



ESCUELA SUPERIOR POLITÉCNICA DEL LITORAL

Facultad de Ingeniería en Electricidad y Computación

“Análisis de arquitecturas existentes para sistemas VoD y diseño de Redes de acceso para proveer el servicio de video en demanda y de valor agregado para el Grupo TV Cable”

TESIS DE GRADO

Previo a la obtención del Título de:

INGENIERO EN ELECTRÓNICA Y TELECOMUNICACIONES

Presentada por:

Byron Efrén Añazco Tenicela

Luís Gabriel Guzmán Zambrano

GUAYAQUIL – ECUADOR

Año: 2007

AGRADECIMIENTO

A todas las personas que colaboraron con la realización de este trabajo, especialmente al Ing. César Yépez Flores, Director de Tesis, por su invaluable ayuda y consejos, y al personal del grupo TV Cable por su colaboración e información que nos supieron brindar para la culminación de este proyecto de tesis.

DEDICATORIA

A NUESTRAS FAMILIAS

A NUESTROS AMIGOS

TRIBUNAL DE GRADUACIÓN

Ing. Holger Cevallos U.
SUBDECANO DE LA FIEC
PRESIDENTE

Ing. César Yépez F.
DIRECTOR DE TESIS

Ing. Washington Medina
VOCAL PRINCIPAL

Ing. Rebeca Estrada
VOCAL PRINCIPAL

DECLARACIÓN EXPRESA

“La responsabilidad del contenido de esta Tesis de Grado, nos corresponde exclusivamente; y el patrimonio intelectual de la misma a la ESCUELA SUPERIOR POLITÉCNICA DEL LITORAL”

(Reglamento de Graduación de la ESPOL)

Byron Añezco Tenicela

Luis Guzmán Zambrano

RESUMEN

El presente trabajo describe el estudio para la implementación de un nuevo servicio para el grupo TV-CABLE. El análisis del nuevo servicio a implementar denominado VOD (Video en demanda) está basado en la red HFC existente teniendo en cuenta la compatibilidad de equipos, los anchos de banda requeridos y demás requerimientos técnicos.

En el capítulo 1 se muestra una perspectiva general de los sistemas de video-bajo-demanda. Se presentan los distintos componentes que integran un sistema VoD, describiendo las políticas de gestión más relevantes utilizadas en cada componente. A continuación se mostrarán las arquitecturas utilizadas en el diseño de los sistemas VoD y como éstas se adaptan a un entorno de gran escala.

En el capítulo 2 se explica el funcionamiento de la arquitectura proxy-tree para la distribución del sistema video sobre demanda, se analiza su

confiabilidad, estabilidad y soporte de errores con respecto a otras tecnologías y su uso para el diseño de nuestra red.

En el capítulo 3 se analiza el protocolo RSTP como protocolo de transporte para los paquetes de video así como la codificación MPG-2 que se utiliza para compresión, la autenticación del equipo Terminal con el servidor por medio de los diferentes dispositivos de la red tales como NC, OM, RPD y DAC.

En el capítulo 4 se describe el esquema del headend y la red de acceso HFC con sus elementos desagregados explicando su funcionamiento y su importancia en la red.

En el capítulo 5 se analizará el espectro de frecuencias utilizados por el grupo TV-CABLE para determinar en que canales se podrá incluir el nuevo servicio (video en demanda), se describirá los equipos que se incrementarán tanto en el headend como en la red de acceso tales como servidores de video, los controladores de red y los equipos terminales (DCT 700) así como también se analizarán los aspectos regulatorios necesarios para la prestación de este servicio que se piensa brindar por el Grupo TvCable.

Para finalizar se realizará una proyección de las áreas comerciales para la instalación de los nuevos STB.

En general, esta tesis de grado tiene como objetivo principal cubrir todos los aspectos necesarios para iniciar la prestación del servicio de video en demanda utilizando elementos de red existentes y cumpliendo con los requerimientos técnicos impuestos por la CONARTEL, con lo que se logra un significativo ahorro en tiempo de instalación y en inversión inicial, beneficiando tanto al proveedor como al usuario.

INDICE GENERAL

RESUMEN.....	VI
ÍNDICE GENERAL.....	IX
ABREVIATURAS.....	XIV
ÍNDICE DE FIGURAS.....	XVIII
ÍNDICE DE TABLAS.....	XXI
ÍNDICE DE MAPAS.....	XXII
INTRODUCCIÓN.....	1
CAPÍTULO 1.....	17
1. Sistemas VoD: Una visión general	17
1.1 Introducción a los sistemas de Video-bajo-Demanda.....	18
1.1.1 Tipos de servicios de VoD.....	20
1.1.2 Contenidos multimedia	23
1.1.3 Componentes de un sistema de VoD.....	26
1.1.4 Requisitos de un sistema de VOD.....	32
1.2 Tecnologías de red.....	37
1.2.1 Red de usuarios	38
1.2.2 Red troncal	41
1.3 Arquitecturas utilizadas en los sistemas de VOD	48
1.3.1 Arquitecturas centralizadas	49
1.3.2 Arquitecturas de servidores independientes.....	55

1.3.3	Arquitecturas basadas en servidores-proxy	57
1.3.4	Arquitecturas distribuidas a nivel de los usuarios	62
1.4	Arquitecturas de VOD a gran escala	64
1.4.1	Requisitos de las arquitecturas LVoD.....	64
1.4.2	Alternativas actuales para los sistemas LVoD.....	71
1.4.3	Sistemas comerciales de VoD.....	77
CAPÍTULO 2.....		81
2.	Proxy-Tree, Arquitectura para sistemas LVoD.....	81
2.1	Introducción	82
2.2	Proxy-Tree, una arquitectura escalable.....	83
2.2.1	Análisis de la escalabilidad en los sistemas de VoD	83
2.2.2	Topología en árbol.....	86
2.2.3	Políticas de gestión de los contenidos en los servidores proxy.....	89
2.2.4	Funcionalidad de la arquitectura Proxy-Tree.....	98
2.3	Modelo analítico de las arquitecturas basadas en servidores proxy.....	104
2.4	Análisis de la escalabilidad de Arquitecturas basadas en Servidores Proxy.....	121
2.4.1	Escalabilidad de la arquitectura de un nivel de servidores proxy.....	121

2.4.2	Escalabilidad de la arquitectura P-Tree.....	127
2.5	Evaluación del rendimiento de la arquitectura Proxy-Tree	131
2.5.1	Análisis del orden del árbol	131
2.5.2	Distribución del almacenamiento de los servidores-proxy	132
2.5.3	Análisis de rendimiento de la arquitectura P-Tree.....	133
2.5.4	Requerimientos de ancho de banda de red.....	137
2.5.5	Comparación de resultados entre diferentes arquitecturas LVoD.....	138
CAPÍTULO 3	141
3.	Seguridad de Red.....	141
3.1	Introducción	142
3.2	Protocolo de transporte RSTP.....	142
3.2.1	Propósito del protocolo RSTP	143
3.2.2	Propiedades del Protocolo.	145
3.2.3	Relación con otros protocolos	148
3.2.4	Mensajes RSTP.....	149
3.3	Codificación del Video	150
3.3.1	Clasificación de codificadores	151
3.3.2	MPG2: Moving Pictures Experts Group 2 (MPEG-2).....	152
3.3.2.1	Descripción General	153
3.3.2.2	Propiedades, Características y Funcionamiento.....	154
3.4	Conectividad de equipos	157

3.4.1	VOD Server MDS 200 Servidor de Video	157
3.4.1.1	Descripción General	158
3.4.1.2	Características Técnicas.....	159
3.4.2	Motorola DCT700 (Set top Box) Equipo Terminal	169
3.4.2.1	Descripción General	170
3.4.2.2	Características Técnicas.....	171
3.4.2.3	Autenticación de los equipos terminales STB con el servidor VOD.....	184
CAPÍTULO 4.....		192
4.	Análisis en redes existentes de TV-CABLE	192
4.1	Introducción	193
4.2	Descripción de la red HFC.....	194
4.2.1	Ubicación de Headends: Esquema de equipamiento	196
4.2.2	Ubicación de Nodos Ópticos y características técnicas	200
4.2.3	Repetidores y niveles aceptables para prestación de servicios.....	202
4.2.3.1	TAPs.....	202
4.2.3.2	Nodo óptico.....	203
4.2.3.3	Patch panel óptico	203
4.2.3.4	Amplificadores ópticos	204
4.2.3.5	Chasis de retorno.....	204
CAPÍTULO 5.....		207

5. Regulación, diseño, planificación y expansión del sistema Ivod en la red.....	207
5.1 Parámetros regualtorios a considerar previo a la instalación de servicios de vod para las ciudades de Guayaquil, Quito y Cuenca.	208
5.1.1 Solicitud a presentar a la conartel para la aprobación de cambios en el headend.....	209
5.2 Servicios prestados por el Grupo TVCABLE y distribución de frecuencias utilizadas	210
5.3 Posible asignación de canales en el espectro para poder prestar el servicio VOD	214
5.4 Proyección de la demanda en 5 años.....	216
5.5 Análisis de áreas comerciales, en donde se distribuiría los nuevos equipos terminales DCT 700	226
CONCLUSIONES.....	222
RECOMENDACIONES.....	225
ANEXO A.....	228
ANEXO B.....	230
ANEXO C.....	232
ANEXO D.....	233
ANEXO E.....	234
ANEXO F.....	235
BIBLIOGRAFÍA.....	236

ABREVIATURAS

ADSL	Asymmetrical Digital Subscriber Loop
AP	Access Point
ASCII	American Standard Code for Information Interchange
ASK	Amplitude-shift keying
ATD	Asynchronous time division
ATM	Asynchronous Transfer Mode
CATV	Cable Television
CBR	Constant Bit Rate
CDMA	Code Division Multiple Access
CMTS	Cable Modem Terminal System
Coax	Coaxial
CONATEL	Consejo Nacional de Telecomunicaciones
Cu	Cobre
DAC	Digital Address Controller
DCT	Marca de Equipo Motorola
DCT	Discrete Cosine Transform
DivX	Formato de Video Basado en MPGE-4
DMA	Time Division Multiple Access
DVD	Digital Versatile Disc

ECC	Error Checking and Correction
ECC	Error Correction Code
FEC	Forward Error Correction
FO	Fibra Óptica
GHz	Gigahertz
GOP	Group Of Pictures
HDSL	High-Speed DSL
HDTV	High Definition Television
HFC	Híbrido Fiber Coaxial
HTTP	HyperText Transfer Protocol
IP	Internet Protocol
ISDN	Integrated Services Digital Network
ISO	Organización mundial de estándares
ISP	Internet Service Provider
ITU	International Telecommunications Union
LAN	Red de área local
LUN	Logical unit number
MAC	Media Access Control
Mbps	Megabits per second
MDF	Multipoint Distribution Frame
MHz	Megahertz
MPGE	Moving Picture Experts Group

MTA	Media Terminal Adapter
NC	Network Controller
NTSC	National Television System Committee
NVOD	Near VOD
OM	Out-of-Band Modulator
OSI	Organización Internacional para la Estandarización
PC	Personal Computer
PPV	Pay per View
PSK	Phase Shift Keying
PTree	Proxy tree
QAM	Quadrature Amplitude Modulation
QOS	Calidad de servicio
QVOD	Quasi video-bajo-demanda
RAB	RAID Advisory Board RAB
RAID	Redundant Array of Independent
RF	Radiofrecuencia
RPD	Return Path Decoder
RSTP	Real Streaming Time Protocol
RTCP	Real-Time Control Protocol
RTP	Real-Time Transport Protocol
RTSP	Real-Time Streaming Protocol
SDH	Synchronous Digital Hierarchy

SEM	SmartStream Encryptor Modulator
SIP	Session Initiation Protocol
STB	Set-Top-Box
STP	Signaling Transfer Point
TCP	Transmission Control Protocol
TDM	Time Division Multiplexing
TVOD	True VOD
UDP	User Datagram Protocol
UPS	Uninterruptible Power Supply
UTF	Unicode Transformation Format
VBR	Variable Bit Rate
VCR	Video Cassette Recorder
VLC	Video LAN Client
VOD	Video on Demand -Video Bajo Demanda
VoIP	Voice over Internet protocol
WAN	Red de area Amplia
WLL	Wireless Local Loop
xDSL	x Digital Subscriber

INDICE DE FIGURAS

Figura 1.1 Principales componentes de un sistema VOD	26
Figura 1.2 Componentes para una instalación VOD en una red ADSL	39
Figura 1.3 Componentes de una red HFC para prestar el servicio de VOD .	40
Figura 1.4 Protocolos para soportar tráfico multimedia sobre TCP/IP	46
Figura 1.5 Arquitectura de VOD centralizada	49
Figura 1.6 Arquitectura VOD centralizada con Servidores Paralelos.....	52
Figura 1.7 Arquitectura Centralizada basada en cluster	54
Figura 1.8 Arquitectura VOD con servidores independientes	56
Figura 1.9 Arquitectura VOD de Servidores Proxy de un nivel	60
Figura 1.10 Arquitectura de VOD Jerárquica de servidores-proxy	62
Figura 1.11 Escalabilidad de la topología Ncube	79
Figura 2.1 Arquitectura P-Tree para un sistema LVOD	87
Figura 2.2 Sistema Jerárquico de catching.....	90
Figura 2.3 Gestión de una petición rechazada en sistema jerárquico.....	92
Figura 2.4 Esquema de almacenamiento Caching y Mirroring	96
Figura 2.5 Diagrama de la gestión de los servidores-proxy	100
Figura 2.6 Distribución de Zipf de popularidad de videos	109
Figura 2.8 Saturación del Sistema Jerárquico Servidores-Proxy	115
Figura 2.9 Esquema de distribución jerárquica de los fallos en los proxys .	118
Figura 2.10. Escalabilidad en los sistemas Proxy	122
Figura 2.11. Escalabilidad de la red principal en la arquitectura P-Tree	129

Figura 2.12 Rendimiento de las arquitecturas de LVoD.....	134
Figura 2.13 Distribución de los servidores a nivel Nacional.....	137
Figura 3.1 Servidor MDS 200.....	158
Figura 3.2 DCT 700 parte frontal	171
Figura 3.3 Esquema de Bloques del equipo Terminal DCT 700	172
Figura 3.4 Parte posterior del Equipo Terminal	173
Figura 3.5 Como conectar TV-DCT700 Cable coaxial 75 Ω	173
Figura 3.6 Conexión de Audio TV-DCT700	174
Figura 3.7 Conexión DCT700-VCR-TV cable coaxial de 75 Ω	174
Figura 3.8 Conexión DCT700-VCR-TV Audio y Video.....	175
Figura 3.9 Constelación para sistema 16-QAM	179
Figura 3.10 Diagrama de bloques del SEM	181
Figura 3.10 Esquema de funcionamiento del DAC6000	186
Figura 3.11 Parte Frontal del NC 1500	188
Figura 3.12 Esquema de Infraestructura para VOD 1.....	189
Figura 3.13 Esquema de Infraestructura para VOD 2.....	189
Figura 3.14 OM 1000 Out-of-Band Modulator.....	191
Figura 3.15 RPD 2000 Retutn Path Decoder.....	191
Figura 4.1 Diagrama esquemático de la topología HFC	193
Figura 4.2 Conversión de Redes	194
Figura 4.3 Esquema General de la Infraestructura del servicio de video....	197
Figura 4.4 Esquema General de la Infraestructura del servicio de telefonía	198

Figura 4.5 Esquema de interconexión de los diferentes servicios	199
Figura 4.6 Esquema básico de nodos.....	201
Figura 4.7 Red de Acceso a Clientes.....	202
Figura 5.8 TAP 4-1.....	203
Figura 4.9 Amplificador Óptico.....	204
Figura 4.10 Esquema de redes de Acceso y niveles de señal.....	205
Figura 5.1 Distribución de los canales en el espectro.....	215
Figura 5.2 Cuadro comparativo de ingresos a nivel mundial	217
Figura 5.3 Crecimiento de usuarios en Guayaquil	218
Figura 5.4 Crecimiento de usuarios en Quito.....	219
Figura 5.5 Crecimiento de usuarios en Cuenca	219
Figura 5.6 Crecimiento de las 3 ciudades a implementar VOD	220
Figura 5.7 Crecimiento porcentual de las ciudades	220

INDICE DE TABLAS

Tabla 1.1 Formatos de Video y sus requisitos de Ancho de Banda.....	25
Tabla 1.2 Número de flujos de Video soportados por las principales tecnologías de red	42
Tabla 1.3 Principales características de sistemas LVOD.....	76
Tabla 2.1 Características ideales para la arquitectura LVoD	82
Tabla 2.2 Ancho de Banda Servidor-Proxy.....	116
Tabla 2.3 Probabilidades de Acierto y Fallo.....	123
Tabla 2.4 Ancho de Banda de Fallo y Acierto.....	123
Tabla 2.5 Ancho de Banda de la Red Principal	124
Tabla 2.6 Ancho de Banda de Fallo y Acierto.....	129
Tabla 2.7 Ancho de Banda de la Red Principal	130
Tabla 2.8. Rendimiento sistemas P-Tree con diferentes topologías	132
Tabla 2.9 Principales de las distintas arquitecturas LVOD	139
Tabla 4.1 Distribución en el espectro de la red HFC.....	195
Tabla 5.1 Distribución en frecuencia de los distintos servicios actuales	212
Tabla 5.2 Distribución de los 3 SEM's actuales	213

INDICE DE MAPAS

Mapa 5.1 Quito Norte.....	228
Mapa 5.2 Quito Centro.....	229
Mapa 5.3 Quito Sur.....	230
Mapa 5.4 Guayaquil Norte	231
Mapa 5.5 Guayaquil Centro	231
Mapa 5.6 Guayaquil Sur	232
Mapa 5.7 Cuenca Este	233
Mapa 5.8 Cuenca Oeste	234

INTRODUCCION

El presente trabajo tiene la finalidad de describir las diferentes arquitecturas actualmente utilizadas para la prestación del servicio de Video en demanda decidiendo cual arquitectura es la más conveniente para que permita una mayor escalabilidad, confiabilidad, aprovechando todos los recursos de la red HFC existente manejada por TVCABLE.

Con la apertura del mercado de telecomunicaciones en el 2002 El Grupo TVcable como empresa ha tratado de ser la pionera en la prestación de nuevos servicios para cubrir todas las necesidades de los clientes enfocándose a dar servicios triple play. Triple play es una integración de servicios (voz, video y datos) en un mismo medio físico, el grupo tvcable presta internet dependiendo del ancho de banda requerido por el cliente, en el caso de la telefonía también tiene la concesión para prestar servicios a través de voz sobre IP y en el caso de video se ofrecen diferentes paquetes dependiendo de las necesidades y gustos del usuario. Un servicio adicional de video que se presta es PPV (pay per view) que ha tenido una gran acogida y es gracias a ello que hemos decidido realizar el análisis para la ampliación de un servicio extra que es video en demanda que permite una mayor interacción entre el usuario y el contenido multimedia.

Teniendo en cuenta la red actual HFC hemos analizado que equipos serian necesarios para poder prestar el servicio de video en demanda teniendo como objetivo principal la compatibilidad de los mismos, así como el ancho de banda requerido y la incorporación de equipos en el headend para que este servicio se pueda prestar.

Algo importante ha recalcar son los permisos necesarios antes de la prestación del servicio, para ello hay que realizar un análisis de la concesión actual y los requerimientos establecidos por la CONARTEL para poner en marcha el proyecto.

Para finalizar se hace una proyección de la penetración del servicio en el mercado actual para estimar el índice de crecimiento de los usuarios y los requerimientos a futuro para mantener el servicio operativo que cada vez permita a los usuarios tener mayores opciones el momento de elegir su contenido multimedia.

Capítulo 1

1. Sistemas VoD: Una visión general

Este capítulo muestra una perspectiva general de los sistemas de video-bajo-demanda. Se presentan los distintos componentes que integran un sistema VoD, describiendo las políticas de gestión más relevantes utilizadas en cada componente. A continuación se mostrarán las arquitecturas utilizadas en el diseño de los sistemas VoD y como éstas se adaptan a un entorno de gran escala.

1.1 INTRODUCCIÓN A LOS SISTEMAS DE VIDEO-BAJO-DEMANDA

El termino Video bajo Demanda (VOD) hace referencia a servicios en los cuales los usuarios son capaces de pedir contenidos multimedia (videos) en cualquier instante de tiempo. Esta tecnología es de vital importancia para diversas aplicaciones multimedia como por ejemplo, aprendizaje a distancia, bibliotecas digitales, videoconferencias, Internet, televisión ó sistemas de video bajo demanda.

En los últimos años, los sistemas de video bajo demanda han sido una de las áreas más activas en la investigación debido a la convergencia de dos factores: el creciente interés de la industria de diversos sectores en desarrollar estos sistemas y su elevada complejidad de diseño e implementación.

Gracias a la reducción de costos de los componentes que integran un sistema de VOD y los avances de la tecnología, los servicios de VOD han alcanzado la madurez necesaria de forma que su implementación y comercialización ya son viables. Esta nueva tecnología ha provocado una revolución en la industria de entretenimiento, atrayendo el interés de las

empresas de cable, como el Grupo TV CABLE en Ecuador, desearon de aumentar su oferta mediante servicios de valor agregado como telefonía, internet y servicios de video pagado como Pay per View.

Respecto a la investigación, los servicios de VOD y su implementación han aportado nuevos retos a la comunidad científica. El diseño de estos sistemas involucra diferentes áreas: psicología (estudio del comportamiento de los usuarios), sistemas de tiempo real, sistemas de archivo de altas prestaciones, calidad de servicio, protocolos de comunicaciones, formatos de compresión, criptografía, sistemas de procesamiento jerárquicos, paralelos ó distribuidos y demás factores que afectan o tienen importancia en este tema.

En este capítulo vamos a realizar una breve visión de los sistemas de video bajo demanda, haciendo hincapié en las características de la información gestionadas (contenidos multimedia), los tipos de servicios ofrecidos a los usuarios de estos sistemas, sus componentes principales y los requisitos específicos de las arquitecturas de VOD.

1.1.1 TIPOS DE SERVICIOS DE VOD

Los sistemas de VOD se pueden clasificar en función del tipo de servicio que ofrecen a los usuarios. La principal característica que distingue el servicio de VOD de otras tecnologías parecidas (como la televisión) es la capacidad de interacción y elección de los usuarios a la hora de escoger qué contenido y cuándo lo quiere reproducir.

Teniendo en cuenta este parámetro los posibles servicios que puede ofrecer un sistema de VOD son: Pay per view (PPV), quasi video-bajo-demanda (QVOD), VOD aproximado (NVOD) y VOD verdadero (TVOD).

- ***Pago por visión***

El servicio de pago por visión (Pay-per-View, PPV), permite al usuario reservar y pagar por programas específicos; este esquema en la actualidad ya se presta por el Grupo TvCable y tiene una gran acogida ante el mercado.

- ***Quasi video bajo demanda***

En el quasi VOD (Quasi-VOD, QVOD), los videos son enviados por el operador a los usuarios a través de la red, solo en el caso de que el número de suscriptores que los han solicitado con anticipación sea lo suficientemente grande. Los usuarios están agrupados por categorías de interés, basándose en políticas de optimización de recursos del sistema. Los usuarios no tienen el control interactivo sobre un canal específico, solo pueden cambiarse de un grupo a otro.

- ***Video bajo demanda aproximado***

En el VOD aproximado (Near-VOD, NVOD), el proveedor trasmite un determinado contenido en intervalos de tiempo regulares (cada 15 minutos por ejemplo). En el momento que un usuario realiza una petición en el sistema, ésta es atendida por el siguiente canal que vaya a transmitir el contenido deseado. Por lo tanto, es posible que la petición no se atienda inmediatamente y que el usuario deba esperar durante un intervalo de tiempo.

- ***Video bajo demanda verdadero***

El servicio de video bajo demanda verdadero (True-VOD, TVOD) es el más completo. El usuario tiene el control total sobre cuándo quiere

visualizar el video y sin estar sujeto a ningún tipo de restricciones por parte del operador. Para seleccionar el contenido a reproducir, el usuario puede elegir cualquiera de los contenidos multimedia disponibles en el catálogo del sistema de VOD.

Los sistemas TVOD suelen soportar todos los comandos disponibles en un VCR (Video Cassette Recorder): reproducir / reiniciar (que permite iniciar o reiniciar la visualización del video), parar (permite parar la visualización del video), pausa (congela la visualización del video), avance y retraso rápido (permite adelantar ó atrasar la visualización del video utilizando una velocidad mayor a la normal), cámara lenta (permite visualizar el video a una velocidad inferior a la normal) y búsqueda (permite buscar una determinada secuencia mediante saltos dentro del video y sin mostrar imagen ó sonido).

El tipo de servicio ofrecido es un parámetro importante en el diseño, ya que a medida que se aumenta la interactividad del usuario también se incrementa la complejidad del sistema de VOD y por lo tanto, el valor agregado del servicio ofrecido a los usuarios.

1.1.2 CONTENIDOS MULTIMEDIA

La mayoría de la funcionalidad específica de los sistemas de VOD deriva de las características particulares del tipo de información (contenidos multimedia) gestionada por estos sistemas.

A diferencia de los tipos de datos tradicionales, los contenidos multimedia tienen una dimensión temporal explícita, y entonces deben ser presentados mediante una frecuencia específica durante un tiempo determinado ó de lo contrario la integridad de la información se perderá.

De todos los contenidos multimedia, el más significativo por sus requisitos y características es el video. Un video consiste en una secuencia de imágenes que son visualizadas a una frecuencia preestablecida, que normalmente suele ser alrededor de 30 imágenes por segundo.

Los contenidos multimedia tienen una naturaleza analógica y para que esta información pueda ser gestionada y almacenada en un ordenador debe ser digitalizada. Sin embargo, su digitalización genera un volumen de

información demasiado grande para ser almacenada ó transmitida eficientemente por la red.

Para reducir los requisitos de los videos, éstos se codifican guardando solo la información correspondiente a los píxeles ó líneas de información consecutivas que son diferentes. Las técnicas de codificación y compresión explotan las redundancias espaciales y temporales del video, las cuales pueden variar de una escena a otra. Por lo tanto, puede ocurrir que las frecuencias de compresión de dichas escenas sean diferentes entre si, provocando con ello diferentes requisitos de ancho de banda entre las distintas partes del video. Este tipo de codificación se denomina VBR (Variable Bit Rate) y puede complicar considerablemente la implementación de los servicios VOD.

Existen codificaciones alternativas que adaptan la calidad del video entre escenas para conseguir una misma frecuencia de compresión a lo largo del video, esta técnica se conoce como CBR (Constant Bit Rate).

La tabla 1.1 muestra los principales formatos de video, junto a sus características más destacadas (canales de video / audio y resolución de la

imagen) y los requisitos desde el punto de vista del servidor / red (ancho de banda).

Se puede constatar que los requisitos exigidos para este tipo de información son muy elevados independientemente del formato utilizado. El almacenamiento y transmisión al cliente de un video de calidad media (2 horas de MPEG-2) precisa una capacidad de almacenamiento de 7.2 GBytes y una infraestructura de red capaz de transmitir hasta 8 Mbits por segundo de información durante dos horas y sin sufrir ningún corte ó retardo.

Tabla 1.1 Formatos de Video y sus requisitos de Ancho de Banda

Formato de Video	Canales	Resolución en pixel	Ancho de Banda
MPGE-4	1 Video 2 Audio	720 x 480	< 1 Mb/s
MPGE-2 HDTV	1 Video	1920 x 1080	19.4 Mb/s
MPGE-2 y Dolby Digital	1 Video 5.1 Audio	720 x 480	6-8 Mb/s
MPGE-2 y Dolby Digital	5 Video 10.2 Audio	5x (720 x 480)	4x5 Mbs/s 11 Mb/s

1.1.3 COMPONENTES DE UN SISTEMA DE VOD

Los sistema de VOD están compuestos por tres componentes básicos: el servidor, la red de transmisión y los usuarios del sistema (ver figura 1.1). A continuación describiremos la funcionalidad de cada uno de estos componentes.

Servidor

El servidor de video almacena los contenidos que pueden ser solicitados por los usuarios. Es el encargado de gestionar el servicio a los clientes, garantizando una cierta calidad de servicio a lo largo del camino que tiene que seguir la información desde el disco hasta los usuarios.

Un servidor de VOD está compuesto por tres subsistemas: El subsistema de control, el subsistema de almacenamiento y el subsistema de comunicación.

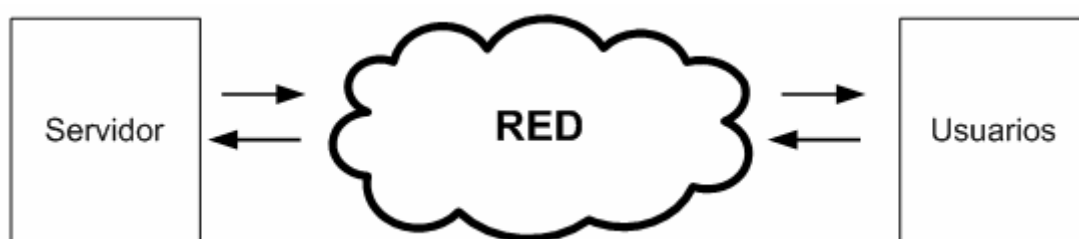


Figura 1.1 Principales componentes de un sistema VOD

- ***Subsistema de control***

El subsistema de control es el encargado de recibir las peticiones de los usuarios y ordenar las acciones que se tienen que llevar a cabo para poder atenderlas.

Este módulo debe decidir si la nueva petición puede ser servida por el sistema sin que ello implique un deterioro de las peticiones activas. Estas decisiones son tomadas por la política de control de admisión en función de los recursos disponibles en el sistema y de los requisitos de la nueva petición.

Otras funciones del módulo de control son la gestión de las estadísticas de utilización del sistema (contabilidad y facturación) y realización de tareas de optimización para incrementar la eficiencia del sistema.

- ***El subsistema de almacenamiento***

Este módulo es el responsable de almacenar y recuperar la información multimedia desde los dispositivos de almacenamiento. Las

principales dificultades a la hora de conseguir este objetivo estriban en el volumen de información que se debe gestionar y que ésta debe ser entregada de acuerdo a las estrictas especificaciones de la calidad de servicio (QOS) requeridas por las aplicaciones de video bajo demanda.

- ***El subsistema de entrega de comunicación***

Es el encargado de planificar la inyección de los contenidos multimedia en la red de transmisión. Este módulo se encarga de gestionar las distintas políticas de servicio que permiten optimizar los recursos de ancho de banda de la red y del servidor.

Red de comunicación

Uno de los principales factores que más han influenciado en el crecimiento de las aplicaciones multimedia es el crecimiento de la red de interconexión. Para permitir a los usuarios acceder a los contenidos multimedia, las redes deben satisfacer al menos dos requisitos: Disponer de mecanismos de transporte para enviar las peticiones y los datos y permitir que la información sea transmitida respetando unos niveles mínimos de rendimiento (calidad de servicio QOS).

La red de comunicación de un sistema de VOD se caracteriza por unos elevados requisitos de ancho de banda (capacidad de transferencia de grandes volúmenes de datos) y grandes velocidades de transmisión.

En un sistema VOD, podemos llegar a encontrar tres niveles de red diferentes: la red principal, la red troncal y las redes locales. Ahora bien, dependiendo de la arquitectura del sistema VOD finalmente utilizada, estos niveles se pueden integrar entre sí, quedando reducidos a únicamente dos niveles la red principal y red troncal.

La red principal es aquella a la cual se conectan los servidores de VOD y sirve punto de conexión de éstos con la red de distribución (red troncal) de los contenidos multimedia a los usuarios. La red troncal (ó *backbone*) permite interconectar la red principal con cada una de las redes de distribución locales (en caso de que éstas existan) ó bien directamente con los usuarios.

Su objetivo es transportar, tan rápido como sea posible, la información generada por los servidores desde la red principal a los usuarios. Las redes locales son las responsables de la conexión final de los usuarios al sistema de VOD. Esta red requiere un ancho de banda inferior respecto a los

otros niveles. El tráfico soportado por las redes de usuario tiene una naturaleza asimétrica, lo cual significa que se necesita un ancho de banda de entrada considerablemente mayor al tráfico de salida.

Clientes

Los usuarios deben soportar la recepción y la visualización sin cortes de los contenidos multimedia, así como soportar los comandos VCR. La interfase entre los usuarios y el sistema de VOD se realiza mediante el STB (Set-Top-Box en este caso DTC-700). Este módulo es el encargado de recibir los comandos del usuario y enviar la señal al servidor a través de la interfase de red.

El STB almacena los contenidos recibidos desde el servidor en unos buffers locales, decodifica los contenidos recibidos en tiempo real y envía las imágenes obtenidas a la pantalla de visualización, con la temporización correcta.

En general los STB constan de 4 componentes principales: Interfase de red, decodificador, buffer y hardware de sincronización.

- ***Interfase de red***

Permite al cliente recibir y enviar información desde ó hacia los servidores.

- ***Decodificador***

Para reducir los requisitos de almacenamiento, ancho de banda de disco y ancho de banda de red, los contenidos multimedia suelen estar codificados. Así se necesita un decodificador en el lugar del cliente para decodificar el video antes de ser presentado al usuario.

- ***Buffer***

Debido a los retrasos introducidos por la red, el tiempo de llegada de la información (video) no puede ser determinado con exactitud. Para conseguir una reproducción sin cortes, el servidor debe garantizar que la siguiente porción del video que se va a visualizar esté disponible antes que el usuario la quiera visualizar. Para lograr este objetivo, el servidor envía datos al usuario en adelanto de forma que se asegure un margen de maniobra que amortigüe los posibles retardos inesperados introducidos por la red de comunicación. Como el usuario

no va a consumir inmediatamente estos datos, estos se tienen que almacenar temporalmente en un buffer hasta que sean requeridos.

- ***Hardware de sincronización***

Los videos están compuestos por un stream de video y un stream de audio independientes. Para poder realizar un reproducción correcta, ambos tipos de información deben ser sincronizados entre si antes de que puedan ser reproducidos.

1.1.4 REQUISITOS DE UN SISTEMA DE VOD

La funcionalidad requerida de un sistema de VOD así como las características de la información gestionada por éstos, imposibilita la utilización de servidores genéricos. Por lo tanto, los servidores de VOD deben ser diseñados teniendo en cuenta una serie de requisitos específicos del tipo de información gestionada. El servicio de una petición para un contenido multimedia requiere un elevado volumen de información, con requerimientos de tiempo real, mantenimiento de la calidad de servicio (QOS) y grandes anchos de banda de transferencia del sistema de almacenamiento y la red de comunicaciones. El conjunto de todos estos requisitos complica el diseño e implementación de los sistemas de VOD y limita considerablemente

el número de usuarios que puede soportar un servidor de VOD. A continuación se describen brevemente cada uno de estos requisitos.

Gran capacidad de almacenamiento

Dada la naturaleza intensiva en almacenamiento de la información multimedia, los requisitos de almacenamiento globales de cientos de contenidos multimedia puede exceder fácilmente un requisito de disco de decenas de Terabytes.

Por ejemplo, un video en formato de televisión de alta definición (HDTV) de dos horas de duración puede requerir hasta 18 Gigabytes (ver tabla 1.1, en el apartado 1.1.2). Por lo tanto, un sistema de VOD compuesto de 200 videos puede requerir aproximadamente unos 3.6 Terabytes de almacenamiento.

Servicio en tiempo real

Para garantizar la reproducción continua de los contenidos multimedia, no es suficiente con que servidor de VOD envíe los datos al usuario y éste los reciba correctamente; sino que esta recepción se debe producir dentro un intervalo de tiempo específico.

Esto implica que todos los componentes del sistema deben tener un control del tiempo máximo permitido para poder realizar cada uno de las operaciones que intervienen en la entrega de información a los usuarios. Además, los distintos componentes que intervienen en el sistema se tienen que sincronizar entre sí para no violar estos requisitos de tiempo. Si esta sincronización no se lleva a cabo es imposible garantizar una calidad de servicio al usuario final.

Es posible suavizar los requisitos en tiempo real de los sistemas de VOD mediante la utilización de buffers intermedios tanto en el servidor como en el cliente y el envío en adelanto de un fragmento del contenido multimedia.

Calidad de servicio (QOS)

Un aspecto clave en cualquier servicio de vídeo es proporcionar una calidad de servicio (QOS) aceptable al usuario. Debido a la naturaleza continua e independiente del tiempo de los contenidos de audio y video, su reproducción requiere un estricto control del momento y la secuencia de recepción de la información por parte del usuario.

Esta calidad de servicio generalmente implica varios aspectos tales como: calidad de la imagen, frecuencia de pérdida de imágenes, sincronización audio y vídeo, entre otros. Algunos de estos parámetros no son fácilmente cuantificables porque dependen de la percepción subjetiva del observador.

La calidad de servicio a nivel del usuario refleja como se suministra el flujo de vídeo original desde un servidor de vídeo remoto. Estos servicios requieren restricciones específicas en el flujo de información desde el servidor al cliente. Por lo tanto, una cuestión importante en VOD es como lograr una correspondencia entre la QOS específica requerida por el cliente con la especificación de una QOS para el servidor de vídeo y la red de transmisión.

Con el objetivo de conseguir unas prestaciones en el sistema que garanticen una QOS aceptable se requiere una fuerte coordinación entre todos los componentes del sistema, desde los servidores a los dispositivos de visualización pasando por las componentes de red. No es suficiente un análisis individual de los componentes del sistema sino que se requiere un diseño unificado que tenga en cuenta todos los componentes. La QOS basada en el análisis de los componentes individuales tropieza con el problema de que los componentes no son independientes entre si. La

solución óptima para una componente no garantiza la mejor solución para todo el sistema y por lo tanto, se requiere un análisis integrado.

Grandes anchos de banda

Los contenidos multimedia requieren el procesamiento de un gran volumen de información de forma periódica y durante grandes periodos de tiempo. Este volumen de información exige grandes anchos de banda en la red de transmisión.

Los requisitos de ancho de banda no se circunscriben exclusivamente a la red de comunicaciones entre el servidor de VOD y los usuarios finales, sino que también involucran al sistema de almacenamiento. Esto implica la utilización de sistemas de almacenamiento complejos basados en sistemas de almacenamiento jerárquicos ó bien la utilización de un conjunto de discos en configuración RAID (la explicación de RAID se la hará mas adelante).

Es importante hacer notar que si no se tiene en cuenta este parámetro en el diseño del sistema de VOD, un incremento en el número de peticiones a gestionar por el sistema, puede aumentar los requisitos de ancho de banda hasta llegar a saturar el sistema de VOD.

1.2 TECNOLOGÍAS DE RED

Dado que un sistema de VOD requiere la transferencia de enormes volúmenes de datos a muy altas velocidades, varios protocolos de comunicación y arquitecturas de red han sido propuestos para conectar los distintos componentes del sistema.

Las tecnologías utilizadas varían considerablemente según el nivel de red que se considere: red principal, red troncal ó red de conexión final con los clientes. Mientras que la red principal requiere grandes ancho de banda, la conexión con los clientes finales tiene requisitos individuales más pequeños (con un ancho de banda de red de 1.5 Mb/s es suficiente para soportar la transmisión de un video en formato MPEG-2).

El criterio más importante para la selección de la tecnología de la red principal es el ancho de banda y el soporte de la gestión de la QOS. En este caso, ATM emerge como la tecnología más importante. Otra alternativa, que permite reutilizar la infraestructura de Internet actual, se basa en la utilización de protocolos específicos (RTP, RTCP, RTSP, RSVP, etc) para soportar la gestión de la calidad de servicio por encima del protocolo TCP/IP sobre redes Ethernet.

A continuación, se describen las distintas alternativas utilizadas en la conexión de los usuarios y en la red troncal y principal.

1.2.1 RED DE USUARIOS

La infraestructura de comunicaciones entre el usuario y la red principal del sistema de VOD se denomina red de los usuarios. Esta red sirve de lazo de unión entre el servidor de VOD y el STB de usuario.

Las principales tecnologías que se utilizan en la conexión de los usuarios a la red troncal de los sistemas de VOD son ADSL y la fibra óptica/cable (red HFC).

- ***Asymmetrical Digital Subscriber Loop (ADSL)***

ADSL se basa en la utilización de redes de cables de cobre. Permite la recepción de datos a altas velocidades, utilizando la infraestructura telefónica actual y con pocas distorsiones.

La instalación de ADSL se compone de un par de unidades (ver figura 1.2), una instalada en el cliente y la otra en la oficina central telefónica. ADSL puede proporcionar al usuario final un ancho de banda de entrada de más 1.5 Mb/s y un ancho de banda de salida de 16 Kb/s para control.

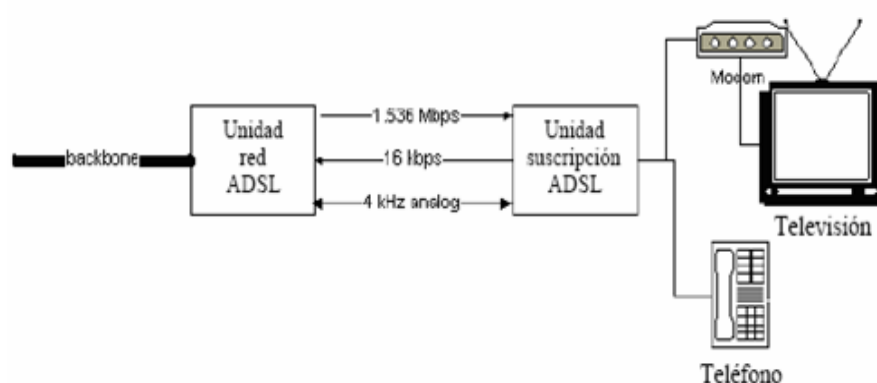


Figura 1.2 Componentes para una instalación VOD en una red ADSL

Extensiones de ADSL (HDSL) pueden alcanzar un mayor ancho de banda, llegando a los 6 Mb/s. Estas características son suficientes para satisfacer los requisitos de ancho de banda y comunicación bidireccional de los servicios de VOD.

- **Cable**

La distribución de información a través del cable se basa en la utilización de la tecnología HCF (Hybrid Fiber Coaxial) que combina el

uso de fibra óptica junto a cables coaxiales. Estas redes de cable han sido utilizadas tradicionalmente para la transmisión de señales analógicas por las compañías de cable de televisión como TVCABLE, pero mediante la utilización de módems también permiten la transmisión de señales digitales.

La topología utilizada para estos sistemas se basa en árboles en fibra, con ramales de cable coaxial, a las cuales se conectan los subscriptores, según muestra la figura 1.3.

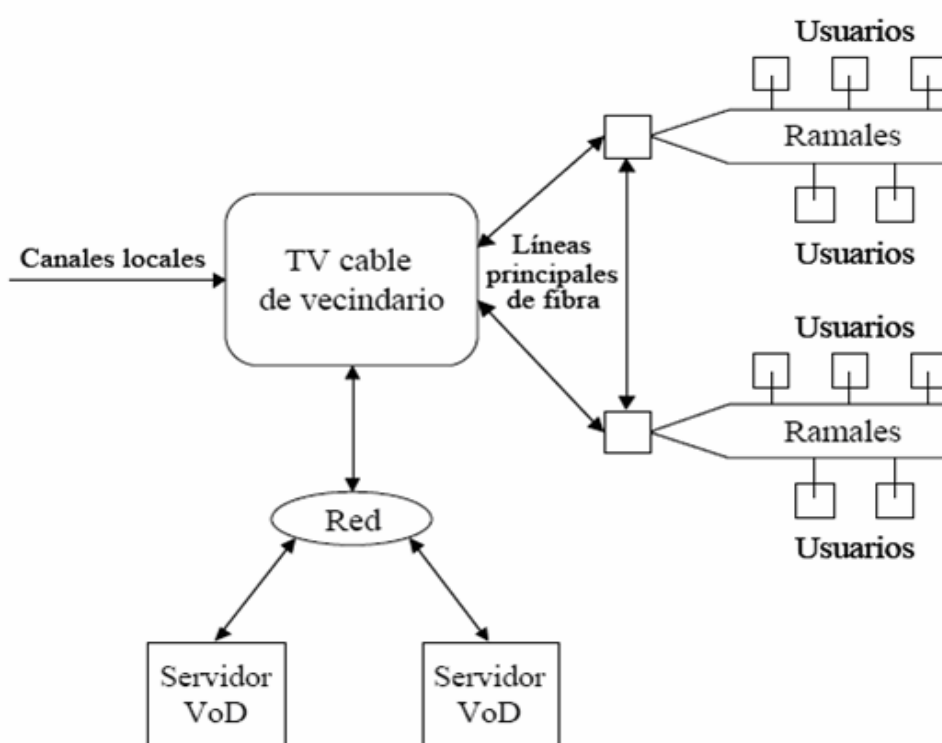


Figura 1.3 Componentes de una red HFC para prestar el servicio de VOD

El ancho de banda total disponible en los ramales suele ser alrededor de 3482 Mb/s, los cuales se dividen entre los canales de entrada (3480 Mb/s) y canales de salida (2 Mb/s).

1.2.2 RED TRONCAL

Las principales tecnologías utilizadas en la red troncal son redes ATM y redes basadas en switches fast-ethernet sobre el protocolo TCP/IP.

En la tabla 1-2, mostramos el número máximo de flujos de video que pueden soportar las distintas tecnologías de red ATM y Ethernet, para los siguientes formatos de video: Televisión de alta definición (HDTV), calidad DVD, estándar de televisión americano y japonés (NTSC, National Televisión System Commiffe), MPEG-1 y MPEG-4 (ó DivX).

A partir de la información de la figura, se puede deducir que el ancho de banda de las redes de transmisión es uno de los factores más importantes a la hora de diseñar un sistema de VOD. Un sistema de VOD medio, capaz de dar servicio a miles de usuarios requiere la utilización de una infraestructura de red compuesta por varios switches Ethernet ó ATM. Hoy en día, las

alternativas que ofrecen una mejor relación costo / prestaciones son los switches Fast-Ethernet ó ATM OC-3. A continuación describimos las características más relevantes de cada una de estas tecnologías.

Tabla 1.2 Número de flujos de Video soportados por las principales tecnologías de red

Formato de Video	Ethernet	FastEthernet	ATM(OC3)	ATM(OC12)	GibaBitEthernet
MPEG-2 HDTV (19.4 Mb/s)	0	5	7	32	51
MPEG-2 DVD (10.8 Mb/s)	0	9	14	57	92
MPEG-2 NTCS (2.7 Mb/s)	3	37	57	230	370
MPEG-1 (1.5 Mb/s)	6	66	103	414	666
MPEG-4 (<1 Mb/s)	10	100	155	622	1000

ATM

ATM es una técnica de conmutación y una tecnología de multiplexación que combina los beneficios de la conmutación de paquetes con los beneficios de la multiplexación de paquetes. Esta técnica de transmisión, está diseñada

para ser un modo de transferencia orientada a conexión, de propósito general, para un rango amplio de servicios. La técnica de multiplexación es por división asincrónica en el tiempo (ATD). ATM es una variación de las técnicas de conmutación de paquetes en la cual se emplean paquetes de tamaño fijo y reducido denominados celdas.

En un nodo de conmutación ATM, el procesamiento de las celdas se limita al análisis de una etiqueta con el número de canal lógico, a partir del cual la celda podrá ser encaminada apropiadamente. Otras funciones más complejas como chequeo y control de errores no son realizadas por la red ATM sino por los usuarios finales en los extremos de la red.

Estas particulares características facilitan una solución razonable para los problemas propuestos por las restricciones asociado con el tráfico de video en tiempo real.

ATM dispone de ciertas características que la hace especialmente interesante para los sistemas de VOD:

- ***Soporta altas velocidades***

Gracias a la limitada funcionalidad de la cabecera, el procesamiento de las mismas es simple y por lo tanto puede ser realizado a muy altas velocidades (de 150 Mb/s a Gb/s). Esto implica retardos muy pequeños de procesamiento y encolamiento.

- ***ATM opera en un modo orientado a conexión***

Antes de que la información sea transferida desde un terminal a la red, se establece una conexión lógica/virtual que permite a la red la realización de la reserva de los recursos necesarios.

- ***Reserva recursos y QOS***

Al pedir una conexión se debe especificar la calidad de servicio que se desea. Esta QOS se define en función de la tasa de error, retardo, etc. y se especifica en el llamado contrato de tráfico. Si el sistema tiene suficientes recursos y la conexión se establece, entonces se realiza un control de tráfico sobre los usuarios, evitando que algún usuario sobrepase los recursos asignados.

- ***Mínimo retardo y variabilidad***

Gracias a que los buffers internos de los nodos de conmutación son pequeños se garantiza un retardo y una variabilidad de los retardos reducidos, tal y como se requiere en los servicios de tiempo real.

Ethernet

La tecnología ATM es costosa, por lo que muchas instalaciones de VOD han preferido la utilización de la infraestructura más económica como la utilizada en las redes de área local basadas en switches fast-ethernet.

El protocolo de comunicaciones más ampliamente utilizado hoy en día en las redes de área local ó en Internet es el protocolo TCP/IP. Este protocolo fue diseñado como protocolo de conmutación de paquetes, cuyo objetivo principal era la entrega de paquetes libres de error desde un remitente a un receptor sin importar cuando lleguen al destinatario.

El IP (Internet protocol) es un protocolo situado en la capa de red del modelo OSI, basado en la conmutación de paquetes y que no está orientado a conexión. Sobre el protocolo IP normalmente son utilizados dos protocolos

de transporte: el protocolo de control de transmisión (TCP) y el protocolo de datagramas (UDP).

El protocolo TCP/IP no permite garantizar una calidad de servicio a los usuarios finales, ni permite la reserva de ancho de banda que garantice la transmisión del flujo de datos durante el periodo de visualización de los contenidos multimedia. Estas características limitan considerablemente la aplicación de este protocolo para soportar servicios de VOD. Para subsanar estas limitaciones se han propuesto un conjunto de protocolos (RTP, RTCP, RSVP y RTSP) soportados encima de TCP/IP que permita soportar el tráfico requerido por las aplicaciones de VOD. En la figura 1.4 mostramos la relación de estos protocolos con los protocolos TCP/IP y UDP.

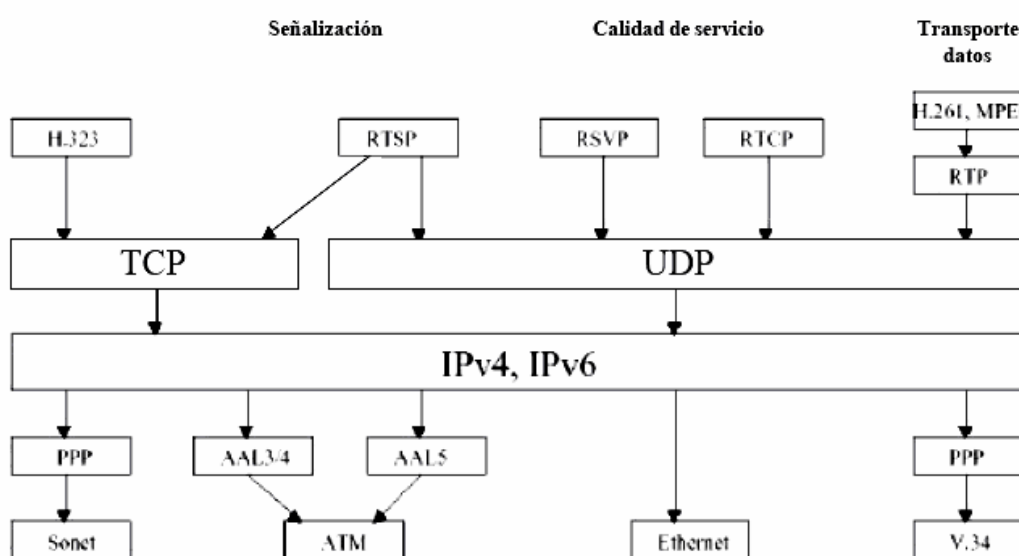


Figura 1.4 Protocolos para soportar tráfico multimedia sobre TCP/IP

- ***RTP (Real-Time Transport Protocol)***

Proporciona un mecanismo para el transporte de datos en tiempo real a través de Internet. RTP ofrece servicios de entrega extremo a extremo para datos con características de tiempo real que son adecuados para aplicaciones distribuidas que transmiten datos a tiempo real. El protocolo ofrece características importantes para las aplicaciones multimedia tales como marca de tiempo y numeración de secuencia de los mensajes e identificación del tipo de datos transmitidos, que permite un tratamiento adecuado por parte de la red.

- ***RTCP (Real-Time Control Protocol)***

Debido a que el protocolo RTP no garantiza la calidad de servicio para las comunicaciones en tiempo real, se requiere un protocolo complementario para controlar la calidad de los datos entregados y control de flujo y congestión. Este protocolo genera la transmisión de informes estadísticos entre el transmisor y receptor en el protocolo RTP, mediante los cuales se identifican el estado de congestión de la red y que consiguen limitar el número de paquetes perdidos (ajuste automático de ancho de banda).

- ***RTSP (Real-Time Streaming Protocol)***

Es un protocolo de nivel de aplicación que ofrece control sobre la entrega de datos a tiempo real. El protocolo se aplica para el control de flujos continuos sincronizados en el tiempo, tanto de audio como de video, y actúa como control remoto de red para los servidores multimedia. RTSP controla los flujos transmitidos por un protocolo de transporte (RTP por ejemplo).

- ***RSVP (Resource ReServation Protocol)***

Protocolo que se encuentra situado encima de la capa de Internet, dentro de la estructura del protocolo TCP/IP, ocupando el lugar de los protocolos de transporte. RSVP proporciona un mecanismo para configurar y gestionar la reserva de ancho de banda en Internet, permitiendo la adaptación de una transmisión a las fluctuaciones de tráfico de las redes.

1.3 ARQUITECTURAS UTILIZADAS EN LOS SISTEMAS DE VOD

En los apartados anteriores hemos descrito los principales componentes y políticas que integran un sistema de video bajo demanda. Sin embargo, a la

hora de implementar un sistema de estas características se pueden adoptar diversas arquitecturas. En este apartado describiremos la organización y características de las principales arquitecturas utilizadas para el diseño e implementación de los sistemas VOD.

1.3.1 ARQUITECTURAS CENTRALIZADAS

Estos sistemas se basan en la conexión de todas las redes de usuarios del sistema a una red principal a la cual se conecta un servidor ó un conjunto de servidores (ver figura 1.5).

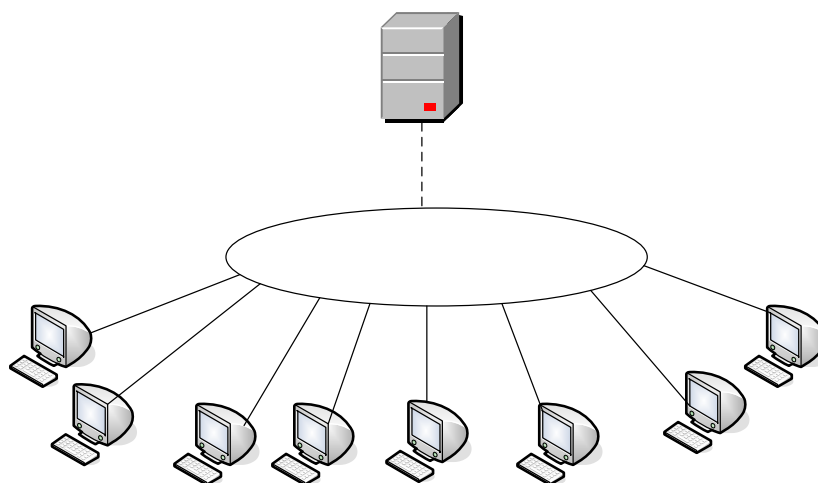


Figura 1.5 Arquitectura de VOD centralizada

Las principales características que definen las configuraciones centralizadas

son la gestión centralizada de todas las peticiones de los usuarios y la utilización de una red principal que es compartida por todos los flujos de información del sistema.

Existen dos categorías de sistemas centralizados en función del número de servidores utilizados: arquitecturas con un único servidor y arquitecturas basadas en múltiples nodos de servicio (ó arquitectura de servidor distribuida).

En la primera configuración con un único servidor, la gestión de los clientes se basa en un único nodo de servicio que centraliza la atención de todas las peticiones. Sin embargo esta aproximación, en general, tiene diversas limitaciones con respecto a la escalabilidad, tolerancia a fallos y disponibilidad del servicio.

Dentro de las arquitecturas centralizadas basadas en múltiples nodos de servicio podemos encontrar dos configuraciones diferentes en función de cómo se organizan los distintos nodos: servidores paralelos (ó array de servidores) ó formando un cluster.

- ***Servidores paralelos ó array de servidores***

Esta arquitectura consiste en un array de servidores, que trabajan de forma similar a un array de discos. Los distintos nodos de servicio no almacenan videos completos, sino que los videos son divididos en trozos y éstos son distribuidos entre los diferentes nodos para lograr una distribución de la carga más homogénea entre todos los servidores.

La figura 1.6 muestra una configuración de una arquitectura de servidores paralelos compuesta por 5 servidores, todos ellos conectados con los usuarios a través de una red de interconexión.

Cada uno de los servidores almacenan un subconjunto de segmentos (v_1, v_2, \dots, v_n) de los videos del catálogo del sistema.

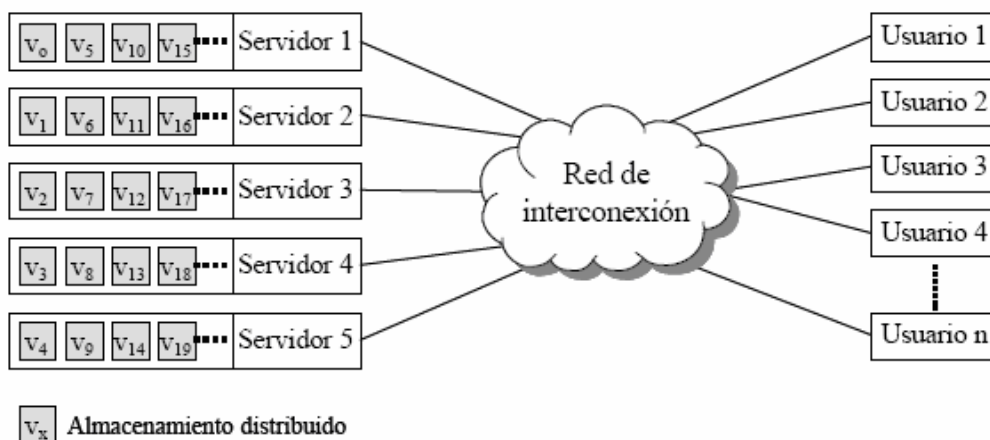


Figura 1.6 Arquitectura VOD centralizada con Servidores Paralelos

Cuando se produce una petición cada nodo de servicio es el responsable de transmitir al usuario los fragmentos de video solicitado que se encuentran en sus discos. El cliente tiene la responsabilidad tanto de fraccionar su petición en las distintas solicitudes a cada uno de los nodos de servicio, como de posteriormente recombinar y sincronizar los distintos flujos de información recibidos, para poder reproducir el contenido multimedia.

Esta arquitectura permite escalar la capacidad del sistema añadiendo nuevos nodos, aunque se requiere realizar una nueva redistribución de los videos que tenga en cuenta los nuevos servidores añadidos.

Otras ventajas asociadas con esta configuración son que permite un balanceo automático de la carga del sistema y que aumenta la tolerancia a fallos respecto a las arquitecturas basadas en un único nodo de servicio. La principal desventaja de estas aproximaciones es que incrementan los requisitos de los STB de los usuarios y complica considerablemente su diseño. Para evitar la utilización de STB demasiados complejos se ha propuesto la utilización de proxies entre los servidores y los clientes. En estos sistemas, el término Proxy hace referencia al módulo del sistema encargado de re-secuenciar y fusionar los datos procedentes de los distintos servidores en un flujo de información coherente para entregárselo al usuario final. El Proxy puede también utilizar información redundante para enmascarar posibles fallos en los servidores.

Además, al no existir un nexo común en la gestión de las peticiones de los usuarios, la utilización de políticas de compartición de recursos se complica considerablemente.

- ***Cluster de servidores***

Una arquitectura basada en cluster consiste en un grupo de nodos conectados entre sí por una red de interconexión (ver figura 1.7). Cada nodo dispone de un disco local conectado a él.

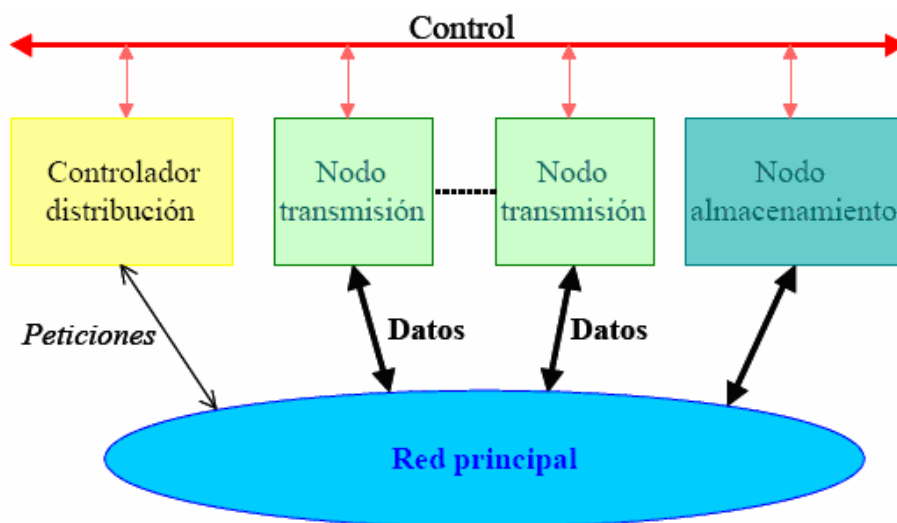


Figura 1.7 Arquitectura Centralizada basada en cluster

Los nodos del cluster se pueden dividir en tres categorías: nodos de transmisión, nodos de almacenamiento y un nodo de control. El nodo control admite las peticiones de la red externa basándose en una estrategia de control de admisión predefinida ó dinámica. Los nodos de almacenamiento, guardan los contenidos de forma similar a los servidores paralelos, proporcionándolos cuando son requeridos a los nodos de transmisión.

Los nodos de transmisión son los encargados de unir los distintos bloques correspondientes a un video, antes de su transmisión al usuario en forma de un único flujo de información.

El principal problema que sufren las arquitecturas centralizadas es el cuello de botella que representa la red principal. La escalabilidad futura del sistema queda limitada por el ancho disponible en esta red.

1.3.2 ARQUITECTURAS DE SERVIDORES INDEPENDIENTES

Una de las soluciones que se ha propuesto para incrementar la escalabilidad de los sistemas de VOD, es la conexión de los usuarios mediante servidores independientes. En estos sistemas, tal y como muestra la figura 1.8, los usuarios están agrupados en segmentos de red cuyo tráfico es independiente entre sí, denominados redes locales, de forma que el ancho de banda del sistema pueda llegar a ser el ancho de banda acumulado de cada una de las redes individuales. Sin embargo, para incrementar el ancho de banda de estos sistemas no es suficiente con únicamente agrupar los usuarios en redes independientes ya que si todos acaban accediendo al mismo servidor y a su red, éstos se convierten en el cuello de botella del sistema, con la consiguiente saturación de todo el sistema.

La clave para que estos sistemas con redes independientes funcionen y obtengan un mejor rendimiento, consiste en que las peticiones se puedan

servir localmente sin la necesidad de acceder a un servidor centralizado. Este objetivo se puede lograr colocando servidores de VOD cerca de las redes locales de los usuarios y replicando todos los contenidos, de forma que éstos no tengan que acceder al servidor central, creando un sistema de servidores independientes ó autónomos.

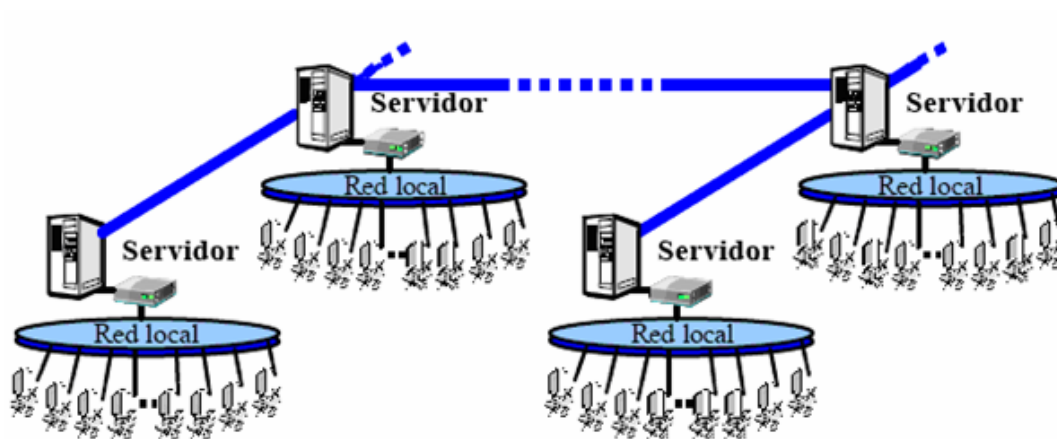


Figura 1.8 Arquitectura VOD con servidores independientes

La principal ventaja de esta arquitectura es que permite una escalabilidad ilimitada mediante la inclusión de nuevos servidores a los cuales se conectarán los nuevos usuarios y que no requiere servidores muy complejos.

Por el contrario, los sistemas de VOD basados en servidores independientes tienen unos elevados costos asociados con el subsistema de almacenamiento debido a que todos los servidores deben replicar los contenidos del catálogo del sistema.

1.3.3 ARQUITECTURAS BASADAS EN SERVIDORES-PROXY

La arquitectura de servidores independientes implica un elevado costo, y por lo tanto, algunas propuestas han optado por reducir el tamaño de los servidores locales, de forma que no almacenen una copia completa de las películas del sistema, sino únicamente los contenidos más populares. Estos servidores locales se denominan servidores-Proxy, al igual que sus homólogos de Internet y se comportan como una caché del catálogo de contenidos almacenado en un servidor principal, el cual contiene todos los videos disponibles en el sistema.

Los sistemas basados en servidores-Proxy de un nivel, identifican a una arquitectura en la cual los servidores-Proxy no están interconectados entre si. Esta arquitectura surge como un compromiso entre las arquitecturas centralizadas (no escalables pero con menores requisitos de

almacenamiento) y las arquitecturas de servidores independientes (escalables, pero con elevados costes de almacenamiento).

Los servidores-Proxy son los encargados de gestionar inicialmente todas las peticiones generadas por los usuarios conectados a sus redes (redes locales), en el caso que la petición no pueda ser atendida localmente debido a que el contenido requerido no se encuentra en la caché, entonces se redirige la petición hacia el servidor principal. Existen diversas políticas para gestionar el contenido de los servidores-Proxy en función de si almacenan los contenidos completos ó solo un fragmento: prefix-caching basado en almacenar en la caché el fragmento inicial (prefijo) de los contenidos de video más populares y segment-caching que archivan en la caché los fragmentos del video más populares.

Existen dos configuraciones básicas que se pueden utilizar a la hora de diseñar un sistema VOD basado en servidores-Proxy. Ambos difieren en la arquitectura utilizada para el servidor principal a la cual se conectan los distintos servidores-Proxy. Tenemos la arquitectura de servidores-Proxy basados en un servidor centralizado y las arquitecturas de servidores-Proxy basadas en un servidor paralelo ó jerárquico.

- ***Basados en un servidor principal centralizado***

La topología general de un sistema basado en servidores-Proxy (ver figura 1.9) se compone de un servidor principal al cual se conectan directamente y a través de una red principal, un conjunto de redes locales con su Proxy. Debido a que solo hay un nivel de servidores-Proxy en la arquitectura, este sistema se suele denominar, sistema basados en servidores-Proxy de un nivel (en contraposición a otras arquitecturas que pueden utilizar diferentes niveles jerárquicos de servidores-Proxy dentro del sistema).

En la figura también se pueden ver los dos tipos de peticiones que tienen estos sistemas: peticiones servidas localmente y las peticiones atendidas por el servidor principal. Las peticiones locales (trazo continuo) son aquellas que pueden servir desde los contenidos almacenados en la caché de los servidores-Proxy. Cuando el servidor-Proxy no dispone del contenido requerido por la petición, ésta se redirige hacia el servidor principal que se encargará de su servicio.

El principal problema que se encuentra en estos sistemas es la escalabilidad limitada derivada de la dependencia de unos componentes centralizados como son el servidor y la red principal. La capacidad de crecimiento del sistema dependerá en última instancia de la capacidad de estos componentes centralizados. De todas formas, siempre disponen de un mayor margen de maniobra comparado con los sistemas centralizados.

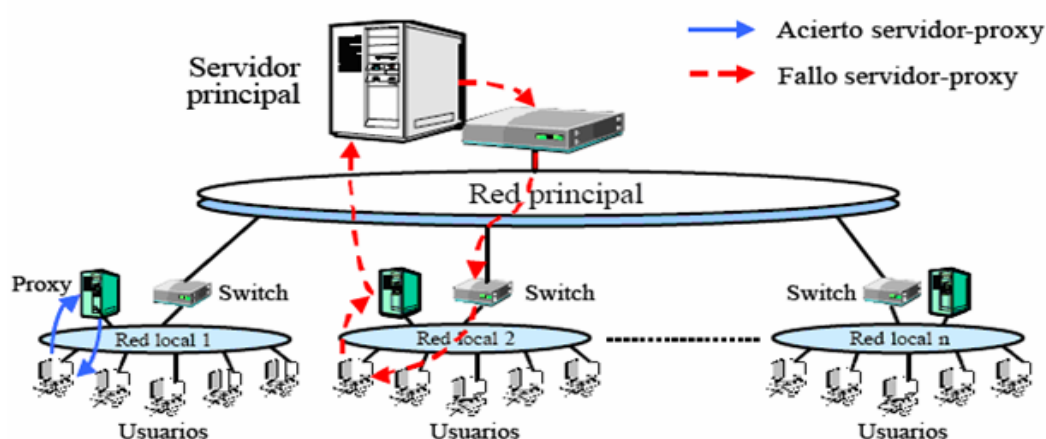


Figura 1.9 Arquitectura VOD de Servidores Proxy de un nivel

- **Basados en un servidor principal paralelo / jerárquico**

Esta aproximación trata de solventar los problemas de escalabilidad del servidor principal-centralizado en la arquitectura de servidores-Proxy de un nivel. En esta arquitectura, de la cual la figura 1.10

muestra una posible configuración, el servidor principal esta diseñado basándose en una red jerárquica ó en árbol, con servidores de VOD en los nodos y enlaces de red en las ramas de la jerarquía. Los nodos de servicio situados en las hojas de la jerarquía son los puntos de acceso para el sistema.

Todos los nodos del sistema solo almacenan un subconjunto de los contenidos del sistema.

Cuando una petición para un contenido llega a un nodo hoja, si el contenido está disponible en su almacenamiento local, el servidor atiende el mismo al cliente.

En caso contrario, la petición se reenvía hacia los niveles superiores de la jerarquía para que sea atendida por otro nodo de servicio de la arquitectura que disponga del contenido requerido.

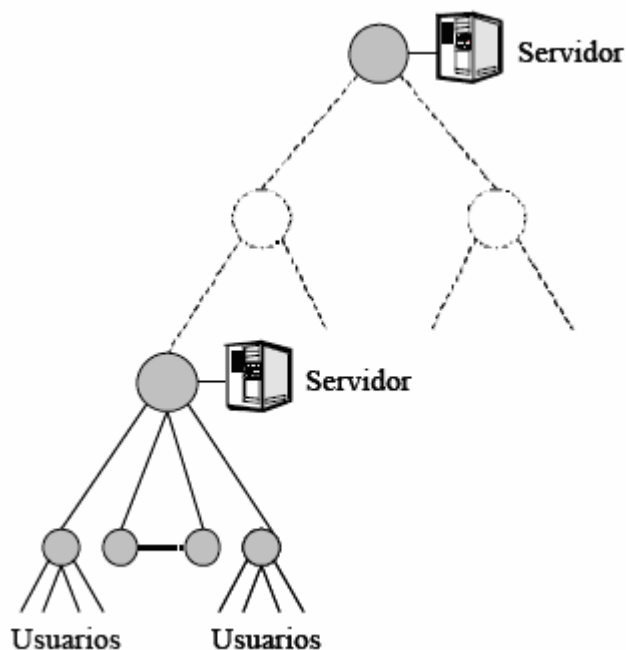


Figura 1.10 Arquitectura de VOD Jerárquica de servidores-proxy

El rendimiento de esta arquitectura es similar a las basadas en servidores-Proxy de un nivel conectadas a un servidor centralizado, pero reduce la probabilidad de saturación del servidor principal e incrementa la capacidad de servicio del sistema.

1.3.4 ARQUITECTURAS DISTRIBUIDAS A NIVEL DE LOS USUARIOS

Las últimas tendencias a la hora de diseñar sistemas de VOD se orientan hacia la adopción de arquitecturas distribuidas, en las cuales la gestión de las

peticiones, así como los contenidos multimedia se distribuyen entre todos los componentes del sistema.

En estos sistemas, los distintos nodos de servicio tienen que colaborar entre sí para poder atender a los usuarios.

Existen diferentes categorías de sistemas distribuidos en función de si existen ó no un nodo maestro encargado de centralizar la gestión del sistema y mantener una copia completa de los contenidos del sistema (arquitecturas de servidores-Proxy de un nivel, por ejemplo).

Una de las primeras opciones en este sentido es la política de servicio de *Chaining* (encadenamiento). Esta política utiliza el contenido de los buffers internos de los STB de los usuarios para a su vez servir peticiones de otros usuarios hacia el mismo contenido.

De esta forma, se crea una cadena de servicio entre los propios usuarios del sistema que permite reducir la carga del servidor de VOD del sistema.

1.4 ARQUITECTURAS DE VOD A GRAN ESCALA

A pesar del atractivo de los servicios de VOD para el público en general, su implementación hasta el momento no ha sido tan grande como se pudiese esperar, debido a la dificultad de diseñar y construir sistemas de VOD de gran escala (LVOD) capaz de atender a decenas de miles de peticiones simultáneas. La construcción de sistemas de video bajo demanda de gran escala está actualmente limitada por dos factores: la capacidad de transmisión simultánea de videos (capacidad de streaming) que puede soportar el servidor y la red de comunicación, y por otro lado, los elevados costes requeridos para su construcción.

Para proveer servicios de VOD capaz de atender a decenas de miles de usuarios es imprescindible el diseño de sistemas de VOD a gran escala eficientes y con un costo aceptable.

1.4.1 REQUISITOS DE LAS ARQUITECTURAS LVOD

A la hora de realizar el diseño de un sistema LVOD, además de tener que proporcionar una alta capacidad de streaming, también son de vital importancia considerar los conceptos de escalabilidad, tolerancia a fallos,

costo, balanceo de carga y compartición de recursos, que describimos a continuación:

- ***Escalabilidad***

En términos generales, la escalabilidad hace referencia a la capacidad del sistema para mantener, sino mejorar, su rendimiento a medida que aumenta el número de clientes.

Ninguna instalación de VOD puede crecer desde 0 a un millón de usuarios de un día para otro. Por lo tanto, sobredimensionar el sistema de VOD teniendo en cuenta los posibles usuarios futuros, puede dar a lugar a que cuando una capacidad adicional se requiera el sistema ya resulte obsoleto debido a los avances en la tecnología.

Los sistemas de VOD deben permitir ajustar su capacidad inicial a los requerimientos de los usuarios para así reducir la inversión inicial.

Pero al mismo tiempo deben conservar intacta capacidad de crecimiento futuro.

La escalabilidad es una de las características más importantes para un sistema VOD, permitiendo que se pueda ajustar el tamaño inicial del sistema a los requerimientos de los usuarios, pero manteniendo la posibilidad de un fácil crecimiento para poder soportar más usuarios ó nuevos servicios.

En general, en los sistemas distribuidos existen dos tipos de escalabilidad: la vertical y la horizontal. Al hablar de escalabilidad vertical nos referimos a incrementar el número de recursos en la misma máquina para conseguir atender a un mayor número de usuarios. La escalabilidad horizontal consiste en incrementar el número de máquinas que integran el sistema.

- **Tolerancia a fallos**

Los sistemas de VOD tienen que continuar dando servicio a los usuarios, incluso si uno ó más componentes de la arquitectura fallan.

En sistemas de gran escala, normalmente enfocados a un público que paga por unos servicios, no es asumible una interrupción del servicio debido a un fallo de alguno de los componentes. Estas interrupciones de servicio, en caso de producirse, pueden producir un grave perjuicio económico y de imagen.

Los sistemas de gran escala son más susceptibles de sufrir algún fallo debido al gran número de componentes (servidores, redes, discos duros) que los forman. Por lo tanto, es importante tener en cuenta la tolerancia a fallos en el diseño de estos sistemas.

Una de las técnicas utilizadas para aumentar la tolerancia a fallos, es la replicación de los componentes más críticos del sistema. Aunque reduce la posibilidad de interrupción del servicio, esta técnica implica una gran inversión en componentes que no estarán en funcionamiento durante la mayor parte del tiempo de vida del componente, y que pueden llegar a quedarse obsoletos incluso antes de ser utilizados. La utilización de componentes de respaldo es muy costosa, por lo cual se suele utilizar cuando no hay más remedio, por ejemplo en sistemas centralizados.

Una alternativa mejor y menos costosa consiste en que los componentes de respaldo formen parte de la propia arquitectura. De esta forma, el posible respaldo para un servidor del sistema es otro servidor de la misma arquitectura. La mejor forma de implementar esta política es logrando que la atención de los usuarios no dependan de un único servidor, sino que puedan ser atendidos por distintos componentes del sistema, en función de las necesidades. La utilización de una arquitectura totalmente distribuida, en la cual la gestión de las peticiones se realiza de forma descentralizada, es la forma más fácil de obtener la tolerancia a fallos sin necesitar componentes de respaldo.

- **Costo**

El VOD a gran escala en el ámbito empresarial está mayormente enfocado a ofrecer servicios de entretenimiento y multimedia a grandes ciudades o la unión de varias ciudades grandes como lo son Guayaquil, Quito y Cuenca.

Este tipo de instalaciones requieren de una gran inversión y, por lo tanto, es imprescindible un control muy estricto de los costos de la arquitectura de VOD.

Desde el punto de vista de la arquitectura de VOD (sin tener en cuenta el Set-Top-Box mediante el cual se conectan los usuarios), los componentes que requieren una mayor inversión son las redes de comunicación y los servidores de video.

El costo de las redes de transmisión depende básicamente del ancho de banda requerido y de la tecnología utilizada. El costo asociado con una red individual se puede incrementar exponencialmente si el ancho de banda requerido es muy grande. Para reducir el costo del sistema de comunicaciones es recomendable que la arquitectura de LVOD no requiera la utilización de redes con anchos de banda muy grandes.

Con respecto a los servidores, los componentes que más influyen en su costo son los asociados con el ancho de banda y el sistema de almacenamiento requerido para soportar el catálogo de contenidos del

sistema. El costo del servidor, al igual que ocurre con la red, depende de la capacidad de servicio deseada. Un servidor más potente necesitará la utilización de técnicas de clustering y discos RAID para poder soportar el ancho de banda requerido, aumentando el costo considerablemente.

En general, parece lógico pensar que se obtiene la mejor relación costo / rendimiento utilizando componentes pequeños y cuyo uso esté generalizado en el mercado. Se deben evitar requisitos que necesiten la utilización de componentes demasiado complejos ó que requieran de las últimas tecnologías disponibles en el mercado.

- **Balanceo de la carga**

La distribución de la carga entre los distintos servidores del sistema, es importante debido a que las peticiones de los usuarios siguen una distribución no uniforme. Esta característica puede provocar un desbalanceo en el volumen de trabajo de los servidores y una pobre utilización de los recursos globales del sistema.

El sistema de VOD debería permitir que parte de la carga de los componentes (servidores ó redes) más saturados, se pueda desviar a otros componentes menos cargados. Una correcta re-distribución de la carga en el sistema, permite reducir la probabilidad de rechazo de servicio a los usuarios.

1.4.2 ALTERNATIVAS ACTUALES PARA LOS SISTEMAS LVOD

Las arquitecturas que se han explicado en la parte inicial del capítulo no cumplen todos los requisitos comentados en el apartado anterior, de forma que la distribución de contenidos multimedia a gran escala en redes de área global sea una realidad.

La construcción de un sistema de LVOD a gran escala está actualmente limitada tanto por la capacidad del servidor (ancho de banda de servicio) como por la capacidad de transmisiones simultaneas que puede soportar una red de comunicaciones (ancho de banda de la red).

Actualmente, la implementación de un sistema de LVOD que pueda soportar un gran número de streams concurrentes, requiere la disposición de varios servidores, que ofrecen la transmisión de video y los servicios de reproducción bajo la forma de un sistema distribuido.

Sin embargo, las distintas aproximaciones para crear esta infraestructura pueden variar desde sistemas completamente centralizados que utilizan hardware dedicado en la capa de red sin almacenamiento intermedio, hasta sistemas completamente descentralizados que replican todo el catálogo de contenidos en servidores cercanos al usuario final.

Cada una de estas arquitecturas tiene sus ventajas e inconvenientes dentro de un entorno de servicio a gran escala que comentamos a continuación:

- **Sistemas centralizados**

Un sistema LVOD centralizado requiere un servidor y una red principal capaces de soportar un gran ancho de banda. Un servidor con estas características puede llegar a ser muy costoso y complejo de diseñar / construir. La disponibilidad de redes capaces de soportar estos volúmenes de tráfico puede estar limitado por la tecnología disponible. Además el sistema resultante no es tolerante a fallos, ni escalable ya

que el crecimiento del sistema esta limitado por la red y el servidor principal.

La única ventaja de esta arquitectura es la alta eficiencia que pueden obtener de las políticas multicast. Al estar todos los usuarios conectados a la misma red, la probabilidad de compartición de streams es la más alta de todas las arquitecturas de VOD. Estos sistemas adolecen de una serie de problemas que dificulta su candidatura a una instalación de gran escala

- ***Servidores independientes***

Esta arquitectura no requiere ni redes grandes, ni servidores complejos para lograr la alta capacidad de streaming requerida por los sistemas de gran escala. Estos sistemas permiten una escalabilidad ilimitada ya que para aumentar la capacidad del sistema únicamente se necesita añadir nuevas redes locales (más los servidores correspondientes) al sistema.

La principal desventaja de esta arquitectura en un entorno de gran escala son los requerimientos de almacenamiento en los servidores (cada servidor necesita una copia entera del catálogo de contenidos), la compartición de recursos esta restringida exclusivamente a los usuarios locales y el balanceo de carga es casi imposible a no ser que las redes independientes estén interconectadas.

- ***Sistemas basados en servidores-Proxy de un nivel***

Como ya hemos visto, uno de los problemas de estos sistemas estriba en la escalabilidad limitada derivada de la dependencia de unos componentes centralizados como son el servidor y la red principal. La capacidad de crecimiento del sistema dependerá de la capacidad de estos componentes centralizados. De todas formas, disponen de un mayor margen de crecimiento comparado con los sistemas centralizados.

Los requisitos globales de ancho de banda de red en estos sistemas, son mayores debido a que las peticiones que no pueden ser atendidas desde la caché de los servidores-Proxy, requieren el doble de recursos de red (red principal + red local) para poder ser atendidas desde el servidor principal.

Por otro lado, al gestionarse la mayoría de las peticiones en los servidores-Proxy, también se reduce la probabilidad de compartición de recursos entre usuarios con respecto a los sistemas centralizados.

- ***Servidores basados en servidores-Proxy paralelos/jerárquicos***

Básicamente estas arquitecturas tienen los mismos parámetros de rendimiento que los sistemas basados en servidores-Proxy de un nivel. La única diferencia estriba en la utilización de un servidor paralelo que aporta múltiples puntos de entrada al servidor y permite reducir el cuello de botella que representa la red principal en los sistemas de servidores-Proxy de un nivel.

En última instancia, la escalabilidad de estas arquitecturas depende de la escalabilidad asociada con el servidor paralelo jerárquico. Por otro lado, parte del ahorro de los costos obtenidos al reducir los requerimientos de almacenamiento se compensan con el costo asociado con el servidor paralelo.

- **Sistemas distribuidos a nivel de usuario**

La principal desventaja con la que se tropiezan estos sistemas es el incremento de la complejidad y costo del Set-Top-Box de los usuarios, al tener que incorporar buffers más grandes y requerir anchos de bandas de red capaces de soportar más de un stream simultáneo.

En la tabla 1.3 mostramos las principales características de las distintas arquitecturas de LVOD. A pesar de que todas estas arquitecturas son capaces de conseguir una capacidad alta de streaming, ninguna de ellas alcanza a cumplir con todos los características deseables para un sistema de video bajo demanda de gran escala.

Tabla 1.3 Principales características de sistemas LVOD

Arquitectura	Alta capacidad de servicio	grandes anchos de banda	tolerante a fallos	Escalable
Centralizada	Si	Si	No	No
Servidores independientes	Si	No	Si	Si
Basados en Servidores-Proxy	Si	Si	No	No
Distribuidos a nivel del cliente	Si	No	Si	Si

Los principales inconvenientes que tienen estas arquitecturas son la insuficiente escalabilidad y el elevado costo requerido para el sistema.

1.4.3 SISTEMAS COMERCIALES DE VOD

Además de las arquitecturas para VOD descritas anteriormente, existen diversas aproximaciones de sistemas de VOD comerciales. A continuación vamos a comentar brevemente las características de tres arquitecturas comerciales: *Kasenna*, *Ncube* y *Media Hawk* que tienen un diseño propietario.

- **KasennaTM**

Kasenna es una plataforma comercial de entrega de video, basada en una arquitectura denominada Kasenna VCD (Video Content Distribution). Este sistema se basa en una arquitectura de servidores-Proxy, interconectados entre ellos. Los servidores-Proxy utilizan una arquitectura interna basada en un cluster.

Cuando un usuario selecciona un video, especificando una URL, se examinan los metadatos (información sobre los contenidos multimedia y su ubicación) correspondientes al video solicitado. Si el objeto multimedia se encuentra en la caché del servidor local, el video es transmitido por éste directamente. En caso contrario, se consulta los metadatos para determinar el servidor más cercano que puede tener el contenido solicitado en su almacenamiento. Una vez localizado el contenido multimedia, este se transmite desde el servidor remoto al usuario, copiándolos a su vez en el servidor local. Para eliminar los retardos iniciales, asociados con la copia del video en el servidor local, se utiliza la técnica de prefix-caching. De esta forma, el usuario puede iniciar inmediatamente la reproducción (utilizando el prefijo del video), mientras se completa la recepción por parte del servidor local del contenido completo transmitido por el servidor remoto.

- **NCubeTM**

Ncube es un servidor de VOD paralelo, basado en una topología en hipercubo. Esta topología aporta al sistema una considerable escalabilidad mediante el incremento de la topología en cubo original, tal y como se muestra en la figura 1.11.

El sistema Ncube alcanza su mayor dimensión con 256 nodos de servicio, lo cual le permite alcanzar una capacidad de servicio de hasta 34.000 streams independientes.

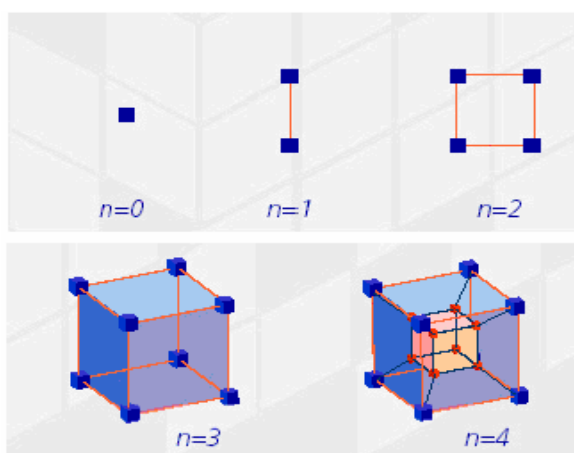


Figura 1.11 Escalabilidad de la topología Ncube

Otras de las características de esta arquitectura son:

Soporta servicios NVOD y VOD.

Es tolerante a fallos, gracias a la utilización de un sistema de almacenamiento basado en la tecnología RAID con discos redundantes.

La compatibilidad con múltiples entornos de red (Gigabit Ethernet, 10/100 baseT, ATM OC- 3, ATM OC-12, etc.), basándose en protocolos RTP/UDP.

No necesita utilizar la replicación de contenidos para garantizar que cualquier usuario pueda acceder a cualquier contenido.

- **Mediallawk Video Server™ (Concurrent Computer Corporation)**

El servidor MediaHawk se puede adaptar a los requisitos de capacidad requeridos por el sistema mediante la inclusión de múltiples módulos de servicio, denominados IVM (Interactive Video Module). Cada módulo IVM es capaz de alcanzar un ancho de banda de servicio de 160 Mb/s.

Los servidores *MediaHawk* pueden formar parte de una arquitectura centralizada para el sistema de VOD o como parte de una arquitectura distribuida. Estos sistemas soportan la tolerancia a fallos mediante la utilización de arrays de discos replicados.

Capítulo 2

2. Proxy-Tree, Arquitectura para sistemas LVOD

En este capítulo se analiza la arquitectura LVOD escalable denominada Proxy-Tree, basada en un sistema jerárquico de servidores-Proxy interconectados mediante una topología en árbol. A continuación se especifican las políticas de gestión de los contenidos (mirroring y caching) que permiten distribuir la gestión del sistema entre todos los componentes del mismo. Por último, se desarrolla un modelo analítico para esta arquitectura, mediante el cual se evaluará las prestaciones de la arquitectura Proxy-Tree y se demostrará su escalabilidad.

2.1 INTRODUCCIÓN

En el presente capítulo vamos a analizar una arquitectura escalable para sistemas de LVOD. En la tabla 2.1 resumimos las principales características que debería poseer una arquitectura LVOD ideal. De las distintas características, las más importantes sin lugar a dudas son las referentes a la escalabilidad, capacidad de servicio y coste final del sistema.

Tabla 2.1 Características ideales para la arquitectura LVOD

Arquitectura	Alta capacidad de Servicio	Grandes Anchos De Banda	Tolerancia a Fallos	Escalable	Coste
Ideal	Si	No	Si	Si	Razonable

El sistema LVOD resultante debe tener un costo razonable, por lo tanto es imprescindible reducir al máximo el tamaño inicial del sistema (creciendo a medida que las necesidades lo requieran), y evitar la utilización de componentes demasiado costosas.

Estas características definen una serie de objetivos que se deben cumplir a la hora de escoger una arquitectura de VOD a gran escala:

1. Alta capacidad de servicio de videos, independiente de la tecnología disponible y adaptable de forma flexible a las necesidades del sistema.
2. Escalabilidad ilimitada con unos costos acotados.
3. Tolerancia a fallos.
4. Capacidad de servicio.

Para cumplir estos objetivos se estudiará una arquitectura distribuida denominada Proxy-Tree, la cual se presentará en el siguiente apartado.

2.2 PROXY-TREE, UNA ARQUITECTURA ESCALABLE

La mayoría de los objetivos que se plantean el momento de elegir una arquitectura dependen de la capacidad de crecimiento del sistema. Por lo tanto, la arquitectura LVOD debe estar orientada en una primera instancia a desarrollar la capacidad de crecimiento del sistema.

2.2.1 ANÁLISIS DE LA ESCALABILIDAD EN LOS SISTEMAS DE VOD

Para lograr un sistema de VOD que sea escalable es necesario conseguir que tanto el ancho de banda de servicio (número de peticiones

independientes que es capaz de atender el servidor) como el ancho de banda del sistema de comunicaciones pueda crecer a medida que el sistema se amplía.

De los principales componentes de un sistema de VOD (servidor y red de transmisión) el ancho de banda del servidor siempre es mayor que el ancho de banda de red, debido a que la tecnología interna del bus ofrece una mejor relación costo / prestaciones y una mayor escalabilidad. Este crecimiento se puede conseguir mediante la inclusión de nuevos discos en configuración de RAID, con la utilización de la metodología de cluster ó la utilización de servidores paralelos (varios servidores independientes conectados a la misma red de servicio). Aún así la creación de estos servidores que puedan gestionar decenas de miles de peticiones simultáneas es compleja y muy costosa debido a las altas prestaciones requeridas.

Por otro lado, el ancho de banda de la red es menor (debido al costo asociado) convirtiéndose en el verdadero cuello de botella del sistema a la hora de su crecimiento, restringiendo considerablemente su tamaño.

La razón de la pobre escalabilidad del sistema de comunicaciones estriba en que su ancho de banda está limitado por la tecnología disponible en cada momento. La capacidad de la red, también está limitada por la infraestructura utilizada y sus únicas posibilidades de crecimiento pasan por la modificación de toda la infraestructura ó la inclusión de nuevas redes conectadas a un servidor central (sistemas basados en servidores-Proxy) ó servidores independientes. La primera alternativa implica un alto costo debido a la necesidad de cambiar toda la infraestructura de transmisión y no garantiza que al cabo de poco tiempo el sistema se vuelva a saturar y se tenga que volver a ampliar de nuevo. La segunda alternativa implica la utilización de arquitecturas que no soportan todas las características exigidas para un sistema LVOD escalable, como ya hemos expuesto en el capítulo anterior.

En definitiva, el tamaño y crecimiento final de los sistemas de VOD dependen, en última instancia, de la capacidad (ancho de banda) de la red de conexión con los clientes finales. Para poder lograr un sistema de transmisión escalable es imperativo prescindir de las limitaciones (físicas) de los sistemas centralizados (una única red ó un único servidor), ya que éstos siempre estarán limitados por la tecnología disponible en cada momento. Por lo tanto, la arquitectura que se debe escoger está orientada hacia los sistemas distribuidos basados en servidores-Proxy.

2.2.2 TOPOLOGÍA EN ÁRBOL

La primera instancia a tomar es la escalabilidad de la arquitectura a seleccionar por lo que en la actualidad podemos encontrar una que esté basada en una topología de árbol. Esta arquitectura es denominada Proxy-Tree ó PTree.

Esta topología aporta al sistema una capacidad de escalabilidad ilimitada, así como una gran flexibilidad a la hora de escoger el tamaño del sistema ó su forma de crecimiento. Esta flexibilidad nos va a permitir una mayor maniobrabilidad a la hora de desplegar el sistema y en su posterior crecimiento. Sin embargo, no todos son ventajas en la topología en árbol. Hay que tener especial cuidado con el nodo inicial del árbol, ya que puede convertirse en el cuello de botella del sistema.

La topología P-Tree, mostrada en la figura 2.1, consta de una serie de niveles (4 en este caso), en función del número de redes locales que componen el sistema de VOD y del orden del árbol (binario, terciario, etc.).

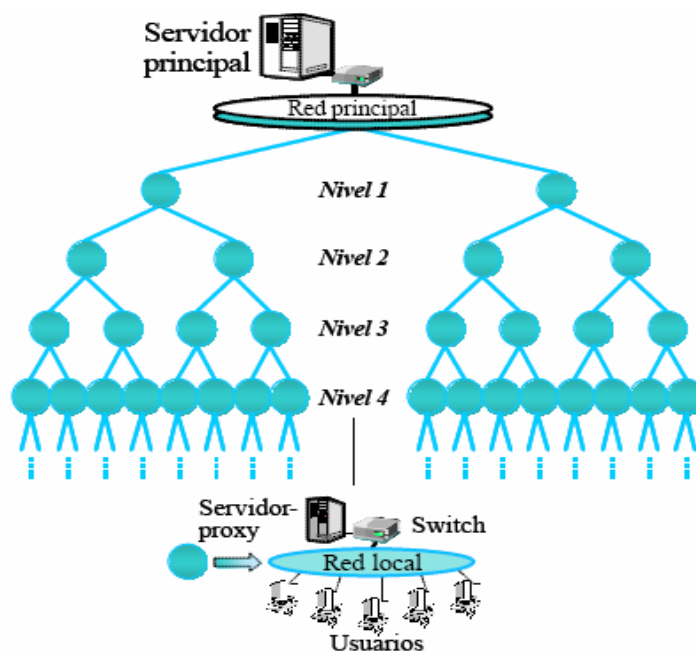


Figura 2.1 Arquitectura P-Tree para un sistema LVOD

Cada nivel de la jerarquía estará compuesto por una serie de redes, denominadas redes locales, que se conectan con el nivel superior del árbol. De cada una de las redes locales de un nivel, se colgarán las redes del siguiente nivel, y así sucesivamente hasta completar el último nivel. En cada red de la topología (salvo la red principal) se conectará un servidor-Proxy (ó también denominado servidor local) y un subconjunto de los usuarios del sistema.

La replicación de todos los contenidos del catálogo del sistema en cada uno de los servidores locales puede resultar económicamente inviable si el catálogo es demasiado grande. Por lo tanto, a diferencia de los sistemas de servidores independientes, esta arquitectura asumirá que los servidores locales dispondrán de una capacidad de almacenamiento limitada para los contenidos multimedia.

Cada uno de los servidores-Proxy de forma individual no puede contener una copia completa de los videos del catálogo del sistema. Para suplir esta carencia, la arquitectura dispondrá de un servidor principal, ubicado en el primer nivel de la topología. Este servidor dispondrá de una copia completa del catálogo de contenidos multimedia.

Esta arquitectura es fácilmente escalable ya que el añadir una nueva red local únicamente implica la inclusión de un switch para separar el tráfico de los dos segmentos de red, sin tener que modificar ningún otro componente del sistema.

La topología en árbol utilizada, nos ofrece una gran versatilidad, debido a que el sistema inicialmente puede estar compuesto por cualquier número de

redes y puede crecer tanto a lo largo, añadiendo nuevos niveles ó bien a lo ancho, incrementando el orden de algunos de los nodos del árbol. Hay que subrayar que la arquitectura Proxy-Tree no requiere ni que los niveles estén completos ni tampoco requiere que todos los niveles del árbol tengan el mismo número de redes conectadas.

2.2.3 POLÍTICAS DE GESTIÓN DE LOS CONTENIDOS EN LOS SERVIDORES-PROXY

En un primer análisis, hemos podido constatar que una topología jerárquica en árbol tiene cualidades, desde el punto de vista de la escalabilidad y versatilidad, ideales para una arquitectura LVOD. No obstante, independientemente del tipo de topología utilizada, se debe demostrar que la arquitectura que se ha seleccionado escala sin saturar el sistema.

Dentro de la arquitectura P-Tree, todos los servidores-Proxy cuyas redes locales se encuentren situadas a una misma distancia del nodo en donde se produjo la petición definen un nivel dentro de un sistema jerárquico de caché. El tamaño de cada nivel de caché vendrá definido como la suma de las capacidades individuales de cada uno de los servidores-Proxy situados dentro de una determinada distancia.

En la figura 2.2 se muestran los tres niveles jerárquicos de caché definidos alrededor de un nodo de la topología, podemos ver que el primer nivel está compuesto exclusivamente por el servidor local, el segundo nivel está compuesto por 3 servidores-Proxy adicionales (situados a distancia 1 del servidor-Proxy en donde se produjo la petición original) y el tercer nivel está compuesto por 5 servidores-Proxy (situados a distancia 2) más el servidor principal.

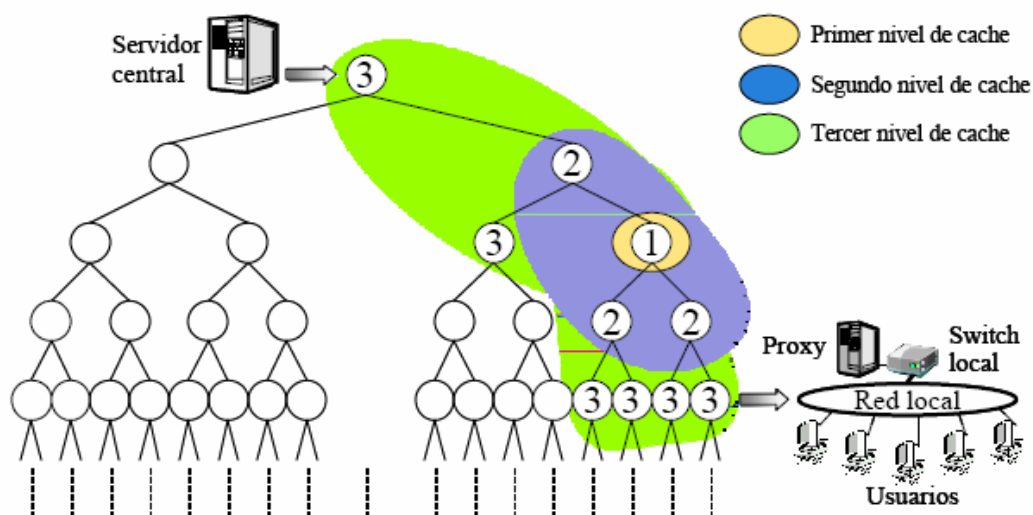


Figura 2.2 Sistema Jerárquico de caching

Este sistema jerárquico propuesto tiene la ventaja de que a medida que una petición no puede ser atendida localmente por su servidor-Proxy (primer nivel del almacenamiento jerárquico), ésta se puede dirigir hacia los servidores-

Proxy vecinos (segundo nivel del almacenamiento jerárquico) que pueden intentar atender la petición. Si ninguno de los servidores vecinos pueden gestionar la petición, entonces ésta se envía hacia los siguientes servidores (tercer nivel del almacenamiento jerárquico), y así sucesivamente hasta que se alcance un servidor-Proxy puede atender la petición ó hasta que se alcance el servidor principal.

Los servidores-Proxy tradicionales, gestionan su almacenamiento como una caché de contenidos multimedia, almacenando en cada momento los videos más populares. Este sistema de gestión de contenidos, implica una alta probabilidad de que todos los servidores-Proxy tengan replicados los mismos videos. Por lo tanto, si una petición no puede ser servida desde su propio servidor-Proxy local, entonces muy probablemente tampoco podrá ser servida por ninguno de los servidores-Proxy del sistema, y se tendrá que acabar recurriendo al servidor principal.

Este comportamiento, queda reflejado en la figura 2.3, en la cual una petición que no ha podido ser atendida por el servidor de un nivel intermedio de la topología, no puede ser atendida por ninguno de los servidores-Proxy y finalmente tiene que acudir al servidor principal

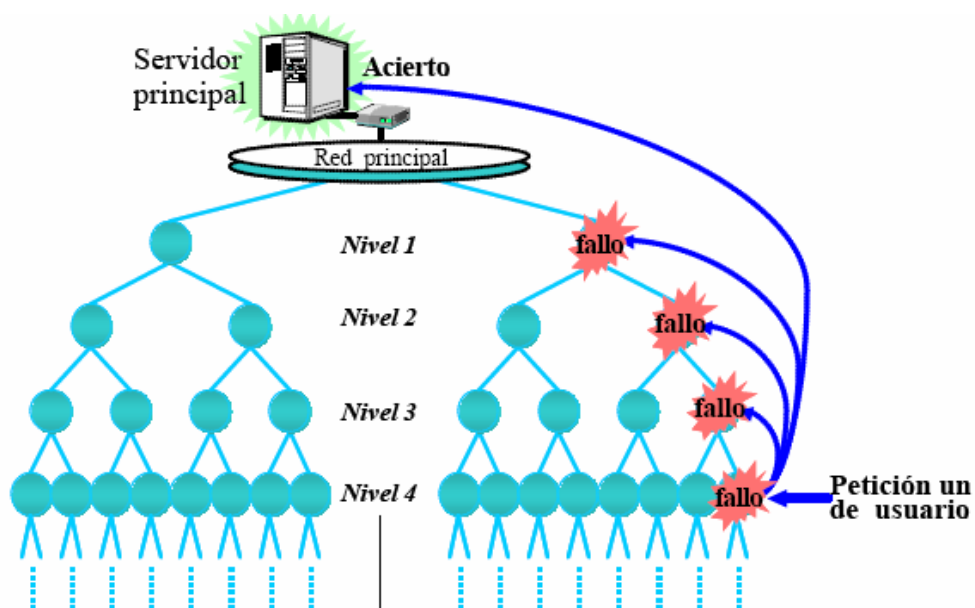


Figura 2.3 Gestión de una petición rechazada en sistema jerárquico

Mediante esta técnica, el almacenamiento dedicado a la caché de una serie de servidores-Proxy adyacentes se gestiona conjuntamente como si se tratase de una única caché distribuida. Las políticas de gestión de esta caché se realiza teniendo en cuenta el estado y los contenidos de todos los servidores-Proxy que la componen. Por ejemplo, las políticas de reemplazo en la caché cooperativa tendrán en cuenta las estadísticas de acceso en todos los servidores-Proxy cooperantes.

Esta técnica permite incrementar considerablemente el tamaño de la caché, pero a costa de reducir la probabilidad de acierto local. Bajo el esquema de caché cooperativo, un único servidor-Proxy no dispone de los contenidos más populares, provocando un aumento del porcentaje de peticiones que se tienen que servir remotamente desde los otros servidores que componen la caché ó desde el servidor principal.

Por otro lado, ésta técnica tampoco termina de subsanar uno de los problemas asociados con las caches: una parte de los contenidos (los menos populares) nunca tendrán la posibilidad de ubicarse en alguna de las caches y se tendrán que servir siempre desde el servidor central. Un caso extremo de caché cooperativa, que permite soslayar este problema, es aquella que tiene suficiente capacidad entre todos los servidores-Proxy involucrados como para almacenar todos los contenidos del catálogo. En este caso, ya no se puede hablar de una caché cooperativa sino de un mirror distribuido.

Mirroring

La técnica de mirroring consiste en replicar los contenidos, sin tener en consideración su frecuencia de acceso. Esta técnica es útil cuando se quiere aportar tolerancia a fallas ó bien reducir la latencia de acceso a información

remota, manteniendo una copia de los contenidos más cerca del usuario final.

En nuestro caso, el mirroring también nos interesa para reducir los requisitos de ancho de banda de las peticiones que se tienen que atender remotamente y poder reducir así la saturación del servidor principal.

En los esquemas más comunes de mirroring, éstos suelen contener una copia completa de la información del sistema. Un ejemplo de este esquema de gestión en VOD es la arquitectura de sistemas independientes, en la cual cada servidor mantiene un mirror de los contenidos del sistema. En el se analizan las diferencias en el rendimiento entre las políticas de caching y replicación (mirroring) de contenidos web, en función de la calidad de servicio, requerimientos de almacenamiento y perfil del tráfico.

Lo que la arquitectura Proxy-tree utiliza es mirroring distribuido, en el cual cada servidor local mantendría una proporción de los contenidos del catálogo del sistema. Mediante este esquema, el sistema dispondrá de una serie de mirrors distribuidos, cada uno de ellos formados por N/V servidores, siendo N

el número de videos del catálogo del sistema y V el número de videos que caben en el mirror parcial de cada servidor Proxy.

Tal y como hemos comentado, utilizando únicamente una política de caching, los contenidos menos populares tienen que ser gestionados en exclusiva por el servidor principal. Este volumen de carga dirigido a un único componente puede provocar a la larga su saturación, limitando la escalabilidad del sistema de VOD. Sin embargo, con la utilización del mirroring distribuido, la mayor parte de esta carga es gestionada por los propios servidores-Proxy mediante los mirrors parciales.

Por otro lado, a medida que se incrementa el número de redes locales que componen el sistema, también se incrementa el número y tamaño de los mirrors distribuidos. De esta forma, este aumento en la capacidad de almacenamiento permite compensar la carga de trabajo adicional asociada con los nuevos usuarios incorporados, logrando un sistema más escalable.

La utilización de mirroring distribuido no es suficiente por si mismo para resolver todos los problemas asociados con un sistema LVOD escalable. Al igual que ocurre con el caching cooperativo, la utilización de un mirror distribuido reduce el porcentaje de peticiones que se pueden servir

localmente e incrementa el número de peticiones que se tienen que atender desde servidores-Proxy remotos. Las peticiones remotas pueden llegar a necesitar el doble del ancho de banda de red de una petición local. Por lo tanto, una excesiva utilización de éstas, puede reducir considerablemente el rendimiento. Para solucionar estos inconvenientes, esta arquitectura propone la utilización conjunta de las dos técnicas: el caching y el mirroring.

Caching + Mirroring

El sistema final que utiliza esta arquitectura para la gestión de los contenidos en la arquitectura P-Tree, consiste en dividir el espacio de almacenamiento disponible en cada uno de los servidores-Proxy en dos partes, ver figura 2.4.

Una porción del espacio de disco seguirá funcionando como una caché, almacenando los videos más populares. El resto del espacio se utilizará para hacer un mirror distribuido de los contenidos que conforman el catálogo del sistema.



Figura 2.4 Esquema de almacenamiento Caching y Mirroring

Cada uno de los esquemas de gestión de contenidos en los servidores-Proxy tiene un papel bien diferenciado. El esquema de caching, se utiliza primordialmente para incrementar el número de peticiones gestionadas localmente. Debido a las características especiales de los contenidos multimedia, unos pocos videos pueden recibir un gran porcentaje de las peticiones de los usuarios, lo cual permite que la política de caching alcance una alta eficiencia con muy pocos recursos. El principal defecto de este esquema de gestión es la replicación de contenidos inherente a la política y su escasa efectividad para la gestión de las peticiones remotas.

El esquema de mirroring esta orientado a reducir la distancia requerida para atender las peticiones remotas, reduciendo el ancho de banda de red consumido por éstas.

Con ambos esquemas, trabajando en colaboración se consigue alcanzar los siguientes objetivos:

1. Reducir la carga del servidor principal. Al utilizar el esquema de caché, incrementamos el número de peticiones servidas localmente, por lo

tanto se reduce el volumen de carga que debe soportar el resto de servidores del sistema. Además, gracias al mirror distribuido, todas aquellas peticiones generadas en nodos situados más allá de una determinada distancia del servidor serán atendidas por los mirrors y no alcanzarán al servidor principal, reduciéndose así su saturación.

2. Reducir la distancia media de servicio. El mirroring distribuido permite reducir la distancia de servicio de aquellas peticiones que han fallado en su propio Proxy. A medida que se incrementa la distancia de servicio también se incrementa el tamaño del mirror distribuido y, por tanto, la probabilidad de acierto también es mayor.
3. Mayor flexibilidad del sistema LVOD. La combinación de dos esquemas de gestión tan complementarios nos permite adaptar la distribución del almacenamiento entre ambos en función de los recursos disponibles y las necesidades del sistema.
4. Incrementar la tolerancia a fallos. El fallo de uno de los nodos del árbol no impide que el resto del sistema pueda seguir funcionando gracias a la utilización de los mirrors distribuidos.

2.2.4 FUNCIONALIDAD DE LA ARQUITECTURA PROXY-TREE

Los servidores-Proxy descritos anteriormente tienen un comportamiento similar a los servidores-Proxy de Internet, los cuales monitorizan las

peticiones de los usuarios locales, capturando y sirviendo directamente aquellas peticiones que hacen referencia a contenidos almacenados en su caché.

La gestión centralizada permite que la implementación de los servidores-Proxy sea más sencilla, pero implica una mayor saturación del servidor principal y una menor tolerancia a fallos.

La integración de los servidores-Proxy en una arquitectura totalmente distribuida implica dotar a los servidores-Proxy de una nueva funcionalidad referida a la gestión distribuida del sistema. Esta nueva funcionalidad incrementa la complejidad de los servidores-Proxy hasta el punto de poder confundirlos con un servidor estándar de VOD.

La principal diferencia entre los servidores-Proxy de la arquitectura P-Tree y los servidores VOD convencionales estriba en la capacidad de almacenamiento limitada y la utilización de una política de gestión de los contenidos basados en el caching de los videos más populares.

El primer nodo de la topología se puede comportar como un nodo más de la arquitectura ó bien realizar la misma funcionalidad que el nodo principal en los sistemas de servidores-Proxy de un nivel. En este último caso, la única

diferencia del servidor principal respecto al resto de servidores-Proxy estriba en que el servidor principal es el único que mantiene una copia completa del catálogo de contenidos. En el caso de que el primer nodo de la topología tenga la misma funcionalidad que cualquier otro nodo del sistema, entonces nos encontramos ante un sistema homogéneo en el cual todos los nodos tienen los mismos componentes y realizan la misma funcionalidad. En la figura 2-5 mostramos el diagrama básico de gestión de los servidores en la arquitectura P-Tree. Como se puede apreciar, al servir una petición en la arquitectura se pueden dar dos casos diferentes: que la petición se pueda gestionar localmente por parte del Proxy asociado con el cliente ó que la petición se deba atender desde un Proxy diferente.

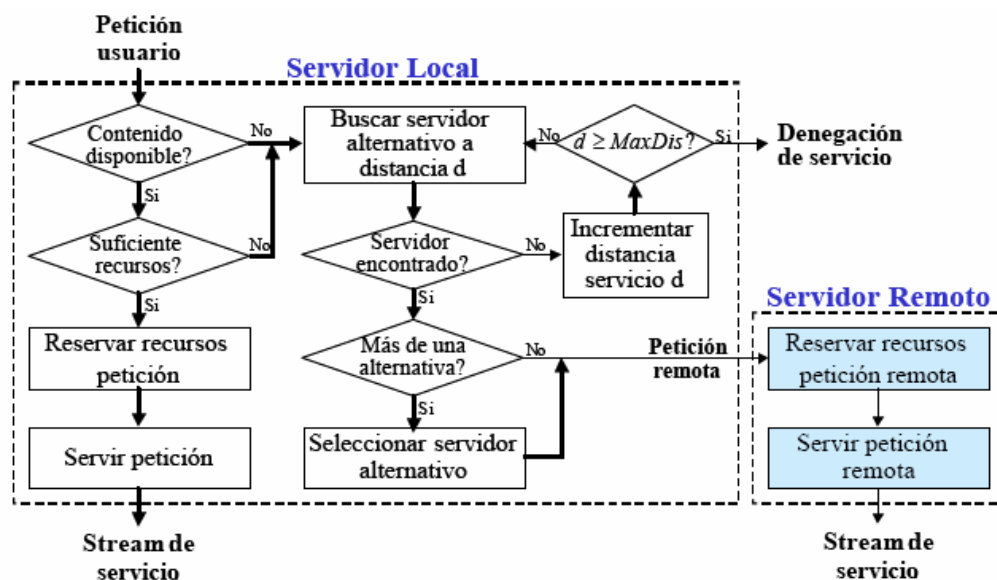


Figura 2.5 Diagrama de la gestión de los servidores-Proxy

Cuando un Proxy ó servidor local recibe una petición de un usuario debe comprobar si el contenido solicitado esta disponible en su almacenamiento y si tiene suficiente recursos (ancho de banda de servicio y de red) para asegurar el servicio de la petición con la calidad de servicio requerida. En el caso que todos los requisitos se cumplan, la petición se atiende por el servidor local a través de la red local y la red del cliente.

Si por cualquier razón (el servidor local no dispone del contenido pedido, ó no hay suficiente ancho de banda de red ó servicio) la petición no se puede atender localmente y es necesario redirigirla hacia otro servidor que si pueda servirla (petición remota).

La selección del servidor alternativo se realizará teniendo en cuenta la distancia del servidor remoto. Primero se comprobará si hay algún candidato entre los servidores adyacentes (a distancia-1 desde el servidor origen), si no es el caso, se analizan los servidores situados a distancia-2 y así sucesivamente hasta que se localiza algún servidor que pueda atender la petición. Al dar más prioridad a los servidores cercanos, se busca reducir el ancho de banda de red consumido por las peticiones remotas ya que los recursos de red consumidos por este tipo de peticiones puede llegar a ser el

doble ó incluso el triple (dependiendo de la distancia entre los dos servidores involucrados) del requerido por una petición local.

La búsqueda de un servidor alternativo requiere disponer de información sobre el estado (contenidos y recursos disponibles) de los servidores adyacentes. Este estado consistirá en información sobre los contenidos almacenados en el servidor remoto, ancho de banda de servicio disponible y ancho de banda disponible en la red local del servidor remoto. Esta información se puede obtener mediante un proceso de encuesta ó bien manteniendo de forma local el estado de cada uno de los servidores adyacentes. Seguidamente se explicará el funcionamiento de ambos métodos.

Mediante el método de encuesta, cada vez que se necesita redirigir una petición, se envía un mensaje de control a cada uno de los servidores situados a una misma distancia del servidor origen. El mensaje de control se utilizará para preguntar a los distintos servidores remotos si tienen suficientes recursos para atender la petición. Los servidores remotos contestan al servidor origen afirmativamente ó negativamente, adjuntando la información precisa para que el servidor local pueda tomar una decisión.

En el caso que todos los servidores remotos de una determinada distancia respondan negativamente, se repite el proceso de encuesta, esta vez con los servidores situados a la siguiente distancia. Si uno ó más servidores responden afirmativamente, el servidor local escoge uno de ellos atendiendo a diferentes criterios.

El método de encuesta puede generar un considerable aumento de los mensajes de control entre los servidores de la arquitectura. Otra alternativa, consiste en disponer de una tabla con el estado de los recursos de todos los servidores del sistema. Esta tabla se actualizará mediante información adjunta a los mensajes de control intercambiados entre los componentes de la arquitectura ó bien bajo demanda en el caso que la información disponible esté obsoleta.

Una vez localizado un servidor adecuado, se le redirige la petición del usuario. A partir de este instante el servicio de la petición se realiza directamente entre el servidor remoto y el cliente, sin que tenga que intervenir el servidor local del usuario. Únicamente, en el supuesto que al servidor local le interese guardar en su caché el contenido servido al usuario,

podrá éste intervenir en la transmisión. En este caso se utilizará una técnica de multicast para servir con el mismo stream al usuario y al servidor local. En el caso de que no haya ningún servidor adecuado se deniega la petición al usuario.

2.3 MODELO ANALÍTICO DE LAS ARQUITECTURAS BASADAS EN SERVIDORES-PROXY

Para evaluar la escalabilidad de una arquitectura de VOD es necesario estudiar como evoluciona el rendimiento del sistema, entendido como la capacidad de servicio del mismo, a medida que crece mediante la inclusión de nuevos nodos de servicio y usuarios. La realización de este análisis implica un cálculo de la capacidad de servicio del sistema y de los requisitos de cada uno de los componentes (redes / servidores) de las arquitecturas estudiadas: un nivel de servidores-Proxy y P-Tree.

Para el cálculo de los distintos parámetros de rendimiento, nos vamos a basar en la definición de un modelo analítico que nos permitirá evaluar la capacidad de servicio y la escalabilidad de la arquitectura P-Tree, y comparar los resultados obtenidos con otras aproximaciones LVOD.

El estudio analítico de la arquitectura P-Tree lo vamos a realizar de forma progresiva. Inicialmente definiremos un modelo para evaluar analíticamente la escalabilidad y el rendimiento de un sistema de servidores-Proxy de un nivel. Posteriormente adaptaremos este modelo para que tenga en cuenta un sistema jerárquico de servidores-Proxy (Proxy-Tree utilizando únicamente caching) y finalmente introduciremos el uso combinado de caching y mirroring dentro de los servidores de la arquitectura.

La evaluación analítica a los sistemas mencionados (basados en servidores-Proxy) requiere la definición de un modelo analítico con las siguientes características: Se supone la utilización de redes no segmentadas en las distintas arquitecturas, cada servidor-Proxy y red local dispondrá de un ancho de banda de B_C Mb/s, tamaño del Proxy suficiente para almacenar V_S películas de los M videos disponibles en el sistema, ancho de banda del servidor principal y de la red principal de B_P Mb/s.

En la realización de este análisis vamos a suponer la utilización de una política de unicast, de forma que cada usuario tiene asignado su propio stream de transmisión para los contenidos solicitados. Este supuesto es válido ya que el estudio de la escalabilidad y la capacidad de servicio del

sistema (entendida como el número de streams independientes que puede gestionar).

Para poder medir la eficiencia de las distintas arquitecturas LVOD, vamos a calcular el *ancho de banda efectivo del sistema*, definido como el número de streams simultáneos que se pueden servir utilizando los mismos recursos (ancho de banda de red y de servicio). El ancho de banda efectivo dividido entre el ancho de banda requerido para servir una petición, nos indicará el número de streams que es capaz de soportar la arquitectura.

Para estimar la capacidad de crecimiento del sistema, evaluaremos el volumen y la distribución del tráfico generado en el sistema. Esta medida nos aportará una idea de la limitación del sistema con respecto al número de usuarios que puede admitir y de su grado de escalabilidad.

Sistema de servidores-Proxy de un nivel

Como ya hemos comentado las arquitecturas de servidores-Proxy de un nivel, se componen de una serie de redes locales conectadas todas directamente al servidor principal mediante su red.

De forma genérica, el ancho de banda efectivo (B_e) de un sistema basado en servidores-Proxy se evaluará como el ancho de banda máximo disponible en el sistema (B_m) menos el ancho de banda adicional (B_{fp}) requerido por los fallos de los servidores-Proxy. Los fallos en los servidores-Proxy tienen que ser gestionados por el servidor principal, lo cual implica utilizar la red principal tanto como la red local. Dichas peticiones remotas requieren el doble de ancho de banda que una petición local y por lo tanto implican una reducción del ancho de banda efectivo del sistema. De este modo:

$$B_e = B_m - B_{fp} \quad (2.1)$$

Este ancho de banda máximo disponible en el sistema se obtendrá como la suma del ancho de banda de todas las redes que forman el sistema LVOD, según la siguiente expresión:

$$B_m = B_p + B_c * n \quad (2.2)$$

Siendo B_p el ancho de banda de la red principal, B_c el ancho de banda de las redes locales y n el número de redes locales.

Para obtener el ancho de banda adicional (B_{fp}) requerido por los fallos de los servidores-Proxy necesitamos calcular la probabilidad de fallo de los mismos. Debido a que los servidores-Proxy en esta arquitectura gestionan los contenidos exclusivamente mediante el esquema de caching, asumiremos que los servidores-Proxy almacenarán en su caché los V_c videos completos más populares.

Para poder definir cuales son los videos más populares del sistema VOD, es necesario modelar la frecuencia de acceso de cada uno de los videos del catálogo del sistema. Según diferentes estudios se puede analizar la frecuencia de acceso a los videos mediante una distribución *Zipf* con un grado de popularidad (*skew factor*) de z . A partir de la distribución *Zipf* y del número de contenidos del catálogo (M) se puede calcular la probabilidad de acceso de cada uno de los contenidos (f_i) mediante la siguiente expresión:

$$f_i = \frac{1}{i^z * \sum_{j=1}^M \frac{1}{j^z}} \quad (2.3)$$

A partir de la expresión la figura 2.6 muestra la distribución de las probabilidades de acceso para cada uno de los contenidos multimedia en un sistema con 100 videos y un grado de popularidad del 0.729. Se puede constatar que los contenidos están ordenados según su popularidad: los videos con un índice menor tienen una mayor frecuencia de acceso. También podemos observar que un reducido grupo de contenidos reciben la mayoría de los accesos. Por ejemplo, los 20 primeros videos reciben más del 53% de las peticiones de los usuarios.

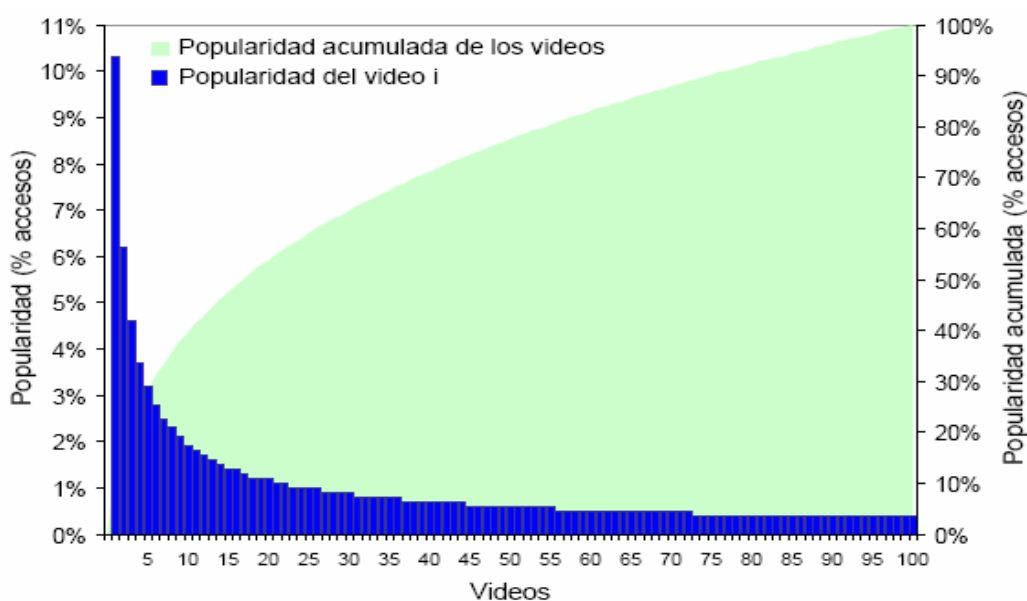


Figura 2.6 Distribución de Zipf de popularidad de videos

Mediante la expresión, podemos calcular la probabilidad de que una petición pueda ser atendida por el servidor-Proxy como la suma de las probabilidades

individuales de los V_c contenidos más populares, y mediante su inverso podemos evaluar la probabilidad de fallo al acceder a un servidor-Proxy (p_{fp}) según la siguiente formula:

$$p_{fp} = 1 - \sum_{i=1}^{V_c} f_i \quad (2.4)$$

De esta forma podemos calcular el ancho de banda adicional requerido en el sistema debido a los fallos de los servidores-Proxy como la probabilidad de fallo en cada Proxy (p_{fp}) multiplicado por el tráfico generado en todas las redes locales del sistema ($B_c * n$), es decir:

$$B_{fp} = n * B_c * p_{fp} \quad (2.5)$$

Esta expresión mide el tráfico enviado por los servidores-Proxy hacia el servidor y la red principal, y por lo tanto, identifica el ancho de banda mínimo

requerido en la red principal para que el sistema no se sature con los fallos en los servidores Proxy.

Sustituyendo las expresiones (2.2) y (2.5) en la expresión (2.1), el ancho de banda efectivo de una arquitectura de servidores-Proxy de un nivel es:

$$B_e = B_p + n * B_c - n * B_c * p_{fp} = B_p + n * B_c * (1 - p_{fp}) \quad (2.6)$$

Sistema jerárquico de servidores-Proxy

Habiendo analizado la efectividad de los sistemas basados en servidores-Proxy de un nivel, a continuación nos planteamos como podemos extender este análisis para contemplar los sistemas basados en una topología jerárquica de servidores-Proxy (arquitectura P-Tree utilizando exclusivamente caching en los servidores).

Para facilitar este estudio asumiremos una topología en árbol completa (todos los niveles están llenos) con L niveles en los cuales el orden (o) identifica el número de redes locales conectadas en cada uno de los nodos del árbol.

Del modelo anterior, el único parámetro que se ve afectado con la modificación de la topología, es el ancho de banda adicional de red requerido debido a los fallos en los servidores-Proxy de las redes locales (expresión 2.5). En esta topología, a medida que una petición no puede ser servida por el Proxy de un nivel, se accederá al siguiente y así sucesivamente. A medida que se accede al siguiente nivel de caché se incrementa el ancho de banda requerido para servir la petición. Este costo estará en función del número de niveles (distancia) que tiene que cruzar la petición desde el cliente hasta el servidor ó el Proxy que la atienda.

En una primera aproximación, para evaluar este costo (Bfp) podríamos asumir que el costo de los fallos de las caches se pueden obtener sumando los fallos producidos en cada uno de los niveles. Así, el costo adicional generado desde un nodo de un determinado nivel, se puede calcular como el porcentaje del tráfico de su red (Bc) que ha fallado en su Proxy y que es atendido desde servidores-Proxy situados a distancia 1 ($Bc * Pfp$) más el porcentaje del tráfico que no se sirve desde los servidores-Proxy de distancia 1 y que se atienden en los servidores-Proxy de distancia 2 ($Bc * Pfp * Pfp$) y así sucesivamente hasta llegar al nivel 0, en el cual se encuentra el servidor principal que será el que atenderá en última instancia la petición. El costo

generado por un nodo de un nivel se tiene que multiplicar por el número de nodos de ese nivel para obtener el ancho de banda requerido en el mismo. A través de este planteamiento, el ancho de banda requerido para atender los servidores-Proxy (B_{fp}) se define como:

$$B_{fp} = B_c * \sum_{l=1}^L o^l \sum_{d=1}^l P_{fp}^d \quad (2.7)$$

Sin embargo, esta expresión no es realista debido a que asume que la probabilidad de acierto en los servidores-Proxy de nivel superior es la misma que en el Proxy de su red local. Esto es totalmente falso, las caches de todos los servidores-Proxy tendrán probablemente los mismos contenidos (los más populares), generando un elevado nivel de redundancia de la información. La replicación de contenidos entre las caches repercute negativamente en la probabilidad de acierto de los servidores-Proxy, ya que si una petición ya ha fallado en un Proxy, muy probablemente volverá a fallar en los restantes. Por lo tanto es más factible que la probabilidad de acierto en los servidores-Proxy situados a distancia mayor de 1 de donde se generó la petición, sea prácticamente nula.

Teniendo en cuenta la reflexión anterior, una forma más realista para calcular el costo adicional en este sistema jerárquico es asumir que las peticiones que no pueden ser atendidas por su Proxy local, lo son desde el servidor principal y por tanto, habrá que considerar un costo proporcional a la distancia (l) que le separa del servidor principal. Por lo tanto, en este caso el ancho de banda requerido para los fallos vendrá definido por:

$$B_{fp} \approx B_c * \sum_{l=1}^L o^l * p_{fp} * l \quad (2.8)$$

Esta fórmula nos demuestra que no tiene sentido esta estructura jerárquica de servidores-Proxy, ya que no aumenta la probabilidad de acierto de los Proxy y sí se incrementa la distancia que tienen que recorrer las peticiones cuando se falla, con el consiguiente aumento de su penalización sobre el ancho de banda efectivo del sistema.

Utilizando la expresión (2.8), en la figura 2-8 mostramos el ancho de banda de red requerido para atender a los fallos en los servidores-Proxy (B_{fp}) a medida que el sistema crece. Podemos comprobar que este ancho de banda

requerido crece mucho más rápido que el ancho de banda de red total ($n \cdot Bc$) disponible en el sistema, lo cual demuestra la escasa viabilidad de esta aproximación para una arquitectura escalable de LVOD. Las características que se tomaron para realizar este gráfico son:

$Bc=100$ $L=2$ $n=12$ $pf=29\%$

Los datos presentados a continuación se hacen referencia a la tabla 2.2

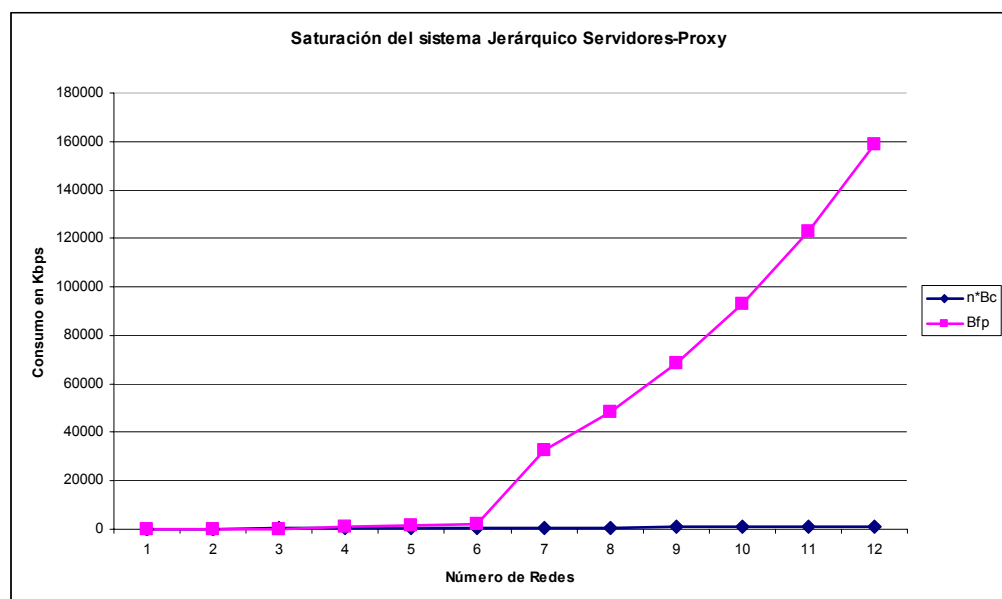


Figura 2.8 Saturación del Sistema Jerárquico Servidores-Proxy

Tabla 2.2 Ancho de Banda Servidor-Proxy		
N. De redes	n*Bc	Bfp
1	100	29
2	200	58
3	300	87
4	400	1044
5	500	1595
6	600	2262
7	700	32886
8	800	48488
9	900	68382
10	1000	93090
11	1100	123134
12	1200	159036

A continuación vamos a estudiar la viabilidad de esta topología mediante la inclusión de un mirror en los servidores-Proxy con el fin de reducir la distancia de servicio de las peticiones remotas y disminuir la saturación del servidor principal.

Arquitectura Proxy-Tree

Bajo el enfoque utilizado para la gestión de los servidores-Proxy en la arquitectura P-Tree, su espacio de almacenamiento no se modifica sino que se distribuye entre dos esquemas diferentes: un porcentaje se continúa

gestionando como una caché para almacenar los contenidos más populares y el resto se utiliza como mirror de una porción de los videos del catálogo del sistema.

Bajo este esquema la probabilidad de acierto de las peticiones que han fallado en su propio servidor-Proxy es mayor. A medida que se aumenta la distancia de servicio también se incrementa el tamaño del mirror distribuido y por lo tanto la probabilidad de que la petición remota sea atendida.

En la figura 2.9 mostramos, sobre la topología de la arquitectura P-Tree, el número de servidores-Proxy alternativos a medida que se incrementa la distancia de servicio. Podemos ver como una petición que falle en su servidor local situado en el segundo nivel de la topología puede acceder a 3 servidores-Proxy situados a distancia 2 (del usuario), si éstos no pueden atender la petición entonces tendrá que acceder a los 6 servidores-Proxy situados a distancia 3 y así sucesivamente hasta que su petición sea atendida por un Proxy ó se alcance el servidor principal.

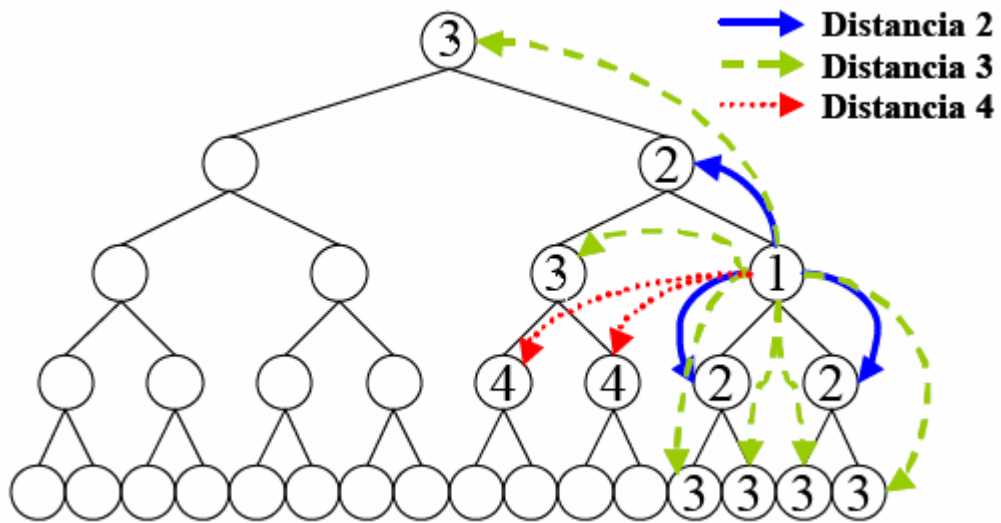


Figura 2.9 Esquema de distribución jerárquica de los fallos en los proxies

Para facilitar el análisis asumiremos que la distribución de las películas en cada uno de los mirror se realiza de forma equitativa y de esta forma la probabilidad de acierto en los mirrors es constante e independientemente de su posición en la jerarquía y se puede aproximar dividiendo la probabilidad total (1) entre el número de contenidos del catálogo del sistema:

$$P_{am} = \frac{1}{N} \quad (2.9)$$

A partir de esta suposición, la probabilidad de fallo de una petición realizada en el nivel l y accediendo a todos los servidores-Proxy situados a una distancia no superior a d , viene dado por la expresión:

$$p_{fm}(l, d) = 1 - (p_{ac} + p_{am} * V_M * Nd(l, d)) \quad (2.10)$$

Siendo:

P_{ac} Probabilidad de acierto en la caché del Proxy ($1-P_{fp}$).

P_{am} Probabilidad media de acierto en un video del mirror del Proxy.

V_m Capacidad del Proxy reservada para el mirror (en videos).

$N_d(l, d)$ Número de servidores a distancia no superior a d del nivel l .

Esta expresión nos indica que la probabilidad de acierto a distancia 1 (Proxy situado en la propia red local en donde se generó la petición) es la probabilidad acumulada de que la petición se pueda servir desde la caché ó desde el mirror del Proxy, mientras que para distancias mayores solo se considera la probabilidad de acierto de los mirrors en los Proxy situados a la

misma distancia (asumimos que la efectividad de la caché para atender peticiones remotas es nula).

De esta forma, el ancho de banda necesario para los fallos en los servidores-Proxy (caché+mirror) sería el siguiente:

$$B_{fp} = B_c * \sum_{v=1}^L o^v * \sum_{d=0}^v p_{fm}(v, d) \quad (2.11)$$

Sustituyendo esta expresión en la expresión (2.1) se obtiene que el ancho de banda efectivo es:

$$B_e = B_m - B_c * \sum_{v=1}^L o^v * \sum_{d=0}^v p_{fm}(v, d) \quad (2.12)$$

2.4 ANÁLISIS DE LA ESCALABILIDAD DE ARQUITECTURAS BASADAS EN SERVIDORES-PROXY

En este apartado vamos a utilizar el modelo analítico definido anteriormente para evaluar la escalabilidad de la arquitectura basada en servidores-Proxy de un nivel.

Durante éste análisis se asume que el sistema homogéneo de LVOD en el cual todos los servidores y redes locales tienen el mismo ancho de banda. Las principales características son las siguientes: el ancho de banda de la red principal y de las redes locales es de 100 Mb/s, se dispone de 100 videos diferentes y un patrón de acceso modelado mediante una distribución Zipf con un factor de skew de 0.729, grado de popularidad más frecuentemente utilizado para modelar los sistemas de VOD.

2.4.1 ESCALABILIDAD DE LA ARQUITECTURA DE UN NIVEL DE SERVIDORES-PROXY

Usando las expresiones (2.5) y (2.6), la figura 2.10 muestra los principales parámetros de comportamiento de los sistemas basados en servidores-Proxy de un nivel cuando el sistema escala. Las gráficas muestran el ancho de

banda efectivo del sistema (gráficas 1, 3 y 5) y el tráfico recibido por la red principal (gráficas 2, 4 y 6) a medida que el sistema crece (mediante el incremento del número de servidores-Proxy conectados al sistema).

Los datos que están graficados se los representa en las tablas 2.3, 2.4 y 2.5

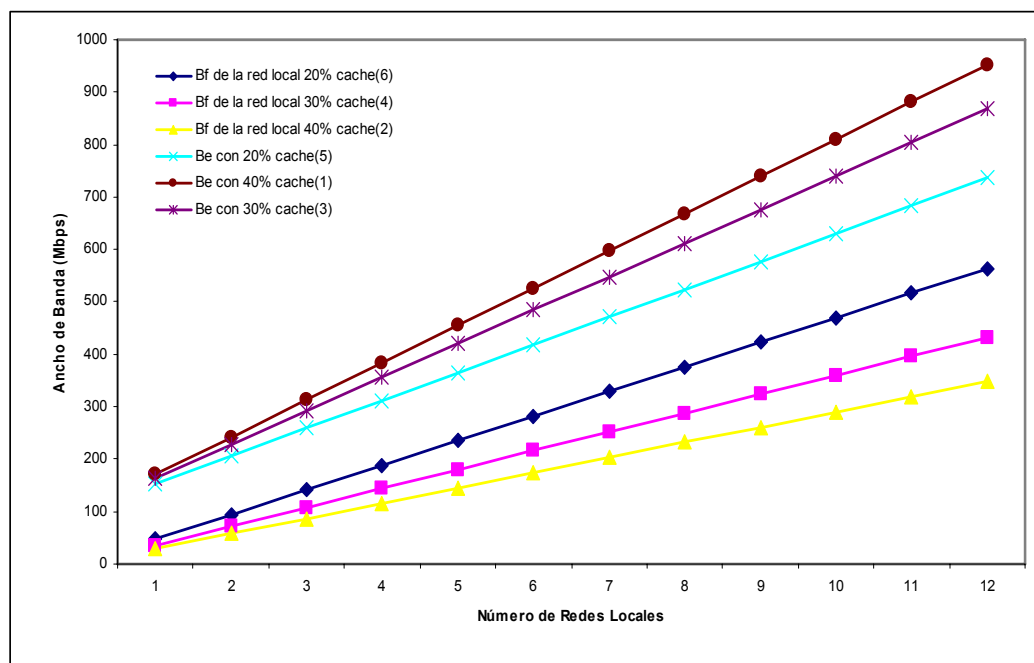


Figura 2.10. Escalabilidad en los sistemas Proxy

Tabla 2.3 Probabilidades de Acierto y Fallo		
	Pacierto	Pfallo
Bfp20%	53	47
Bfp30%	64	36
Bfp40%	71	29

Tabla 2.4 Ancho de Banda de Fallo y Acierto			
N. Redes	Bfp20%	Bfp30%	Bfp40%
1	47	36	29
2	94	72	58
3	141	108	87
4	188	144	116
5	235	180	145
6	282	216	174
7	329	252	203
8	376	288	232
9	423	324	261
10	470	360	290
11	517	396	319
12	564	432	348

Tabla 2.5 Ancho de Banda de la Red Principal			
N. Redes	Bfp20%	Bfp30%	Bfp40%
1	153	164	171
2	206	228	242
3	259	292	313
4	312	356	384
5	365	420	455
6	418	484	526
7	471	548	597
8	524	612	668
9	577	676	739
10	630	740	810
11	683	804	881
12	736	868	952

Mediante este estudio queremos analizar el comportamiento del sistema, en concreto la saturación de la red, cuando crece. Para poder medir la influencia de la capacidad de almacenamiento de los servidores-Proxy sobre la escalabilidad, se han obtenido resultados para distintos tamaños de caché. Así, la figura muestra los resultados utilizando servidores-Proxy con capacidad para almacenar el 20%, 30% y 40% de los videos del sistema.

En los resultados mostrados podemos comprobar como a medida que aumentamos el número de servidores-Proxy de la arquitectura, se aumenta el ancho de banda efectivo del sistema (gráficas 1, 3 y 5) y por ende el número de usuarios que se pueden servir.

Sin embargo, éste incremento en la capacidad del sistema se consigue a costa de incrementar los requisitos de ancho de banda de la red y del servidor principal. Así, podemos observar en las gráficas 2, 4 y 6, que el ancho requerido por la red principal se incrementa linealmente con el número de redes locales del sistema.

El aumento de los requisitos de los componentes centralizados de la arquitectura (servidor principal y red principal) ocurre independientemente de la capacidad de almacenamiento de los servidores-Proxy. Tal y como podemos observar en la figura, el aumento de la capacidad de almacenamiento únicamente reduce la pendiente de crecimiento de la saturación de éstos componentes. Además, la efectividad del nuevo almacenamiento añadido es cada vez menor debido a que los nuevos contenidos incluidos en la caché tienen cada vez menor frecuencia de acceso.

Como consecuencia de estos resultados podemos concluir que la escalabilidad de las arquitecturas de servidores-Proxy de un nivel esta limitada por la capacidad de servicio del servidor principal y el ancho de banda de la red principal.

Con respecto a la eficiencia de estas arquitecturas hacemos constar la importancia del tamaño de la caché sobre el rendimiento final del sistema. Utilizando servidores-Proxy con capacidad para el 20% de las películas (gráficas 5 y 6), la mayor parte del ancho de banda efectivo obtenido (736 Mb/s para un sistema con 12 redes locales), proviene del ancho de banda de

la red principal (564 Mb/s). En sistemas con servidores-Proxy con una capacidad pequeña, resulta más rentable la conexión de los clientes directamente a la red principal que utilizar redes locales con servidores-Proxy. Un sistema centralizado con los mismos recursos tendría un ancho de banda efectivo de 564 Mb/s, con el ahorro del costo de todas las redes locales.

Sin embargo, cuando la capacidad de los servidores-Proxy es lo suficientemente grande (tamaños del 30% y 40%), se incrementa el acierto en la caché del Proxy y se reduce el acceso a la red del servidor principal.

Por ejemplo, utilizando servidores-Proxy con tamaño del 40% con 12 redes locales (gráficas 1 y 2) se obtienen anchos de banda efectivos de 952 Mb/s, requiriendo solamente una red principal con capacidad de 348 Mb/s.

2.4.2 ESCALABILIDAD DE LA ARQUITECTURA P-TREE

Para demostrar que la arquitectura escogida es escalable, tenemos que garantizar que la ampliación del sistema no modifica los requerimientos

(ancho de banda) de los elementos existentes ó bien que el aumento de los requerimientos esta acotado.

Para cumplir este objetivo tenemos que verificar que el volumen de trabajo / tráfico soportado por cualquiera de los componentes del sistema queda acotado a medida que el sistema escala.

Ahora bien, en los sistemas jerárquicos el elemento que puede recibir una mayor carga es el que se encuentra situado en el nivel superior de la jerarquía, es decir el servidor y la red principal en nuestro caso. Por lo tanto, inicialmente debemos demostrar que la red principal y el servidor principal no se saturan cuando el sistema crece.

En la figura 2.11 mostramos el ancho de banda efectivo (B_e) según la expresión (2.11) y la sobrecarga de la red principal (S_p) definida por la expresión (2.12), para sistemas basados en topologías con árboles binarios y utilizando servidores-Proxy con una capacidad para el 30% y 40% de los contenidos del sistema.

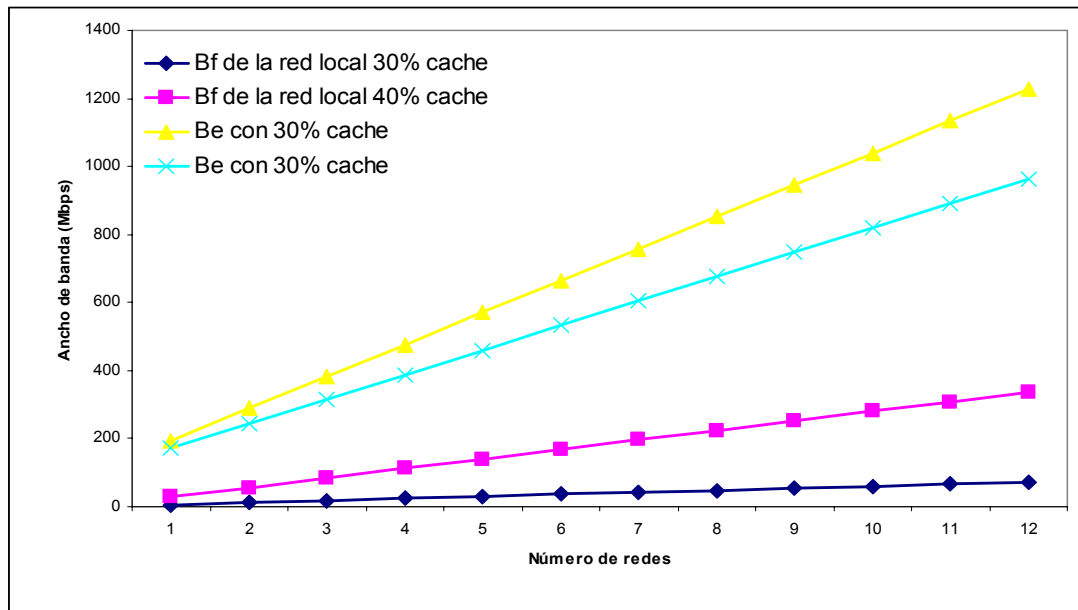


Figura 2.11. Escalabilidad de la red principal en la arquitectura P-Tree

N. Redes	Bfp30%	Bfp40%
1	6	28
2	12	56
3	18	84
4	24	112
5	30	140
6	36	168
7	42	196
8	48	224
9	54	252
10	60	280
11	66	308
12	72	336

Tabla 2.7 Ancho de Banda de la Red Principal		
N. Redes	Bfp30%	Bfp40%
1	194	172
2	288	244
3	382	316
4	476	388
5	570	460
6	664	532
7	758	604
8	852	676
9	946	748
10	1040	820
11	1134	892
12	1228	964

Si analizamos el comportamiento de la arquitectura se puede observar que a medida que se incrementan el número de redes locales conectadas, el incremento del ancho de banda de las redes locales no es tan pronunciado, lo que permite una mejor escalabilidad que en un sistema basado en servidores Proxy. Como ejemplo: con 12 redes tenemos 72 Mb/s con servidores-Proxy con tamaño del 30% y en 336 Mb/s para servidores-Proxy con un tamaño del 40%, mientras que con servidores Proxy se tienen 432 con un caché del 30 % y 348 con un caché del 40%.

2.5 EVALUACIÓN DEL RENDIMIENTO DE LA ARQUITECTURA PROXY-TREE

Una vez verificada la escalabilidad de la arquitectura P-Tree, es necesario evaluar su rendimiento, comparándolo con otras arquitecturas similares. Este análisis se realizará utilizando el modelo analítico desarrollado en el apartado anterior, en concreto, el ancho de banda efectivo se calculará utilizando la expresión (2.12).

Antes de poder evaluar el rendimiento de la arquitectura P-Tree es necesario ajustar una serie de parámetros para poder optimizar la utilización de sus recursos. Los parámetros principales que hay que tener en cuenta en la arquitectura son el orden del árbol de la topología (binario, terciario, etc.) y la distribución del almacenamiento del Proxy entre los esquemas de caché y mirroring.

2.5.1 ANÁLISIS DEL ORDEN DEL ÁRBOL

En la tabla 2.8 mostramos el rendimiento de varios sistemas P-Tree con una capacidad para el 40% de los videos del catálogo y utilizando diferentes órdenes para el árbol de la topología.

Tabla 2.8. Rendimiento sistemas P-Tree con diferentes topologías

Sistema P-Tree	Ancho de banda efectivo	Ancho banda red principal
Árbol binario	89.144 Mb/s	75 Mb/s
Árbol terciario	88.791 Mb/s	84 Mb/s
Árbol cuaternario	87.418 Mb/s	90 Mb/s

Como podemos observar el menor ancho de banda utilizado es cuando se tiene un árbol binario. Pero en nuestro caso el servicio se dará en un árbol terciario ya que existen tres ciudades en las que actualmente se piensa brindar el servicio, pero al momento de aumentar la escalabilidad se lo hará de forma binaria así la utilización de recursos será óptima.

2.5.2 DISTRIBUCIÓN DEL ALMACENAMIENTO DE LOS SERVIDORES-PROXY

Uno de los parámetros que tiene más incidencia sobre el rendimiento de la arquitectura P-Tree es la distribución de la capacidad del Proxy entre los dos esquemas de gestión de contenidos. Una distribución incorrecta puede afectar al rendimiento del sistema. Si no se asigna la suficiente capacidad al esquema de mirroring, puede implicar que las peticiones remotas deban servirse desde servidores-Proxy lejanos (requiriendo más recursos de ancho

de banda red). Por el contrario, si no se le asigna la suficiente capacidad a la caché se reducirá el número de peticiones que se atienden localmente y consecuentemente se incrementa el número de peticiones remotas. Por lo que para nuestro diseño hemos escogido un caché del 40% y un mirroring distribuido del 30% dejando un 30% al servidor principal obteniendo un mejor nivel de performance de la red.

2.5.3 ANÁLISIS DE RENDIMIENTO DE LA ARQUITECTURA P-TREE

Una vez seleccionado la distribución óptima de almacenamiento dedicada a cada uno de los esquemas de gestión de los contenidos, ya podemos evaluar el rendimiento de las arquitecturas P-Tree.

En la figura 2.12 se puede observar que las prestaciones que se obtienen para este sistema, entendida como el ancho de banda efectivo del mismo. Para la arquitectura de servidores-Proxy de un nivel suponemos que la red principal del sistema en todo momento tiene la suficiente capacidad como para atender todos los fallos de los servidores-Proxy. Por lo tanto, el ancho

de banda de esta red crece cuando se amplia el sistema y podemos comparar las prestaciones obtenidas por nuestra arquitectura con las obtenidas con un sistema clásico de Proxy de un nivel.

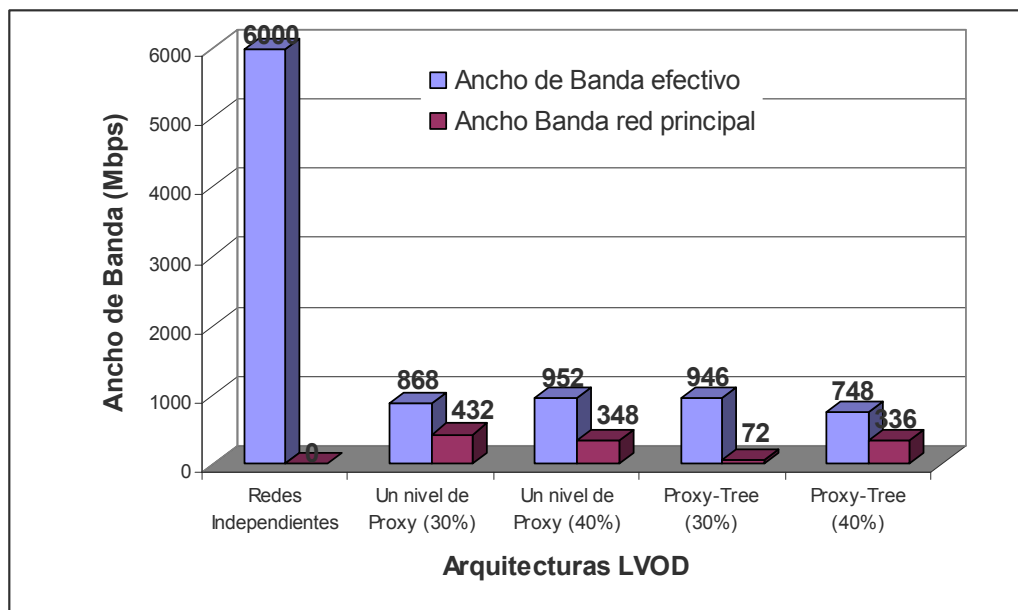


Figura 2.12 Rendimiento de las arquitecturas de LVOD

A primera vista se puede observar una pérdida de prestaciones de la arquitectura P-Tree. La reducción del ancho de banda efectivo del sistema jerárquico con respecto al mismo sistema pero con un nivel de servidores-Proxy depende del tamaño de los servidores-Proxy. Así, la pérdida de

rendimiento se puede cuantificar en un 21% (952 Mb/s versus 748 Mb/s) para un sistema con 12 redes locales y un Proxy del 40%.

La diferencia de prestaciones entre ambas arquitecturas se puede justificar por el nivel de distribución de la gestión de ambos sistemas. La arquitectura de servidores-Proxy de un nivel únicamente esta compuesta por 2 niveles, el primer nivel compuesto por los distintos servidores-Proxy y el segundo por el servidor principal. Estos dos niveles permiten limitar la distancia de servicio máxima de las peticiones a 2 redes (peticiones atendidas desde el servidor principal). De esta forma, se obtiene una arquitectura más centralizada y por lo tanto, más eficiente.

En cambio, la gestión de los sistemas P-Tree esta más distribuida entre los distintos servidores-Proxy y la arquitectura puede estar constituida por más niveles.

Ambas características provocan que la distancia de servicio pueda ser mayor que la correspondiente de la arquitectura de servidores-Proxy de un nivel.

Además, la utilización del esquema de gestión de mirroring reduce la capacidad del esquema de caching lo cual reduce el porcentaje de peticiones que se pueden servir localmente (comparadas respecto a los servidores-Proxy de un nivel, que dedican todo su almacenamiento a caching).

Los resultados obtenidos por la arquitectura de servidores independientes identifica el rendimiento óptimo (6.000 Mb/s) para un sistema en el cual, al igual que ocurre con los sistemas centralizados, todas las peticiones se pueden servir localmente. Comparando los resultados obtenidos mediante la arquitectura de servidores independientes y las arquitecturas con los contenidos distribuidos se puede comprobar que estas últimas ofrecen una menor efectividad.

Esta reducción en el rendimiento de las arquitecturas distribuidas respecto a las centralizadas ó de servidores independientes es exclusivamente achacable a los recursos de red requeridos por los sistemas distribuidos debido a las peticiones que no se pueden servir localmente (mayor distancia de servicio).

2.5.4 REQUERIMIENTOS DE ANCHO DE BANDA DE RED

Otro factor importante a la hora de elegir un sistema como el propuesto para nuestro proyecto en el Grupo TvCable, es el tamaño que se tiene que utilizar para las redes del sistema.

ARQUITECTURA A NIVEL NACIONAL

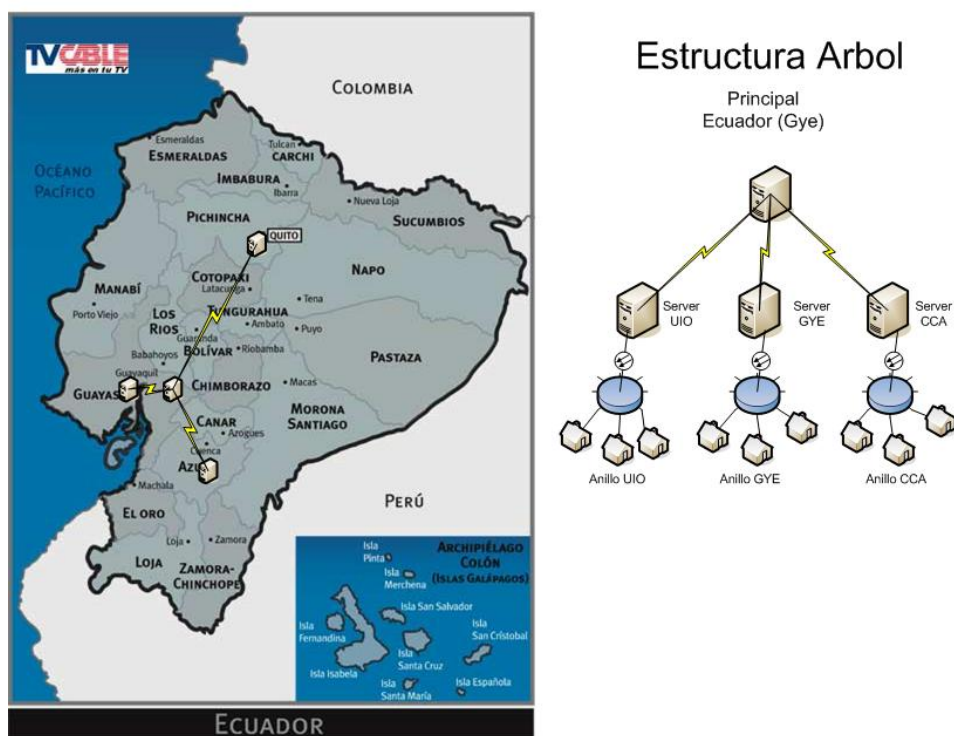


Figura 2.13 Distribución de los servidores a nivel Nacional

En la figura 2.13 se puede observar nuestra propuesta para el diseño de la red en forma de árbol. Actualmente estos enlaces ya existen y pertenecen a

la infraestructura del grupo TvCable como tal, por lo que la implementación de este nuevo servicio a nivel de utilización de enlaces de interurbanos no representa mayor complicación y es por esta razón que los costos en los que incurriría hacer enlaces interurbanos no se aplican a nuestro proyecto, así como los anchos de bandas requeridos para la comunicación entre los servidores.

2.5.5 COMPARACIÓN DE RESULTADOS ENTRE DIFERENTES ARQUITECTURAS LVOD

Por último, en la tabla 2.9 resumimos los principales parámetros de un sistema de LVOD, utilizando las siguientes arquitecturas: centralizada, servidores independientes, un nivel de servidores-Proxy y Proxy-Tree.

El cálculo del número máximo de usuarios independientes que soporta el sistema se ha realizado suponiendo la utilización de políticas de servicio unicast y la utilización de los contenidos multimedia con el formato MPEG-1 y una duración aproximada de 90 minutos, por lo tanto para cada video se

requiere un ancho de banda 1.5 Mb/s y almacenamiento de aproximadamente 1GByte.

Tabla 2.9 Principales de las distintas arquitecturas LVOD

Parámetros	Centralizado	Servidores Independientes	Un nivel de servidores-proxy	Proxy-Tree
<i>Redes locales</i>	0	0	12	12
<i>Redes principales</i>	1	12	1	1
<i>Servidores</i>	1	12	1	1
<i>Servidores-proxy</i>	0	0	12	12
<i>Capacidad Servidores-proxy</i>	0	0	40%	40%
<i>Almacenamiento requerido (100 videos x 1GB)</i>	100 GBytes	1.200 GBytes	480 GBytes	480 GBytes
<i>Ancho banda de la red principal</i>	6.000 Mb/s	0 Mb/s	348 Mb/s	336 Mb/s
<i>Ancho banda redes locales</i>	0 Mb/s	500 Mb/s	500 Mb/s	500 Mb/s
<i>Ancho de banda efectivo</i>	6.000 Mb/s	6.000 Mb/s	952 Mb/s	748 Mb/s
<i>Número máximo usuarios soportados</i>	90.000	90.000	60.000	59.333
<i>Escalabilidad</i>	limitada	ilimitada	limitada	ilimitada

De los resultados obtenidos podemos deducir que de las principales arquitecturas escalables (sin tener en cuenta los sistemas centralizados), la que ofrece los mejores resultados es la de servidores independientes, pero a costa de incrementar el volumen de almacenamiento requerido hasta alcanzar los 1.2 Terabytes.

Las dos arquitecturas basadas en servidores-Proxy, obtienen resultados parecidos (952 Mb/s y 748 Mb/s, para el sistema de servidores-Proxy de un nivel y P-Tree respectivamente). La arquitectura de un nivel de servidores-Proxy se caracteriza por requerir todavía un servidor central complejo y una red principal con un ancho de banda de 348 Mb/s.

Mientras, la arquitectura P-Tree permite una **mayor escalabilidad** y únicamente requiere la utilización de **redes locales y servidores-Proxy de menor complejidad**.

Con respecto a la arquitectura de servidores independientes el sistema P-Tree permite una **mayor tolerancia a fallos** del sistema, un mejor **balanceo de la carga** entre los distintos servidores (si un servidor esta saturado, puede redirigir sus peticiones hacia los servidores-Proxy vecinos) y permite un **mayor potencial de efectividad de las políticas multicast** (el número de usuarios que pueden acceder a un determinado contenido es del orden de 4 veces superior comparado con las arquitecturas de servidores independientes).

Capítulo 3

3. SEGURIDAD DE RED

En este capítulo se analiza el protocolo RSTP como protocolo de transporte para los paquetes de video así como la codificación MPG-2 que se utiliza para compresión, la autenticación del equipo Terminal con el servidor por medio de los diferentes dispositivos de la red tales como NC, OM, RPD y DAC.

3.1 INTRODUCCIÓN

La conectividad es lo más esencial para el diseño y selección de hardware ya que dependiendo de esto se podrá utilizar la red HFC que actualmente existe en el headend del Grupo TVCable así como también cada nodo de su infraestructura. Si lo analizamos globalmente se necesitan dos equipos que son los más importantes para poder realizar la descarga de video estos son: El servidor de Video donde se encontrará alojado el contenido multimedia que se ofrecerá a los clientes y del otro lado se encuentra el equipo terminal ubicado en cada casa el cuál recibirá un requerimiento y lo enviará hacia el servidor para que éste empiece la descarga en tiempo real. A continuación se describirá las características más importantes así como porqué se escogieron los equipos para realizar el diseño de este proyecto. Algo importante a resaltar el momento de hablar de la conectividad también lo es el protocolo a utilizarse ya que basado en este es la forma en que los paquetes son enviados hacia cada equipo terminal ya que necesitan hablar “el mismo idioma” para poder intercambiar información.

3.2 PROTOCOLO DE TRANSPORTE RSTP

Real Streaming Time Protocol (RSTP) es un protocolo a nivel de aplicación para el control sobre la entrega de datos con propiedades de tiempo real.

RSTP controla la demanda y entrega de datos en tiempo real, tanto audio como video. La fuente de los datos puede contener ambos, datos en vivo o videos guardados. La intención del protocolo es el control de los datos en la entrega de múltiples sesiones, mantiene un medio escogiendo diferentes canales para la entrega tales como UDP, multicast UDP y TCP, los mecanismos de entrega de los medios está basado en RTP

3.2.1 PROPÓSITO DEL PROTOCOLO RSTP

Real Streaming Time protocol (RSTP) establece y controla flujos de datos solos o varios de tiempo sincronizado tales como audio y video. Los paquetes stream no se entregan por si solos, ellos van entrelazados con los medios multimedia tantos como sean posibles. En otras palabras RSTP actúa como un “control remoto de la red” para los servidores multimedia.

Una sesión RSTP no está de ninguna manera atada a una conexión a nivel de transporte como una conexión TCP. Durante una sesión de RSTP, un cliente RSTP puede abrir y cerrar muchas conexiones de transporte al servidor al emitir demandas de RSTP. Alternativamente, se puede usar un protocolo de transporte orientado a conexión tal como UDP.

Los paquetes multimedia controlados por RSTP pueden usar RTP, pero el funcionamiento de RTSP no depende de los mecanismos de transporte usados para llevar los paquetes. El protocolo es intencionalmente similar en sintaxis y funcionamiento a HTTP, para que los mecanismos de la extensión a HTTP también puedan funcionar en la mayoría de los casos se ha agregado RSTP. Sin embargo, RSTP difiere en aspectos importantes de HTTP tales como:

- RSTP introduce nuevos métodos y tiene un identificador de protocolo diferente.
- Un servidor RSTP necesita mantener estado por defecto en casi todos los casos, opuesto a HTTP que es sin estado.
- Un servidor RSTP y el cliente puede emitir demandas.
- Los datos son llevados fuera de banda por un protocolo diferente
- RTSP se define para usar ISO 10646 (UTF-8) en lugar de ISO 88559-1 consistente con HTTP.

El protocolo soporta las diferentes operaciones:

Invitación de servidores multimedia a una conferencia:

Un servidor multimedia puede ser “invitado” para unirse a una conferencia existente, o para tocar contenidos multimedia en una presentación o para grabar todos o un subconjunto de los contenidos multimedia en una presentación. Este modo es útil para distribuir aplicaciones de enseñanza.

Adición de contenidos multimedia en una presentación existente:

Particularmente para las presentaciones en vivo, es útil si el servidor puede decirle al cliente sobre contenido multimedia adicional que se vuelve disponible.

3.2.2 PROPIEDADES DEL PROTOCOLO.

RSTP tiene las siguientes propiedades:

Extendible:

Nuevos métodos y parámetros pueden ser fácilmente agregados a RSTP de ello los fabricantes se aprovechan para agregar campos que permiten mejorar la seguridad al momento de realizar una descarga.

Seguridad:

RSTP re-usa mecanismos de seguridad Web. Todos los mecanismos de autenticación de son directamente aplicables. Uno también puede rehusar los mecanismos de seguridad de la capa de transporte o red.

Transporte independiente:

RSTP puede usar tanto un datagrama de un protocolo inestable como el UDP (no orientado a conexión), o un datagrama de un protocolo estable (RDP) o un protocolo de paquetes como el TCP (orientado a conexión) cuando se lleva a cabo la fiabilidad a nivel de aplicación.

Capacidad de multi-servidor:

Cada paquete multimedia dentro de una presentación puede residir en un servidor diferente. El cliente establece automáticamente varias sesiones con diferentes servidores multimedia. La sincronización de los medios es realizada a nivel de transporte.

Separación de los paquetes de control e iniciación de conferencia:

Los paquetes de control se niegan a ser invitados por un servidor multimedia a una conferencia. El único requisito es que el protocolo de iniciación de la conferencia provea o se lo pueda usar para crear un identificador único de la conferencia.

Conveniente para las aplicaciones profesionales:

RSTP soporta corrección a nivel de paquete a través de SMPTE para permitir edición digital remota.

Presentación neutral:

El protocolo no impone una presentación en particular y puede llevarse al tipo de formato a ser usado.

Proxy y cortafuego amigable:

El protocolo deber ser manejado por ambas capas la de aplicación y la de transporte para los cortafuegos. Un cortafuego puede necesitar un método de inicio para abrir un agujero para los paquetes multimedia UDP.

Apropiado control del servidor:

Si un cliente puede empezar un contenido multimedia, debe poder detener dicho contenido. Los servidores no deben empezar la entrega de los contenidos multimedia de tal manera que los clientes no puedan detenerla.

Negociación de transporte:

El cliente puede negociar la prioridad del método del transporte a necesitarse para procesar paquetes multimedia continuos.

Negociación de capacidad:

Si las características básicas son inválidas, debe haber algunos mecanismos para que el cliente pueda determinar cuales métodos no deben ser implementados. Esto le permite al cliente presentar al usuario una interfaz apropiada. Por ejemplo, si buscar no es permitido, la interfaz del usuario debe poder deshabilitar el movimiento del indicador de la posición a mover.

3.2.3 RELACIÓN CON OTROS PROTOCOLOS

RSTP tiene algunos conflictos con respecto a la funcionalidad con HTTP, pero también puede actuar recíprocamente con HTTP en donde el contacto inicial con el contenido del streaming es a menudo realizado por un cliente. Las especificaciones del protocolo actual apunta a permitir diferentes hand-off points entre un servidor VOD y un servidor de multimedia implementado en RSTP.

Re-usando la funcionalidad del HTTP este tiene ventajas en por lo menos dos áreas a saber: la seguridad y Proxy. Los requisitos son muy similares, de este modo tiene la habilidad de adoptar HTTP para trabajar en caché, Proxy y autenticación.

Mientras la mayoría de los medios de comunicación de tiempo real usan RTP como un protocolo de transporte, RSTP no se ata a RTP.

3.2.4 MENSAJES RSTP

RSTP es un protocolo basado en texto y usa ISO 10646 los caracteres están codificados en UTF-8. Las líneas son terminadas por CLRF.

Los protocolos basados en texto hacen más fácil la adición de parámetros opcionales. Como los números de los parámetros y la frecuencia de los comandos son bajos, la eficiencia en el procesamiento no es un problema. Los protocolos basados en texto, hechos cuidadosamente, también permiten la fácil implementación de búsquedas de prototipos de lenguajes encriptados tales como Tcl, Visual Basic y Perl.

Las demandas contienen métodos, el objetivo del método es operar más allá de los parámetros al describir el método. Los métodos son impotentes hasta que sean nombrados. Los métodos también son designados para requerir el estado de mantenimiento del servidor multimedia.

3.3 CODIFICACIÓN DEL VIDEO

La codificación de video sirve para pasar señales de video analógico a señales de video digital. La mayoría de codificadores comprimen la información, para que pueda ser almacenada o transmitida ocupando el mínimo espacio posible, para conseguirlo se aprovecha que las secuencias de video tienen redundancia en las dimensiones espacial y temporal. Por lo tanto eliminando información redundante se consigue codificar la información de manera más óptima.

Para eliminar la información redundante en el plano temporal se utiliza la predicción por compensación de movimiento, donde se estima el movimiento entre bloques sucesivos de la imagen. Los bloques son formados por grupos de píxeles, generalmente de 8×8 o 16×16 . Posteriormente se codifica la información de los vectores de movimiento y del error de predicción respecto al bloque anterior.

El funcionamiento del codificador de vídeo es el siguiente: Se separan las señales de luma (Y) y croma (C). Se busca el error de estimación y se hace la DCT. Los coeficientes se cuantifican y se codifican entrópicamente (VLC). Se multiplexan los coeficientes y se pasan al buffer. El buffer controla la

calidad de la señal (se busca que el flujo de bits saliente del buffer no sea variable, ya que la señal está pensada para ser transmitida en un canal con una velocidad estable). La imagen cuantificada se reconstruye mediante Q-1 y la IDCT, para futuras referencias de predicción y estimación de movimiento.

La realización de la transformada DCT y la cuantificación sobre los bloques independientes, pueden ocasionar discontinuidades visibles en los bordes de los bloques provocando el conocido efecto bloque. Con lo cual, los nuevos estándares de codificación de video, como por ejemplo el H.264/MPEG-4 AVC, incorporan algoritmos de filtrado capaces de disminuir este efecto. Uno de los métodos de filtrado más innovadores es el Adaptive deblocking filter.

3.3.1 CLASIFICACIÓN DE CODIFICADORES

Existen varios tipos de codificadores estos básicamente tienen un nivel de calidad y según esta calidad necesitan mayores recursos de red para poder ser visualizados por los cliente en sus equipos terminales. Tal y como se habló en el capítulo uno y dos existen varios tipos como por ejemplo:

- MPEG-2 HDTV
- MPEG-2 DVD

- MPEG-2 NTCS
- MPEG-1
- MPEG-4

A continuación describiremos a fondo la codificación MPEG-2 que es el formato en la que actualmente el Grupo TvCable recibe los videos desde otros proveedores a nivel mundial y al pasar los años se ha convertido en su Standard.

3.3.2 **MPG2:** MOVING PICTURES EXPERTS GROUP 2 (MPEG-2)

Es la designación para un grupo de estándares de codificación de audio y video acordado por MPEG (grupo de expertos en imágenes en movimiento), y publicados como estándar ISO 13818. MPEG-2 es por lo general usado para codificar audio y video para señales de transmisión, que incluyen televisión digital terrestre, por satélite o cable. MPEG-2. Con algunas modificaciones, es también el formato de codificación usado por los discos SVCD's y DVD's comerciales de películas.

3.3.2.1 DESCRIPCIÓN GENERAL

MPEG-2 es similar a MPEG-1, pero también proporciona soporte para video entrelazado (el formato utilizado por las televisiones.) MPEG-2 video no está optimizado para bajas tasas de bits (menores que 1 Mb/s), pero supera en desempeño a MPEG-1 a 3 Mb/s y superiores.

MPEG-2 introduce y define Flujos de Transporte, los cuales son diseñados para transportar video y audio digital a través de medios impredecibles e inestables, y son utilizados en transmisiones televisivas. Con algunas mejoras, MPEG-2 es también el estándar actual de las transmisiones en HDTV. Un decodificador que cumple con el estándar MPEG-2 deberá ser capaz de reproducir MPEG-1.

MPEG-2 audio, mejora a MPEG-1 audio al alojar la codificación de programas de audio con más de dos canales. MPEG-2 admite que sea hecho retro-compatibile, permitiendo que decodificadores MPEG-1 audio puedan decodificar la componente estéreo de los dos canales maestros, o en una manera no retro-compatibile, la cual permite a los codificadores hacer un mejor uso del ancho de banda disponible. MPEG-2 soporta varios formatos de audio, incluyendo MPEG-2 AAC.

3.3.2.2 PROPIEDADES, CARACTERÍSTICAS Y FUNCIONAMIENTO

MPEG-2 es utilizado para la codificación genérica de imágenes en movimiento y el audio asociado que crea un flujo de video mediante tres tipos de datos de marco (cuadros intra, cuadros posteriores predecibles y cuadros predecibles bi-direccionales) arreglados en un orden específico llamado “La estructura GOP”(GOP = Group Of Pictures o grupo de imágenes).

Generalmente el material originado es una secuencia de video a una resolución de píxeles pre-fijada a 25 o 29,97 cuadros por segundo con sonido.

MPEG-2 admite flujos de video escaneado de manera tanto progresiva como entrelazada. En flujos de escaneo progresivo, la unidad básica de codificación es un campo.

El flujo MPEG-2 esta hecho de una serie de cuadros de imágenes codificadas. Las tres maneras de codificar una imagen son: intra-codificado (I cuadro), predecible posterior (P cuadro) y predecible bi-direccional (B cuadro).

La imagen del video es separada en dos partes: luminancia (Y) y croma (también llamada señales de diferencia de color U y V) a su vez, son divididos en "Macro-bloques" los cuales son la unidad básica dentro de una imagen. Cada macro-bloque es dividido en cuatro 8X8 bloques de luminancia. El número de bloques de croma 8X8's depende del formato de color de la fuente. Por ejemplo en el formato común 4:2:0 hay un bloque de croma por macro-bloque por cada canal haciendo un total de seis bloques por macro-bloque.

En el caso de los cuadros I, la verdadera información de imagen pasada a través del proceso codificador descrito abajo, los cuadros P y B primero son sujetos a un proceso de "compensación de movimiento", en el cual son correlacionados con la imagen previa (y en el caso del cuadro B, la siguiente). Cada macro-bloque en la imagen P o B es entonces asociada con un área en la imagen previa o siguiente que este bien correlacionada con alguna de éstas. El "vector de movimiento" que mapea el macro-bloque con su área correlacionada es codificado, y entonces la diferencia entre las dos áreas es pasada a través del proceso de codificación descrito abajo. Cada bloque es procesado con una transformada coseno discreta (DCT) 8X8. El coeficiente DCT resultante es entonces cuantificado de acuerdo a un esquema predefinido, reordenado a una máxima probabilidad de una larga hilera de

ceros, y codificado. Finalmente, se aplica un algoritmo de codificación Huffman de tabla fija.

Los cuadros I codifican redundancia espacial, mientras que los cuadros B y P codifican redundancia temporal. Debido a que los marcos adyacentes son a menudo bien co-relacionados, los cuadros P pueden ser del 10% del tamaño de un cuadro I, y el cuadro B al 2% de su tamaño.

La secuencia de diferentes tipos de marcos es llamada “la estructura de grupos de imágenes” (GOP). Hay muchas estructuras posibles pero una común es la de 15 marcos de largo, y tiene la secuencia I_BB_P_BB_P_BB_P_BB_P_BB_. Una secuencia similar de 12 marcos es también común. La relación de cuadros I, P y B en “la estructura GOP es determinado por la naturaleza del flujo de video y el ancho de banda que constriñe el flujo, además el tiempo de codificación puede ser un asunto importante. Esto es particularmente cierto en las transmisiones en vivo y en ambientes de tiempo real con Fuentes de cómputo limitados, un flujo que contenga varios cuadros B puede tomar tres veces más tiempo para codificar que un archivo que sólo contenga cuadros I.

La tasa de bit de salida de un codificador MPEG-2 puede ser constante (CBR) o variable (VBR), con un máximo determinado por el reproductor – por ejemplo el máximo posible en un DVD de película es de 10.4 Mb/s. Para lograr una tasa de bits constante el grado de cuantificación es alterado para lograr la tasa de bits requerida. Incrementar la cuantificación hace visible un defecto cuando el video es descodificado, Generalmente en la forma de “amosaicamiento”, donde las discontinuidades en los fillos de los macrobloques se hace más visible como reducción de la tasa de bits.

3.4 CONECTIVIDAD DE EQUIPOS

Una vez descrito el codificador y el protocolo utilizado para la conectividad entre los equipos que están en cada headend y los equipos que utiliza cada cliente para tener acceso al nuevo servicio de VOD nos queda por hablar de los equipos que se van a utilizar al momento de implementar la infraestructura en cada headend.

3.4.1 VOD SERVER MDS 200 SERVIDOR DE VIDEO

En el capítulo dos ya se describieron las características que deberían tener los servidores que se vayan a escoger para poder soportar la demanda las

peticiones que hagan los usuarios. Principalmente la característica que permita a los servidores disponerlos en arquitectura Proxy-tree para mejorar niveles de escalabilidad y en general aprovechar al máximo recursos de red.

3.4.1.1 DESCRIPCIÓN GENERAL

El servidor de VOD que se ha escogido para el diseño de nuestro proyecto de tesis es de la marca SeaChange, en los anexos C y D se pueden observar las características técnicas de diferentes VOD Server, y del porqué se escogió la marca SeaChange que en la actualidad es una de las empresas más importantes que se han dedicado al desarrollo de equipos que soporten altos niveles de tráfico reduciendo al máximo pérdidas y optimizando todos los recursos de red. En la Figura 3.1 podemos el equipo que se utilizará.



Figura 3.1 Servidor MDS 200

Una de las características más importantes que se necesitan al momento de la elección de equipos es la convergencia de los mismos sin importar el

fabricante de los mismos. Ya que en la actualidad como sabemos el Grupo TvCable presta varios servicios como lo son internet, televisión, telefonía y servicios de valor agregado como lo es PPV, y para la implementación de los mismos se tiene una gama diferente de equipos desde fabricantes conocidos como Cisco, Motorota Safari, entre otros. Con la inclusión de los equipos que proponemos en el diseño de infraestructura y demás componentes debemos asegurar que los servicios que ya se prestan no se vean afectados o que el servicio que se piensa prestar (VOD) no tenga inconveniente alguno de incompatibilidad.

3.4.1.2 CARACTERÍSTICAS TÉCNICAS

El servidor VOD MDS 200 tiene muchas características importantes que en la actualidad se requieren por su demanda.

Una de ellas es que soporta RAID-5. A continuación se describirá los niveles de RAID y la importancia de esta característica para el diseño de nuestro proyecto.

RAID:

El término RAID es un acrónimo del inglés "Redundant Array of Independent Disks". Significa matriz redundante de discos independientes. RAID es un método de combinación de varios discos duros para formar una única unidad lógica en la que se almacenan los datos de forma redundante. Ofrece mayor tolerancia a fallos y más altos niveles de rendimiento que un sólo disco duro o un grupo de discos duros independientes.

Una matriz consta de dos o más discos duros que ante el sistema principal funcionan como un único dispositivo. Un RAID, para el sistema operativo, aparenta ser un sólo disco duro lógico (LUN). Los datos se desglosan en fragmentos que se escriben en varias unidades de forma simultánea.

La elección de los diferentes niveles de RAID va a depender de las necesidades que se tengan por parte de la empresa en este caso el Grupo TvCable con lo que respecta a factores como seguridad, velocidad, capacidad, coste, etc. Cada nivel de RAID ofrece una combinación específica de tolerancia a fallos (redundancia), rendimiento y coste, diseñadas para satisfacer las diferentes necesidades de almacenamiento. La mayoría de los niveles RAID pueden satisfacer de manera efectiva sólo uno o dos de estos

criterios. No hay un nivel de RAID mejor que otro; cada uno es apropiado para determinadas aplicaciones y entornos. De hecho, resulta frecuente el uso de varios niveles RAID para distintas aplicaciones del mismo servidor. Oficialmente existen siete niveles diferentes de RAID (0-6), definidos y aprobados por el RAID Advisory Board (RAB). Luego existen las posibles combinaciones de estos niveles. Los niveles RAID 0, 1, 0+1 y 5 son los más populares.

RAID 0: Disk Striping "La más alta transferencia, pero sin tolerancia a fallos".

También conocido como "separación ó fraccionamiento/ Striping". Los datos se desglosan en pequeños segmentos y se distribuyen entre varias unidades. Este nivel de "array" o matriz no ofrece tolerancia al fallo. Al no existir redundancia, RAID 0 no ofrece ninguna protección de los datos. El fallo de cualquier disco de la matriz tendría como resultado la pérdida de los datos y sería necesario restaurarlos desde una copia de seguridad. Por lo tanto, RAID 0 no se ajusta realmente al acrónimo RAID. Consiste en una serie de unidades de disco conectadas en paralelo que permiten una transferencia simultánea de datos a todos ellos, con lo que se obtiene una gran velocidad en las operaciones de lectura y escritura. La velocidad de transferencia de datos aumenta en relación al número de discos que forman el conjunto. Esto representa una gran ventaja en operaciones secuenciales con ficheros de

gran tamaño. Por lo tanto, este array es aconsejable en aplicaciones de tratamiento de imágenes, audio, video o CAD/CAM, es decir, es una buena solución para cualquier aplicación que necesite un almacenamiento a gran velocidad pero que no requiera tolerancia a fallos. Se necesita un mínimo de dos unidades de disco para implementar una solución RAID 0.

RAID 1: Mirroring "Redundancia. Más rápido que un disco y más seguro"

También llamado "Mirroring" o "Duplicación" (Creación de discos en espejo). Se basa en la utilización de discos adicionales sobre los que se realiza una copia en todo momento de los datos que se están modificando. RAID 1 ofrece una excelente disponibilidad de los datos mediante la redundancia total de los mismos. Para ello, se duplican todos los datos de una unidad o matriz en otra. De esta manera se asegura la integridad de los datos y la tolerancia al fallo, pues en caso de avería, la controladora sigue trabajando con los discos no dañados sin detener el sistema. Los datos se pueden leer desde la unidad o matriz duplicada sin que se produzcan interrupciones. RAID 1 es una alternativa costosa para los grandes sistemas, ya que las unidades se deben añadir en pares para aumentar la capacidad de almacenamiento. Sin embargo, RAID 1 es una buena solución para las aplicaciones que requieren redundancia cuando hay sólo dos unidades disponibles. Los servidores de archivos pequeños son un buen ejemplo. Se necesita un mínimo de dos unidades para implementar una solución RAID 1.

RAID 0+1/ RAID 0/1 ó RAID 10: "Ambos mundos"

Combinación de los arrays anteriores que proporciona velocidad y tolerancia al fallo simultáneamente. El nivel de RAID 0+1 fracciona los datos para mejorar el rendimiento, pero también utiliza un conjunto de discos duplicados para conseguir redundancia de datos. Al ser una variedad de RAID híbrida, RAID 0+1 combina las ventajas de rendimiento de RAID 0 con la redundancia que aporta RAID 1. Sin embargo, la principal desventaja es que requiere un mínimo de cuatro unidades y sólo dos de ellas se utilizan para el almacenamiento de datos. Las unidades se deben añadir en pares cuando se aumenta la capacidad, lo que multiplica por dos los costes de almacenamiento. El RAID 0+1 tiene un rendimiento similar al RAID 0 y puede tolerar el fallo de varias unidades de disco. Una configuración RAID 0+1 utiliza un número par de discos (4, 6, 8) creando dos bloques. Cada bloque es una copia exacta del otro, de ahí RAID 1, y dentro de cada bloque la escritura de datos se realiza en modo de bloques alternos, el sistema RAID 0. RAID 0+1 es una excelente solución para cualquier uso que requiera gran rendimiento y tolerancia a fallos, pero no una gran capacidad. Se utiliza normalmente en entornos como servidores de aplicaciones, que permiten a los usuarios acceder a una aplicación en el servidor y almacenar datos en sus discos duros locales, o como los servidores web, que permiten a los

usuarios entrar en el sistema para localizar y consultar información. Este nivel de RAID es el más rápido, el más seguro, pero por contra el más costoso de implementar.

RAID 2: "Acceso paralelo con discos especializados. Redundancia a través del código Hamming"

El RAID nivel 2 adapta la técnica comúnmente usada para detectar y corregir errores en memorias de estado sólido. En un RAID de nivel 2, el código ECC (Error Correction Code) se intercala a través de varios discos a nivel de bit. El método empleado es el Hamming. Puesto que el código Hamming se usa tanto para detección como para corrección de errores (Error Detection and Correction), RAID 2 no hace uso completo de las amplias capacidades de detección de errores contenidas en los discos. Las propiedades del código Hamming también restringen las configuraciones posibles de matrices para RAID 2, particularmente el cálculo de paridad de los discos. Por lo tanto, RAID 2 no ha sido apenas implementado en productos comerciales, lo que también es debido a que requiere características especiales en los discos y no usa discos estándares. Debido a que es esencialmente una tecnología de acceso paralelo, RAID 2 está más indicado para aplicaciones que requieran una alta tasa de transferencia y menos conveniente para aquellas otras que requieran una alta tasa de demanda I/O.

RAID 3: "Acceso síncrono con un disco dedicado a paridad"

Dedica un único disco al almacenamiento de información de paridad. La información de ECC (Error Checking and Correction) se usa para detectar errores. La recuperación de datos se consigue calculando el O exclusivo (XOR) de la información registrada en los otros discos. La operación I/O accede a todos los discos al mismo tiempo, por lo cual el RAID 3 es mejor para sistemas de un sólo usuario con aplicaciones que contengan grandes registros.

RAID 3 ofrece altas tasas de transferencia, alta fiabilidad y alta disponibilidad, a un coste intrínsecamente inferior que un Mirroring (RAID 1). Sin embargo, su rendimiento de transacción es pobre porque todos los discos del conjunto operan al unísono. Se necesita un mínimo de tres unidades para implementar una solución RAID 3.

RAID 4: "Acceso Independiente con un disco dedicado a paridad."

Basa su tolerancia al fallo en la utilización de un disco dedicado a guardar la información de paridad calculada a partir de los datos guardados en los otros discos. En caso de avería de cualquiera de las unidades de disco, la

información se puede reconstruir en tiempo real mediante la realización de una operación lógica de O exclusivo. Debido a su organización interna, este RAID es especialmente indicado para el almacenamiento de ficheros de gran tamaño, lo cual lo hace ideal para aplicaciones gráficas donde se requiera, además, fiabilidad de los datos. Se necesita un mínimo de tres unidades para implementar una solución RAID 4. La ventaja con el RAID 3 está en que se puede acceder a los discos de forma individual.

RAID 5: "Acceso independiente con paridad distribuida."

Este array ofrece tolerancia al fallo, pero además, optimiza la capacidad del sistema permitiendo una utilización de hasta el 80% de la capacidad del conjunto de discos. Esto lo consigue mediante el cálculo de información de paridad y su almacenamiento alternativo por bloques en todos los discos del conjunto. La información del usuario se graba por bloques y de forma alternativa en todos ellos. De esta manera, si cualquiera de las unidades de disco falla, se puede recuperar la información en tiempo real, sobre la marcha, mediante una simple operación de lógica de O exclusivo, sin que el servidor deje de funcionar.

Así pues, para evitar el problema de cuello de botella que plantea el RAID 4 con el disco de comprobación, el RAID 5 no asigna un disco específico a esta misión sino que asigna un bloque alternativo de cada disco a esta misión de escritura. Al distribuir la función de comprobación entre todos los discos, se disminuye el cuello de botella y con una cantidad suficiente de discos puede llegar a eliminarse completamente, proporcionando una velocidad equivalente a un RAID 0.

RAID 5 es el nivel de RAID más eficaz y el de uso preferente para las aplicaciones de servidor. Comparado con otros niveles RAID con tolerancia a fallos, RAID 5 ofrece la mejor relación rendimiento-coste en un entorno con varias unidades. Gracias a la combinación del fraccionamiento de datos y la paridad como método para recuperar los datos en caso de fallo, constituye una solución ideal para los entornos de servidores en los que gran parte del E/S es aleatoria, la protección y disponibilidad de los datos es fundamental y el coste es un factor importante. Este nivel de array es especialmente indicado para trabajar con sistemas operativos multiusuarios. Se necesita un mínimo de tres unidades para implementar una solución RAID 5. Los niveles 4 y 5 de RAID pueden utilizarse si se disponen de tres o más unidades de disco en la configuración, aunque su resultado óptimo de capacidad se obtiene con siete o más unidades. RAID 5 es la solución más económica por

megabyte, que ofrece la mejor relación de precio, rendimiento y disponibilidad para la mayoría de los servidores.

Una vez descritos los niveles de RAID y su utilidad cabe recalcar que el servidor MDS 200 es de tipo modular por lo que si se necesita mayor capacidad al aumentar los requerimientos se puede aumentar la capacidad del mismo.

Otras de las características que tiene son:

- Tiene 8 puertos GigabitEthernet
- Capacidad de 5.5 Gbps de streaming
- 32 GB DRAM caché memory
- 6 TB de almacenamiento físico.
- 1500 a 3200 horas de Videos Formato MPEG2
- Temperatura de Operación: 10°C a 30°C
- Soporta Fuentes de Poderes de 110V y 220V

Estas características son muy importantes por lo que se describió en el capítulo 2 con respecto a capacidades de almacenamiento que se tendrán para separar el almacenamiento de la parte Proxy y la distribución de mirror

para mejorar las probabilidades de acierto al momento de recibir un requerimiento de un cliente. Algo importante de acotar es que las películas en la práctica no son de tamaños fijos lo que para este estudio se está asumiendo para realizar de forma analítica el análisis del sistema. Si asumimos que las películas y demás contenido multimedia es de 2 horas en el formato establecido MPEG-2, el número de películas que se podrán acceder es de 750 si tomamos el mínimo de hora que se pueden almacenar en la configuración más básica del mismo.

Con respecto a la temperatura de operación, en la actualidad los headends del Grupo TvCable cubren todas las especificaciones ya que se debe prever el estado de los equipos para evitar un recalentamiento de los mismos y esto traiga problemas mayores como la suspensión temporal del servicio.

3.4.2 MOTOROLA DCT700 (SET TOP BOX) EQUIPO TERMINAL

Talvez este equipo sea la parte más importante en toda la infraestructura que se va a armar para prestar el servicio, ya que es este el que finalmente va a estar en cada casa y lo más importante que este equipo es el que se va a reemplazar por el que actualmente se encuentra distribuido por todo el país, el STB que se va a utilizar va a ser comprado de forma masiva y si no cuenta

con las características necesarias para brindar los servicios que se están prestando actualmente y con el nuevo servicio, la inversión que se hará será en vano, lo que representaría un pérdida muy grande para la empresa.

3.4.2.1 DESCRIPCIÓN GENERAL

En el capítulo 1 apartado 1.1.3 vimos que los STB deben de tener ciertas características para poder brindarle un buen nivel de QOS y así tenga una buena acogida en el mercado actual que ofrece varios productos. En el anexo E se muestra varios modelos y marcas de Set Top Box, características tales como modulación, sincronización, almacenamiento del buffer, interfaz gráfica y costo definen al DCT700 de Motorota el más indicado para nuestro diseño.

En general los STB (set top box) constan de 4 componentes principales: Interfase de red, decodificador, buffer y hardware de sincronización. Cada componente será descrito en el siguiente apartado a detalle para conocer este equipo de forma exhaustiva.

En la figura 3.2 se puede observar la parte frontal del DTC700.



Figura 3.2 DCT 700 parte frontal

3.4.2.2 CARACTERÍSTICAS TÉCNICAS

Si empezamos a describir las características que tiene el DCT700 debemos empezar por las 4 partes primordiales que lo conforman.

En el gráfico 3.3 podemos observar un esquema de bloques de las cuatro partes que conforman el STB y que permiten que al momento de que se tenga un requerimiento de video este sea atendido de forma rápida y eficaz. Así como también las partes necesarias para cuando exista una petición por parte de un usuario esta tenga que llegar hasta el servidor para poder gestionarse.

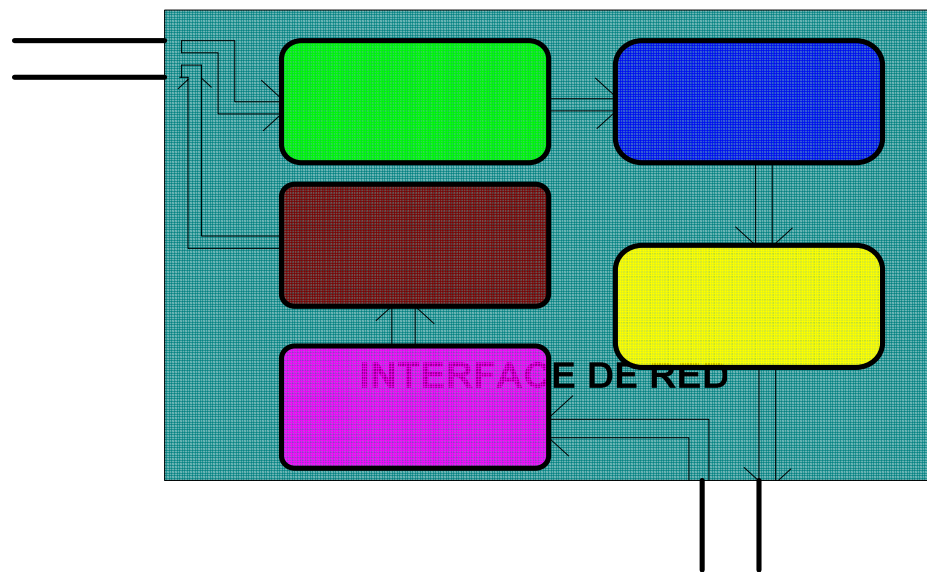


Figura 3.3 Esquema de Bloques del equipo Terminal DCT 700

La primera característica es la interfase de Red que permite al cliente recibir y enviar la información. En este caso la interfase que se tiene en estos equipos y por la característica de la red de acceso al cliente es un cable coaxial con un conector de 75Ω . En la figura 3.4 (RF IN) podemos observar la interfase de red del DCT700.



Figura 3.4 Parte posterior del Equipo Terminal

Como podemos observar en el conector que indica RF IN va conectado con la red de acceso que el Grupo TvCable provee a sus clientes. Los demás conectores tienen su diferente utilidad.

Se pueden realizar dos diferentes configuraciones. La primera es directamente a un Televisor ilustrada en las figuras 3.5 y 3.6, la segunda configuración es con un VCR intermedio entre el equipo Terminal o STB (DCT700) y el Televisor que se ilustra en la figura 3.7 y 3.8.

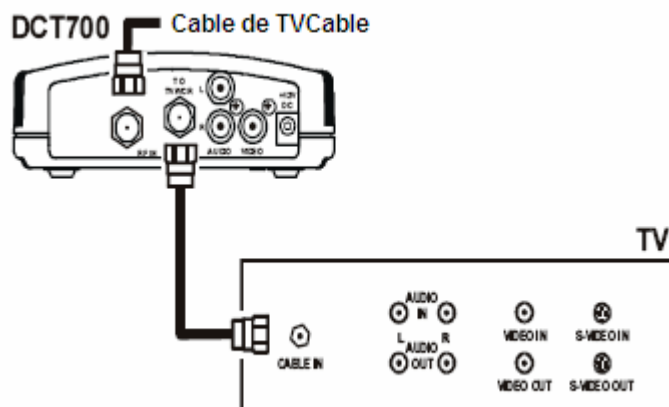


Figura 3.5 Como conectar TV-DCT700 Cable coaxial 75 Ω

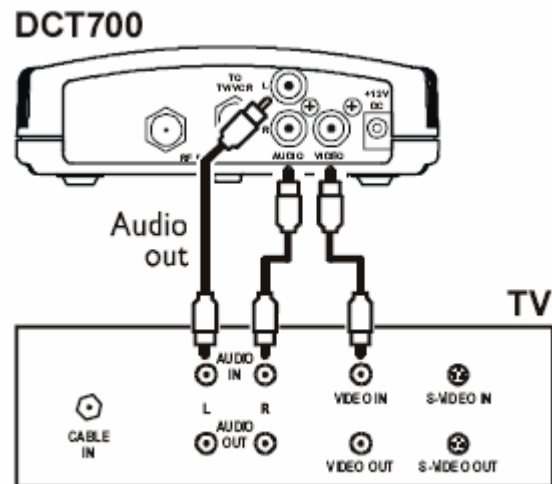


Figura 3.6 Conexión de Audio TV-DCT700

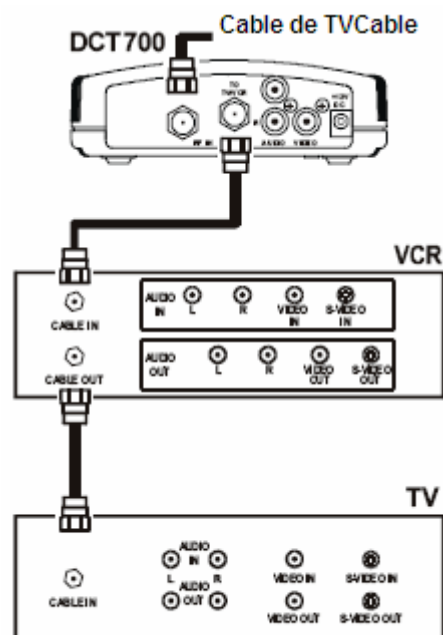


Figura 3.7 Conexión DCT700-VCR-TV cable coaxial de 75 Ω

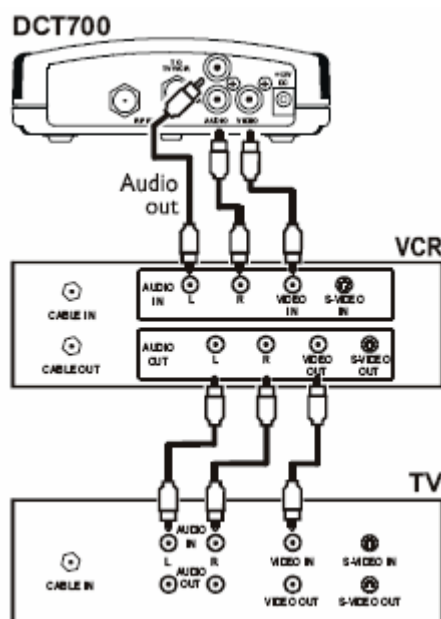


Figura 3.8 Conexión DCT700-VCR-TV Audio y Video

Una vez descrito las diferentes configuraciones que se pueden armar por cualquier usuario según su necesidad hablaremos de las demás características que tiene el DCT700 por el cual fue escogido.

Si hablamos de todo sistema de comunicación ya sea este analógico o digital, sabemos que la información que viaja a través de una red no puede ir en banda base, esto quiere decir que viaje tal y como fue capturada de su medio, ya sea sonido, video o datos. Para optimizar el ancho de banda del canal por el cual viaja la información se debe de utilizar modulaciones que dependiendo del tipo de información que vaya a viajar se debe escoger. En

capítulos anteriores se estudió ampliamente los tipos de formatos de video en este caso el formato con el que trabaja el grupo TvCable es MPGE-2, ya que este le permite una buena calidad de señal y presentación hacia los usuarios, además los proveedores internacionales de video, como por ejemplo la cadena Discovery trabajan bajo este mismo formato lo que permite llevar un estándar ante los usuarios. Hay que tomar en cuenta que los cálculos que se han realizado y que se van a realizar están basados en este formato de video.

Ahora si hablamos a nivel de modulación el DCT 700 puede trabajar con 64/254 QAM y FEC. A continuación haremos una descripción de lo que significa de estos sistemas de modulación para tener más en cuenta su importancia.

Modulación QAM (Modulación de Amplitud en Cuadratura)

Es una modulación digital avanzada que transporta datos cambiando la amplitud de dos ondas portadoras. Estas dos ondas, generalmente sinusoidales, están desfasadas entre si 90° en la cual una onda es la portadora y la otra es la señal de datos. Se utiliza para la transmisión de datos a alta velocidad por canales con ancho de banda restringido.

La modulación QAM consiste en modular en amplitud (ASK) de forma independiente, dos portadores que tienen la misma frecuencia pero que están desfasadas entre sí 90°. La señal modulada QAM es el resultado de sumar ambas señales ASK. Estas pueden operar por el mismo canal sin interferencia mutua porque sus portadoras están en cuadratura.

La fórmula matemática de una señal modulada en QAM se la puede observar en la expresión 3.1:

$$a_n \cos \omega t + b_n \sin \omega t \quad (3.1)$$

Las amplitudes de las dos señales moduladas en ASK (a y b), toman de forma independiente los valores discretos a_n y b_n correspondientes al total de los "N" estados de la señal moduladora codificada en banda base multinivel (Expresión 3.2)

$$N = n \times m. \quad (3.2)$$

Una modulación QAM se puede reducir a la modulación simultanea de amplitud $ASK_{n,m}$ y fase $PSK_{n,m}$ de una única portadora, pero solo cuando los estados de amplitud $A_{n,m}$ y de fase $H_{n,m}$ que esta dispone, mantienen con las amplitudes de las portadoras originales a_n y b_n la relaciones que se indican en la expresión 3.3:

$$\text{QAM} \rightarrow A_n(\cos wt) + B_m(\sin wt) = A_{n,m}(\cos wt - H_{n,m}) \quad (3.3)$$

Donde $A_n(\cos wt)$ y $B_m(\sin wt)$ están moduladas en ASK, $A_{n,m}$ esta modulada en ASK y $(\cos wt - H_{n,m})$ es una expresión modulada en PSK .

Estas expresiones se deducen fácilmente a partir de las siguientes expresiones 3.4, 3.5 y 3.6.

$$\text{QAM} \rightarrow A \cos(wt - h) = A \cos wt \times \cos h + A \sin wt \times \sin h \quad (3.4)$$

$$\text{QAM} \rightarrow A \cos(wt - h) = (A \cos h) \times \cos wt + (A \sin h) \times \sin wt \quad (3.5)$$

$$\text{QAM} \rightarrow A \cos(wt - h) = A \cos wt + b \sin wt; \text{ por lo tanto: } a=A \cos h; b=A \sin h$$

$$(3.6)$$

A partir de la explicación anterior lo que se busca es la inmunidad que tiene la señal modulada en cuanto a las perturbaciones y al ruido de la línea, es mayor cuanto más separados estén los puntos del diagrama de estados. Se trata pues de buscar una constelación de puntos con unas coordenadas de amplitud y fase que hagan máxima la separación entre ellos.

De aquí se creó la modulación M-QAM en las cuales dependiendo del número n de bits se puede crear 2^n estados, de los cuales los más utilizados son 16-QAM, 64-QAM y 256-QAM que quiere decir que se han utilizado 4, 6 y 8 bits respectivamente. En la figura 3.9 se muestra la constelación 16-QAM y su respectiva codificación de los estados en código binario.

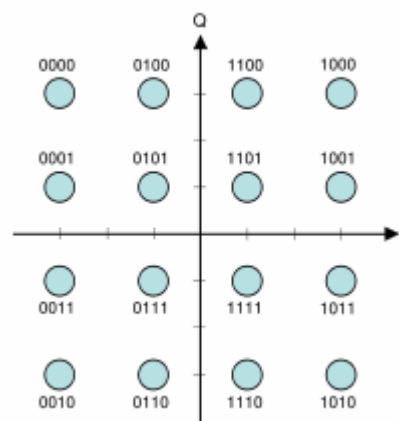


Figura 3.9 Constelación para sistema 16-QAM

En el DCT se puede configurar su funcionamiento para que según las necesidades que se tengan trabaje a 64-QAM o 256-QAM. Si el STB (DCT-700) puede manejar señales moduladas 64 o 256 QAM necesita un equipo que del lado del Headend que maneje esta misma modulación. Este equipo es el SEM (SmartStream Encryptor Modulator) a continuación haremos una breve descripción de este equipo que maneja principalmente los canales que se utilizan por lo diferentes servicios que se prestan.

SEM (SmartStream Encryptor Modulator)

En el anexo F se observa los detalles técnicos de las diferentes marcas que ofrecen el servicio de Modulador. Y aunque el costo sea un poco mayor, los beneficios que prestan este representan un peso mayor para su elección. El SEM provee las bases para las aplicaciones como VOD, inserción de programas digitales y soporte broadcast. El SEM soporta múltiples entradas MPEG-2, permitiendo la recepción simultanea de diferentes servidores VOD, encoders locales, multiplexadores o receptores digitales vía Gigabit Ethernet o DVB-ASI. En total, el SEM puede procesar más de 900Mbps de datos MPEG-2 a través de todas sus entradas.

En la figura 3.10 podemos observar el diagrama de bloques del funcionamiento del SEM.

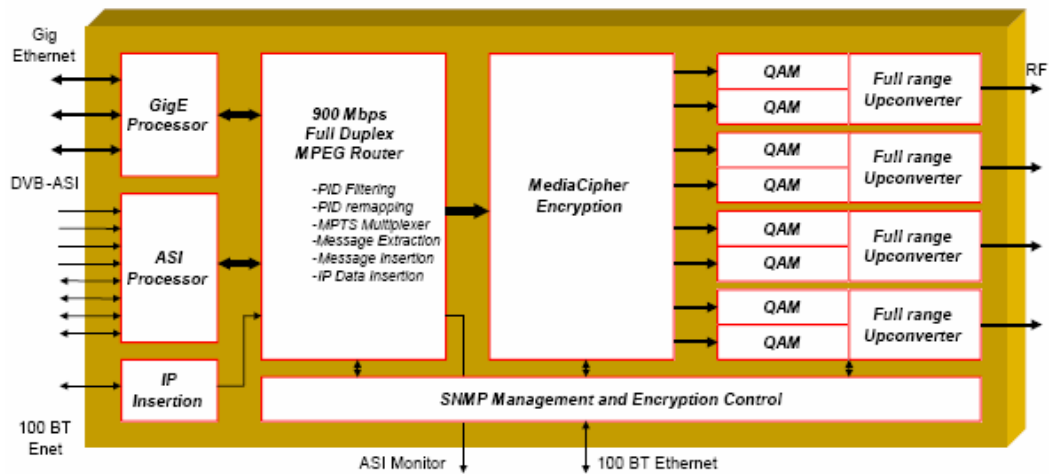


Figura 3.10 Diagrama de bloques del SEM

Como se puede observar en el diagrama el SEM puede aceptar entradas de hasta 8 interfaces ASI y 3 interfaces Gigabit Ethernet, estas entradas son agregadas en un multiplexador que provee la salida para el procesamiento de la encriptación en RF. El bloque RF consiste de 8 moduladores QAM y 4 RF upconverter, cada uno de estos lo multiplexan en su canal RF correspondiente

FEC (Forward Error Correction)

Es un tipo de mecanismo de corrección de errores que permite su corrección en el receptor sin retransmisión de la información original. Se utiliza en sistemas sin retorno o sistemas en tiempo real donde no se puede esperar a la retransmisión para mostrar los datos.

La posibilidad de corregir errores se consigue añadiendo al mensaje original unos bits de redundancia. La fuente digital envía la secuencia de datos al codificador, encargado de añadir dichos bits de redundancia. A la salida del codificador obtenemos la denominada palabra código. Esta palabra código es enviada al receptor y éste, mediante el decodificador adecuado y aplicando los algoritmos de corrección de errores, obtendrá la secuencia de datos original. Los dos principales tipos de codificación usados son:

Códigos bloque. La paridad en el codificador se introduce mediante un algoritmo algebraico aplicado a un bloque de bits. El decodificador aplica el algoritmo inverso para poder identificar y, posteriormente corregir los errores introducidos en la transmisión.

Códigos convolucionales. Los bits se van codificando tal y como van llegando al codificador. Cabe destacar que la codificación de uno de los bits está enormemente influenciada por la de sus predecesores. La decodificación para este tipo de código es compleja ya que en principio, es necesaria una gran cantidad de memoria para estimar la secuencia de datos más probable para los bits recibidos. En la actualidad se utiliza para

decodificar este tipo de códigos el algoritmo de Viterbi, por su gran eficiencia en el consumo de recursos.

Estas características QAM y FEC que el dispositivo DCT700 tiene, ayudan ampliamente en la reproducción en tiempo real reduciendo la posibilidad de errores al momento de estar reproduciendo un video.

El buffer y el hardware de Sincronización están embebidos dentro del equipo Terminal permitiendo que el cliente tenga una calidad de video y audio excelente al momento de descargar el video. Algo importante de nombrar es la comunicación constante entre el servidor y el equipo Terminal, ya que en cualquier instante el usuario puede tener algún otro requerimiento como pausar, retroceder o avanzar en el video y la respuesta desde el servidor debe de ser instantánea para que no existan cortes en la reproducción del video.

Una vez escogidos los equipos terminales, servidor, y el equipo modulador es importante recalcar que el servicio que se ofrece tanto de VOD como la de HFC es necesario los equipos DAC 6000 y el NC 1500 que serán descritos a continuación.

3.4.2.3 AUTENTICACIÓN DE LOS EQUIPOS TERMINALES STB CON EL SERVIDOR VOD

En la actualidad en nuestro país es muy importante el tema de seguridad de contenidos, ya que existen muchos dispositivos que diariamente se fabrican dentro o fuera del país que violan la seguridad y permiten tener acceso a servicios o información que muchas personas no deberían tener.

El grupo TvCable tiene redes de acceso ubicados por todo el país lo que imposibilita el control de cada punto de acceso físico a su red, por lo que se deben tomar otras medidas de seguridad para que la información que manejan, ya sea voz, video o datos no sea adquirida por terceras personas.

El primer nivel de seguridad que se propone para asegurar el acceso de personas al servicio de VOD del cuál se está diseñando la red y demás componentes necesarios para la prestación del mismo es:

Autenticación de la MAC ADDRESS del dispositivo Terminal

Cada dispositivo en el mundo entero ya sean computadoras, routers, etc., tienen una dirección única que lo identifica ante el mundo. Esta dirección contiene datos muy importantes como: fabricante, modelo, etc.

Una vez que se quiera agregar un nuevo usuario al servicio de video en demanda se debe agregar a una base de datos donde conste información básica como número de contrato, nombre y la MAC Address del equipo que se encontrará en su hogar, de esta forma si alguna persona se conecta en un punto de acceso y tratase de obtener algún video de forma ilegal los elementos que conforman la infraestructura no le permitirán que este tenga acceso.

Este tipo de autenticación va de la mano con los protocolos que fueron descritos en capítulos anteriores que prestan un nivel de seguridad básico dentro de nuestro sistema de comunicaciones.

Pero este sistema de autenticar la MAC de los equipos no es del todo seguro ya que actualmente existen dispositivos que permiten duplicar MACs de cualquier dispositivo, es más en dispositivos como los AP (Access Point) tiene la opción de clonar MACs de las computadoras que están conectadas a su red. Por esta razón dentro de los equipos que se proponen para la infraestructura están el DAC 6000 y NC 1500. A continuación se describirá la utilización de estos dispositivos en la infraestructura.

DAC 6000 (DIGITAL ADDRESS CONTROLLER)

DAC 6000 es un equipo fabricado por Motorola específicamente para brindar seguridad al momento de prestar un servicio. Cuenta con varios sistemas de comprobación hacia los equipos terminales y es compatible con una gama muy amplia de equipos motorola, entre esos el DCT 700. La forma de utilización de este equipo es una especie de Firewall ante los intrusos, se puede observar un esquema de funcionamiento en el siguiente esquema (figura 3.10).

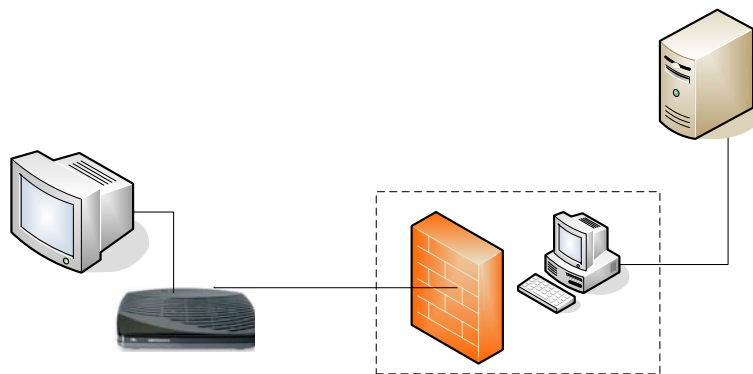


Figura 3.10 Esquema de funcionamiento del DAC6000

Presenta un ambiente gráfico que permite un mejor manejo, además es multiusuario lo cuál permite tener varias sesiones para ingresar a su base de datos. Un dato interesante en este equipo es que se pueden ejecutar scripts

que permiten que desde estaciones remotas se puedan leer o grabar cambios dentro de su base de datos.

Este equipo es muy importante al momento de la facturación ya que es este donde se alojan los logs o reportes de los requerimientos que se han tenido y con estos datos se realizará la facturación, por estas razones es que debe de contar con los controles de seguridad necesarios para que el acceso de ciertos usuarios no pueda modificar los registros que se están llevando. El software que maneja este equipo es propietario de motorola lo que permite en mejor control con respecto a accesos ya que si alguien intenta conectar un equipo que no sea motorola este lo detectará inmediatamente bloqueando su acceso.

NC 1500 (Network Controller)

Si ya dijimos que gracias al DAC (digital Address Controller) podemos controlar el acceso de los equipos Terminal STB en este caso los DCT 700 falta un equipo que nos permita la comunicación ente los DCT y el VOD Server. Un equipo que nos permita asignar indicadores a los STB de una forma interactiva para poder controlar de mejor manera los equipos. El NC 1500 (figura 3.11) es un equipo Motorola que nos permite manejar la

comunicación entre cada STB y los contenidos multimedia, en este caso el servidor VOD, también es el encargado de enviar por separado la información hasta cada STB por separado.



Figura 3.11 Parte Frontal del NC 1500

Los DCT utilizan el NC 1500 para habilitar sesiones interactivas con cualquier tipo de aplicación, como lo es con el servidor VOD. Este equipo junto al OM 1000 y RDP 2000 sirve como el camino de bajada y subida para los datos que se reciben o envían hacia un determinado DCT.

Las gráficas (figura 3.12 y 3.13) siguientes muestran la forma de interconexión que van a tener con los equipos que hasta ahora hemos descrito.

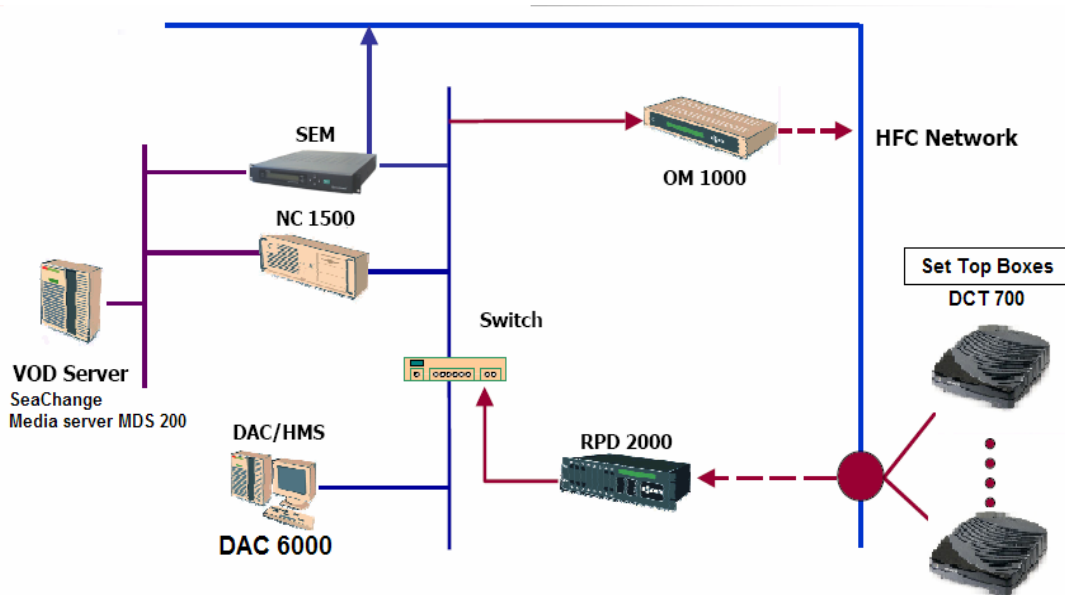


Figura 3.12 Esquema de Infraestructura para VOD 1

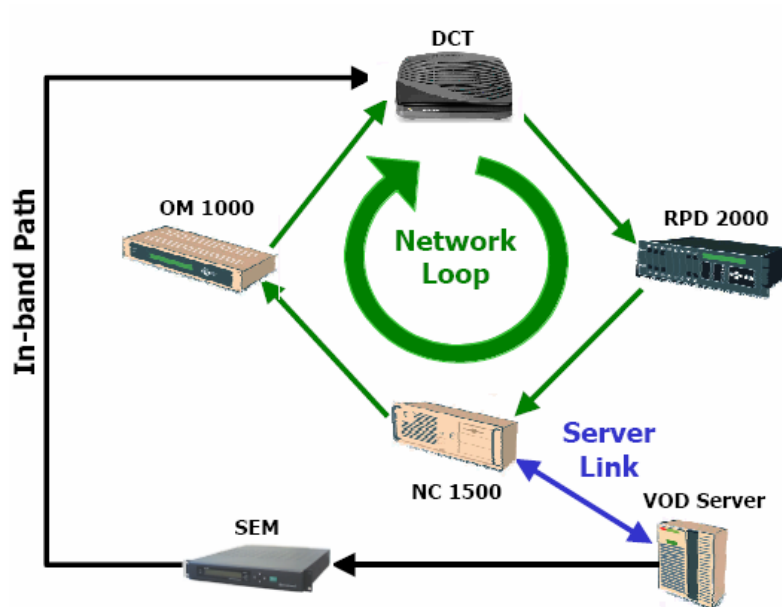


Figura 3.13 Esquema de Infraestructura para VOD 2

Talvez estas figuras sean unas de las más importantes en nuestro proyecto de tesis, ya que este el esquema que van a tener todos los equipos que estén conectados en nuestro diseño. Cuando se tenga un requerimiento de un DCT este será recibido por el NC 1500 a través del RDP 2000, dicho requerimiento será validado por el DAC 6000, analizando si este equipo puede realizar dicha operación. Una vez que este requerimiento sea validado será enviado hacia el servidor de VOD MDS 200 el mismo que empezará a enviar información con el NC 1500 a través del OM 1000 enviará al DCT indicado la información que anteriormente se requirió. Cada requerimiento que se tenga de un DCT será manejada por el NC 1500 y luego validada por el DAC 6000. Los equipos RDP 2000 y OM 1000 independientemente manejan cada requerimiento ya sea de upstream o downstream respectivamente. La funcionalidad del SEM será explicada de forma más detallada en el capítulo 4 ya que es parte de la red actual de servicios que brinda el Grupo TvCable.

A continuación se muestran los equipos OM 1000 y RPD 2000 para que se pueda tener una idea de su estructura y dimensiones físicas en la figura 3.14 y 3.15 respectivamente.



Figura 3.14 OM 1000 Out-of-Band Modulator



Figura 3.15 RPD 2000 Return Path Decoder

Capítulo 4

4. ANALISIS EN REDES EXISTENTES DE TV-CABLE

En este capítulo 4 se describe el esquema del headend y la red de acceso HFC con sus elementos desagregados explicando su funcionamiento y su importancia en la red.

4.1 INTRODUCCIÓN

Esta red inició su construcción en el año 1986. Hoy en día, es la red más extensa del país para este tipo de servicio, con cobertura en las ciudades de Ambato, Cuenca, Guayaquil, Ibarra, Loja, Machala, Manta, Portoviejo, Quito, Riobamba, Salinas y Tulcán.

La red tiene una topología tipo árbol. Se comunica por medio de Fibra óptica desde el headend hasta los nodos y desde allí por cable coaxial hasta el usuario. Actualmente existen 53 nodos instalados en la ciudad de guayaquil, 23 en Quito y 23 en Cuenca. A continuación se detalla en un gráfico esquemático de su topología en la figura 4.1

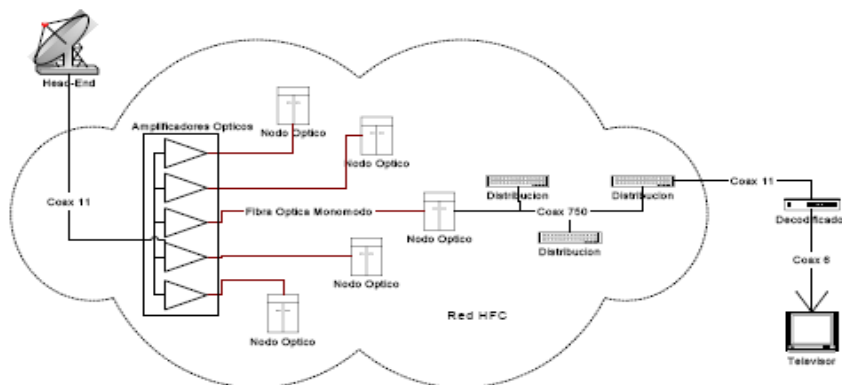


Figura 4.1 Diagrama esquemático de la topología HFC

Sobre estas redes se han implementado una serie de servicios permitiéndola llamar una red convergente. En la actualidad se brindan servicios triple-play (voz, video y datos figura 4.2) con un gran índice de penetración de cada uno de los servicios brindados y es por esta razón que el Grupo TVCable está interesado en la implementación de nuevos servicios como lo es VOD para acaparar mas clientes en el mercado de telecomunicaciones.

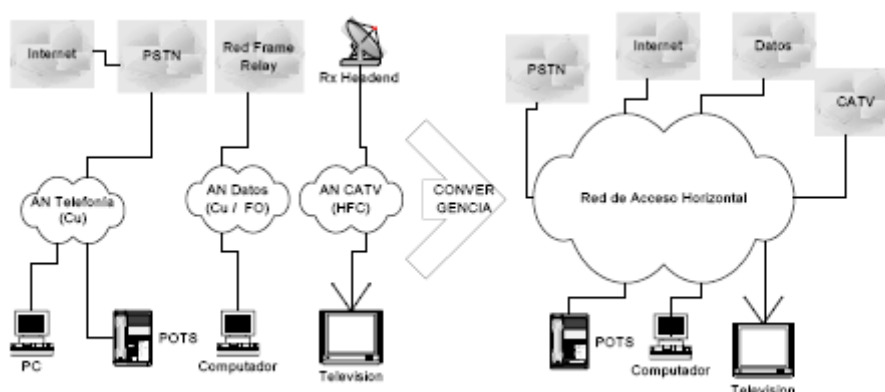


Figura 4.2 Conversión de Redes

4.2 DESCRIPCIÓN DE LA RED HFC

A nivel nacional existen en general dos tipos de redes HFC, la basada en fibra óptica de dos hilos uno de retorno y otro de envío, mientras que en la mayoría de las ciudades pequeñas no existe el retorno por lo tanto no se

puede prestar los servicios especiales como datos, voz, PPV (pay per view) y VOD (video en demanda). Es por eso que nuestro estudio se ha enfocado en las ciudades en donde existe el canal de retorno como lo son Guayaquil, Quito y Cuenca.

Estos servicios multimedia se pueden lograr al segmentar el gran ancho de banda que provee la fibra óptica, el cual está repartido como indica la tabla 4.1. Se puede alcanzar velocidades de hasta 27Mbps en el downstream y 10Mbps en el upstream con una eficiencia espectral de 4.5 y 1.6 bits/hz

Tabla 4.1 Distribución en el espectro de la red HFC

Servicio	Frecuencia	Ancho de Banda	Transmisión	Modulación
Video	55,25 MHz - 575 MHz	6MHz/canal (máx. 86)	Broadcast	VSB
Cablemodem	597 MHz	6 MHz	Upstream	QPSK
Cablemodem	31 MHz	6 MHz	Downstream	256QAM
PPV	8,9 MHz	6 MHz	Upstream	
PPV	106,5 MHz	6 MHz	Downstream	

La red de fibra óptica, originalmente tipo árbol, fue modificada en una parte para que, mediante arriendo de infraestructura, sea utilizada como backbone

para la red de datos de Suratel, cerrando un anillo. Actualmente se utiliza un 40% de la capacidad física total de la red.

En estas tres ciudades existe muchas similitudes con respecto a los equipos que se encuentran en cada headend diferenciándose entre si principalmente por su capacidad.

4.2.1 UBICACIÓN DE HEADENDS: ESQUEMA DE EQUIPAMIENTO

Para las ciudades de Guayaquil, Quito y Cuenca existe un Headend en cada ciudad en el cual se encuentra toda la infraestructura central para brindar los servicios de: Telefonía, PPV, Internet, y video. La infraestructura en cada ciudad es muy parecida por lo que hablaremos de manera en general de un sólo esquema.

Si desglosamos cada servicio que presta el grupo TvCable podemos entender de una mejor manera su infraestructura.

Para empezar describiremos el servicio de video con el que actualmente cuenta el grupo TvCable. En la figura 4.3 podemos observar los elementos que están involucrados.

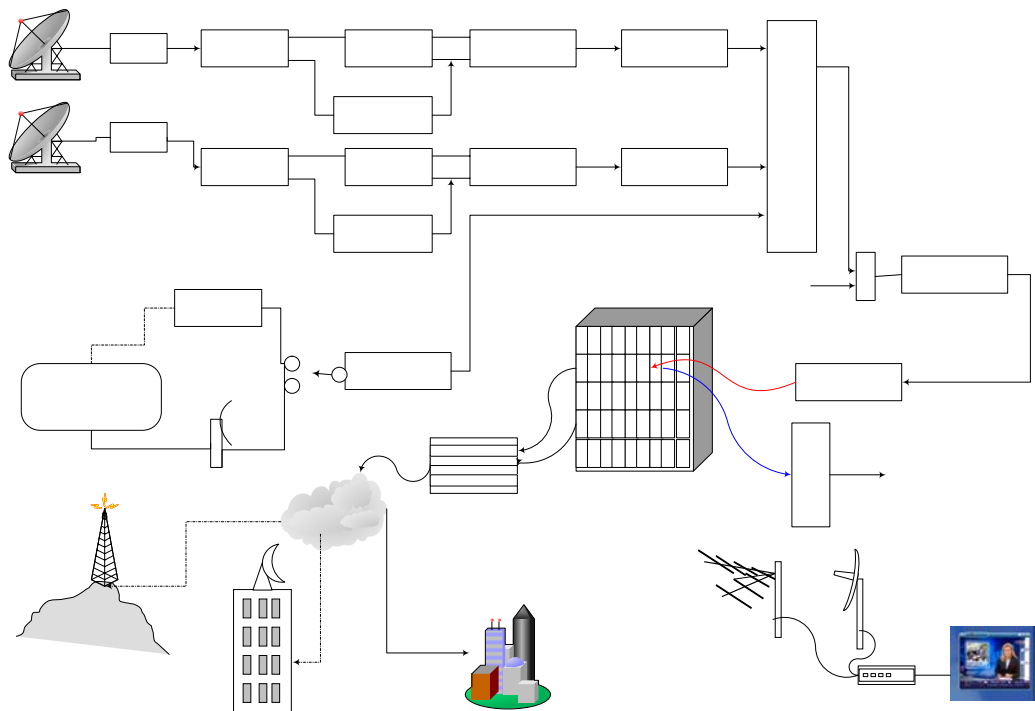


Figura 4.3 Esquema General de la Infraestructura del servicio de video

Existen dos fuentes de video que se presenta a los usuarios, los canales locales y los canales que se compran a nivel mundial y que generalmente se los adquiere mediante satélite.

ANTENA PAS5

ANTENA INTELSAT

LNB

RECEPTOR

VIDEO

CHANNEL MATIC

PROC. AUDIO SAP-ST-MON

AUDIC

TX OPTICO

FIBRA OPTICA

PROCESADORES

CANALES LOCALES CERRO DEL CARMEN

Los canales que provienen de señal satelital son codificadas y moduladas para su transmisión por la red de TvCable, estas señales son ingresadas en el conmutador junto con las señales de televisión local para ser enviadas como un solo paquete, son amplificadas y enviadas por fibra óptica hasta la red coaxial y finalmente al decodificador para ser vista por el usuario de acuerdo a su paquete de películas contratado.

Ahora se describirá el servicio de telefonía con el que actualmente cuenta el grupo TvCable. En la figura 4.4 podemos observar los elementos que están involucrados.

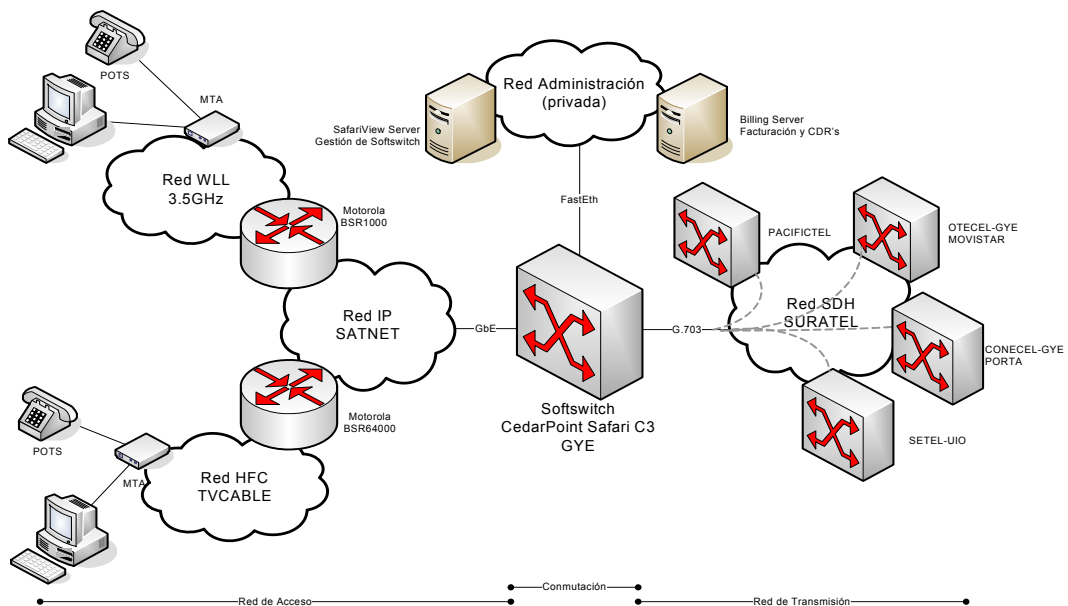


Figura 4.4 Esquema General de la Infraestructura del servicio de telefonía

El elemento esencial para prestar el servicio de telefonía es el SAFARI que es en sí la central telefónica en el cual están alojados todos los números telefónicos, existen dos formas de acceder a los clientes los cuales son la red HFC de TvCable o la red WLL. En este gráfico también podemos observar la interconexión que se tiene con los demás proveedores de telefonía ya sea fija o móvil.

Ahora se describirá el servicio de internet y la interconexión de los diferentes servicios con el que actualmente cuenta el grupo TvCable. En la figura 5.5 podemos observar los elementos que están involucrados.

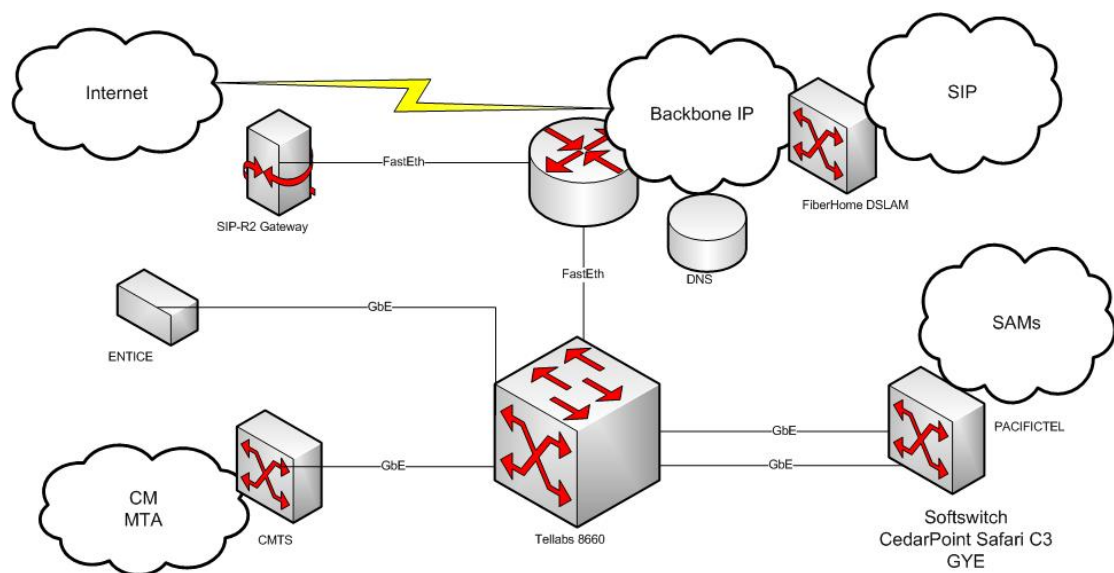


Figura 4.5 Esquema de interconexión de los diferentes servicios

El CMTS (Cable MODEM Terminal System) es el equipo en donde todos los MTA o CM que se encuentran en cada usuario se autentica y así puedan tener servicio de internet.

El safari (central telefónica) se interconecta con el CMTS y con los demás equipos principales a través del TELLABS que es la cabecera del Headend, las interfaces con los que cuenta pueden ser FastEthernet, GigaEthernet o ATM.

4.2.2 UBICACIÓN DE NODOS ÓPTICOS Y CARACTERÍSTICAS TÉCNICAS

En la ciudad de Guayaquil existen diferentes nodos ópticos los cuales están distribuidos por toda la ciudad.

El esquema básico de un nodo se puede observar en la figura 4.6

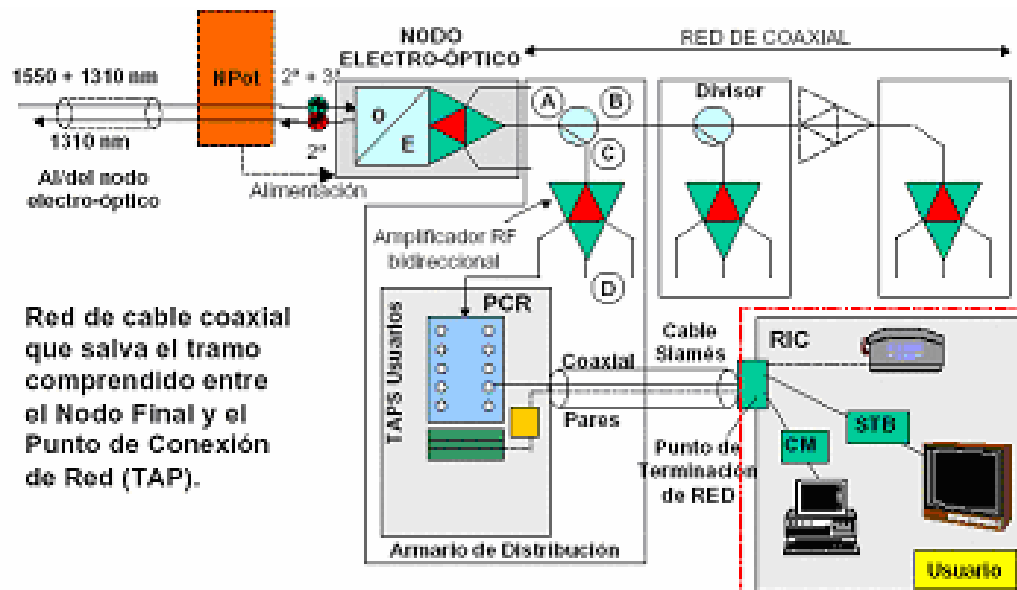


Figura 4.6 Esquema básico de nodos

Como se observa en la gráfica la fibra óptica que sale desde el headend con la información ya multiplexado en los canales de frecuencia que se deben utilizar según el servicio y su distribución que ya está establecida, llega a un convertidor opto-eléctrico el mismo que transforma las señales lumínicas en pulsos eléctricos que ahora viajarán a través del cable coaxial. Esta señal será amplificada cada cierta distancia para garantizar un cierto nivel de potencia que permita brindar servicios sensibles a errores como PPV, VOD o Internet.

4.2.3 REPETIDORES Y NIVELES ACEPTABLES PARA PRESTACIÓN DE SERVICIOS

Para la prestación de servicios se necesita una serie de equipos que permitan convertir la señal Óptica que viaja por la fibra óptica en RF que viaja por el cable coaxial y repetirla hasta las tomas en donde los clientes se puedan conectar. Para apreciar de mejor manera los elementos que se utilizan para todo es proceso se muestra la siguiente figura 4.7.

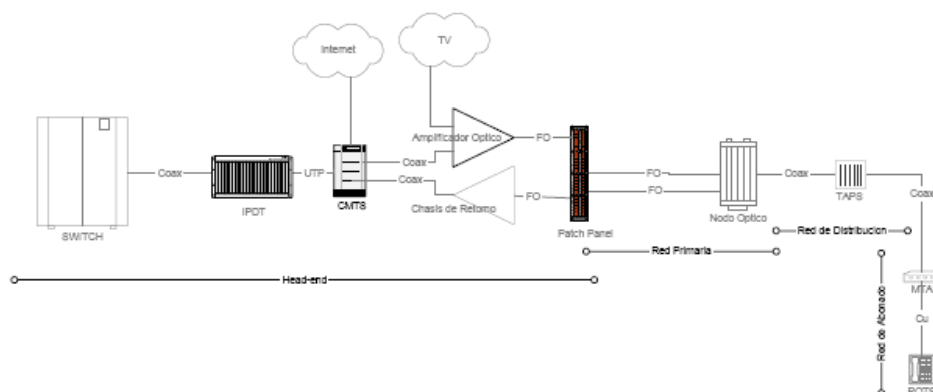


Figura 4.7 Red de Acceso a Clientes

4.2.3.1 TAPs.

Son elementos de línea pasiva (figura 5.8), que se caracterizan por mantener la señal íntegra, ya que de ahí parte la alimentación hacia el usuario. Se conecta con el nodo óptico a través de un coaxial 750

TAP 4-1**Figura 5.8 TAP 4-1**

4.2.3.2 NODO ÓPTICO

Consta de un receptor RM-8 y un transmisor óptico RPTV. El receptor RM-8 recibirá la señal desde el headend en forma de luz y la convertirá en una señal eléctrica en RF, la misma que será aplicada al módulo principal para la distribución al resto de amplificadores. El transmisor óptico RPTV recibirá las señales de datos IP desde el MTA y las señales de retorno del servicio de CATV desde el coaxial de la red de distribución, las convertirán en señales ópticas y las enviará en la fibra de retorno hacia el headend. Los equipos son marca General Instruments.

4.2.3.3 PATCH PANEL ÓPTICO

Recibe todos los cables de fibra de la red en la calle y provee los conectores para las fibras ópticas hacia los amplificadores ópticos y el chasis de retorno. Es un elemento pasivo

4.2.3.4 AMPLIFICADORES ÓPTICOS

Recibirán las señales RF provenientes de los moduladores de TV, del módulo de PPV Data Commander y del CMTS y los convertirá a señales ópticas por medio de luz láser a 10dB. Los amplificadores ópticos son marca General Instruments modelo AM-750AT. Cada módulo puede manejar hasta 6 nodos ópticos (Figura 5.9).



Figura 4.9 Amplificador Óptico

4.2.3.5 CHASIS DE RETORNO

Son los convertidores óptico-electrónicos de la señal de retorno de los nodos. Estarán a cargo de enviar la información correspondiente al servicio de CATV a los módulos Data Commander y los datos IP de retorno al CMTS por medio de cables coaxiales. Son marca Omnistar.

Una vez descrito todos los elementos que conforman la red de acceso hacia los clientes, hablaremos sobre los niveles de decibelios que se necesitan para brindar los diferentes servicios de TvCable. En la figura 4.10 se muestran todos los elementos de la red HFC y los niveles mínimos para garantizar un cierto nivel de QOS que permita ganar usuarios día a día.

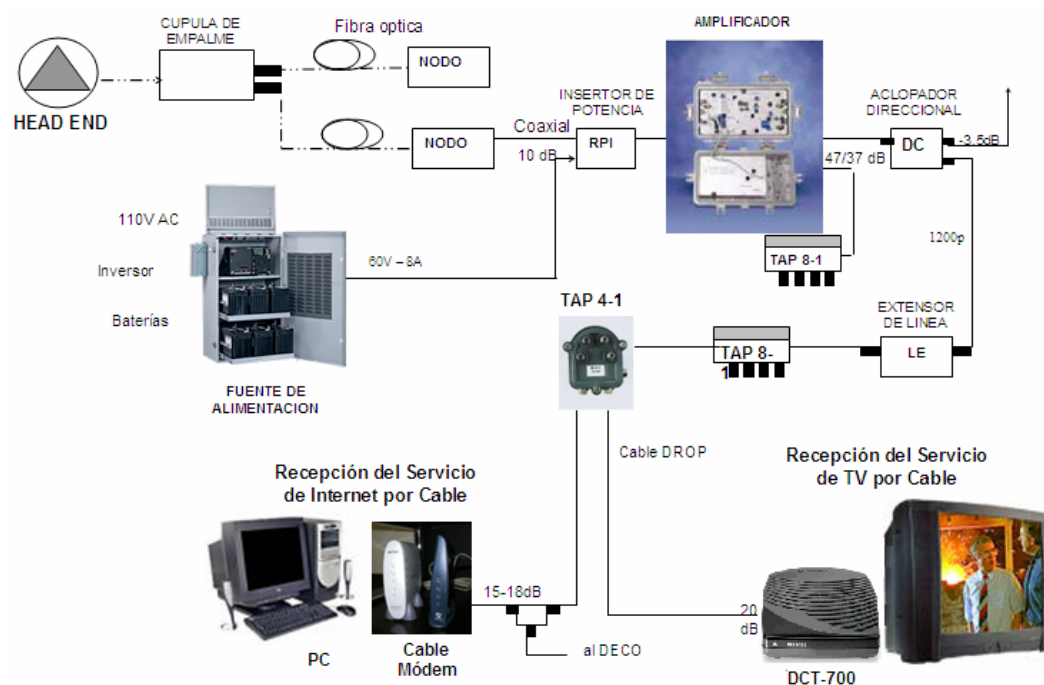


Figura 4.10 Esquema de redes de Acceso y niveles de señal

Como se puede observar en la figura se debe tener niveles de 20 dB a la entrada del DCT-700 para que la señal de video sea nítida.

Cabe recalcar que en la actualidad, el departamento de redes han tenido varios problemas con el servicio de telefonía ya que se han presentado varios problemas de degradación de la señal, se han cumplido una serie de cronogramas de mantenimientos correctivos en cada nodo óptico y repetidor en toda la red HFC que se encuentra en la ciudad de Guayaquil. En muchos casos se han hecho necesario cambiar algunos equipos ya que no cumplían con los parámetros mínimos (señal / ruido). Estos parámetros se registran en el grafico anterior, los mismos que se mantienen en constante monitoreo ya que son los puntos críticos en el servicio de Video en Demanda.

Además se ha realizado el cambio de cable coaxial, ya que este no es el mismo al que se instalaba hace dos años. Los conectores que se utilizan tienen un mejor desempeño y disminuye la atenuación de la señal siendo óptia para el usuario final.

Capítulo 5

5. REGULACIÓN, DISEÑO, PLANIFICACION Y EXPANSIÓN DEL SISTEMA LVOD EN LA RED

Se analizará el espectro de frecuencias utilizados por el grupo TV-CABLE para determinar en que canales se podrá incluir el nuevo servicio (video en demanda), se describirá los equipos que se incrementarán tanto en el headend como en la red de acceso tales como servidores de video, los controladores de red y los equipos terminales (DCT 700).

Para finalizar se realizará una proyección de las áreas comerciales para la instalación de los nuevos STB.

5.1 PARÁMETROS REGULATORIOS A CONSIDERAR PREVIO A LA INSTALACIÓN DE SERVICIOS DE VOD PARA LAS CIUDADES DE GUAYAQUIL, QUITO Y CUENCA.

El grupo TVCable fue fundado en 1986 y desde entonces ha brindado servicios de televisión codificada para las principales ciudades del Ecuador. En la actualidad presta diferentes servicios de telecomunicación y desde hace 9 años atrás se ha venido prestando el servicio de PPV en varias ciudades del Ecuador. Con lo que respecta al marco regulatorio, en el artículo 2 de la constitución de la empresa constan los servicios que puede prestar el Grupo TVCable hacia sus clientes, incluyendo los servicios de video en demanda como PPV y VOD. Por lo que no sería necesaria la modificación de la constitución de la empresa para la prestación de este nuevo servicio.

Un parámetro importante para la prestación de servicios codificados por cable es la utilización del espectro de frecuencia que está establecida en institutos internacionales como la RTPSTC y nacionales como la CONARTEL. Se establece que el espectro no puede superar los 900 MHZ en lo que respecta a la utilización del espectro de frecuencia del cable coaxial. Este valor está contemplado en la infraestructura de los equipos que en la actualidad existen en los headend de cada ciudad.

5.1.1 SOLICITUD A PRESENTAR A LA CONARTEL PARA LA APROVACIÓN DE CAMBIOS EN EL HEADEND

Previo a la prestación del servicio de Video en Demanda en las ciudades de Guayaquil, Quito y Cuenca se debe de llenar una solicitud que se detalla en el anexo A.

En la solicitud que se va a presentar se deben de especificar los datos del Ing. Que hará la presentación de dicha solicitud. En el grupo TVCable existe un departamento encargado de la parte de redes e infraestructura que sería el responsable de realizar dicha solicitud y por ende los datos que se necesitan. Con respecto a las características técnicas de los equipos, a lo largo del desarrollo de la tesis ya se han descrito las características de todos los equipos necesarios para la prestación de servicios de Video en Demanda y con respecto a la programación que se desea presentar, esta será de contenido variado ya que solo es necesario que el Grupo TVCable cuente con los derechos de la publicación de los videos que se estimen de interés hacia los usuarios finales.

En la instalación del nuevo servicio de VOD en demanda no se va a requerir la instalación de antenas, ya que se utilizará los paquetes ya contratados por el Grupo TVCable con las diferentes cadenas internacionales de Televisión.

Con lo que respecta a los equipos a utilizarse en la transmisión y recepción las características técnicas están descritas a lo largo del desarrollo de esta tesis.

En el tema de la programación que se va a presentar puede ser de gran variedad ya que este tipo de servicio permite ofrecer una variedad de contenidos. La condición que se debe de tener en cuenta previo a la publicación de los contenidos es tener los derechos necesarios como se ha venido trabajando con PPV desde 1.998.

5.2 SERVICIOS PRESTADOS POR EL GRUPO TVCABLE Y DISTRIBUCIÓN DE FRECUENCIAS UTILIZADAS

Los servicios que ofrece el Grupo TvCable son los siguientes:

- canales TV terrestres, satélite y producción propia: habitualmente en diferentes opciones de contratación, que implican distintos canales y precios.

- vídeo a la carta (PPV): selección de películas de entre las ofrecidas.
- datos e Internet mediante módems de cable.
- telefonía básica y RDSI, con opción de acceso a Internet por medio de la tecnología de VOIP.
- alquiler de líneas y fibras: de interés para comercios y oficinas que necesitan de líneas dedicadas, tanto de telefonía como de datos.

En general, estos servicios requieren de una transmisión bidireccional por la red. La transmisión de señal hacia el abonado se dice que se lleva a cabo mediante el canal descendente o directo, mientras que las que parten del abonado se realizan a través del canal ascendente o de retorno.

En última instancia, las señales llegan a los abonados mediante coaxial, transmitidas en Radiofrecuencia (RF). El espectro de la señal transmitida se divide en diferentes segmentos, transportando cada uno de ellos la información relativa a uno de los servicios prestados. El plan de frecuencias definido para una red de cable debe seguir la normativa del Reglamento Técnico y de Prestación del Servicio de Telecomunicaciones por Cable (RTPSTC).

El canal directo ofrece los servicios de TV analógica y digital, telefonía y datos, y cubre el rango de 86..862 MHz. El canal de retorno, que permite la prestación de servicios interactivos, telefonía, datos y contratación instantánea de servicios PPV (Pay Per View, Pago Por Visión) corresponde al rango 5.65 MHz. Es de destacar que la información de gestión de los equipos circula por el cable junto con el resto de servicios, tanto en el sentido directo como en el retorno.

La tabla 5.1 muestra la distribución en frecuencia de los distintos servicios de la red del grupo TvCable.

Tabla 5.1 Distribución en frecuencia de los distintos servicios actuales

Descendente (MHz)	Servicio	Capacidad total
86-128	Audio digital	145 Canales
	Gestión de equipos	1 Canal
	Gestión terminal de abonado	1 Canal
128-568	Canales de TV analógica	55 Canales (8 MHz/canal)
568-606	Telefonía	240 líneas / zona 500 hogares (6 MHz)
	Datos	30Mbps / zona 500 hogares (6 MHz)
606-862	Canales de TV digital DVB-C	197 Canales de 4Mbps
Ascendente (MHz)	Servicio	Capacidad total
5-8.3	Sin servicio	-
8.3-65	Telefonía	240 líneas / zona 500 hogares (6 MHz)
	Datos	2,56Mbps / zona 500 hogares (1,8 MHz)
	Retorno terminal de abonado	1 Canal
	Gestión de equipos	1 Canal

Para realizar esta distribución se utilizan los SEM's que son los que proveen de encriptación y modulación a nivel de QAM y QPSK.

Cada SEM trabaja en un rango de 66 MHz, y dentro de este rango de frecuencia existen subcanales. En la actualidad el Grupo TvCable cuenta con 3 SEM's que tienen la siguiente distribución.

Tabla 5.2 Distribución de los 3 SEM's actuales

SEM-1		SEM-2		SEM-3	
Canal 1	567	Canal 1	639	Canal 1	693
Canal 2	573	Canal 2	645	Canal 2	699
Canal 3	579	Canal 3	651	Canal 3	705
Canal 4	585	Canal 4	657	Canal 4	711
Canal 5	603	Canal 5	663	Canal 5	717
Canal 6	609	Canal 6		Canal 6	723
Canal 7	627	Canal 7		Canal 7	729
Canal 8	633	Canal 8		Canal 8	735

Como podemos apreciar aún se tiene la disponibilidad de introducir equipos SEM adicionales para completar el ancho de banda establecido por los Organismos de control Internacional y Nacional como lo son la CONARTEL y la SENATEL, además actualmente existe la capacidad de el SEM-2 de 3 portadoras más.

5.3 POSIBLE ASIGNACIÓN DE CANALES EN EL ESPECTRO PARA PODER PRESTAR EL SERVICIO VOD

La distribución de frecuencia de los equipos que se necesita para la prestación de estos servicios debería quedar de la siguiente forma.

- Para el equipo RPD-2000 cuya funcionalidad ya se la explicó con anterioridad debe de manejar la frecuencia 8.09 MHz
- Para el equipo OM-1000 necesita manejar dos frecuencias ya que necesita manejar los tanto analógico como digital. Para la parte digital se manejará la frecuencia 75.25 MHz y para la parte analógica se deberá manejar la frecuencia 104.2 MHz.
- Para la recepción de los DCT 700 la frecuencia que se asignaría es 669.25 Mhz.
- Los canales analógicos serán desde el canal 2 (55.25Mhz) al canal 87 (556.25Mhz)
- El canal 116 (745.25Mhz) se utilizará para la calibración de los equipos que se encuentran distribuidos. Este canal se utiliza por el departamento de redes principalmente.

- Los canales digitales empiezan en 567Mhz hasta 735Mhz.
- La portadora del Downstream del Internet es de 597 Mhz.
- Las portadoras de los CM (upstream) es de 23 a 32 Mhz.

En la tabla 5.2 se describió la distribución de las frecuencias que utiliza el SEM. Ahora se lo hará de una manera gráfica para que se pueda apreciar de una mejor manera las portadoras que se van a distribuir.

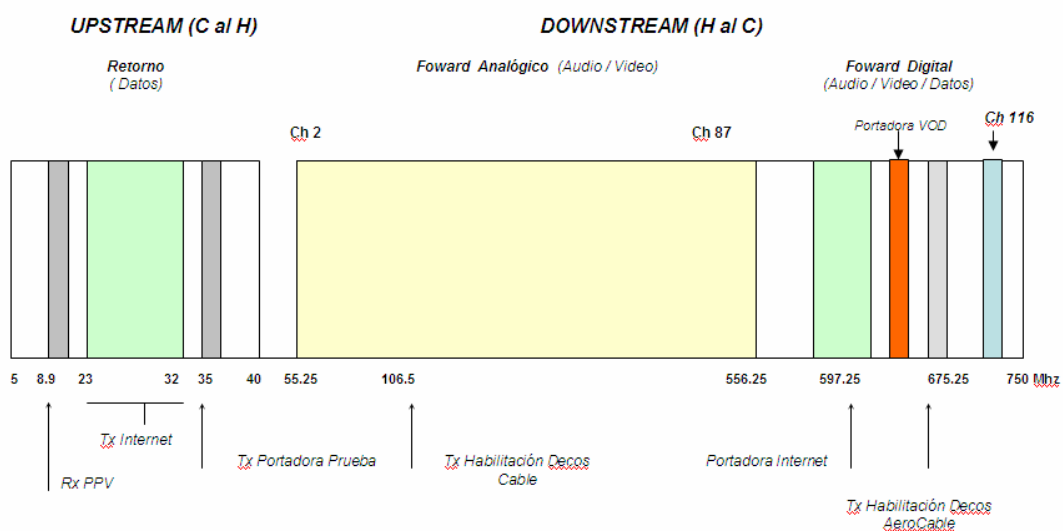


Figura 5.1 Distribución de los canales en el espectro

Las asignaciones de frecuencias que se han descrito aseguran el buen funcionamiento de todos los servicios y permitirán el funcionamiento del nuevo servicio de Video en Demanda. La portadora VOD que se señaló permitirá realizar las descargas desde el servidor MDS200 hacia cada terminal. Este contenido se presentará en cualquier canal que esté designado para mostrar el contenido que se ha requerido por el cliente.

5.4 PROYECCIÓN DE LA DEMANDA EN 5 AÑOS

Si hacemos una mirada al mundo de cual ha sido su reacción a la inserción del servicio del Video en Demanda nos podemos dar cuenta que la aceptación ha sido muy favorable lo cual ha traído a mayores inversiones para desarrollar nuevos y mejores sistemas que permitan prestar video en Demanda a más gente presentando más opciones.

El gráfico 5.1 podemos observar cual ha sido el crecimiento a nivel mundial de los abonados que cuentan con el servicio de Televisión por Cable (barras azules) y el crecimiento de los usuarios que han adquirido servicios de video en demanda (barras moradas).

Como se puede observar en la gráfica el año 1998 la televisión por cable ha percibido ingresos de 1300 millones de dólares mientras que los servicios de Video en Demanda tan solo de 500 millones de dólares. Al transcurrir los años esta diferencia de ingresos se ha acortado por la gran demanda hasta tal punto de tener casi los mismos ingresos.

Es por esta razón que el grupo TvCable está muy interesado en la implementación de este nuevo servicio para poder percibir mayores ingresos y estar en la vanguardia en el campo de las telecomunicaciones.

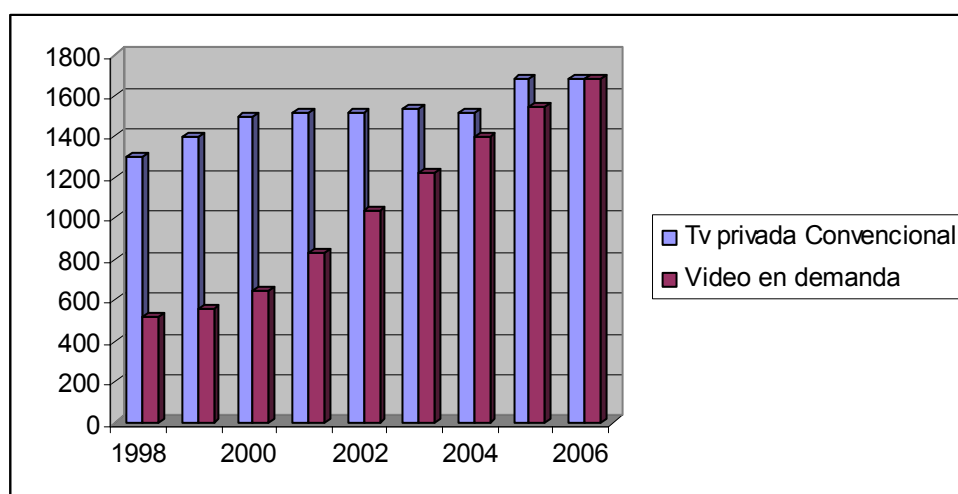


Figura 5.2 Cuadro comparativo de ingresos a nivel mundial

Actualmente el grupo TvCable cuenta con el servicio de PPV, este servicio se presta desde el año 1998 con una gran acogida.

De acuerdo a servicios anteriores y al ver como estos tenían su impacto en los consumidores se define una tasa de crecimiento anual del 5% al 8% dependiendo de publicidad y planes que se ofrezcan.

El grupo TvCable tiene aproximadamente 109000 usuarios a nivel nacional. En la ciudad de Guayaquil hay 53800 abonados, en Quito 48700 y Cuenca 6500 abonados. El objetivo inicial al introducir este nuevo servicio es del 8% a nivel de las 3 ciudades más grandes del Ecuador que se desglosa a continuación.

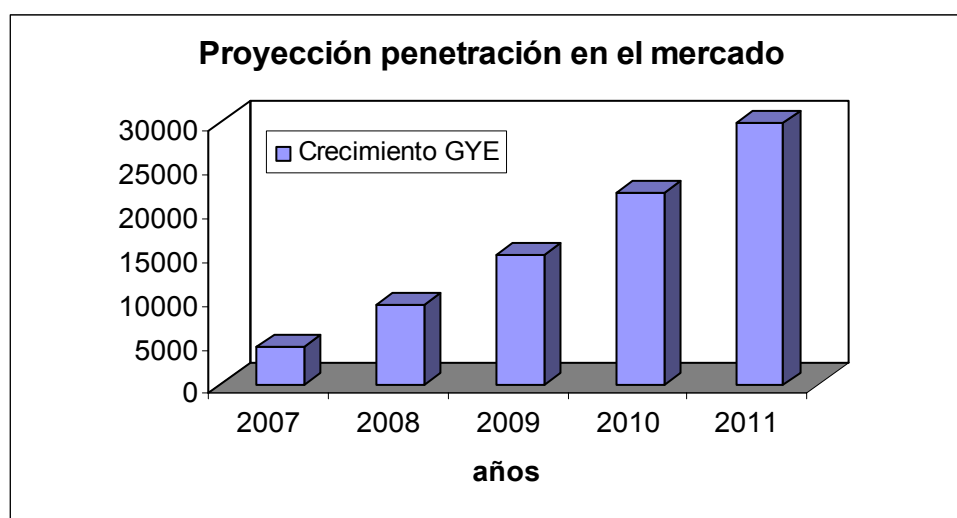


Figura 5.3 Crecimiento de usuarios en Guayaquil

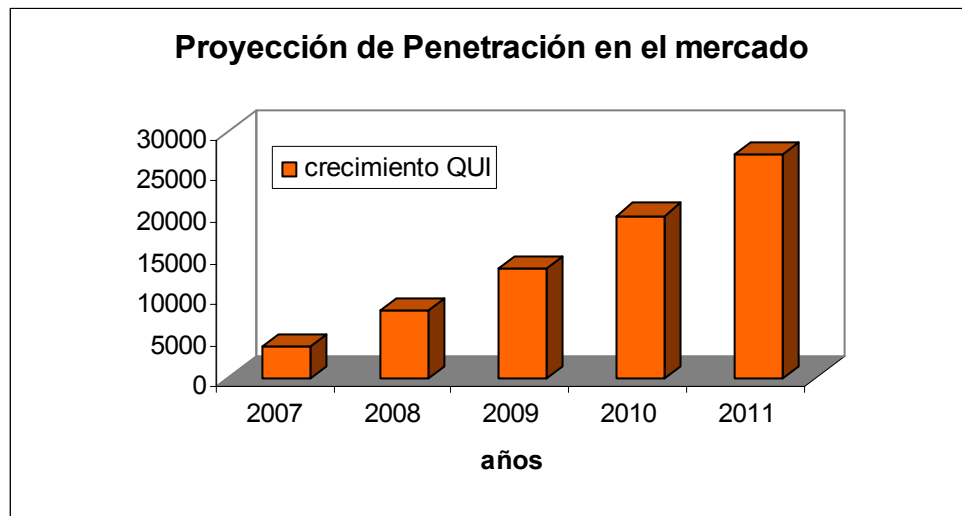


Figura 5.4 Crecimiento de usuarios en Quito

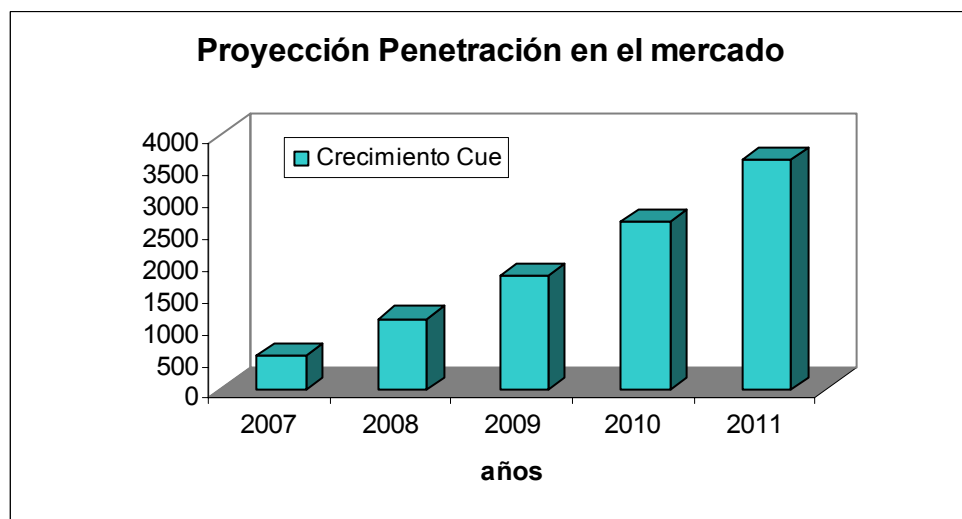


Figura 5.5 Crecimiento de usuarios en Cuenca

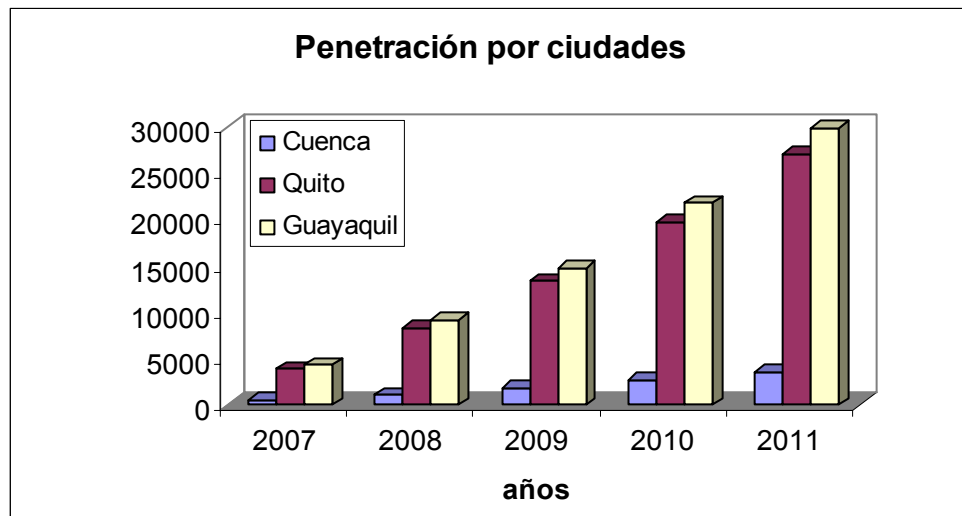


Figura 5.6 Crecimiento de las 3 ciudades a implementar VOD

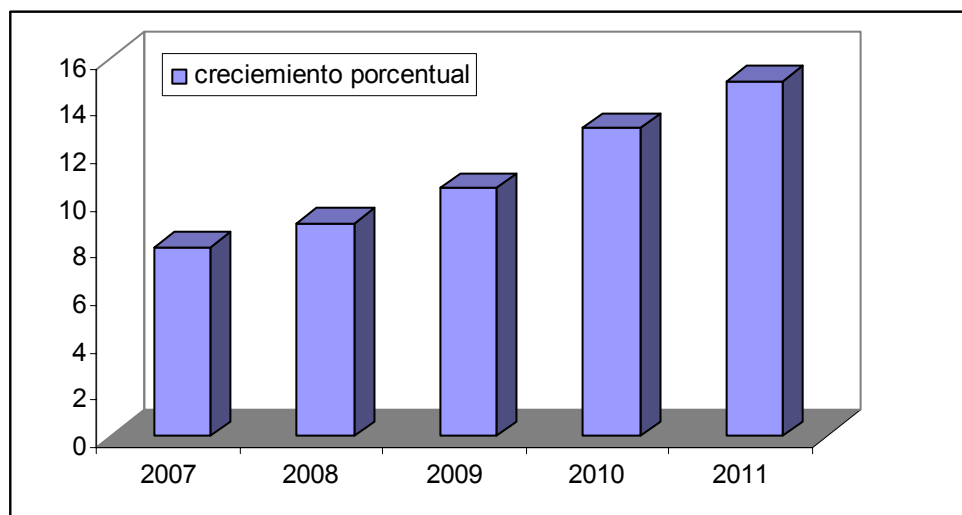


Figura 5.7 Crecimiento porcentual de las ciudades

En la figura 5.7 se determina el crecimiento porcentual al que se introducirá en el mercado los equipos DCT-700 para brindar el servicio VOD, se estima para el primer año una introducción del mercado del 8% y para el quinto año se espera finalizar con un crecimiento del 15% anual. Este cálculo se ha estimado sobre un valor actual de clientes, estimando que su crecimiento no variaría en los próximos 5 años. A continuación se hará los cálculos de ancho de banda necesarios para los enlaces interurbanos (WAN) entre las redes locales (Guayaquil, Quito, Cuenca) y pueda soportar los requerimientos que se presentaran al momento de la implementación del servicio.

Los cálculos que se realizaron en el capítulo 2 con respecto al ancho de banda necesario para prestar el servicio de VOD serán de gran ayuda en nuestra proyección. Hemos escogido para iniciar una gama de 100 videos de calidad MPEG-2, que es la calidad que en la actualidad el Grupo TvCable brinda a sus clientes con los paquetes que se pueden contratar. El número de redes con las que se va a trabajar será de 3 que corresponden a cada ciudad en la que se piensa prestar el servicio de VOD a nivel nacional.

En la gráfica 2.6 se puede observar que con un caché del 40% se tiene una probabilidad de acierto del 0.71 si sumamos 40 videos más de mirroring la probabilidad de acierto al tener 80 videos es del 0.92, es decir, la probabilidad de fallo (petición que se realiza al servidor principal) es del 0.08. Los cálculos de ancho de banda para los enlaces interurbanos se basan en esta probabilidad de fallo, descrita a continuación.

Tomando como referencia la población total de usuarios VOD descrita en la figura 5.6 y asumiendo que un 2.3% de dicha población estarán realizando una petición al mismo tiempo del 20% de los videos que no se encuentran incluidos en el servidor local ya sea por caching o mirroring tenemos los siguientes resultados:

QUITO.

En cinco años se tendrá una población de 27028 usuarios, el 2.3% de este universo es:

$$x = 27028 * 0.023$$

$$x = 621$$

621 usuarios estarán pidiendo un requerimiento de un video, de los cuales un 8% de ellos no se encuentran en el servidor local, haciendo la petición al servidor principal, este número de requerimientos es:

$$x = 621 * 0.08$$

$$x = 49$$

49 son los requerimientos que atenderá el servidor principal, como la tasa de transferencia para un video MPEG-2 es de 1.5Mbps para garantizar una calidad de video excelente, el total de ancho de banda requerido para el enlace Guayaquil- Quito sería:

$$x = 49 * 1.5$$

$$x = 77Mbps$$

CUENCA.

En cinco años se tendrá una población de 3606 usuarios, el 2.3% de este universo es:

$$x = 3606 * 0.023$$

$$x = 85$$

85 usuarios estarán pidiendo un requerimiento de un video, de los cuales un 8% de ellos no se encuentran en el servidor local, haciendo la petición al servidor principal, este número de requerimiento es:

$$x = 85 * 0.08$$

$$x = 6$$

6 son los requerimientos que atenderá el servidor principal, como la tasa de transferencia para un video MPEG-2 es de 1.5Mbps para garantizar una calidad de video excelente, el total de ancho de banda requerido para el enlace Guayaquil- Cuenca sería:

$$x = 6 * 1.5$$

$$x = 10Mbps$$

GUAYAQUIL.

En cinco años se tendrá una población de 29859 usuarios, el 2.3% de este universo es:

$$x = 29859 * 0.023$$

$$x = 686$$

686 usuarios estarán pidiendo un requerimiento de un video, de los cuales un 8% de ellos no se encuentran en el servidor local, haciendo la petición al servidor principal, este número de requerimiento es:

$$x = 686 * 0.08$$

$$x = 54$$

54 son los requerimientos que atenderá el servidor principal, como la tasa de transferencia para un video MPEG-2 es de 1.5Mbps para garantizar una calidad de video excelente, el total de ancho de banda requerido para este enlace sería:

$$x = 54 * 1.5$$

$$x = 82Mbps$$

Se ha elegido el 2.3% por estudios realizados a nivel internacional en cuales se determina la posibilidad de que un número n de usuarios de un universo accedan al mismo tiempo a un servicio de este tipo. Las especificaciones de ancho de banda que se requieren para el funcionamiento de VOD están

dentro de la disponibilidad que tiene TvCable en sus enlaces y que les pertenece.

Ahora que se la realizado los cálculos para los requerimientos del ancho de banda en los enlaces troncales, se realizará un pequeño análisis sobre los requerimientos de la red interna (LAN). Los requerimientos de video que se hagan dentro de cada ciudad deberán ser atendidos en primera instancia por el servidor que se encuentra en cada headend, y si este no se encuentra deberá ser atendido por el servidor de nivel superior (cabecera). Lo que queda por describir es cuantos requerimientos puede soportar la red HFC de cada ciudad, y la respuesta está en la calidad de señal que se tenga en cada nodo y repetidor de la señal.

5.5 ANÁLISIS DE ÁREAS COMERCIALES, EN DONDE SE DISTRIBUIRÍA LOS NUEVOS EQUIPOS TERMINALES DCT 700

Luego de haber determinado el alcance que se desea de tener en las ciudades que se ha pensado la implementación de este nuevo servicio queda por decidir las áreas en las que se empezará con la distribución de los STB

(DCT700) los mismos que una vez que se cuente con los equipos necesarios en cada headend podrán empezar a brindar el servicio.

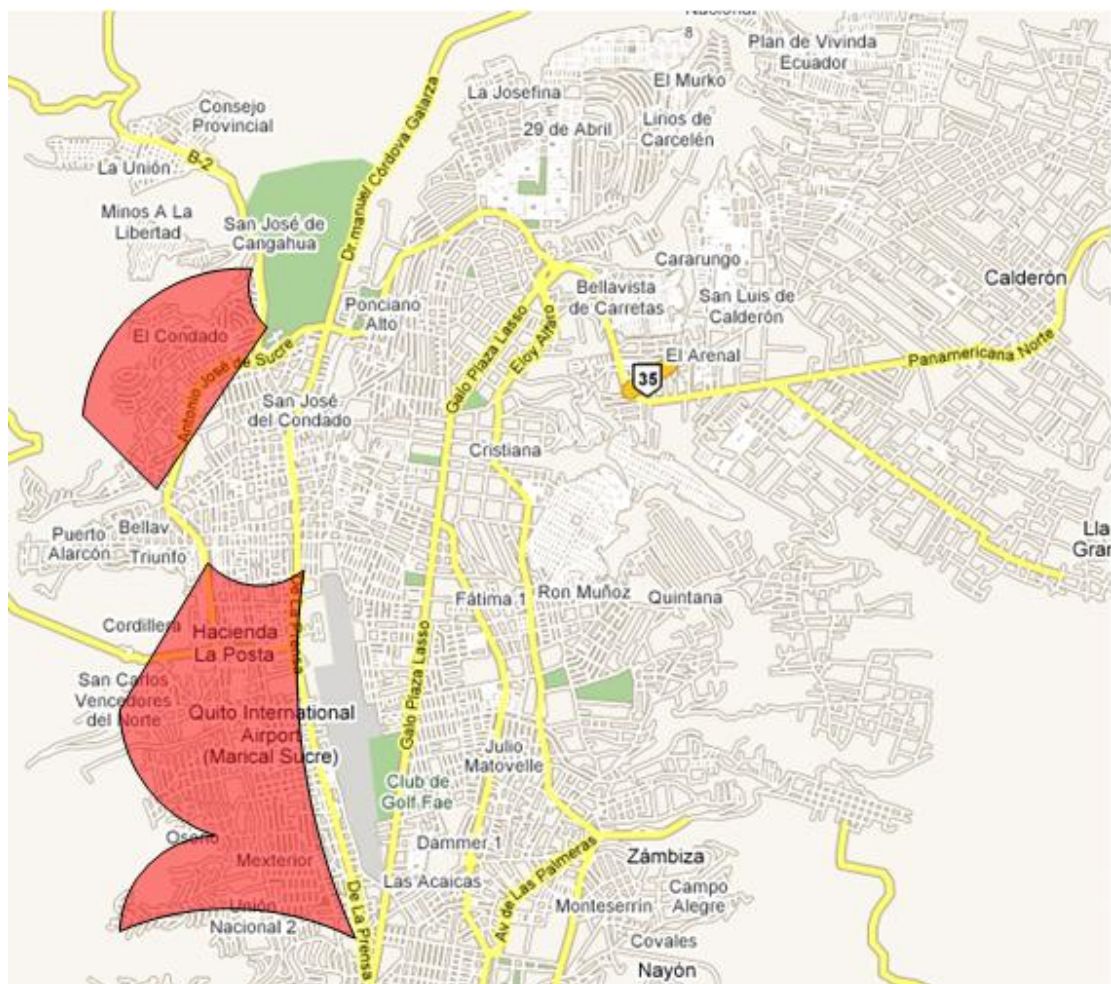
Algo que se ha venido recalando durante todo el desarrollo de la tesis es la importancia de la compatibilidad de los equipos que se están instalando actualmente con los que han existido desde antes para la prestación del servicio de PPV.

Para hacer este análisis nos hemos basado de las peticiones que actualmente existen del servicio de Pay per View en las tres ciudades y de esto se ha determinado los lugares estratégicos en los que se debe empezar a cambiar los equipos terminales. Para ser más específicos en cada ciudad.

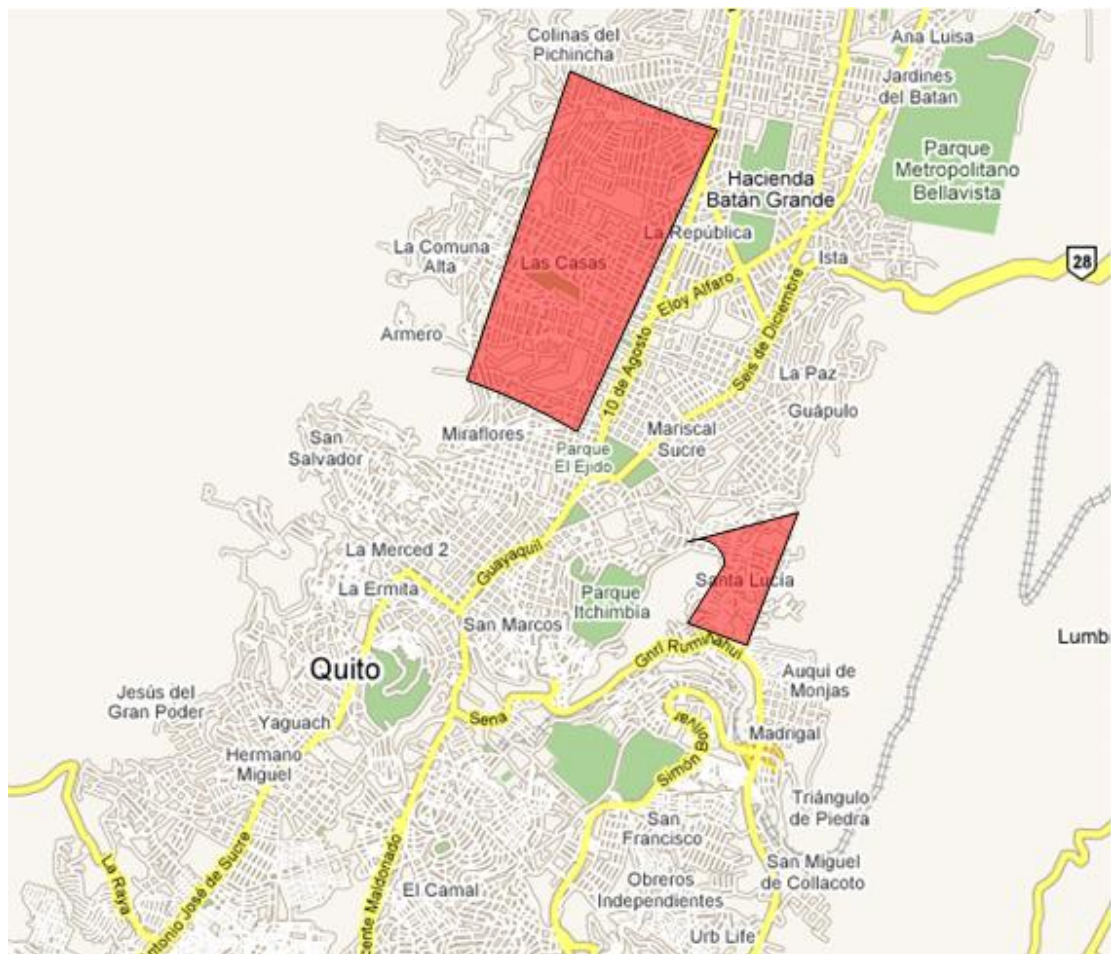
Para las ciudades de Guayaquil de Quito y Cuenca se ha determinado que el 90% de las peticiones que se tienen por servicios de video pagado en la actualidad se reciben de las zonas que se marcan en los planos que corresponden a Guayaquil, Quito y Cuenca respectivamente. Las zonas que se describen a continuación se han estimado las más atractivas para empezar con la distribución masiva de los DCT que es el equipo terminal necesario para brindar el servicio de Video en Demanda, si existe algún requerimiento para brindar el servicio en una zona no descrita con anterioridad no existe ningún problema para el cambio del equipo ya que el

servicio estaría listo a brindarse en el momento que se habilite el servidor en el Headend.

En la ciudad de Quito las zonas residenciales que obtienen la mayor parte de los requerimientos de Pay Per View son las siguientes: iñaquito, la mañosca, la gasca, el condado, el bosque, aeropuerto, el batan, granda centeno, las casas, y se encuentran sectorizadas y descritas a continuación



Mapa 5.1 Quito Norte



Mapa 5.2 Quito Centro

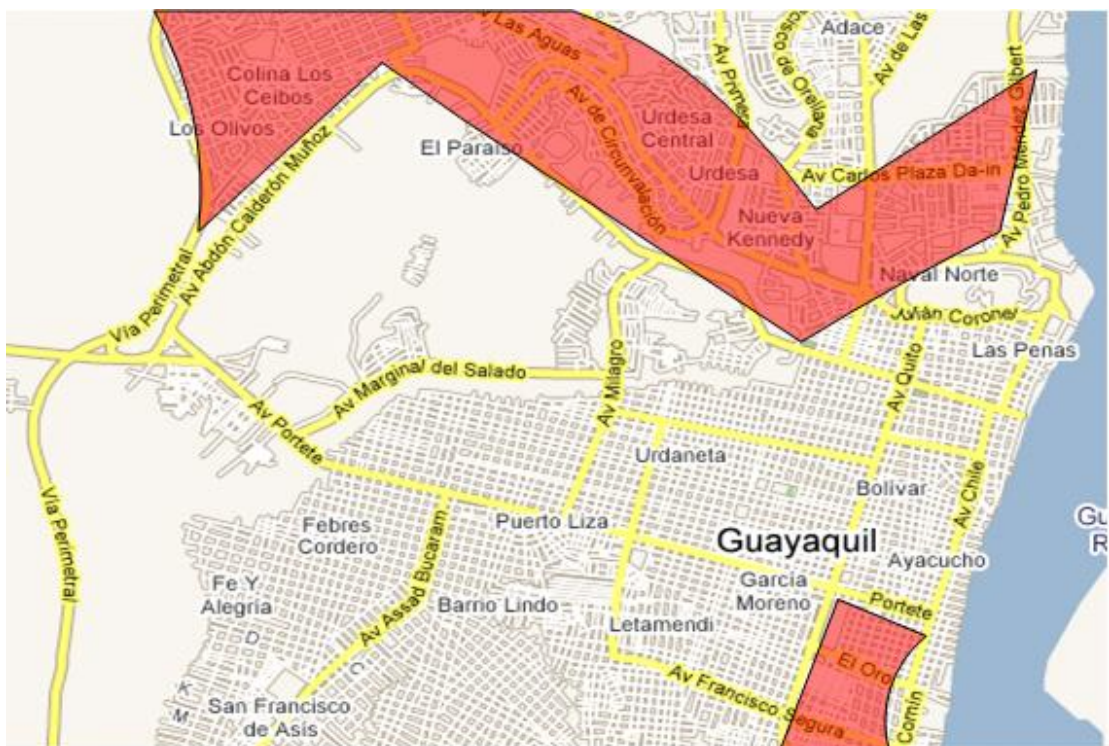


Mapa 5.3 Quito Sur

En la ciudad de Guayaquil las zonas residenciales que obtienen la mayor parte de los requerimientos de Pay Per View son las siguientes: Samanes, Garzota, Puntilla, Ceibos, Los olivos, Urdesa, Urdesa Central, Kennedy, Nueva Kennedy, Centenario, Alborada, y se encuentran sectorizadas y descritas a continuación



Mapa 5.4 Guayaquil Norte

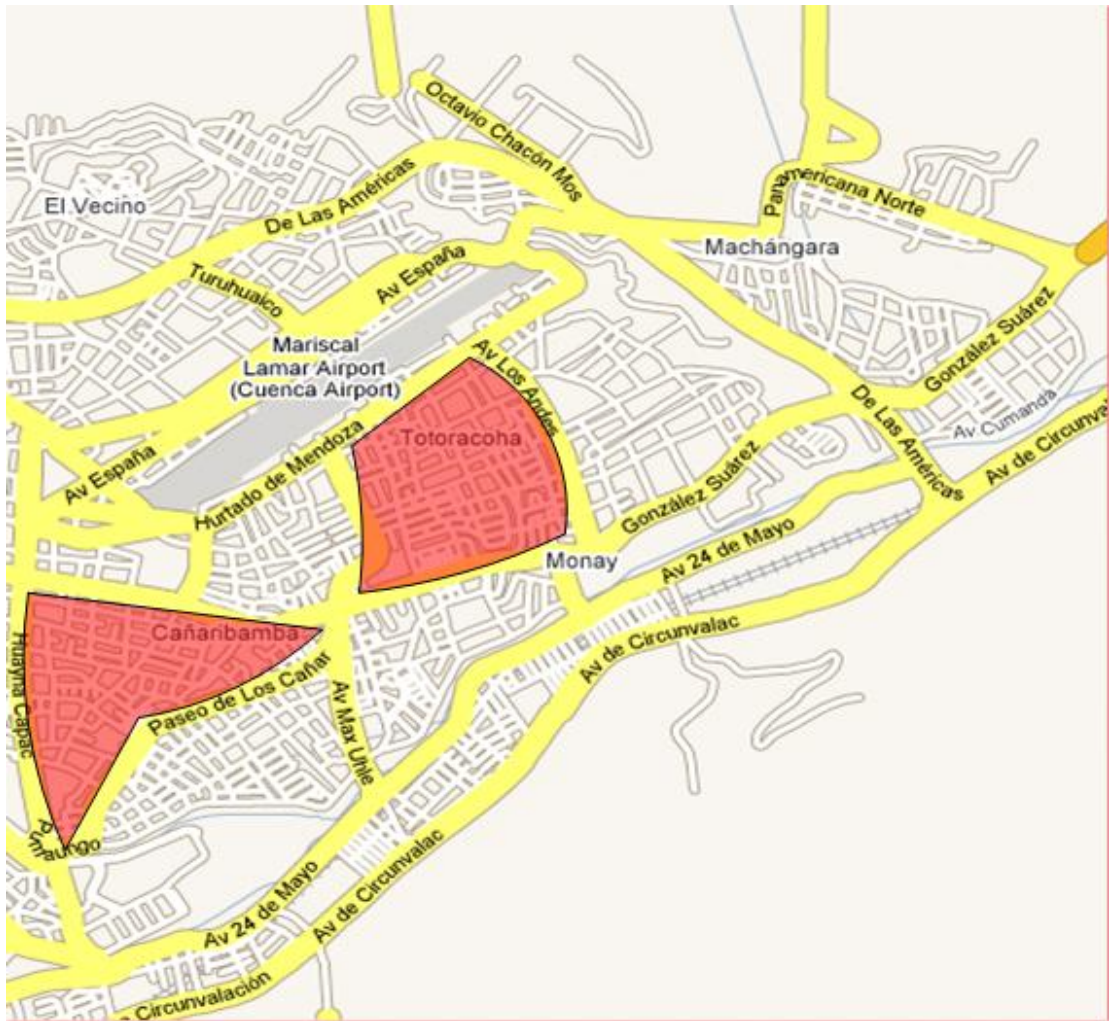


Mapa 5.5 Guayaquil Centro

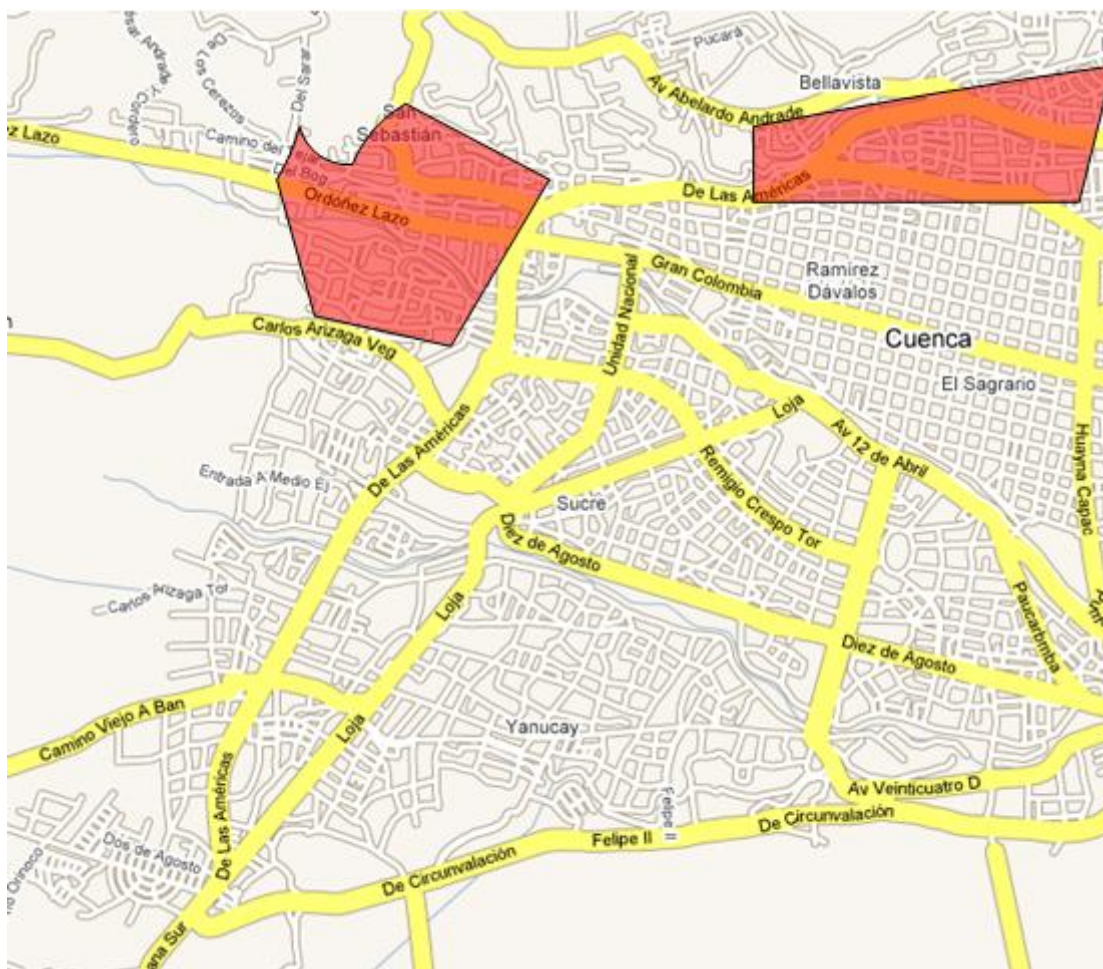


Mapa 5.6 Guayaquil Sur

En la ciudad de Cuenca las zonas residenciales que obtienen la mayor parte de los requerimientos de Pay Per View son las siguientes: puertas del sol, chuyallamba, cdla. los ingenieros, las fincas, el sector del estadio, totoracocha, Cañaribamba, y se encuentran sectorizadas y descritas a continuación



Mapa 5.7 Cuenca Este



Mapa 5.8 Cuenca Oeste

Ya que los índices que se han estimado durante los próximos 5 años la distribución de los equipos se debe hacer de igual manera de forma paulatina para que se pueda alcanzar a brindar el servicio a todos los clientes que deseen adquirir este nuevo servicio, que se anticipa con una gran acogida en el mercado actual.

CONCLUSIONES

1. El proyecto expuesto en esta tesis de grado propone una solución innovadora para la prestación del servicio de Video en Demanda. En el pasado, los clientes para solicitar un video tenían que notificarlo telefónicamente a su call center para su aprobación, con la implementación de VOD el requerimiento se realiza directamente al equipo remoto.
2. La integración de las empresas TVCable, Satnet y Suratel en el grupo Corporativo TV Cable permitirá un rápido desarrollo del servicio VOD (Video en demanda), ya que utilizará en la mayoría de casos la infraestructura de red ya instalada en las principales ciudades del país para proveer este servicio y además gracias a esta integración ya se pueden ofrecer servicios Triple Play con mayores beneficios.
3. Se realizó el estudio de tráfico para dimensionar los enlaces entre los servidores cuando se realiza los requerimientos entre servidores interurbanos y las peticiones locales.

4. Se iniciará la prestación del servicio de VOD (Video en Demanda) por medio de la red HFC. Debido a que la red está ampliamente difundida y tiene una sólida infraestructura esta permite una rápida instalación y puesta en marcha. Esto permitirá al grupo TVCable entrar de una manera altamente competitiva al mercado de las telecomunicaciones en cuanto a cobertura y calidad de servicio.
5. El servicio de VOD a través de la red HFC permitirá brindarle al usuario la provisión de todos los servicios de telecomunicaciones a través de una sola plataforma. Esto reduce significativamente el costo de instalación, consecuentemente un menor precio para el usuario.
6. Los equipos descritos en esta tesis fueron escogidos basándose en aspectos tanto técnicos como financieros. El Server VOD y el equipo remoto DCT-700 son parte fundamental de la puesta en marcha, por lo que se realiza una descripción técnica de cada equipo propuesto y de la plataforma que estos pueden manejar.
7. La coordinación en el momento de instalación de los equipos es básica para evitar perjudicar al usuario con prolongados cortes en el servicio de video o los otros servicios con los que cuenta el Grupo TvCable.

8. La distribución de frecuencias que se ha destinado para la ubicación de los equipos que son necesarios para la prestación del servicio VOD garantizan el óptimo funcionamiento de los equipos, representando esto buenos niveles de calidad de video que se ofrecerá a los clientes.

9. Los índices de expansión que se han presentado al momento de realizar el estudio para dimensionar la infraestructura necesaria se ha basado en países en donde el servicio de VOD ha tenido gran acogida por lo que se estima que podrán soportar los requerimientos que se presenten por los clientes del Grupo TvCable.

10. La arquitectura Proxy-tree que se ha escogido garantiza un buen funcionamiento del servicio de video en demanda garantizando la escalabilidad y confiabilidad del mismo y una mejor administración previendo posibles problemas en la red.

RECOMENDACIONES

1. Es necesaria la organización adecuada del diseño para cumplir con las obligaciones con el CONARTEL en lo que respecta a plazos de entrega y contenido de los informes técnicos.
2. Es conveniente escoger equipos de plataforma abierta para la operación y el mantenimiento. Esto permite tener un solo sistema de información (software) para la gestión de todos los equipos de la red, lo cual facilita la labor de los operadores y del personal de monitoreo.
3. Para la instalación de los equipos dentro de cada uno de los headends se recomienda cumplir con los requerimientos mínimos que estos necesitan como lo son la temperatura, el voltaje de entrada (tierra) y la humedad los mismos que garantizaran el buen funcionamiento de todos los equipos y la prolongación de vida de los mismos.

4. El equipo que se ha propuesto como equipo terminal (DCT700) se ha elegido por su alta funcionalidad para la prestación de servicios, pero si en un futuro se estima el cambio del mismo por la variación de servicios no sería un problema para los equipos que se encuentran en cada headend ya que los otros equipos que se describieron como el servidor de video, podría trabajar sin problemas con el nuevo equipo terminal.

5. Algo importante de recalcar es la importancia de la calidad de la señal que se necesita en el punto donde se conecta cada usuario para poder proveer una gran calidad de video, ya que si la calidad de la señal que viaja por la red HFC se degrada no se podrá cubrir los niveles mínimos necesarios y se tendrá problemas con el servicio de VOD, telefonía y demás servicios que son sensibles a variaciones y retardos.

A N E X O S

ANEXO A

FORMATO PARA AUTORIZACIÓN DE INCREMENTO DE CANALES O MODIFICACIÓN DEL HEAD END EN LOS SISTEMAS DE TELEVISIÓN POR CABLE

PARA ESTUDIOS DE INGENIERÍA DE INCREMENTO DE CANALES EN LOS SISTEMAS DE TELEVISIÓN POR CABLE
1. Descripción del profesional
2. Características técnicas de los equipos que se necesitan para el incremento de canales, tanto en recepción como en retransmisión
3. Características de la programación de los canales a incrementarse y antenas, de acuerdo al cuadro 1.
Cuadro N° 1:
PARA ESTUDIOS DE INGENIERÍA DE INCREMENTO DE CANALES EN LOS SISTEMAS DE TELEVISIÓN POR CABLE
1. DECLARACIÓN DEL PROFESIONAL: El profesional debe declarar que el Estudio de Ingeniería, planos de equipos e instalaciones y demás documentación técnica, los presenta bajo su responsabilidad; demostrará que su especialización se encuentra dentro del campo de la Electrónica y/o Telecomunicaciones; indicará claramente su nombre y número de afiliación al Colegio Profesional correspondientes; y

declarará que conoce las disposiciones contempladas en la Ley de Radiodifusión y Televisión, su Reglamento y las normas técnicas pertinentes.

2. CARACTERÍSTICAS TÉCNICAS DE LOS EQUIPOS QUE SE NECESITAN PARA EL INCREMENTO DE CANALES

3. CARACTERÍSTICAS DE LA PROGRAMACIÓN DE LOS CANALES A INCREMENTARSE Y ANTENAS

A N E X O B

DATASHEET DE EQUIPOS NECESARIOS

- **Servidor Multimedia SeaChange MDS 200**

http://www.schange.com/Downloads/ondemand/MediaServer_MDS201_QS.pdf

- **Equipo Terminal Motorola DCT 700**

<http://broadband.motorola.com/consumers/products/dct700/downloads/DCT700.pdf>

- **Motorola Digital Address Controler 6000**

http://broadband.motorola.com/catalog/product_documents/DAC6000.pdf

- **Motorola Network Controler NC 1500**

http://broadband.motorola.com/catalog/product_documents/NC1500_9.17.pdf

- **Motorola SmartStream Encryptor Modulator**

http://broadband.motorola.com/catalog/product_documents/SEMV8_538835-001-a.pdf

- **Motorola Out-of-Band Modulator (OM 1000)**







http://broadband.motorola.com/catalog/product_documents/OM1000_june02.pdf

- **Motorola Return Path Decoder (RPD 2000)**

http://broadband.motorola.com/catalog/product_documents/RPD2000_may02.pdf







ANEXO C

CUADRO DE CARACTERÍSTICAS TÉCNICAS DE LOS SERVIDORES VOD

VOD Vendors	Servidor VOD	Formato de Video	capacidad	IP OUTPUT	Costo
	TOUCAN 100	MPEG-2 TS, MPEG-4 CBR & VBR STREAM SD & HD Max Live input Stream:20 Mbps	up to 2,25 TB, 3 sata slots 1000h a 5 Mbps, 2500h a 2Mbps RAID 0,1,5	2 Giga Ethernet RJ-45 Protocolo: MPEG-TS sobre UDP Multicast o Unicast 60 stream simultáneos	\$ 10000
	Cisco CDE-200	MPEG vía RTP/RSTP, incluye MPEG1/2/4 H.264, MOV, y 3GPP	6 TB, 12 sata disk	2 Giga Ethernet 1.4 Gbps output por RTSP-UDP 1.2 Gbps output por RTSP-TCP	\$ 20000
	Vision 280	MPEG-2 TS, MPEG-4 (H.264), VC1 CBR & VBR, SDTV, HDTV	Up to 3TB per server Up to 3300 hours of content at 3.75 Mbps	Dual Gigabit Ethernet 480x3.75 Mbps MPEG-2 streams 900x2 Mbps MPEG-4 (H.264) streams	\$ 12000
	B-1	MPEG-2 and MPEG-4 (H.264)	8TB, 16 sata disk	80 Gbps of output per chassis	\$ 22000
	Orbit 2x	MPEG-1 MPEG-2 MPEG-4 AVC/H.264	128 GB - 3 TB, solid-state NAND flash • up to 2 500 HDTV streams @ 8.0 Mbps • up to 5 400 SDTV streams @ 3.7 Mbps	2 puertos Giga Ethernet	\$ 8000
	MDS 200	MPEG-1 MPEG-2 MPEG-4 AVC/H.264	5.5 Gbps de streaming 32 GB DRAM caché memory 6 TB almacenamiento 1500 a 3200 horas de video MPEG-2	8 puertos Giga Ethernet	\$ 8000

ANEXO D

TABLA DE CALIFICACIÓN DE LAS DIFERENTES MARCAS VOD

VOD Vendors	Servidor VOD	Formato de Video	capacidad	IP OUTPUT	Costo	Total
 anevia	TOUCAN 100	5	3	3	4	15
 arroyo	Cisco CDE-200	5	5	4	3	17
 bitbond	Vision 280	5	4	4	4	17
 BRADBUS	B-1	5	5	4	2	16
 edgeware	Orbit 2x	5	3	3	5	16
 SeaChange	MDS 200	5	5	4	5	19

Calificación a los diferentes servidores VOD de acuerdo a su formato de Video, Capacidad, IP Output y costo. Donde 5 es el más rentable y 1 el menos rentable.




ANEXO E

CUADRO COMPARATIVO ENTRE DIFERENTES MARCAS Y TIPOS DE SET TOP BOX

MARCA	SERIE	EQUIPO	SOPORTA MPEG-2	SOPORTA QAM 256 Downstream	SOPORTA QPSP Upstream	Sincronización Audio Video	Buffer de Almacenami ento	Interfaz Gráfica Amigable	Compatibilidad SEM	COSTO
Motorola	DCT 700		SI	SI	SI	SI	SI	SI	SI	300
	DCT 2000		SI	SI	SI	SI	SI	SI	SI	750
	DCT 1000		SI	SI	SI	NO	NO	NO	SI	250
Chengdu Newstar	SHIDA		SI	SI	SI	NO	SI	SI	NO	400
General Instrument	Cft 2200		SI	SI	SI	NO	NO	NO	SI	240

ANEXO F

CUADRO COMPARATIVO ENTRE DIFERENTES MARCAS Y TIPOS DE MODULADORES

MARCA	SERIE	IMAGEN	SOPORTA MPEG-2	INTERFACES	REQUERIMIENTOS SOPORTADOS	SOPORTA DOCSIS	COMPATIBLE CON DCT 700	SOPORTA IPTV SERVICIOS	MODULACION 64/256 QAM	COSTO
Motorola	SEM		SI	3 GIGABIT PORT 8 ASI PORT 4 RF OUTPUT	1524	SI	SI	SI	SI	4000
Arris	D5 Edge QAM		SI	2 GIGABIT PORT 8 ASI PORT 2 RF OUTPUT	1488	SI	SI	SI	SI	3500
Chengdu Newstar	QAM		SI	1 ASI PORT 1 10BASET PORT 2 RF OUTPUT	1200	NO	SI	NO	SI	2700

BIBLIOGRAFÍA

- Ernest Tunmann, “Hybrid Fiber Optic/Coaxial (HFC) Networks”, Flatiron Publishing , First Edition, 1995.
- Shih-Fu Chang, Dimitris Anastassiou, Alexandros Eleftheriadis, John V. Pavlik “Video on Demand Systems: Technology, Interoperability and Trials”,Kluwer Academic Publishers Group, 1997
- Michael E Lewis, “Design of an advanced development model optical disk-based redundant array of independent disks (RAID) high speed mass storage subsystem” Rome Laboratory, 1997
- Haykin Simon, “Sistemas de Comunicación”,Limusa Wileyl, 2002
- RSTP, “ Real Time Streaming Protocol (RTSP)” IETF, <http://tools.ietf.org/html/rfc2326>
- Motorola, “Broadband Communication”, <http://broadband.motorola.com>