



UNIVERSIDAD NACIONAL
AUTÓNOMA DE
MÉXICO

UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

**PROGRAMA DE MAESTRÍA Y DOCTORADO EN
INGENIERÍA**

FACULTAD DE INGENIERÍA

**“Análisis y Diseño de Técnicas de Calidad de Servicio en
Redes Inalámbricas de Banda Ancha”**

T E S I S

QUE PARA OPTAR POR EL GRADO DE

MAESTRO EN INGENIERÍA

INGENIERÍA ELÉCTRICA-TELECOMUNICACIONES

P R E S E N T A :

ING: JAVIER ORTIZ CASTRO



TUTOR:

DR. VICTOR RANGEL LICEA

2006

JURADO ASIGNADO:

Presidente: **Dr. Javier Gómez Castellanos**

Secretario: **Dr. Ramón Gutiérrez Castrejón**

Vocal: **Dr. Víctor Rangel Licea**

1^{er}. Suplente: **Dr. Miguel Moctezuma Flores**

2^{do}. Suplente: **Dr. Héctor Benítez Pérez**

Ciudad Universitaria, México, D.F.

TUTOR DE TESIS:

DR. VÍCTOR RANGEL LICEA

FIRMA

AGRADECIMIENTOS

A Dios, por permitirme alcanzar una meta más en mi vida.

A la Facultad de Ingeniería de la UNAM, por permitirme cursar mis estudios de maestría.

Al Consejo Nacional de Ciencia y Tecnología por el apoyo económico durante mis estudios de posgrado.

A la DGAPA por el apoyo recibido a través del proyecto PAPIIT IN 110805.

Al Dr. Víctor Rangel Licea por la paciencia y el apoyo constantes durante la elaboración de este trabajo.

Al Dr. Javier Gómez Castellanos, que junto con el Dr. Rangel han iniciado el grupo de investigación en redes inalámbricas. Gracias por sus enseñanzas.

A los sinodales: Dr. Ramón Gutiérrez Castrejón, Dr. Miguel Flores Moctezuma y Dr. Héctor Pérez Benítez, por el tiempo dedicado para la revisión de este trabajo.

A los profesores del posgrado, en especial a los Doctores Oleksandr Martynyuk, Jorge Rodríguez Cuevas e Ismael Martínez quienes dejaron una huella en mí.

Al Servicio Sismológico Nacional, por las facilidades para la culminación de este trabajo.

Al Dr. Carlos Valdés por creer en mí y por el apoyo recibido para la terminación de esta tesis. Mi más sincero agradecimiento.

A mis compañeros y amigos del SSN: Jorge Estrada, José Luis Cruz, Arturo Cárdenas, Jesús Pérez, José Santiago, Tan Li Yi, Caridad Cárdenas, Iván Rodríguez, Alfredo Yáñez, Enedina Martínez, Ángeles Gutiérrez, Bernardino Rubí y Casiano Jiménez. Gracias por ayudarme a terminar esta tesis.

A los ingenieros Roberto Quaas, Enrique Guevara, Gilberto Castelán, Miguel Franco, César Morquecho y Javier González quienes me motivaron enormemente para iniciar mis estudios de posgrado.

Finalmente quiero agradecer a todas aquellas personas que de una u otra manera me ayudaron durante mi estancia en el posgrado y durante la elaboración de esta tesis. A todos gracias.

Dedicatorias

Este trabajo está dedicado a mi mamá María Isabel Castro Medrano, por su paciencia, amor y cariño durante toda mi vida.

A mi papá y mis hermanos.

A Cris, por su gran amistad.

A todos mis compañeros y amigos de la maestría, a Maribel†, Isaac, Rafa, Meche, Judith, Toño y en especial a Oscar y a Karen, quienes hicieron más agradables las largas horas de estudio durante la maestría.

A mis amigos Jazmín Vázquez, Ana Alarcón, y Miguel Luna por su amistad, paciencia y apoyo.

ÍNDICE

ÍNDICE	i
ÍNDICE DE FIGURAS	iv
ÍNDICE DE TABLAS.....	vi
RESÚMEN.....	1
CAPITULO I	
INTRODUCCIÓN.....	3
1.1 Sistemas de tercera y cuarta generación	3
1.2 Sistemas de acceso inalámbrico de banda ancha	6
1.3 Planteamiento del problema y objetivos.....	8
1.4 Modelado de redes	9
1.5 Estructura de la tesis.....	10
CAPITULO II.....	
ESTADO ACTUAL DE LOS SISTEMAS BWA	11
2.1 Introducción.....	11
2.2 Necesidades de servicio, escenarios para el desarrollo y requerimientos de la arquitectura	12
2.2.1 Servicios BWA y requerimientos de la portadora	12
2.2.2 Escenarios y arquitectura	12
2.3 Retos en redes inalámbricas fijas	13
2.4 Canales inalámbricos de banda ancha	14
2.4.1 Pérdidas en la línea y retardos de dispersión.....	15
2.4.2 Factor K.....	15
2.5 Capa Física, Capa MAC y protocolo de radio enlace.....	16
2.5.1 La capa Física	16
2.5.2 La capa MAC y RLP	19
2.6 Antenas múltiples en BWA.....	20
2.6.1 Ventajas de usar antenas múltiples.....	20
2.6.2 Codificación espacio-temporal y diversidad en la transmisión	21
2.6.3 Multiplexaje espacial	22
2.7 Retos en el futuro y tendencias de la industria.....	23
2.7.1 Retos a futuro.....	23
2.7.2 Estándares	23
2.7.3 Tendencias a futuro	25
CAPÍTULO III.....	
DESCRIPCIÓN DEL PROTOCOLO DE COMUNICACIONES IEEE 802.16.....	26
3.1 Introducción.....	26
3.2 Aspectos de diseño.....	27
3.2.1 Control de acceso al medio (MAC).....	27
3.2.2 La capa física	28

a) 10-66 GHz.....	28
b) 2-11 GHZ	29
3.3 La capa física.....	29
3.4 Detalles de la capa MAC.....	31
3.4.1 Subcapas de convergencia de servicio específico	32
3.4.2 Subcapa de parte común	32
3.4.2.1 Introducción y arquitectura general	32
3.4.2.2 Formatos MAC PDU.....	33
3.4.2.3 Transmisión de MAC PDUs.....	34
3.4.2.4 La capa física y la estructura de trama.....	34
3.4.2.5 Control del radio enlace.....	35
3.4.2.6 Calidad de Servicio (QoS).....	36
3.4.2.7 Solicitud de ancho de banda y asignación de oportunidades de transmisión o grants	38
3.4.2.8 Adquisición del canal.....	41
3.4.2.9 Ranging inicial y negociación de las capacidades básicas de la SS	42
3.4.2.10 Autenticación y registro de la SS.....	43
3.4.2.11 Conectividad IP	43
3.4.2.12 Establecimiento de conexión.....	44
3.4.2.13 Subcapa de privacidad	44
3.4.2.14 Asociaciones de seguridad.....	44
3.4.2.15 Métodos de encriptación	45
 CAPÍTULO IV	
DISEÑO E IMPLEMENTACIÓN DEL MODELO DE SIMULACIÓN DEL PROTOCOLO IEEE 802.16	46
4.1 Introducción.....	46
4.2 Diseño de la Unidad Estación Suscritora (SS)	48
4.2.1 Enlaces.....	48
4.2.2 Descripción de las funciones de los módulos de la estación suscritora ...	49
4.2.3 Módulo de generación de paquetes Src_type_1	49
4.2.3.1 Proceso Bwa_gen_root.....	49
4.2.3.2 Proceso Bwa_gen_src	50
4.2.4 Módulo de eliminación de paquetes Downstream_sink.....	53
4.2.4.1 Proceso bwa_down_sink	53
4.2.5 Módulo para la realización de las funciones MAC	55
4.2.5.1 Proceso bwa_wm_mgr.....	55
4.2.5.2 Proceso bwa_wm_child	59
4.3 Diseño de la estación base (BS).....	63
4.3.1 Descripción de las funciones de los módulos de la estación base	63
4.3.2 Módulo HE_MAC.....	64
4.3.2.1 Proceso bwa_be_mgr	64
4.3.2.2 Proceso bwa_be_txctl.....	67
4.3.2.3 Algoritmo empleado para la asignación de slots en la red BWA (Scheduler)	69

CAPÍTULO V	
ANÁLISIS DEL COMPORTAMIENTO DINAMICO	78
5.1 Escenarios de tráfico.....	78
5.2 Análisis de desempeño	81
5.2.1 Escenario de red para tráfico con servicio UGS y BE.	81
5.2.2 Escenario de red para tráfico rtPS.....	84
5.2.3 Escenario de red para tráfico rtPS y UGS	85
5.2.4 Escenario para tráfico UGS, rtPS y BE	87
5.2.5 Escenario para tráfico UGS, rtPS y BE con pesos ponderados	91
CAPÍTULO VI	
CONCLUSIONES.....	95
6.1 Discusiones finales	95
6.2 Contribuciones	95
6.3 Trabajo futuro	96
6.4 Conclusiones finales	96
REFERENCIAS BIBLIOGRÁFICAS	98
Artículos y libros.....	98
Páginas de internet.....	100
GLOSARIO DE TÉRMINOS Y ACRÓNIMOS.....	102
APÉNDICE A	
MUESTRAS DE VIDEO UTILIZADAS EN LAS SIMULACIONES	105
Descripción general de MPEG 4.....	105
Trazas de los videos empleados durante las simulaciones.....	109

ÍNDICE DE FIGURAS

Figura 1.1. Comparación de las velocidades de las tecnologías 2G, 2.5G y 3G	4
Figura 1.2. Arquitectura de un sistema de acceso inalámbrico de banda ancha.....	7
Figura 2.1. Crecimiento del ancho de banda disponible para usuarios	12
Figura 2.2. Factor K como función de la distancia entre el transmisor y el receptor.....	15
Figura 2.3. Nivel de señal en el receptor de sistemas con a) 1 input 1-output, 2) 1 input-2 output and c) 2 input 3-output	21
Figura 2.4. Niveles de señal e interferencia para un sistema con a) 1-input 1-output, y b) 2 input 3 output.....	22
Figura 2.5. Esquema de un sistema de multiplexaje espacial	23
Figura 2.6. Acceso inalámbrico de banda ancha.....	24
Figura 3.1. Estructura de la subtrama de subida	30
Figura 3.2. Estructura de la subtrama de subida	31
Figura 3.3. Formato TC PDU.....	31
Figura 3.4. Formato del encabezado genérico para un MAC PDU.....	33
Figura 3.5. Tiempo mínimo entre la recepción y la aplicación de un mapa UL-MAP.....	35
Figura 3.6. Transición a un perfil más robusto.....	37
Figura 3.7. Transición a un perfil menos robusto	37
Figura 3.8. Diagrama de inicialización y registro de un nodo con la estación Base	41
Figura 4.1. Niveles jerárquicos de las simulaciones con Opnet	47
Figura 4.2. Nodo correspondiente a la estación suscriptora	48
Figura 4.3. Generador de tráfico	49
Figura 4.4. Proceso Bwa_gen_root	50
Figura 4.5. Proceso Bwa_gen_source.....	51
Figura 4.6. Módulo de eliminación de paquetes Downstream sink	53
Figura 4.7. Máquina de estados del proceso bwa_down_sink	54
Figura 4.8. Módulo MAC.....	55
Figura 4.9. Proceso bwa_wm_mgr	56
Figura 4.10. Proceso bwa_wm_child	60
Figura 4.11. Nodo correspondiente a la estación base	63
Figura 4.12. Módulo HE_MAC.....	64
Figura 4.13. Proceso bwa_be_mgr.....	64
Figura 4.14. Proceso bwa_be_txctl	68
Figura 4.15. Algoritmo de calendarización propuesto para redes IEEE 802.16	70
Figura 4.16. Tipos de retardo de los paquetes de tráfico rtPS durante la transmisión	75
Figura 4.17. Gráfica de los diferentes tipos de retardos de los paquetes de tráfico rtPS ...	75
Figura 4.18. Retardos de acceso promedio para tráfico de voz	76
Figura 4.19. Retardos de acceso promedio para tráfico BE cuando se emplean diferentes seeds en las simulaciones	77
Figura 5.1. Encapsulamiento de tramas de video.....	79
Figura 5.2. Distribución de tráfico de Internet.....	80
Figura 5.3. Configuración de la red con tráfico VoIP e Internet	82
Figura 5.4. Throughput para los diferentes tipos de tráfico	82
Figura 5.5. Retardos de acceso vs tiempo para 111 y 112 SS.....	82

Figura 5.6. Probabilidad acumulada vs retardos de acceso	83
Figura 5.7. Retardos de acceso vs tiempo (135 SS)	83
Figura 5.8. Red con generadores de tráfico de video	84
Figura 5.9. Utilización y throughput del sistema para tráfico para rtPS	84
Figura 5.10. Tasa pérdida de paquetes y retardos de acceso promedio para rtPS.....	85
Figura 5.11. Access delay para 31 y 32 usuarios	85
Figura 5.12. Red con generadores de voz y video	86
Figura 5.13. Throughput para el escenario con generadores de tráfico UGS y rtPS.....	86
Figura 5.14. Pérdida de paquetes de video para una red con tráfico de voz y video	86
Figura 5.15. Retardos de acceso para 60 nodos.....	86
Figura 5.16. Periodos de ocupación del canal ascendente en diferentes intervalos de tiempo	87
Figura 5.17. Paquetes descartados durante la simulación con 60 nodos activos en la red.....	88
Figura 5.18. Resultados del throughput de la red con tráfico UGS, rtPS y BE	88
Figura 5.19. Retardos de acceso para tráfico de tipo BE con 108 y 114 usuarios	88
Figura 5.20. Retardos de acceso promedio para tráfico de tipo BE	88
Figura 5.21. Retardos de acceso para el tráfico rtPS	89
Figura 5.22. Pérdida de paquetes para el tráfico rtPS.....	89
Figura 5.23. Probabilidad acumulada de los retardos de acceso para tráfico rtPS	89
Figura 5.24. Retardos de acceso promedio para el tráfico rtPS	89
Figura 5.25. Retardos de acceso promedio para el tráfico UGS	90
Figura 5.26. Resultados del throughput cuando se usan pesos ponderados	91
Figura 5.27. Retardos de acceso para el tráfico de voz para una red con 108 y 114 usuarios	92
Figura 5.28. Pérdida de paquetes de voz.....	92
Figura 5.29. Retardos de acceso promedio para tráfico de voz	92
Figura 5.30. Pérdida de paquetes de tráfico rtPS.....	93
Figura 5.31. Retardos de acceso promedio para tráfico rtPS.....	93
Figura 5.32. Retardos de acceso de los paquetes rtPS para 96 y 102 nodos activos.....	93
Figura 5.33. Retardos de acceso para 120 y 126 nodos activos.....	94
Figura 5.34. Retardos de acceso promedio del tráfico de Internet	94
Figura A.1. Arquitectura de un sistema MPEG4	106
Figura A.2. Ejemplo de una escena MPEG4	107
Figura A.3. Principio de animación facial utilizado en MPEG4	108
Figura A.4. GOP con $n = 9$	109
Figura A.5. Ejemplo de reordenamiento de un grupo de imágenes para su transmisión ..	109
Figura A.6. Trazas de video empleadas durante las simulaciones.....	110

ÍNDICE DE TABLAS

Tabla 4.1. Características de cada estado del proceso bwa_wm_child	60
Tabla 4.2. Eventos lógicos para el proceso bwa_be_mgr	64
Tabla 4.3. Eventos factibles para el proceso bwa_be_mgr	65
Tabla 4.4. Eventos lógicos para el proceso bwa_be_tx_tcl	67
Tabla 4.5. Estados Idle y TX_Busy del proceso bwa_be_txtcl	68
Tabla 4.6. Eventos factibles para el proceso bwa_be_txctl.....	68
Tabla 4.7. Tabla de respuesta a eventos en el proceso bwa_be_txctl	69
Tabla 4.8. Prioridades de programación en las redes 802.16	73
Tabla 5.1. Parámetros de las simulaciones	78
Tabla 5.2. VoIP CODECS: G.711 y G.723.1	78
Tabla 5.3. Datos de las muestras de las películas empleadas en las simulaciones.....	80

RESÚMEN

La demanda de usuarios que navegan por Internet ha estado creciendo en forma exponencial desde su consolidación a mediados de los 80's. cifras recientes indican que el número de personas en el ciberespacio ya pasó de los 600,00,00 usuarios. Esta gran demanda ha motivado a la industria y a centros de investigación a buscar nuevas tecnologías de comunicaciones de datos para que todos tengamos acceso a este servicio desde cualquier lugar, a la hora que sea y que se pueda transmitir toda clase de servicios y aplicaciones (e.g. voz, video, multimedia, Internet a gran velocidad, etc.).

Una de estas tecnologías de radio frecuencia que pretende satisfacer la necesidad de comunicación que hoy en día se requiere, se le conoce como Acceso inalámbrico de Banda Ancha (Broadband Wireless Access-BWA), la cual queda dentro de la categoría de las redes inalámbricas de cobertura metropolitana (Wireless-MAN). A diferencia de las redes celulares actuales (GSM y CDMA), las redes BWA tienen un ancho de banda muy superior con velocidades de transmisión de datos por arriba de los 100 Mbps.

En la actualidad existen dos propuestas diferentes para redes BWA: ETSI BRAN e IEEE 802.16. La primera propuesta (ETSI Broadband Radio Access Networks ó HIPERLAN), es un protocolo europeo, el cual se pretende que opere en la banda de los 5 GHz y 17 GHz, y lograr velocidades de transmisión de datos de 25-54 Mbps (para la banda de 5GHz y de 155 Mbps a 17 GHz).

La propuesta del protocolo 802.16 operará en a banda 2.5-11 GHz para sistemas fijos y móviles con línea de vista (o LOS por sus siglas en inglés) y 11-66 GHz para sistemas fijos solamente sin LOS. Este protocolo es mucho más robusto que el protocolo ETSI BRAN, al ser una propuesta que se deriva del estándar americano de televisión por cable para la transmisión de datos (DOCSIS: Data Over Cable Service Interface Specification), el cual cuenta con 6 tipos diferentes de reservación (o Calidad de Servicio QoS) y modulación (QPSK, 16-QAM y 64 QAM), para la transmisión de servicios en tiempo real (e.g. Voz sobre IP, video telefonía) y la transmisión de tráfico en ráfagas (e.g. video conferencias, tráfico web a gran velocidad, etc.). Las velocidades de transmisión del protocolo IEEE 802.16 se espera que oscilen entre los 33 Mbps y 134 Mbps aproximadamente, y se tiene pensado que lleguen hasta los 310Mbps en la segunda etapa de su desarrollo, cuando se disponga de sistemas radioeléctricos que utilicen esquemas de modulación de orden superior.

El estándar 802.16 fue aprobado en 2002 y desde entonces ha estado en un proceso de constante actualización.

En este trabajo se presenta el diseño y el análisis de desempeño de una técnica de asignación de slots para la transmisión de datos (denominado Scheduling algorithm). El algoritmo propuesto está basado en el protocolo MAC 802.16 y se centra en el canal ascendente, el cual es el factor limitante en las redes BWA y es crítico en el aprovisionamiento de servicios. Aún cuando el protocolo IEEE 802.16 propone diversas clases de servicio para diferentes tipos de aplicaciones, no sugieren como asignar el ancho de banda para satisfacer a aquellos servicios con requerimientos de servicio crítico, tales como, voz y flujo de audio y video.

Este algoritmo combina diferentes técnicas de calidad de servicio para satisfacer las diferentes demandas de tráfico que tienen tanto una tasa de bits constante como variable.

Las simulaciones realizadas estuvieron basadas en un modelo de programación basado en OPNET Modeler. Los resultados obtenidos a través de las simulaciones demuestran que este algoritmo puede proveer servicios en tiempo real con retardos de acceso muy cortos aún durante los periodos de congestión de la red.

CAPITULO I

INTRODUCCIÓN

1.1 Sistemas de tercera y cuarta generación

Durante los últimos años uno de los temas más importantes en el área de la computación y las comunicaciones, ha sido la tecnología inalámbrica. En este tiempo la tecnología ha atraído a muchos usuarios y ha sufrido numerosos cambios, incluyendo la conectividad a Internet. Sin embargo estos cambios parecerán insignificantes en comparación con los cambios que se sucederán en los próximos años.

Además, parece ser que la tecnología inalámbrica ha alcanzado un punto en el que tanto los fabricantes como los investigadores ya se están preparando para el siguiente nivel. Se espera que el siguiente nivel de tecnología inalámbrica ofrezca un mayor ancho de banda, mayor seguridad y confiabilidad, haciéndola mas adecuada para multimedia, comercio electrónico, videoconferencia y otras aplicaciones avanzadas. Esas aplicaciones incluirán video en demanda, comercio electrónico móvil y navegación en Internet inalámbrica.

A finales de la década de 1970 y principios de 1980, las comunicaciones inalámbricas comenzaron a hacerse más comunes. La tecnología de esta primera generación era totalmente analógica, basada en circuitos y adecuada únicamente para comunicación por voz. En los últimos años, los dispositivos inalámbricos comerciales han usado una tecnología conocida como de segunda generación, la cual es digital, basada en circuitos, de banda angosta, adecuada para voz y para una transmisión de datos limitada.

Recientemente apareció la siguiente generación llamada 3G, adecuada para voz y aplicaciones que requieran la transmisión de gran cantidad de datos, incluyendo multimedia y comercio electrónico. 3G está asociada principalmente con redes móviles, sin embargo puede usarse con redes inalámbricas fijas, tales como LMDS (Servicio Local de Distribución Multipunto) y MMDS (Servicio Multicanal de Distribución Multipunto).

3G promete velocidades de transmisión de hasta 2.05 Mbits por segundo en aplicaciones estacionarias, 384 kbits por segundo para usuarios móviles a velocidades bajas y 128 kbits por segundo para usuarios que viajan dentro de vehículos. 3G es considerablemente

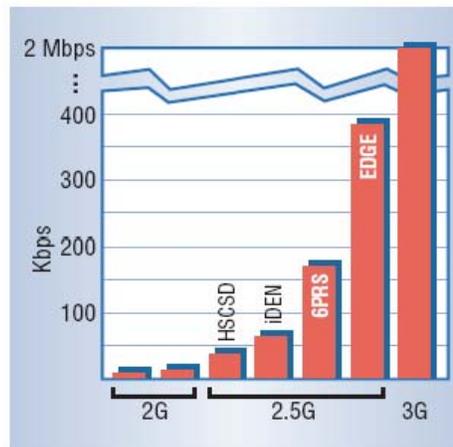


Figura 1.1. Comparación de las velocidades de las tecnologías 2G, 2.5G y 3G.

más rápida que las tecnologías 2 y 2.5G (Ver Figura 1.1).

La tecnología de 3G abarca 3 estándares principales: W-CDMA (Wide-Band Code-Division Multiple Access), CDMA2000, y TD-CDMA (Time-Division CDMA). Estos estándares tienen algunos aspectos en común, sin embargo no son completamente compatibles a nivel de interfase.

Los estándares básicos de 3G han sido desarrollados principalmente por el sector privado más que por organizaciones encargadas de establecer estándares. Sin embargo, la Unión Internacional de Telecomunicaciones (ITU) ha adoptado el IMT-2000 para estandarizar formalmente los servicios de 3G inalámbricos y así ofrecer un conjunto de servicios consistente a través de todo el mundo proporcionando las bases para desarrollar nuevas mejoras al sistema.

A diferencia de 3G, la tecnología WLAN como ETSI's, HiperLan (High Performance radio LAN), y la tecnología líder del mercado IEEE 802.11b (también llamada Wi-Fi) proporciona servicios de red a través de redes de tipo corporativo.

Además de las redes 3G existen las redes de datos inalámbricas, tales como WLAN (Wireless Local Area Network). Sin embargo, las redes 3G no compiten directamente contra las redes WLAN, de hecho se puede decir que las redes LAN y las redes móviles son servicios complementarios destinados a diferentes mercados y a diferentes ambientes. Sin embargo, algunos analistas de la industria dicen que muchos usuarios potenciales de 3G que tiene acceso a WLAN preferirían tener servicios inalámbricos a través de ésta última. La principal ventaja de las redes WLAN sobre 3G es que el costo de instalación es bajo. Sin embargo la tecnología WLAN no es la más adecuada para una cobertura amplia, de hecho su uso es para interiores más que para exteriores.

A pesar de que en Japón desde el año 2001 se implementó una red 3G de cobertura nacional y de que en otros países apenas se está iniciando la implementación de esta tecnología, se espera que este tipo de redes continúe implementándose en otros países durante los próximos años. No obstante, muchas empresas ya están pensando en la

siguiente generación de comunicaciones, la cual será conocida como 4G, y de la cual se espera que ofrezca un servicio global de *roaming* a través de múltiples redes inalámbricas y móviles, por ejemplo de una red celular a una red basada en satélite o a una red de cobertura local de banda ancha inalámbrica.

Con esta característica, los usuarios tendrán acceso a diferentes servicios, mayor cobertura, y la conveniencia de un solo dispositivo y una sola cuenta con las ventajas de acceso total y un mayor acceso inalámbrico más confiable aún si fallase una o más de las redes. Las redes de 4G también soportarán la interoperabilidad de IPv6, para un acceso a Internet móvil sin mantener un enlace con una estación en particular y tasa de hasta 50-100 Mbps y aún más.

Algunas de las diferencias de 4G con respecto a 3 G, son las siguientes:

- | | |
|---|---|
| • 3G | • 4G |
| • Compatible con 2G. | • Extiende la capacidad de 3G en un orden de magnitud.. |
| • Redes de conmutación por circuito y por paquetes. | • Redes conmutadas únicamente por paquetes. |
| • Combinación de elementos existentes y mejorados. | • Todos los elementos son digitales. |
| • Velocidades de hasta 2 Mbps. | • Mayor ancho de banda (hasta 100 Mbps). |

Dentro de esta cuarta generación se encuentran las tecnologías de acceso inalámbrico de banda ancha (BWA). Aún cuando la tecnología inalámbrica 4G ya es una realidad aún quedan por resolver algunos de los principales requerimientos técnicos de estos sistemas, como por ejemplo:

- Acceso
- Handoff
- Coordinación de recursos para añadir nuevos usuarios
- Soporte de multicasting
- Soporte de calidad de servicio
- Seguridad inalámbrica y autenticación
- Fallo en la red, respaldo
- Establecer cuotas y mecanismos de cobranzas

Se espera que la implementación y penetración de sistemas 4G que ayude a cerrar la brecha en atención médica, educación, información y otras áreas y contribuir sustancialmente a la solución de problemas ambientales en una sociedad avanzada. Específicamente en una sociedad penetrada por los sistemas 4G la atención médica y el

diagnóstico remoto serán muy comunes, la revisión médica hecha por especialistas y la prescripción de medicamentos será posible en casa y en áreas rurales, todo esto basado en la tecnología para la transmisión de imágenes de alta resolución y cirugía remota, se podrán realizar hospitales virtuales sin médicos residentes. Los cuidados médicos también pueden enfatizarse, por ejemplo, un individuo puede llevar consigo sensores en la muñeca, del tamaño de un reloj y esos datos transmitirse constantemente a un hospital diariamente, y con base en esos datos hacer un diagnóstico, el cual será reenviado al individuo. En la educación, la enseñanza a distancia puede mejorarse utilizando técnicas multimedia, entre lugares remotos y las grandes ciudades. Esta tecnología puede ayudar a la preservación del medio ambiente ya que se pueden utilizar sensores equipados con circuitos (chips) para la comunicación inalámbrica y de esta manera monitorear la flora, y la fauna.

1.2 Sistemas de acceso inalámbrico de banda ancha

Con la desregulación del sistema telefónico en muchos países, en la actualidad a los competidores de la compañía telefónica arraigada con frecuencia se les permite ofrecer voz local y servicio de alta velocidad de Internet. Ciertamente hay mucha demanda. El problema es que el tendido de fibra óptica, cable coaxial ó incluso cable de par trenzado categoría 5 a millones de casas y oficinas es extremadamente costoso, es por ello que ha habido un gran incremento en la demanda de servicios de acceso inalámbrico de banda ancha (BWA). Estas tecnologías BWA están emergiendo de manera muy rápida mostrando varias ventajas. Estas incluyen una implementación rápida, alta escalabilidad, bajos costo de mantenimiento y mejoramiento, además de una inversión granular de acuerdo a los requerimientos del mercado.

El rápido incremento en la demanda de acceso a Internet de alta velocidad y servicios de multimedia para pequeños usuarios ha creado una demanda de última milla a servicios de acceso inalámbrico de banda ancha (BWA).

Al mismo tiempo, con el rápido crecimiento de los servicios inalámbricos de datos y las aplicaciones multimedia, se espera que BWA provea servicios a diferentes clases de tráfico con diferentes requerimientos de calidad de servicio.

Un sistema fijo BWA incluye al menos una estación base (BS) y una o más estaciones suscriptoras (SS). La estación base es un nodo central y las estaciones suscriptoras son nodos remotos colocados a diferentes distancias de la estación base. El canal de bajada ó *downstream*, se define como la dirección en que fluyen los datos desde la BS hasta la SS, es decir, se trata de un canal *broadcast*, mientras que el canal ascendente ó *upstream* es un medio compartido de acceso múltiple. En el canal de subida, el tiempo es dividido en ranuras denominadas *mini-slots*, los cuales proporcionan acceso múltiple por división de tiempo (TDMA); mientras que el canal de bajada se utiliza un esquema de multiplexaje por división de tiempo. Cada SS puede enviar voz y datos utilizando interfaces comunes, como teléfono, ethernet, video, VoD y otros servicios con diferentes requerimientos en cuanto a calidad de servicio (Ver Figura 1.2).

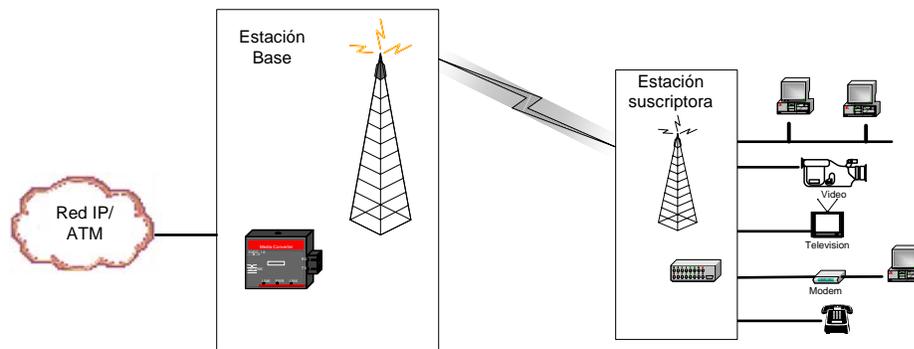


Figura 1.2. Arquitectura de un sistema de acceso inalámbrico de banda ancha.

En un inicio, cada desarrollador de tecnología BWA, como los LMDS, diseñaba su propio sistema. Esta falta de estándares significaba que el hardware y el software no se podían producir en masa por lo que los precios eran altos y la aceptación baja. Sin embargo, la industria se dio cuenta de que tener un estándar, era el elemento clave que faltaba, por lo que se le pidió a la IEEE que formara un comité compuesto por personal tanto de compañías clave como de centros de investigación para redactar el estándar. El grupo de trabajo creado con el nombre de 802.16 tiene por objetivo “desarrollar estándares y recomendar prácticas para respaldar el desarrollo y optimización de sistemas de acceso inalámbrico de banda ancha fijos”. El trabajo se inició en julio de 1999 y el estándar final se aprobó en abril de 2002. Oficialmente el estándar se llama “Air Interface for Fixed Broadband Wireless Access Systems” (Interfaz de Aire para Sistemas Fijos de Acceso Inalámbrico de Banda Ancha), también conocido como WiMAX (Worldwide Interoperability for Microwave Access ó Interoperabilidad Mundial para Acceso por Microondas). Al igual que otros estándares 802, el 802.16 estuvo influido fuertemente por el modelo OSI incluyendo las subcapas, terminología, primitivas de servicios y más. Justo un año más tarde, en marzo de 2003, se ratificó una nueva versión, el 802.16a, y fue entonces cuando WiMAX empezó a cobrar relevancia como una tecnología de banda ancha inalámbrica. También se pensó en enlaces fijos, pero llega a extender el rango alcanzado desde 40 a 70 Km operando en la banda de 2 a 11 GHz, parte del cual es de uso común y no requiere licencia para su operación. Es válido para topologías punto a multipunto y, opcionalmente para redes en malla, y no requiere línea de vista. Emplea las bandas de 3.5 GHz y 10.5 GHz, válidas internacionalmente, que requieren licencia (2.5-2.7 en EE.UU.) y las de 2.4 GHz y 5.725-5.825 GHz que son de uso común y no requieren disponer de licencia alguna.

El pasado 7 de diciembre de 2005, el IEEE aprobó el estándar del WiMAX móvil, el 802.16e, que permite utilizar este sistema de comunicaciones inalámbricas con terminales en movimiento. Muchos fabricantes de hardware y operadores estaban esperando a esta decisión para empezar a desplegar redes de WiMAX. Ahora ya pueden hacerlo.

Lo que ocurría en la práctica es que pocos se atrevían a invertir en WiMAX bajo el único estándar aprobado hasta ahora, el 802.16d, que sólo sirve para aquellos terminales que están en un punto fijo. Ahora ya saben qué especificaciones técnicas debe tener el hardware del WiMAX móvil, con lo que es posible diseñar infraestructuras mixtas fijo-

móviles. En Corea se ha materializado las ventajas de un WiMAX móvil trabajando en 2,3 Ghz y se le ha acuñado el nombre de WiBRO (Wireless Broadband).

1.3 Planteamiento del problema y objetivos

Como se ha mencionado anteriormente, el desarrollo de nuevas tecnologías trae consigo el planteamiento de una serie de problemas que se tienen que resolver para lograr dar un servicio adecuado a cada aplicación, uno de estos problemas es el relacionado con la calidad de servicio. La calidad de servicio se refiere al manejo adecuado de los recursos de una red con el fin de cumplir con los niveles de desempeño que satisfagan las expectativas del usuario. Por ejemplo si un usuario desea tener acceso a una red el 99.9 % del tiempo con un round-trip-delay de 80 ms ó menos, el proveedor del servicio deberá reservar los recursos suficientes (ancho de banda a través de los dispositivos de acceso, como routers y switches) para garantizar que la red satisfará aquellas expectativas.

Un manejo efectivo de las colas es fundamental en muchos esquemas de calidad de servicio propuestos para redes. Un manejador de colas (Puede ser hardware, software ó firmware, dependiendo del esquema) consiste de un conjunto de 3 a 8 colas, un programador de paquetes, y una política para descartar paquetes. Cada clase de servicio usualmente tiene una cola y todas las colas envían paquetes hacia el mismo enlace de transmisión. La tarea del programador de colas es segregar los paquetes de tal manera que los paquetes más sensibles (como voz sobre IP) sean expedidos, los paquetes que son sensibles a las pérdidas (contraseñas) no sean indebidamente descartados y todos los demás paquetes sean transmitidos de acuerdo a la prioridad de clase. La navegación en Internet y el correo electrónico no son sensibles a los retardos ó pérdidas, pero la navegación es un proceso interactivo y debe tener una prioridad mayor que el correo electrónico.

Cuando llega un paquete nuevo a una fila a una velocidad menor a la que es expedido, la cola normalmente permanecerá vacía. Naturalmente, cuando los paquetes arriban más rápido (conforme aumenta la tasa de arribos), la cola se puede llenar. Si esto ocurre, los paquetes enviados comenzarán a perderse porque el número de paquetes ha excedido la capacidad de la cola. El tiempo que le toma transmitir un paquete es proporcional a su longitud, de manera que un paquete que llega justo detrás de un paquete grande, tardará el mismo tiempo que si llegara justo detrás de 20 ó 25 paquetes de menor tamaño.

Los programadores de paquetes ó “*schedulers*” regulan el retardo en las colas a través de diferentes disciplinas, como primera entrada-primera salida. Sin embargo, aún cuando los programadores se aseguran de que los paquetes de las colas de menor prioridad no pasen desapercibidos, no pueden evitar las pérdidas cuando llegan más paquetes de alta prioridad de los que puede acomodar. Con el fin de regular las pérdidas, el programador de colas utiliza también una política para descartarlas.

Un aspecto importante que se debe hacer notar, es que los manejadores de colas no toman la alta prioridad como “no descartable”; por ejemplo un paquete de voz deja de ser útil si se sobrepasa su tiempo de retardo, y por consiguiente el programador de colas deberá descartarlo.

El protocolo 802.16 define cuatro clases de servicio, sin embargo, no define cómo asignar las reservaciones a cada uno de ellos y cómo expedirlos en el programador, lo cual se ha dejado a los fabricantes de equipos para que puedan diferenciar sus productos y ofrecer soluciones adaptadas a diferentes entornos de uso.

De acuerdo a lo anterior, los objetivos de este trabajo son los siguientes:

- Proponer un esquema de calidad de servicio para el protocolo 802.16 capaz de satisfacer los requerimientos de diferentes tipos de tráfico cumpliendo con los siguientes puntos:
- Combinar eficientemente la solicitud de un ancho de banda determinado y su asignación.
- Elaborar diferentes escenarios para determinar el tamaño óptimo de los paquetes de datos de las diferentes clases de tráfico.
- Determinar el retardo óptimo de los paquetes de diferentes clases de servicio.

1.4 Modelado de redes

La simulación de modelos de redes, es un método muy popular para el análisis del desempeño de redes. Generalmente, hay 2 formas de simular una red: a través de un modelo analítico y a través de una simulación en computadora. El primero se refiere a un análisis matemático que caracteriza a la red como un conjunto de ecuaciones. La principal desventaja es la sobre simplificación de la red y la incapacidad de simular la naturaleza dinámica de una red. Así, el estudio de un sistema complejo siempre requiere de un paquete de simulación de eventos discretos, que pueda calcular el tiempo que debería ser asociado con los eventos en una situación real. El software de simulación es una herramienta de gran ayuda especialmente para las redes actuales que poseen arquitecturas y topologías muy complejas.

La elaboración de un modelo de simulación para este protocolo no es una tarea fácil de resolver, por lo que es necesario utilizar un simulador de redes, que cuente con una extensa librería de funciones y módulos de recepción y transmisión de mensajes. El simulador elegido para esta tarea es el software Opnet (Optimized Network Engineering Tool) en su versión 11, el cual proporciona un entorno de desarrollo comprensible, para la especificación, simulación y análisis del desempeño de redes de comunicaciones.

El siguiente paso en este trabajo será la elaboración del modelo de la red BWA, que cuente con generadores de tráfico con diferente prioridad, para a continuación comenzar con los análisis de los diferentes esquemas de QoS.

Una vez que se cuente con un modelo apropiado de red, se elaborará un *scheduler*. Para probar la eficiencia de este *scheduler* se elaborarán diferentes escenarios. En cada uno de estos escenarios se simularán los diferentes tipos de servicio con diferentes tamaños de paquetes de datos y se analizará el desempeño de la red.

Finalmente se hará una comparación de los resultados obtenidos en cada una de las simulaciones con el fin de determinar qué combinación de esquema de QoS junto con el tamaño de los paquetes de datos y los tiempos de retardo demuestran un mejor

desempeño para este tipo de redes.

1.5 Estructura de la tesis

En el capítulo 2 se presenta una descripción del estado del arte de los sistemas inalámbricos de banda ancha.

En el capítulo 3 se describe el funcionamiento del protocolo IEEE 802.16. Se menciona la tanto la operación en la capa física como en la capa MAC. Se describe la arquitectura general del sistema, el formato de las tramas, inicialización del sistema, etc.

El capítulo 4 describe el modelo de red utilizado en las simulaciones. Se menciona tanto la estructura de la red como el funcionamiento de cada uno de los elementos que la componen. Adicionalmente se explica la manera en la que opera el algoritmo implementado, el cual se programó dentro de los procesos.

El capítulo 5 muestra los resultados obtenidos durante las simulaciones. Se presentan las gráficas de los resultados y la interpretación de cada una de ellas con el fin de entender mejor el desempeño de la red.

En el capítulo 6 tenemos la discusión final de los resultados obtenidos con las simulaciones. Asimismo se mencionan algunas de las mejoras que podrían hacerse al algoritmo implementado.

ESTADO ACTUAL DE LOS SISTEMAS BWA

2.1 Introducción

En los últimos años ha habido un rápido crecimiento en la demanda de acceso de alta velocidad a Internet y servicios de voz tanto por parte de pequeñas empresas como de zonas residenciales. Esto ha creado una demanda por servicios de banda ancha de última milla. Las velocidades típicas de un canal de banda ancha para zonas residenciales y pequeñas oficinas u hogares están entre 5 y 10 Mbps para el enlace descendente y entre 0.5 y 2 Mbps para el canal ascendente. Esta asimetría proviene de la naturaleza y predominio del tráfico web. Los servicios de voz y de videoconferencia exhiben un tráfico simétrico. Aunque es difícil predecir cómo evolucionarán los servicios de Internet y por tanto el tráfico resultante, se sabe que la demanda de altas velocidades y la calidad de los servicios de última milla en banda ancha ciertamente se incrementarán en un futuro cercano.

Los servicios de banda ancha se ofrecen actualmente a través de sistemas como xDSL, cable y sistemas inalámbricos de banda ancha (BWA). Cada una de estas tecnologías tiene diferente costo, desempeño, ventajas y desventajas. Actualmente los sistemas del cable y ADSL están siendo empleados en gran escala, sin embargo, BWA está emergiendo como una tecnología de acceso con varias ventajas. Algunas de estas ventajas, con respecto a estas tecnologías son que la distancia no es una limitante debido a que no es necesaria la instalación de cables, la alta escalabilidad y los bajos costos de mantenimiento y actualización. A pesar de esto todavía quedan varios aspectos que es necesario resolver para que los sistemas BWA tengan una buena penetración en el mercado, como por ejemplo, la eficiencia espectral, la escalabilidad de red, una instalación sencilla de las antenas para que la puedan realizar los usuarios y la operación confiable del sistema cuando las antenas no tienen línea de vista (NLOS). Los servicios inalámbricos en la banda de 24-48 GHz como los servicios LMDS (Local Multipoint Distribution Services) son los más adecuados para oficinas ubicadas en grandes edificios, donde se puede tener línea de vista por encima de los árboles de la zona. A continuación se describen las características de los sistemas que operan en las bandas libres y menores a 3 GHz (Servicios de distribución multipunto MDS, servicios de distribución

multicanal multipunto MMDS y los sistemas de comunicaciones inalámbricas WCS), donde es factible la operación NLOS y la penetración a través de follaje. En la Figura 2.1, se muestra el crecimiento del ancho de banda disponible para los sistemas de banda ancha inalámbricos.

2.2 Necesidades de servicio, escenarios para el desarrollo y requerimientos de la arquitectura

2.2.1 Servicios BWA y requerimientos de la portadora

Los servicios típicos de BWA incluyen acceso a Internet, servicio de voz multilínea, audio y video en tiempo real. Por tanto, se requiere el manejo de diferentes clases de tráfico y la capacidad para manejar Calidad de Servicio (QoS).

Entre los requerimientos que debe satisfacer un proveedor de servicios de BWA se incluye el satisfacer las regulaciones de la FCC en cuanto a la emisión de potencia y a la interoperabilidad de radio, escalabilidad usando una arquitectura celular donde se pueda dividir la célula para incrementar el throughput por kilómetro cuadrado, bajos costos en los equipos del usuario y en la infraestructura, gran cobertura y capacidad por celda con el fin de reducir aún más los costos de infraestructura, y finalmente la portabilidad.

2.2.2 Escenarios y arquitectura

A continuación se discuten 3 escenarios de operación diferentes: superceldas, macroceldas y microceldas.

Superceldas. En Con este escenario, se puede cubrir una gran área de servicio con un radio de hasta 30 millas (48 km). La antena situada en la estación base transceptora (ó BTS por sus siglas en inglés) se instala típicamente a una altura de 1000 pies como

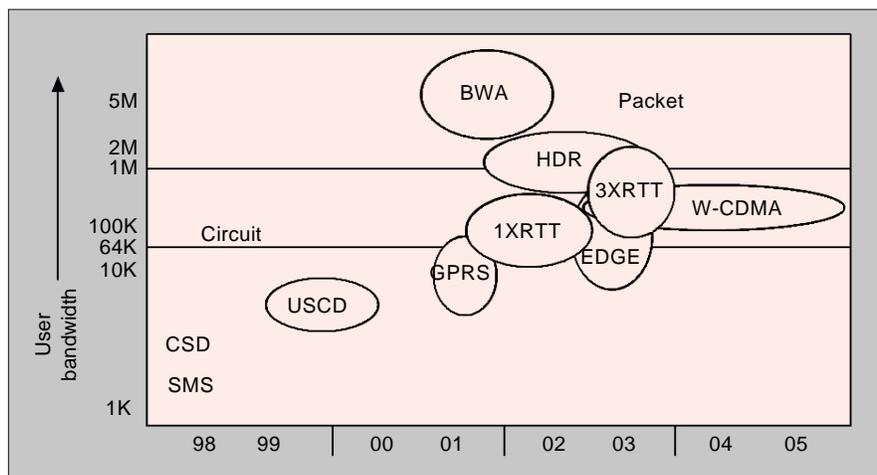


Figura 2.1. Crecimiento del ancho de banda disponible para usuarios.

máximo (304.8 m), y en cada receptor es necesaria la instalación de una antena de alta ganancia con línea de vista hacia la estación base. Esta es una configuración para una célula única, y por tanto no es escalable. Es posible el reuso de frecuencias con una misma polarización en una misma celda si se sectoriza la celda. Debido a la transmisión con línea de vista, pueden obtenerse valores de la relación de la señal portadora a ruido (C/N) de hasta 30 dB, lo cual hace posible el uso de técnicas de modulación de alto orden. Desafortunadamente no se puede garantizar una gran cobertura debido a la necesidad estricta de que exista línea de vista entre las antenas.

Macro celdas. Las macro celdas utilizan típicamente la arquitectura celular con reuso espacial de frecuencias entre células. La altura de la antena de la estación base es significativamente más baja que en el caso de la supercelda, típicamente entre 50 y 100 pies. En este caso las antenas de los suscriptores se pueden colocar cerca del nivel de las azoteas. Si la antena de la estación base transceptora se encuentra a una altura muy baja, se pueden inducir severas pérdidas y la pérdida de la línea de vista. Con este tipo de configuración es posible tener células con un radio de hasta 5 millas (8km). Debido a la propagación sin línea de vista (NLOS) y a la interferencia co-canal ó CCI (*co-channel interference*) de otras celdas, se manejan niveles significativamente más bajos en la relación señal a ruido de la portadora (C/N) y de la portadora a interferencia (C/I) comparados con aquellos de las superceldas. Por lo anterior, se deberán utilizar modulaciones de menor orden. En esta configuración también se deben utilizar antenas direccionales en la estación remota, es decir como parte del equipo de los suscriptores. Esta arquitectura es escalable tanto en capacidad como en cobertura. La gran área de cobertura que se obtiene en esta configuración se logra gracias a que es posible la propagación de la señal aún sin línea de vista.

Microceldas. Las microceldas son similares a las macroceldas, con la diferencia de que se utilizan celdas mucho menores (típicamente celdas con un radio de 1 milla). La antena en la estación base, ó BTS, es más baja que en la macrocelda, típicamente debajo del nivel de los edificios cercanos y puede tener de 20 a 40 pies de altura (6 a 12m). Esta arquitectura soporta portabilidad, por tanto, las antenas en la estación suscriptor son omnidireccionales y trabajan en interiores. La pequeñez del tamaño de las celdas proporciona suficiente margen al enlace para proveer una buena cobertura en interiores.

2.3 Retos en redes inalámbricas fijas

A continuación se discutirán los principales retos en las redes inalámbricas fijas y se discutirán las principales diferencias entre las redes de banda ancha fijas y las redes inalámbricas móviles actuales. Se puede decir que esencialmente la calidad y la velocidad de transmisión son significativamente más altas en el caso de las redes fijas que en el caso de las redes móviles. Dado que la potencia de transmisión se mantiene fija y se aumentan los requerimientos para obtener una alta tasa de transmisión de datos, el cálculo de enlace decrece en 15 dB. Los requerimientos para obtener una alta calidad, incrementan la necesidad de márgenes de desvanecimiento, en las relaciones C/N y C/I, en aproximadamente 15 dB en cada una. Tomando en cuenta que el uso de antenas direccionales proporciona una ganancia de aproximadamente 15 dB en el cálculo de enlace, se obtiene una desventaja aproximada del enlace de 15 dB en contra del ruido y

una desventaja de 15 dB contra la interferencia CCI. La pérdida de 15 dB contra el ruido significa una menor cobertura ó una disminución en el radio de la célula (1/5 del radio de la célula), mientras que la disminución en contra de la interferencia CCI indica que se requiere un factor de reuso de frecuencias mucho mayor (20-30 en lugar del factor de 3 utilizado en redes móviles) y por consiguiente una disminución de 1/6 en la capacidad de la celda. En consecuencia, se hace necesario el uso de nuevas y sofisticadas tecnologías para tanto para la capa física como para la de radio enlace con el fin de mantener la cobertura y mantener un factor de reuso de frecuencia de 3. El uso de antenas múltiples, que se discutirán más adelante, proporciona ventajas significativas, en términos del enlace, contra el ruido y la interferencia CCI, por tanto, parece ser la mejor solución para satisfacer estos requerimientos. Una buena cobertura repercute en el costo de la red ya que se reducen los costos de instalación. Otro factor, que influye notablemente en la economía, es el hecho de contar con capacidad extra en la célula, debido a que se retarda la necesidad de dividir las celdas.

La eficiencia espectral de una red inalámbrica se mide en bits por segundo por Hertz por celda (BHC). La eficiencia espectral se puede incrementar a través del reuso de frecuencias y técnicas de modulación de alto orden. Sin embargo, el reuso de frecuencias en un ambiente multicelda provoca interferencia CCI. En un ambiente de una (super) celda el reuso de frecuencia en ángulo es fuente de CCI y depende de la magnitud de los lóbulos laterales de las antenas de la estación base y de la dispersión de los sectores de reuso. Si la interferencia CCI se trata como ruido blanco aditivo Gaussiano, la fórmula de Shannon para el límite teórico en BHC puede ser escrito como:

$$BHC = \frac{L}{mK} \log_2 \left(1 + \frac{C}{N+I} \right)$$

donde K es el factor de reuso espacial, L el factor de reuso angular, m un factor de encabezado para contabilizar el exceso de ancho de banda y la banda de guarda, y $C/(N+I)$ es la razón de la señal portadora a la señal interferencia más ruido. En un sistema macro/micro celda K es igual al tamaño del cluster y $L=1$. En los sistemas supercelda $K=1$ y L es el número de veces que un canal es reusado angularmente. Reduciendo K en los sistemas macro/micro celda ó incrementando L en los sistemas supercelda, se debería incrementar la eficiencia espectral BHC. Sin embargo, al hacer lo anterior simultáneamente se incrementa la interferencia CCI y por consiguiente se decrementa la relación $C/(N+I)$, teniéndose que reducir el orden de modulación. En la práctica la mejor relación entre K, L y $C/(N+I)$ depende de varios factores como la tasa de error propuesta, las condiciones de propagación, la razón C/N, los lóbulos laterales de la antena y los esquemas de diversidad. Los valores típicos BHC de los esquemas celulares actuales como GSM están entre 0.2 y 0.3. Para los sistemas fijos de BWA, los requerimientos de BHC están en el rango de 2-2.5, lo que implica que los sistemas BWA necesitan un incremento significativo en BHC.

2.4 Canales inalámbricos de banda ancha

La transmisión inalámbrica está limitada por el espectro electromagnético disponible y es atenuada por pérdidas por trayectoria, por interferencia, y por propagación multitrayectoria, lo cual es causa de un desvanecimiento de la señal (*fading*) y retardos de dispersión

(*Delay Spread*). A causa de estas limitaciones, los sistemas inalámbricos ofrecerán mayores retos que los sistemas por cable. A continuación se discuten los canales BWA, enfocando el rango situado por debajo de los 3GHz.

2.4.1 Pérdidas en la línea y retardos de dispersión

Las pérdidas de la línea en los canales BWA dependen del tipo de terreno, mientras que la dispersión de retardos depende en gran medida de las características de las antenas. Los estudios que se han hecho reportan retardos de dispersión RMS para antenas direccionales en entornos suburbanos de aproximadamente 75 ns, mientras que con antenas omnidireccionales se han obtenido retardos de dispersión de hasta 175 ns en los mismos lugares. La razón para que exista esta diferencia en los retardos para las antenas omnidireccionales y direccionales es que en el caso de las antenas omnidireccionales los ecos con mayores retardos tienden a llegar en ángulo desde puntos más alejados de la trayectoria directa y son más atenuados por los lóbulos laterales.

2.4.2 Factor K

La ganancia de trayectoria de un canal BWA puede ser representada como si se tuviera una componente fija y una componente variable correspondiente a la dispersión). La razón de la energía promedio en la componente fija al promedio de la energía en la componente de dispersión es llamada factor K. El valor del factor K tiene implicaciones significativas en el diseño del sistema y en su desempeño. Generalmente se encuentra que el factor K en aplicaciones inalámbricas fijas, puede ser muy bajo, lo que se debe a que las antenas en la estación base y en las estaciones remotas se encuentran a poca altura. La figura 2.2 muestra mediciones del factor K hechas por el grupo Smart Antenas Research Group en la universidad de Stanford. En estas mediciones realizadas en la banda de los 2.4 GHz, la antena transmisora estaba entre 10 y 20m de altura respectivamente, mientras que la antena receptora tenía un lóbulo principal de 50° a 3 dB y estaba colocada a 3m de altura. En estas mediciones se observó que el factor K disminuye significativamente con el

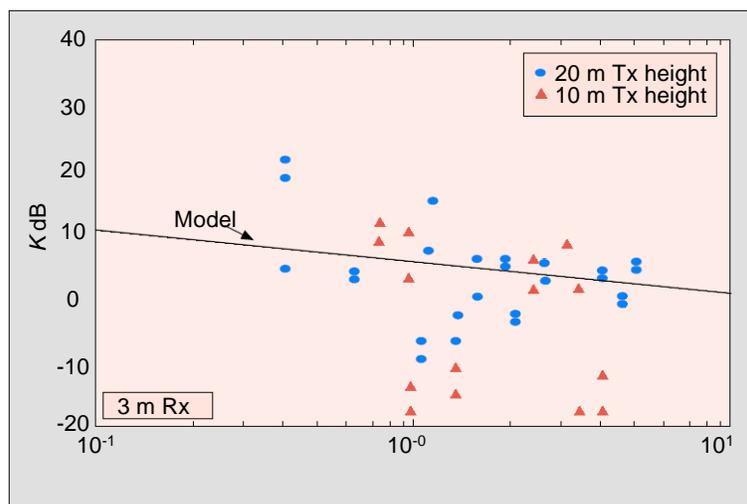


Figura 2.2. Factor K como función de la distancia entre el transmisor y el receptor.

incremento de la distancia que hay entre el transmisor y el receptor. En la gráfica de la figura 2.2 , el factor K se ha promediado tanto en tiempo como en frecuencia. En la práctica, las fluctuaciones significativas en el factor K, pueden deberse tanto a las condiciones de viento como al tráfico. Asimismo, en la figura 2.2 se muestra un modelo teórico para la media del factor K en función de la distancia asumiendo una altura de 20 m para la antena transmisora y 3 m para la receptora. Como se puede apreciar de la gráfica, los resultados prácticos coinciden en gran medida con los datos del modelo.

En resumen, en el diseño de un sistema BWA fijo, se deben asumir valores muy bajos del factor K (es decir condiciones de atenuación Rayleigh) con el fin de conseguir una gran cobertura de la celda y una operación confiable en los límites de la célula.

2.5 Capa Física, Capa MAC y protocolo de radio enlace

En esta sección se discuten algunos puntos relacionados con la capa física, la capa MAC y el protocolo de radio enlace (RLP) de un sistema BWA.

2.5.1 La capa Física

Modulación. A continuación se consideran 3 formatos de modulación: modulación con portadora única ó *Single-Carrier* (SC), modulación con ecualización, acceso múltiple por división de código en secuencia directa (DS-CDMA) con receptor rake y multiplexaje por división de frecuencia ortogonal con entrelazado y codificación (OFDM), y de manera muy breve, se menciona una técnica nueva de modulación en banda ultra ancha (UWBM.).

Modulación de portadora única con ecualización. Existen diferentes opciones de modulación con portadora única, cada una con diferente desempeño y diferentes ventajas y desventajas. La ecualización de máxima probabilidad produce un desempeño óptimo pero computacionalmente son muy caros. La ecualización con decisión realimentada ó DFE es considerada una opción muy atractiva en la práctica. Existen alternativas más simples que incluyen ecualizadores lineales como MMSE (*Minimum Mean Square Error*). La ecualización lineal, sin embargo, no explota apropiadamente la diversidad de frecuencias cuando el canal tiene retardos de dispersión. En la práctica, la complejidad computacional de los ecualizadores de una sola portadora y la complejidad requerida para adaptar un ecualizador pueden imponer límites al desempeño de los sistemas de una sola portadora en el caso de la transmisión de datos a alta velocidad.

DS-CDMA. DS-CDMA utiliza una secuencia de código extendida que multiplica en banda base a los símbolos transmitidos. Este código extendido tiene una tasa de transmisión más alta que la tasa de símbolos, con lo cual, los símbolos se extienden en frecuencia. Con el fin de explotar la diversidad en frecuencia puede utilizarse un receptor rake. Al incrementarse la velocidad de transmisión, se incrementa la tasa de símbolos y el chip rate, lo que permite al sistema resolver ligeras diferencias en los retardos de la trayectoria física, pero al mismo tiempo requiere una complejidad computacional mucho mayor.

Multiplexaje por división ortogonal de frecuencia. OFDM elimina la necesidad de ecualización insertando un intervalo de guarda (prefijo cíclico) que es una copia de la

última parte del símbolo OFDM y debe ser suficientemente larga como para acomodar el retardo de dispersión más largo. El transmisor y el receptor emplean la transformada inversa de Fourier (IFFT) y la transformada de Fourier (FFT) respectivamente, y la ecualización se reduce a simples multiplicaciones de escalares en una base tono por tono. En OFDM, la diversidad de frecuencia se obtiene codificando e intercalando entre tonos. Para un incremento en el retardo extendido y ó en la tasa de transmisión, el prefijo cíclico debe incrementarse proporcionalmente para que permanezca más tiempo que la respuesta a impulso del canal. En resumen, la simplicidad en la ecualización, parece favorecer a OFDM sobre SC y DS-CDMA desde el punto de vista de la complejidad.

Modulación en ultra banda ancha. Recientemente UWBM ha sido de gran interés en el medio de las comunicaciones inalámbricas de banda ancha. En UWBM, se usan trenes de pulsos utilizando modulación por posición de pulsos para enviar la información. Se puede utilizar un receptor rake para utilizar la diversidad de trayectorias. El resultado de los pulsos transmitidos a través del espectro de banda ultra ancha significa que UWBM puede ser capaz de coexistir con otros sistemas de banda angosta, ya que la energía de interferencia por sistema puede ser pequeña e incrementar únicamente el ruido de fondo. Recientemente, UWBA ha recibido gran atención por parte de la FCC. Actualmente algunas compañías están haciendo esfuerzos para comercializar esta tecnología, de hecho todavía no existe un consenso con respecto a las áreas de aplicación de UWBM. Sin embargo todavía no existe un consenso en cuanto a las áreas de aplicación de UWBM.

Consideraciones de Hardware. Desde un punto de vista de la complejidad, OFDM parece ser más atractivo que SC y DS-CDMA. Sin embargo en la práctica, las señales OFDM hacen más sensible al sistema con respecto a las no linealidades de los amplificadores de potencia. Además, en el caso de OFDM, los amplificadores de potencia son más caros. Por tanto, decidir cuál de los sistemas SD, DS-CDMA u OFDM debe usarse depende de dos factores: el costo del equipo requerido para procesar las señales de transmisión y recepción y el costo del amplificador de potencia.

Codificación de canal. La codificación de canal añade redundancia a los datos transmitidos para permitir al receptor la corrección de errores. Como se mencionó anteriormente, en el caso de un canal OFDM, la codificación de canal en combinación con entrelazado proporciona diversidad en frecuencia. La codificación típica de un canal BWA emplea esquemas de codificación Reed-Solomon y códigos convolucionales; los códigos Reed-Solomon son utilizados como códigos externos y los convolucionales como códigos internos. Algunas otras técnicas como decodificación iterativa pueden proporcionar ganancia adicional.

Sincronización. La sensibilidad al tiempo y a la frecuencia de los sistemas SC y DS-CDMA teóricamente es la misma en tanto se utilice el mismo ancho de banda y el mismo throughput. En contraste OFDM es más sensible a los errores de sincronización que SC y DS-CDMA.

Adaptación de enlace. En los sistemas BWA las condiciones del canal pueden variar significativamente debido a la dispersión y a la atenuación de la señal (fading). En consecuencia, es deseable adaptar tanto la modulación como el esquema de codificación

a las condiciones actuales del canal. Mientras que las redes de voz están diseñadas para manejar una tasa fija de bits, los servicios de datos están diseñados para una tasa variable. Las redes de voz están diseñadas para entregar datos a cierta velocidad aún en los límites de la célula, el cual es el peor caso, sin embargo, la mayoría de los usuarios generalmente tienen condiciones del canal más favorables. Por tanto, las redes de datos pueden tomar ventaja de la codificación y de la modulación adaptiva para mejorar el throughput total. En un esquema típico de modulación adaptiva, es posible tener una variación del orden de modulación (tamaño de la constelación) y de la tasa de codificación para la corrección de errores. Durante la operación el receptor debe enviar información al emisor acerca de la calidad de la recepción de la señal, la cual es utilizada para controlar la adaptación. La modulación adaptiva puede utilizarse tanto en el enlace de ascendente como en el descendente.

La adaptación puede hacerse de varias maneras: para solo un usuario en específico, para un usuario y un tiempo específicos ó también puede depender de la calidad de servicio.

Acceso múltiple. En TDMA, el acceso múltiple se hace asignando diferentes ranuras de tiempo a diferentes usuarios de manera que cada usuario tiene su turno para transmitir y recibir, dicho de otra manera, el tiempo de transmisión está dividido en pequeñas ranuras de tiempo de acceso secuencial. Cada usuario tiene su turno para transmitir y recibir, siguiendo un arreglo de tipo *round-robin*. En las redes de datos, donde la utilización del canal puede cambiar de manera abrupta debido a la transmisión por ráfagas, TDMA está modificado con esquemas basados en reservación donde los slots de tiempo se asignan únicamente si el usuario tiene datos para transmitir. En CDMA todos los usuarios transmiten al mismo tiempo utilizando diferentes secuencias cuasi-ortogonales. Mientras que teóricamente hay poca diferencia en cuanto a capacidad entre TDMA y CDMA, CDMA ofrece ventajas en la implementación en cuanto al reconocimiento de las señales y en cuanto a la diversidad de la interferencia. Sin embargo, en redes BWA, CDMA de extensión fija (*fixed spreading*) no es muy atractivo debido a que el factor de dispersión es muy alto (típicamente mayor a 32). Por ejemplo, para una velocidad de transmisión de 10 Mb/s se requiere un ancho de banda de 160 MHz para un factor de esparcimiento de 32. En la tercera generación de sistemas móviles, para un enlace con una alta velocidad de transmisión de datos, el factor de dispersión es de 4 con el fin de mantener el ancho de banda de 4 MHz. Ese factor de esparcimiento tan bajo hace a CDMA muy semejante a TDMA. Sin embargo en la práctica se han intentado algunas otras soluciones tales como la modulación CDMA multicódigo.

TDD vs FDD. Actualmente en la industria de BWA se están debatiendo los méritos de las tecnologías TDD (*Time Division Duplexing*) y FDD (*Frequency Division Duplexing*) en redes punto multipunto. Se puede decir que FDD es el legado de la industria de comunicaciones inalámbricas fijas, en enlaces punto a punto establecidos originalmente para transportar tráfico de voz analógica, el cual es en gran medida simétrico y predecible. Por otro lado, TDD está siendo usado en el diseño de redes punto-multipunto para transportar datos digitales, los cuales son asimétricos e impredecibles. Mientras que TDD requiere un solo canal para comunicaciones full duplex, los sistemas FDD requieren un par de canales para comunicarse, uno para el enlace descendente y otro para el enlace ascendente. En TDD la separación entre la transmisión y la recepción ocurre en el dominio del tiempo, mientras que en FDD ocurre en el dominio de la frecuencia. FDD está

diseñado para manejar tráfico con requerimientos de ancho de banda constantes en ambas direcciones de comunicación, por tanto, TDD permite un mejor manejo de la asimetría del tráfico al variar la asignación de los tiempos de transmisión de subida y de bajada. Dado que el tráfico de Internet es muy variable en el tiempo debido a que opera por ráfagas, el ancho de banda ascendente y descendente debe variar con la demanda de los usuarios, lo cual favorece a TDD. TDD requiere un tiempo de guarda igual al retardo de propagación de ida y de regreso entre la estación base y las estaciones remotas. Por lo anterior este tiempo de guarda se incrementa de acuerdo a la distancia del enlace. En FDD también se requiere suficiente aislamiento de las frecuencias de los canales de los enlaces de subida y de bajada. En resumen se puede decir que FDD es más fácil de implementar sin embargo es la solución menos eficiente.

2.5.2 La capa MAC y RLP

La capa MAC y la capa RLP (*Radio Link Protocol*) trabajan junto con la capa física con el fin de proporcionar la mejor QoS en términos de throughput, de retardo, y de *jitter* permitido a todos los usuarios. La principal tarea de la capa MAC es asociar los requerimientos de transporte y de QoS con las diferentes aplicaciones y servicios, asignar prioridades y programar apropiadamente las transmisiones en los enlaces de ascendente y descendente. En consecuencia, un protocolo inalámbrico MAC debe, proporcionar servicios diferenciados de calidad de servicio, asignación dinámica de ancho de banda y calendarizar ó programar la transmisión de datos en ráfagas. Una característica importante de la capa MAC es la retransmisión, lo que permite la operación con altas tasas de error y con mejor reuso de frecuencia, incrementando la robustez y mejorando el desempeño TCP. Las principales funciones de la capa MAC son:

- Controlar la programación de la transmisión de datos de subida y de bajada, lo que permite soportar múltiples servicios de flujo de datos (QoS) por cada enlace usuario-estación base.
- Controlar la admisión de nuevas conexiones para asegurar que una adecuada capacidad del canal está disponible para acomodar los requerimientos de calidad de servicio de cada nuevo flujo, y vigilar que se cumplan las restricciones de servicio como sería el verificar que un usuario está autorizado para recibir la QoS solicitada.
- Inicialización y mantenimiento del enlace como selección del canal, sincronización registro y aspectos de seguridad.
- Proveer los medios necesarios para el transporte integrado de voz y datos. Los requerimientos típicos de la transmisión de datos son ancho de banda en demanda, bajas tasas de error, y diferenciación de servicios. Los requerimientos para voz, son un ancho de banda garantizado y pérdidas limitadas, retardos y jitter.
- Proveer los medios necesarios para fragmentación, petición de repetición automática (ARQ), modulación adaptiva y codificación.

Para el caso de las redes inalámbricas, las características MAC más importantes son las siguientes:

- Fragmentación de PDUs (*Packet Data Units*) en paquetes más pequeños. Esto ayuda a reducir la tasa de errores y a limitar la latencia en los paquetes de voz.
- Retransmisión a nivel de fragmentos de PDUs.
- Proveer los medios necesarios para programar múltiples esquemas de modulación y codificación.
- Mantenimiento del enlace inalámbrico específico y su control, por ejemplo el control de potencia y la codificación y modulación adaptivas.

2.6 Antenas múltiples en BWA

2.6.1 Ventajas de usar antenas múltiples

Como se ha mencionado previamente, los sistemas BWA fijos, se enfrentan a dos problemas principales: proveer una tasa de transmisión de datos alta y una alta calidad en el acceso inalámbrico, a través de canales de que presentan desvanecimientos, todo ello con una calidad comparable a aquella que proporcionan los servicios por cable.

Los altos requerimientos en cuanto a calidad surgen debido a que los sistemas inalámbricos BWA compiten con los cablemódems y con DSL asíncrono (ADSL), los cuales operan en canales fijos y por tanto proporcionan una gran calidad. Estos requerimientos de alta calidad constituyen la mayor diferencia con respecto a las redes celulares móviles existentes, donde los clientes están acostumbrados a aceptar bajos niveles de QoS. Además, en las redes celulares móviles, los requerimientos para la tasa de transmisión de datos son mucho menores que en las redes BWA fijas. El uso de múltiples antenas en los sitios de transmisión y recepción de un enlace inalámbrico en combinación con el procesamiento de las señales y codificación es un medio que promete satisfacer todos estos requerimientos. Hay que hacer notar que en los sistemas BWA fijos, a diferencia de los sistemas celulares móviles, es posible usar antenas múltiples en el CPE.

Las ventajas que proporciona el uso de antenas múltiples en la BTS y en la CPE son las siguientes:

- Ganancia del arreglo. Antenas múltiples pueden combinar señales de manera coherente para incrementar la razón C/N y por consiguiente mejorar la cobertura. La combinación coherente puede utilizarse tanto en el transmisor como en el receptor y requiere un conocimiento de las condiciones del canal. Debido a que es difícil conocer las condiciones del canal en el lado del transmisor, es más fácil obtener una ganancia del arreglo en el receptor.
- Diversidad en la ganancia. La diversidad espacial a través de múltiples antenas puede usarse para combatir el desvanecimiento de la señal y mejorar significativamente la confiabilidad del enlace. La diversidad en la ganancia se puede obtener tanto en el transmisor como en el receptor. Los códigos de espacio-tiempo desarrollados recientemente permiten una diversidad en la transmisión sin el conocimiento del canal por parte del transmisor.
- Supresión de interferencia. Se pueden utilizar antenas múltiples para suprimir la CCI y con ello incrementar la capacidad de la célula.

- Multiplexaje de ganancia. El uso de antenas múltiples tanto en el transmisor como en el receptor permite establecer diferentes enlaces de comunicación paralelos con del mismo ancho de banda, lo cual lleva a un incremento lineal (con el número de antenas) en la tasa de transmisión de datos.

Resumiendo, el empleo de antenas múltiples tanto en BTS como en CPE puede mejorar la capacidad celular y la confiabilidad del enlace.

2.6.2 Codificación espacio-temporal y diversidad en la transmisión

Dos de los mayores problemas de los sistemas de comunicación inalámbricos son el desvanecimiento causado por la adición destructiva de señales debido a multi trayectorias en el medio de propagación y la interferencia proveniente de otros usuarios. La diversidad proporciona al receptor varias réplicas (idealmente independientes) de la señal transmitida y es un medio muy poderoso para combatir la interferencia y el efecto de desvanecimiento. Algunas formas comunes de diversidad son, diversidad en tiempo (debido a la dispersión Doppler) y diversidad en frecuencia (debido a la dispersión de retardos). En años recientes, se ha hecho muy popular el uso de diversidad espacial (ó de antenas). La diversidad de recepción, es decir, el uso de múltiples antenas en el receptor, es un tema muy estudiado actualmente. El uso de múltiples antenas para transmitir, en combinación con la codificación en espacio y en tiempo (diversidad de transmisión) ha sido un área de gran interés para la investigación, lo cual ha estado motivado por las aplicaciones inalámbricas móviles, debido a que es difícil emplear múltiples antenas en un dispositivo celular portátil. La codificación en espacio-tiempo (diversidad de transmisión) es un método que conlleva diversidad de ganancia sin el conocimiento del canal por parte del transmisor, lo anterior se logra codificando en espacio (con antenas) y en tiempo. Con el fin de demostrar el impacto de la diversidad de ganancia en el desempeño de un sistema BWA, la figura 2.3 muestra el nivel de señal en un sistema BWA, con y sin diversidad de antenas. Podemos ver claramente los desvanecimientos profundos que aparecen cuando hay una sola antena, no aparecen en el caso en el que hay múltiples antenas; de aquí que la diversidad hace que el canal tenga menos desvanecimientos, lo cual es de importancia fundamental en los sistemas BWA fijos, donde se pueden presentar en gran medida desvanecimientos mientras que las condiciones del canal cambian muy lentamente. La figura 2.4 muestra los niveles de señal y de interferencia para un sistema con una sola antena y para otro sistema con múltiples antenas respectivamente. Podemos ver que

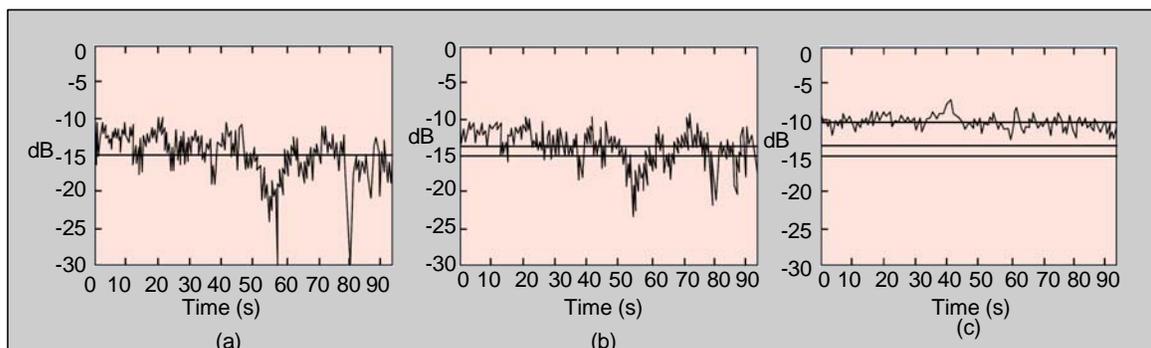


Figura 2.3. Nivel de señal en el receptor de sistemas con a) 1 input 1-output, b) 1 input-2 output and c) 2 input 3-output.

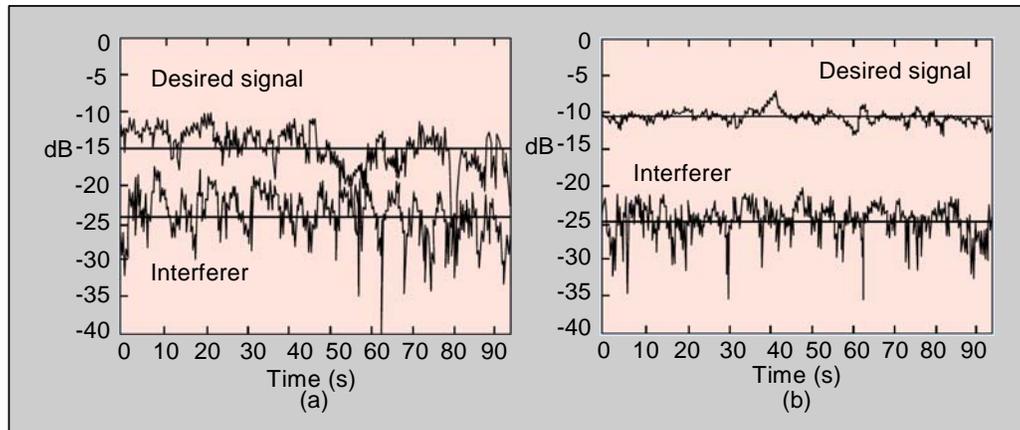


Figura 2.4. Niveles de señal e interferencia para un sistema con a) 1-input 1-output, y b) 2 input 3 output.

cuando hay múltiples antenas, el nivel de la señal es más alto y varía mucho menos. En consecuencia, es posible hacer un reuso de frecuencia más agresivo en el caso en el que se tengan antenas múltiples, lo cual mejora la capacidad de la célula. También podemos notar que en el caso de los sistemas BWA fijos, es más difícil explotar la diversidad en tiempo, ya que el canal está cambiando lentamente y sería necesario el entrelazado en periodos largos de tiempo, lo cual no es consistente cuando se tienen requerimientos de retardo muy estrictos.

La solicitud automática de repetición (ARQ por sus siglas en inglés) nos permite realizar cierta diversidad en tiempo. Es probable que la diversidad en frecuencia esté disponible en el caso de los sistemas BWA fijos, pero no es una fuente confiable de diversidad, ya que no se pueden garantizar grandes retardos de dispersión, especialmente cuando se utilizan antenas direccionales donde los retardos de dispersión son reducidos significativamente. Las mediciones hechas por la universidad de Stanford muestran que en una macrocelda de un ambiente MMDS, las antenas de media onda espaciadas adecuadamente son suficientes para asegurar una ganancia en diversidad espacial. Por tanto, es factible la diversidad espacial ya que además es una forma de diversidad bastante confiable.

2.6.3 Multiplexaje espacial

El multiplexaje espacial es una técnica que produce un incremento en la tasa de transmisión utilizando múltiples antenas en ambos extremos del enlace inalámbrico. Este incremento se obtiene sin utilizar un mayor ancho de banda y sin aumentar el consumo de potencia. La idea básica es que el uso de múltiples antenas en el transmisor y en el receptor proporcionan múltiples trayectorias espaciales paralelas dentro del mismo ancho de banda y permiten un incremento lineal (con el número de antenas) en la capacidad cuando hay una gran dispersión. En el caso de los sistemas BWA fijos, el uso de multiplexaje espacial parece ser particularmente interesante debido a que se requiere una alta tasa de transmisión de datos. En la figura 2.5 se muestra un esquema de multiplexaje espacial. Debido a que la ganancia de multiplexaje es obtenida únicamente si hay un medio de dispersión adecuado, ó equivalentemente, si hay suficientes retardos de

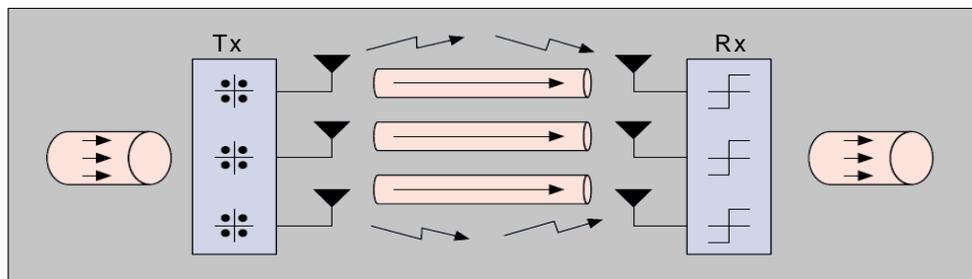


Figura 2.5. Esquema de un sistema de multiplexaje espacial.

dispersión, el multiplexaje espacial trabaja mejor con antenas omnidireccionales ya que las antenas direccionales tienden a limitar la contribución de trayectorias múltiples, reduciendo así los retardos de dispersión.

2.7 Retos en el futuro y tendencias de la industria

2.7.1 Retos a futuro

Anteriormente se ha mencionado que en las redes inalámbricas fijas BWA, aún cuando el transmisor y el receptor están localizados en puntos fijos, una altura baja de las antenas, tanto transmisora como receptora, la influencia del viento, el movimiento del tráfico y el follaje representan un medio bastante hostil. Frecuentemente se han visto desvanecimientos temporales de cerca de 30 dB a una tasa de hasta 1-2 Hz en la banda menor a 3 GHz. Ya que los sistemas BWA compiten básicamente con cable modems y ADSL donde el canal es estático y no presenta desvanecimientos (*fading*), es necesario que estos sistemas proporcionen una gran calidad además de una transmisión de datos a alta velocidad. El mayor reto, es en consecuencia diseñar un enlace inalámbrico en un medio que favorece el desvanecimiento de las señales y que tenga la calidad de una línea de cable. Como se discutió previamente, una posible solución es emplear antenas múltiples. Las recientes actividades de la industria sugieren que el uso de antenas múltiples en combinación con OFDM (MIMO OFDM) es una posible solución para las redes BWA fijas. El segunda mayor desafío ocurre en la capa MAC, en donde es crucial que la capa MAC soporte técnicas más sofisticadas de la capa física tales como la modulación y la codificación adaptativa, la codificación en espacio-tiempo y el multiplexaje espacial.

2.7.2 Estándares

Actualmente los estándares de acceso de banda ancha son objeto de discusión en muchos países de Europa, así como en Japón y Estados Unidos. Diferentes regiones y diferentes países utilizan términos diferentes cuando se refieren a estos estándares. En Europa son conocidos como *Broadband Radio Access Networks* (BRAN); en Estados Unidos, como *Local Multipoint Distribution Services* (LMDS), IEEE 802.16 y sistemas inalámbricos de banda ancha (BWA). En Canadá y en otros países, también son conocidos como *Local Multipoint Communication Systems* (LMCS). Sin embargo, sus

aplicaciones son variadas: fijas y móviles, de área local y de área amplia e incluyen aplicaciones muy prometedoras, como por ejemplo:

- Acceso a Internet de alta velocidad.
- Comunicaciones de datos en dos direcciones (*peer to peer* ó cliente/servidor).
- Telefonía pública ó privada.
- Servicios multimedia de 2 vías, como video conferencia y video comercio.
- Transmisión de video en modo *broadcast*.

El acceso a banda ancha consiste en lo que se ha llamado HIPERACCESS (*High-Performance, Radio Access*), HIPERLAN e HIPERLINK, como se muestra en la figura 2.6.

Mientras que las definiciones ETSI indican una topología jerárquica bien definida, es posible conectar directamente redes LAN inalámbricas ya sea al *backbone* ó a las terminales HIPERACCESS.

Los sistemas HIPERACCESS proporcionan servicio principalmente a pequeñas oficinas, casas (*Small Office/Home Office, SOHO*) y a empresas pequeñas y medianas proporcionando acceso a una variedad de servicios de telecomunicaciones como voz, datos y servicios multimedia, con tasas de transmisión que van desde los 2 Mb/s hasta los 25 Mb/s. El uso de HIPERACCESS será primordialmente como red de acceso remoto de banda ancha. El espectro en frecuencias podrá estar en el rango de los 2-40 MHz.

La segunda aplicación, llamada HIPERLAN, proporciona acceso local con una QoS controlada para aplicaciones de banda ancha (por ejemplo Internet y videoconferencia) a computadoras portátiles, para usarse dentro de edificios ó campus, utilizando principalmente una conexión inalámbrica en la banda libre de los 5 GHz.

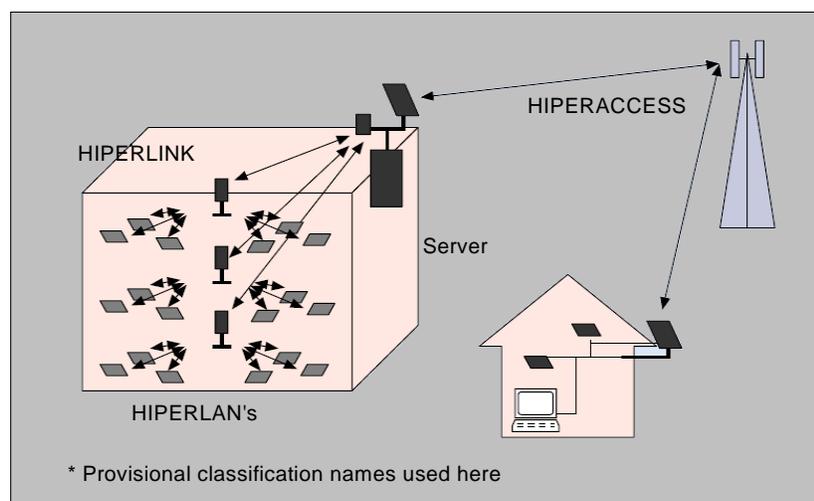


Figura 2.6. Acceso inalámbrico de banda ancha.

La tercera aplicación de BRAN es llamada HIPERLINK, y es principalmente una conexión de red a red vía radio, la cual dará servicio a una variedad de protocolos y a los tipos de tráfico antes mencionados. Esta aplicación tendrá una tasa de transmisión de hasta 155 Mb/s en la banda de los 17 GHz.

El protocolo IEEE802.16, cubre otros aspectos además de HIPERACCESS, incluyendo *WirelessMAN* y *Wireless High-Speed Unlicensed Metropolitan Area Networks* (HUMAN), lo cual incluye frecuencias desde 2 hasta 66 GHz. Anteriormente las capas física y MAC del protocolo 802.16 tenían 2 variantes, una de ellas basada en DOCSIS. Actualmente el comité está trabajando en otras soluciones que reflejen las aplicaciones de Internet coexistiendo con las nuevas tecnologías de acceso inalámbrico.

La Unión Internacional de Telecomunicaciones (ITU) ha iniciado un grupo de trabajo ITU JRG8A-9B que está a cargo de la estandarización de los sistemas de acceso inalámbrico de banda ancha. Este grupo recibe la información del grupo 802.16 y BRAN, y está tratando de lograr un consenso global en esta tecnología desde el punto de vista de ITU.

2.7.3 Tendencias a futuro

Las tendencias futuras en BWA serán las siguientes:

- Evolución de BWA de aplicaciones principalmente empresariales ó de negocios a aplicaciones residenciales.
- El ancho de banda en demanda será un tema clave, así como una utilización más eficiente del espectro.
- Sistemas reconfigurables con mayor adaptabilidad para soportar multibanda, múltiples estándares y múltiples portadoras.
- Convergencia de los servicios de banda ancha inalámbricos y móviles.

Con la nueva asignación del espectro hecha por la ITU para la cuarta generación de comunicaciones móviles, la convergencia de BWA y 4G móvil, será el foco de actividad en las comunicaciones inalámbricas.

DESCRIPCIÓN DEL PROTOCOLO DE COMUNICACIONES IEEE 802.16

3.1 Introducción

El estándar 802.16, completado en abril de 2002, define las especificaciones para la interfase de las redes de área metropolitana inalámbricas WirelessMAN™. La elaboración de este estándar marcó el inicio de la llegada del acceso inalámbrico de banda ancha tanto a casas como a oficinas. Gracias a esta nueva herramienta las pequeñas redes locales de casas y oficinas estarán en comunicación con las grandes redes de telecomunicaciones de todo el mundo.

El estándar define el acceso a una red de área metropolitana o MAN a través del uso de antenas exteriores, las cuales permitirán comunicarse vía radio con las estaciones base (BSs). Esta tecnología es una alternativa a los accesos de red vía cable como por ejemplo, los enlaces de cable coaxial utilizados en los cable módems, los enlaces de fibra óptica y los enlaces DSL (Digital Subscriber Line). Las redes inalámbricas tienen la capacidad de dar servicio a una gran área geográfica sin la necesidad de instalar una línea de cables hasta algún sitio específico. Debido a lo anterior estas redes han probado que su implementación es más económica y que puede llevar a un incremento en el número de accesos a la tecnología de banda ancha. Estos sistemas ya han estado en uso por varios años, sin embargo el desarrollo de un nuevo estándar, es un indicativo de que la industria está alcanzando madurez en ese sentido y establece las bases del éxito de la segunda generación de estos equipos.

En este escenario, los usuarios que estén dentro de algún edificio con servicio wireless MAN accederán a la red de la manera que se ha estado haciendo tradicionalmente. Por ejemplo para la transmisión de datos los usuarios se conectarán a la red Ethernet del edificio (estándar IEEE 802.3), o a alguna red Lan inalámbrica (estándar IEEE 802.11). Sin embargo, el diseño fundamental del estándar permitirá eventualmente la extensión de los protocolos de red wireless MAN directamente hasta el usuario final. Por ejemplo, algún día la estación central BS compartirá datos desde la capa MAC con una computadora individual en una casa. Los enlaces desde la estación base al receptor dentro de las casas y del receptor casero hacia la laptop, podrán utilizar capas físicas completamente diferentes, pero el diseño de la capa MAC Wireless MAN podrá dar servicio a esa

conexión proporcionando una buena calidad de servicio (QoS). Actualmente los miembros del grupo de trabajo del estándar IEEE 802.16 están trabajando en la evolución de la operación fija a la portabilidad y movilidad. La enmienda IEEE 802.16e corrige la especificación base para habilitar no sólo la operación fija, sino también la portátil y la móvil. Los grupos de trabajo de IEEE 802.16f e IEEE 802.16g se encargan de las interfaces de administración de la operación fija y móvil. En un escenario totalmente en movimiento, los usuarios podrán desplazarse mientras tienen acceso de datos de banda ancha o a una sesión de transmisión en tiempo real de multimedia. Todas estas mejoras ayudarán a hacer que WiMAX sea una solución aún mejor para el acceso de Internet para economías en crecimiento.

El estándar 802.16 fue diseñado para evolucionar como un conjunto de interfaces aéreas basadas en un protocolo MAC común, pero con especificaciones de la capa física dependiente de la frecuencia que se esté usando y de las regulaciones asociadas. El estándar, tal y como fue aprobado en 2001, está diseñado para manejar frecuencias desde 10 hasta 66 Ghz. En ese rango de frecuencias hay una parte importante del espectro disponible a nivel mundial, sin embargo, es en ese mismo rango en donde las microondas presentan retos importantes para su desarrollo. El proyecto enmienda IEEE 802.16a extiende esta interfase aérea para dar servicio a bajas frecuencias, en la banda de 2-11 GHz, incluyendo el espectro libre y con licencia. Comparado con altas frecuencias, esta banda espectral ofrece la oportunidad de llegar a más usuarios y de una manera más económica aunque a velocidades de transmisión más bajas. Esto sugiere que estos servicios estarán orientados hacia casas habitación o pequeñas y medianas empresas.

3.2 Aspectos de diseño

3.2.1 Control de acceso al medio (MAC)

El protocolo MAC IEEE 802.16, está diseñado para acceso inalámbrico de banda ancha con conexiones punto-multipunto. Esto deja ver claramente la necesidad de tener velocidades de transmisión muy altas en ambas direcciones: tanto ascendente (hacia la estación base BS) como descendente (desde la estación base BS hacia las terminales). Los algoritmos de acceso y asignación de ancho de banda deben ser capaces de acomodar a cientos de terminales por canal, donde cada terminal está compartida por múltiples usuarios finales. La naturaleza de los servicios requeridos por los usuarios es muy variable. Estos servicios incluyen voz y datos utilizando multiplexaje por división de tiempo (TDM), tráfico de internet (IP) y servicios de voz sobre IP (VoIP). Con el fin de soportar esta variedad de servicios, la capa MAC 802.16 debe soportar tanto tráfico continuo como por ráfagas. Adicionalmente se espera poder ofrecer una calidad de servicio diferente a cada uno de los distintos tipos de tráfico utilizados. El estándar 802.16 MAC proporciona un amplio rango de categorías de servicios similares a los de ATM, y otras nuevas categorías como la tasa de tramas garantizada o GFR (guaranteed frame rate).

El protocolo 802.16 MAC debe cumplir una gran variedad de requerimientos backhaul, incluyendo tanto el modo de transferencia asíncrona (ATM) como protocolos basados en paquetes. Las subcapas de convergencia son usadas para mapear el tráfico de capas de

transportación específicas a una MAC que es lo suficientemente flexible como para transportar de manera eficiente cualquier tipo de tráfico. A través de técnicas como supresión de encabezados (header supresión), empaquetamiento (packing) y fragmentación, las subcapas de convergencia y MAC trabajan conjuntamente para transportar el tráfico de una manera que resulta más eficiente que el mecanismo de transporte original.

En la interfase entre la capa MAC y física (PHY) también se manejan aspectos de la eficiencia de transporte. Por ejemplo, los esquemas de modulación y codificación están especificados en un perfil de ráfagas que puede ser ajustado adaptativamente en cada ráfaga con cada estación suscriptor. La capa MAC puede utilizar un perfil que utilice eficientemente el ancho de banda cuando existan las condiciones favorables de enlace y cambiar a otro perfil más confiable aunque menos eficiente cuando así se requiera con el fin de dar el mejor servicio mientras haya un enlace disponible.

El mecanismo de petición-concesión o request-grant está diseñado para ser escalable, eficiente y auto-correctivo. El sistema de acceso 802.16 no pierde eficiencia cuando se presentan múltiples conexiones por terminal, múltiples niveles de calidad de servicio por terminal, y un gran número de usuarios multiplexados estadísticamente. Este sistema tiene las ventajas de poder utilizar una gran variedad de mecanismos para hacer peticiones de ancho de banda. Esto permite balancear la estabilidad del acceso sin contención o reservado con la eficiencia del acceso orientado a conexión.

El estándar 802.16 menciona los mecanismos de asignación de ancho de banda y las categorías de calidad de servicio. Sin embargo, los detalles relativos a la implementación de dichos servicios no están estandarizados y proporcionan mecanismos importantes para la diferenciación del producto. Además de las tareas fundamentales de distribuir el ancho de banda y transportar datos, la capa MAC incluye una subcapa de privacidad que proporciona la autenticación con la red de acceso y el establecimiento de la conexión con el fin de evitar que usuarios no autorizados hagan uso del servicio. Esta subcapa proporciona también un mecanismo de intercambio de llaves y encriptación de datos para mantener la privacidad de los mismos.

Con el fin de ajustarse al entorno físico y poder satisfacer los diferentes requerimientos de servicio en las frecuencias de entre 2 y 11 GHz, el estándar 802.16 se está actualizando para proporcionar mecanismos de solicitud automática de repetición (ARQ) y servicios de red de malla, además de las arquitecturas de red de punto-multipunto.

3.2.2 La capa física

a) 10-66 GHz

En el diseño de las especificaciones de la capa física para 10-66 GHz, se estableció la comunicación con línea de vista como una necesidad primordial. De acuerdo a lo anterior, se seleccionó una modulación con una sola portadora (single-carrier modulation) y la interfase aérea se denominó "WirelessMAN-SC". Sin embargo todavía permanecen varios retos de diseño que aún no han sido resueltos. Debido a la arquitectura punto-multipunto, la estación base transmite básicamente una señal TDM, con slots o ranuras de tiempo

correspondientes a cada estación dispuestos en forma serial. El acceso en la dirección ascendente se realiza por medio de acceso múltiple por división de tiempo o TDMA. En cuanto a la transmisión tanto de subida como de bajada, se seleccionó un diseño que permite tanto el duplexaje por división de tiempo (TDD) como el duplexaje por división de frecuencia. En TDD el canal de subida y de bajada comparten un canal pero no transmiten simultáneamente mientras que en FDD el canal de subida y de bajada operan en canales separados y en algunas ocasiones simultáneamente. Ambas alternativas soportan perfiles adaptativos en ráfagas, en los cuales las opciones de codificación y modulación pueden asignarse dinámicamente de acuerdo a las ráfagas de tráfico.

b) 2-11 GHZ

Las bandas de 2-11GHz, tanto las que requieren licencia como aquellas que son libres, están descritas en el documento IEEE Project 802.16a. El diseño de este estándar está orientado por la necesidad de operación sin línea de vista (NLOS por sus siglas en inglés: Non-Line Of Sight). Por ejemplo, en el caso de las aplicaciones residenciales, las antenas localizadas en las azoteas de las casas pueden estar muy bajas y no tener una línea de vista franca hacia la antena de la estación base, debido a la obstrucción de los árboles. Además, también se espera que la señal se propague por multitrayectorias, ya que el hecho de reinstalar una antena en un mejor sitio puede no resultar económico, tanto por el hardware como por los costos de reinstalación.

Las 3 especificaciones para la interfaz aérea en la banda de 2-11GHz son:

- WirelessMAN-SC2: Esta usa un formato de modulación con portadora única.
- WirelessMAN-OFDM: Este utiliza un multiplexaje por división de frecuencia ortogonal con una transformada de 256 puntos. El acceso se realiza por medio de TDMA. Esta interfase aérea es obligatoria para las bandas libres.
- WirelessMAN OFDMA: Utiliza acceso múltiple por división de frecuencia con una transformada de 2048 puntos. En este sistema el acceso múltiple se realiza a través del uso de múltiples portadoras y de un sólo receptor.

3.3 La capa física

Las especificaciones de la capa física para la banda de 10-66 GHz, definen que se utilizará una modulación con portadora única con un perfil adaptativo para tráfico en ráfagas. En este sistema los parámetros de transmisión, tales como la modulación y la codificación, se pueden ajustar individualmente para cada estación suscriptora trama por trama. En estas especificaciones se definen también las variantes TDD y FDD. También se especifica un ancho de banda de 20 a 25 MHz (asignación para EU) y de 28 MHz (asignación Europea).

El mecanismo de corrección de errores utilizado (FEC), es el Reed-Solomon GF (256), con un tamaño de bloque variable. Este algoritmo está acompañado de un bloque convolucional interno para transmitir de una manera robusta datos críticos tales como control de trama y control de acceso inicial. Las opciones de FEC se complementan con transmisiones que utilizan un desplazamiento de fase en cuadratura (QPSK), modulación por amplitud en cuadratura de 16 estados (16-QAM) y una modulación QAM de 64

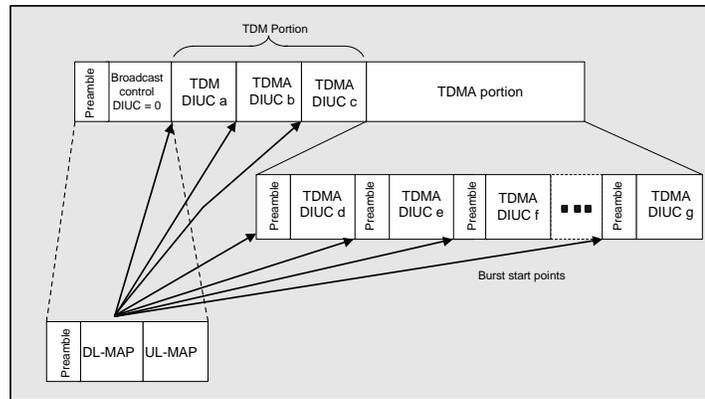


Figura 3.1. Estructura de la subtrama de subida.

estados (64-QAM) para formar perfiles de ráfagas con una eficiencia y robustez variables. Si el último bloque FEC no se utiliza completamente puede ser acortado. El acortamiento de bloques tanto en el enlace de subida como de bajada es controlado por la estación base. La estación base lo comunica implícitamente tanto en el mapa de subida (UL-MAP) como en el de bajada (DL-MAP).

El sistema utiliza tramas o frames de 0.5, 1 o 2 ms. Este frame está dividido en slots físicos con el fin de asignar el ancho de banda e identificación de las transiciones de la capa física. Se definió un slot físico como aquel que ocupa 4 símbolos QAM. Cuando se utiliza TDD, el subframe de subida sigue al subframe de bajada utilizando una portadora de la misma frecuencia. En el caso de FDD, las subtramas de subida y de bajada ocurren al mismo tiempo pero se transmiten a diferentes frecuencias, la subtrama de bajada se muestra en la figura 3.1.

La subtrama del enlace de bajada comienza con una sección de control que contiene el mapa de bajada DL-MAP, con la descripción de esa trama de bajada que está siendo enviada así como el mapa de subida UL-MAP que será transmitido a futuro. El mapa de bajada especifica el momento en que ocurren transiciones en la capa física, tales como cambios en FEC o en la modulación, en la subtrama de bajada. La subtrama de bajada típicamente contiene una porción TDM inmediatamente después de una sección de control de trama. Los datos de bajada son transmitidos a cada SS utilizando un ancho de banda negociado previamente. Los datos son transmitidos en orden de robustez decreciente, con el fin de permitir a las estaciones suscriptoras recibir sus datos antes de que se presente otro perfil de ráfagas que pudiera causarles la pérdida de sincronía con el enlace de bajada.

En los sistemas FDD, la porción TDM puede estar seguida de un segmento TDMA que incluye un preámbulo extra al inicio de cada nuevo perfil de ráfagas. Esta característica permite a las SSs soportar mejor la comunicación half-duplex. En un sistema FDD programado eficientemente con múltiples SSs utilizando half dúplex puede ocurrir que muchas de las estaciones deban transmitir en la trama antes de que reciban algún dato. Debido a la naturaleza half dúplex, estas SSs pierden sincronía con el enlace de bajada, pero ésta puede ser recuperada gracias al preámbulo TDMA.

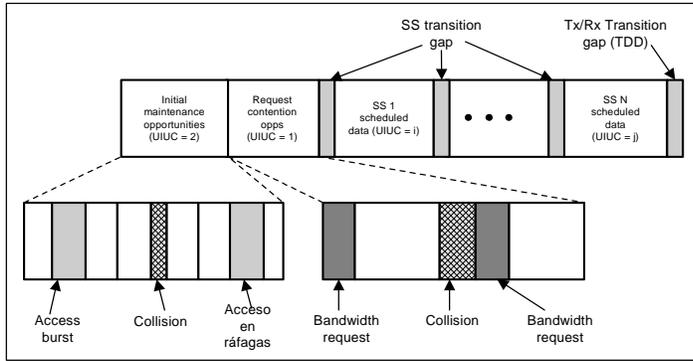


Figura 3.2. Estructura de la subtrama de subida.

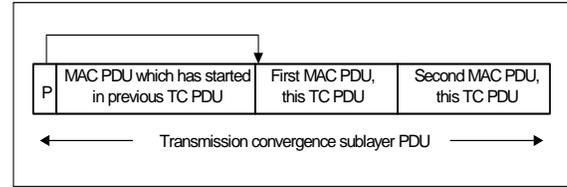


Figura 3.3. Formato TC PDU.

La demanda de ancho de banda varía enormemente por la variedad de servicios que pueden estar activos. Debido a esa variación, la combinación y la duración de los perfiles de tráfico, así como la presencia o ausencia de la porción TDMA pueden variar dinámicamente de una trama a la siguiente. Ya que cada destinatario está indicado en los encabezados MAC y no en el DL-MAP, todas las SS escuchan todas las porciones de la subtrama de bajada que son capaces de recibir. Esto quiere decir que una SS full dúplex recibirá los mensajes de todas las ráfagas que tengan un perfil de una robustez igual o mayor que la que negoció con la estación base.

Una subtrama ascendente para la capa física de 10-66 GHz se muestra en la figura 3.2. A diferencia de los enlaces descendente el mapa UL proporciona ancho de banda a través oportunidades de transmisión a estaciones SS específicas. Las estaciones SS transmiten en el slot que tienen asignado en el mapa utilizando el perfil especificado en el código de uso de intervalo de subida (Uplink Interval Usage Code UIUC). La subtrama de subida también puede contener slots de contención para un acceso inicial al sistema y solicitudes de acceso broadcast o multicast.

Entre la capa física y la capa MAC está la capa de convergencia de transmisión (TC). Esta capa se encarga de transformar las unidades de datos de protocolo (PDUs) MAC en bloques FEC de longitud fija (con la posibilidad de añadir un bloque acortado al final) en cada ráfaga. La capa TC tiene un PDU del tamaño justo para colocarse en el bloque FEC actual. Este MAC PDU comienza con un apuntador que indica dónde comienza el siguiente encabezado MAC PDU dentro del bloque FEC (ver figura 3.3).

El formato TC PDU permite la resincronización con el siguiente MAC PDU en el caso en el que al bloque FEC previo tenga errores no recuperables. Sin la capa TC un receptor, ya fuera SS o BS, podría perder completamente el resto de una ráfaga después de la ocurrencia de un error.

3.4 Detalles de la capa MAC

El control de acceso al medio incluye subcapas de convergencia de servicios específicos que sirven de interfase con las capas superiores. Las subcapas de convergencia se encuentran sobre la subcapa común la cual realiza las funciones clave de la capa MAC.

Bajo esta capa se encuentra la subcapa de privacidad.

3.4.1 Subcapas de convergencia de servicio específico

El estándar 802.16 define 2 subcapas de convergencia de servicio específico para fines de mapeo desde y hacia las conexiones MAC 802.16. La subcapa de convergencia ATM está definida para servicios ATM, y la subcapa de convergencia de paquetes está definida para servicios de mapeo de paquetes de servicios tales como IPv4, IPv6, Ethernet y redes de área local virtuales (VLAN). La tarea principal de estas subcapas es la de clasificar unidades de servicio de datos (SDUs) hacia la conexión MAC apropiada, preservar o habilitar QoS y permitir la asignación de ancho de banda. El mapeo toma varias formas dependiendo del tipo de servicio. Además de estas funciones básicas, las capas de convergencia también desarrollan funciones más sofisticadas como la supresión de encabezados y la reconstrucción con el fin de mejorar la eficiencia del enlace aéreo.

3.4.2 Subcapa de parte común

3.4.2.1 Introducción y arquitectura general

En general, el 802.16 MAC está diseñado para soportar una arquitectura de punto multipunto con una estación base BS manejando múltiples sectores independientes simultáneamente. En el enlace de bajada, los datos hacia las estaciones son multiplexados utilizando TDM, mientras que en la subida se emplea TDMA.

El MAC 802.16 está orientado a conexión. Todos los servicios, incluyendo aquellos que no lo requieren están mapeados a una conexión. Esto proporciona un mecanismo para solicitar ancho de banda, calidad de servicio y parámetros de tráfico, transportando y enrutando datos a la subcapa de convergencia apropiada y para todas aquellas acciones asociadas con las cláusulas del servicio. Las conexiones están referenciadas con un identificador de conexión de 16 bits (CID), y puede requerir oportunidades de transmisión continua o en demanda, como se describirá más adelante.

Cada estación o SS tiene una dirección estándar MAC de 48 bits, que sirve principalmente como un identificador del equipo ya que la dirección utilizada durante la operación es el CID. Después de registrarse en una red, el SS tiene asignadas 3 conexiones de control en cada dirección. Estas 3 conexiones reflejan los 3 diferentes requerimientos de calidad de servicio usados por los diferentes niveles de control. La primera de ellas es la conexión básica (Basic Connection), la cual es utilizada para la transferencia de mensajes cortos, tanto aquellos mensajes MAC de tiempo crítico, como aquellos de control del radio enlace (RLC). La segunda conexión es llamada conexión de administración principal (primary management connection), y es usada para transmitir mensajes más largos y tolerantes a retardos como aquellos utilizados en la autenticación y establecimiento de la conexión. Por último, la conexión de administración secundaria (secondary management connection) es usada para la transferencia de mensajes de control basados en estándares, tales como el Protocolo de Configuración del Host Dinámico (DHCP), y el Protocolo Simple de Manejo de Red (SNMP). Además de estas conexiones de control, las SSs tienen asignadas conexiones de transporte para los servicios contratados. Las conexiones de transporte son

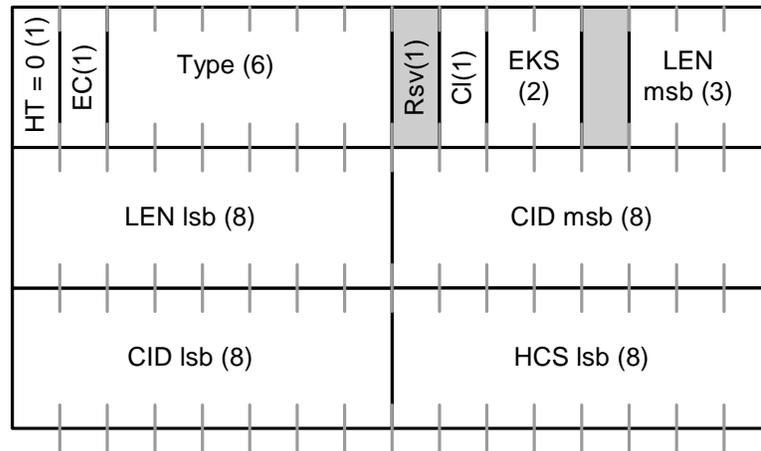


Figura 3.4. Formato del encabezado genérico para un MAC PDU.

asignadas por pares a cada SS y son unidireccionales para proporcionar diferentes calidades de servicio y parámetros de tráfico.

El MAC reserva conexiones adicionales para otros propósitos. Una conexión está reservada para el acceso inicial basado en contención. Otra conexión está reservada para transmisiones broadcast de bajada, así como para la señalización de los slots de contención. Adicionalmente, hay conexiones multicast para poleo. Los SS pueden recibir las instrucciones de reunir grupos de poleo multicast y asociarlas con estas conexiones de poleo.

3.4.2.2 Formatos MAC PDU

Los MAC PDUs, son las unidades empleadas para transmitir datos entre las capas MAC de la BS y de las SSs. Un MAC PDU consiste en un encabezado MAC de longitud fija, una parte de carga útil del mensaje o payload de tamaño variable y opcionalmente un parte de chequeo por redundancia cíclica (CRC). Existen 2 formatos de encabezados MAC. Los encabezados se distinguen por el campo HT (Header Type) que puede ser: el encabezado genérico (figura 3.4) y el encabezado para petición de ancho de banda.

Con excepción de los MAC PDUs para demanda de ancho de banda, que no contienen carga útil, los MAC PDUs contienen ya sea mensajes de control MAC o datos de la capa de convergencia.

Tres tipos de subencabezados MAC pueden estar presentes. El subencabezado de administración de oportunidades de transmisión (*grants*) es usado por las estaciones suscriptoras para enviar sus requerimientos de ancho de banda. El subencabezado de fragmentación contiene información que indica la presencia y orientación en la carga útil de fragmentos de SDUs. Por último, el subencabezado de empaquetamiento (*packing*) que se usa para indicar que múltiples SDUs están contenidos en un solo PDU. Los subencabezados de administración de *grants* y fragmentación pueden estar insertados en MAC PDUs inmediatamente después del encabezado genérico si así se indica en el

campo type. El subencabezado de empaquetamiento se puede insertar antes de cada MAC SDU si así se indica en el campo type.

3.4.2.3 Transmisión de MAC PDUs

El MAC IEEE 802.16 soporta varios protocolos de capas superiores, tales como ATM o IP. Los SDUs MAC provenientes de las subcapas de convergencia correspondientes son formateados de acuerdo al formato MAC PDU, posiblemente con fragmentación y/o empaquetamiento, antes de ser transferidas a través de una o más conexiones de acuerdo al protocolo MAC. Después de transmitirse a través del enlace, se reconstruyen los MAC SDUs a partir de los MAC PDUs, de tal manera que las modificaciones hechas por el protocolo de la capa MAC son completamente transparentes para la entidad receptora.

El 802.16 aprovecha al máximo los procesos de fragmentación y empaquetamiento al incorporarlos con el proceso de asignación de ancho de banda maximizando la flexibilidad, eficiencia y efectividad de ambos. La fragmentación es el proceso mediante el cual un MAC SDU es dividido en uno o más fragmentos MAC PDU. El empaquetamiento es el proceso a través del cual, múltiples MAC SDUs son empaquetados como una sola carga útil MAC PDU. Ambos procesos pueden ser iniciados ya sea por la estación Base para una conexión de bajada o por una SS para una conexión de subida.

El protocolo 802.16 permite el uso tanto de la fragmentación como del empaquetamiento simultáneamente para usar de manera más eficiente el ancho de banda.

3.4.2.4 La capa física y la estructura de trama

El IEEE 802.16 soporta tanto TDD como FDD. En FDD se puede transmitir tanto de manera continua como por ráfagas. El tráfico en ráfagas en un enlace de bajada, ya sea FDD o TDD, permite el uso de técnicas para mejorar la capacidad que son más robustas y avanzadas que aquellas utilizadas con el tráfico continuo. Un ejemplo de lo anterior son los sistemas de antenas más avanzados, y perfiles de ráfagas adaptativos a nivel del suscriptor.

El MAC elabora la subtrama de bajada iniciando con una sección de control de trama que contiene los mensajes DL-MAP y UL-MAP. Estos indican las transiciones físicas en el enlace de bajada así como las asignaciones de ancho de banda y los perfiles de ráfagas en el enlace de subida.

El DL-MAP es válido para el frame actual y siempre tiene el largo de 2 bloques FEC. La primera transición FEC se transmite en el primer bloque FEC, para permitir un tiempo de procesamiento adecuado. Tanto en sistemas TDD como FDD, el UL-MAP proporciona asignaciones que no van más allá del siguiente mapa. Sin embargo, el UL-MAP puede asignar slots en la trama actual siempre y cuando se cumpla con el tiempo de procesamiento y con el retardo de ida y vuelta. En la figura 3.5, se muestra el tiempo mínimo entre la recepción y la puesta en práctica del UL-MAP para un sistema FDD.

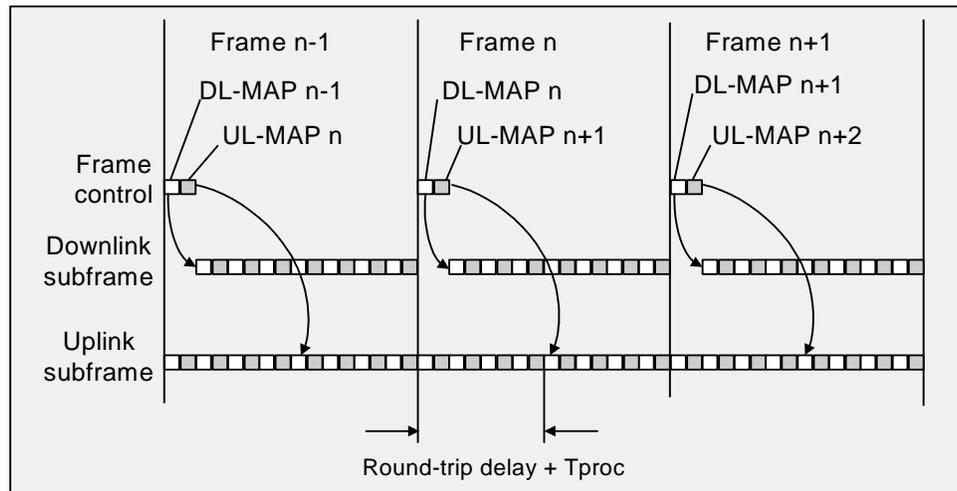


Figura 3.5. Tiempo mínimo entre la recepción y la aplicación de un mapa UL-MAP.

3.4.2.5 Control del radio enlace

La tecnología avanzada de la capa física del protocolo 802.16, requiere un control del radio enlace igualmente avanzado (RLC). Particularmente la capa física debe tener la capacidad para hacer transiciones de un perfil de ráfagas a otro. Además debe hacer las funciones tradicionales RLC tales como el control de potencia y ranging.

El control de radio enlace (RLC) comienza con envíos periódicos broadcast que realiza la BS con los perfiles de ráfagas que se han elegido para los enlaces ascendente y descendente.. Los perfiles de ráfagas particulares que se usan en un canal son elegidos con base en un número de factores, tales como la lluvia en la zona y las capacidades del equipo. Cada uno de los perfiles de ráfagas en el enlace descendente está etiquetado con un Código de Uso del Intervalo Descendente o DIUC por sus siglas en inglés (Downlink Interval Usage Code). Aquellos perfiles para el canal ascendente están etiquetados con un Código de Uso de Intervalo Ascendente o UIUC (Uplink Interval Usage Code).

Durante el acceso inicial, la SS hace un ajuste inicial de potencia y ranging usando mensajes de solicitud de ranging (RNG-REQ). Los ajustes requeridos en las SSs, tanto en el tiempo de transmisión, como en la potencia, son enviados en los mensajes de respuesta de ranging (RNG-RSP). Para ajustes posteriores de potencia y de ranging, la BS puede transmitir mensajes no solicitados de RNG-RSP indicándole a la SS que debe ajustar nuevamente el tiempo o la potencia de transmisión.

Durante el ranging inicial, la SS también solicita ser atendida en el enlace descendente a través de un perfil de ráfagas particular transmitiendo el DIUC elegido a la estación base. Esta elección esta basada en la medición de la calidad de la señal de bajada que hace la SS al inicio y durante el ranging. La estación base puede aceptar o rechazar esta petición en la respuesta de ranging (RNG-RSP). Similarmente, la estación base monitorea la calidad de la señal ascendente que recibe de la SS. La estación base ordena a la SS utilizar un determinado perfil de ráfagas al incluir éste en el UIUC junto con las asignaciones o *grants* en el mensaje

UL-MAP.

Después de determinar los perfiles de ráfagas ascendente y descendente entre la BS y un SS en particular, el RLC continúa monitoreando y controlando los perfiles de ráfagas. Las condiciones ambientales adversas, tales como la lluvia pueden forzar a la SS a solicitar un perfil más robusto. Alternativamente, si existen condiciones ambientales excepcionalmente buenas, se puede permitir a la SS operar con un perfil de ráfagas más eficiente. El RLC continúa adaptando los perfiles ascendente y descendente de cada SS, siempre tratando de lograr el mejor balance entre robustez y eficiencia. Debido a que la BS tiene el control y monitorea directamente la calidad de la señal ascendente, el protocolo para cambiar el perfil de ráfagas de subida para un SS es bastante simple; la BS simplemente especifica el UIUC asociado del perfil y lo añade a las oportunidades de transmisión concedidas a la SS en una trama. Este mecanismo elimina la necesidad de un mensaje de confirmación, ya que la SS recibirá siempre tanto el UIUC como el *grant* asignado o ninguno de los dos. De esta manera no hay oportunidades para que aparezca una diferencia entre los perfiles que manejan tanto la BS como la SS.

En el enlace descendente, la SS es la entidad que monitorea la calidad de la señal recibida y en consecuencia sabe cuando ha llegado el momento de cambiar su perfil de ráfagas descendente. La BS, sin embargo, es la entidad que controla los cambios. Existen 2 métodos disponibles para que la SS solicite un cambio en el perfil de ráfagas descendente, dependiendo del modo en que esté operando, ya sea en el modo de asignación por conexión (GPC) o asignación por estación suscriptor (GPSS). El primer método debería de aplicarse únicamente a las SS operando en modo GPC (aunque en realidad depende del algoritmo de asignación de *grants*). Si la SS está trabajando en el modo de asignación por conexión, la BS puede asignar periódicamente un intervalo de mantenimiento a la SS. La SS puede utilizar los mensajes RNG-REQ para solicitar un cambio en el perfil descendente de ráfagas. En el segundo método la SS envía un mensaje de petición de cambio del perfil de ráfagas descendente o DBPC-REQ (downlink burst profile change request). En este caso, que siempre es una opción cuando la SS opera en el modo GPSS y que puede ser una opción para las SS operando en el modo GPC, la BS responde con un mensaje de respuesta de cambio de perfil de ráfagas de bajada o DBPC-RSP (Downlink Burst Profile Change Response) confirmando o denegando el cambio.

Debido a que a veces ocurren errores irre recuperables en algunos bits, en ocasiones pueden perderse mensajes completos. Debido a lo anterior los protocolos para cambio de perfil de ráfagas de bajada deben ser estructurados cuidadosamente. El orden en que se llevan a cabo las acciones para un cambio en el perfil de ráfagas es diferente cuando la transición es a un perfil más robusto que cuando se cambia a uno menos robusto. El estándar toma ventaja del hecho de que una SS siempre escucha, además del perfil que negoció, a los otros perfiles más robustos del enlace. La figura 3.6 muestra la transición a un perfil de ráfagas más robusto mientras que la figura 3.7 muestra la transición a uno menos robusto.

3.4.2.6 Calidad de Servicio (QoS)

Cada conexión de subida es mapeada en un servicio de calidad de servicio o asignación de slots. Cada servicio programado está sujeto a un conjunto de reglas impuestas en el “scheduler” o calendarizador de slots de la BS responsable de la asignación de la capacidad

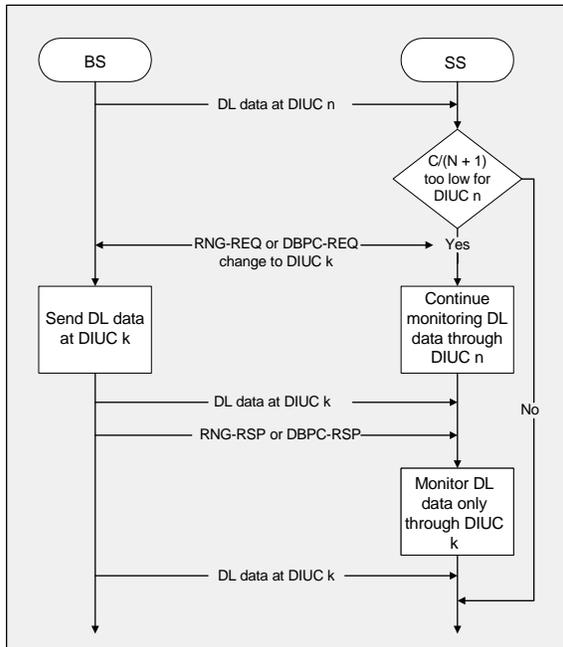


Figura 3.6. Transición a un perfil más robusto.

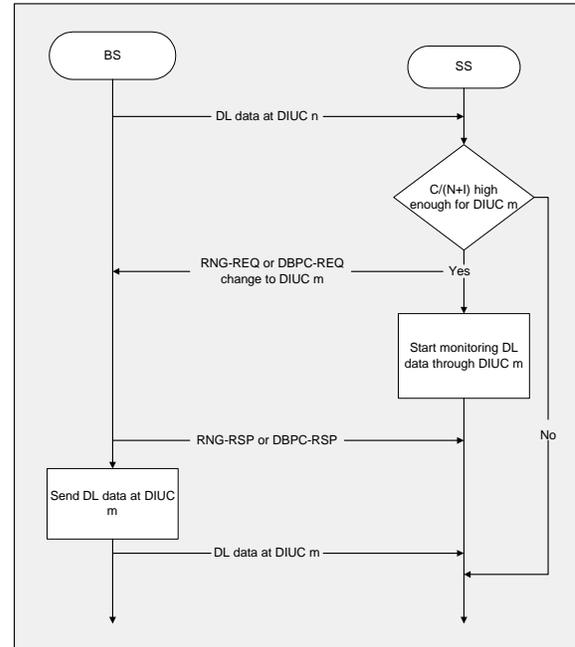


Figura 3.7. Transición a un perfil menos robusto.

ascendente y del protocolo de petición y concesión de *slots* o *request-grant* entre la SS y la BS. Las especificaciones detalladas de las reglas y el servicio de calendarización utilizado para una conexión ascendente en particular son negociadas al momento de establecer la conexión.

Los servicios de calendarización en IEEE 802.16 están basados en aquellos definidos para cable módems en el estándar DOCSIS [19].

UGS: El servicio garantizado no solicitado o UGS (Unsolicited Grant Service) es un servicio de tasa de bits constante y está diseñado para transportar datos de los servicios que generan un número fijo de unidades de tráfico periódicamente. Aquí la estación base calendariza o programa oportunidades de transmisión de manera regular, de manera anticipada y del tamaño negociado con anterioridad, durante el establecimiento de la conexión, sin que haya una petición explícita de parte de la SS. Esto elimina la sobrecarga y el periodo de latencia de las peticiones de ancho de banda con el fin de cumplir con los requerimientos de retardo y con el *delay jitter* del servicio que se trate. Un límite práctico para el delay jitter es aquel establecido por la duración de la trama. Si se requiere satisfacer un jitter más restringido, entonces se deberán utilizar búffers a la salida. Los servicios que normalmente deben utilizar conexiones con servicio UGS incluyen ATM con tasa de bits constante y E1/T1 sobre ATM.

Cuando se utiliza el servicio *UGS*, el subencabezado de control del *grant* (Grant Management subheader) incluye tanto el bit *poll-me* así como el bit SI (Slip Indicator). La bandera SI permite a la estación suscriptora reportar que el número de elementos en la cola de transmisión ha excedido cierto límite debido a factores como la pérdida de grants o la pérdida de sincronía entre el reloj del sistema IEEE 802.16 y la red externa. La BS después de detectar la bandera SI puede asignar capacidad adicional a esa SS con el fin de que la cola recupere su longitud normal. Este mecanismo se utiliza debido a que a las estaciones que trabajan con servicio

UGS no les está permitido utilizar las oportunidades de acceso aleatorio para solicitar ancho de banda.

rtPS: El Servicio de Consulta en Tiempo Real o rtPS por sus siglas en inglés (real-time Polling Service), es un servicio de tasa de bits variable en tiempo real. Este servicio está diseñado para satisfacer las necesidades de servicios que son de naturaleza dinámica. A las estaciones con tráfico rtPS se les asignan oportunidades para solicitar ancho de banda de manera periódica con el fin de satisfacer sus requerimientos en tiempo real. Debido a que la estación suscriptora debe hacer una petición explícita para cada trama o frame, el overhead del protocolo y la latencia se incrementan, pero cada estación es atendida de acuerdo a las necesidades reales de la conexión. El servicio rtPS es adecuado para conexiones que dan servicios como VoIP o video o audio en tiempo real (streaming video o streaming audio).

nrtPS: El Servicio de Consulta no en Tiempo Real o nrtPS (non-real-time Polling Service) es un servicio de tasa de bits variable que no es en tiempo real. Es casi idéntico al servicio rtPS, excepto que las conexiones pueden utilizar las oportunidades de transmisión de acceso aleatorio o slots de contención para solicitar ancho de banda. Los servicios utilizados en este tipo de conexiones toleran retardos más largos y no son sensibles al delay jitter. El servicio nrtPS es adecuado para el acceso a Internet cuando no se tiene una tasa mínima garantizada y para conexiones ATM GFR.

BE: El servicio de mejor esfuerzo o BE (Best Effort) también está definido en el protocolo 802.16. Para este servicio no se garantiza un throughput ni tampoco un retardo específico. La SS envía una petición solicitando ancho de banda en las ranuras de acceso aleatorio u oportunidades de transmisión dedicadas. La ocurrencia de oportunidades dedicadas está sujeta a la carga de la red y la SS no puede basarse en su presencia.

3.4.2.7 Solicitud de ancho de banda y asignación de oportunidades de transmisión o grants

El protocolo IEEE 802.16 MAC define 2 clases de SS, las cuales se distinguen por su habilidad de aceptar oportunidades de transmisión. El primer grupo recibe el nombre de "Asignación por Conexión o GPC por sus siglas en inglés (Grant per Connection), mientras que las segundo grupo se llama Asignación por Estación Suscriptora o GPSS (Grant per Subscriber Station). Ambos tipo de estaciones suscriptoras solicitan un ancho de banda por cada conexión. Gracias a esto el algoritmo programador de slots o "scheduler" puede tomar en cuenta la calidad de servicio o QoS cuando hace las asignaciones de ancho de banda.

Generalmente una estación suscriptora cuenta con diversas conexiones activas durante su operación. Para las estaciones operando en modo GPC esto puede implicar que en algunos casos sea necesario reprogramar la transmisión del tráfico generado por alguna de sus conexiones o, inclusive, descartarla. Lo anterior se debe a que para una conexión específica es posible que no se garantice el que siempre cuente con reservaciones, o bien, que éstas no sean suficientes. Las SS que pertenecen al grupo de GPC, reciben grants destinados específicamente para una conexión y no la pueden utilizar para ningún otro fin. Los protocolos de administración y control, tales como RLC utilizan el ancho de banda asignado específicamente para las conexiones de control.

Las estaciones del tipo GPSS reciben un ancho de banda que busca satisfacer al conjunto de necesidades de todas sus conexiones, ya que las asignaciones realizadas por la BS se generan con base en el CID básico de la estación suscriptora. Estas estaciones necesitan manejar de manera inteligente el QoS. Cuando una SS recibe oportunidades de transmisión no lo asignará necesariamente a aquella conexión que solicitó ancho de banda. La SS puede establecer a cual o cuales conexiones de las que atiende podrán contar con la reservación o parte de ella mejorando la eficiencia en la transmisión, ya que pudieran realizarse decisiones inmediatas conforme varía la demanda del acceso al canal inalámbrico con cada una de las conexiones establecidas a través de la estación suscriptora. Por ejemplo, podría suceder que después de la última petición de ancho de banda cambie la situación de QoS. Cuando la SS reciba el ancho de banda, lo destinará a aquella conexión que tenga una prioridad mayor de QoS y al mismo tiempo hará una nueva solicitud de ancho para reemplazar al que se utilizó de la conexión que tenía menor QoS. La SS puede usar una porción del ancho de banda para reaccionar más rápidamente a las condiciones ambientales. Un ejemplo de lo anterior sería el envío de mensajes DBPC-REQ (downlink burst profile change request).

Las dos clases de SS permiten un intercambio entre la simplicidad y la eficiencia. La necesidad de hacer una petición por un mayor ancho de banda para RLC junto con el hecho de que se pueden tener varias conexiones en cada SS hace al modo GPC menos eficiente y escalable que GPSS. Adicionalmente, la habilidad de una estación en modo GPSS de reaccionar más rápidamente las necesidades de la capa física y de las conexiones mejora el desempeño del sistema. El modo GPSS es la única clase de SS permitidas para operar en el rango de 10-88 GHz.

Con ambas clases de grants, la MAC 802.16 utiliza un protocolo de autocorrección de errores más que uno de confirmaciones. Este método permite utilizar un ancho de banda menor. Además, aquellos protocolos que hacen uso de confirmaciones requieren más tiempo de operación y por consiguiente ocasionan retardos mayores.

En algunas ocasiones, algunas conexiones pueden solicitar oportunidades de transmisión a la estación sin recibir respuesta alguna. Hay varias razones por las cuales el ancho de banda solicitado por una SS para una conexión puede no estar disponible. Algunas de ellas son las siguientes:

- La estación base no recibe la petición debido a errores irre recuperables en la capa física o a que se presentaron colisiones en la zona de contención.
- La estación suscriptora no reconoce el grant asignado debido a errores irre recuperables en la capa física.
- La estación base no tiene ancho de banda disponible para responder a esa demanda.
- La estación suscriptora GPSS destina el ancho de banda para otra función.

Los protocolos autocorrectivos manejan cada una de las anomalías anteriores de la misma manera. Las SS hacen una petición y a continuación esperan un lapso de tiempo que varía dependiendo de la QoS que maneje la conexión. Si después de que transcurre el tiempo máximo de espera (Timeout) no se recibe ninguna respuesta por parte de la BS (o inmediatamente, si el ancho de banda fue utilizado por la SS para otro propósito), la SS simplemente hace la petición nuevamente. Por razones de eficiencia, la mayoría de las

peticiones de ancho de banda son incrementales. Es decir, después de cada petición sin respuesta, se hará una nueva petición solicitando más ancho de banda para la conexión de que se trate. Sin embargo, para que el mecanismo de petición y asignación de ancho de banda (request/grant) con autocorrección funcione adecuadamente, las peticiones de ancho de banda ocasionalmente deberán ser agregadas. Esto quiere decir que la SS deberá informar a la BS de sus necesidades totales de ancho de banda en ese momento determinado. Al hacer esto, la BS renueva su percepción de las necesidades de las SSs sin el uso de un protocolo más complicado para la asignación del ancho de banda extra.

La estación base tiene una gran variedad de maneras distintas de solicitar ancho de banda. En la solicitud de ancho de banda se combinan características de los requerimientos de los diferentes tipos de tráfico, tales como el determinismo en las consultas unicast, y la eficiencia de la asignación de ancho de banda no solicitado. Por ejemplo, cuando hay una demanda continua de ancho de banda como en las transmisiones que las que se requiere una tasa de bits constante (CBR), tales como T1/E1, la SS no necesita solicitar ancho de banda. En estos casos la BS debe contemplar la asignación de ancho de banda sin recibir solicitudes de manera periódica.

Para acortar el ciclo normal de la consulta (polling) una SS que tenga una conexión UGS activa puede utilizar el bit de poleo poll-me en el subencabezado de management. Al hacer esto le hará saber a la BS que necesita ser interrogada nuevamente ya que sus necesidades han cambiado debido a una conexión nueva. La estación base puede ahorrar ancho de banda haciendo un sondeo ("polling") a aquellas SS con servicios UGS que tengan habilitado el bit de poleo (poll-me).

Una manera más convencional de solicitar ancho de banda es enviar una petición de ancho de banda MAC PDU. Ésta consta simplemente del encabezado de una petición de ancho de banda sin carga útil o payload. Las SSs GPSS pueden enviar esta petición en cualquier asignación de ancho de banda que tengan disponible. Las terminales GPC pueden enviar esta petición en un intervalo ya sea de sondeo o de asignación de datos de su conexión básica. Un método muy semejante para solicitar ancho de banda adicional para una misma conexión consiste en hacer una petición anidada (piggy back) en el subencabezado de management dentro del MAC PDU.

Además de la consulta individual a una SS, la BS puede hacer un sondeo broadcast. Cuando la Estación Base no cuenta con ancho de banda suficiente para proporcionarlo a cada una de las Estaciones Suscriptoras a las que atiende, puede emplear Consultas realizadas a un grupo de SSs. Estos grupos se identifican mediante CIDs exclusivos y las oportunidades que reciben se especifican mediante el mapa descriptor del canal ascendente. Las estaciones Suscriptoras involucradas sólo envían solicitudes si lo necesitan, participando en contenciones con las otras SS de su grupo. Debido a que se encuentran sujetas a la ocurrencia de colisiones, emplean el algoritmo exponencial backoff tanto para determinar el minislot en el que enviarán su solicitud como para reprogramarla en el caso de haberse presentado una colisión. Las solicitudes se considerarán exitosas sólo si en algún futuro mapa descriptor del canal ascendente se incluye una reservación para la CID en cuestión. El parámetro Random Access Timeout, incluido en el mensaje Descriptor del Canal Ascendente UCD, indica cuántos mapas esperará la Estación Suscriptora que ha hecho una solicitud en contención antes de considerarla fracasada. Debido al retardo no determinístico que se genera a causa de las colisiones y los intentos repetidos,

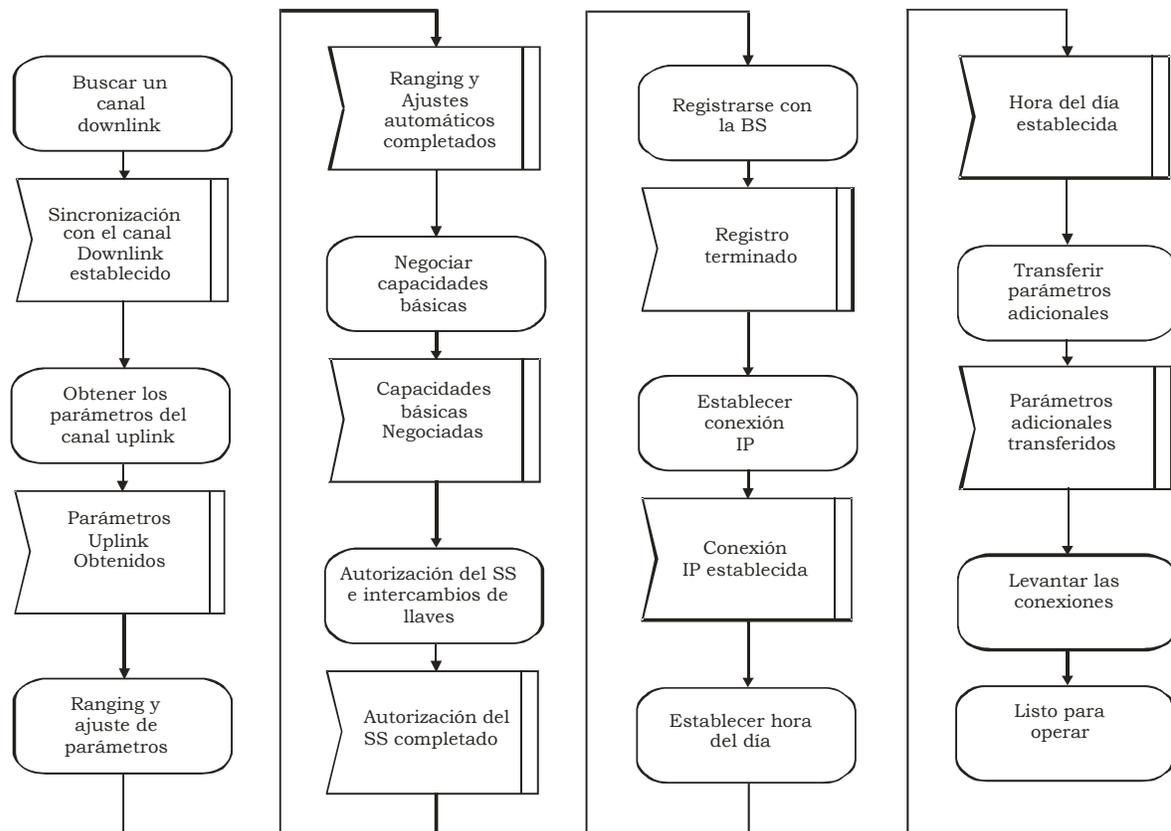


Figura 3.8. Diagrama de inicialización y registro de un nodo con la estación Base

las peticiones basadas en contención son permitidas únicamente para las clases de servicio que tienen un QoS muy bajo.

3.4.2.8 Adquisición del canal

El protocolo MAC incluye un procedimiento de inicialización diseñado para eliminar la necesidad de una configuración manual. Después del proceso de instalación, una SS comienza a explorar su lista de frecuencias de trabajo con el fin de encontrar un canal en operación. Cada UCD cuenta con un parámetro de identificación (ID) que permite distinguir a los canales ascendentes entre sí. Al recibir el UCD, la estación suscriptora almacenará la información correspondiente y verificará la disponibilidad del canal. En caso de no poder usarlo, la SS buscará otro canal de acuerdo con los UCD que reciba. Al agotar los canales ascendentes, se reiniciará el proceso de la búsqueda de canal descendente. También puede estar programada para registrarse con una estación base específica utilizando el BS ID que es enviado periódicamente por cada BS. Esta función es muy útil en lugares con una infraestructura muy densa. En estos lugares una estación suscriptora puede escuchar a una estación secundaria debido a una atenuación de la señal primaria. También puede suceder que la SS alcance a captar un lóbulo lateral de la antena de la Estación Base más cercana.

Después de decidir el canal o los canales que se utilizarán para intentar establecer comunicación, la SS tratará de sincronizarse con el enlace descendente. Para ello tratará de detectar de manera periódica los preámbulos de cada trama. Después de que la SS ha

sincronizado correctamente la capa física, comenzará a procesar la información proveniente de los mensajes periódicos UCD y DCD. La SS deberá ser capaz de recibir correctamente los mensajes UCD y UL-MAP periódicamente para que la conectividad a través del canal ascendente que esté usando se considere normal, de lo contrario se eliminarán los parámetros correspondientes al canal y se reiniciará el proceso de obtención de los parámetros. Gracias a la información obtenida a través de esos mensajes, la SS podrá conocer los esquemas de modulación y de corrección de errores utilizados en la señal portadora.

3.4.2.9 Ranging inicial y negociación de las capacidades básicas de la SS

El proceso de Ranging consiste en la sincronización de la Estación Suscriptora con la Estación Base en el canal ascendente con el fin de que las transmisiones de la estación suscriptora estén alineadas al inicio de un minislot. Inmediatamente después de saber qué parámetros deberá usar para las transmisiones iniciales de ranging, la SS explorará los mensajes UL-MAP presentes en cada trama buscando oportunidades para hacer su ranging inicial. La estación suscriptora utiliza un algoritmo exponencial backoff para determinar el slot de ranging que deberá utilizar para enviar el mensaje de petición de ranging. Durante un intervalo de Mantenimiento Inicial, la SS enviará un mensaje de RNG-REQ considerando los retrasos de la capa física. La potencia con la que se envíe este mensaje será mínima y se incrementará en mensajes sucesivos en caso de que no se obtenga respuesta de la estación base. Cuando la BS reciba el mensaje RNG-REQ exitosamente, contestará enviando la respuesta RNG-RSP, incluyendo los CID de las conexiones básica y primaria de la Estación Suscriptora correspondiente, así como cualquier otra información para el ajuste de potencia y sincronización en la SS. A partir de entonces, los ajustes que tenga que realizar la SS se llevarán a cabo en la región de Mantenimiento de Estación mediante mensajes RNG-REQ. El proceso se repite hasta que la Estación Base notifique que el Ranging se ha completado exitosamente.

Hasta este punto, todas las transmisiones se han hecho utilizando el perfil de ráfagas más robusto y menos eficiente. Con el fin aprovechar mejor la capacidad del sistema, la estación suscriptora informará a la estación base de sus características particulares en la capa física mediante un mensaje SBC-REQ. Esta información contempla:

- Número de CIDs ascendentes que soporta la SS. Deben ser al menos tres, correspondientes a las Conexiones Básica, Primaria y Secundaria.
- Cantidad de parámetros considerados en las capacidades físicas de la SS.
- Demoduladores y moduladores en la Estación Suscriptora. Puede considerarse QPSK, 16-QAM o 64-QAM.
- Tipo de Código FEC usado en las transmisiones ascendente y descendente.
- Número máximo de transacciones PKM (Control de Llaves de Privacidad), MCA (Asignación a grupos Multicast) y DSA, DSC o DSD (Adición, Modificación o Eliminación Dinámica de Servicio) concurrentes en espera.
- Propiedades de la asignación de ancho de banda como soporte a asignaciones por Conexión o por Estación Suscriptora, o comunicaciones half o full dúplex.
- Versión del protocolo IP.

- Soporte a CRC en mensajes MAC.
- Soporte a Consultas para grupos Multicast.

La estación base, en respuesta, puede aceptar o denegar el uso de alguna de las capacidades reportadas por la SS.

3.4.2.10 Autenticación y registro de la SS

Los procesos involucrados en la autenticación e intercambio de llaves son:

- Autenticación de la identidad de la Estación Suscriptora por parte de la Estación Base.
- La Estación Base proporciona la Llave de Autorización (AK) con la que se obtiene la Llave de Cifrado de Llave (KEK) y las llaves de autenticación de mensajes.
- La Estación Base asigna las identidades y propiedades de asociaciones de seguridad primarias y estáticas para las cuales la Estación Suscriptora está autorizada a obtener información.

Cada SS contiene tanto un certificado digital X.509 instalado en el momento de fabricación como un certificado del fabricante. Estos certificados establecen un enlace entre la dirección MAC de 48 bits de la SS y su llave pública RSA. Los certificados son enviados a la Estación Base por la estación suscriptora en los mensajes de petición de autorización y de Información de autenticación. La red tiene la capacidad de verificar la identidad de la estación suscriptora revisando los certificados y puede subsecuentemente revisar el nivel de autorización de la estación suscriptora. Si la estación suscriptora está autorizada para agregarse a la red, la estación base responderá a su petición con una contestación de autorización conteniendo la llave de autorización (AK) encriptada con la llave pública SS, una secuencia numérica de 4 bits para identificar cada generación de AKs, la duración de la llave, las identidades (SAID, identificador de Asociación de Seguridad) de las SA (Asociaciones de Seguridad) primarias o estáticas para las cuales la SS está autorizada para obtener información.

Después de haber concluido la etapa de autenticación exitosamente, la SS se registrará en la red. El objetivo de la etapa de registro es establecer la conexión de control secundario y determinar las capacidades relacionadas con el establecimiento de la conexión y la operación MAC. Para conseguirlo, la SS envía un mensaje REG-REQ y recibe un REG-RSP que incluye el CID de Control Secundario. Adicionalmente, puede informarse la versión del protocolo IP empleado en caso de que lo haya solicitado la SS. La versión empleada por defecto es IPv4. La estación base espera un mensaje TFTP- CPLT (Configuración de Archivo TFTP Completada) para considerar finalizado el proceso de registro.

3.4.2.11 Conectividad IP

Después del proceso de registro, la Estación Suscriptora adquiere la dirección IP que le asigna la estación Base mediante el protocolo DHCP a través de un canal secundario y establece la hora del día a través del protocolo de tiempo de Internet.

La configuración de la fecha y la hora sólo son necesarias para el registro de eventos para fines operativos del sistema. No es necesario proteger el intercambio de esta información, y la resolución empleada es de un segundo. Se emplea la Conexión de Control Secundario y el

protocolo UDP.

El servidor DHCP también proporciona la dirección del servidor TFTP al cual la Estación suscriptora puede solicitarle un archivo de configuración. Este archivo proporciona una interfaz estándar para la información de configuración proporcionada por un vendedor específico.

3.4.2.12 Establecimiento de conexión

El protocolo 802.16 utiliza el concepto de flujos de servicio para definir el transporte unidireccional de paquetes ya sea de manera ascendente o descendente. Los flujos de servicio están caracterizados por un conjunto de parámetros de QoS tales como el jitter y la latencia. Con el fin de utilizar los recursos de la red de una manera eficiente, tales como ancho de banda y memoria, el 802.16 adopta un modelo de activación en 2 fases en el cual los recursos asignados a un flujo de servicios particular pueden no estar comprometido sino hasta que se activa el flujo. Cada servicio activo o admitido es mapeado a una conexión MAC con una identificación CID única.

En general, los flujos de servicio en IEEE 802.16 son preaprovisionados y el establecimiento del flujo de servicio es iniciado por la BS durante la inicialización de la SS. Sin embargo, el flujo de servicio puede establecerse dinámicamente ya sea por la BS o por la SS. La SS típicamente inicia el flujo de servicio únicamente si hay una conexión señalizada dinámicamente, por ejemplo una conexión virtual conmutada (SVC) de una red ATM. El establecimiento de un flujo de servicio se hace a través de un protocolo de handshaking de 3 vías, en el cual la petición para el establecimiento de un flujo de servicio es contestada y a su vez la contestación es confirmada.

Adicionalmente al establecimiento dinámico de servicio, el IEEE 802.16 también soporta cambios dinámicos en el servicio, en los cuales los parámetros del flujo de servicio son renegociados, los cambios en el flujo de servicio también siguen un protocolo de 3 vías similar al anterior.

3.4.2.13 Subcapa de privacidad

EL protocolo de privacidad IEEE 802.16 está basado en el control de llaves de privacidad o PKM (Privacy Key Management) de las especificaciones DOCSIS + BPI, pero ha sido mejorado para ajustarse de una manera consistente dentro del protocolo IEEE 802.16 y para acomodar métodos criptográficos más fuertes, tales como el Estándar de Encriptación Avanzada.

3.4.2.14 Asociaciones de seguridad

El control de llaves de privacidad o PKM por sus siglas en inglés, está desarrollado alrededor del concepto de asociaciones de seguridad o SAs. La SA es un conjunto de métodos criptográficos y del material de llaves asociado; esto es, contiene la información acerca de qué algoritmos aplicar, qué llaves se deben utilizar, etc. Cada SS establece al menos una SA durante la inicialización. Cada conexión, con excepción del las conexiones de control básico y primario, es mapeada a una SA ya sea al momento de establecer la conexión o dinámicamente durante la operación.

3.4.2.15 Métodos de encriptación

Actualmente, el protocolo PKM utiliza certificados digitales X.509 con encriptación pública RSA para autenticación de la SS e intercambio de llaves. Para la encriptación de tráfico, es obligatorio el uso del Estándar de Encriptación de Datos (DES) corriendo en modo de cadena de bloque cifrado (CBC) con una llave de 56 bits. El vector de inicialización CBC depende del contador de frames y difiere de frame a frame. Con el fin de reducir el número de operaciones con llaves públicas durante la operación normal, las llaves de encriptación de transmisión son intercambiadas utilizando 3DES junto con una llave de intercambio derivada de la llave de autorización.

Los mensajes del protocolo PKM son autenticados utilizando el protocolo HMAC (Hashed Message Authentication Code) con SHA-1. Además, la autenticación de mensajes en funciones vitales MAC, como el establecimiento de la conexión se hace con el protocolo PKM.

DISEÑO E IMPLEMENTACIÓN DEL MODELO DE SIMULACIÓN DEL PROTOCOLO IEEE 802.16

Basados en el modelo de red, se elaboró un modelo de simulación en Opnet, el cual consta de una estación base y un número variable de estaciones suscriptoras con generadores de tráfico de diferentes tipos. Algunas de estas estaciones generan paquetes de tráfico de Internet, algunas otras generan tráfico de voz (VoIP) y algunas otras generan tráfico de video (mpeg 4). A continuación se describe con detalle cada uno de los elementos que componen la red, tanto en la estación base como en las estaciones suscriptoras, así como el algoritmo que emplea la estación base para satisfacer los diferentes requerimientos de calidad de servicio de cada una de las estaciones suscriptoras de la red.

4.1 Introducción

Con el fin de simular el comportamiento de una red inalámbrica de banda ancha, se implementó un modelo de simulación del protocolo MAC 802.16 utilizando Opnet Modeler Package V. 11.0. Opnet Modeler es un simulador de sistemas de comunicación disponible comercialmente, creado y mantenido por Opnet Technologies, Inc [42]. Opnet Modeler fue diseñado para ayudar en el diseño y pruebas de protocolos de comunicación y para simular el desempeño de redes. Opnet proporciona una interfase gráfica jerárquica para la definición de sistemas de comunicaciones y protocolos. En el nivel jerárquico más alto se encuentra la red, en el siguiente nivel se encuentran los nodos y finalmente los procesos. Una red se construye gráficamente conectando los nodos de comunicación a través de enlaces de comunicación. Los nodos se construyen definiendo modelos de procesos conectados entre sí. Un proceso individual típicamente representa diferentes capas de la pila de protocolos, pero es lo suficientemente flexible para ser definido por el usuario. Los modelos de procesos se construyen especificando diagramas de estado con transiciones condicionales ó automáticas. El comportamiento durante cada estado está definido por eventos y por un código en proto-C que puede ser modificado por el usuario (ver figura 4.1).

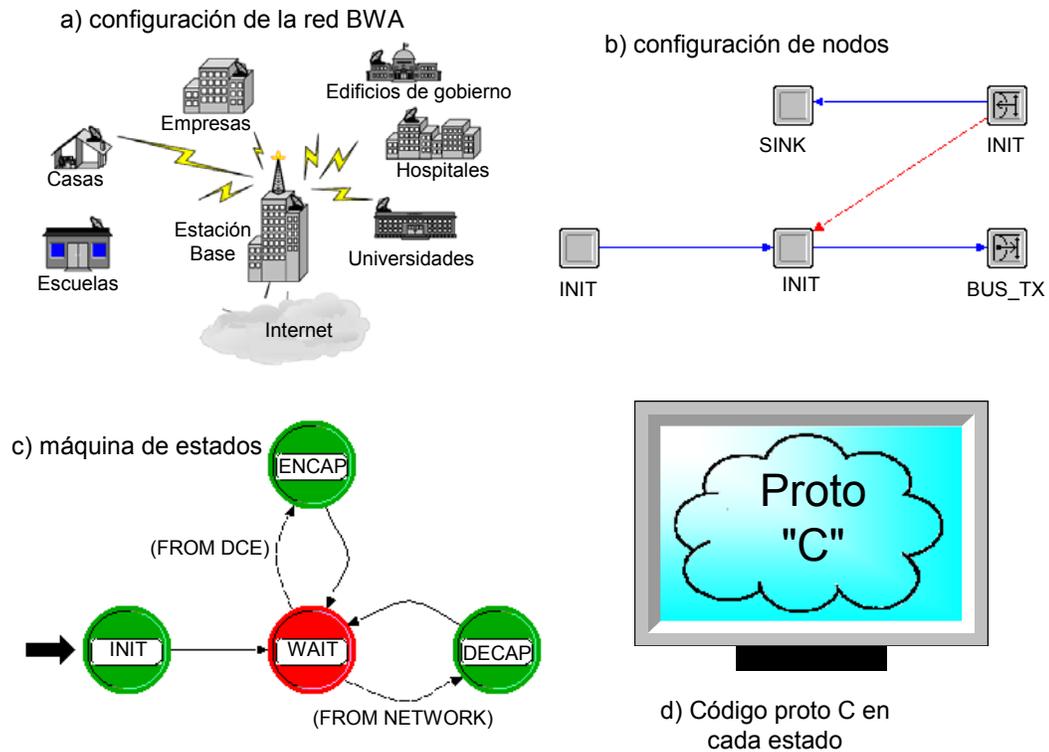


Figura 4.1. Niveles jerárquicos de las simulaciones con Opnet.

Como se mencionó anteriormente, el código que se utiliza para definir los modelos está escrito en proto C, es decir, es un lenguaje parecido al lenguaje C con funciones especializadas disponibles para facilitar la funcionalidad de las simulaciones de sistemas de comunicación. Este código, además de que puede ser modificado por el usuario, incluye llamadas para establecer y manejar colas, retardos de proceso, manejo de eventos, generación aleatoria de números basada en una variedad de distribuciones de probabilidad, construcción y encapsulamiento de paquetes y recolección estadística de datos.

Opnet Modeler tiene una extensa librería estándar. Esta librería incluye modelos comunes de tráfico, algunos modelos de protocolos estándares que han sido utilizados y validados tanto por la industria como por instituciones académicas, distribuciones estándares de variables aleatorias y de estadísticas. Los datos almacenados de diferentes variables pueden ser superpuestos en una sola gráfica. También se puede tener una variedad de representaciones de las variables almacenadas, desde el despliegue discreto de datos, valores promedio, desviaciones estándar, etc.

Adicionalmente, Opnet Modeler tiene un módulo de radio que proporciona la habilidad de modelar los efectos de propagación en las capas superiores de la pila de protocolos. Esto incluye la habilidad de modelar los retardos de transmisión y propagación, cálculos de línea de vista, cálculos de enlace incluyendo la habilidad de especificar la interferencia, los desvanecimientos por lluvia y los modelos de difracción. Los enlaces modelados pueden

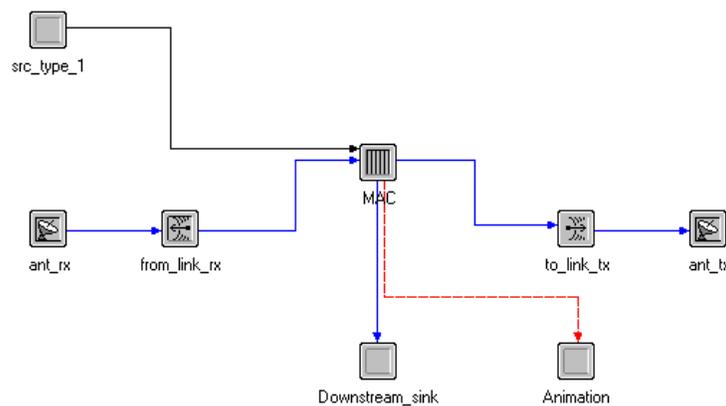


Figura 4.2. Nodo correspondiente a la estación suscritora.

ser modificados para crear modelos específicos de propagación adecuados al problema que se requiere estudiar.

4.2 Diseño de la Unidad Estación Suscritora (SS)

La Estación Suscritora (SS) es el nodo que representa al equipo transceptor del usuario. Está conformado por módulos encargados de la generación y eliminación de paquetes de información, del envío y recepción al y desde el canal inalámbrico, del cumplimiento de las funciones propias de la subcapa MAC, así como de animaciones, característica disponible en la realización de simulaciones en Opnet. La arquitectura del nodo SS se muestra en la figura 4.2.

4.2.1 Enlaces

Los enlaces que interconectan a los nodos pueden ser de los siguientes tipos:

- **Flujo de paquetes.** Transporta flujos de paquetes (colecciones de información de tamaño variable) desde la salida de un módulo generador de datos hasta la entrada de un módulo destino. Un flujo sólo puede conducir información en un solo sentido. Aunque la mayoría de los módulos pueden contar con varios flujos de paquetes, existen ciertas restricciones en las que un módulo generador sólo puede tener flujos en dirección a otros módulos y no desde otros.
- **Cable estadístico.** Envían información acerca del valor de algún parámetro desde algún nodo hacia otro. También pueden emplearse para informar sobre el estado de un módulo que los origina. Aunque estos datos pueden considerarse sólo para fines informativos en otros módulos, también son capaces de ser usados, junto con algún valor umbral, para la activación de procesos o interrupciones específicas. Los módulos cola y procesador son los únicos que pueden recibir flujos desde este tipo de conexiones.
- **Asociación lógica de transmisión/recepción.** Establece relaciones entre módulos receptores y transmisores que pertenecen a un mismo enlace. Tiene por objeto el indicar que ambos módulos representan un mismo enlace y no sean separados por asignaciones automáticas que ocurren en el editor de proyectos al tratar de identificar los módulos transmisores y receptores.

4.2.2 Descripción de las funciones de los módulos de la estación suscriptora

Los módulos que intervienen en la SS son los que se indican a continuación:

- **Generación y eliminación de paquetes de información.** Pueden incorporarse tantas fuentes de información como se desee, de acuerdo con el tipo de tráfico por transmitir (intermitente, constante, de paquetes de longitud variable, en tiempo real, etc.), en nuestro caso, se utilizó una sola fuente por nodo y se modificó el código de cada generador para la generación de cada tipo de tráfico (Internet, voz y video). Los módulos que realizan estas funciones en la SS son los siguientes:
 - Src_type_1
 - Downstream_sink
- **Procesamiento de paquetes por enviar y recibidos, y realización de las funciones MAC.** Consiste en una cola de recepción de paquetes que cuenta con un proceso administrativo y uno hijo para cumplir con las funciones más importantes de la capa MAC de acuerdo con el protocolo IEEE 802.16.
 - MAC
- **Transmisión y recepción.** Se encargan de las funciones de la capa física del modelo OSI y están compuestas por receptores y transmisores capaces de identificar señales con características específicas, así como de antenas transmisoras y receptoras. Estos módulos son los siguientes:
 - Ant_rx, Ant_tx. Son las antenas receptora y transmisora respectivamente.
 - From_link_rx, To_link_tx. Corresponden al receptor y al transmisor.
- **Funciones adicionales.** Éstas corresponden a la generación de animaciones durante una simulación en Opnet.
 - Animation

4.2.3 Módulo de generación de paquetes Src_type_1

El módulo Src_type_1 (ver figura 4.3) es el encargado de la generación de paquetes de información cuyo formato, longitud y tasa de transmisión son definidas mediante variables promovidas y especificadas por el usuario, o bien establecidas internamente. Está conectado al módulo MAC mediante una asociación lógica de transmisión/recepción cuyo flujo se origina en Src_type_1.

4.2.3.1 Proceso Bwa_gen_root

El módulo generador de paquetes es un módulo de tipo procesador que cuenta con un proceso interno consistente en el proceso administrativo, cuya máquina de estados se muestra en la figura 4.4 y que se auxilia de un proceso hijo, descrito más adelante.



Figura 4.3. Generador de tráfico.

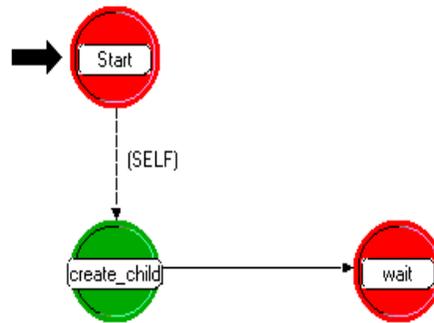


Figura 4.4. Proceso Bwa_gen_root.

Las funciones específicas de cada estado en el proceso administrativo (Bwa_gen_root) se describen a continuación:

Estado Start

Ejecutivas de Entrada

- 1) Obtiene el instante en el que se produce la generación de paquetes actual, que puede emplearse en una autointerrupción (SELF) cuyo inicio es igual al obtenido previamente.
- 2) En caso de que el número de nodos activos en la red (variable sim_active_nodes) siendo simulada sea mayor a 1, se activa la generación de tráfico intermitente mediante la programación de la interrupción SELF. También es necesario que se especifique que la generación de paquetes esté activa (variable de estado promovida traffic_generator_active) mediante el valor 1.

Estado Create_child

Ejecutivas de Entrada

- 1) En caso de que esté activa la generación de paquetes (traffic_generator_active), se invoca el proceso hijo (bwa_gen_src) por cada una de las instancias generadoras de tráfico del tipo representado por este módulo con que cuente el nodo SS.

Estado Wait

Es un estado al que procede automáticamente el módulo una vez que ha completado la invocación de los procesos hijos correspondientes.

4.2.3.2 Proceso Bwa_gen_src

El proceso hijo Bwa_gen_src se encarga de las funciones del tráfico intermitente mediante un estado activo y otro inactivo de transmisión de paquetes. Se basa en la máquina de estados mostrada en la figura 4.5.

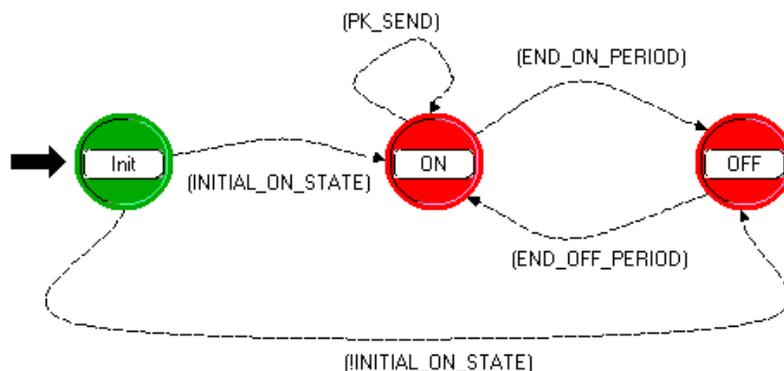


Figura 4.5. Proceso Bwa_gen_source.

Estado Init

Ejecutivas de Entrada

- 1) Identificación del módulo que invocó al proceso hijo así como el nombre del nodo al que pertenece
- 2) Se realiza la inicialización de variables que describen el comportamiento de la generación de paquetes. Éstas son variables promovidas, esto es, definidas en los parámetros de simulación:
 - a) Distribución de la duración del periodo activo de transmisión (`bwa_gen_src.on_period_dist`). Función de densidad de probabilidad empleada para determinar la duración del periodo en el que el módulo transmite información.
 - b) Distribución del periodo de transmisión de información (`bwa_gen_src.on_period_arrival_dist`). Función de densidad de probabilidad empleada para determinar el tiempo que transcurre durante la generación de información por transmitir.
 - c) Distribución de la duración del periodo inactivo de transmisión (`bwa_gen_src.off_period_dist`). Función de densidad de probabilidad empleada para determinar la duración del periodo en el que el módulo se encuentra inactivo, es decir, no transmite información.
 - d) Distribución de la longitud de los paquetes (`bwa_gen_src.pksize_dist`). Función de distribución de probabilidad empleada para determinar la longitud de los paquetes por transmitir.
 - e) Prioridad de los mensajes (`msg_priority`). Parámetro empleado para establecer preferencias cuando el nodo incluye varios módulos generadores de tráfico con distintas prioridades
 - f) Duración promedio del periodo entre la recepción de un mensaje y el siguiente (`bwa_gen_src.mean_interarrival_time`).
 - g) Duración promedio del periodo activo de transmisión (`bwa_gen_src.mean_on_time`).
 - h) Duración promedio del periodo inactivo de transmisión (`bwa_gen_src.mean_off_time`).

- i) Longitud promedio de los paquetes (`bwa_gen_src.mean_pktsize`).
 - j) Indicador de uso de distribución de paquetes (`SS_use_packet_distribution`). Especifica si se usa alguna de las distribuciones de probabilidad predefinidas para paquetes de 64, 128, 256, 512, 1024 y 1518 bytes.
 - k) En el caso del tráfico rTPs, se emplea una distribución uniforme para generar un valor entre 1 y 10, y de acuerdo al valor generado, se selecciona uno de entre diez archivos disponibles con 1500 trazas de video cada uno, correspondientes a un minuto de reproducción con 25 imágenes por segundo. Esas muestras se almacenan en una lista a la que se le asigna un apuntador y se guarda para leer las muestras más adelante.
- 3) Cálculo de parámetros iniciales como el Factor de Carga (Burstiness Factor) y de la Carga Media (Mean Load) en Mbps con base en los valores de las duraciones promedio de los periodos activo e inactivo de transmisión así como del tiempo entre la llegada de paquetes sucesivos.
 - 4) Generación de apuntadores de variables de estado para las distribuciones de los promedios de duración de los estados activo e inactivo, del tiempo entre llegadas de paquetes sucesivos, y de la longitud de los paquetes generados.
 - 5) Generación del instante de programación para la creación del paquete. Por defecto, el estado inicial del módulo es inactivo. Se genera la duración del estado de inactividad con base en la distribución especificada por `off_period_dist` y la media `mean_off_time`. Al valor obtenido se asocia la transición `INITIAL_ON_STATE`, que permite el paso hacia el estado ON u OFF. Aún cuando esta interrupción puede conducir tanto a ON como a OFF, se ha modificado la ejecutiva de manera que el estado siguiente siempre sea OFF, o bien, que la interrupción `INITIAL_ON_STATE` tenga el valor de False.

Estado ON

Ejecutivas de Entrada

- 1) En caso de que se haya activado la interrupción `END_OFF_PERIOD` desde el estado OFF, se calcula inicialmente el instante en el que finalizará la permanencia en el estado ON, con base en la distribución y en la duración del periodo activo.
- 2) También se calcula el tiempo requerido para enviar un paquete con base en la distribución del periodo de recepción de información. En caso de que el instante calculado ocurra después de que haya finalizado el periodo activo, no se producirá la programación del envío. El instante de envío se asigna a la interrupción `PK_SEND_CODE`.

Ejecutivas de Salida:

- 1) Estas ejecutivas se activan cuando ocurre la interrupción `PK_SEND_CODE`, es decir, cuando ha llegado el momento de generar un paquete. Se crea un paquete con sus campos y longitud respectiva y se envía.
- 2) En caso de que se haya especificado alguna distribución de paquetes mediante la variable `SS_use_pk_dist`, se obtiene la longitud correspondiente en bits, de lo contrario, se genera la longitud en bits de acuerdo con la distribución de longitud de paquete, `pktsize_dist`.
- 3) Para el caso de servicio rTPS, se lee una muestra de la lista de datos asignada previamente y se añaden los encabezados de las capas RTP y UDP.

Posteriormente se fragmenta de acuerdo al número de bits asignados para los paquetes PDU de la capa IP, más adelante se encapsulan para formar los paquetes MAC_PDU, de ser necesario se añaden los bytes de padding y finalmente se calcula el tamaño de la trama de video en la capa MAC_ethernet.

- 4) El paquete se crea con base en el formato de CATV, según se especifica en la variable CATV_MSG_PK_FORMAT y se asigna la longitud calculada previamente.
- 5) Se obtienen y actualizan las estadísticas que pueden usarse en análisis posteriores, bits enviados, paquetes enviados, bits globales enviados y paquetes globales enviados.
- 6) Se envía el paquete hacia el flujo de salida del módulo generador.

Estado OFF

Ejecutivas de Entrada

- 1) Cálculo del instante de finalización de permanencia en el estado OFF con base en la distribución del periodo inactivo, definido de acuerdo con off_period_dist. Este valor se relaciona con la interrupción OFF_PERIOD_CODE

4.2.4 Módulo de eliminación de paquetes Downstream_sink

Es un módulo procesador encargado de recibir los paquetes provenientes del módulo MAC a través de un flujo de paquetes. Los paquetes recibidos son procesados para obtener información relativa al tipo de información que contienen, y otros parámetros que permiten la obtención de estadísticas (ver figura 4.6).

El módulo Downstream_sink basa su funcionamiento en una sola máquina de estados, Bwa_down_sink, descrita a continuación (ver figura 4.7).

4.2.4.1 Proceso bwa_down_sink

Estado Init

Ejecutivas de Entrada

- 1) Inicializa las variables de estado dall_ete_gshandle y dall_thru_gshandle para el manejo de las estadísticas de Retraso Ethernet y Throughput, respectivamente.
- 2) Obtiene el instante para el inicio de captura de estadísticas con base en la variable stats_start_time.

Estado Wait

Es un estado en el que permanece el módulo hasta que se produzca cualquiera de las interrupciones PK_ARR o END.



Downstream_sink

Figura 4.6. Módulo de eliminación de paquetes Downstream sink.

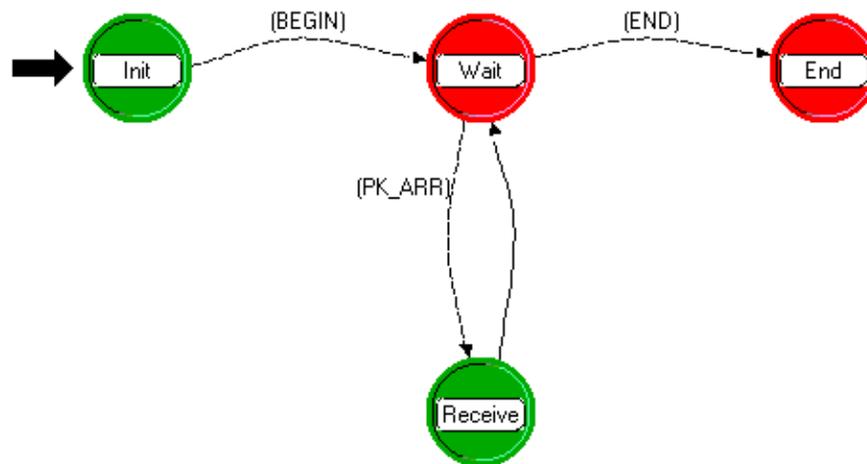


Figura 4.7. Máquina de estados del proceso bwa_down_sink.

Estado Receive

Ejecutivas de Entrada

- 1) Al recibir un paquete, se obtienen los siguientes parámetros:
 - a) Identificador de aplicación (app_id).
 - b) Tipo de fuente generadora (source_type).
 - c) Instante de generación (creation_time).
 - d) Estadísticas recolectadas (collect_stats).
 - e) Longitud (pksize).
- 2) Se calcula el retraso de la transmisión como una diferencia entre el instante actual y el de la creación del paquete.
- 3) Se destruye el paquete y se calcula el tiempo efectivo de obtención de estadísticas como una diferencia entre el instante actual y el de inicio de captura de estadísticas.
- 4) Se actualizan las estadísticas de retraso ethernet, retraso ethernet al cuadrado, paquetes recibidos y bits recibidos.
- 5) Se calculan y almacenan los valores del retraso ethernet y de rendimiento (Throughput).
- 6) Aunque no se emplean, también se definen el cálculo de retrasos y de paquetes y bits recibidos distinguiendo hasta 6 tipos de tráfico, además del correspondiente a voz.

Estado End

Ejecutivas de Entrada

- 1) Se exportan los siguientes valores en forma de estadísticas escalares:
 - a) Carga global ofrecida en el canal descendente [kbps] y [paquetes/seg].
 - b) Latencia promedio de acceso para todas las fuentes en el canal descendente [mseg].

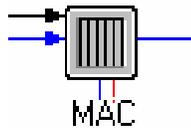


Figura 4.8. Módulo MAC.

- c) Desempeño (Throughput) para todas las fuentes en el canal descendente [kbps].
- d) Desviación estándar de la latencia en todas las fuentes en el canal descendente [mseg].
- e) Coeficiente de variación de la latencia en todas las fuentes en el canal descendente

4.2.5 Módulo para la realización de las funciones MAC

El módulo MAC es una cola que lleva a cabo el procesamiento de los paquetes de información recibidos desde capas superiores en el nodo SS, así como de aquellos provenientes del canal de transmisión inalámbrico descendente (ver figura 4.8). Este procesamiento incluye, principalmente, la agregación de información adicional (encabezados) a los paquetes, programación de oportunidades para el acceso al canal inalámbrico, envío de información durante oportunidades garantizadas por la BS, así como la resolución de colisiones durante periodos de contención. Está conformado por un proceso administrativo, `Bwa_wm_mgr`, que hace referencia a un proceso hijo, `Bwa_wm_child`.

4.2.5.1 Proceso `bwa_wm_mgr`

Los enlaces desde y hacia otros módulos son los indicados a continuación:

- Flujo de asociación lógica de transmisión/recepción desde el módulo generador de paquetes `Src_type_1`.
- Flujo de paquetes hacia el módulo procesador de estadísticas de paquetes recibidos desde el canal inalámbrico descendente y eliminador de dichos paquetes, `Downstream_sink`.
- Flujo de paquetes desde el módulo receptor de radio, `From_link_rx`.
- Flujo de paquetes hacia el módulo transmisor de radio, `To_link_tx`.
- Flujo a través de un cable estadístico hacia el módulo para la generación de animaciones durante la ejecución de simulaciones, `Animation`.

El proceso administrativo `bwa_wm_mgr` es clave en la simulación de las funciones de la subcapa de Control de Acceso al Medio del protocolo IEEE 802.16. La implementación en el modelo Opnet contempla las funciones de ranging, recepción y procesamiento de tráfico de aplicación desde capas superiores, y recepción y procesamiento de los paquetes de sincronización, UCD, MAP y de tráfico en el canal descendente, provenientes de la BS. La

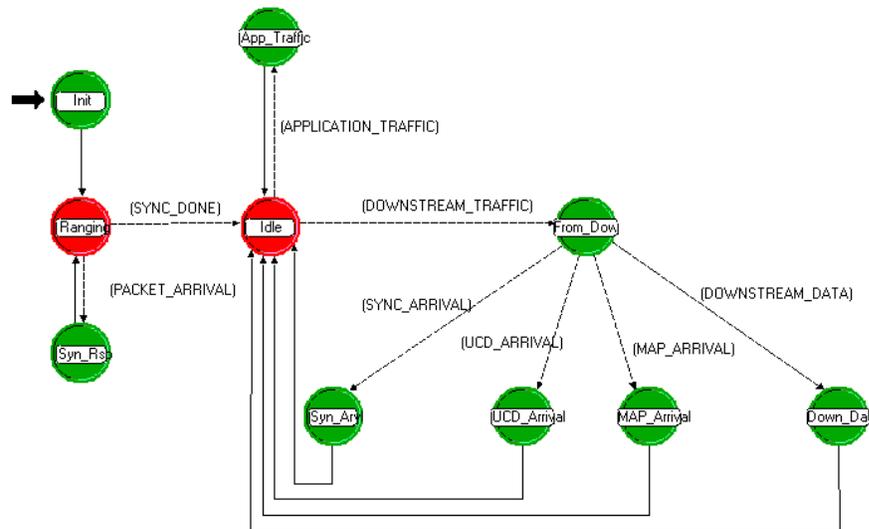


Figura 4.9. Proceso bwa_wm_mgr.

máquina de estados del proceso Bwa_wm_mgr es el que se muestra en la figura 4.9.

Estado Init

Ejecutivas de Entrada

- 1) Obtiene los atributos de simulación iniciales como el instante de Ranging, y el tamaño de los paquetes de solicitud de Ranging y el tamaño de los paquetes de control MAC.
- 2) Genera la dirección e identificador del nodo. Ambos valores son enteros, únicos y consecutivos para cada estación suscriptora.
- 3) Cuando las SS se agrupan en distintas prioridades, se sigue el siguiente procedimiento:
 - a) La primera SS en cada una de las prioridades obtiene la cantidad de estaciones que pertenecen a su prioridad. Se determina la cantidad máxima de estaciones por prioridad, Priority_Spacing, así como el valor inicial de los identificadores de SS, SID_start, ambos parámetros se obtienen desde el archivo de variables de ambiente bwa_mac_default_streams.ef. Se verifica que la topología actual cumpla con dichas especificaciones.
 - b) Se asigna una identificación a cada estación de acuerdo con la prioridad a la que pertenece. Consiste en un número consecutivo que inicia en 0. Cada una de las prioridades se identifica con el parámetro PRI (Priority Identifier).
- 4) En caso de que no se definan prioridades entre las SS, sólo se asigna la identificación a cada nodo.
- 5) Obtención de los parámetros de la estación suscriptora que incluyen:
 - a) Prioridad.
 - b) Tipo de servicio.
 - c) Máscara de política de solicitud.

- d) Modo de fragmentación.
 - e) Máxima tasa de transmisión.
 - f) Tasa de transmisión ascendente.
 - g) Tasa de transmisión descendente.
 - h) Longitud máxima de paquetes.
 - i) Longitud máxima de flujo de información (Burst).
 - j) Intervalo de solicitud unicast.
 - k) Intervalo máximo de asignación no solicitada (UGS).
 - l) Intervalo mínimo de asignación no solicitada (UGS).
- 6) Se establece la potencia inicial de transmisión de la SS mediante el siguiente procedimiento:
- a) Cálculo de la distancia entre la SS y la BS.
 - b) Obtención de la frecuencia inferior, del ancho de banda, del nivel de ruido (noise figure) y la potencia del canal de transmisión ascendente, así como de las ganancias de las antenas de la SS y la BS, con base en los parámetros de simulación.
 - c) Cálculo de la ganancia en la recepción de la estación suscriptora, el ruido ambiental, el ruido térmico, la relación señal a ruido (SNR) de la BS y la SNR efectiva en la SS.
 - d) Cálculo de las pérdidas de trayectoria considerando un terreno sinuoso con poca vegetación o un terreno plano con alta densidad de árboles, las compensaciones de la frecuencia y las alturas de las antenas de transmisión y recepción así como la atenuación (shadow fading).
 - e) Cálculo de la potencia mínima de transmisión en la SS con base en la recomendación de SNR mínima establecida en el protocolo IEEE 802.16, el ruido térmico, las ganancias de las antenas en la SS y la BS, así como las pérdidas por trayectoria.
 - f) Asignación del valor de la potencia mínima de transmisión al nodo SS correspondiente y al canal de transmisión.
- 7) Envío de una solicitud de ranging hacia la BS mediante el siguiente procedimiento:
- a) Generación del paquete de Ranging mediante el formato CATV_RNG_REQ_PK_FMT que incluye el SID (identificador de estación suscriptora).
 - b) Agregación al paquete de Ranging de la información relativa a QoS, que consiste en:
 - a. Prioridad.
 - b. Tipo de servicio.
 - c. Máscara de política de solicitud.
 - d. Modo de fragmentación.
 - e. Máxima tasa de transmisión.
 - f. Tasa de transmisión ascendente (Committed Bit Rate Up).
 - g. Tasa de transmisión descendente (Committed Bit Rate Down).
 - h. Longitud máxima de paquetes.

- i. Longitud máxima de un flujo de información (Burst).
 - j. Intervalo de solicitud unicast.
 - k. Intervalo máximo de asignación no solicitada (UGS).
 - l. Intervalo mínimo de asignación no solicitada (UGS).
- c) Encapsulamiento del mensaje de control Ranging mediante la agregación de información sobre el Tipo de paquete (CMTS_MSG_TYPE_RNG_REQ).
 - d) Agregación del encabezado MAC (CMTS_SYNC_MAC_HD_FMT).
 - e) Envío directo de la trama hacia la BS.
- 8) Ejecución de la interrupción SYNC_DONE para indicar la finalización del proceso de Ranging.

Estado Ranging

Ejecutivas de Salida

- 1) Verificación de la ocurrencia de las interrupciones SYNC_DONE y PACKET_ARRIVAL, correspondientes a la finalización del proceso de sincronización y de recepción del paquete de respuesta desde la BS, respectivamente.

Estado Syn_Rsp

Ejecutivas de Entrada

- 1) Obtiene el paquete proveniente de la BS y verifica que corresponda con el formato CMTS_MAC_FRAME_FMT.
- 2) Desencapsula el mensaje de control y verifica que tenga el formato CMTS_MNG_PKT_FMT; de no ser así, lo destruye.
- 3) Comprueba que el destinatario del mensaje sea la BS, o bien, que sea un mensaje Broadcast. De ser así, verifica que tenga sea del tipo CATV_RNG_RSP_PK_FMT.
- 4) Comprueba que el SID del mensaje corresponda con el del nodo. Si es así, obtiene del mensaje el valor del ajuste de tiempo.
- 5) Se invoca el proceso hijo, Bwa_wm_child, para registrar el valor del retraso hacia la BS.
- 6) Destrucción del mensaje y del paquete recibido.

Estado Idle

Ejecutivas de Salida

- 1) Verifica que se haya activado una interrupción correspondiente a la recepción de un paquete, PACKET_ARRIVAL, e identifica su procedencia, APPLICATION_TRAFFIC o DOWNSTREAM_TRAFFIC.

Estado App_Traffic

Ejecutivas de Entrada

- 1) Al entrar a este estado, el módulo invoca el proceso hijo, Bwa_wm_child, para el procesamiento del tráfico de aplicación por enviar hacia la BS.

Estado From_Downstream

Ejecutivas de Entrada

- 1) Verifica que el paquete recibido tenga el formato CMTS_MAC_FRAME_FMT, de no ser así, lo desecha.
- 2) Identifica el tipo de mensaje incluido en el paquete realizando la comparación de su formato con el preestablecido para cada tipo, según se indica a continuación:
 - a) Paquete de información de aplicación proveniente de la BS, CMTS_DATA_PKT_FMT.
 - b) Paquete de control, CMTS_MNG_PKT_FMT, que puede pertenecer a alguno de los siguientes:
 - a. Mapa descriptor del canal ascendente, CMTS_MAP_PK_FMT.
 - b. Paquete de sincronización, CMTS_SYNC_PK_FMT.
 - c. Paquete descriptor del canal ascendente, CMTS_UCD_PK_FMT.
- 3) En cualquiera de los cuatro casos anteriores, recepción de paquete de datos, mapa, de sincronización, o UCD, se activa la interrupción DOWNSTREAM_DATA, MAP_ARRIVAL, SYNC_ARRIVAL o UCD_ARRIVAL, respectivamente.

Estado Down_Data

Ejecutivas de Entrada

- 1) Comprueba que el destino del mensaje sea igual que la dirección de la estación actual.
- 2) En caso de que el destinatario sea la SS actual, envía el paquete hacia el módulo Downstream_sink para la recolección de estadísticas.

Estado MAP_Arrival

Ejecutivas de Entrada

- 1) Procesa el paquete mapa mediante la invocación del proceso hijo Bwa_wm_child.

Estados Syn_Arvl y UCD_Arrival

Ejecutivas de Entrada

- 1) No se encuentra implementado ningún proceso en ninguno de estos estados, salvo la destrucción del paquete.

4.2.5.2 Proceso bwa_wm_child

El proceso hijo Bwa_wm_child, mostrado en la figura 4.10, se encarga de procesar y redirigir los paquetes provenientes de la BS de tipo MAP así como los paquetes generados por las fuentes de información de aplicación. Cada vez que el módulo MAC de una SS recibe alguno de estos paquetes, invoca a este proceso desde los estados App_Traffic o MAP_Arrival. El proceso contempla una cola FIFO (First In, First Out) mediante la cual se atienden los paquetes conforme se reciben. Pueden incorporarse funciones que permitan configurar distintos niveles de prioridad para cumplir con los requerimientos de QoS necesarios para el funcionamiento de aplicaciones en tiempo real como por ejemplo, voz o videoconferencia. En nuestro caso en particular, sólo se tiene una fuente generadora de tráfico por cada nodo, por lo que no fue necesaria la implementación de algoritmos de calidad de servicio en los nodos. También se incluyen funciones específicas para la

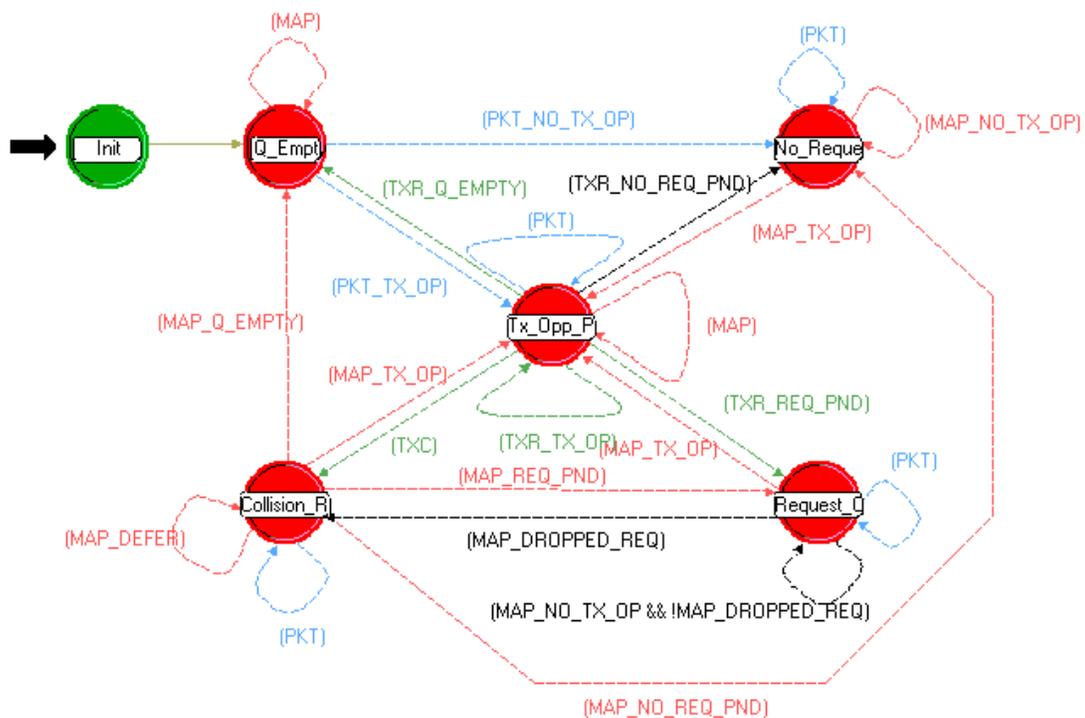


Figura 4.10. Proceso bwa_wm_child.

resolución de colisiones, en específico, el algoritmo exponencial vuelta atrás o Exponential Back-up.

Cada uno de los estados en el proceso hijo corresponde a situaciones específicas de la cola de procesamiento, de la existencia de transmisiones pendientes, de solicitudes sin resolver y de la ocurrencia previa de colisiones. La tabla 4.1 resume las características de

Estado	Cola vacía	Transmisiones pendientes	Solicitudes sin resolver	En deferencia
Init	Sí	No	No	No
Q_Empty	Sí	No	No	No
No_Request_Outstanding	No	No	No	No
Tx_Opp_Pending	No	Sí	No	No
Request_Outstanding	No	No	Sí	No
Collision_Resolution	No	No	--	Sí

Tabla 4.1. Características de cada estado del proceso bwa_wm_child.

cada estado

A continuación se describen las funciones de cada uno de los estados de la tabla 4.1.

Estado init

Ejecutivas de Entrada

- 1) Inicialización de las variables de estado enlistadas a continuación.
 - a) Transmisión programada, tx_scheduled.
 - b) Cola no vacía, queue_occupied.
 - c) Solicitud sin respuesta, req_pending.
 - d) Tiempo de espera de respuesta, ack_wait_time.
 - e) Número de intentos de transmisión, attempts.
 - f) En deferencia tras colisión, defer.
 - g) Ventana actual del algoritmo backoff, backoff.
 - h) En espera de confirmación después de una transmisión de información, data_ack_pending.
 - i) Inicio de la ventana del algoritmo backoff, data_backoff_start.
 - j) Fin de la ventana del algoritmo backoff, data_backoff_end.
 - k) Longitud en bytes del paquete encriptado, encryption_sz.
 - l) Paquetes en contención, contention_hold.
 - m) Paquetes enviados, pk_snt.
 - n) Slots de contención transcurridos, cs_passed.
 - o) Slots de contención pendientes, cs_remaining.
 - p) Longitud del encabezado de la carga útil, payload_header_size.
 - q) Longitud del encabezado de la carga útil con solicitud, payload_header_size_with_req.
 - r) Longitud del encabezado de la carga útil con solicitud de confirmación, payload_header_size_with_ack_reqed.
 - s) Longitud del encabezado de paquetes concatenados, concat_header_size.
 - t) Longitud del encabezado de paquetes concatenados con solicitud de confirmación, concat_header_size_with_ack_reqed.
 - u) Longitud del paquete de solicitud de asignación de ancho de banda, request_packet_size.
 - v) Longitud de la información de solicitud de asignación de ancho de banda, request_data_length.
- 2) Obtención de los parámetros SID, dirección de la SS, tiempo de retraso por propagación, y, en caso de aplicar, de QoS, fragmentación, piggyback y UGS.

Estado Q_Empty

En este estado no existe información por transmitir en la cola de espera, solicitudes pendientes ni transmisiones previamente programadas. Sólo son válidos los eventos de recepción de mapas, MAP_EVENT, o de paquetes, PKT_EVENT.

Ejecutivas de Salida

- 1) Durante la permanencia en el presente estado se identifica el tipo de evento de acuerdo con la interrupción generada. Los eventos válidos son PKT_EVENT o MAP_EVENT.
- 2) En caso de que haya ocurrido un PKT_EVENT, se verifica que esté habilitado y activo el servicio UGS. De ser así, se busca un minislot disponible para iniciar la transmisión, produciéndose la interrupción PKT_TX_OP que conduce al estado Tx_Opp_Pending. De lo contrario, iniciará un periodo de deferencia hasta encontrar alguna oportunidad válida. Lo anterior se realiza mediante el algoritmo exponencial Backoff. La interrupción generada es PKT_NO_TX_OP hacia el estado No_Request_Outstanding.
- 3) Si ocurre un MAP_EVENT, al producirse la interrupción MAP, se almacenará la información contenida en el paquete y, posteriormente, se desechará. El proceso permanecerá en el estado Q_Empty.

Estado No_Request_Outstanding

Aunque en este estado existen paquetes en la cola de espera, no existen solicitudes de asignación de ancho de banda pendientes de ser resueltas ni se han realizado previamente transmisiones. Tampoco se ha activado la deferencia (producida por la ejecución del algoritmo Backoff). Los eventos válidos son la recepción de un paquete de aplicación, PKT_EVENT, que, además de agregarlo a la cola de espera, no genera ninguna acción adicional, y la de un paquete mapa, MAP_EVENT, en el que se busca alguna oportunidad válida para la transmisión de los paquetes acumulados en la cola.

Ejecutivas de Salida

- 1) Si ocurre un MAP_EVENT, al producirse la interrupción MAP, se buscará una oportunidad para transmitir, si no se encuentra ninguna se permanecerá en este estado, en caso contrario, se pasará al estado Tx_Opp_Pending.

Estado Tx_Opp_Pending

Ejecutivas de Salida

- 1) Para salir de este estado, se debe enviar un mensaje que se tiene pendiente, ya sea en una región de contención ó reservada.

Estado Collision_Resolution

Cuando se transmite en una región de contención es preciso esperar una respuesta de la estación base, dicha respuesta se recibe a través de un mapa, cuando no se recibe hay que deferir, eso se hace en este estado mientras que se aplica el algoritmo exponential backoff.

Ejecutivas de Salida

- 1) Si ocurre el evento MAP_EVENT, se debe buscar una confirmación de la transmisión, si ésta ocurre indica que no hubo colisiones y que la estación base recibió el mensaje enviado, de no recibirse, se deberá deferir nuevamente utilizando el algoritmo exponential backoff.

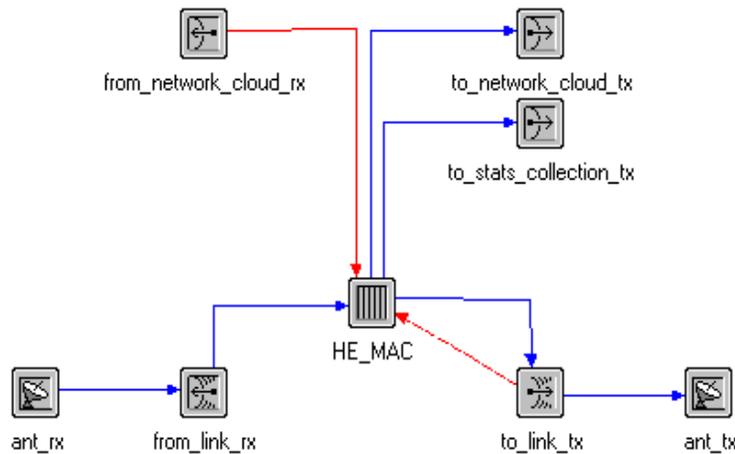


Figura 4.11. Nodo correspondiente a la estación base.

Estado Request_Outstanding

Ejecutivas de Salida

- 1) Para salir de este estado, se debe enviar un mensaje que se tiene pendiente, ya sea en una región de contención ó reservada.

4.3 Diseño de la estación base (BS)

La estación base es el nodo que representa al equipo que proporciona conectividad, dirige y controla a las estaciones suscriptoras. La estación base está conformada por diversos módulos, cada uno de los cuales realiza una tarea específica, como el envío y la recepción al y desde el canal inalámbrico, el envío y recepción desde y hacia la red de cable, el cumplimiento de las funciones propias de la subcapa MAC y la recolección de estadísticas (ver figura 4.11).

4.3.1 Descripción de las funciones de los módulos de la estación base

Los módulos que intervienen en la BS son los que se indican a continuación:

- **Procesamiento de paquetes por enviar y recibidos, y realización de las funciones MAC.** Cuenta con un proceso administrativo con el cual se encarga de controlar, y asignar el ancho de banda, tanto del canal ascendente como descendente, con el fin de proporcionar calidad de servicio y cumplir con las funciones más importantes de la capa MAC de acuerdo con el protocolo IEEE 802.16.
 - HE_MAC
- **Transmisión y recepción.** Se encargan de las funciones de la capa física del modelo OSI y están compuestas por receptores y transmisores, tanto alámbricos como inalámbricos, capaces de identificar señales con características específicas, así como de antenas transmisoras y receptoras
 - Ant_rx, Ant_tx

Nombre del estado	Evento lógico	Factibilidad
Ranging	Down_data_traffic	factible
	Upstream_traffic	no factible
	Ranging_request	factible
	MTN_MAP_time	no factible
	NXT_MAP_time	no factible
	SYNC_time	no factible
	UCD_time	no factible
	Transmitter_free	suprimido
	Sync_done	factible
	Idle	Down_data_traffic
Upstream_traffic		factible
Ranging_request		no factible
MTN_MAP_time		factible
NXT_MAP_time		factible
SYNC_time		factible
UCD_time		factible
Transmitter_free		suprimido
Sync_done		no factible
Idle		Traffic_arrival
	Transmitter_status_indication	factible
Wait	Traffic_arrival	factible
	Transmitter_status_indication	factible

Tabla 4.3. Eventos. factibles para el proceso bwa_be_mgr.

Factibilidad de eventos por estado

La tabla 4.3 muestra todos los eventos lógicos y su factibilidad de acuerdo al estado en que se encuentre el proceso. Los eventos son factibles si existe la posibilidad de que ocurran bajo las circunstancias en que está el estado del proceso. Los eventos suprimidos son eventos que pueden ocurrir, pero que no aportan ninguna información útil.

Estado INIT

Ejecutivas de entrada:

Se encarga de obtener las coordenadas de la estación base, crear el proceso TX_Controller, crear las listas con los elementos de información, leer parámetros de simulación, como número de nodos, y crear las listas de calidad de servicio para cada tipo de tráfico, tamaño de los mapas, tamaño de los slots y encabezados e iniciar las variables de estado que se utilizarán como contadores durante la simulación.

Durante este estado se programa el momento en el que terminará el ranging. La variable End_of_Ranging es un parámetro de simulación disponible en el archivo de variables de ambiente. Se crea e invoca el proceso bwa_be_txctl con el fin de completar la inicialización.

Una vez realizadas las tareas anteriores se pasa al estado de ranging.

Estado Ranging

Posibles acciones de acuerdo al tipo de evento:

- 1) En caso de que llegue un paquete para el canal descendente, se deberá destruir el paquete mientras no se termine de hacer el ranging. Se mantiene el estado de ranging.
- 2) Si ocurre un evento de tipo de solicitud de ranging (ranging_request) se pasa al estado Rng_rcvd.
- 3) Si se termina el tiempo de sincronización (Sync_done), entonces se pasa al estado Rng_complete.

Estado Rng_rcvd

En este estado se creará un mensaje de respuesta de ranging (ranging response) con el sid del nodo que hizo la petición, junto con un cálculo del retardo de propagación desde la SS hasta la BS para la corrección del tiempo. Se enviará este mensaje y se destruye la petición. Se hace una transición hacia el estado de ranging.

Estado Rng_Complete

Ejecutivas de entrada

Al terminarse el tiempo de sincronización (Sync_done), se crea y se envía el primer mapa. Se programa una interrupción para enviar el mapa siguiente. Se pasa al estado Idle.

Estado Idle

Las transiciones de estado dependen del tipo de evento lógico:

- 1) Al recibirse tráfico de la fuente de tráfico descendente se produce el evento Down_data_traffic, se produce la transición SEND_DATA y se pasa al estado Down_data.
- 2) Al recibirse tramas del canal ascendente, ocurre el evento lógico Upstream traffic, se produce la transición RCV_DATA y se pasa al estado Upstream.
- 3) En el momento en el que sea necesario enviar un mapa de mantenimiento, ocurre el evento MTN_MAP_Time y se pasa al estado Mtn_MAP con la transición SEND_MTN_MAP ó SEND_STATION_MTN_MAP.
- 4) Si se genera el evento NXT_MAP_time ocurre la transición SEND_NXT_MAP, y se pasará al estado MAP_time.
- 5) Al generarse el evento SYNC_Time, es tiempo de enviar un mensaje de sincronía, se produce la transición SEND_SYNC y se pasa al estado SYNC_Time.
- 6) Al generarse el evento UCD_Time, es tiempo de enviar un mensaje de tipo UCD, se produce la transición SEND_UCD y se pasa al estado UCD_Time.

Estado Down_data

Si se recibe tráfico del canal descendente (Down_data_traffic), se encapsula en un frame MAC y se le asigna la dirección destino en caso de que no la tenga y se transmite. Se pasa al estado Idle.

Estado Upstream

Se revisa la estructura de la trama recibida, si se recibe una trama con una petición de ancho de banda, se agrega ésta a la lista de peticiones, en caso de que la trama contenga datos, se pasa al módulo de estadísticas. Una vez hecho lo anterior, se regresa al estado Idle.

Estado Mtn_MAP

Se crea y se envía un mapa de mantenimiento. Una vez hecho lo anterior, se programa una interrupción para enviar el siguiente mapa. Se regresa al estado Idle.

Estado MAP_Time

Si el proceso llegó a este estado por la transición SEND_NXT_MAP, entonces, se crea un mapa regular, se envía y se programa la interrupción para enviar el siguiente mapa. Una vez hecho lo anterior se regresa al estado Idle.

Estado Sync_Time

Se genera un mensaje de sincronía, se envía, y programa la ocurrencia de la siguiente interrupción SYNC_Time; al terminar se regresa al estado Idle.

Estado UCD_Time

Se genera y envía un mensaje de UCD, programa la ocurrencia de la siguiente interrupción SYNC_Time, una vez hecho lo anterior, se regresa al estado Idle.

4.3.2.2 Proceso bwa_be_txctl

Como se mencionó anteriormente, el proceso bwa_be_txctl (ver figura 4.14) maneja la transmisión de datos y los mensajes de control generados por el proceso bwa_be_manager. La tabla 4.4 muestra los eventos lógicos que pueden ocurrir en este proceso, así como su descripción y el tipo de evento. La tabla 4.5 muestra describe los estados Idle y TX_Busy, es decir los estados no forzados del proceso bwa_be_txctl. Los estados no forzados son aquellos en los que permanece el proceso después de haber invocado y respondido una interrupción.

Nombre del evento	Descripción del Evento	Tipo de evento
Process_Create	Evento inicial para permitir la inicialización	Invocación por el proceso bwa_be_mgr
Traffic_arrival	Recepción de datos ó mensaje de tráfico	Invocación a través de la interrupción Statwire del bwa_be_mgr
Transmitter_status_indication	Señal indicadora del estado del transmisor (libre u ocupado)	Interrupción Statwire

Tabla 4.4. Eventos lógicos para el proceso bwa_be_tx_tcl.

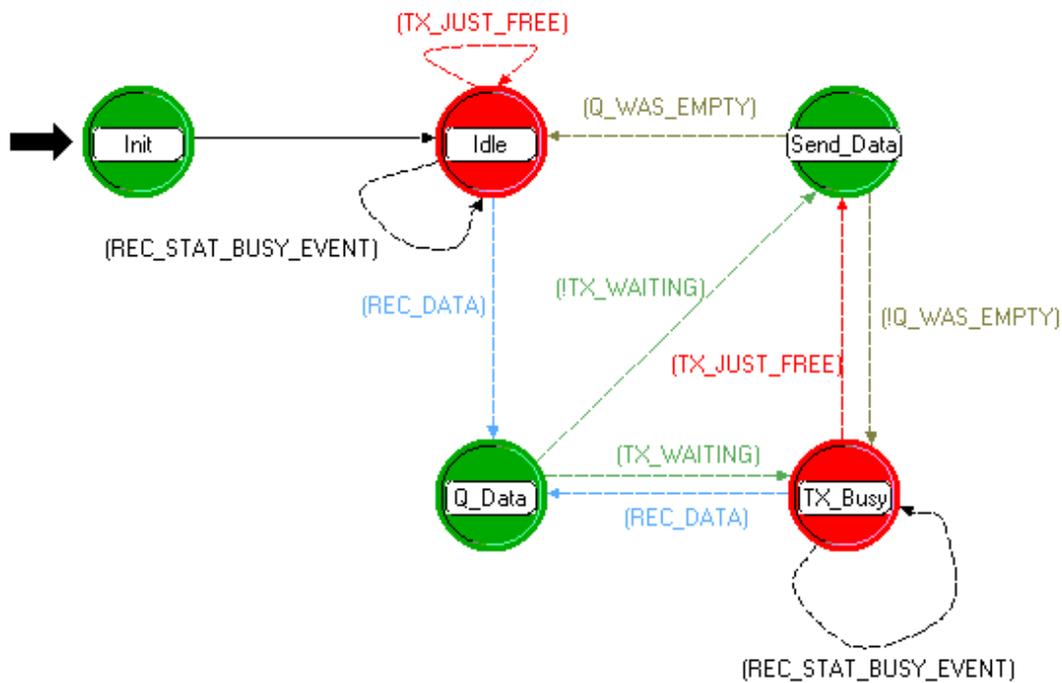


Figura 4.14. Proceso bwa_be_txctl.

Nombre del estado	Descripción del estado
Idle	La cola está vacía y el transmisor está libre. Este estado puede recibir datos ó mensajes de tráfico.
TX_Busy	Periodo en espera para que el transmisor se encuentre disponible

Tabla 4.5. Estados Idle y TX_Busy del proceso bwa_be_txctl.

Nombre del estado	Evento lógico	Factibilidad
Idle	Traffic_arrival	factible
	Transmitter_status_indication	factible
Wait	Traffic_arrival	factible
	Transmitter_status_indication	factible

Tabla 4.6. Eventos factibles para el proceso bwa_be_txctl.

Factibilidad de eventos por estado y respuesta ante cada evento

La tabla 4.6 muestra todos los eventos lógicos y su factibilidad de acuerdo al estado en que se encuentre el proceso. Los eventos son factibles si existe la posibilidad de que ocurran bajo las circunstancias en que está el estado del proceso. Los eventos suprimidos son eventos que pueden ocurrir, pero que no aportan ninguna información útil.

Estado Actual	Evento lógico	Condición	Acción	Estado siguiente
None	Process_Create	Begsim_Interrupt		Init
Init			Hacer el proceso de inicialización	Idle
Idle	Traffic Arrival	Transmisor ocupado	Colocar el paquete en la cola	Tx_Busy
		Transmisor libre	Enviar paquete	Idle
	Transmitter Indication		ninguna	Idle
TX_Busy	Traffic Arrival	En espera de transmisión del paquete previo	Colocar el paquete en la cola	TX_Busy
	Transmitter Indication	Transmisor libre y hay más paquetes en la cola	Enviar paquete	TX_Busy
		Transmisor libre y no hay paquetes en espera	Ninguna	Idle

Tabla 4.7. Tabla de respuesta a eventos en el proceso bwa_be_txctl.

A cada evento corresponde una acción dependiendo del estado actual en que se encuentre el proceso; dicha acción puede ir acompañada de un cambio de estado, tal y como lo muestra la tabla 4.7.

4.3.2.3 Algoritmo empleado para la asignación de slots en la red BWA (Scheduler)

Como se ha mencionado, los tipos de tráfico UGS, rtPS y nrtPs tienen requerimientos específicos; por ejemplo, UGS y rtPS tienen un tiempo límite (también llamado jitter) para llegar a su destino. Los paquetes que lleguen después de ese plazo son inservibles, sin embargo, estos 2 tipos de tráfico pueden soportar la pérdida de paquetes (packet loss). Por el contrario, el tráfico de tipo nrtPS, no acepta la pérdida de paquetes pero en cambio puede aceptar retardos de duración considerable. Con el fin de garantizar la atención a estos tres tipos de servicios con sus requerimientos específicos, se implementó un nuevo algoritmo de programación de servicio que permite priorizar el tráfico y además combina las técnicas de Round Robin, Early Deadline First. Con este algoritmo se espera poder satisfacer los requerimientos de los de los servicios que requieran una transmisión con una tasa de bits tanto constante (CBR) como variable (VBR).

Antes de que la red comience a operar, se debe configurar la estación base para que admita a cada uno de los usuarios. Durante la configuración se le informa a la estación base, cual será el tipo de tráfico que transmitirá cada usuario. Esta información es útil durante la operación, pues cuando la estación base recibe una petición de un nodo que solicita ancho de banda para transmitir, revisa que la petición recibida no esté sobrepasando el ancho de banda máximo que se puede asignar a dicho suscriptor y que estaba acordado previamente. Cada petición que se recibe en la estación base, es clasificada y agregada en una lista de servicios que pueden ser UGS, rtPS, nrtPS ó BE, tal como se describe en el estándar 802.16 y como se muestra en la figura 4.15.

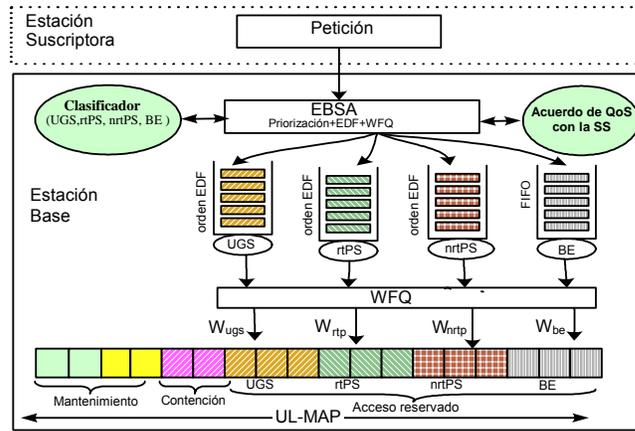


Figura 4.15. Algoritmo de calendarización propuesto para redes IEEE 802.16.

Si la petición recibida corresponde al servicio UGS, entonces el algoritmo asignará periódicamente grants de tamaño fijo a ese suscriptor. Esos grants serán colocados en una cola de tráfico UGS ordenada utilizando una planificación EDF, en la cual, se le da prioridad a los grants con el retardo permitido más bajo. Si la petición recibida se refiere a un servicio rtPS ó nrtPS, entonces se programará el envío periódico de un slot de poleo ó *utxop* por sus siglas en inglés (*unicast transmission opportunity*), en el cual, cada nodo especificará sus requerimientos de ancho de banda. Estas oportunidades de transmisión serán añadidas a las colas de rtPS ó de nrtPS, respectivamente y ordenadas utilizando EDF. Para ambos casos se utilizará el tiempo de poleo máximo permitido como parámetro principal para asignar prioridades en la cola. En cuanto a las peticiones de servicio BE, estas serán colocadas y ordenadas en una cola utilizando un esquema FIFO.

Finalmente, el algoritmo asignará los grants ó *utxop* utilizando un esquema de encolamiento justo ponderado ó WFQ. En este esquema se asigna a cada cola de tráfico un peso diferente. El ancho de banda de salida asignado a cada tipo de tráfico (W_{ugs} , W_{rtp} , W_{nrtp} y W_{be}) estará de acuerdo al peso asignado a cada cola. La fracción del ancho de banda total que se asigna a cada tipo de servicio puede ser configurada por el operador de la red ó también puede ser asignada dinámicamente de acuerdo al tráfico que se tenga en la red en un momento determinado. En este trabajo, con el fin de probar el desempeño del algoritmo, se utilizaron únicamente asignaciones fijas propuestas de antemano.

A continuación se presentan cada una de las etapas involucradas en el algoritmo desarrollado.

1. Cada vez que una estación suscriptorra se conecta ó actualiza el estado de su conexión se establece el vector de parámetros siguiente:

$$SS_n = \{QoS_i, s, \lambda, j, Poll, p, G\} \tag{1}$$

Donde:

QoS_i es el tipo de servicio requerido por la estación suscriptorra SS y que está definido como sigue:

$$QoS_i = \{0 \text{ para } UGS, 1 \text{ para } rtPS, 2 \text{ para } nrtPS\} \tag{2}$$

“**s**” es el estado de la conexión, el cual se va incrementando en 1 por cada UL-MAP enviado.

“**λ**” es el tiempo entre arribos ó intervalo de poleo medido en mapas o UL-Frames, por ejemplo, si el tiempo de interarribo es 10 ms y la duración del mapa es UL-Frame = 2 ms, entonces $\lambda = 10\text{ms} / 2\text{ms}/\text{mapa} = 5$ mapas.

“**j**” es el jitter también expresado en mapas

“**Poll**” es el bit de poleo

“**p**” es la prioridad de transmisión de cada usuario

“**G**” es el tamaño del grant (o número de slots solicitados)

Cuando el SSs_n es igual al tiempo de interarribo ($SSs_n = SS\lambda_n$), entonces será el momento en que la BS otorgue una oportunidad de transmisión unicast ó un grant de poleo (*utxop*) de acuerdo al bit de poleo.

2. Obtener el número máximo de prioridades para el servicio QoS_i

$$Max_p_i = \max\{SSj_1_QoS_i, SSj_2_QoS_i, \dots, SSj_n_QoS_i\} \quad (3)$$

donde $SSj_n_QoS_i$ se refiere al jitter del usuario SSn que requiere un servicio QoS_i .

3. Obtener la prioridad de transmisión de cada estación suscriptora SSn con servicio QoS_i utilizando la expresión (4) del siguiente código.

```

if  $SSs_n\_QoS_i = SS\lambda_n\_QoS_i$  and  $Poll = 1$  (se requiere un grant de poleo)
     $SSp_n\_QoS_i = Max\_p_i$  (se asigna a  $SSp$  la máxima prioridad de transmisión)
Else if  $SSs_n\_QoS_i - SS\lambda_n\_QoS_i > 0$  (Se requiere asignar slots de transmisión)
     $SSp_n\_QoS_i = SSs_n\_QoS_i - SSj_n\_QoS_i - SS\lambda_n\_QoS_i + Max\_p_i$  (4)
    (Se asigna una prioridad de acuerdo al estado actual, al retardo máximo aceptable y al tiempo de interarribo)
Else (no se requiere un grant en el mapa actual, por tanto la prioridad de esa SS deberá ser cero)
     $SSp_n\_QoS_i = 0$ 

```

4. Programar primero las regiones de mantenimiento, tal y como se indica en el estándar, y a continuación programar los minislots que se usarán en la región de contención (los cuales se utilizan para solicitar ancho de banda), utilizando un “asignador de slots de contención” ó CSA por sus siglas en inglés (Contention Slot Allocator) [18]. Si hay algunos SSn que requieran oportunidades de transmisión en el mapa actual, entonces, estos slots deberán programarse primero. Los usuarios de UGS que tengan habilitado el bit de poleo (bit $Poll = 1$), utilizarán el *utxop* para indicarle a la BS que reestablezca su servicio. Los usuarios con servicio *rtPS* y *nrtPS* siempre tendrán habilitado el bit de poleo y utilizarán esta *utxop* para comunicar sus requerimientos de ancho de banda.
5. En el espacio restante del UL-Frame, se deberán programar los slots solicitados por cada estación suscriptora, iniciando con el servicio UGS. El algoritmo asignará grants UGS mientras tenga solicitudes pendientes por atender y

mientras tenga espacio disponible. Al terminar de atender a las estaciones con servicio UGS, atenderá a las estaciones con servicio rtPS de manera similar a como lo hizo con UGS. A continuación atenderá a las peticiones nrtPS de manera similar a las anteriores y finalmente asignará slots de tipo BE. Si no se asignan pesos ponderados a cada tipo de tráfico, el espacio disponible para cada tipo de servicio corresponderá al espacio no utilizado en el mapa. Bajo este esquema, al incrementarse continuamente el número de usuarios más allá del punto de saturación, se observará que a medida que se incremente el número de usuarios con un tipo de servicio con una prioridad más alta (UGS o rtPS), se dejará de atender a los usuarios con una prioridad más baja (rtPS y BE). Con el fin de tener una situación más justa en la programación de slots, se asigna un peso ponderado a cada tipo de servicio (W_{UGS} , W_{rtP} , W_{nrtPS} y W_{be}). En el caso del tráfico UGS, que no utiliza fragmentación, puede ocurrir, que al irse asignando y agotando los slots disponibles de un mapa determinado, al final queden algunos slots que no sean suficientes para satisfacer la demanda de un usuario que requería un grant. En este caso, se contabilizará el número de slots UGS no utilizados en este mapa y se añadirán a los slots UGS del siguiente mapa. De esa manera se respetará el ancho de banda asignado al tráfico UGS. En el caso en el que haya 2 o más SS_n con la misma prioridad, se utilizará la técnica de Round Robin para atender a estos usuarios. Para los servicios rtPS, si el espacio disponible en el UL-Frame es menor que el número de slots solicitados G , entonces se debe utilizar la técnica de fragmentación. La estación base debe calcular cuántos bytes del mensaje original podrán transmitirse con los slots asignados y recalculan el número de slots que se necesitan para transmitir el resto del mensaje, con este valor se actualiza el valor de G . Con ello se evita utilizar el mecanismo de piggyback para solicitar ancho de banda adicional, ya que de otra manera habría que esperar a que la estación suscriptora solicitara nuevamente los slots que le hacen falta para transmitir la parte restante del mensaje, con lo cual tomaría más tiempo satisfacer las demandas de los usuarios. En el caso del tráfico BE, que tiene menores restricciones de tiempo, se permite utilizar tanto fragmentación como piggyback. Puede darse el caso en que el ancho de banda requerido por un tipo de tráfico sea superior al ancho de banda asignado mientras que para otro tipo de servicio ocurra lo contrario, si esto ocurriera, se estarían desperdiciando slots en los mapas. Para evitar esta situación, después de asignar los slots UGS, rtPS, nrtPS y BE de acuerdo al peso ponderado correspondiente, se hace una nueva reasignación de los slots que quedaron libres en el mapa. Durante esta segunda asignación de slots, ya no se utiliza una restricción en el ancho de banda. La reasignación de slots libres se hace con el mismo orden en que se hizo la primera vez, es decir: UGS, rtPS, nrtPS y BE. Si después de la segunda asignación de slots, aún quedan slots libres en el mapa, éstos se pueden utilizar como slots de contención. Por cada oportunidad de transmisión asignada en su totalidad a una estación suscriptora colocada en el UL-MAP, se debe actualizar el estado de la estación suscriptora, esto quiere decir que cuando hay fragmentación, el estado de la SS se actualiza hasta que se asignan los slots requeridos para la última parte del mensaje, dicha actualización se realiza de la siguiente manera

QoS	SS	QoS	s	λ	j	Poll	p	Grant
	1	VoIPa	1	5	5	0	0	13
	2	VoIPa	5	5	5	0	25	13
U	3	VoIPa	10	5	5	0	30	13
G	4	VoIPb	30	30	15	1	30	11
S	5	VoIPb	31	30	30	0	1	11
	6	VoIPb	60	30	30	1	30	11
	7	VoIPc	15	15	15	0	15	9
	8	VoIPc	30	15	15	0	30	9
	9	VoIPc	31	15	15	0	loss	9
	10	MPEGa	20	20	20	1	20	1
r	11	MPEGa	21	20	20	1	1	17
t	12	MPEGa	22	20	20	1	2	9
P	13	MPEGb	20	20	20	1	20	1
S	14	MPEGb	21	20	20	1	1	139
	15	MPEGb	40	20	20	1	20	139
	15	MPEGb	41	20	20	1	loss	139
n	16	FTP	50	50	100	1	100	1
r	17	FTP	51	50	100	1	1	365.
t	18	FTP	100	50	100	1	50	120
P	19	FTP	150	50	100	1	100	50
S	20	FTP	151	50	100	1	101	20.

Tabla 4.8. Prioridades de programación en las redes 802.16.

$$SS_{s_n} - QoS_i = SS_{s_n} - QoS_i - SS\lambda_n - QoS_i \tag{5}$$

La tabla 4.8 muestra un ejemplo con diferentes nodos que manejan diferentes servicios. Por ejemplo, el tipo de servicio “VoIPa”, corresponde al códec G711, el cual genera un frame de voz de 80 bytes cada 10 ms y que en la capa física del protocolo 802.16 solicita una oportunidad para transmitir 13 minislots, los servicios “VoIPb” y “VoIPc” corresponden a los codecs G729 y G723 respectivamente. Cuando un nodo registra un servicio UGS, el estado (s) de la SS es 1 y este valor se incrementa en 1 por cada UL-MAP enviado por la estación base. Cuando $s = \lambda$, el algoritmo primero revisa el bit de poleo (Poll). Para el caso en el que $Poll = 0$, se asume que se requiere un grant de tamaño $G = 13$ slots, el algoritmo determina la prioridad de transmisión para este usuario, como sigue:

$$SSp_2_UGS = SS_{s_2} - QoS_0 - SS_{j_2} - QoS_0 - SS\lambda_n - QoS_i + Max_p_i$$

$$= 5 - 5 - 5 + 30 = 25.$$

La prioridad más alta para el servicio UGS es $Max_p_{UGS} = 30$. Así pues, si SS_2 necesita enviar un dato debe esperar hasta que los usuarios que tengan una prioridad más alta (como SS_3) sean programados. Si el bit de poleo $Poll = 1$, quiere decir que se requiere un bit de poleo ó utxop, de manera que el algoritmo programa una utxop con la prioridad máxima (ver SS_4). Si la estación SS_4 se activa, utilizará el grant asignado para indicar a la

estación base que se debe restablecer el servicio UGS. La estación base, después de haber recibido esta indicación, establece el valor del bit de poleo $Poll = 0$, y el algoritmo obtiene la nueva prioridad de transmisión para este usuario (SS_5 muestra este hecho). Si se incrementa el tráfico en la red, llegará un momento en que la congestión sea tal, que las prioridades de transmisión sean muy altas, si las prioridades llegaran a ser mayores que Max_p_i , el algoritmo envía grants nulos a los usuarios con el tipo de servicio UGS, debido a que sus paquetes llegarán después del plazo límite y ya no serán de utilidad. Un grant nulo indica a la estación suscriptora SS que debe destruir el paquete y esperar al siguiente (ver SS_9). Cabe hacer notar que si $s = \lambda$, hay algunos servicios tendrán una prioridad más alta que otros (ver SS_2 y SS_7), esto se debe a que un servicio con un jitter muy corto requiere una prioridad de transmisión más alta que otro usuario que requiera un servicio con un jitter más largo.

Para el servicio rtPS, el algoritmo EBSA usa los mismos pasos para programar los grants de la misma manera que lo hace para el servicio UGS, con la excepción de que el slot de poleo ó $utxop$ se usará para solicitar una petición de ancho de banda de un tamaño predeterminado. La estación base, después de recibir la petición del usuario con este servicio, actualiza su vector y programa el grant de acuerdo a su prioridad. En la tabla 4.8 se incluyen 2 servicios de tipo “MPEG4a” y “MPEG4b”, los cuales demandan en promedio 76 kbps y 444kbps (en la capa de aplicación) respectivamente. Estas tasas de transmisión fueron calculadas utilizando un tamaño pequeño para los frames I, P, y B para el video con formato “MPEG4a” y un tamaño medio para los frames I, P y B para el video con formato “MPEG4b”, a una velocidad de 25 fps (frames per second) (ver apéndice A). De manera que si se recibe un frame P, (por ejemplo de 417 bytes para MPEG4a o de 2934 para MPEG4b), el tamaño del grant requerido es de 17 minislots para el MPEG4a y de 139 minislots para MPEG4b, como se ve en la tabla para los nodos SS_{11} y SS_{14} . La tabla 4.8, también muestra el caso cuando el número de minislots requeridos es mayor que el número de minislots que quedan en el mapa después de haber programado el tráfico con servicio UGS. Si al momento de programar SS_{11} hubiera únicamente 10 minislots disponibles en el UL-Frame actual, estos minislots se programan para SS_{11} y en el siguiente mapa se programarán los slots restantes (más los slots adicionales debido a la fragmentación y a los encabezados), tal y como se muestra para el usuario SS_{12} .

Gracias a la manera en que este algoritmo da servicio a los usuarios con tráfico rtPS, el retardo de acceso obtenido en las trazas de video es casi uniforme (en redes en las que no hay una gran variación en el tamaño de los frames). El retardo de acceso de los paquetes, es el tiempo que transcurre desde que el paquete es generado hasta el momento en el que es recibido por la estación base. Este retardo también se puede calcular como la suma de los retardos de poleo más el tiempo que tarda la estación base en asignar los slots requeridos para transmitir los paquetes de video (grant delay), esto se puede ver gráficamente en la figura 4.16. En condiciones ideales, la variabilidad de los retardos de acceso es muy baja, esto se debe a que si el retardo de poleo es alto, el grant delay es bajo, y de manera inversa, cuando el retardo de poleo es bajo, normalmente toma más tiempo la asignación de los slots solicitados. Este retardo constante indica que en algunas ocasiones se hace el poleo de la estación, pero como todavía no tiene una prioridad muy alta, puede atender a otros usuarios que tienen una prioridad más alta, retardando la asignación de slots. Por el contrario, si hubo un retardo

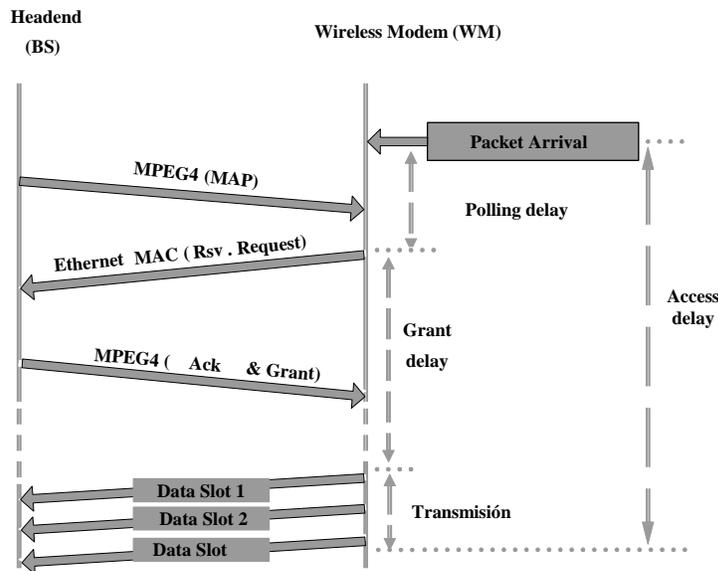


Figura 4.16. Tipos de retardo de los paquetes de tráfico rtPS durante la transmisión.

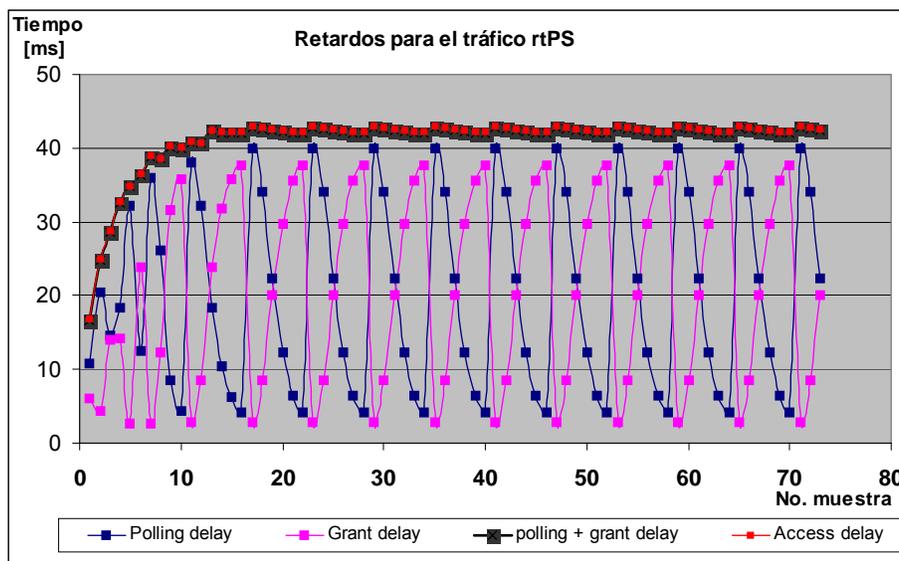


Figura 4.17. Gráfica de los diferentes tipos de retardos de los paquetes de tráfico rtPS.

muy grande en el poleo, la prioridad de esa estación será alta, por lo que deberá ser atendido lo más pronto posible. Este hecho se observa en la figura 4.17.

El servicio de tipo nrtPS se programa de la misma manera que el de tipo rtPS, con la excepción de que se permite la llegada de paquetes con retrasos muy largos. Para este tipo de servicio se utilizó un tipo de servicio FTP, en donde los usuarios envían un paquete de datos con un tamaño de 5.12 kbytes, el cual requiere 4 paquetes de 1024 bytes TCP PDU's y requiere un grant de 365 minislots (ver SS₁₇). Debido a que este es un grant muy grande, la estación base, normalmente envía grants cortos, de acuerdo a la capacidad máxima y mínima permitida para este usuario con base en el servicio contratado por el usuario. Cada vez que el usuario recibe un grant corto, envía la fracción del mensaje que se ajusta a ese grant y además hace una petición de piggyback en la que se indica el

tamaño del mensaje restante (ver SS₁₈ y SS₁₉). Si después de algunos grants el mensaje no se ha transmitido completamente y el retardo de transmisión acumulado para este mensaje es más grande que el máximo retardo registrado para este servicio, la estación base comienza a enviar grants consecutivamente hasta que el mensaje haya sido transmitido en su totalidad, ya que el servicio puede soportar retardos bastante largos, pero no soporta la pérdida de paquetes, tal y como se muestra para el usuario SS₂₀, donde el máximo retardo registrado es de 100 UL-frames y el retardo actual es de 101 UL-frames.

A continuación se hace un breve análisis del comportamiento general de la red, ya que el análisis más detallado de cada tipo de escenario se hace en el capítulo siguiente. En general se encontró que en una red con múltiples fuentes de tráfico, tanto de VoIP (G-711 y G-723) e Internet, los resultados alcanzan un estado estable y a partir de ese momento varían muy poco en función del tiempo. Para una red con 57 usuarios (50% de la capacidad del canal ascendente) el retardo de acceso promedio para los paquetes de tráfico G-723 (como se describe en la sección 5.1) fue de 1.178 ms al cabo de un minuto. Cuatro minutos después, este retardo fue de 1.177 ms, lo cual representa una variación de 0.11%. Si se hace un análisis similar cuando la red tiene 72 usuarios activos (64% de ocupación del canal), los retardos son de 1.052 ms y 1.050 ms para 60 y 300 segundos respectivamente, esto equivale a una variación del 0.14%. Cuando la red se empieza a saturar, con 84 nodos activos (75% de su capacidad), los retardos de acceso promedio para 60 y 300 segundos son de 1.127 ms y 1.125 ms, lo que equivale a una variación del 0.18% (Ver figura 4.18).

Un análisis similar para el tráfico G-711 muestra que cuando la red tiene 57 usuarios activos, la variación en los retardos va de 0.6599 ms a 0.658 ms., lo que equivale al 0.27%. Si la red tiene 72 nodos generando tráfico los retardos de acceso promedio serán de 5.43 ms al cabo de un minuto y de 5.45 ms al cabo de 5 minutos, indicando una variación del 0.51%. Finalmente cuando la red esté operando al 75% de su capacidad (84 nodos activos), se tendrán retardos de 6.251 ms al cabo de un minuto y de 6.276 ms después de 5 minutos, lo cual equivale a una variación del 0.4% (Ver figura 4.18).

Para el tráfico BE, que es menos sensible a los retardos de acceso, el análisis presentó una variación del 0.24%, 0.51% y 0.41% para 57, 72 y 84 usuarios, respectivamente, lo

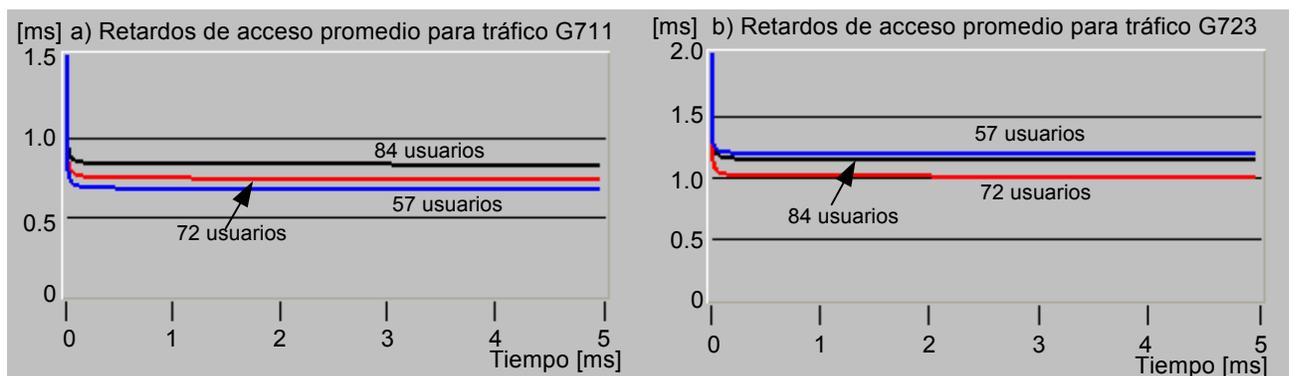


Figura 4.18 Retardos de acceso promedio para tráfico de voz.

cual indica que con estos tres casos, una vez que la red alcanza el estado estable, los resultados ya no varían y por consiguiente se utilizará un tiempo de simulación de 60 segundos en los escenarios presentados en el siguiente capítulo.

Otro de los parámetros involucrados en las simulaciones, es aquel que está relacionado con la generación de números aleatorios. Este parámetro se llama “seed”, al cambiarlo de una simulación a otra, cambiarán los valores para la generación de tráfico, sin embargo, estos valores seguirán cumpliendo con la misma función de densidad de probabilidad.

Para evaluar el impacto del *seed* se hicieron varias simulaciones en las que se asignaron diferentes valores entre 100 y 700. Como se mencionó anteriormente, el tráfico de tipo BE es el tráfico que sufre los mayores retardos de acceso promedio debido a la prioridad que tiene dentro del algoritmo *scheduler*. Las gráficas de la figura 4.19 muestran los retardos de acceso promedio para el tráfico BE. En estas simulaciones, el retardo de acceso promedio fue de 5.1 ms, y a partir de ese valor tendremos una variación de $\pm 3\%$.

Las simulaciones analizadas en el siguiente capítulo tuvieron una duración de 60 segundos y en ellas se empleó un solo *seed*, ya que de acuerdo a los resultados analizados anteriormente, podemos concluir que después de un minuto se pueden tomar los resultados de acceso promedio como representativos del estado estable y también podemos concluir que podemos elegir cualquier valor de *seed* y los resultados que se obtengan serán representativos del comportamiento de la red con una variación mínima de $\pm 3\%$.

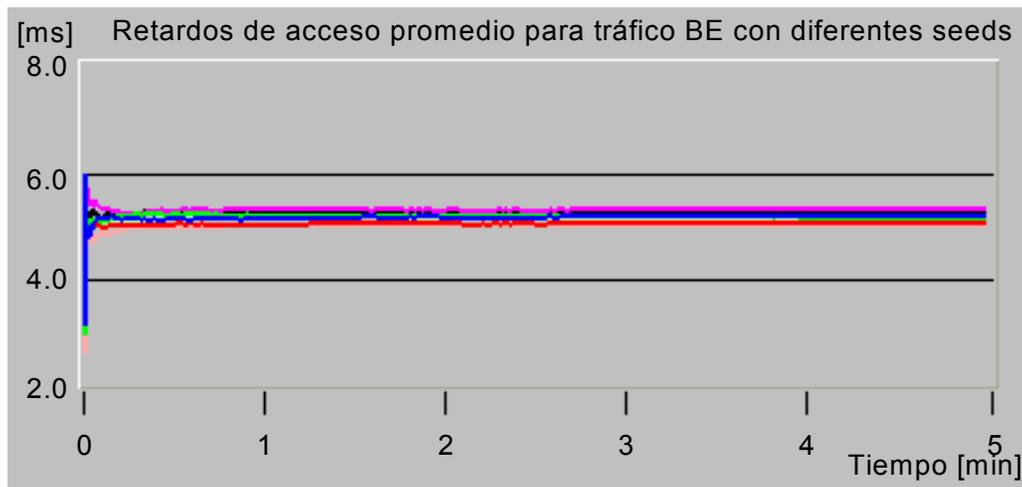


Figura 4.19. Retardos de acceso promedio para tráfico BE cuando se emplean diferentes *seeds* en las simulaciones.

ANÁLISIS DEL COMPORTAMIENTO DINÁMICO

En este capítulo se presenta un estudio sobre el comportamiento dinámico del protocolo de comunicaciones IEEE 802.16, cuando diferentes técnicas de calidad de servicio son utilizadas, de acuerdo a las aplicaciones que se le esté dando a la red.

5.1 Escenarios de tráfico

En todas las simulaciones se empleó un canal ascendente con una capacidad de 9.6 Mbps y un canal descendente con una capacidad de 22 Mbps. En las simulaciones únicamente se emplearon 4 fuentes de tráfico, “G-711-UGS”, “G-723-UGS”, “mpeg4-rtPS” y tráfico “Internet-BE” para el análisis de desempeño. En la tabla 5.1 se muestran los parámetros utilizados en el modelo de simulación.

1) VoIP – G723. Este tipo de tráfico emula al códec de voz “G-723.1”, el cual, de acuerdo a la ITU, a la IETF y al foro VoIP, es el códec preferido para aplicaciones de telefonía a través de Internet [16]. Este códec genera datos a una tasa de 5.3 kbps ó 6.3 kbps dependiendo del modo de operación, donde cada 30 ms se generan y codifican paquetes de 20 bytes para una tasa de 5.3 kbps ó 24 bytes para 6.3 kbps. Después de añadir los encabezados como se muestra en la tabla 5.2, se obtiene una tasa de 38.4 kbps en la capa física.

Parámetro	Valor
Velocidad del canal ascendente	9.6 Mbps
Velocidad del canal descendente (16-QAM, 7 MHz Ch.)	22.4 Mbps
Slots de contención por mapa	7 slots
Tamaño de los minislots en cada mapa	16 bytes
Duración del mapa	2 ms =150 minislots
Periodo de simulación por cada corrida	60s
Distancia desde la SS más cercana/alejada a la BS elegida aleatoriamente	0.1 - 5 km
Reed Solomon (grants cortos / grants largos)	6 bytes / 10 bytes
Tamaño límite entre los grants cortos y largos	245 bytes
Máximo número de usuarios en la red BWA	200
Tipo de tráfico analizado	VoIP, Video & Internet

Tabla 5.1. Parámetros de las simulaciones.

2) VoIP-G.711. El códec G.711 fue utilizado debido a la alta tasa de transmisión que requiere con el fin de forzar a la red BWA y también debido a que este códec será utilizado ampliamente para tener una gran calidad en las comunicaciones por voz ya que es la recomendación del estándar de la ITU-T H.323, G.711 [16]. Este códec de audio utiliza modulación por codificación de pulsos ó PCM para generar datos a una tasa de 64 kbps. Los datos creados se encapsulan en una trama de VoIP de 80 bytes de longitud cada 10 ms que requieren 166.4 kbps en la capa física (ver tabla 5.2).

3) Tráfico de video mpeg4. Para este tipo de tráfico se utilizaron las trazas de video de 10 películas diferentes, las cuales se digitalizaron en un formato QCIF (Quarter Common Intermediate Format) de mpeg4 con una resolución de 176 x 144 pels a razón de 25 imágenes por segundo. Las muestras fueron tomadas de [33] y fueron encapsuladas como se muestra en la figura 5.1. Las películas seleccionadas son las mostradas en la tabla 5.3. En el apéndice A se muestran las distribuciones de tráfico para estas películas.

	G.711 - 64 kbps	G.723.- 5.3 kbps
Frame size [ms]	10	30
Voice frame [bytes]	80	20
RTP [bytes]	12	12
UDP [bytes]	8	8
IP [bytes]	20	20
LLC [bytes]	3	3
SNAP [Bytes]	5	5
Ethernet MAC [bytes]	18	18
IEEE 802.16 MAC	6	6
PHY: (Prea+GB+FEC)	10+FEC	10+FEC
Total PacketSize	202 bytes or G=13 slots	86 bytes or G=9 slots
Net rate at MAC / PHY	116.8 / 166.4 kbps	22.9 / 38.4 kbps

Prea = Prea mble, GB = Guarband, and FEC = 6* No_CodeWords

Tabla 5.2. VoIP CODECS: G.711 y G.723.1.

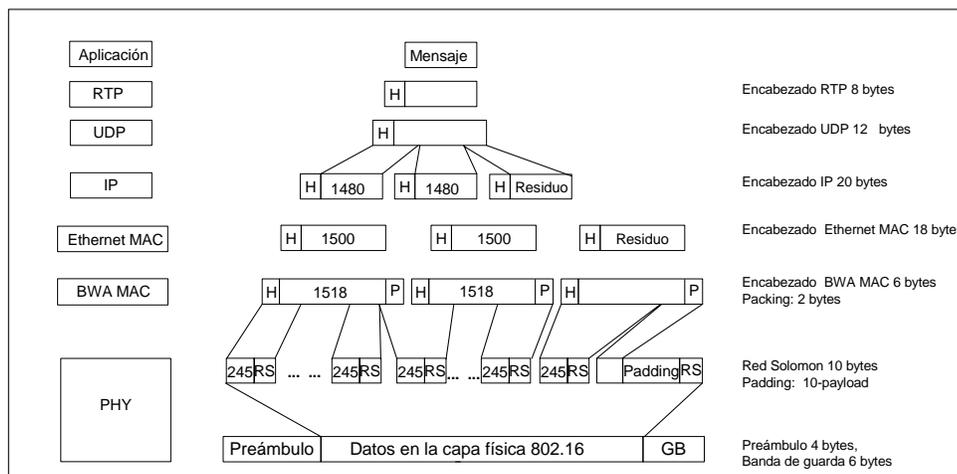


Figura 5.1. Encapsulamiento de tramas de video.

	Nombre de la película	Media de las muestras [bytes]	Tasa de datos de usuario (bps)	Slots requeridos para la media	Tasa de transmisión requerida [bps]
1	Aladdin	297.61	59523	25	80000
2	Die Hard_III	587.06	117411	44	140800
3	Futurama	1106.30	221260	70	224000
4	Jurassic Park I	684.74	136947	50	160000
5	Mr. Bean	437.91	87582	34	108800
6	Robin Hood	460.18	92035	36	115200
7	Silence Of The Lambs	1871.20	374240	130	416000
8	Start Trek First Contact	209.22	41843	19	60800
9	Star Wars IV	530.59	106117.	40	128000
10	Simpsons	1464.60	292920.	103	329600
	promedio	764.90	152980.0	55.1	176320

Tabla 5.3. Datos de las muestras de las películas empleadas en las simulaciones.

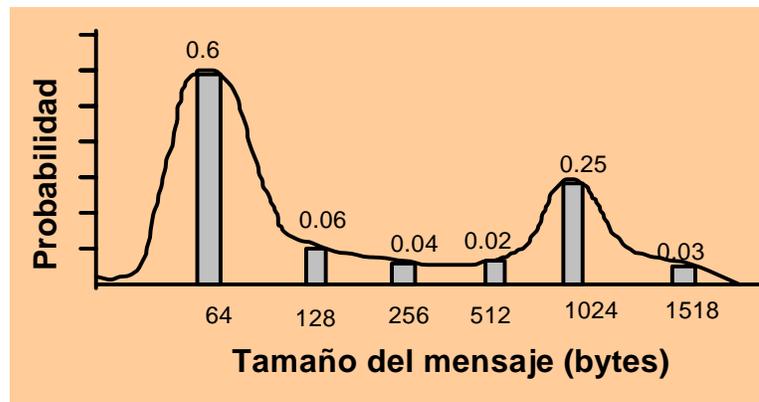


Figura 5.2. Distribución de tráfico de Internet.

4) Tráfico de Internet IP. La distribución de tráfico de Internet utilizada es la introducida por el grupo de trabajo IEEE 802.14 [16]. La distribución del tamaño de los mensajes es la siguiente: 60 % es para mensajes de 64 bytes, 6% para paquetes de 128 bytes, 4% para paquetes de 256 bytes, 2% para paquetes de 512 bytes, 25 % para paquetes de 1024 bytes y 3% para paquetes de 1518 bytes (ver figura 5.2). El tiempo entre llegadas de los paquetes varía de acuerdo a la tasa de transmisión a utilizar por el usuario. Por ejemplo, para una tasa de transmisión de 32 kbps y un tamaño promedio de mensaje de 368.1, el cual se obtiene de la expresión (5.1), se obtiene un tiempo entre llegadas de 92 ms.

$$\bar{L} = \sum_0^{\infty} L * Prob_L \tag{5.1}$$

Donde:

L = tamaño del mensaje

\bar{L} = tamaño promedio de los mensajes

$$\bar{L} = 0.64 * 0.6 + 128 * 0.06 + 256 * 0.04 + 512 * 0.02 + 1024 * 0.25 + 1518 * 0.03 = 368.1$$

$$tiempo_entre_llegadas_a = \frac{1}{\frac{paq}{seg}} = \frac{1}{\frac{32000[bits/seg]}{368.1 * 8[bits]}} = 92.025[ms]$$

5.2 Análisis de desempeño

Para obtener el desempeño de una red BWA basada en diferentes técnicas de calidad de servicio se utilizaron diferentes escenarios de tráfico. En realidad, un operador de redes BWA, deberá seleccionar las técnicas de QoS que se demanden en una zona en particular, por ejemplo en una zona de un nivel socioeconómico medio-bajo, los servicios que más se demandarán son los del tipo BE, es decir, navegación por Internet, FTP y correo electrónico, entre otros, los cuales no requieren transmisión en tiempo real. Por consiguiente, un canal uplink puede ser configurado para soportar solamente tráfico BE. En otras zonas, de un nivel socioeconómico medio-alto, es necesario que un canal sea utilizado no solamente para la transmisión de datos, sino también para video, voz y datos a gran velocidad; por consiguiente un canal se puede configurar para soportar video (con rtPS), voz (con UGS), y datos a gran velocidad (con nrtPS). En esta sección se presentan diferentes escenarios que mostrarán el comportamiento dinámico de un canal cuando es configurado para soportar varios niveles de calidad de servicio y asignaciones de ancho de banda. Las primeras simulaciones muestran el comportamiento de la red cuando el ancho de banda se asigna de acuerdo al tipo de tráfico que esté en la red. En la última simulación se asigna un ancho de banda fijo a cada tipo de tráfico y se comparan los resultados con los obtenidos previamente.

5.2.1 Escenario de red para tráfico con servicio UGS y BE.

Para el primer escenario se configuró una red de tal manera que una tercera parte de los nodos de la red transmitiera tráfico VoIP-G723 (es decir, los nodos 1, 4, 7, 10, etc.), otra tercera parte generara tráfico VoIP-G711 (los nodos 2, 5, 8, etc.) y la parte restante (los nodos 3, 6, 9, etc.) tráfico de Internet (BE) a 32 kbps (figura 5.3). El análisis de resultados para esta simulación se hará en términos de la velocidad real de transporte por tiempo de trama (throughput), el retardo de acceso (access delay) y la probabilidad acumulada.

La figura 5.4 presenta la utilización del canal ascendente y el throughput alcanzado con cada flujo de servicio diferente. En la gráfica de la figura podemos apreciar que la máxima utilización del sistema es de 9.2 Mbps, lo que corresponde a un poco más del 95% de la capacidad del canal ascendente (de 9.6 Mbps), el otro 5% corresponde a los slots de

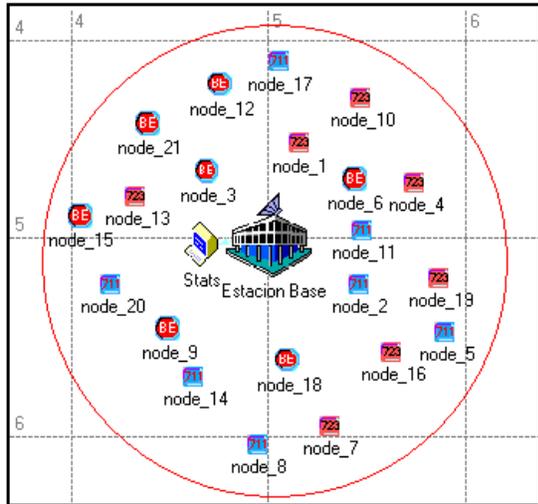


Figura 5.3. Configuración de la red con tráfico VoIP e Internet.

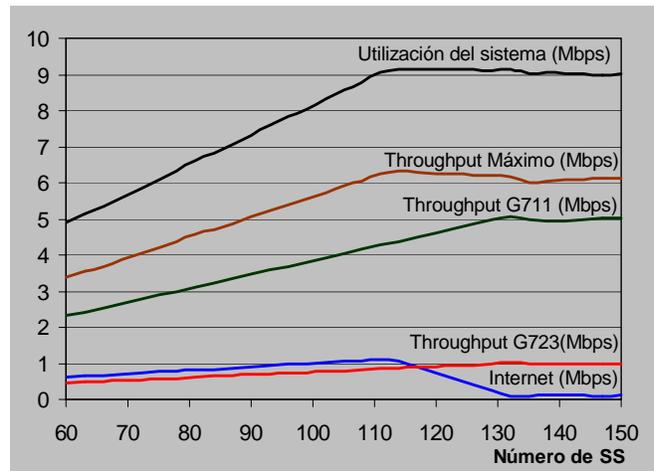


Figura 5.4. Throughput para los diferentes tipos de tráfico.

contención (recordemos que 7 de cada 150 slots son utilizados para contención en cada mapa), de esta utilización, el 66% (6.3 Mbps) fue utilizado para la transmisión de datos de usuario y corresponde al throughput máximo. El resto (29%) fue consumido por el protocolo de enlace de radio (RLP) del estándar 802.16, y la codificación en la capa física (PHY). El número máximo de nodos que soporta un canal ascendente de 9.6 Mbps se puede calcular utilizando la expresión siguiente:

$$MaxSS = 3 * \frac{Utilización_del_sistema}{VoIP_{G711} + VoIP_{G723} + Internet} = 3 * \frac{9600 * 143 / 150}{166.4 + 38.4 + 38.4} \approx 112 \quad (5.2)$$

En las gráficas de la figura 5.5 podemos apreciar que el retardo de acceso para el tráfico de Internet es mucho mayor para 112 usuarios que para 111, esto se debe a que con 111

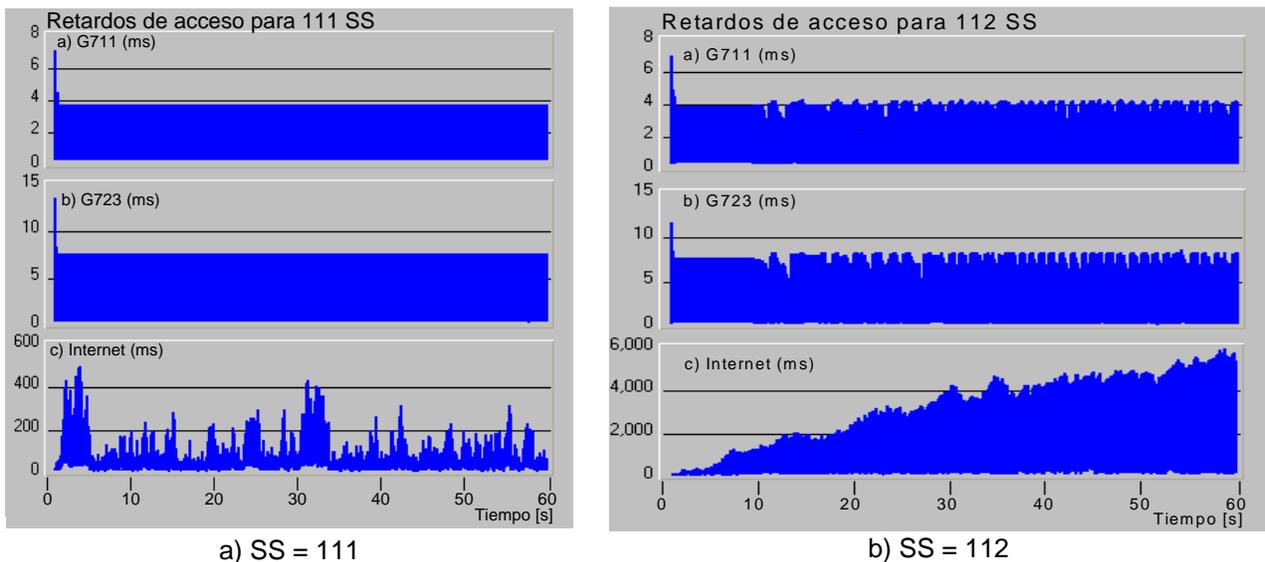


Figura 5.5. Retardos de acceso vs tiempo para 111 y 112 SS.

usuarios activos todavía no se alcanza a saturar el canal, lo que sí ocurre con 112 usuarios, sin embargo, a pesar de que los retardos son mayores, el throughput para Internet (y el throughput total de la red) es mayor para 112 usuarios, por lo que podemos decir que 112 usuarios sería el valor óptimo que una red BWA puede alcanzar (37 usuarios generando tráfico VoIP G-711, 38 usuarios generando tráfico VoIP G-723 y 37 usuarios de Internet). A partir de ese punto, el tráfico de Internet comienza a decrecer (ver Figura 5.4). Este es el resultado esperado, ya que el algoritmo programa al tráfico UGS con una prioridad más alta que al tráfico de Internet. Esto da también como resultado un retardo de acceso menor para los servicios UGS, como se aprecia en las gráficas de la figura 5.5, de la figura 5.6 y de la figura 5.7. En la figura 5.6.b podemos observar que el 100 % de las tramas de VoIP fueron transmitidas por debajo de los 4 ms para los flujos de tráfico de tipo G-711 y de los 8 ms para los flujos de tipo G-723. Sin embargo, sólo el 0.4 % de las tramas de Internet tuvieron retardos de acceso por debajo de los 10 ms. También podemos observar, que aún cuando el tráfico de VoIP sigue aumentando, el throughput total del sistema disminuye ligeramente; esto tiene su explicación en el hecho de que aún se transmiten paquetes de Internet y las oportunidades de transmisión para estos paquetes son tan pequeñas, que los encabezados son mayores que la información transmitida, lo que da como resultado que el throughput total disminuya ligeramente.

Aún durante los periodos de congestión, al algoritmo implementado es capaz de proporcionar retardos de acceso suficientemente bajos para los flujos UGS. De la ecuación (5.2), el máximo número de flujos de voz sobre IP (con igual número de nodos G-711 y G-723 y sin que haya tráfico de Internet) que el canal ascendente puede soportar es de 44 flujos de G-711 y 44 de G-723. La gráfica de la figura 5.6.c muestra los resultados de esa simulación, ahí observamos que para tráfico de VoIP, el 100% de las tramas transmitidas en la red BWA tuvieron un retardo de acceso de menos de 4 ms para el tráfico de tipo G-711 y de menos de 12 ms para el de tipo G-723. En la transmisión de tráfico VoIP, no hubo pérdida de paquetes; sin embargo, las gráficas de la figura 5.6.d y de la figura 5.7 muestran el desempeño de la red cuando está sobrecargada con 45 nodos generando tráfico de tipo G-711, 45 de tipo G-723 y 45 de Internet. Los resultados

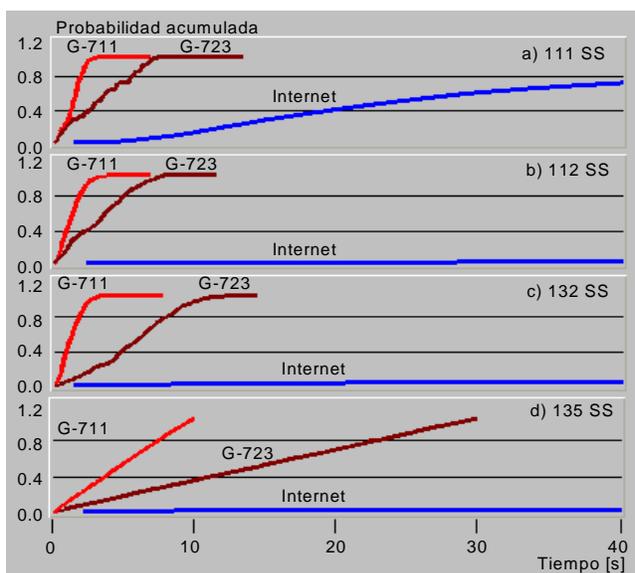


Figura 5.6. Probabilidad acumulada vs retardos de acceso.

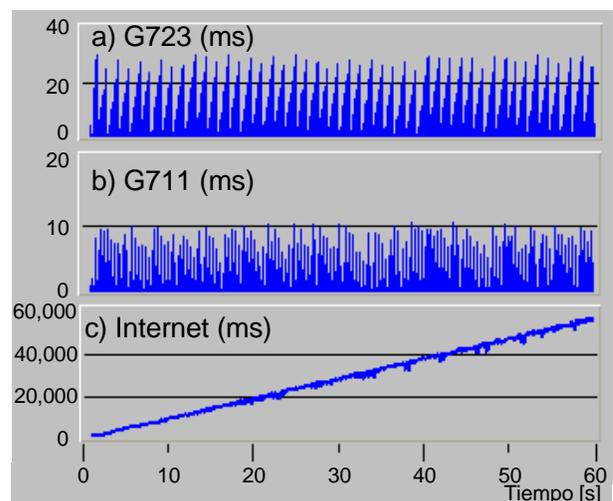


Figura 5.7. Retardos de acceso vs tiempo (135 SS).

muestran que para el tráfico G-711 hay una pérdida de paquetes del 5.4%, mientras que para el tráfico G-723 la pérdida de paquetes es del 2.82%. Esta pérdida de paquetes se debe a que los retardos son mayores que el jitter permitido y por tanto son descartados, debido a que aún si los paquetes pudieran ser enviados serían inútiles debido a los retardos. Sin embargo, esta tasa de pérdida de paquetes aún es aceptable para la comunicación utilizando VoIP.

5.2.2 Escenario de red para tráfico rtPS

El siguiente escenario fue uno que contenía únicamente nodos que generaban tráfico de video ó rtPS (ver figura 5.8). Si suponemos, de acuerdo a la tabla 5.3, que en promedio se necesitan 176.320 kbps por cada fuente de tráfico rtPS, entonces, de acuerdo a la ecuación (5.2), tenemos que $9152/176.32 = 51$ nodos. Sin embargo, en la gráfica de la figura 5.9, observamos que el throughput para 51 nodos es de tan solo 8.2 Mbps, en lugar de los 9.15 Mbps que se esperaban. Esto se debe principalmente a la pérdida de paquetes que tiene lugar debido a los retardos, en la figura 5.10, observamos que para 51 nodos, la pérdida de paquetes es cercana al 12%. En [39] se presenta un estudio acerca de la percepción de la calidad de video en función de la pérdida de paquetes y se propone una técnica para el mejoramiento del video recibido. En nuestro caso consideraremos como aceptable una pérdida de paquetes cercana al 5%. Además de tomar en cuenta la pérdida de paquetes, se deben tomar en cuenta los retardos de acceso que sufre cada paquete durante la simulación. Debido a la gran variabilidad en el tamaño de las muestras de video, puede ocurrir que en instantes cercanos de tiempo, varias estaciones intenten transmitir tramas de video demasiado grandes, con lo cual la estación base tardará demasiado tiempo en atenderlas y en consecuencia el retardo de acceso se incrementará. En [13], se asume que las tramas de video no deben tener un retardo mayor a 150 ms, por tanto, si tomamos 100 ms como el valor límite del retardo de acceso, debemos revisar que los retardos sufridos por las trazas durante la simulación sean menores que ese valor. Si algún paquete sufre un retardo mayor, deberá ser descartado. En la figura 5.11 se

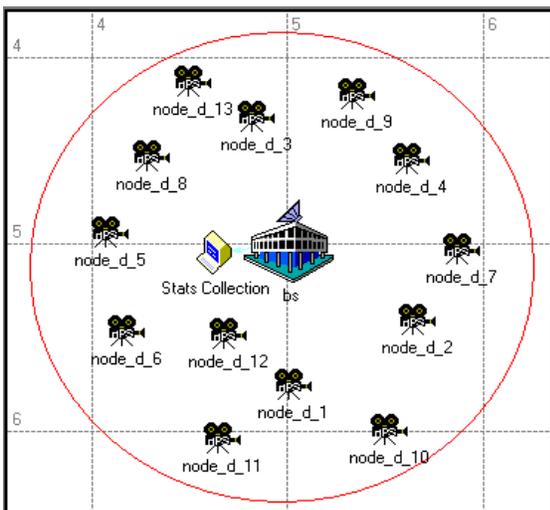


Figura 5.8. Red con generadores de tráfico de video,

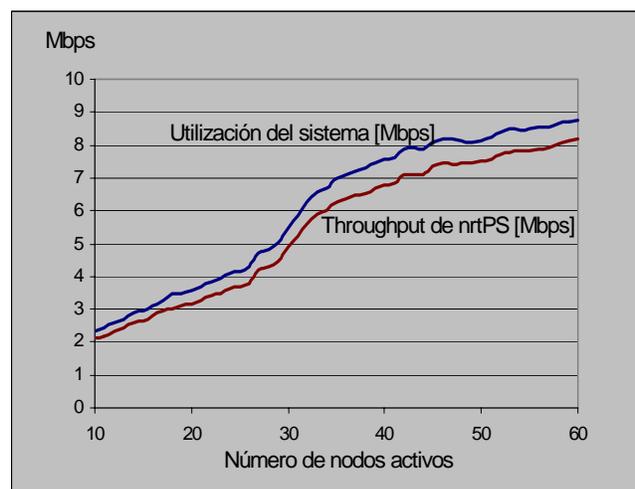


Figura 5.9. Utilización y throughput del sistema para tráfico para rtPS

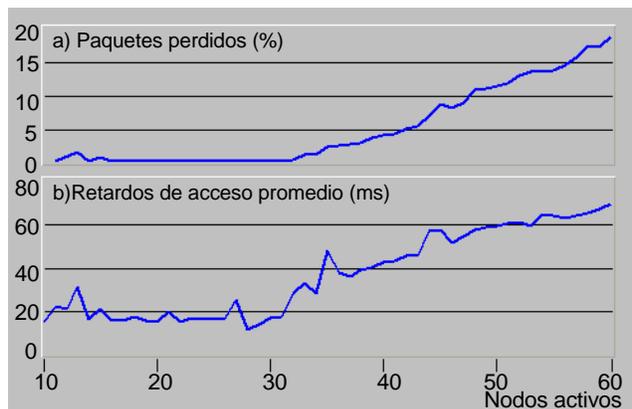


Figura 5.10. Tasa pérdida de paquetes y retardos de acceso promedio para rtPS.

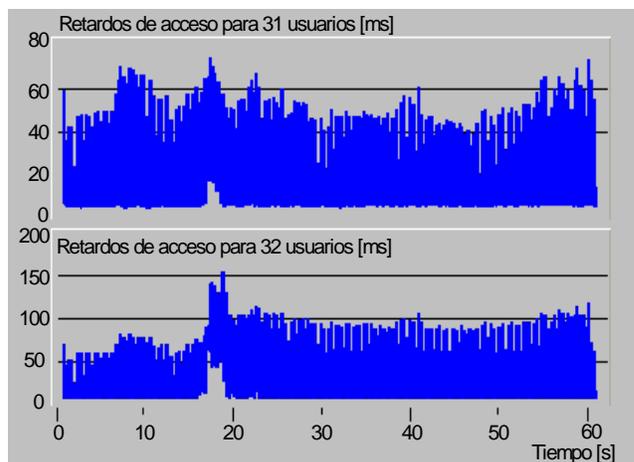


Figura 5.11. Access delay para 31 y 32 usuarios.

muestran los retardos de acceso para 31 y para 32 usuarios. De la comparación de ambas gráficas, observamos que el número de nodos más adecuado es 31, es decir, el 60% de la cantidad de nodos que se había calculado inicialmente. Asimismo, en las gráficas de la figura 5.10 observamos que para 31 nodos activos, no hay paquetes perdidos y que el retardo de acceso promedio es de 16 ms. En la misma figura observamos que para 32 nodos activos se comienzan a perder algunos paquetes. La simulación reportó que esta pérdida de paquetes para 32 nodos activos corresponde al 0.26%, el cual no es un valor significativo, por lo que se puede afirmar que el servicio rtPS todavía puede ser soportado.

Una vez que hemos observado el desempeño de la red para el servicio rtPS, observaremos cómo trabajaría una red que transporte tráfico de voz y de video simultáneamente.

5.2.3 Escenario de red para tráfico rtPS y UGS

Con este escenario se ejecutó una serie de 26 simulaciones, iniciando con 24 nodos activos y finalizando con 96. En cada simulación se fueron agregando 3 usuarios más, uno por cada tipo de tráfico (G-711, G-723 y video). De acuerdo a la ecuación (5.2) tenemos que el número máximo de usuarios que puede soportar la red con esta configuración de tráfico es de 72 ($3 \cdot 9152 / (38.4 + 166 + 176.320) = 72$). De acuerdo a este resultado, esperaríamos que al tener 72 nodos activos en la red se saturara el canal, sin embargo, de acuerdo a la gráfica de la figura 5.13, observamos que a partir de 57 usuarios la pendiente del throughput de tráfico rtPS comienza a disminuir. Como ya se había comentado anteriormente, en la transmisión de video no sólo hay que considerar el throughput, sino también los retardos de acceso y la pérdida de paquetes. De acuerdo a las gráficas de la figura 5.14 y de la figura 5.15, para 60 nodos activos, la pérdida de paquetes es del 4.43% y el retardo promedio es de 43 ms respectivamente. En la gráfica de la figura 5.15.a vemos que cuando en la red hay 60 nodos activos, algunos paquetes, principalmente aquellos generados entre 14 y 27 segundos, tienen retardos de acceso mayores a 100 ms. Esto se debe principalmente a que el 20% de los nodos de video seleccionaron la película cuyas muestras de video son más grandes que las muestras del resto de las películas. Además, ésta película genera sus paquetes más grandes en el intervalo comprendido entre los 14 y los 27 segundos. Esto

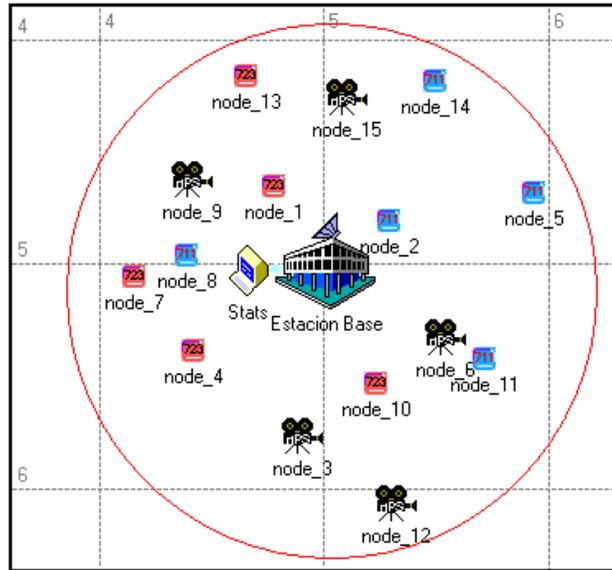


Figura 5.12. Red con generadores de voz y video.

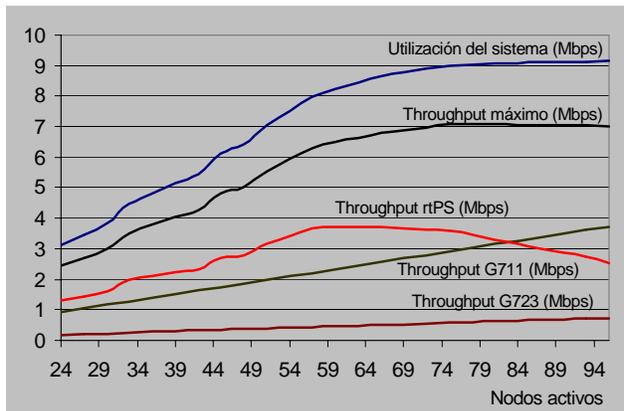


Figura 5.13. Throughput para el escenario con generadores de tráfico UGS y rtPS.



Figura 5.14. Pérdida de paquetes de video para una red con tráfico de voz y video.

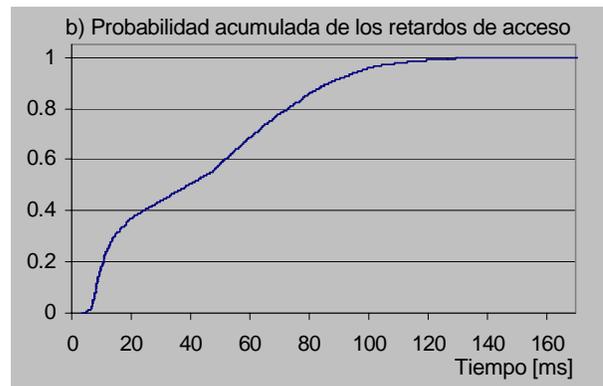
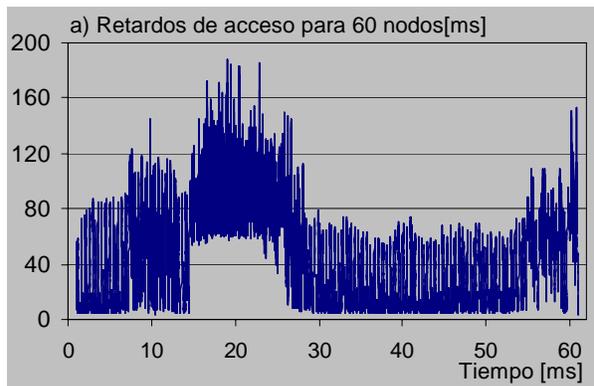


Figura 5.15. Retardos de acceso para 60 nodos.

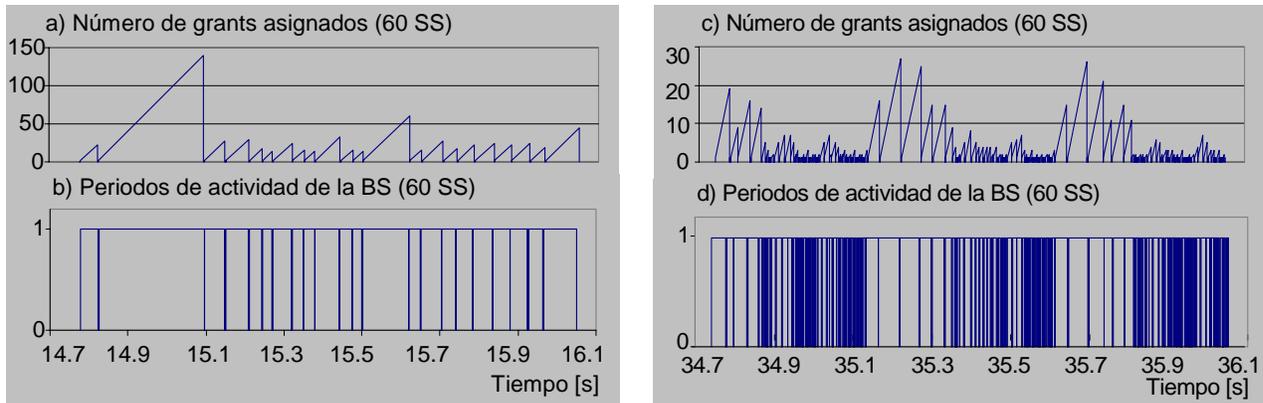


Figura 5.16. Periodos de ocupación del canal ascendente en diferentes intervalos de tiempo.

quiere decir que en ese lapso de tiempo, se requiere un mayor número de oportunidades de transmisión o grants, la estación base no puede asignar los slots demandados lo suficientemente rápido y en consecuencia se obtienen retardos mayores. En las gráficas de la figura 5.16, se muestran dos periodos de tiempo durante los cuales se pueden apreciar diferentes instantes de actividad e inactividad de la estación base, así como el número de slots que se asignan durante esos periodos de actividad. El primero de esos intervalos de tiempo está comprendido entre 14.77 y 16.05 s mientras que el segundo lo está entre 34.77 y 36.05 s. Durante los periodos de actividad (valor de 1 en la figura 5.16 b y d), se están asignando slots a los nodos en los mapas del canal ascendente, mientras que durante los periodos de inactividad (valor de 0 en la misma figura) se dejan slots libres en el mapa; esto indica que no hay peticiones pendientes por atender de tráfico rtPS y que en el canal ascendente, habrá breves instantes de tiempo durante los cuales ninguna estación transmita ningún paquete de video. Si comparamos las gráficas a y b con las gráficas c y d de la figura 5.16, observaremos que en el primer intervalo de tiempo hay menos transiciones entre los estados de actividad e inactividad y que se asigna un número mayor de slots. En la figura 5.17 podemos confirmar que la pérdida de paquetes de video ocurre de manera uniforme en todos los nodos y en el lapso de tiempo comprendido entre 14 y 27 segundos. Además de lo anterior, la gráfica de la figura 5.15.b muestra que el número de paquetes que tienen un retardo menor a 100 ms es mayor al 96% de las muestras recibidas. Si consideramos como paquetes perdidos a aquellos que sufran un retardo mayor a 100 ms, entonces se obtiene una pérdida de paquetes total del 5.43%. Si se hace un cálculo similar para 57 nodos activos en la red, se obtiene una pérdida menor al 4%, el cual es un mejor valor para la transmisión de video en este tipo de redes. Ahora que hemos visto el comportamiento de la red con tráfico UGS y rtPS, analizaremos su comportamiento en un ambiente más real, para esto se agregarán nodos que generen tráfico de tipo BE.

5.2.4 Escenario para tráfico UGS, rtPS y BE

Para este escenario al igual que los anteriores, se usaron grupos de 6 generadores de tráfico: uno de G-711 y uno de G-723 para voz, uno de mpeg4 para video y 3 fuentes de tráfico de Internet. Para calcular el número máximo de nodos que podría soportar la red, se

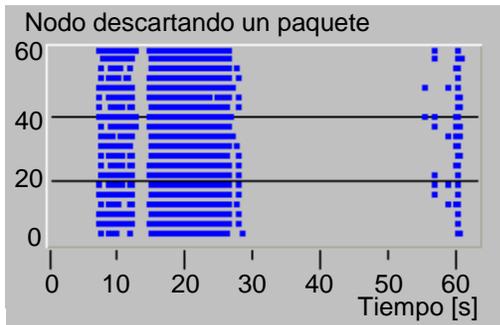


Figura 5.17. Paquetes descartados durante la simulación con 60 nodos activos en la red.

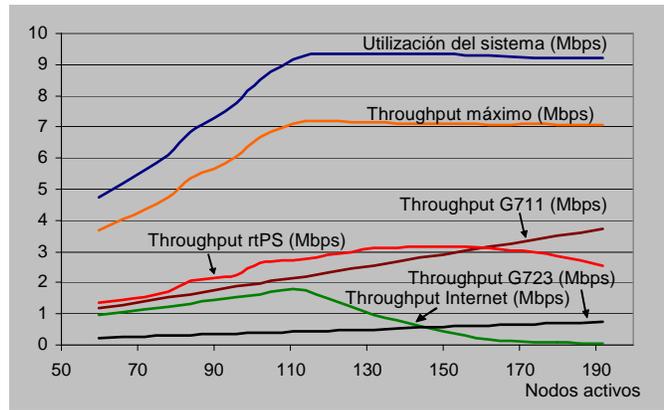


Figura 5.18. Resultados del throughput de la red con tráfico UGS, rtPS y BE.

utilizó nuevamente la ecuación (5.2) de acuerdo a la cual tenemos que: el número máximo de nodos activos que la red puede soportar es de 110 ($6 \cdot 9152 / (38.4 + 38.4 + 38.4 + 38.4 + 166.4 + 176.320) = 110.63$). Para este escenario se corrieron 23 simulaciones comenzando con 60 nodos, es decir 10 grupos de 6 nodos y se fue agregando un grupo de 6 nodos hasta llegar a completar 192 nodos. Las gráficas de los resultados del throughput de las simulaciones se muestran en la Figura 5.18.

En la gráfica se observa como a partir de 114 usuarios, el sistema se encuentra saturado. A partir de ese punto, si se sigue incrementando el número de nodos activos, tanto el throughput total, como la utilización del sistema permanecen casi constantes. Las ligeras variaciones presentadas, se deben principalmente a la utilización de los slots de contención.

En cuanto al tráfico de Internet, se observa que el valor del throughput es prácticamente el mismo cuando se cambia de 108 a 114 usuarios activos. Sin embargo, si comparamos los retardos de acceso para esas simulaciones, veremos que aumentan considerablemente (ver Figura 5.19). El retardo de acceso promedio para el tráfico de Internet, que se muestra en la Figura 5.20, para 108 usuarios es de 0.686 s, mientras que para 114 usuarios el retardo aumenta a 2.3 s. Al seguir aumentando el número de nodos activos en la red, el throughput

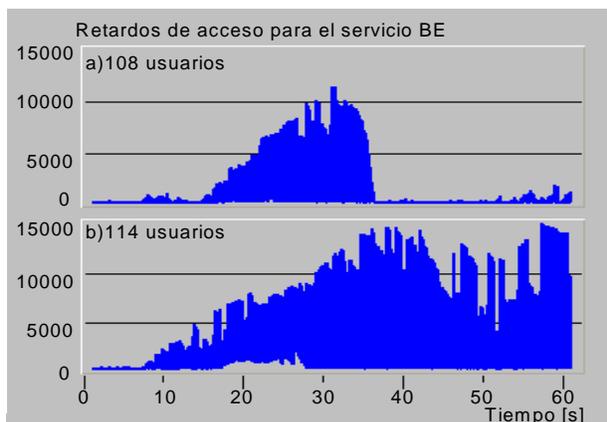


Figura 5.19. Retardos de acceso para tráfico de tipo BE con 108 y 114 usuarios.

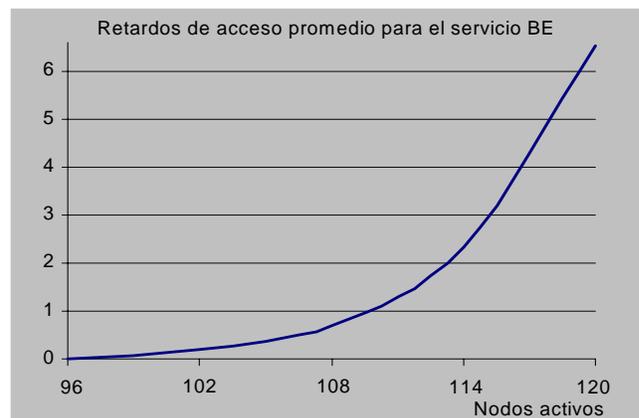


Figura 5.20. Retardos de acceso promedio para tráfico de tipo BE.

de Internet comienza a disminuir y los retardos de acceso promedio siguen incrementándose cada vez más.

Ahora analizaremos el comportamiento del tráfico rtPS. El algoritmo programador de oportunidades de transmisión asigna una prioridad más alta al tráfico rtPS que al tráfico BE. De acuerdo a la ecuación (5.2), el sistema se satura con 110 usuarios, no obstante, todos los paquetes generados cuando hay 114 usuarios activos son recibidos íntegramente y con retardos de acceso menores a 100 ms (ver figura 5.21). De la gráfica de la figura 5.22 observamos que la pérdida de paquetes para tráfico rtPS, se presenta por primera vez cuando hay 126 nodos activos en la red. A estos paquetes perdidos (no enviados) debemos sumar los paquetes descartados por tener un retardo de acceso mayor a 100 ms. Del total de paquetes recibidos el 99.9 % tuvo un retardo menor a 100 ms, por lo que 126 usuarios aún es un número adecuado para soportar tráfico de video.

Para 132 nodos activos el número de paquetes perdidos representa apenas el 0.92% de los paquetes generados y el número de paquetes que llegaron con un retraso mayor a 100 ms representa el 1.1 %. Por tanto la red aún puede dar servicio rtps cuando está sobrecargada con 132 nodos activos, a pesar de que hubo algunos paquetes cuyos retardos de acceso rebasaron los 150 ms.

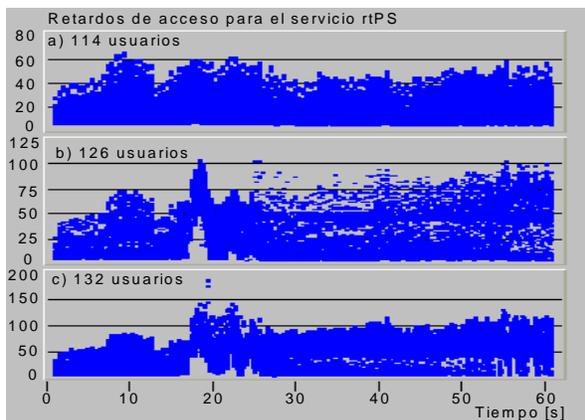


Figura 5.21. Retardos de acceso para el tráfico rtPS.



Figura 5.22. Pérdida de paquetes para el tráfico rtPS.

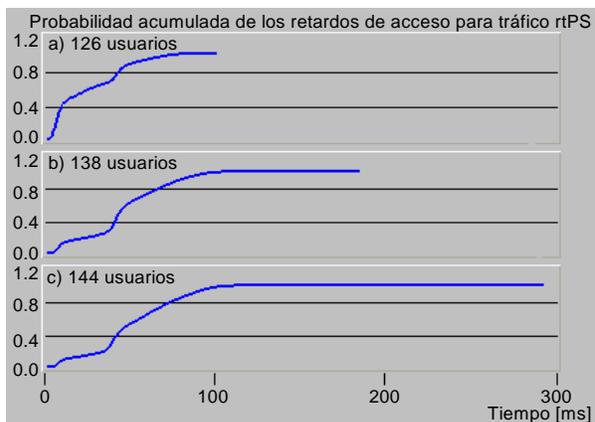


Figura 5.23. Probabilidad acumulada de los retardos de acceso para tráfico rtPS.



Figura 5.24. Retardos de acceso promedio para el tráfico rtPS.

Cuando la red tiene 138 nodos activos, el 1.66 % tiene retardos mayores a 100 ms, y los paquetes perdidos representan el 2.2 %, por lo que podemos afirmar que la red aún puede dar servicio rtPS bajo estas condiciones. Por el contrario, cuando se tienen 144 nodos activos, el 4.29 % son descartados y 1.66% de los paquetes recibidos tienen retardos mayores a 100 ms. De lo anterior, podemos concluir que el límite para que una red pueda dar servicio de tipo rtPS cuando está sobrecargada, con esta configuración de nodos, es de 138. Cabe hacer notar, el hecho de que el límite teórico, para saturar la red con tráfico de voz y video en esta red es de 144. Sin embargo, en la figura 5.18, observamos que todavía hay tráfico de Internet que se transmite en la red. Si analizamos el tamaño de las muestras de las 10 películas, el data rate necesario para transmitir el promedio de la media de cada película es de 176.32 kbps. Sabemos también, que las 10 películas son elegidas aleatoriamente y en este caso específico, la tasa de transmisión necesaria para transmitir el promedio de la media de las películas seleccionadas fue de 147.2 kbps. A pesar de lo anterior, el resultado teórico obtenido con 176.32 kbps fue bastante cercano al obtenido durante la simulación.

A continuación analizaremos el comportamiento del tráfico UGS en esta red. Como ya se mencionó el tráfico UGS tiene la prioridad más alta que los demás tipos de tráfico, por lo que se espera que tenga retardos menores y menor pérdida de paquetes que el tráfico de tipo rtPS. Como se vió en la figura 5.18.a medida que se aumenta el número de nodos activos, el throughput para tráfico UGS, tanto G-711 como G-723, también aumenta linealmente. En la misma gráfica se puede observar como el throughput obtiene este aumento permanente, que se sigue manteniendo aún cuando el sistema está saturado, a expensas de los demás tipos de tráfico que tienen menor prioridad. Sabemos, de acuerdo a la ecuación (5.2) que el sistema se encuentra saturado con 114 usuarios. También sabemos que el límite teórico para dar servicio a usuarios rtPS es 144. En la gráfica de la figura 5.25 se muestran los retardos de acceso del tráfico de tipo UGS, tanto G-711 como G-723, cuando hay 114, 144 y 192 usuarios activos.

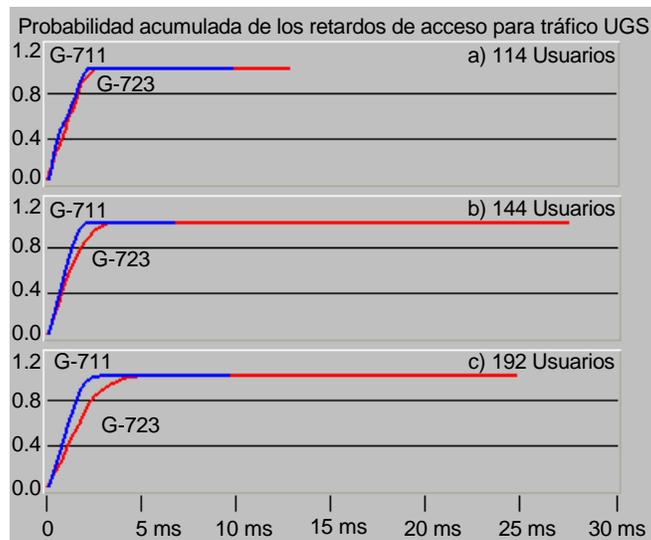


Figura 5.25. Retardos de acceso promedio para el tráfico UGS.

5.2.5 Escenario para tráfico UGS, rtPS y BE con pesos ponderados

Este escenario es exactamente igual al escenario anterior. Se usaron los mismos tipos de generadores de tráfico, es decir, grupos con un generador G-711 y otro G-723 para voz, uno de mpeg4 para video y 3 fuentes de tráfico de Internet. La diferencia con el escenario anterior se encuentra en el algoritmo de asignación de slots. En las simulaciones anteriores no se asignaron pesos ponderados a cada tipo de tráfico. En esta simulación se asignó un peso de 0.40 a UGS y de 0.35 a rtPS. Esto quiere decir que se reservó el 40% del ancho de banda para el tráfico de voz, el 35% para tráfico de video y el ancho de banda restante se asignó al tráfico de Internet.

De acuerdo a la ecuación 5.2, el número total de usuarios de voz, tomando en cuenta tráfico G-723 y G-711, que puede soportar la red es de 35 ($0.4 * (2 * 9152 / (166.4 + 38.4)) = 35.75$). Debido a que en las simulaciones se utilizaron grupos de 6 usuarios, el número total de usuarios para saturar el ancho de banda asignado a voz es de 107.

$$MaxSS = 6 * \frac{Utilización_del_sistema}{VoIP_{G711} + VoIP_{G723}} = 6 * \frac{9152 * 0.4}{166.4 + 38.4} = 107.27$$

En la gráfica de la figura 5.26 podemos ver que efectivamente, cuando hay 108 usuarios en la red, se satura el ancho de banda asignado a voz. De acuerdo a las gráficas de la figura 5.27, para 108 usuarios activos, el 100% de los paquetes tuvo retardos menores a 10 ms mientras que para 114 nodos activos, los retardos fueron menores 10 ms para G-711 y de 30 ms para G-723. Las simulaciones con menos de 108 usuarios activos no mostraron pérdida de paquetes de voz. Para 108 usuarios, la simulación reportó que el total de los paquetes perdidos es de 0.118% para el tráfico de tipo G-711 y de 0.281% para el tipo G-723, mientras que para 114 usuarios el total de paquetes perdidos es de 5.59% para G-711 y de 6.86% para G-723 (ver figura 5.28). Por tanto el máximo número de nodos de voz que puede soportar una red con esta configuración es de 108.

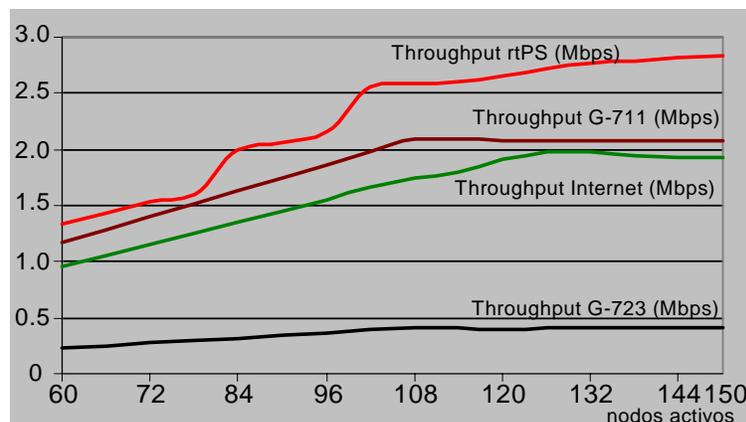


Figura 5.26. Resultados del throughput cuando se usan pesos ponderados.

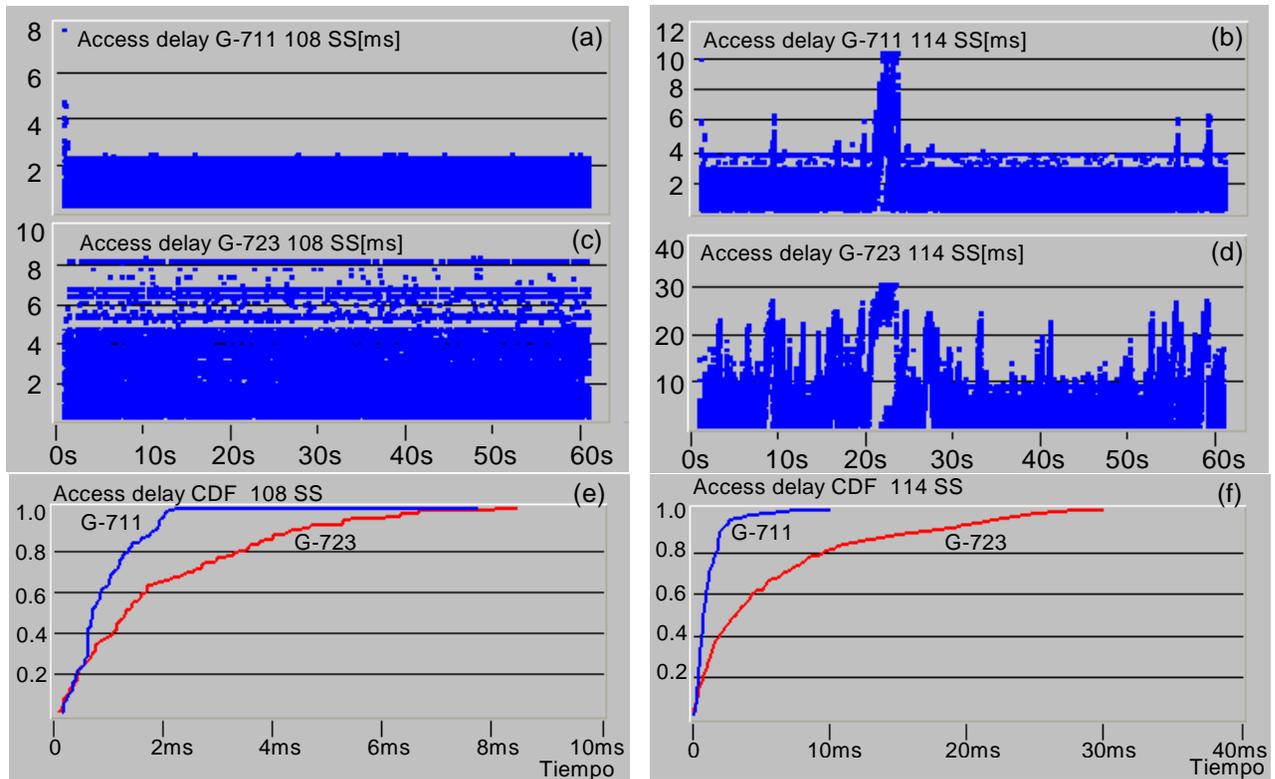


Figura 5.27. Retardos de acceso para el tráfico de voz para una red con 108 y 114 usuarios.

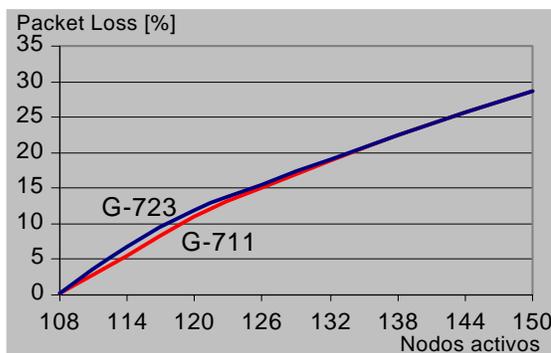


Figura 5.28. Pérdida de paquetes de voz.

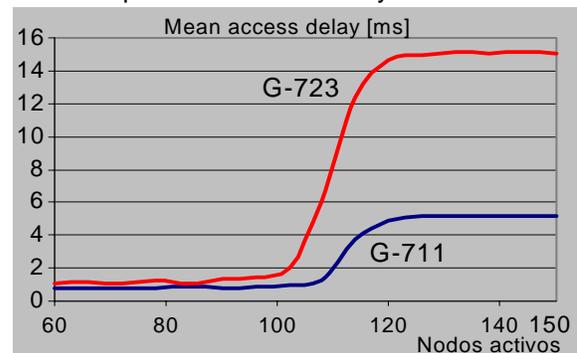


Figura 5.29. Retardos de acceso promedio para tráfico de voz.

Para video tenemos asignado el 35 % del ancho de banda. El número de slots requeridos para transmitir una muestra con un tamaño que sea el promedio de las muestras promedio de cada película es de 55.1. De manera que para video se utilizan 55 slots por muestra y un slot de poleo, o sea 56 slots por muestra. Calculando la tasa de transmisión requerida para transmitir este servicio tenemos:

$$56 \left[\frac{\text{slots}}{\text{muestra}} \right] * 25 \left[\frac{\text{muestras}}{\text{segundo}} \right] * 16 \left[\frac{\text{bytes}}{\text{slot}} \right] * 8 \left[\frac{\text{bits}}{\text{byte}} \right] = 179,200 \left[\frac{\text{bits}}{\text{segundo}} \right]$$

Ahora calculamos el número máximo de usuarios activos que puede haber en la red para utilizar el ancho de banda asignado al tráfico rtPS:

$$MaxSS = 6 * \frac{9152 * 0.35}{179.2} \approx 107.25$$

Las gráficas de la figura 5.30 y de la figura 5.31 muestran que para 102 usuarios todos los paquetes generados llegaron íntegramente a su destino con un retardo promedio de 75.57 ms. En la figura 5.26 podemos apreciar que a partir de 102 usuarios activos en la red, el throughput para tráfico de video ya no aumenta de manera significativa. Estos resultados muestran que para 108 usuarios activos hubo una pérdida de paquetes del 0.86%, mientras que el retardo de acceso promedio fue de 201 ms. Sin embargo, aún cuando el retardo promedio y el throughput comprueban los resultados de la ecuación 5.2, es necesario observar el retardo de acceso que sufren los paquetes de manera individual durante las simulaciones. Para 102 usuarios activos sabemos que el retardo de acceso promedio fue de 75.57 ms, sin embargo únicamente el 83.57 % de los paquetes generados tuvo retardos menores a 100 ms, y durante los breves periodos de congestión se llegaron a registrar retardos de hasta 594 ms (figura 5.32). En las simulaciones observamos que para 96 nodos activos, el retardo de acceso promedio fue de 12.66 ms y el mayor retardo registrado fue de 85 ms. De acuerdo a estos resultados, el mayor número de nodos activos que puede soportar esta red es de 96 ó de manera equivalente, el 35% del ancho de banda será capaz de dar servicio de manera confiable a 16 usuarios que generen tráfico de video. Comparando este resultado con el número de usuarios obtenidos con la ecuación 5.2 vemos que es el 89% del valor calculado inicialmente.

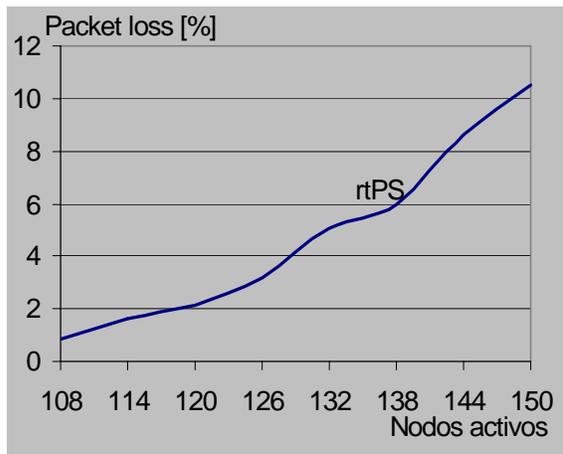


Figura 5.30. Pérdida de paquetes de tráfico rtPS.

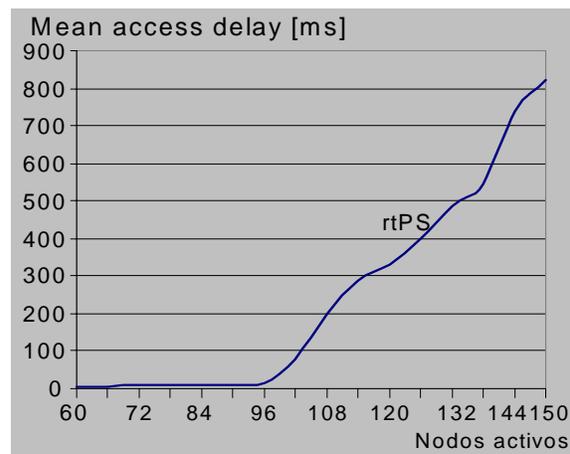


Figura 5.31. Retardos de acceso promedio para tráfico rtPS.

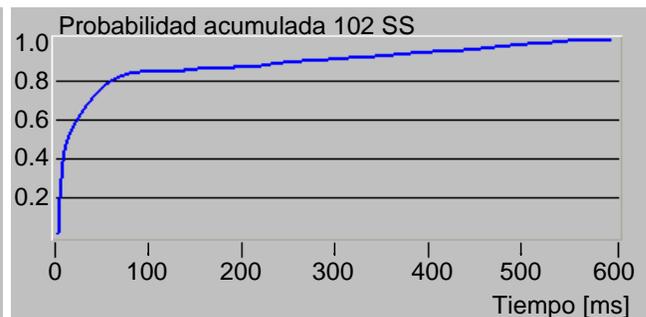
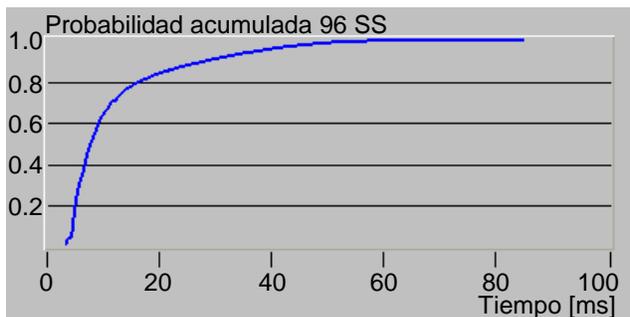


Figura 5.32. Retardos de acceso de los paquetes rtPS para 96 y 102 nodos activos.

Ahora para el tráfico de Internet nos queda el 25% del ancho de banda disponible. Este ancho de banda se utilizará completamente cuando se tengan 119 usuarios:

$$MaxSS = 6 * \frac{9152 * 0.25}{38.4 * 3} \approx 119.16$$

Esto quiere decir, que cuando se tengan 108 ó 114 usuarios activos habrá ancho de banda asignado al tráfico BE que no estará siendo ocupado mientras que el ancho de banda asignado al tráfico de voz y a rtPS estará totalmente agotado.

De acuerdo a la gráfica de la figura 5.26 se observa que el máximo throughput para BE, se obtiene cuando hay 126 nodos activos en la red. A partir de ese punto, el throughput se mantiene casi constante aunque tiene ligeras variaciones debido a la fragmentación y empaquetamiento de los paquetes.

Aún cuando el tráfico de tipo BE no es sensible a los retardos, el retardo obtenido durante las simulaciones es bastante bajo. Para 120 usuarios activos las simulaciones reportaron un retardo de acceso promedio de 63.25 ms, mientras que para 126 el retardo fue de 434 ms (ver figura 5.33 y figura 5.34).

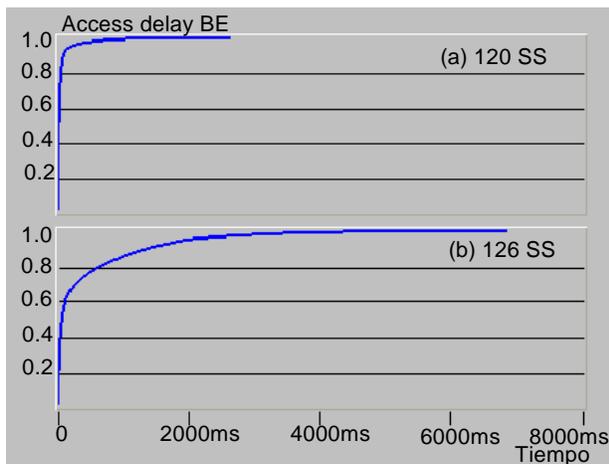


Figura 5.33. Retardos de acceso para 120 y 126 nodos activos.

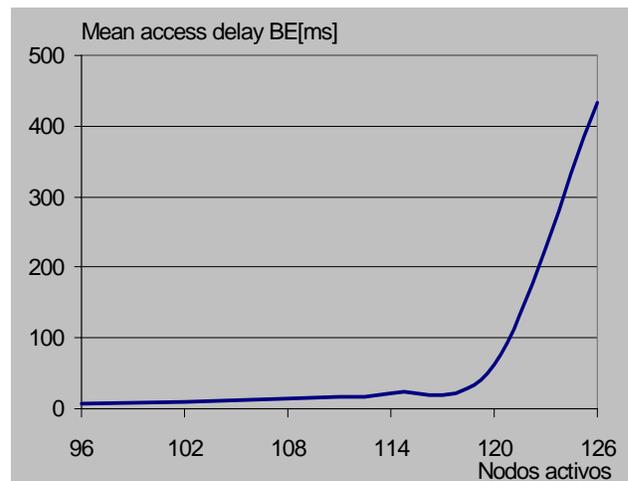


Figura 5.34. Retardos de acceso promedio del tráfico de Internet.

CAPÍTULO VI

CONCLUSIONES

6.1 Discusiones finales

El modelo propuesto para una red BWA con diferentes tipos de tráfico permite una asignación eficiente del ancho de banda entre aquellos nodos activos en la red. Estos usuarios pueden tener diferentes requerimientos de calidad de servicio. Sin embargo todavía hay pequeños detalles que se podrían mejorar.

A lo largo de todo el trabajo también quedaron de manifiesto las ventajas de utilizar un software de simulación, sin él hubiera sido imposible obtener los resultados que se obtuvieron. Con este software se simulan los eventos reales y se les asocia un instante de tiempo, tal y como ocurre en la vida real. Los resultados que se obtuvieron con estas simulaciones son de escenarios particulares. Sin embargo si se quisiera conocer el desempeño de una red bajo otras demandas de tráfico diferentes a las utilizadas aquí habría que cambiar las condiciones de los escenarios. Para ello sólo es necesario cambiar algunos de los parámetros de la simulación. Gracias a esto se pueden conocer el comportamiento que tendría la red sin necesidad de implementarla físicamente y de esa manera se puede tener una idea de la viabilidad de los sistemas BWA.

6.2 Contribuciones

Con este trabajo se ha obtenido un modelo de red que maneja Calidad de Servicio, con el cual se pueden modelar diferentes escenarios de tráfico. Con él será posible predecir el comportamiento dinámico de una red bajo diferentes demandas de tráfico antes de implementarlo físicamente. El modelo obtenido cumple efectivamente con los requerimientos establecidos en el estándar IEEE 802.16.

En esta tesis, hemos demostrado que es posible el manejo de diferentes tipos de tráfico en redes BWA. Cada tipo de tráfico tiene diferentes requerimientos en cuanto a jitter, intervalos de poleo, prioridad, etc. Con este modelo es posible también hacer simulaciones

asignando una porción del ancho de banda a cada tipo de tráfico, de acuerdo a las necesidades de cada usuario.

6.3 Trabajo futuro

En un futuro se espera poder realizar una asignación dinámica de pesos ponderados de acuerdo al número de usuarios y al tipo de tráfico de cada uno de ellos. Al asignar dinámicamente el ancho de banda a cada tipo de tráfico, se espera asignar el peso que vaya más de acuerdo al número de usuarios activos que se tengan en cada momento en la red. Con ello se tendrá un uso más eficiente del ancho de banda. Como se mencionó anteriormente el estándar define cuatro tipos diferentes de tráfico. De ellos, todavía hace falta por incorporar el tráfico nrtPS al modelo. Con la incorporación del tráfico nrtPS se tendrán cubiertos todos los requerimientos de los tipos de tráfico definidos en el estándar IEEE 802.16 y se tendrá un algoritmo para operar apropiadamente en una red inalámbrica de este tipo.

Adicionalmente a la asignación de pesos ponderados es necesario también contar con políticas de admisión las cuales permitan asegurar la calidad de servicio. Al incrementar el número de usuarios de cualquier tipo de tráfico el throughput total se mantiene en un máximo. Sin embargo el continuo aumento en el número de nodos activos en la red ocasiona que el throughput vaya aumentando al igual que la pérdida de paquetes. Ante tal situación es preferible negar el acceso a algún usuario cuando la red está saturada con el fin de continuar ofreciendo un buen servicio al resto de los usuarios que ya están activos.

Otra mejora del sistema podría ser la de un mejor manejo de los slots de contención, ya que en los primeros momentos de la simulación, varios usuarios comienzan a generar paquetes en el mismo instante de tiempo. Esto ocasiona que tenga que utilizar los slots de contención y si son bastantes usuarios, se puede tener una pérdida de paquetes durante el inicio. Asimismo, durante todo el tiempo que dura la simulación, los nodos que genera tráfico de tipo BE utilizan esta región para solicitar la asignación de ancho de banda. En las simulaciones realizadas, los slots de contención se asignaron en una cantidad fija. Esta cantidad podría hacerse variable de acuerdo al número de usuarios en la red.

Además de las simulaciones realizadas, se pueden hacer utilizando video mpeg4, pero con una mejor calidad, ó en formato CIF (Common Intermediate Format), para tener una imagen más grande.

6.4 Conclusiones finales

En este trabajo se implementaron técnicas de Calidad de Servicio para redes BWA 802.16. El algoritmo presentado es práctico, compatible con los requerimientos de calidad de servicio y fácil de implementar.

El algoritmo implementado se puede utilizar en dos modos. En el primero de ellos se asignan pesos ponderados a cada tipo de tráfico. De esa manera se tiene una mayor versatilidad en el manejo de calidad de servicio y se asegura un mejor aprovisionamiento del servicio que el usuario tenga contratado.

Cuando no se utilizan pesos ponderados es posible proporcionar una alta prioridad de transmisión a los servicios que tienen un jitter más bajo. Con ello se asegura el dar servicios en tiempo real tales como VoIP con retardos de acceso muy bajos aún durante periodos de congestión, reservando un ancho de banda mínimo para los servicios nrtPs y BE.

Al utilizar pesos ponderados se proporciona un servicio continuo a un número fijo de usuarios de cada tipo. Al hacer esto, se asegura que aún durante los periodos de congestión, aquellos servicios con una prioridad más baja aún tendrán ancho de banda disponible para seguir transmitiendo.

Se espera que este trabajo sirva como base para estudios posteriores para estudios más profundos de calidad de servicio en este tipo de redes.

REFERENCIAS

BIBLIOGRÁFICAS

Artículos y libros

- [1] Bölcksei, Helmut; Paulraj, Arogyaswami J; Hari, K.V.S.; Nabar, Rohit U.; Lu, W.W. "Fixed Broadband Wireless Access: State of the Art, Challenges, and Future Directions". *IEEE Communications Magazine*. Volume 39, Issue 1. January 2001. Pages 100-108.
- [2] Bragg, Arnold W. "Quality of Service: Old Idea, New Options". *IEEE IT Professional*. Volume 1, Issue 5. September- October 1999. Pages 37-44.
- [3] El-Sayed, Mohamed; Jaffe, Jeffrey. "A View of Telecommunications Network Evolution". *IEEE Communications Magazine*. Volume 40, Issue 12. December 2002. Pages 74-81.
- [4] Eklund, Carl; Marks, R. B.; Stanwood, K. L.; Wang, S.; "IEEE Standard 802.16: A Technical Overview of the WirelessMAN™ Air Interface for Broadband Wireless Access". *IEEE Communications Magazine*. Volume 40, Issue 6. June 2002. Pages 98-107.
- [5] IEEE 802.16-2001, "IEEE Standard for Local and Metropolitan Area Networks — Part 16: Air Interface for Fixed Broadband Wireless Access Systems," Apr. 8, 2002
- [6] Evans, B.G.; Baughan, K.; "Visions of 4G". *Electronics & Communications Engineering Journal*, Volume 12, Issue 6, December 2000. Pages 293-303.
- [7] GuoSong Chu; Deng Wang; Shunliang Mei. "A QoS Architecture for the MAC Protocol of IEEE 802.16 BWA System". *Communications, Circuits and Systems and West Sino Expositions, IEEE 2002 International Conference on*. Volume 1, 29 June-1 July 2002 pages 435-439 vol 1.
- [8] Hantro Products Oy. "MPEG4 Codec overview". Hantro Products. Finland. March 2001. Oulu, Finland.

- [9] Hawa, Mohammed; Petr W., David. "Quality of Service Scheduling in Cable and Broadband Wireless Access Systems". Quality of Service, 2002. Tenth IEEE International Workshop on. 15-17 May 2002 Pages 247-255.
- [10] Hertrich Daniel."MPEG4 Video Transmission in wireless LANs. Basic QoS support on the data link layer of 802.11b". Inédita. TU Berlin. Minor thesis, October 12, 2002. Pages 37.
- [11] Huidobro, José Manuel "Wimax. Un estandar emergente". Antena de Telecomunicación, Septiembre 2004. Páginas 67-68.
- [12] Klaus-Dieter Schunke, "IP Telephony- An Introduction", Braunschweig Technical University; Institute for Communications Technology. EuroCableLabs. 9 december 1998. Pages 19.
- [13] Landsiedel Olaf and Minden Gary "MPEG-4 for Interactive Low-delay Real-time Communication", Technical Report, information and Telecommunication Center, University of Kansas, ITTC-FY2004-TR-23150-10, December 2003.
- [14] Parry,Richard. "Overlooming 3G"; Potentials, IEEE, Volume 21 Issue 4. Oct-Nov 2002. Pages 6-9.
- [15] Raivio, Y. "4G- Hype or reality", 3G Mobile Communication Technologies, 2001, 26-28 March 2001, Conference Publication No. 477, IEEE 2001. Pages 346-350.
- [16] Rangel, Victor; Edwards, Robert M.; Tzerefos, Polychronis; Schunke Klaus-Dieter. "Delivery of Low Rate Isochronous Streams over the Digital Video Broadcasting/Digital Audio-Visual Council Cable Television Protocol", *IEEE Transactions on Broadcasting*, Vol. 48, Issue 4, December 2002. Pages:307-316.
- [17] Rangel, V; Edwards, R. "Performance Analysis and Optimisation of the Digital Video Broadcasting/Digital Audio Visual Council Cable Modem Protocol for the Delivery of Isochronous Streams". Global Telecommunications Conference, 2001. GLOBECOM '01. IEEE. Volume 1, 25-29 Nov. 2001. Pages 430-434 vol1.
- [18] V. Rangel and R. M. Edwards, "Performance Evaluation of Adaptive Contention Slot Allocators for CATV Networks based on the European Cable Communications Protocol DVB/ DAVIC", Journal of the SCTE, vol. 24, Sep. 2002, pp 24-30
- [19] SCTE DSS 00-05, Data-Over-Cable Service Interface Specification (DOCSIS) SP-RF1v1.1-I05-000714, "Radio Frequency Interface 1.1 Specification," July 2000.
- [20] Tachikawa; Keiji "A Perspective on the Evolution of Mobile Communications";IEEE Communications Magazine. Volume 41, Issue10, October 2003. Pages 66-73.
- [21] Tanenbaum, Andrew S., "Redes de computadoras", cuarta edición, Pearson Educación, México 2003.

- [22] Varshney, Upkar; Jain, Radhika. "Issues in Emerging 4G Wireless Networks". Computer, Volume 34, Issue 6. June 2001. Pages 94-96.
- [23] Xinjie Chang. "Network Simulations with Opnet". Simulation Conference Proceedings, 1999 Winter. Volume1, 5.8 Dec. 1999 Pages 307-314 vol. 1.

Páginas de internet

- [24] <http://users.ece.gatech.edu/~jxie/4G/index.html>, Mobility Management in 4G Wireless Systems; Supported by: National Science Foundation (NSF)
- [25] <http://www.eurotechnology.com/index.html>
- [26] <http://www.mobileinfo.com/Default.asp>
- [27] <http://whatis.techtarget.com/>
- [28] <http://www.3gtoday.com/technology/index.html>
- [29] <http://www.3gpp.org/>
- [30] <http://www.3gpp2.org/>
- [31] <http://www.whatis.org/tg2k/t1g2k.html> American National Standard for Telecommunications. Telecom Glossary 2000
- [32] <http://grouper.ieee.org/groups/802/16/pub/background.html>
- [33] <http://trace.eas.asu.edu/TRACE/ltvt.html> "Trace Files Movies"
- [34] <http://wimaxxed.com/> "The exclusive news source for everything WiMax"
- [35] <http://www.intel.com/netcomms/technologies/wimax/experiences.htm> "WiMAX Experiences-Case Studies & Videos"
- [36] <http://www.cdt.luth.se/projects/siram/mP4/> "Peppar MPEG4 Player- MP4"
- [37] <http://www.chiariglione.org/mpeg/index.htm> "The MPEG Home Page"
- [38] <http://archive.dstc.edu.au/RDU/staff/jane-hunter/video-streaming.html> "A Review of Video Streaming over the Internet"

- [39] <http://www.cs.wpi.edu/Resources/Techreports/99-19/index.html>
“Video Redundancy –A Best-Effort Solution to Network Data Loss”
- [40] <http://www.cit.cornell.edu/atc/itsupport/streaming.shtml>
“Audio and Video Streaming”
- [41] <http://www.microsoft.com/windows/windowsmedia/howto/articles/OptimizingLowBitRate.aspx> “Optimizing Low Bit Rate Audio”
- [42] <http://www.opnet.com/>
- [43] <http://www.cudenver.edu/Resources/ITS/troubleshootHowTo/FAQ/videoServicesFAQ.htm>
- [44] <http://www.cse.wustl.edu/~jain/cis788-99/ftp/compression/index.html>
“Video Compression: MPEG-4 and Beyond”
- [45] http://www.telefonica.es/sociedaddelainformacion/pdf/publicaciones/movilidad/capitulo_3.pdf
- [46] http://www.rcru.rl.ac.uk/radioresearch/ray_trace/delay.htm
“delay spread”

GLOSARIO DE TÉRMINOS Y ACRÓNIMOS

1xRTT	1 time Ratio Telephone Technology.
3GPP	third generation partnership project.
Access delay	Tiempo transcurrido entre la solicitud de acceso y el acceso efectivo a una red.
ADSL	Asymmetric Digital Subscriber Line ("Línea de Abonado Digital Asimétrica"). Consiste en una línea digital de alta velocidad, apoyada en el par trenzado de cobre que lleva la línea telefónica convencional o línea de abonado.
AMPS	Advanced Mobile Phone System. (Sistema Telefónico Móvil Avanzado) es un sistema de telefonía móvil de primera generación (1G, voz analógica) desarrollado por los laboratorios Bell. Se implementó por primera vez en 1982 en Estados Unidos. Se llegó a implantar también en Inglaterra y en Japón, aunque con otros nombres, TACS y MCS-L1 respectivamente.
Backhaul:	Red de retroceso. Conexión de baja, media o alta velocidad que conecta a computadoras u otros equipos de telecomunicaciones encargados de hacer circular la información. Los backhaul conectan redes de datos, redes de telefonía celular y constituyen una estructura fundamental de las redes de comunicación. Usado para interconectar redes entre sí utilizando diferentes tipos de tecnologías alámbricas o inalámbricas.
BHC	Bits per second per Hertz per Cell.
BS	Base Station ó Estación Base.
BTS	Base Transceiver Station ó Estación Base Transceptora.
Burst	En cualquier sistema basado en TDMA (Acceso Múltiple por División de Tiempo) existe la necesidad de transmitir información en ranuras de tiempo. Esto involucra el envío de pequeños bloques de información conocidos como bursts ó ráfagas.
C/I	Razón de la señal portadora a interferencia (Carrier-to-Interference Ratio).
C/N	Razón de la señal portadora a ruido (Carrier-to-Noise ratio).
CCI	Co-Channel Interference ó Interferencia Co-canal.
CDMA	Code Division Multiple Access ó Acceso Múltiple por División de Código.
cdmaOne	Se refiere al protocolo de interfase inalámbrica ITU IS-95 (CDMA) que fue estandarizado en 1993.
CID	Connection Identifier. Es un término usado en las redes 802.16 para especificar una dirección en la capa MAC que identifica a una estación suscriptora con su equivalente en la estación base.
CPE	Customer Premises Equipment. Equipo en las Instalaciones de Cliente.
CS	Convergence sublayer. Subcapa de convergencia.
DFE	Decision-Feedback Equalization.
DECT	Digital Enhanced Cord-Less Telecommunication. Telecomunicaciones Digitales Mejoradas e Inalámbricas. El estándar DECT define una tecnología de acceso por radiofrecuencia para comunicaciones inalámbricas digitales.
Delay jitter	Es la fluctuación ó variación en el retardo de extremo a extremo (End to End) que ocurre entre un paquete y el siguiente del mismo flujo ó servicio.
Delay Spread	Es el término con que se designa a la diferencia en tiempo entre la primera y la última reflexión en arribar a un receptor.
DIUC	Downlink Interval Usage Code. Es un término utilizado en las redes IEEE 802.16 para describir un IUC (Interval Usage Code) específico en el enlace descendente.

DO	Data Only.
DSA-ACK	Dynamic Service Addition Acknowledge Primary Management.
DSC-ACK	Dynamic Service Change Acknowledge Primary Management.
DS-CDMA	Direct Sequence Code-Division Multiple Access.
DSC-REQ	Dynamic Service Change Request Primary Management.
DSC-RSP	Dynamic Service Change Response Primary Management.
DSL	Digital Subscriber Line.
EC	Encryption Control.
EDGE	Enhanced Data Rates for GSM Evolution. Es una versión mejorada de GSM.
Fading	Fenómeno que consiste en la pérdida de señales de telecomunicaciones.
FCC	Federal Communications Comision.
FDD	Frequency Division Duplex.
FEC	Forward Error Correction.
Frequency diversity	Transmisión y recepción en las cuales la misma señal de información es transmitida y recibida simultáneamente en 2 o más frecuencias independientes.
FDMA	Frequency Division Multiple Access.
GMSK	Gaussian Minimum Shift Keying.
GFR	Guaranteed Frame Rate.
GMSK	Gaussian Minimum Shift Keying.
GPRS	General Packet Radio Service.
GSM	Global System for Mobile communication.
IUC	Interval Usage Code. Término utilizado en redes 802.16 para identificar un perfil de ráfagas que puede ser utilizado en un enlace de subida ó de bajada.
IMT-2000	International Mobile Telecommunications-2000. Es el estándar global para la tercera generación de comunicaciones inalámbricas de acuerdo a la definición de la ITU.
IP	Internet Protocol.
ITU	International Telecommunications Union.
Jitter	Variación en los tiempos de llegada entre los paquetes.
LAN	Local Area Network.
LMDS	Local Multipoint distribution services.
LOS	Line Of Sight.
MAC	Medium Access Control.
MAN	Metropolitan Area Network.
MDS	Multipoint Distribution Services.
Mesh network	Una red de malla (también llamada red de "varios saltos"), es una red que emplea un arreglo que puede ser parcial ó completo. En la topología de malla completa, cada nodo está conectado directamente a cada uno de los otros. En la topología de malla parcial, algunos nodos están conectados al resto, pero algunos de estos nodos sólo están conectados a aquellos nodos con quienes intercambian mayormente datos.
MMDS	Multichannel Multipoint Distribution Services.
MMSE	Minimum Mean Square Error.
Multiplexaje estadístico	Es un tipo de multiplexaje que explota la no coincidencia de picos de velocidad, transportando altas velocidades instantáneas en una salida de baja velocidad. Es similar al multiplexaje por división de tiempo (TDM), excepto que, en lugar de asignar arbitrariamente un slot de tiempo a cada señal, cada señal es asignada a un slot de acuerdo a su prioridad y necesidad. Este tipo de multiplexaje asegura que los slots no serán desperdiciados (a diferencia de TDM) pero en lugar de eso, se pierde más tiempo. También puede haber pérdida de datos debido a overflow.
MWIF	Mobile Wireless Internet Forum.
NFS	Network File System.
NLOS	Non-Line-Of-Sight.
NTSC	Es un sistema de codificación y transmisión de televisión analógica desarrollado en Estados Unidos en torno a 1940, y que se emplea en la actualidad en la mayor parte de América y Japón, entre otros países. El nombre viene del comité de expertos que lo desarrolló, el National Television System(s) Committee.
NTT DoCoMo	Es el principal operador telefónico en Japón y parte de la compañía Japonesa NTT (Nippon Telegraph and Telephone Corporation) la cual provee servicios inalámbricos

	<p>incluyendo celular, paging y servicios de telefonía marítima y aérea. DoCoMo es la abreviación de la frase "Do Communications Over the Mobile Network", pero también significa "en cualquier lugar en cualquier momento" en Japonés.</p>
OFDM	Orthogonal Frequency-Division Multiplexing.
PAL	Son las siglas de Phase Alternating Line (en español línea alternada en fase). Es el nombre con el que se designa al sistema de codificación empleado en la transmisión de señales de televisión analógica en color en la mayor parte del mundo. Es de origen alemán y se utiliza en la mayoría de los países africanos, asiáticos y europeos, además de Australia y algunos países latinoamericanos. El sistema PAL surgió en el año 1963, de manos del Dr. Walter Bruch en los laboratorios de Telefunken en su intento por mejorar la calidad y reducir los defectos en los tonos de color que presentaba el sistema NTSC. No obstante, los conceptos fundamentales de la transmisión de señales han sido adoptados del sistema NTSC.
PDU	Protocol Data Unit.
QCIF	Quarter Common Intermediate Format.
RAKE Receiver	Es un receptor diseñado para compensar los efectos de las multitrayectorias. Esto se hace utilizando diversos "sub-receptores" cada uno retrasado ligeramente con el fin de sintonizar cada uno de los componentes multitrayectoria. Cada componente es decodificada de manera independiente pero recombinada más adelante. Esto se hace con el fin de obtener el máximo provecho de cada trayectoria de transmisión.
RLP	Radio Link Protocol.
RTT	Radio Telephone Technology.
SAID	Security Association Identifier.
SC	Single-Carrier.
SDU	Service Data Units.
SMS	Short Message Services. Servicio disponible en los sistemas de telefonía móvil que permite el envío de mensajes cortos de texto.
SOHO	Small Office/Home Office.
TDD	Time Division Duplex.
TDMA	Time Division Multiple Access.
Throughput	Velocidad real de transporte por tiempo de trama.
Time diversity	En este tipo de transmisiones, un dato se transmite varias veces. Se utiliza cuando el canal de transmisión sufre de efecto Doppler. Esto puede ocurrir como resultado de un movimiento rápido relativo entre el transmisor y el receptor.
TLV	Type-Length-Value.
UIUC	Término usado en las redes 802.16 para describir un IUC específico en el enlace de subida ó Uplink.
UWBM	Ultra-Wideband Modulation.
UMTS	Universal Mobile Telecomm-unications System También llamado WCDMA..
VLAN	Virtual Local Area Network.
WCS	Wireless Communications Services.
WCDMA	Wideband Code Division Multiple Access.

APÉNDICE A

MUESTRAS DE VIDEO UTILIZADAS EN LAS SIMULACIONES

Descripción general de MPEG 4

Mpeg 4 es el sucesor de los estándares de compresión de audio y video MPEG1 y MPEG2, los cuales fueron desarrollados por MPEG (Moving Pictures Expert Group, parte de la organización Internacional de Estándares ISO) para hacer posible el envío de audio y video en ráfagas en medios en los cuales se cuenta con un ancho de banda limitado ó poco espacio de almacenamiento en disco

MPEG1 fue creado en 1992 y define la compresión de video (junto con 2 canales de audio incluidos) con una resolución de 352 x 240 pixeles y una tasa de 25 cuadros por segundo en formato PAL ó 30 cuadros por segundo en formato NTSC. La tasa de transmisión para MPEG1 está definida como 170 kbps, lo cual equivale a una tasa de compresión de datos de aproximadamente 100:1 comparado con el formato Europeo PAL. Debido a la baja resolución de este estándar no ha sido un sustituto adecuado para los estándares de TV.

En 1994 se desarrolló un nuevo estándar llamado MPEG2. Éste soporta resoluciones mucho más altas que MPEG1 (“baja” con resolución MPEG1, la “media” de 720 x 480, “alta 1440” con 1440 x 1152 y “alta 1920” de 1920 x 1080 pixeles”) y soporta 5 canales de audio multiplexados. MPEG2 es el estándar de compresión utilizado en los sistemas de video digital DVD.

El formato de audio MPEG1/2 capa 3, mejor conocido como MP3, se hizo muy popular en el mundo de la música debido a la tasa de compresión tan alta que se logra sin perder mucha calidad en el audio.

El formato MPEG4, fue liberado en 1999, éste incluye mucho más que solamente compresión de audio y video, describe estándares para codificar por ejemplo:

- Objetos de audio y video.
- Descripciones de escenas con BIFS (Binary Format Scenes), un formato que describe las relaciones espacio-temporales entre los diferentes objetos de una escena.

- Voz sintética (por medio de fonemas).
- Parámetros de movimiento facial (para una voz sintética).
- Música sintética (similar a MIDI).
- Texto con parámetros tales como tipo de letra, tamaño, posición, color, etc.

Así, cada escena puede contener diferentes objetos, cada uno con su posición dentro de la escena, descrita mediante coordenadas en 2D ó 3D y con su propia sincronización de tiempo. La Figura A.1 muestra la arquitectura general de un sistema de codificación y decodificación MPEG4. En la Figura A.2 se muestra como se define una escena en el formato MPEG4. Aún existe un problema sin resolver, y es precisamente la generación de esos objetos a partir de un escenario completo de audio y video. Sin embargo MPEG4 define un estándar para transmitir con estos objetos por separado.

Imaginemos por ejemplo, un programa de noticias que se transmite por TV. Actualmente, la imagen se transmite píxel por píxel, y el audio es transmitido muestra por muestra (si se usa transmisión digital). Podríamos utilizar MPEG1 ó MPEG2, para comprimir el audio y el video, con ello perderíamos un poco de calidad en la imagen, pero a cambio ahorraremos ancho de banda.

MPEG4 ofrece posibilidades totalmente nuevas: En el estudio de producción, el reportero podría ser un objeto que es transmitido en MPEG2, como video comprimido, su voz es sincronizada, comprimida y transmitida como un objeto separado. El fondo no se transmite ni como video ni como imagen, sino que únicamente se transmite el color de fondo, por ejemplo "azul". Y el texto de las noticias que lee el locutor también se transmite simultáneamente como texto ASCII, lo cual hace posible para nosotros grabar todo el noticiero y más tarde buscar un tema en especial simplemente buscando la palabra de

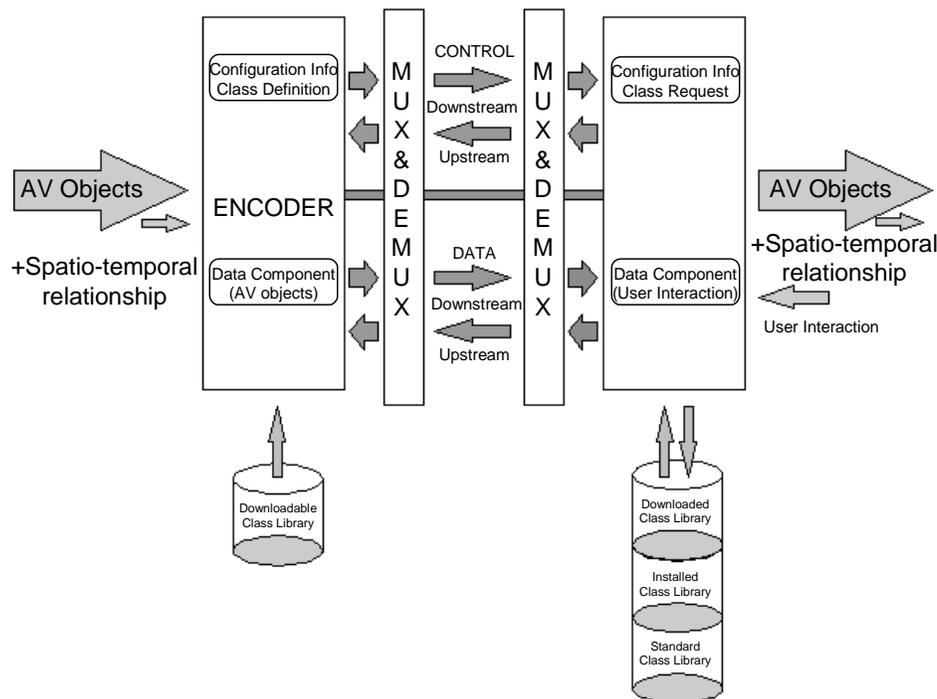


Figura A.1. Arquitectura de un sistema MPEG4

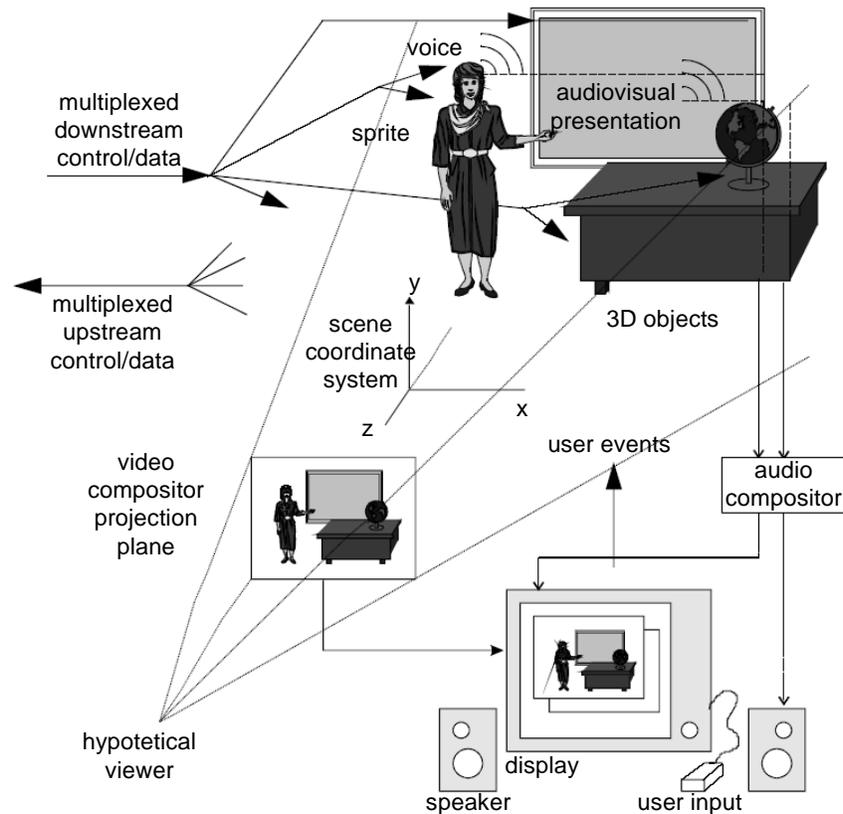


Figura A.2. Ejemplo de una escena MPEG4.

interés para nosotros (dicha información es llamada metadato). La TV podría además mostrar el texto mientras vemos el noticiero, de tal manera que podemos apagar el objeto “audio” y simplemente leer el texto. O también podemos apagar el objeto de video “locutor”, entonces podemos ver únicamente el fondo con el texto y escuchar el audio sin ver la imagen del locutor.

Para reducir la cantidad de datos transmitidos de una manera aún más significativa, el locutor podría ser sustituido por uno sintético, cuyos movimientos faciales estén controlados por las propiedades fonéticas del texto ASCII transmitido. Esto hace innecesaria la transmisión del objeto “locutor” ahorrando con ello ancho de banda. El objetivo que se persigue es que sea prácticamente indistinguible el orador real del orador sintético, lo cual puede lograrse insertando partes de la imagen del orador real. Todo esto sin embargo está todavía en desarrollo y probablemente nunca sea perfecto.

Para tener una mejor impresión de uno de los varios métodos de codificación de MPEG4, la Figura A.3 muestra el principio de animación facial. La cara contiene un conjunto de parámetros de animación. Inicialmente, el objeto rostro muestra una cara genérica con una expresión neutral. Cuando se reciben los parámetros de animación, la cara se anima de acuerdo a estos parámetros con el fin de expresar ciertas emociones o por ejemplo para mover la boca mientras habla.

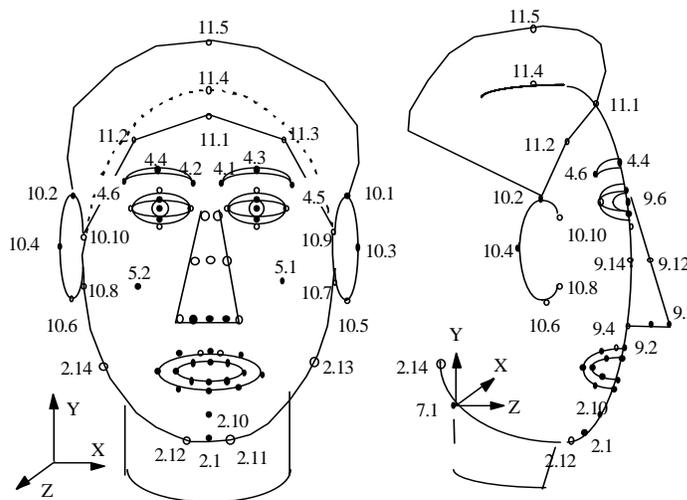


Figura A.3. Principio de animación facial utilizado en MPEG4.

En cuanto a compresión de video, MPEG4 ofrece audio y video a tasas extremadamente bajas (5-64 kbps), el límite máximo para un flujo de datos de video en MPEG4 es de 4 Mbps.

Para este trabajo, únicamente la transmisión de video fue relevante, otros tipos de objetos MPEG4 no fueron de interés debido a que el video es, por mucho, el tipo de tráfico que consume un mayor ancho de banda, y probablemente el objeto más usado del estándar MPEG4.

El método de compresión de video está basado en la estimación y compensación de movimiento, el mismo método que el utilizado en MPEG2. Básicamente, uno de cada n video frames contiene una imagen completa (usualmente $n = 12$), y el espacio entre estas imágenes se rellena con frames que contienen principalmente compensación de movimiento, es decir, vectores que describen el error estimado de bloques de movimiento de las imágenes completas. El estándar define la manera como se calcula esta estimación, de manera que cada codificador y decodificador estiman de manera idéntica, de manera más precisa, podemos decir que se tienen los siguientes tipos de frames:

Frames I (Intra-coded): Un frame I contiene una imagen codificada semejante a JPEG.

Frames P (predicted): Un frame P contiene principalmente errores de compensación para vectores de movimiento (MC Motion Compensation) y pequeños elementos de imágenes de intra código para completar la imagen.

Frames B (bidirectionally predicted): cada frame de tipo B contiene Compensación de Movimiento referido al frame anterior y al siguiente, ya sean I o P.

Estos diferentes tipos de frames se transmiten con cierto orden en el flujo MPEG4. Este orden que llevan los frames es llamado GOP por sus siglas en inglés (Group of Pictures),

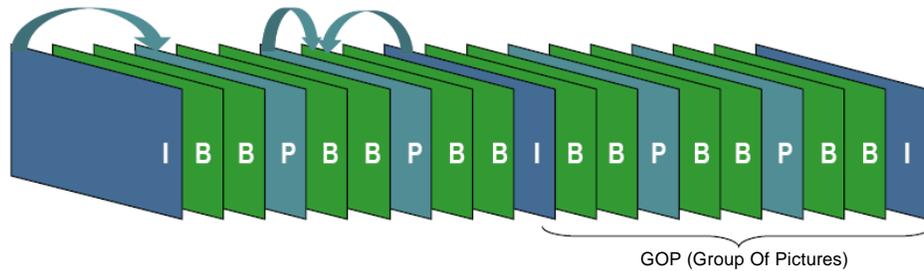


Figura A.4. GOP con n = 9.

el GOP más utilizado tiene 12 frames en el siguiente orden: I-B-B-P-B-B-P-B-B-P-B-B (Figura A.4).

Debido a que los frames B requieren de los frames precedentes y subsecuentes, el orden de los frames en el GOP se modifica para el almacenamiento o transmisión de flujos de video MPEG4. Este reordenamiento se hace de una manera tal, que todos los frames necesarios para la decodificación de otro frame ya están disponibles el momento de la decodificación de ese frame. Ese un ejemplo de ese reordenamiento es mostrado en la Figura A.5.

Trazas de los videos empleados durante las simulaciones.

A continuación se muestran las gráficas de los valores de las muestras de cada imagen empleadas durante las simulaciones. Cada una de las gráficas corresponde a un segmento de video de un minuto de duración en el cual se emplearon 1500 muestras.

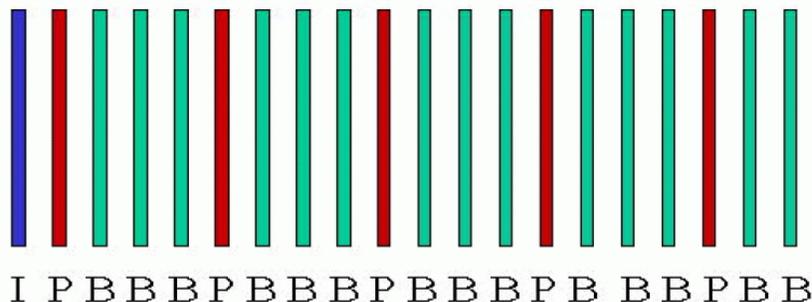


Figura A.5. Ejemplo de reordenamiento de un grupo de imágenes para su transmisión.

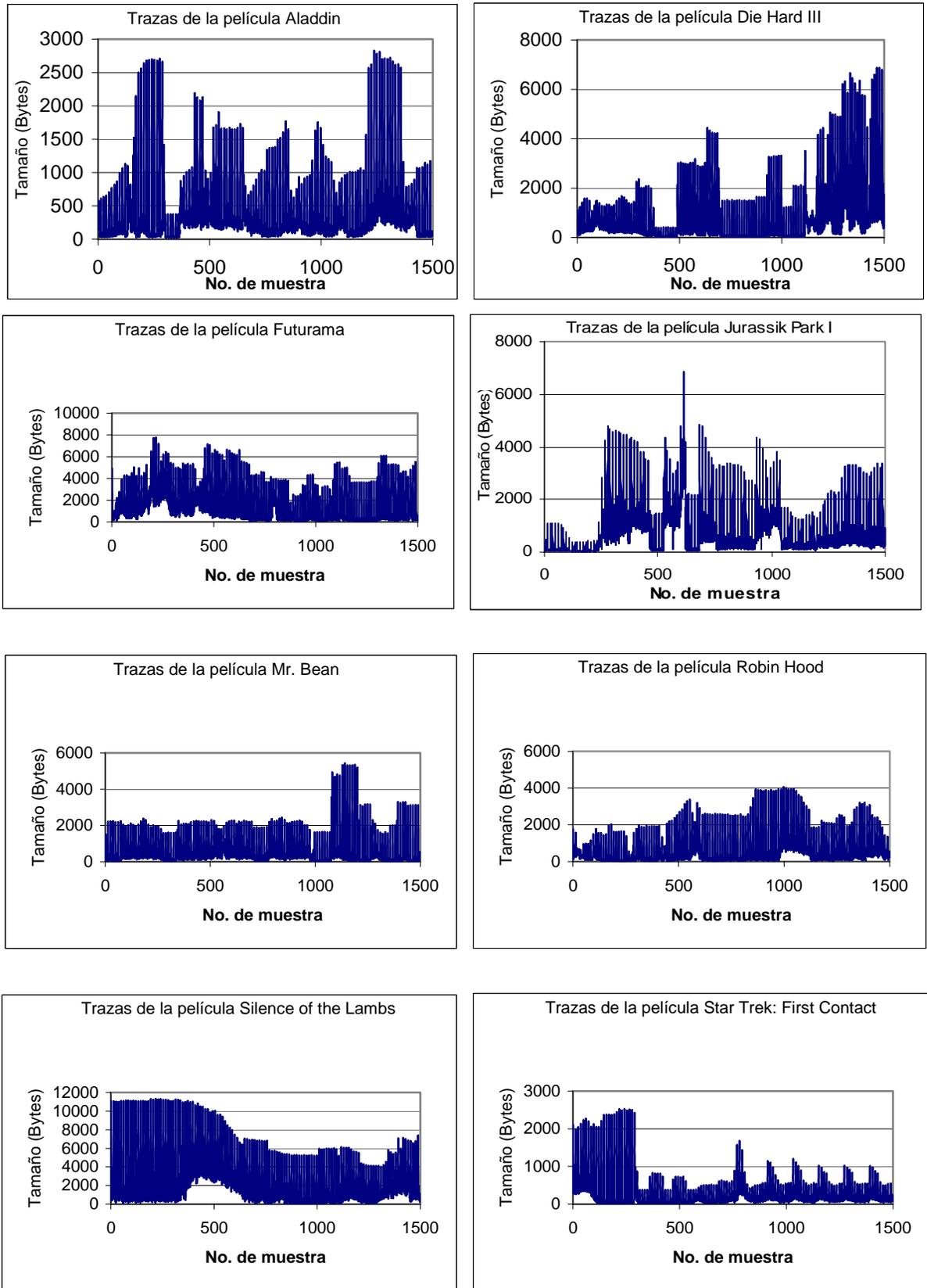


Figura A.6. Trazas de video empleadas durante las simulaciones.

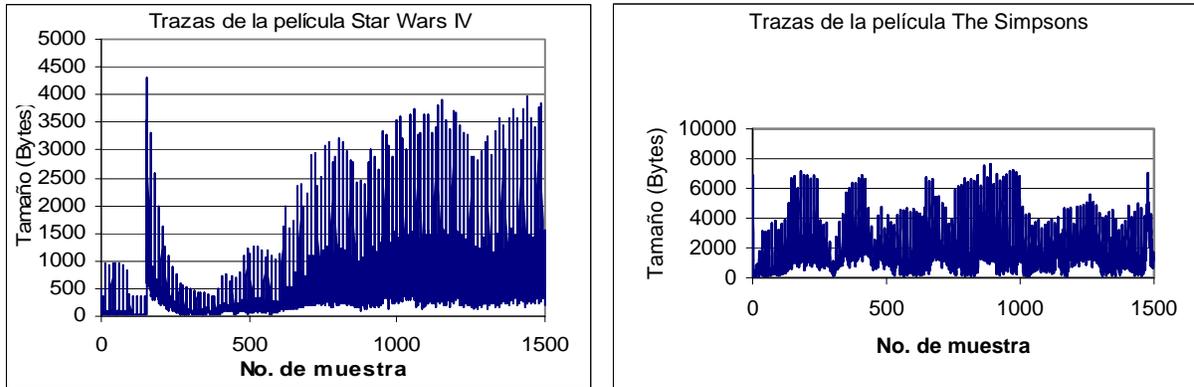


Figura A.6. Trazas de video empleadas durante las simulaciones (continuación).