicaciones con requisitos de retardo muy ajustados, tales como aplicaciones de difusión multimedia en tiempo real, telefonía IP o plataformas de juego en red en tiempo real, ya que garantizaría un valor máximo para el retardo.

Finalmente, se han mostrado los experimentos de simulación realizados para ilustrar el significado de los diferentes parámetros de diseño del algoritmo propuesto, así como su impacto en las prestaciones del sistema completo.

Cuida de los pequeños gastos;
un pequeño agujero hunde un
barco.

Benjamin Franklin

Capítulo 7

Incremento del ahorro energético de una E-PON

7.1. Trabajos relacionados

Como se ha venido introduciendo en los capítulos anteriores, existe una gran atención actualmente en el ahorro energético en el campo de las redes de comunicaciones. En el tema concreto de las redes ópticas de acceso (FTT-X), el uso de soluciones basadas en PON ya es de por sí una estrategia eficiente en cuanto a que los componentes que las conforman son en su mayor parte pasivos (fibras ópticas, combinadores/divisores pasivos) y no necesitan de alimentación eléctrica. A partir de ahí, el trabajo de optimización del consumo se centra en el diseño de los dispositivos terminales, OLT y ONU, y en los protocolos de comunicaciones y el uso que hacen de los recursos electrónicos y ópticos en dichos terminales.

Una solución de ahorro muy aceptada en el campo telemático es diseñar los protocolos de forma que se pueda permitir tener los dispositivos en un modo de “bajo consumo” o “*sleep*” mientras no se están utilizando, que suele ser una parte importante del tiempo en el que están encendidos [125]. Ya se ha comentado en el capítulo 5 cómo trabajan por ejemplo las redes basadas en *Energy Efficient Ethernet* con dos modos de funcionamiento, por lo que no es necesario extenderse aquí de nuevo con su descripción. Simplemente cabe resaltar que dicho nivel de “bajo consumo” ha de ser suficiente para mantener ciertos servicios del dispositivo activos, como por ejemplo:

- mantener información sobre su estado de sincronización,
- mantener activas las memorias de almacenamiento temporales del sistema,
- atender peticiones para volver a un estado de funcionamiento completo,
- etc.

En trabajos como [126] se propone el uso de la capa de control para realizar una planificación que optimice el consumo energético a la hora de establecer las rutas en la red, sabiendo que los dispositivos pueden estar en cada uno de los dos modos, y no activando más dispositivos de los estrictamente necesarios. Este artículo añade a la idea de eficiencia energética, el uso de GMPLS para la configuración óptima del encaminamiento. Se estudia tanto una solución centralizada como una solución distribuida de anuncio de enlaces en modo *sleep* para la planificación de rutas o para tener que despertarlos si es necesario.

En el caso concreto de las redes PON, pueden encontrarse diversos artículos en la literatura analizando el efecto de implementar en sus dispositivos los dos modos de funcionamiento típicos: *dormido* o *activo*, similar a cómo se han definido para EEE. Por ejemplo en [127] puede encontrarse la descripción de la aplicación de los dos modos de funcionamiento a las ONUs de una TDM-PON, tanto para Ethernet-PON (EPON), como para Gigabit-PON (GPON): un modo de funcionamiento normal mientras que la ONU está transmitiendo, y un modo de funcionamiento que permite ahorrar consumo (modo *sleep*). El artículo describe además las modificaciones que han realizarse en la arquitectura de los dispositivos terminales, en función de la cantidad de potencia que se quiere ahorrar en el modo *sleep* y de la rapidez con la que la ONU puede volver a estar sincronizada y completamente operativa de nuevo.

Parece obvio reseñar que el ahorro de energía que se consiga va a tener como contrapartida la degradación en mayor o menor medida de otros aspectos de las prestaciones de servicio de la red. Es sencillo intuir que los dispositivos terminales deberán diseñarse para aceptar una mayor cantidad de almacenamiento de datos que pudiesen llegar mientras el dispositivo en cuestión está en el modo de bajo consumo, o modificar los protocolos de planificación de las ranuras de tiempo de servicio. Por lo tanto, es necesario primero estudiar hasta qué punto afecta la implementación de políticas de ahorro a los parámetros más críticos, como por ejemplo, el retardo que sufren los paquetes en la PON en este nuevo paradigma.

En [128] a su vez, se describen los modos de funcionamiento con sus respectivas máquinas de

estado para el OLT y las ONUs de redes 10G-EPON y GPON de 10 Gbps (XG-PON). Los autores validan los modelos mediante simulación para diferentes situaciones de carga de tráfico agregado, obteniendo una estimación de los valores de los parámetros de calidad de servicio más importantes como pueden ser el retardo y el jitter.

Dado que en estos sistemas el consumo de energía está relacionado directamente con la carga de tráfico, otra forma de reducirlo es la modificación de la arquitectura electro-óptica de los dispositivos. En el trabajo de [129] se propone realizar modificaciones en la arquitectura física interna del OLT. Dado que es el OLT el que tiene que agregar todo el tráfico de la red y además gestionar la señalización, va a disfrutar de intervalos en modo *sleep* inferiores, con lo que su consumo será mucho mayor que el de las ONUs. En este artículo se presenta una estructura diferente de la arquitectura de tarjetas de la OLT mediante conmutadores (*switches*) optomecánicos 2x2 en cascada para poder apagar cada tarjeta de la OLT, la cual se comunica con un grupo de ONUs de la PON, en función del tráfico recibido. En función de si trabaja en EPON o GPON, se necesitan tiempos de configuración de los switches de menos de 50 ms ó $125\mu s$ respectivamente, consiguiendo alrededor del 20 % de ahorro para un perfil de tráfico diario dado.

De entre los estudios previos consultados se extrae también la idea de que es importante adaptar los tiempos de servicio y la velocidad de transmisión a la carga de tráfico que soporta la red. En [130] los autores proponen una solución adaptada a la carga de tráfico con tres estados: un primer estado que coincide con el modo *sleep* de bajo consumo, y dividir el estado de actividad en dos modos de transmisión, para adaptarse a condiciones de alta o baja carga, reconfigurando la PON para trabajar a 10Gbps ó a 1Gbps respectivamente.

El artículo [131] también aborda el tema de adaptar los tiempos de cada modo a la situación de carga. En este trabajo se detalla la planificación del DBA para los ciclos en el modo *sleep*, adaptando estos tiempos de forma que la ONU pueda estar en el modo de bajo consumo un solo ciclo (*Tciclo*) o varios si es posible. Aunque los autores no describen cómo afecta el modelo a la calidad del servicio (en cuanto a cómo afecta al retardo por ejemplo), el análisis de cómo el algoritmo se adapta a la carga de tráfico está muy detallado y describe todos los parámetros para valores típicos de este tipo de escenarios.

El objetivo de la contribución incluida dentro de este capítulo consiste en proporcionar una solución que permita conseguir que la ONU permanezca en modo *sleep* el tiempo óptimo para conseguir el mayor ahorro energético posible. Para ello se propone la adaptación dinámica de la planificación de las ranuras TDM en función de las condiciones de carga de tráfico instantáneas de

la red, para conseguir que la ONU tenga que esperar en modo de bajo consumo un tiempo mayor de lo normal que sea comparable al tiempo de transmisión de los paquetes recibidos. Esta adaptación dinámica supondrá un tiempo mayor de almacenamiento de los paquetes en la cola de las ONUs, por lo que el modelo introduce una limitación de este tiempo en función del retardo máximo admisible. De esta forma se propone un modelo de diseño que permita garantizar unas mínimas prestaciones de calidad de servicio con el mayor ahorro energético posible. A continuación se detalla la formulación del diseño y los parámetros del modelo, así como los resultados obtenidos mediante la simulación del mismo.

7.2. Formulación del problema

En esta sección se analiza la formulación analítica del problema de ahorro energético para una red PON de N ONUs. En el diseño del modelo analítico del problema, se asumen las mismas condiciones que se utilizaron para el diseño del modelo del capítulo 6 y que están descritas detalladamente en la sección 6.2.1. La descripción de los parámetros más importantes del sistema que se usarán para este modelo se resume en la tabla 7.1:

Parámetro	Descripción
N	Número de ONUs
d	distancia a OLT en km
RTT	retardo en la comunicación ONU-OLT-ONU
C	Velocidad de la línea
T_{ciclo}	Tiempo de ciclo entre dos instantes de transmisión consecutivos de la misma ONU
T_{proc}	Tiempo de procesado del mensaje <i>Report</i>
T_{sync}	Tiempo de re-encendido de la ONU
T_{guard}	Tiempo de guarda entre transmisiones TDM
T_{rep}	Tiempo de transmisión del mensaje <i>Report</i>
$E[B]$	Tamaño medio de paquete
$E[X] = \frac{1}{\mu}$	tiempo medio de servicio por paquete
λ_i	tasa media de tráfico ofrecido por la i -ésima ONU
$\rho_i, i = 1, \dots, N$	Carga media de tráfico ofrecido por la i -ésima ONU
ρ_T	carga de tráfico ofrecida total de la red
T_{wait_i}	Tiempo de espera de la ONU i -ésima en cada T_{ciclo}
T_{b_i}	Tiempo de guarda de la ONU i -ésima respecto de la ONU anterior

Tabla 7.1: Valores de los parámetros de diseño de la simulación de una PON con EEE

De las expresiones y suposiciones descritas en la sección 6.2.1 cabe recalcar las siguientes:

- El análisis corresponde al comportamiento en estado estacionario de algoritmo de asignación de ventanas de transmisión con TDM dinámico, una vez sincronizadas las ONUs y medidos sus RTTs.
- Las ONUs se encuentran todas a la misma distancia, d km, del OLT.
- El tráfico medio ofrecido por cada ONU a lo largo del tiempo es $\rho_i = \frac{\lambda_i}{\mu_i} = \lambda_i E[X]$, $i = 1, \dots, N$. Esto significa que la i -ésima ONU recibe de sus usuarios conectados un tráfico agregado que sigue una distribución de Poisson con una tasa media λ_i paquetes/seg. y un tiempo de servicio que sigue una distribución exponencial, cuyo valor medio es $E[X]$ segundos/paquete.
- La carga de tráfico total ofrecida a la red es ρ_T , esto es, la suma de todas las cargas de tráfico ofrecidas independientemente por cada ONU, ρ_i . Se asume también que se cumple que la carga total ofrecida ha de ser menor que la unidad:

$$\rho_T = \sum_{i=1}^N \rho_i < 1 \quad (7.1)$$

- Usando las asunciones anteriores, se puede calcular la duración del siguiente intervalo de transmisión de la ONU i -ésima como:

$$E[V_i(k+1)] = \frac{1}{\mu} E[N_i(k+1)] = \rho_i E[T_i(k)] \text{ segundos} \quad (7.2)$$

- Y de la misma manera el tiempo de ciclo T_{ciclo} entre dos intervalos consecutivos de la misma ONU:

$$E[T_i(k)] = \sum_{i=1}^N E[V_i(k)] + N(T_{guard} + T_{report}) \text{ segundos} \quad (7.3)$$

- y, finalmente, resolviendo el sistema de ecuaciones anterior para $E[T]$ se obtiene:

$$E[T] = \frac{NT_0}{1 - \sum_i \rho_i} = \frac{NT_0}{1 - \rho_T} \text{ segundos} \quad (7.4)$$

donde T_0 es el tiempo de espera entre las ventanas de transmisión de dos ONUs consecutivas diferentes más el tiempo de transmisión del mensaje de *Report*.

Para poder implementar los modos de funcionamiento de una ONU para el ahorro energético (modo *activo*, que cuando la ONU esté transmitiendo o recibiendo datos, y modo *sleep*, o de bajo consumo), es necesario introducir en la ONU ciertas modificaciones en los bloques electró-ópticos [127], de forma que permitan recibir las indicaciones del OLT para poner el sistema en modo *sleep* o activarla, según sea necesario.

Para la formulación del modelo analítico, el tiempo que la ONU esté en este modo lo denominaremos T_{sleep} . Durante este tiempo, sólo se mantienen activos los bloques imprescindibles (tales como bloques de memoria, relojes del sistema, etc.), con lo que el consumo de potencia es mínimo.

Cuando la ONU se disponga a transmitir datos de nuevo, es necesario volver a poner todo el sistema en su correcto estado de funcionamiento, no siendo posible transmitir hasta que el sistema no esté de nuevo estable y los relojes correctamente sincronizados con la OLT. Dependiendo de la arquitectura electrónica elegida este tiempo de estabilización, que denominaremos T_{sync} puede variar, en función de si la arquitectura mantiene o no el estado de los bloques del sistema, aún cuando se encuentre en modo *sleep*. En [127] se proponen tres tipos diferentes de arquitectura, y tanto los tiempos de estabilización como la potencia consumida en cada una de las opciones se recogen en la tabla 7.2.

La principal diferencia entre las tres arquitecturas que propone el trabajo de [127] está en si la ONU ha de recuperar la sincronización con el reloj de la OLT para recibir las nuevas ventanas de transmisión para el TDM dinámico o no. En la primera opción, la ONU apaga la mayor parte de sus circuitos electrónicos durante el modo *sleep* consumiendo el mínimo de energía previsto para mantener encendida la circuitería imprescindible, pero consume un tiempo muy grande en volver a resincronizarse y estabilizarse. En las últimas dos opciones, el estado de temporización respecto a la OLT se mantiene en la ONU. La diferencia principal entre ellas es que la segunda necesita añadir a la circuitería de una ONU típica, un oscilador local de alta precisión, que permita mantener el estado del reloj del OLT durante el tiempo que la ONU pasa a bajo consumo. La tercera de ellas no necesita introducir bloques nuevos en el dispositivo, pero sin embargo, ha de mantener parte de la circuitería de control encendida para mantener la sincronización con el OLT, lo que supone un consumo de potencia ligeramente mayor a cambio de reducir el coste de fabricación del equipo.

Para el resto de este capítulo supondremos que la tecnología implementada es la correspon-

diente a la tercera opción dado que, a pesar de consumir 200 mW más en el estado *sleep*, es la más sencilla y barata de fabricar adaptando los dispositivos existentes. Por lo tanto el tiempo T_{sync} se reduce al tiempo mínimo necesario para resincronizar la ONU a los tiempos de trama de la OLT. En GPON, dicho tiempo mínimo es de $125\mu s$, mientras que en EPON es menor dado que los tiempos de trama son menores, pero usaremos $T_{sync} = 125\mu s$ como referencia, de forma que el análisis del modelo de reducción de consumo permita una cierta compatibilidad temporal con ambos sistemas que se implementan habitualmente en las redes PON.

Finalmente, para simplificar el cálculo de la potencia consumida, utilizaremos la suposición de que durante el tiempo de sincronización ($T_{sync} = 125\mu s$), la ONU consume la misma potencia nominal que mientras transmite datos, por lo que el tiempo que la ONU está en el modo *activo* será $T_{ON} = T_{sync} + V_i$, siendo V_i la duración de la ventana de transmisión asignada. En la figura 2.3(b) se describen los tiempos comentados en un ciclo de transmisión del canal ascendente de una ONU. A lo largo del tiempo, el cálculo de la potencia media consumida por la ONU puede expresarse como:

$$P_{avg} = P_{ON} \sum_{k=0}^{\infty} T_{ON}[k] + P_{sleep} \sum_{k=0}^{\infty} T_{sleep}[k] \quad (7.5)$$

donde P_{ON} y P_{sleep} son respectivamente los valores nominales de potencia del modo *activo* y del modo *sleep*, que pueden verse en la tabla 7.2 [127].

El diagrama de temporización de una ONU queda descrito como puede observarse en la Fig.7.1

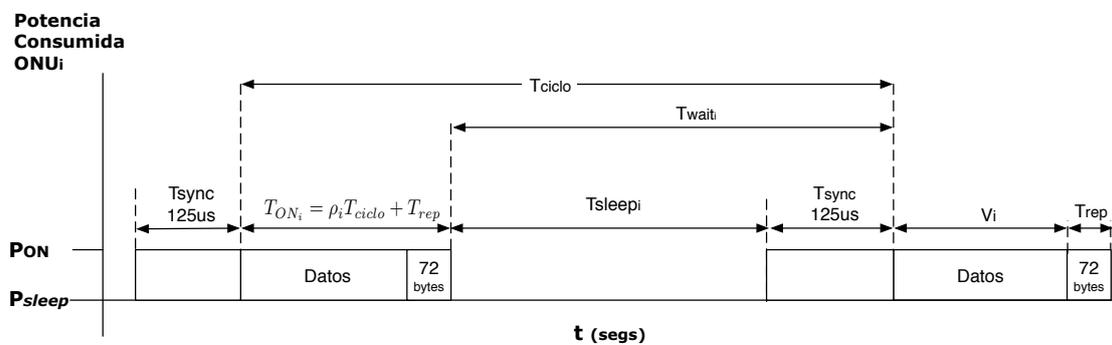


Figura 7.1: Esquema básico de temporización de los cambios de estado de una ONU.

Modo de Funcionamiento	T_{sync}	Potencia en modo <i>sleep</i> (W)	Potencia total ó P_{ON} (W)
Opción 1	5.125ms	0.75	3.85
Opción 2	125 μ s	1.08	3.85
Opción 3	125 μ s	1.28	3.85

Tabla 7.2: Valores de T_{sync} y de potencia nominal consumida en una ONU en cada uno de las opciones de funcionamiento

7.3. Aumento del porcentaje de ahorro energético

El objetivo de la propuesta que se describe a continuación podría resumirse en el hecho de conseguir aumentar el tiempo total en que la ONU está en el modo *sleep* de bajo consumo. Sin embargo, surge el problema de que aquellos paquetes que se reciban en la ONU durante dichos períodos sufrirán un mayor retardo de espera hasta poder ser transmitidos por el canal ascendente, de la misma manera que el OLT no podría enviar tráfico destinado a dicha ONU porque en el modo *sleep* no lo podría recibir.

Por tanto, es necesario mantener una relación de compromiso entre los tiempos de los modos *activo* y de *sleep* para que el retardo medio que sufren los paquetes en la ONU no supere un cierto valor máximo admisible. Por tanto, se propone desde el OLT una gestión de las ranuras de transmisión eficiente que mantenga a la ONU en estado *sleep* durante un tiempo T_{sleep} mayor en cada ciclo de operación con relación al tiempo de actividad T_{ON} .

Para simplificar la formulación, consideraremos que el tiempo de espera T_{sleep} se contabiliza desde el final de la última transmisión, que se corresponde con el paquete de *Report* de la ONU, hasta el comienzo de transmisión del primer paquete del siguiente intervalo de transmisión (véase Fig. 7.2).

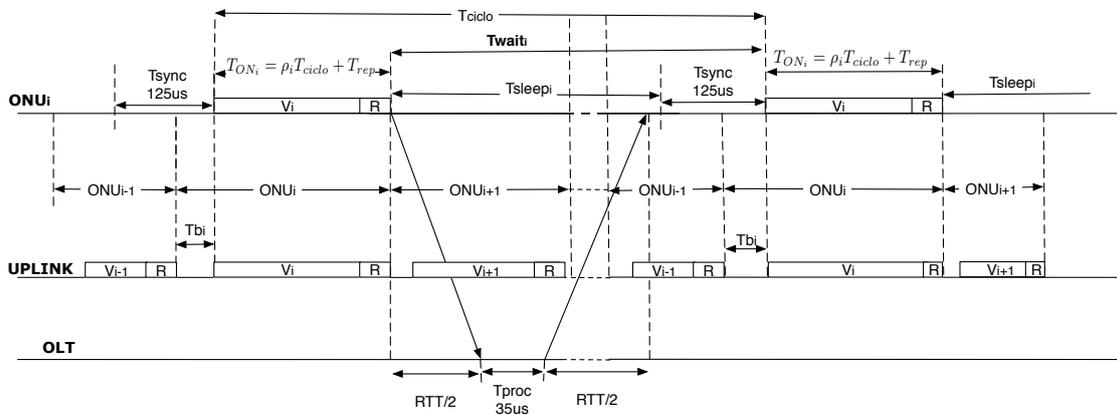


Figura 7.2: Esquema básico de la gestión del enlace ascendente en EPON con EEE. Nomenclatura del modelo.

Para que pueda implementarse el modo de bajo consumo o *sleep* de EEE en EPON, ha de esperarse un tiempo adecuado entre una y otra transmisión de la misma ONU. Durante este tiempo de espera entre dos ranuras consecutivas asignadas a la misma ONU, los paquetes se almacenarán en las colas de espera de la ONU, provocando que en la siguiente reserva, la ráfaga de paquetes sea mayor para aprovechar el tiempo invertido en volver a encender y estabilizar los dispositivos de transmisión de la ONU. Este tiempo, en adelante T_{sleep} , debe de cumplir la condición de que:

1. en general, ha de ser superior al tiempo mínimo necesario para que funcione el "polling" de la PON correctamente.

$$T_{sleep} > T_{min} = RTT + T_{proc} \quad (7.6)$$

2. y, ha de transcurrir el tiempo suficiente para que la ONU se haya estabilizado de nuevo tras el período de *sleep*

$$T_{sleep} > T_{sync} = 125\mu s \quad (7.7)$$

Con esta condición además se garantiza que el tiempo de ciclo (T_{ciclo}) sea suficiente para garantizar el buen funcionamiento del protocolo MPCP. Tal y como se vio en la sección 6.2.1, en la expresión 5.8, dicho tiempo de ciclo depende proporcionalmente del tiempo de guarda (dado que $T_0 = T_{guard} + T_{proc}$ y T_{proc} es un valor constante), T_{guard} es el tiempo que las ONUs han de esperar para evitar colisionar entre sí al usar el canal ascendente. Además, de los resultados numéricos del capítulo anterior se puede observar que los tiempos de ciclo son, en general, pequeños, lo que supone que la ONU va a transmitir muy frecuentemente. Otro efecto de que los tiempos de ciclo sean cortos es que el número de paquetes que habrán llegado a la ONU es pequeño y por tanto el tiempo que están transmitiendo también lo será.

Si dicho tiempo de guarda (T_0) se aumentase, el tiempo de ciclo aumentaría, permitiendo enviar ráfagas de paquetes más largas, pero que compensarían el gasto de potencia que supone activar la ONU (potencia empleada durante T_{sync}). Por lo tanto, el objetivo principal de la propuesta es definir un tiempo de guarda variable que permite modificar el tiempo de ciclo de la ONU y adaptarse al tráfico recibido de forma que se consiga una relación óptima entre los tiempos de T_{sleep} y T_{ON} .

Para ello, se define T_b como el nuevo tiempo de guarda entre períodos de transmisión de

las ONU en el canal ascendente. Este tiempo lo gestionará el OLT de forma que se adapte a la intensidad del tráfico en cada ONU en el momento de asignarle la ventana de transmisión TDM. Como se ha descrito, T_b será siempre mayor o igual que T_0 , lo que supone que los tiempos de ciclo que se obtendrán, que denotaremos por T_{ciclo}^{ext} , también serán mayores o iguales que en el caso de MPCP:

$$T_b > T_{guard} \quad (7.8)$$

$$T_{ciclo}^{ext} > T_{ciclo} \quad (7.9)$$

El tiempo medio de ciclo extendido T_{ciclo}^{ext} de cada ONU podrá de esta manera calcularse como:

$$E[T_{cycle}^{ext}] = \frac{NT_b}{1 - \rho_T} \quad (7.10)$$

Con este esquema de organización de las ventanas temporales de TDM, la ONU va a estar en estado de bajo consumo un tiempo T_{sleep_i} mayor, por lo que al tiempo que pasa entre el inicio de dos ventanas de transmisión se le denominará tiempo de ciclo extendido, $T_{ciclo_i}^{ext}$. Además, se considera que está encendida y consumiendo su potencia nominal máxima un tiempo $T_{ON} = T_{sync} + V_i$ de cada ciclo. El porcentaje de estos períodos de tiempo respecto del ciclo completo y las potencias nominales de la ONU en cada momento permiten calcular la potencia media, P_{avg} , consumida en un T_{ciclo} medio.

$$E[P] = P_{avg} = P_{ON} \frac{T_{ON}}{T_{ciclo}} + P_{sleep} \frac{T_{sleep}}{T_{ciclo}} \quad (7.11)$$

$$P_{avg} = P_{ON} \frac{T_{sync} + E[V_i]}{E[T_{ciclo_i}^{ext}]} + P_{sleep} \frac{T_{ciclo_i}^{ext} - (T_{sync} + E[V_i])}{E[T_{ciclo_i}^{ext}]} \quad (7.12)$$

siendo $E[V_i] = E[\rho_i]E[T_{ciclo_i}^{ext}]$

7.4. Estimación del ahorro energético

Para ofrecer un valor comparativo con el sistema tradicional donde no existen los dos modos de funcionamiento de EEE, se puede calcular el porcentaje de potencia consumida por cada ONU, P_{cons_i} , respecto de la potencia nominal máxima de la ONU, P_{ON} , que se correspondería con el caso de que estuviese encendida la ONU constantemente.

$$P_{cons_i} = \frac{P_{avg_i}}{P_{ON_i}} = \frac{T_{sync}}{E[T_{ciclo_i}^{ext}]} + E[\rho_i] + \frac{P_{sleep_i}}{P_{ON_i}} \left(1 - \frac{T_{sync_i}}{E[T_{ciclo_i}^{ext}]} + E[\rho_i]\right) \quad (7.13)$$

Para simplificar las expresiones se asume que todas las ONUs son iguales y ofrecen la misma carga de tráfico media al sistema, con lo cual $E[\rho_i] = \rho_i = \frac{\rho T}{N}$ y $E[T_{ciclo_i}^{ext}] = T_{ciclo}^{ext}$. De la misma manera las potencias nominales de funcionamiento P_{ON_i} y P_{sleep_i} y la fracción de potencia consumida P_{cons_i} también serán iguales en todas ellas y se puede simplificar la expresión 7.13 como:

$$P_{cons} = \frac{P_{avg}}{P_{ON}} = \frac{T_{sync}}{T_{ciclo}^{ext}} + \frac{\rho T}{N} + R_p \left(1 - \frac{T_{sync}}{T_{ciclo}^{ext}} + \frac{\rho T}{N}\right) \quad (7.14)$$

donde $R_p = \frac{P_{sleep}}{P_{ON}}$ es la relación entre el consumo de potencia de la ONU en el modo de bajo consumo o *sleep* y la potencia consumida en el modo activo habitual.

Reescribiendo la ecuación 7.14, se obtiene la relación de compromiso entre la duración de los ciclos y el porcentaje de potencia consumida deseable respecto del que se obtendría con la ONU encendida continuamente.

$$T_{ciclo}^{ext} = \frac{T_{sync}(1 - R_p)}{P_{cons} - \frac{\rho T}{N} - R_p \left(1 - \frac{\rho T}{N}\right)} \quad (7.15)$$

En este punto es importante reseñar que aumentar el tiempo de ciclo entre dos ranuras TDM de transmisión de la misma ONU conlleva una degradación de prestaciones del servicio en cuanto al retardo medio de los paquetes que la ONU agrega al enlace ascendente, dado que tendrán que esperar, en general, más tiempo para poder ser transmitidos.

7.5. Diseño del tiempo de espera óptimo

Según el estudio de IPACT realizado en [124] y asumiendo por simplicidad del análisis, que el tráfico de entrada sigue una distribución de Poisson, el retardo medio Dq que sufren los paquetes desde que llegan a una ONU hasta que son transmitidos es de $Dq \simeq \frac{3}{2}T_{ciclo}$.

De esta manera, el ahorro energético máximo que puede conseguirse vendrá limitado por la máxima degradación en los valores de retardo que sea admisible. Si el retardo medio de espera en cola en la ONU que se puede permitir para un servicio dado es muy bajo, los tiempos de ciclo medios deberán ser menores. Esto conlleva que la ONU ha de estar más tiempo encendida y menos tiempo en el modo *sleep*, reduciéndose proporcionalmente el ahorro energético que podría conseguirse.

Por ello, se propone la búsqueda del período de espera óptimo que, en función de las condiciones de tráfico de entrada a la ONU, permita conseguir la mayor proporción de ahorro sin superar un valor máximo para el retardo medio que sufren los paquetes antes de ser transmitidos hacia el OLT.

El valor del retardo medio óptimo, estará relacionado con la duración de ciclo deseado para conseguir un cierto nivel de ahorro energético. Esta relación se consigue, como se ha visto en la sección 7.4, porque el consumo de potencia de la ONU está relacionado también, con la duración de los ciclos (ec.7.15). Dado que, a su vez, la duración del tiempo de espera T_b incrementa proporcionalmente la duración de los ciclos según la ec. 7.10, podemos reescribir las ecuaciones involucradas de manera que consigamos relacionar el retardo de los paquetes con el tiempo de espera T_b de acuerdo a las siguientes expresiones:

$$Dq = \frac{3}{2}T_{cycle}^{ext} = \frac{3}{2} \frac{NT_b}{1-\rho_T} \quad (7.16)$$

$$T_b = \frac{2}{3} \frac{Dq}{N} (1 - \rho_T) \quad (7.17)$$

De esta forma se puede diseñar un tiempo de espera T_b que garantice que el retardo medio en cola será como máximo Dq . Sin embargo, el hecho de que el cálculo de T_b se realice en función de ρ_T supone que si las condiciones de la carga de tráfico se conocen a lo largo del tiempo, el retardo

instantáneo que sufren los paquetes en la ONU puede tener mucha varianza respecto del valor medio deseado. Este efecto se puede contrarrestar si T_b se calcula adaptándose a las condiciones de tráfico actuales de la red. De la misma manera que la ONU debe estimar el tamaño de la siguiente reserva TDM para transmitir, el OLT puede realizar una estimación de las condiciones de tráfico actuales en cada ONU, de forma que les asigne tiempos de espera adecuados para reducir la varianza de retardo, asegurando además que las condiciones de consumo de potencia también se adaptarán a dicha variabilidad del tráfico.

Para conseguir esto, se propone obtener T_b en función de un estimador de la carga de tráfico actual $\hat{\rho}$ que a su vez se calcula con una cierta memoria en función del estado de la carga de tráfico en instantes anteriores mediante la expresión:

$$\hat{\rho}_i[k] = \alpha \frac{V_i[k]}{T_{ciclo}[k]} + (1 - \alpha)\hat{\rho}_i[k - 1] \quad (7.18)$$

donde k es el ciclo para el cual se calcula la estimación y α es un parámetro de ajuste del peso de los nuevos valores estimados con respecto a los anteriores, y cuyo valor está comprendido entre 0 y 1. El valor de la carga de tráfico instantánea para la ONU i -ésima en un determinado ciclo de transmisión se puede calcular como la relación entre el tiempo que el OLT asignó a la ONU para la transmisión de paquetes (V_i) en el k -ésimo ciclo respecto de la duración de dicho k -ésimo ciclo (T_{ciclo}):

La carga de tráfico total estimada será la suma de las estimaciones de cada una de las ONUs,

$$\hat{\rho}[k] = \alpha \sum_{i=1}^N \frac{V_i[k]}{T_{ciclo}[k]} + (1 - \alpha)\hat{\rho}[k - 1] \quad (7.19)$$

Un ejemplo de funcionamiento del estimador de la carga de tráfico puede verse en la figura Fig. 7.3, donde la intensidad del tráfico agregado por la ONU sigue un patrón variable en el tiempo (con una variación sinusoidal) y con un valor del parámetro de estimación de $\alpha=0.75$.

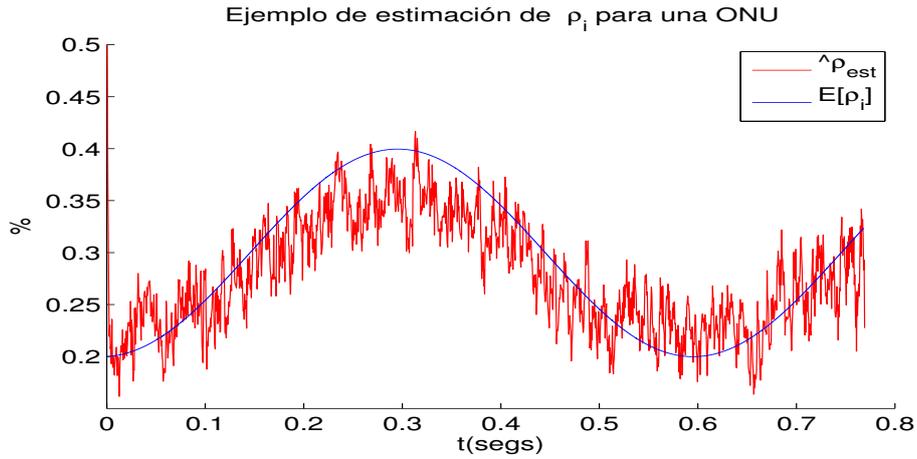


Figura 7.3: Ejemplo de cálculo del tráfico estimado de una ONU con $\alpha = 0,75$

Con este estimador, el cálculo en cada ciclo de T_b sería:

$$T_b[k] = \frac{2}{3} \frac{Dq}{N} (1 - \hat{\rho}[k]) \quad (7.20)$$

Finalmente, podemos usar por lo tanto la relación implícita que existe entre el retardo medio deseado Dq de los paquetes en sus respectivas ONUs y el porcentaje de potencia consumida respecto del máximo nominal P_{cons} , en cada ciclo k para cada ONU mediante la siguiente ecuación, que resume las ecuaciones del análisis anterior:

$$P_{cons_i}[k] = \frac{T_{sync}}{\frac{2}{3}Dq} + R_p + (1 - R_p)\hat{\rho}_i[k] \quad (7.21)$$

Esta expresión a su vez está acotada por el valor de potencia mínimo que se consume en la red debido a los términos que no dependen del objetivo de retardo sino de la carga estimada de la red y de la duración de la estabilización y sincronización a cambiar desde el modo *sleep* al modo activo. El porcentaje mínimo de potencia que se va a consumir en la red viene dado por la ecuación:

$$PC_{min} = \frac{\rho T}{N} + R_p(1 - \frac{\rho T}{N}) \quad (7.22)$$

por lo que podemos reescribir la ecuación 7.21 como:

$$P_{cons_i}[k] = \frac{T_{sync}}{\frac{2}{3}D_q} + PCmin \quad (7.23)$$

La figura 7.4 muestra el conjunto de curvas de consumo de potencia en función de la carga de la red para diferentes valores de retardo medio objetivo en una PON de 16 ONUs y el resto de valores de los parámetros de diseño recogidos en la tabla 7.3.

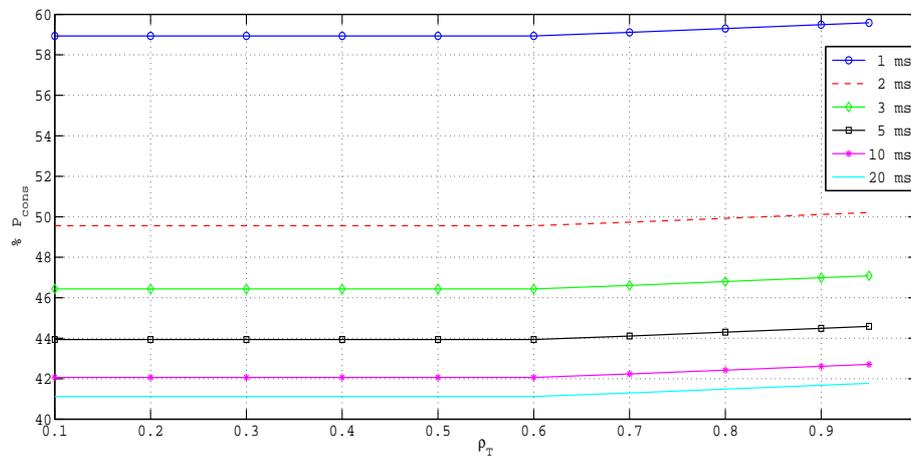


Figura 7.4: Ejemplo de cálculo del porcentaje de potencia consumida respecto del máximo en función de la carga para diferentes valores objetivo del retardo medio máximo

De la ecuación 7.21 se puede obtener una expresión homóloga que relaciona el retardo que se va a conseguir en función del porcentaje medio de potencia que se desea como objetivo de consumo en la red respecto de la potencia de máxima, P_{ON} :

$$D_q = \frac{3}{2} \frac{T_{sync}}{P_{cons_i} - PCmin} \quad (7.24)$$

La figura 7.5 muestra las curvas del retardo medio esperable en la red en función de la carga para el mismo escenario de una red PON de 16 ONUs y parámetros con valores tales como los que se muestran en la Tabla 7.3.

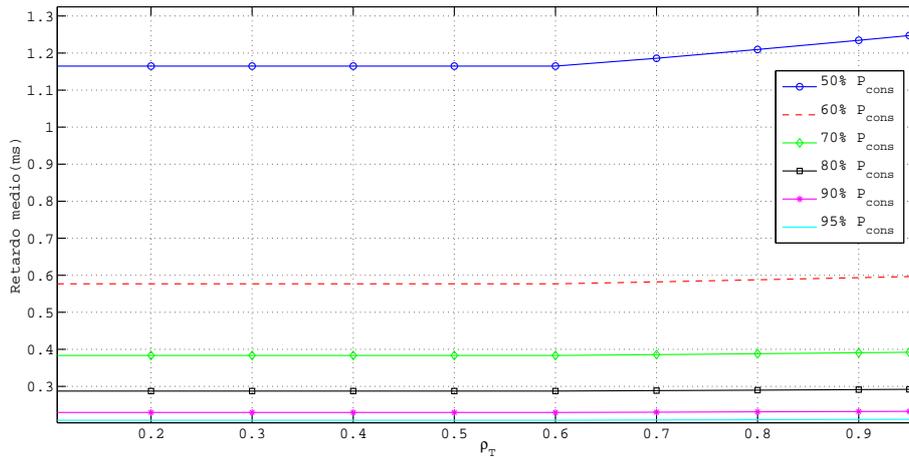


Figura 7.5: Ejemplo de cálculo del retardo medio esperable en función de la carga para varios objetivos de consumo de potencia

7.6. Resultados numéricos obtenidos mediante la simulación de una PON de 8 ONUs

En esta sección utilizaremos en primer lugar los valores de las potencias nominales para los modos de bajo consumo y encendido usados en [127] y que pueden verse en la tabla 7.2.

El resto de parámetros de la simulación pueden verse en la tabla 7.3. El tráfico de entrada se ha generado como un tráfico de Poisson pero con una tasa variable en cada ONU para depurar el correcto funcionamiento del estimador de carga $\hat{\rho}$. De esta forma, las ONUs reciben un tráfico exponencial con una tasa que varía de forma sinusoidal con períodos de variación distintos entre sí, y con valores de amplitud entre $0,4\frac{\rho_T}{N}$ y $\frac{\rho_T}{N}$ para que en el peor de los casos no se supere la carga de tráfico máxima del enlace. El tamaño medio de los paquetes también se ha configurado variable con una distribución exponencial de media $E[B]=1250$ bytes, pero acotada entre un valor mínimo de 64 bytes y un valor máximo de 1518 correspondientes a valores típicos de tramas Ethernet (Fig.7.6). De esta forma el tiempo medio de servicio de un paquete trabajando con una velocidad de línea de $C = 1$ Gbps es de $E[X] = 10\mu s$.

En las figuras 7.7 y 7.8 puede observarse el resultado de la simulación con los parámetros de la tabla 7.3.

La figura 7.7 muestra la potencia que se consume en la red PON a medida que relajamos la condición del retardo medio máximo asumible. A medida que dicho valor de retardo aumenta, es posible mantener la ONU en el modo *sleep* durante más tiempo, de forma que los paquetes esperan hasta que se conforma una ráfaga para su transmisión sin superar el retardo medio máximo

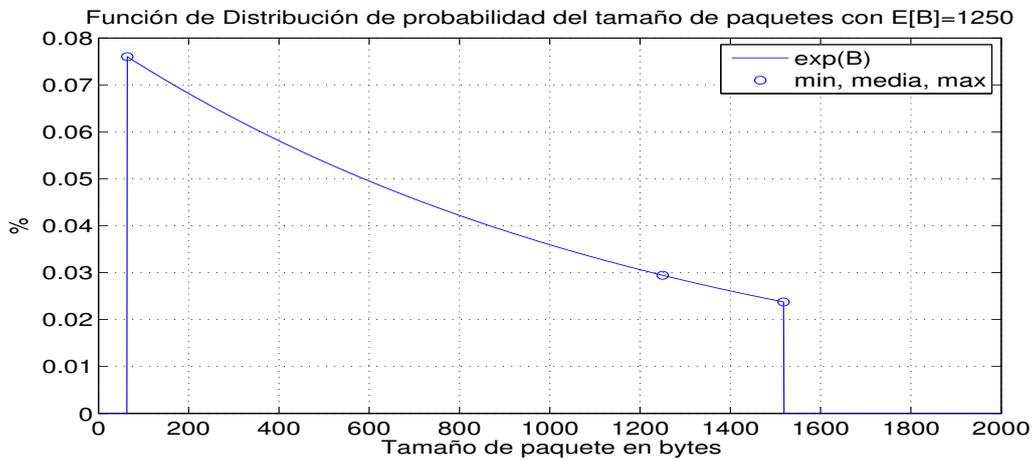


Figura 7.6: Distribución de probabilidad del tamaño de paquete, con tamaño medio de 1250 bytes, mínimo de 64 bytes y máximo de 1518 bytes.

Parámetro	Valor
N	16 ONUs
distancia a OLT	10km
RTT	20 μ s
C	1Gbps
T_{proc}	35 μ s
T_{sync}	125 μ s
T_{guard}	1,5 μ s
T_{rep}	0,448 μ s
$E[B]$	1000 bytes
$E[X]$	10 μ s
	64 < B < 1518 bytes
α	0,25

Tabla 7.3: Valores de los parámetros usados en la simulación de una red PON con EEE

permitido. Las diferentes curvas corresponden a diferentes situaciones de carga total de tráfico agregado en el enlace de subida. Para un retardo medio de 1ms se consigue un ahorro energético de aproximadamente el 50 % del valor máximo nominal de consumo ($P_{ON} = 3,28W$), con un margen de variación del 10 % (desviación típica de la curva para una carga de tráfico de $\rho_T = 0,99$) de la ONU. Si se aumenta el retardo que pueden asumir los distintos servicios de la red, el consumo de energía se acerca al valor de consumo mínimo ($P_{sleep} = 1,28$).

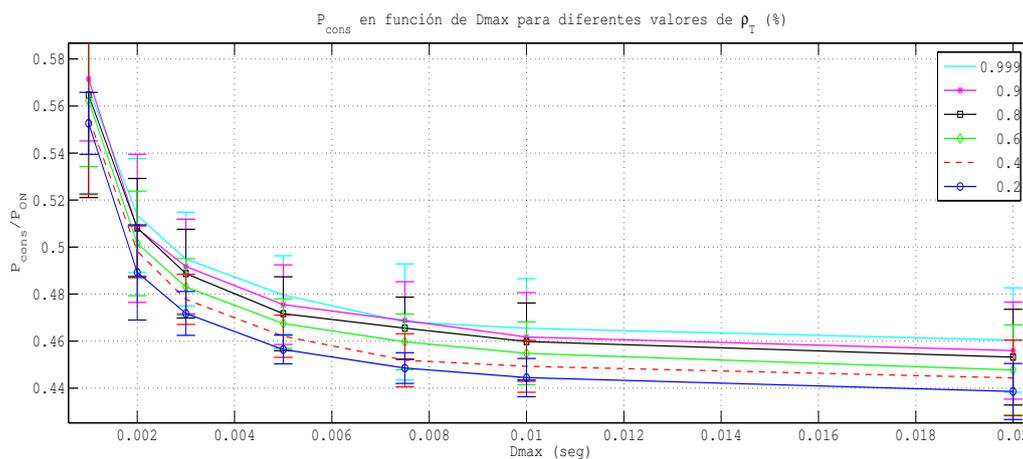


Figura 7.7: P_{cons} en función de D_q para varios valores de ρ_T

La segunda figura (Fig.7.8) muestra los mismos resultados pero desde el punto de vista de la influencia de la carga de tráfico en las ONUs en la potencia consumida. Cada una de las curvas indica la potencia que se consume teniendo por objetivo un valor medio de retardo D_{max} . Mientras la carga de tráfico es cercana o inferior a la carga mínima requerida para que la temporización TDM se vea influida por el tráfico más que por el sistema de *polling* de MPCP (véase la ec. 6.14), el aumento de carga no hace aumentar el tráfico considerablemente. Sin embargo, para cargas altas, el consumo aumenta proporcionalmente al aumento de los tiempos de ciclo de cada una de las ONUs (tal y como ocurría en la Fig. 7.7).

7.7. Conclusiones y trabajos futuros

Como se ha comentado al principio del presente capítulo, el simple hecho de usar redes ópticas pasivas ya supone de por sí un ahorro energético considerable frente a otro tipo de soluciones para la red de acceso, dado que parte del equipamiento de la red no necesita de alimentación eléctrica extra para su correcto funcionamiento. Esto no ocurre sin embargo en los equipos terminales, ya que tanto OLT como ONUs requieren alimentación.

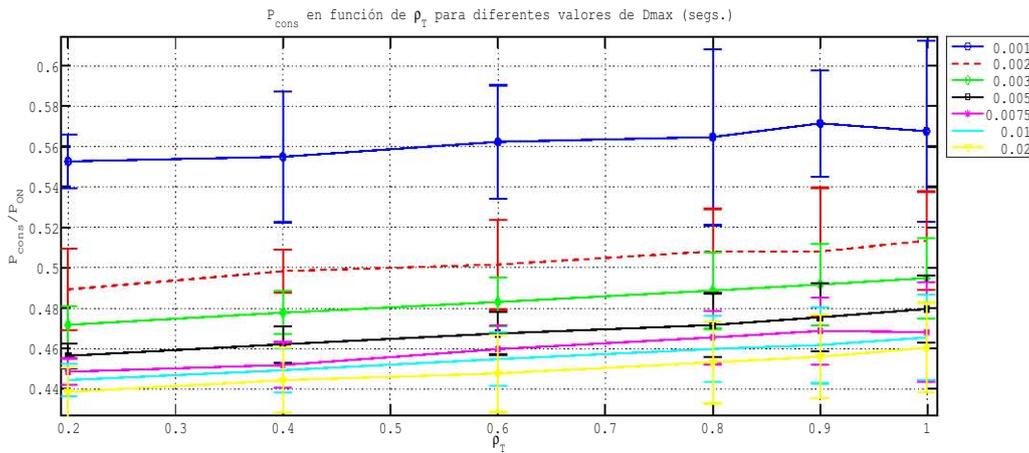


Figura 7.8: P_{cons} en función de ρ_T para varios valores de D_q

La principal conclusión de esta propuesta es que es posible aumentar el porcentaje de ahorro que se consigue en una red PON implementando dos estados de funcionamiento en las ONUs: el modo *activo*, cuando la ONU transmite o recibe datos, y el modo *sleep*, de bajo consumo donde las ONUs tienen la mayor parte de su electrónica desactivada mientras no se esté usando.

Este protocolo de ahorro de energía en dos estados puede conseguir un consumo energético aún menor si se aumentan los períodos en los que la ONU está en modo *sleep*. Con ello se consigue agrupar los paquetes a transmitir en ráfagas mayores, pero cuya duración no suponga un gasto de energía mayor de lo deseado.

Esta modificación se consigue sin embargo, a costa de la degradación en las prestaciones del retardo medio que sufren los paquetes en función de la carga de tráfico que agrega la ONU al canal de subida hacia el OLT. Para que la calidad de servicio (basada en retardo) que proporciona la red no se vea seriamente afectada, es necesario acotar el ahorro de energía que se desea conseguir.

De esta manera, puede diseñarse la red y los tiempos en los que está en cada uno de los dos modos de funcionamiento siguiendo una sencilla regla de compromiso entre la potencia media que se desea que consuma la red frente al máximo retardo medio asumible en función de la carga y del tipo de servicio que se esté ofreciendo.

Parte III

Conclusiones y Trabajos Futuros

Capítulo 8

Conclusiones y Trabajos Futuros

8.1. Conclusiones Finales

Para terminar el presente documento, se recogen en este capítulo las principales conclusiones extraídas de los trabajos de investigación que se han realizado durante la elaboración de esta Tesis Doctoral.

Tras estudiar los modelos y soluciones propuestos en las distintas contribuciones, se llega a la conclusión general de que no se puede conseguir una mejora global de todos los parámetros y prestaciones de la red óptica simultáneamente, dado que la mayor parte de ellos están correlacionados, existiendo multitud de relaciones de compromiso entre la mejora de un grupo de ellos a costa de la degradación de otros.

El resto de conclusiones específicas han sido descritas al final de cada capítulo, pero se destacarán las más importantes a modo de resumen, clasificadas en función de su ámbito de aplicación:

1. Respecto a los estudios realizados acerca de la propagación de errores cabe resaltar las siguientes conclusiones:
 - **Las cadenas de Markov de tiempo continuo (CTMC) son una herramienta válida para el modelado de este tipo de topologías y escenarios de red** para el estudio de la propagación y reparación de errores en redes ópticas gracias a varios factores:
 - el estado de la red puede caracterizarse en función del estado de los nodos y de los eventos que van sucediendo en cada plano,

- es posible modelar por separado los distintos planos de la red. En el caso estudiado, el plano de control y el de datos,
 - y, los cambios de estado responden a la ocurrencia de eventos que pueden modelarse en función de sus tasas y que no suceden simultáneamente.
- **El estudio de las probabilidades resultantes** del modelo CTMC ofrece información relativa a parámetros de fiabilidad que son de interés por el operador de red y por los ISPs, como el tiempo medio entre fallos (MTTF) y los tiempos medios de reparación (MTTR).
 - La **relación entre la fiabilidad deseada a lo largo del tiempo y las tasas de ocurrencia de los eventos de fallo y recuperación** permiten al operador de red y/o al ISP ajustar sus herramientas, mecanismos y protocolos de actuación, adecuando dichas tasas para cumplir con los objetivos previstos.
 - En el caso de la topología en anillo estudiada, la conclusión que destaca en la aplicación del modelo CTMC a este escenario, es la necesidad que tiene el operador de red de usar herramientas de reparación de errores en el plano de control adecuadas, de forma que pueda asegurar que la tasa de reparación sea muy rápida: de una magnitud **al menos dos órdenes de magnitud más rápida** que la tasa de fallos.
 - Por tanto, finalmente se concluye que para poder asegurar la fiabilidad máxima en una red, es más interesante el **esfuerzo en la prevención de la propagación de los errores a nodos adyacentes**, antes que en la reparación posterior de errores graves o que hayan desembocado en un grupo muy grande de nodos afectados.
2. En relación a las contribuciones propuestas para mejorar la QoS ofrecida basada en retardo podemos concluir que:
- dentro de la calidad de servicio, existen diversos parámetros a tener en cuenta, cuya optimización global es un problema complejo ya que normalmente están muy correlacionados entre sí. La solución de otorgar una determinada prioridad superior a un cierto conjunto de tráfico supondrá la degradación de las prestaciones del tráfico de menor prioridad.
 - En anillos ópticos, el número de nodos y la distancia al nodo de destino aumentan proporcionalmente el retardo medio. Usando la **técnica adecuada de acceso al medio compartido**, se pueden conseguir mejoras que garanticen un acceso justo de todos los

nodos para ofrecerles similares oportunidades de acceso y por tanto similares garantías de QoS. Como se ha visto en el capítulo 4, una **técnica de "bypass", que evite la conversión OEO en los nodos permite reducir los valores de jitter**, cuando se usa conjuntamente con un sistema de QoS que otorgue mayor prioridad al tráfico que ya ha sufrido algún bloqueo al intentar atravesar un nodo intermedio.

- En redes PON, el acceso al medio se puede planificar de forma que todas las ONUs dispongan de suficiente tiempo para transmitir el tráfico prioritario de forma garantizada cada vez que son consultadas por la OLT. A continuación, sólo un subconjunto de ellas reciben intervalos garantizados para el resto de su tráfico en función del objetivo de QoS basada en retardo que se quiera conseguir. En conclusión, al disponer de períodos de transmisión más cortos pero más habituales para el tráfico de alta prioridad **se consigue que su retardo y jitter no se vean influenciados por el tráfico de baja prioridad de las ONUs vecinas**.
- En general, **el retardo medio que sufre el tráfico es una de las características que pueden verse degradadas más fácilmente al intentar optimizar otros parámetros de la red**, como se comenta a continuación. Por ejemplo al priorizar el ahorro de energía como métrica en la asignación de rutas.

3. Finalmente, cabe resaltar las siguientes conclusiones relacionadas con las contribuciones propuestas para el ahorro de energía

- La primera conclusión que deriva siempre en este tipo de estudios, y no por su trivialidad carece de importancia, es que **pequeñas modificaciones** realizadas en entornos o protocolos de red concretos que estén siendo muy desplegados **permiten un ahorro global importante**. Este hecho tiene en la actualidad un interés capital en la comunidad, tanto científica como industrial, a nivel internacional.
- Sin embargo, como ya se ha comentado previamente, priorizar el consumo de energía suele conllevar una degradación del resto de parámetros del sistema, principalmente los relacionados con la QoS como puede ser el retardo y su variabilidad (*jitter*). Como se ha visto en el capítulo 5, **optimizar el coste en función del consumo energético global aumenta la carga media de los enlaces de la red y el retardo medio** que sufren los flujos de tráfico.
- Tal y como se ha descrito en el capítulo 7, para cumplir el objetivo de QoS en la red PON, usando dos modos de funcionamiento (activo y *sleep*), es necesario asumir un

consumo de energía mínimo, y viceversa: **marcar una cota máxima de consumo supone una degradación importante de las prestaciones de retardo.**

8.2. Trabajos Futuros

A pesar de que existe un elevado número de contribuciones en la literatura y muchos grupos de investigación trabajando en proyectos relacionados con los temas que abarca este trabajo de tesis doctoral, el campo de investigación en dichos temas no está en absoluto cerrado.

De hecho, la penetración de las redes ópticas en entornos cada vez más cercanos al usuario final está poniendo de actualidad la posibilidad de su optimización en todos los aspectos de la tecnología, desde los dispositivos de comunicación ópticos a los servicios ofrecidos, pasando por todas las capas de diseño: medios de transmisión, diseño de topologías, protocolos de comunicaciones, consumo de energía, etc. tal y como se ha ido comentando en los diferentes capítulos de este trabajo.

Para finalizar, cabe resaltar brevemente los principales temas cuya investigación y/o desarrollo previstos a continuación de la realización de la presente tesis doctoral:

- En relación con el estudio de la fiabilidad y los modelos de propagación de errores, se puede estudiar la manera de aplicar las CTMC a otro tipo de topologías de red, adecuando la nomenclatura de diseño de la máquina de estados, así como la aplicación de estos estudios a otros entornos de "grandes redes".
- En redes ópticas pasivas (PONs) se abren varias líneas de investigación posibles a partir de los resultados de esta Tesis Doctoral:
 1. En primer lugar, se extendería el estudio de las contribuciones relacionadas tanto con el ahorro energético como con la provisión de QoS en PON, utilizando en las simulaciones modelos de tráfico ofrecido más complejos. Dichas contribuciones han sido realizadas teniendo en cuenta diferentes valores de carga de tráfico y de tamaños de paquete, pero siempre siguiendo modelos de tráfico de Poisson, para simplificar el estudio analítico. Pero sería interesante validar estas nuevas prestaciones con tráfico que siga modelos más complejos (tráfico a ráfagas, modelos con distribuciones de cola pesada, tráfico extraído de modelos reales, etc.).

2. Estas contribuciones realizadas en redes ópticas pasivas se han diseñado y analizado intentando generalizar su uso en diferentes tipos de PON. De este modo, se prevé la modificación del simulador utilizado de forma que permita la validación de estos modelos utilizando escenarios de nueva generación como pueden ser redes de largo alcance (*Long Reach PON*) y de anchos de banda del orden de 40-100 Gbps.
 3. Y finalmente, queda como línea futura el estudio analítico de un modelo conjunto que permita optimizar simultáneamente el ahorro de energía y la QoS inter-ONUs en PON.
- En relación con el ahorro de energía, está previsto el estudio y medición de las curvas de consumo en interfaces ópticos de un nodo WDM, permitiendo la actualización de los resultados del capítulo 5 para redes ópticas, teniendo en cuenta estos valores medidos y las posibilidades de agregación de canales que permiten este tipo de nodos, usando modelos de 4 y 10 canales para proporcionar 40 - 100Gbps en redes malladas WDM.

Queda una utópica línea futura final, que a todo investigador de este tipo de entornos le gustaría al menos iniciar, y que consistiría en la implementación real de un prototipo de nodo que incorporase las modificaciones necesarias en el hardware de red para la implementación de los prototipos diseñados.

En el caso concreto de esta tesis, la implementación de la mayor parte de resultados y diseños realizados podría hacerse mediante modificaciones sencillas del firmware y software de control en los nodos de la red, pero ciertos aspectos como los relacionados con el cambio de políticas de acceso en los nodos del anillo propuestas en el capítulo 4, necesitan modificaciones en el hardware pasivo (espejos, fibras) y activo (electrónica de control de conmutadores, conversión OEO y otros elementos) de los nodos, cuyo coste no es asumible en esta fase de la investigación.

Lista de Contribuciones

A continuación se presenta la lista de contribuciones y publicaciones científicas más relevantes realizadas dentro del marco de trabajo de esta Tesis Doctoral:

1. Seoane, I., Rodríguez de los Santos, G., Hernández, J. A., Urueña, M., Romeral, R., Cuevas, Á., Larrabeiti, D. "*Analysis of delay mean and variance of collision-free WDM rings with segment recirculation of blocked traffic*". Photonic Network Communications. 2011-06-01. Ed. Springer Netherlands. Issn: 1387-974X. Pages: 278-287, vol. 21, Issue: 3. URL: <http://dx.doi.org/10.1007/s11107-010-0299-2>

Impacto de la contribución: Esta contribución ha sido publicada en la revista científica Photonic Network Communications, posición 50 (Q3) del área de Telecomunicaciones en el índice de referencia JCR de 2010, con factor de impacto 0.60.

Resumen de la contribución: En las topologías en anillo de redes ópticas WDM (Wavelength Division Multiplexed - Multiplexado por División de Longitud de Onda) basadas en sistemas TT-FR (*Tunable-Transmitter Fixed-Receiver* – Transmisor Sintoni-zable y Receptor Fijo), cada nodo dispone de una longitud de onda dedicada (home channel ó canal local), que ha de ser compartida como canal ascendente por el resto de nodos que se quieren comunicar con él. En esta contribución se propone y se analiza un mecanismo de control de acceso al medio - MAC (*Medium Access Control*) que evita colisiones mediante la re-circulación de tráfico del canal ascendente del anillo en sentido contrario cuando se detecta congestión, usando los segmentos de fibra que componen la red a modo de unidades de almacenamiento. En concreto, esta contri-

bución recoge los resultados del análisis de las prestaciones de este mecanismo MAC implementado en conjunto con dos políticas de atención al tráfico que se refleja:

- Política de “*bypass*”: se aplica al tráfico reflejado una mayor prioridad sobre el tráfico ofrecido a la red por el nodo anterior, que permita su re-envío directo cuando llega al nodo por la fibra del sentido contrario.
- Política de almacenamiento y reenvío (“*store-and-forward*”): en la que el tráfico reflejado llega al nodo anterior y se le aplica la conversión OE (“óptica-electrónica”) necesaria para agregarlo a la cola de espera de dicho nodo junto con su propio tráfico entrante.

2. Seoane, I., Calle, E., Hernández, J.A., Segovia, J., Romeral, R., Vilà, P., Urueña, M., Manzano, M. “*Failure propagation in GMPLS optical rings: CTMC model and performance analysis*”. *Optical Switching and Networking*. Volume 9, Issue 1, January 2012, Pages 39-51, ISSN 1573-4277, 10.1016/j.osn.2011.04.002. URL: <http://www.sciencedirect.com/science/article/pii/S1573427711000269>

Impacto de la contribución: Esta contribución ha sido publicada en la revista científica *Optical Switching and Networking*, posición 32 (Q2) del área de Telecomunicaciones en el índice de referencia JCR, con factor de impacto 0.98

Resumen de la contribución: Esta contribución presenta un modelo de Cadena en Tiempo Continuo de Markov (CTMC - *Continuous-Time Markov Chain*) para caracterizar la propagación de fallos en redes ópticas en anillo basadas en GMPLS (*Generalized MultiProtocol Label Switching*), para proporcionar a los operadores de red y proveedores de servicio de un mecanismo de diseño para ofrecer los objetivos de fiabilidad deseados en sus sistemas.

La contribución distingue entre dos tipos de fallos, los que afectan únicamente al plano de control de la red y los que se producen en cambio en el plano de datos y que afectan a ambos planos en el nodo. Se asume para el modelo que los errores en el plano de datos siguen un modelo de propagación a través de sus vecinos cercanos, similar al modelo de propagación epidémico SID (*Susceptible – Infected – Disabled*).

3. Seoane, I.; Hernandez, J.A.; Reviriego, P.; Larrabeiti, D.; , “*Energy-aware flow allocation algorithm for Energy Efficient Ethernet networks*”. *SoftCOM 2011, 19th International Conference on Software, Telecommunications and Computer Networks (SoftCOM)*, 15-17

Sept. 2011. Split-Hvar (Croatia). URL: <http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=6064376>

Impacto de la contribución: Esta contribución ha sido publicada como un artículo presentado en el congreso IEEE SoftCom'11 celebrado en Split-Hvar y publicada online en el sistema de hemeroteca del IEEE, IEEEExplore.

Resumen de la contribución: El estándar IEEE802.3az-*Energy Efficient Ethernet* (EEE), no ha resultado tan eficiente como se preveía en un principio debido a que los tiempos de transición entre los estados de consumo de energía de los nodos (“activo” y bajo consumo o “sleep”) son muy largos. De hecho, EEE sólo se acerca a la solución óptima cuando la carga de los enlaces es o muy baja o muy alta, pero no en régimen de carga medio. Esta contribución presenta el algoritmo EEE-FA (*Energy Efficient Ethernet Flow Allocation*), orientado a optimizar la asignación y distribución de las demandas, mediante el cálculo de la mejor ruta posible en base al impacto del encaminamiento en el consumo energético de la red. Los resultados de la contribución presentan el ahorro que proporciona EEE-FA frente a un algoritmo de encaminamiento por camino más corto. Sin embargo, este ahorro se consigue a costa de la degradación en el retardo que sufre cada flujo de tráfico y en el incremento de la carga global de la red.

4. Isaac Seoane, José Alberto Hernández, Ricardo Romeral, Manuel Urueña, Eusebi Calle, Marc Manzano. “*A CTMC-based characterisation of the propagation of errors in GMPLS Optical Rings*”. WGN9 (IX Workshop in G/MPLS networks), Girona, Spain. Julio (2010)

Impacto de la contribución: Finalmente, esta contribución sirvió como toma de contacto con el tema de la propagación de errores en anillo, por lo que su temática, es similar a la descrita en la número 2 de esta lista.

Resumen de la contribución: Sin embargo, este artículo presentaba una nomenclatura y descripción del modelo de propagación de errores para la topología en anillo, distintos de los recogidos finalmente en el artículo número 2, y que no podían generalizarse fácilmente para proporcionar los resultados deseados, y que fue desechada en fases posteriores del trabajo de investigación. Pero sirvió para proporcionar, como punto de partida, unos primeros resultados que validasen la corrección de las hipótesis realizadas para abordar la solución final correcta.

Referencias

- [1] Martin Maier. *Optical Switching Networks*. Cambridge University Press, 2008.
- [2] B. Mukherjee. *Optical WDM Networks*. Springer, 2006.
- [3] A. Shami, M. Maier, and C. Assi. *Broadband Access Networks*. Springer, 2009.
- [4] R. Ramaswami, K. Sivarajan, and G. Sasaki. *Optical networks: a practical perspective*. Morgan Kaufmann, Oxford, Elsevier, Inc. ,3rd ed, 2010.
- [5] Tarek S. El-Bawab. *Optical switching*. Springer, 2006.
- [6] P. Cochrane, R. Heckingbottom, and D. Heatley. The hidden benefits of optical transparency. *Communications Magazine, IEEE*, 32(9):90–97, sep 1994.
- [7] M. Yoo, C. Qiao, and S. Dixit. Optical burst switching for service differentiation in the next-generation optical internet. *Communications Magazine, IEEE*, 39(2):98–104, feb 2001.
- [8] J. Teng and G. N. Rouskas. A comparison of the jit, jet, and horizon wavelength reservation schemes on a single obs node. In *In Proc. of the First International Workshop on Optical Burst Switching*, page 2003, 2003.
- [9] P. Newman, W. L. Edwards, R. Hinden, E. Hoffman, F. Ching Liaw, T. Lyon, and G. Minshall. Ipsilon flow management protocol specification for ipv4, May 1996.
- [10] E. Mannie. Generalized Multi-Protocol Label Switching (GMPLS) Architecture. RFC 3945 (Proposed Standard), October 2004.
- [11] R. Romeral. *Mecanismos de protección en escenarios IP-MPLS multidominio*. PhD thesis, Universidad Carlos III de Madrid. Dpto. Ing. Telemática, 2007.

- [12] Ieee standard for information technology–telecommunications and information exchange between systems local and metropolitan area networks–specific requirements part 17: Resilient packet ring (rpr) access method and physical layer specifications. Technical report, 20 2011.
- [13] D. Stamatelakis W. D. Grover. Cycle-oriented distributed preconfiguration: Ring-like speed with mesh-like capacity for self-planning network restoration. *IEEE International Conference on Communications (ICC) '98*, june 1998.
- [14] W. Grover D. Stamatelakis. Ip layer restoration and network planning based on virtual protection cycles. *IEEE JSAC*, 10(18), 2000.
- [15] W.D. Grover J. Doucette, P.A. Giese. Combined node and span protection strategies with node-encircling p-cycles. *Proceedings.5th International Workshop on Design of Reliable Communication Networks, 2005. (DRCN 2005)*, 16-19 October 2005.
- [16] Gangxiang Shen Wayne D. Grover. Extending the p-cycle concept to path-segment protection. In *ICC2003*, 2003.
- [17] D. Schupke C. Gruber. Capacity-efficient planning of resilient networks with p-cycles. In *10th International Telecommunication Network Strategy and Planning Symposium*, 2002.
- [18] A. Autenrieth D. Schupke, C. Gruber. Optimal configuration of p– cycles in wdm networks. In *IEEE ICC 2002*, 2002.
- [19] S. Sarkar, S. Dixit, and B. Mukherjee. Hybrid wireless-optical broadband access network (woban): A review of relevant challenges. *IEEE/OSA Journal of Lightwave Technology, Special Issue on Convergence of Optical Wireless Access Networks*, 25:3329–3340, 2007.
- [20] G. Kramer and G. Pesavento. Ethernet passive optical network (epon): building a next-generation optical access network. *Communications Magazine, IEEE*, 40(2):66 –73, feb 2002.
- [21] Chang-Joon Chae, E. Wong, and R.S. Tucker. Optical csma/cd media access scheme for ethernet over passive optical network. *Photonics Technology Letters, IEEE*, 14(5):711 –713, May 2002.
- [22] E. Wong and M.A. Summerfield. Performance analysis of baseband carrier-sense circuit in optical csma networks. *Photonics Technology Letters, IEEE*, 14(5):708 –710, may 2002.

- [23] B. Skubic, Jiajia Chen, J. Ahmed, L. Wosinska, and B. Mukherjee. A comparison of dynamic bandwidth allocation for epon, gpon, and next-generation tdm pon. *Communications Magazine, IEEE*, 47(3):S40 –S48, march 2009.
- [24] A. R. Dhaini, C. M. Assi, M. Maier, and A. Shami. Dynamic wavelength and bandwidth allocation in hybrid tdm/wdm epon networks. *Journal of Lightwave Technoly*, (25):277–286, 2007.
- [25] A. Banerjee, Y. Park, F. Clarke, H. Song, S. Yang, G. Kramer, K. Kim, and B. Mukherjee. Wavelength-division-multiplexed passive optical network (wdm-pon) technologies for broadband access: a review [invited]. *Journal of Optical Networking*, 4(11):737–758, 2005.
- [26] Ieee standard for information technology- telecommunications and information exchange between systems- local and metropolitan area networks- specific requirements part 3: Carrier sense multiple access with collision detection (csma/cd) access method and physical layer specifications amendment: Media access control parameters, physical layers, and management parameters for subscriber access networks. Technical report, 2004.
- [27] Ieee standard for information technology - telecommunications and information exchange between systems - local and metropolitan area networks - specific requirements part 3: Carrier sense multiple access with collision detection (csma/cd) access method and physical layer specifications amendment 1: Physical layer specifications and management parameters for 10 gb/s passive optical networks. Technical report, 30 2009.
- [28] K. Christensen, P. Gunaratne, B. Nordman, and A. George. The next frontier for communications networks: Power management. *Computer Communications*, 27, No. 18, pp. 1758-1770, December 2004(18):1758–1770, December 2004.
- [29] Barath Raghavan and Justin Ma. The energy and emergy of the internet. *Proceedings of the ACM Workshop on Hot Topics in Networks (HotNets)*, November 2011.
- [30] K. Hinton, J. Baliga, M.Z. Feng, R.W.A. Ayre, and R.S. Tucker. Power consumption and energy efficiency in the internet. *Network, IEEE*, 25(2):6 –12, march-april 2011.
- [31] W. Vereecken, W. Van Heddeghem, M. Deruyck, B. Puype, B. Lannoo, W. Joseph, D. Colle, L. Martens, and P. Demeester. Power consumption in telecommunication networks: overview and reduction strategies. *Communications Magazine, IEEE*, 49(6):62 –69, june 2011.

- [32] J. Baliga, R. Ayre, K. Hinton, and R.S. Tucker. Energy consumption in wired and wireless access networks. *Communications Magazine, IEEE*, 49(6):70–77, june 2011.
- [33] Chengchen Hu, Chunming Wu, Wei Xiong, Binqiang Wang, Jiangxing Wu, and Ming Jiang. On the design of green reconfigurable router toward energy efficient internet. *Communications Magazine, IEEE*, 49(6):83–87, june 2011.
- [34] P. Mahadevan, S. Banerjee, P. Sharma, A. Shah, and P. Ranganathan. On energy efficiency for enterprise and data center networks. *Communications Magazine, IEEE*, 49(8):94–100, august 2011.
- [35] Asfandiyar et al Qureshi. Cutting the electric bill for internet- scale systems. *Proceedings of the ACM SIGCOMM 2009 Conference on Data Communication - SIGCOMM '09. Barcelona, Spain, 2009*.
- [36] Congzheng Han, T. Harrold, S. Armour, I. Krikidis, S. Videv, P.M. Grant, H. Haas, J.S. Thompson, I. Ku, Cheng-Xiang Wang, Tuan Anh Le, M.R. Nakhai, Jiayi Zhang, and L. Hanzo. Green radio: radio techniques to enable energy-efficient wireless networks. *Communications Magazine, IEEE*, 49(6):46–54, june 2011.
- [37] Eunsung Oh, B. Krishnamachari, Xin Liu, and Zhisheng Niu. Toward dynamic energy-efficient operation of cellular network infrastructure. *Communications Magazine, IEEE*, 49(6):56–61, june 2011.
- [38] Yan Chen, Shunqing Zhang, Shugong Xu, and G.Y. Li. Fundamental trade-offs on green wireless networks. *Communications Magazine, IEEE*, 49(6):30–37, june 2011.
- [39] R. Bolla, R. Bruschi, F. Davoli, and F. Cucchietti. Energy efficiency in the future internet: A survey of existing approaches and trends in energy-aware fixed network infrastructures. *Communications Surveys Tutorials, IEEE*, 13(2):223–244, quarter 2011.
- [40] S.J.B. Yoo. Energy efficiency in the future internet: The role of optical packet switching and optical-label switching. *Selected Topics in Quantum Electronics, IEEE Journal of*, 17(2):406–418, march-april 2011.
- [41] Yi Zhang, P. Chowdhury, M. Tornatore, and B. Mukherjee. Energy efficiency in telecom optical networks. *Communications Surveys Tutorials, IEEE*, 12(4):441–458, quarter 2010.

- [42] Ming Xia, M. Tornatore, Yi Zhang, P. Chowdhury, C. Martel, and B. Mukherjee. Greening the optical backbone network: A traffic engineering approach. In *Communications (ICC), 2010 IEEE International Conference on*, pages 1–5, may 2010.
- [43] P. Reviriego, B. Huiszoon, V. López, R.B. Coenen, J.A. Hernández, and J.A. Maestro. Improving energy efficiency in IEEE 802.3ba high-rate ethernet optical links. *Selected Topics in Quantum Electronics, IEEE Journal of*, 17(2):419–427, march-april 2011.
- [44] J. Baliga, R. Ayre, K. Hinton, W.V. Sorin, and R.S. Tucker. Energy consumption in optical ip networks. *Lightwave Technology, Journal of*, 27(13):2391–2403, july 1, 2009.
- [45] R.S. Tucker, R. Parthiban, J. Baliga, K. Hinton, R.W.A. Ayre, and W.V. Sorin. Evolution of wdm optical ip networks: A cost and energy perspective. *Lightwave Technology, Journal of*, 27(3):243–252, feb. 1, 2009.
- [46] IEEE 802.3az Energy Efficient Ethernet Task Force. IEEE 802.3az energy efficient ethernet standard. Draft Amendment to IEEE Std 802.3-2008, 2010.
- [47] K. Christensen, P. Reviriego, B. Nordman, M. Bennett, M. Mostowfi, and J. A. Maestro. IEEE 802.3az: The road to energy efficient ethernet. *IEEE Communications*, 48(11):50–56, November 2010.
- [48] P. Reviriego, K. Christensen, J. Rabanillo, and J. A. Maestro. An initial evaluation of energy efficient ethernet. *IEEE Communications Letters*, 2011.
- [49] James P.G. Sterbenz, David Hutchison, Egemen K. Çetinkaya, Abdul Jabbar, Justin P. Rohrer, Marcus Schöller, and Paul Smith. Resilience and survivability in communication networks: Strategies, principles, and survey of disciplines. *Computer Networks*, 54(8):1245–1265, 2010.
- [50] A. Haider and R. Harris. Recovery Techniques in Next Generation Networks. *IEEE Communications, Surveys and Tutorials*, 9(3):2–17, 2007.
- [51] M. Pickavet, P. Demeester, D. Colle, D. Staessens, B. Puype, L. Depre, and I. Lievens. Recovery in multilayer optical networks. *Journal of Lightwave Technology*, 24(1):122–134, jan. 2006.
- [52] M. Vigoureux, B. Berde, L. Andersson, T. Cinkler, L. Levrau, M. Ondata, D. Colle, J. Fernandez-Palacios, and M. Jager. Multilayer traffic engineering for GMPLS-enabled networks. *IEEE Communications Magazine*, 43(7):44–50, jul. 2005.

- [53] P. Demeester, M. Gryseels, K. Van Doorselaere, A. Autenrieth, C. Brianza, G. Signorelli, K. Van, M. Ravera, C. Brianza Italtel, G. Signorelli Sirti, R. Clemente, M. Ravera Cselt, A. Jajszczyk, D. Janukowicz Itti, G. Kalbe Belgacom, Y. Harada, S. Ohta Ntt, A. G. Rhissa Int, and Correspondence Piet Demeester. Resilience in a Multi-Layer Network. *IEEE Communications Magazine*, 37:70–76, 1999.
- [54] P. Cholda and A. Jajszczyk. Recovery and Its Quality in Multilayer Networks. *Lightwave Technology, Journal of*, 28(4):372–389, feb.15, 2010.
- [55] A. Jajszczyk and P. Rozycki. Recovery of the control plane after failures in ASON/GMPLS networks. *IEEE Network*, 20(1):4–10, 2006.
- [56] Adrian Farrel and Igor Bryskin. *GMPLS: Architecture and Applications*. Morgan-Kaufmann (Elsevier), 2004.
- [57] G. Li, J. Yates, D. Wang, and C. Kalmanek. Control plane design for reliable optical networks. *IEEE Communications Magazine*, 40(2):90–96, Feb. 2002.
- [58] Olivier Crochat, Jean-Yves Le Boudec, and Ornan Gerstel. Protection Interoperability for WDM Optical Networks. *IEEE/ACM Trans. Netw.*, 8(3):384–395, 2000.
- [59] M. Kurant and P. Thiran. Survivable routing of mesh topologies in IP-over-WDM networks by recursive graph contraction. *IEEE Journal on Selected Areas in Communications*, 25(5):922–933, 2007.
- [60] C. Liu and L. Ruan. A new survivable mapping problem in IP-over-WDM networks. *IEEE Journal on Selected Areas in Communications*, 25(3):25–34, 2007.
- [61] K. Thulasiraman, T. Lin, M. Javed, and G.L. Xue. Logical topology augmentation for guaranteed survivability under multiple failures in IP-over-WDM optical networks. *Optical Switching and Networking*, 7(4):206–214, 2010.
- [62] A. Narula-Tam, E. Modiano, and A. Brzezinski. Physical Topology Design for Survivable Routing of Logical Rings in WDM-based Networks. *IEEE Journal on Selected Areas in Communications*, 22(8):1525–1538, oct 2004.
- [63] A. Jaekel, S. Bandyopadhyay, and Y. Aneja. Logical Topology Design for WDM Networks Using Survivable Routing. In *IEEE International Conference on Communications, 2006. ICC '06*, volume 6, pages 2471–2476, jun. 2006.

- [64] Peera Pacharintanakul and David Tipper. The effects of multi-layer traffic on the survivability of IP-over-WDM networks. In *ICC'09: Proceedings of the 2009 IEEE international conference on Communications*, pages 2354–2359, Piscataway, NJ, USA, 2009. IEEE Press.
- [65] K. Lee and E. Modiano. Cross-Layer Survivability in WDM-Based Networks. In *IEEE INFOCOM 2009*, pages 1017–1025, apr. 2009.
- [66] Sheng Huang, Ming Xia, C.U. Martel, and B. Mukherjee. A Multistate Multipath Provisioning Scheme for Differentiated Failures in Telecom Mesh Networks. *Journal of Lightwave Technology*, 28(11):1585–1596, jun. 2010.
- [67] Jing Zhang and Biswanath Mukherjee. A review of Fault Management in WDM Mesh Networks: Basic Concepts and Research Challenges. *IEEE Network*, 18(2):41–48, 2004.
- [68] J. Doucette and W.D. Grover. Capacity design studies of span-restorable mesh transport networks with shared-risk link group (SRLG) effects. In *SPIE Optical Networking and Communications Conference (Opticomm 2002)*, volume 4874, pages 25–38. Citeseer, 2002.
- [69] L. Shen, X. Yang, and B. Ramamurthy. Shared risk link group (SRLG)-diverse path provisioning under hybrid service level agreements in wavelength-routed optical mesh networks. *IEEE/ACM Transactions on Networking (TON)*, 13(4):918–931, 2005.
- [70] D. Coudert, F. Huc, F. Peix, and M.E. Voге. Reliability of connections in multilayer networks under shared risk groups and costs constraints. In *IEEE International Conference on Communications, 2008. ICC'08*, pages 5170–5174. IEEE, 2008.
- [71] Gerard O'Reilly, Ahmad Jrad, Ramesh Nagarajan, Theresa Brown, and Stephen Conrad. Critical Infrastructure Analysis of Telecom for Natural Disasters. In *Telecommunications, Network Strategy and Planning Symposium, 2006. NETWORKS 2006. 12th International*, pages 1–6, 2006.
- [72] Yasuichi Kitamura, Youngseok Lee, Ryo Sakiyama, and Koji Okamura. Experience with Restoration of Asia Pacific Network Failures from Taiwan Earthquake. *IEICE Transactions*, 90-B(11):3095–3103, 2007.
- [73] Michael Lesk. The New Front Line: Estonia under Cyberassault. *IEEE Security and Privacy*, 5:76–79, 2007.

- [74] Reka Albert, Hawoong Jeong, and Albert-Laszlo Barabasi. Error and attack tolerance of complex networks. *Nature*, 406(6794):378–382, July 2000.
- [75] Liang Zhao, Kwangho Park, and Ying-Cheng Lai. Attack vulnerability of scale-free networks due to cascading breakdown. *Phys. Rev. E*, 70(3):035101, Sep 2004.
- [76] Clémence Magnien, Matthieu Latapy, and Jean-Loup Guillaume. Impact of Random Failures and Attacks on Poisson and Power-Law Random Networks. *ACM Computing Surveys 2010*, 2010. (to appear).
- [77] Tomasz Bilski. Disaster’s Impact on Internet Performance—Case Study. In Andrzej Kwiecień, Piotr Gaj, and Piotr Stera, editors, *Computer Networks*, volume 39 of *Communications in Computer and Information Science*, pages 210–217. Springer Berlin Heidelberg, 2009.
- [78] E.G. Coffman, Z. Ge, Vishal Misra, and Don Towsley. Network Resilience: Exploring Cascading Failures within BGP. In *Allerton Conference on Communication, Control and Computing*, October 2002.
- [79] Sebastian Neumayer, Gil Zussman, Reuven Cohen, and Eytan Modiano. Assessing the Vulnerability of the Fiber Infrastructure to Disasters. In *INFOCOM*, pages 1566–1574, 2009.
- [80] S. Neumayer and E. Modiano. Network Reliability with Geographically Correlated Failures. In *Proceedings IEEE INFOCOM, 2010*, pages 1–9, mar. 2010.
- [81] K.K. Nguyen, B. Jaumard, and A. Agarwal. A distributed and scalable routing table manager for the next generation of IP routers. *IEEE Network*, 22(2):6–14, 2008.
- [82] A. Capello, S. Milani, C. Moriondo, G. Rossi, P. Salamandra, M. Perrone, and M. Barone. Non-stop forwarding behaviour and performance in high-end ip routers for isp’s backbone networks. In *5th International Workshop on Design of Reliable Communication Networks, 2005. (DRCN 2005). Proceedings.*, pages 279–285, oct 2005.
- [83] J. Perello, S. Spadaro, J. Comellas, and G. Junyent. An Analytical Study of Control Plane Failures Impact on GMPLS Ring Optical Networks. *IEEE Communications Letters*, 11(8):695–697, aug. 2007.
- [84] O. Komolafe and J. Sventek. Impact of GMPLS Control Message Loss. *Journal of Lightwave Technology*, 26(14):2029–2036, jul. 2008.

- [85] M. Ruiz, J. Perello, L. Velasco, S. Spadaro, and J. Comellas. An analytical model for GMPLS control plane resilience quantification. *IEEE Communications Letters*, 13(12):977–979, dec. 2009.
- [86] Jorge M. Finochietto, Javier Aracil, Ángel Ferreiro, Juan Pedro Fernández-Palacios Giménez, and Óscar González de Dios. Migration Strategies Toward All Optical Metropolitan Access Rings. *J. Lightwave Technology*, 25(8):1918–1930, 2007.
- [87] A. Ganesh, L. Massoulie, and D. Towsley. The effect of network topology on the spread of epidemics. In *INFOCOM 2005, IEEE*, volume 2, pages 1455–1466, mar. 2005.
- [88] Deepayan Chakrabarti, Yang Wang, Chenxi Wang, Jurij Leskovec, and Christos Faloutsos. Epidemic thresholds in real networks. *ACM Trans. Inf. Syst. Secur.*, 10(4):1–26, 2008.
- [89] Marc Barthélemy, Alain Barrat, Romualdo Pastor-Satorras, and Alessandro Vespignani. Dynamical patterns of epidemic outbreaks in complex heterogeneous networks. *Journal of Theoretical Biology*, 235(2):275–288, 2005.
- [90] T. G. Lewis. *Network Science: Theory and Applications*. Wiley Publishing, 2009.
- [91] E. Calle, J. Ripoll, J. Segovia, P. Vilà and, and M. Manzano. A multiple failure propagation model in GMPLS-based networks. *IEEE Network*, 24(6):17–22, November-December 2010.
- [92] I. Chlamtac, A. Fumagalli, L. G. Kazovsky, and P. Poggiolini. A multi-Gbit/s WDM optical packet network with physical ring topology and multi sub-carrier header encoding. In *Proc. of European Conference and Exhibition on Optical Communication*, Sept. 1993.
- [93] M. Ajmone Marsan, A. Fumagalli, E. Leonardi, and F. Neri. R-daisy: an all-optical packet network. In *Proc. European Symposium on Advanced Networks and Services*, March 1995.
- [94] M. Ajmone Marsan, A. Bianco, E. Leonardi, M. Meo, and F. Neri. MAC protocols and fairness control in WDM multirings with tunable transmitters and fixed receivers. *IEEE/OSA J. Lightwave Technology*, 14(6):1230–1244, June 1996.
- [95] M. Ajmone Marsan, A. Bianco, E. Leonardi, A. Morabito, and F. Neri. All-optical WDM multi-rings with differentiated QoS. *IEEE Comm. Magazine*, 37(2):58–66, February 1999.

- [96] I. M. White, M. S. Rogge, K. Shrikhande, and L. G. Kazovsky. A summary of the HORNET project: A next-generation metropolitan area network. *IEEE J. Selected Areas in Communications*, 21(9):1478–1494, 2003.
- [97] K. S. Kim, H. Okagawa, K. Shrikhande, and L. G. Kazovsky. Unslotted optical CSMA/CA MAC protocol with fairness control in metro WDM ring networks. In *Proc. of IEEE Globecom*, volume 3, pages 2370–2374, Nov 2002.
- [98] N. Bouabdallah, A. L. Beylot, E. Dotaro, and G. Pujolle. Resolving the fairness issues in bus-based optical access networks. *IEEE J. Selected Areas in Communications*, 23(8):1444–1457, 2005.
- [99] N. Bouabdallah and H. Perros. Cost-effective single-hub WDM ring networks: A proposal and analysis. *Computer Networks*, 51(13):3878–3901, Sept. 2007.
- [100] Ken Christensen. Green networks: Reducing the energy consumption of networks. <http://www.csee.usf.edu/christen/energy/madridTalk10.pdf>.
- [101] C. Gunaratne, K. Christensen, S. Suen, and B. Nordman. Reducing the energy consumption of ethernet with an adaptive link rate (alr). *IEEE Transactions on Computers*, 57(4):448–4561, April 2008.
- [102] Erol Gelenbe and Simone Silvestri. Reducing power consumption in wired networks. In *ISCIS*, pages 292–297. IEEE, 2009.
- [103] Joseph Chabarek, Joel Sommers, Paul Barford, Cristian Estan, David Tsiang, and Steve Wright. Power awareness in network design and routing. In *In Proc. IEEE INFOCOM*, 2008.
- [104] Erol Gelenbe and Simone Silvestri. Optimisation of power consumption in wired packet networks. pages 717–729, 2009.
- [105] P. Reviriego, J. A. Hernández, D. Larrabeiti, and J. A. Maestro. Performance evaluation of energy efficient ethernet. *IEEE Communications Letters*, 13(9):697–699, Sept. 2009.
- [106] J. C. C. Restrepo, C. Gruber, and C. M. Machuca. Energy profile aware routing. In *In First International Workshop on Green Communications IEEE International Conference on Communications (ICC)*, June 2009; Dresden, Germany.

- [107] IEEE 802.1aq D3.0. Ieee draft standard for local and metropolitan area networks : Virtual bridged local area networks – amendment 9: Shortest path bridging”, June 2010.
- [108] D. Allan, P. Ashwood-Smith, N. Bragg, J. Farkas, D. Fedyk, M. Ouellete, M. Seaman, and P. Unbehagen. Shortest path bridging: Efficient control of larger ethernet networks. *Communications Magazine, IEEE*, 48(10):128 –135, october 2010.
- [109] János Farkas and Zoltán Arató. Performance analysis of shortest path bridging control protocols. In *Proceedings of the 28th IEEE conference on Global telecommunications, GLOBECOM’09*, pages 4191–4196, Piscataway, NJ, USA, 2009. IEEE Press.
- [110] G. Pesavento and M. Kelsey. PONs for the broadband local loop. *Lightwave*, 16(10):68–74, Sept. 1999.
- [111] B. Lung. PON architecture futureproofs FTTH. *Lightwave*, 16(10):104–107, Sept. 1999.
- [112] M.P. McGarry, M. Maier, and M. Reisslein. Ethernet PONs: A survey of dynamic bandwidth allocation (DBA) algorithms. *IEEE Communications Magazine*, 42(8):8 – 15, Aug. 2004.
- [113] G. Kramer, B. Mukherjee, and G. Pesavento. Interleaved polling with adaptive cycle time (IPACT): A dynamic bandwidth distribution scheme in an optical access network. *Photonic Network Communications*, 4(1):89–107, 2002.
- [114] M. Ma, Y. Zhu, and T. H. Chen. A bandwidth guaranteed polling MAC protocol for Ethernet Passive Optical Networks. In *Proc. of IEEE INFOCOM*, volume 1, pages 22–31, San Francisco, CA, USA, 2003.
- [115] L. Zhang, E.-S. An, C.-H. Youn, H.-G. Yeo, and S. Yang. Dual DEB-GPS scheduler for delay-constraint applications in Ethernet Passive Optical Networks. *IEICE Trans. Communications*, E86-B(5):1575–1584, May 2003.
- [116] C. M. Assi, Y. Ye, S. Dixit, and M. A. Ali. Dynamic bandwidth allocation for Quality of Service over Ethernet PONs. *IEEE J. Selected Areas in Communications*, 21(9):1467–1477, Nov. 2003.
- [117] S.-I. Choi and J.-D. Huh. Dynamic bandwidth allocation for multimedia services over Ethernet PONs. *ETRI Journal*, 24(6):465–468, Dec. 2002.

- [118] B. Moon and H. Song. Hierarchical Ethernet-PON design with QoS supporting by eligible reporting. In *Proc. of Asia-Pacific Conf. in Communications*, pages 1–5, Busan, Aug. 2006.
- [119] C.G. Park, D.H. Han, and B. Kim. Packet delay analysis of dynamic bandwidth allocation scheme in an Ethernet PON. *Lecture Notes in Computer Science*, 3420:161–168, 2005.
- [120] Chul Geun Park, Dong Hwan Han, and Kee Wook Rim. Packet delay analysis of symmetric gated polling system for DBA scheme in an EPON. *Telecommunication Systems*, 30:13–34, 2005.
- [121] T. Berisa, Z. Ilic, and A. Bazant. Absolute delay variation guarantees in passive optical networks. *IEEE/OSA J. Lightwave Technology*, (99), 2011.
- [122] F.T. An, Y.L. Hsueh, K.S. Kim, I.M. White, and L.G. Kazovsky. A new dynamic bandwidth allocation protocol with quality of service in ethernet-based passive optical networks. In *Proc. of IASTED Int. Conf. Wireless and Optical Communications (WOC 2003)*, volume 3, pages 165–169, 2003.
- [123] S. Bhatia, D. Garbuzov, and R. Bartos. Analysis of the gated IPACT scheme for EPONs. In *IEEE Proc. of Int. Conf. Communications*, volume 6, pages 2693–2698, 2006.
- [124] B. Lannoo, L. Verslegers, D. Colle, M. Pickavet, M. Gagnaire, and P. Demeester. Analytical model for the IPACT dynamic bandwidth allocation algorithm for EPONs. *OSA J. Optical Networking*, 6(6):677–688, 2007.
- [125] K. J. Christensen, C. Gunaratne, B. Nordman, and A. D. George. The next frontier for communications networks: power management. *Computer Communications*, 27(18):1758 – 1770, 2004.
- [126] I. Cerutti, N. Sambo, and P. Castoldi. Sleeping link selection for energy-efficient gmpls networks. *Lightwave Technology, Journal of*, 29(15):2292 –2298, aug.1, 2011.
- [127] Shing-Wa Wong, L. Valcarenghi, She-Hwa Yen, D.R. Campelo, S. Yamashita, and L. Kazovsky. Sleep mode for energy saving pons: Advantages and drawbacks. In *GLOBECOM Workshops, 2009 IEEE*, pages 1 –6, 30 2009-dec. 4 2009.
- [128] B. Skubic and D. Hood. Evaluation of onu power saving modes for gigabit-capable passive optical networks. *Network, IEEE*, 25(2):20 –24, march-april 2011.

-
- [129] Jingjing Zhang, Ting Wang, and N. Ansari. Designing energy-efficient optical line terminal for tdm passive optical networks. In *Sarnoff Symposium, 2011 34th IEEE*, pages 1 –5, may 2011.
- [130] R. Kubo, J.-i. Kani, H. Ujikawa, T. Sakamoto, Y. Fujimoto, N. Yoshimoto, and H. Hadama. Study and demonstration of sleep and adaptive link rate control mechanisms for energy efficient 10g-epon. *Optical Communications and Networking, IEEE/OSA Journal of*, 2(9):716 –729, september 2010.
- [131] Jingjing Zhang and N. Ansari. Toward energy-efficient 1g-epon and 10g-epon with sleep-aware mac control and scheduling. *Communications Magazine, IEEE*, 49(2):s33 –s38, february 2011.

Acrónimos

AON All Optical Networks

APON ATM PON

ASK Amplitude Shift Keying

ASON Automatized Switched Optical Network

ASTN Automatic Switched Transport Network

BER Bite Error Rate

BPON Broadband PON

CTMC Continuous Time Markov Chain)

DBA Dynamic Bandwidth Allocation

DSL Digital Subscriber Lines

EEE Energy Efficient Ethernet

EEE-FA Energy Efficient Ethernet - Flow Allocation

EPON Ethernet PON

FDL Fiber Optical Delays

FIFO First Input First Output

FPT First-Passage Times

FSC	Fiber Switching Capability
FTTB	Fiber To The Building
FTTC	Fiber to the Curb
FTTH	Fiber to the Home
FWA	Fiber Wireless Access
GMPLS	Generalized Multiprotocol Label Switching
GPON	Gigabit PON
HFC	Hybrid Fiber Coax
IETF	Internet Engineering Tasking Force
ISC	Interfaz Switching Capabilities
ISI	Inter Symbolic Interference
ISP	Internet Service Provider
JET	Just Enough Time
L2SC	Layer 2 Switching Capability
LAN	Local Area Network
LDP	Label Distribution Protocol
LED	Light Emitting Diode
LSC	Lambda Switching Capability
LSP	Label Switched Path
LSR	Label switching routers
MAC	Medium Access Control
MAN	Metropolitan Area Networks
MPλS	Multi-Protocol Lambda Switching

MPCP Multi Point Control Protocol

MPLS Multiprotocol Label Switching)

MTTF Mean Time To Failure

MTTR Mean Time To Repair

NNI Node-to-Node Interface

NRZ Non-Return to Zero

OADM Optical ADD-Drop Multiplexer

OBS Optical Burst Switching

OCDMA Optical Code Division Multiplexing Access

OCS Optical Circuit Switching

OEO Optical-Electrical-Optical Conversion

OLT Optical Line Terminal

ONU Optical Network Unit

OPS Optical Packet Switching

OSK ON-OFF Shift Keying

OSN Optical Switching Networks

OSPF Open Shortest Path First

OUNI Optical User-Network Interface

OXC Optical Cross Connect

PDH Plesiochronous Digital Hierarchy

PML Path Merge LSR

PON Passive Optical Networks

PSC Packet Switching Capability

PSL Path Switching LSR

QoS Quality of Service

ROADM Reconfigurable OADM

RPR Resilient Packet Ring

RSVP-TE Resource Reservation Protocol with Traffic Engineering extension

RTT Round Trip Time

SAN Storage-Area Networks

SCMA SubCarrier Multiple Access

SDH Synchronous Digital Hierarchy

SDM Space Division Multiplexing

SHR Self Healing Rings

SID Susceptible-Infected-Disabled

SLA Service Level Agreement

SNR Signal to Noise Ratio

SONET Synchronous Optical Network

SPB Shortest Path Bridging

SRR Synchronous Round Robin

STP Spanning Tree Protocol

TDM Time Division Multiplexing

TDMA Time Division Multiple Access

TDMC Time Division Multiplexing Capable

TM Terminal Multiplexer

TT-FR Tunable-Transmitter Fixed-Receiver

WAN Wide Area Networks

WDM Wavelength Division Multiplexing

WDMA Wavelength Division Multiple Access

