Universidad de San Carlos de Guatemala
Facultad de Ingeniería
Escuela de Ingeniería Mecánica Eléctrica

CONVERGENCIA DEL PROTOCOLO DE INTERNET (IP) Y EL PROTOCOLO DE SEÑALIZACIÓN DE TELEFONÍA (SS7)

Luis Alejandro Lira Loarca
Asesorado por la Inga. Ingrid Rodríguez de Loukota

Guatemala, abril de 2014

UNIVERSIDAD DE SAN CARLOS DE GUATEMALA

FACULTAD DE INGENIERÍA

CONVERGENCIA DEL PROTOCOLO DE INTERNET (IP) Y EL PROTOCOLO DE SEÑALIZACION DE TELEFONíA (SS7)

TRABAJO DE GRADUACIÓN

PRESENTADO A LA JUNTA DIRECTIVA DE LA
FACULTAD DE INGENIERÍA
POR

LUIS ALEJANDRO LIRA LOARCA
ASESORADO POR LA INGA. INGRID RODRÍGUEZ DE LOUKOTA

AL CONFERÍRSELE EL TÍTULO DE

INGENIERO EN ELECTRÓNICA

GUATEMALA, ABRIL DE 2014

UNIVERSIDAD DE SAN CARLOS DE GUATEMALA

FACULTAD DE INGENIERÍA

NÓMINA DE JUNTA DIRECTIVA

DECANO	Ing. Murphy Olympo Paiz Recinos
VOCAL I	Ing. Alfredo Enrique Beber Aceituno
VOCAL II	Ing. Pedro Antonio Aguilar Polanco
VOCAL III	Inga. Elvia Miriam Ruballos Samayoa
VOCAL IV	Br. Walter Rafael Véliz Muñoz
VOCAL V	Br. Sergio Alejandro Donis Soto
SECRETARIO	Ing. Hugo Humberto Rivera Pérez

TRIBUNAL QUE PRACTICÓ EL EXAMEN GENERAL PRIVADO

DECANO	Ing. Murphy Olympo Paiz Recinos
EXAMINADOR	Ing. Luis Eduardo Durán Córdova
EXAMINADOR	Ing. Julio Rolando Barrios Archila
EXAMINADOR	Ing. José de León Escobar
SECRETARIA	Inga. Marcia Ivónne Véliz Vargas

HONORABLE TRIBUNAL EXAMINADOR

En cumplimiento con los preceptos que establece la ley de la Universidad de San Carlos de Guatemala, presento a su consideración mi trabajo de graduación titulado:

CONVERGENCIA DEL PROTOCOLO DE INTERNET (IP) Y EL PROTOCOLO DE SEÑALIZACION DE TELEFONÍA (SS7)

Tema que me fuera asignado por la Dirección de la Escuela de Ingeniería Electrónica, con fecha 18 de febrero de 2011.

Luis Alejandro Lira Loarca

Guatemala 21 de noviembre del 2012

Ingeniero
Carlos Eduardo Guzmán Salazar
Coordinador del Área de Electrónica
Escuela de Ingeniería Mecánica Eléctrica
Facultad de Ingeniería, USAC.

Estimado Ingeniero Guzmán.

Me permito dar aprobación al trabajo de graduación titulado: **"CONVERGENCIA DEL PROTOCOLO DE INTERNET (IP) Y EL PROTOCOLO DE SEÑALIZACIÓN DE TELEFONÍA (SS7)"**, del señor **Luis Alejandro Lira Loarca**, por considerar que cumple con los requisitos establecidos.

Por tanto, el autor de este trabajo de graduación y, yo, como su asesora, nos hacemos responsables por el contenido y conclusiones del mismo.

Sin otro particular, me es grato saludarle.

Atentamente,

Inga. Ingrid Rodríguez de Loukota
Colegiada 5,356
Asesora

Ingrid Rodríguez de Loukota
Ingeniera en Electrónica
Colegiado 5356

UNIVERSIDAD DE SAN CARLOS
DE GUATEMALA

FACULTAD DE INGENIERIA

Ref. EIME 65. 2013
Guatemala, 26 de ABRIL 2013.

Señor Director
Ing. Guillermo Antonio Puente Romero
Escuela de Ingenierìa Mecànica Elèctrica
Facultad de Ingenierìa, USAC.

Señor Director:

Me permito dar aprobación al trabajo de Graduación titulado: CONVERGENCIA DEL PROTOCOLO DE INTERNET (IP) Y EL PROTOCOLO DE SEÑALIZACIÓN DE TELEFONÍA (SS7), del estudiante Luis Alejandro Lira Loarca que cumple con los requisitos establecidos para tal fin.

Sin otro particular, aprovecho la oportunidad para saludarle.

Atentamente,
ID Y ENSEÑAD A TODOS

Ing. Carlos Eduardo Guzmán Salazar
Coordinador Área Electrónica

UNIVERSIDAD DE SAN CARLOS DE GUATEMALA
DIRECCION ESCUELA DE INGENIERIA MECANICA ELECTRICA
FACULTAD DE INGENIERIA

sro

UNIVERSIDAD DE SAN CARLOS
DE GUATEMALA

FACULTAD DE INGENIERIA

REF. EIME 65. 2013.

El Director de la Escuela de Ingeniería Mecánica Eléctrica, después de conocer el dictamen del Asesor, con el Visto Bueno del Coordinador de Área, al trabajo de Graduación del estudiante; LUIS ALEJANDRO LIRA LOARCA titulado: CONVERGENCIA DEL PROTOCOLO DE INTERNET (IP) Y EL PROTOCOLO DE SEÑALIZACIÓN DE TELEFONÍA (SS7), procede a la autorización del mismo.

Ing. Guillermo Antonio Puente Romero

UNIVERSIDAD DE SAN CARLOS DE GUATEMALA
DIRECCION ESCUELA DE INGENIERIA MECANICA ELECTRICA
FACULTAD DE INGENIERIA

GUATEMALA, 25 DE SEPTIEMBRE 2,013.

Universidad de San Carlos
De Guatemala

Facultad de Ingeniería
Decanato

Ref. DTG.D.186-2014

El Decano de la Facultad de Ingeniería de la Universidad de San Carlos de Guatemala, luego de conocer la aprobación por parte del Director de la Escuela de Ingeniería Mecánica Eléctrica, al trabajo de graduación titulado: **CONVERGENCIA DEL PROTOCOLO DE INTERNET (IP) Y EL PROTOCOLO DE SEÑALIZACIÓN DE TELEFONÍA (SS7),** presentado por el estudiante universitario: **Luis Alejandro Lira Loarca,** autoriza la impresión del mismo.

IMPRÍMASE.

Ing. Murphy Olympo Paiz Recinos
Decano

UNIVERSIDAD DE SAN CARLOS DE GUATEMALA
DECANO
FACULTAD DE INGENIERIA

Guatemala, abril de 2014

/cc

ACTO QUE DEDICO A:

Dios	Por ser la fuente de sabiduría que me ha permitido alcanzar las metas propuestas en la vida.
Mis padres	Luis Alberto Lira y Eugenia Loarca de Lira. Cuyo amor y apoyo incondicional me han dado la confianza y la oportunidad de salir adelante.
Mi tía	Celia Loarca. Por su amor y cuidado a mi vida.
Mis hermanos	Pablo y Andrea Lira. Por estar presentes en cada una de las etapas de mi vida brindándome su apoyo.
Mis amigos	Por su constante motivación y compañerismo, cada uno de ustedes ha sido importante en todo momento.

AGRADECIMIENTOS A:

Mis padres	Por su constante ayuda, dedicación y perseverancia para llevar a la culminación mis estudios. Padres abnegados a quienes amo, admiro y respeto.
Mi familia	Por ser un ejemplo de vida, por sus consejos y cuidado.
Mi asesora	Inga. Ingrid de Loukota. Por la dedicación, orientación y asesoría académica que me brindó durante mis estudios y la elaboración del presente trabajo.

ÍNDICE GENERAL

ÍNDICE DE ILUSTRACIONES

FIGURAS

TABLAS

LISTA DE SÍMBOLOS

Símbolo	Significado
%	Porcentaje

GLOSARIO

1G Telefonía móvil de primera generación. Sistema de comunicación móvil, capaz de transmitir únicamente voz análoga.

2G Telefonía móvil de segunda generación. Sistema de comunicación móvil, capaz de transmitir voz en forma digital, llenando necesidades de cifrado y servicios agregados.

3G Telefonía móvil de tercera generación. Sistema de comunicación móvil, con capacidades agregadas como acceso a internet desde el celular, video llamada, correo electrónico.

3GPP 3rd Generation Partnership Project, colaboración de grupos de asociaciones de telecomunicaciones, conocidos como miembros organizativos cuyo objetivo es asentar las especificaciones de un sistema global de comunicaciones de tercera generación.

AH Authentication Header. Campo de la trama de IP que provee la autenticación de los datos de origen.

Alignment Error Rate	Función de monitoreo en MTP2 utilizado para asegurar la sincronía de las unidades transmitidas.
AMPS	Advance Mobile Phone System. Sistema de telefonía móvil análogo desarrollado por laboratorios Bell introducido en América en 1978.
AMR	Adaptive Muti Rate. Esquema de compresión de audio optimizada para la codificación de la voz en tecnología 2G.
ANSI	American National Standards Institute. Organización privada sin fines de lucro que supervisa el desarrollo bajo estándares consensuados para productos, servicios, sistemas y personal en los Estados Unidos de América.
AOL	America On Line. Compañía americana dedicada a los servicios web con una amplia gama de marcas y ofertas para el público global.
ARPA	Advance Research Project Agency. Agencia de investigación del departamento de defensa de los Estados Unidos de América.
ATM	Asynchronous Transfer Mode. Técnica de conmutación estándar, diseñada para unificar las redes de telefonía y computación.

BISDN	Broadband ISDN. Norma que fue concebida para su uso con aplicaciones avanzadas. SDH / SONET y ATM nacieron fuera de la norma ISDN de banda ancha.
Bit	Es una contracción del término *Binary digit* (Dígito Binario). Es la unidad de información más pequeña que una computadora puede procesar.
BSD	Bussiness Service Database. Permite a los suscriptores almacenar instrucciones para el proceso de sus llamadas.
Byte	Un conjunto de *bits* (unos y ceros) de una longitud específica representando un valor en el esquema de codificación de una computadora.
Call forward	Característica de algunas redes de telefonía que permiten a una llamada entrante ser redireccionada a un tercero.
Called Party Address	Se define de este modo a la persona o dispositivo que recibe una llamada dentro de una red de telefonía.
Calling Party Address	Se define de este modo a la persona o dispositivo que realiza (inicia) una llamada dentro de una red de telefonía.

CCITT	Comité Consultatif International Téléphonique et Télégraphique. Comité creado en 1956 (luego renombrado a ITU-T) encargado de la estandarización en telecomunicaciones.
CCS	Common Channel Signaling. Red de señalización independiente de los circuitos de voz creada durante los años 60, durante el inicio de la digitalización de las redes de telefonía.
CDMA	Code Division Multiple Access. Método de acceso a canal de comunicación mediante la división de código.
CERN	Conseil Européenne pour la Recherche Nucléaire. Organización internacional cuyo propósito es administrar el laboratorio de partículas físicas.
Chunks	Unidad de información en los paquetes de SCTP que consiste en un encabezado y un contenido.
Codificación	Proceso por el cual la información se codifica en símbolos digitales en el transmisor y se decodifican en el receptor para maximizar la precisión de la información presentada al usuario.
Convergencia	El punto en el cual todos los dispositivos de red interconectados comparten información en común de las rutas óptimas del tráfico. Entre más rápida sea la

convergencia, más rápido se recobra de una falla de red.

COPS Common Open Policy Service. Protocolo base para comunicar la información de las políticas entre los servidores de políticas y los *routers* que se encuentren en el marco de RAP.

Corriente DC La corriente directa (CD) o corriente continua (CC) es aquella cuyas cargas eléctricas o electrones fluyen siempre en el mismo sentido en un circuito eléctrico cerrado, moviéndose del polo negativo hacia el polo positivo de una fuente de fuerza electromotriz (FEM).

CPU Central Process Unit. El componente de una computadora que ejecuta la lógica computacional y las funciones de decisión.

CRC Cyclic Redundancy Code. Es un código de detección de error comúnmente usado en redes digitales y dispositivos de almacenamiento de información para detectar errores en la información original.

Cuantización Conversión de una señal análoga en un formato digital a través de muestreo y codificación en valores discretos de la amplitud de la señal.

Datagrama Mensaje de longitud máxima fija enviado sin verificación alguna que asegure su exactitud, envío o

secuencia correcta con respecto a los mensajes relacionados, que lleva la dirección completa de destino utilizada para el enrutamiento.

DCME Digital circuit multiplication equipment. Es un equipo de compresión capaz de manejar voz, datos y señales de fax. Este equipo se utiliza como un medio para aumentar la capacidad de los sistemas de transmisión digital.

Decibel Unidad de medida de la potencia del sonido o de la fuerza de una señal, generalmente expresado como la relación entre la señal transmitida y una fuente de señal de referencia. El decibel fue inventado para medir el nivel de ganancia o pérdida en los sistemas de transmisión de telefonía.

DECnet Grupo de productos de comunicaciones (incluyendo un conjunto de protocolos) desarrollado y soportado por Digital Equipment Corporation.

DQDB Distributed Queue Dual Bus. Técnica de acceso de la red de área metropolitana (MAN) definido por el estándar IEEE 802.6 para el servicio de datos conmutados multimegabit (SMDS).

Drivers Un *driver (*software*)* proporciona instrucciones para cambiar el formato o la interpretación de los comandos de software para la transferencia desde y

	hacia dispositivos periféricos y la unidad de procesamiento central (CPU).
DSI	Técnica de compresión para concentrar más conversaciones de voz en una línea. DSI digitaliza la voz de manera que no se transmiten *bits* cuando no se está hablando. Tan pronto como el habla comienza, los *bits* fluyen de nuevo.
E.164	Estándar de la ITU-T en la que se normaliza que el número máximo de dígitos para números internacionales, servicios globales, redes y grupos de aplicaciones debe ser de 15 dígitos (excluyendo el prefijo internacional).
Emulador	Dispositivo o programa de ordenador que puede actuar como si se trata de otro dispositivo o programa.
Encriptación	El cifrado es el uso de un algoritmo para ocultar el significado de una pieza de información por lo que no puede ser leído y entendido.
End-to-end	Conexión entre el sistema origen y el sistema destino.
Erlang	Medida de tráfico telefónico. Un *Erlang* es el uso de un circuito de comunicación durante una hora.

Ethernet Red de área local (LAN), norma oficialmente conocida como IEEE 802.3 y que opera a través del cable par trenzado y por cable coaxial a velocidades a partir de 10 Mbps.

FDDI Fiber Distributed Data Interface. FDDI es un estándar ANSI (X3T12) para una red de fibra óptica de 100 Mbps utilizando una topología de anillo doble con una técnica de paso de *token* para el control de acceso al medio.

Fibra óptica Tecnología en la que la luz se utiliza para transportar información de un punto a otro. Hilo muy fino de material transparente, vidrio o materiales plásticos, por el que se envían pulsos de luz que representan los datos a transmitir. El haz de luz queda completamente confinado y se propaga por el interior de la fibra con un ángulo de reflexión por encima del ángulo límite de reflexión total, en función de la ley de Snell.

FM Frecuencia Modulada. Técnica de modulación en la cual la frecuencia de la señal portadora es desplazada por una cantidad proporcional a la de la señal moduladora que contiene la información a transmitir.

Fotón El fotón es una partícula de luz. Durante cientos de años se pensó en la luz como una onda. En 1905, Einstein descubrió que bajo ciertas circunstancias, la

energía de una onda de luz sólo se producía en cantidades específicas o cuantos. Estos cuantos son llamados fotones.

Frame Un *frame* es un paquete. Es un término genérico específico para una serie de protocolos de comunicaciones de datos. Un *frame* de datos es una unidad lógica de datos, que normalmente es un fragmento de un conjunto mucho mayor de datos.

Frame relay Estándar de acceso definido por el UIT-T en la Recomendación I.122. Servicios *Frame Relay* son aquellos proporcionados por las compañías de telecomunicaciones empleando una forma de conmutación de paquetes similar a una versión simplificada de redes X.25. Los paquetes son en forma de *frames*, que son de longitud variable.

Full duplex Proceso de funcionamiento de un circuito de modo que los extremos pueden transmitir y recibir simultáneamente.

Gateway Puerta de enlace. Un *gateway* es un dispositivo repetidor electrónico que intercepta y dirige las señales eléctricas de una red a otra. En redes de datos, usualmente ejecutan procesos de conversión de código y protocolos. De acuerdo al modelo OSI, un *gateway* es un dispositivo que provee servicio a las 7 capas del modelo.

GPS Global Positioning System. Constelación de 24 satélites en órbita que permiten determinar con precisión la posición en cualquier lugar de la Tierra con una precisión de un metro.

Handshake La serie de señales entre un ordenador y otros dispositivos periféricos que establecen los parámetros necesarios para transferir datos.

HDLC High level Data Link Control. Protocolo de capa de enlace (capa 2 del modelo de referencia OSI) estándar de punto a punto y comunicaciones punto a multipunto.

Header La porción de un mensaje que contiene la información que guía el mensaje al destino correcto. Esta información puede contener las direcciones del remitente y del receptor, el nivel de prioridad, las instrucciones de ruta y los pulsos de sincronización.

Hops Viaje individual que los paquetes recorren muchas veces, de *router* a *router*, en su camino hacia sus destinos.

Host Dispositivo inteligente conectado a una red.

HSSI High Speed Serial Interface. Interfaz de comunicación serial de datos optimizados para altas

velocidades de hasta 52 Mbps. Utilizada para conectar un *switch* ATM a un enlace T-3.

Hub En redes de área local es el núcleo de una topología de red en estrella.

Hub and Spoke Topología de red semejante a una estrella en la que existe un nodo central por el cual fluye todo el tráfico de los nodos conectados a él.

IEEE 802.3 Protocolo de red de área local comúnmente conocido como *Ethernet. Ethernet* tiene una velocidad de 10 Mbps o 100 Mbps.

IEEE 802.5 Un estándar de especificación de capa física de una LAN con un método de acceso de paso de *token* en una topología de anillo.

IETF Internet Engineering Task Force. Formado en 1986, cuando la Internet estaba evolucionando de un experimento del Departamento de Defensa en una red académica. El IETF es uno de los dos órganos técnicos de trabajo de la junta de actividades de internet.

IGRP Interior Gateway Routing Protocol. Un protocolo de enrutamiento de vector de distancia desarrollado por Cisco Systems para su uso en grandes redes

	heterogéneas. Aprende las mejores rutas a través de una LAN o internet.
Interface	Interfase. Punto de demarcación física entre dos dispositivos en los que las señales eléctricas, conectores y el *handshake* se definen. Los procedimientos, códigos y protocolos que permiten a dos entidades que interactúan en un intercambio significativo de información.
Interleave	Intercalado. Técnica de comunicación de datos que se utiliza en conjunción con códigos correctores de errores para reducir el número de errores en ráfagas sin ser detectadas.
Intranet	Una red basada en protocolos TCP / IP (internet) que pertenece a una organización y es accesible sólo por miembros de la organización, los empleados u otras personas con autorización.
IP Tunnel	Técnica para portar cualquier protocolo por medio de un paquete TCP/IP.
IP v4	Internet Protocol Version 4. La versión actual del protocolo de internet.
IP v6	Internet Protocol Version 6. El nuevo protocolo de internet diseñado para sustituir y mejorar el presente protocolo que se llama IP v4.

IPX Internet Packet Exchange. Protocolo de comunicaciones Novell NetWare que se utiliza para mover datos entre el servidor y / o programas de la estación de trabajo que se ejecutan en diferentes nodos de la red.

ISDN Integrated Services Digital Network. Es un conjunto de normas internacionales establecidas por la ITU-T (International Telecommunications Union-Telecommunications) para un circuito de conmutación de red digital que permite el acceso a cualquier tipo de servicio (por ejemplo, voz, datos y vídeo).

IS-IS Intermediate System to Intermediate System. El protocolo IS-IS utiliza un algoritmo de estado de enlace para proporcionar servicios de enrutamiento de TCP/IP y OSI. Se determina el mejor camino para paquetes TCP/IP y OSI a través de la red. Mantiene a los *routers* informados del estado de la red y los sistemas disponibles.

ITU International Telecommunication Union. Organización con sede en Ginebra, Suiza. Es el organismo más importante de telecomunicaciones que fija las normas en el mundo. Está formado actualmente por tres grandes sectores que se establecieron en 1992: sector de radiocomunicaciones (ITU-R), el sector de desarrollo

	de las Telecomunicaciones (ITU-D) y el sector de normalización de las telecomunicaciones (ITU-T).
ITU-T Y.101	Norma de la ITU-T donde se formaliza los términos y definiciones de Infraestructura de Información global.
LAN	Local Area Network. Red de comunicaciones que conecta las computadoras personales, estaciones de trabajo, impresoras, servidores de archivos y otros dispositivos dentro de un edificio o un campus.
LAP-B	Link Access Procedure Balanced. El más común de los enlaces de datos del protocolo de control utilizado para la interfaz X.25.
LAP-D	Link Access Procedure-D. También llamado protocolo de acceso de enlace para el canal D. protocolo a nivel de enlace diseñado para ISDN. Diferente del LAPB (LAP-Balanced) en la secuencia de tramas.
Layer	En el modelo de referencia OSI, una colección de funciones de procesamiento relacionado con la red que forman parte de un nivel de una jerarquía de funciones.
Límite de Shannon	El límite teórico de la información a través de un canal.

Link	El foro ATM define un enlace (*link*) como una entidad que define una relación topológica (incluyendo la capacidad de transporte disponible) entre dos nodos de diferentes subredes.
Meta señalización	El canal de meta-señal es un canal de gestión que se utiliza para establecer señalización de circuitos virtuales entre la red y cada uno de los dispositivos individuales en una interfaz multipunto.
Modelo OSI	Open Systems Interconnection. El modelo de referencia OSI es el único marco internacionalmente aceptado de estándares para la comunicación entre los diferentes sistemas de varios vendedores.
Modulación	El proceso de convertir las señales de voz o datos para su transmisión por un medio físico.
MTU	Maximum Transmission Unit. La unidad más grande posible de datos que se pueden enviar en un medio físico determinado. Ejemplo: el MTU de e*thernet* es 1 500 *bytes*.
Multicast	Envío de mensajes a un grupo seleccionado de estaciones de trabajo en una LAN, WAN o internet.
Multimedia	Es la combinación de múltiples formas de medios de comunicación en la transmisión de información. Permite a las personas comunicarse con medios de

	comunicación integrados: audio, video, texto, gráficos, fax y telefonía.
Multiplexación	Intercalar o transmitir simultáneamente dos o más mensajes en un solo canal.
NIC	Network Interface Card. Una tarjeta de red funciona con el software de red y sistema operativo del ordenador para transmitir y recibir mensajes en la red.
NNI	Network Node Interface. Interfaz entre dos segmentos de la red pública de los equipos (a diferencia de la UNI, que representa la interfaz de usuario de red).
NSAP	Network Service Access Point. El punto en el que el servicio de red del modelo OSI está a disposición de la entidad de transporte.
OSPF	Open Shortest Path First. Utiliza el estado de enlace y protocolos de puerta interior para crear un mapa de la red en cada *router* para luego utilizar el algoritmo Dijkstra del camino más corto y determinar la ruta óptima entre dispositivos de red.
Overhead	En las comunicaciones es toda la información como el control, ruteo y comprobación de errores que se suman a la transmisión de datos de usuario. Incluye

información que lleva el estado de la red o las instrucciones de funcionamiento, la información de enrutamiento de red, así como retransmisiones de mensajes de datos por el usuario que se han recibido con error.

PABX Private Automatic Branch Exchange. Una central telefónica ubicada en las instalaciones de un cliente que establece circuitos principalmente de grado de voz entre los usuarios individuales (extensiones) y la red telefónica conmutada sin necesidad de operadora.

Paquetes Término genérico para un conjunto de datos, por lo general en forma binaria, organizados de una manera específica para su transmisión.

Paridad Un proceso para detectar si los bits han sido alterados durante su transmisión.

Payload Carga de información. La parte que representa la información útil para el usuario.

PCM Pulse Code Modulation. Método de codificación de una señal analógica en una señal digital.

Peer to peer Comunicación entre dos entidades que operan dentro de la misma capa de protocolo de un sistema.

Pila de protocolo	Es básicamente una colección de módulos de software que juntos se combinan para producir el software que permite que el protocolo funcione, es decir, para permitir las comunicaciones entre dispositivos informáticos diferentes.
Plesiócrono	Redes plesiócronas involucran múltiples circuitos digitales síncronos funcionando a velocidades de reloj diferentes.
PM	Phase Modulation. Modulación de fase. Técnica de modulación que cambia la fase de la señal portadora de acuerdo a la señal de información que se le adiciona.
Polarización de antena	La polarización de una antena indica la dirección del campo eléctrico en una onda radiada. Es un factor que determina la cantidad de energía que recibe una antena en la recepción de la señal.
Protocolo	Es un conjunto de normas que regulan el formato de los mensajes que se intercambian entre los ordenadores y las personas.
Protocolo internet	Parte de la familia de protocolos TCP/IP que describen el software que controla la dirección de Internet de los nodos, las rutas de los mensajes salientes y reconoce los mensajes entrantes. Se

	utiliza en puertas de enlace para conectar las redes a nivel de capa de red y superiores.
QoS	Quality of Service. Calidad de servicio es una medida del servicio que las telecomunicaciones (voz, datos y/o vídeo) prestan a un abonado.
Ráfaga	En la comunicación de datos, una secuencia de señales, ruido, o interfaz tomada como una unidad de acuerdo con algún criterio o medida específica.
Receptor	Cualquier componente de un dispositivo de telecomunicaciones que decodifica una señal codificada en su forma deseada.
Retardo de Propagación	El tiempo que tarda una señal en viajar desde el emisor hasta el receptor a través de un circuito. El retardo de propagación es un factor de la velocidad finita en la que las señales electromagnéticas pueden viajar a través de un medio de transmisión.
RFC	Request For Comment. El desarrollo de estándares TCP/IP, procedimientos y especificaciones se realiza a través de este mecanismo. Los RFC son los documentos que progresan a través de varias etapas de desarrollo, bajo el control de la IETF, hasta que son finalizados o descartados.

RFC 1483	RFC Multiprotocol Encapsulation over ATM Adaptation Layer 5. En este RFC se describen dos métodos de encapsulación para transportar tráfico de interconexión sobre ATM AAL5.
RFC 2581	RFC TCP Congestion Control. En este RFC se definen los algoritmos de control de congestión de TCP.
RIP	Routing Information Protocol. Se basa en algoritmos de vector de distancia que miden el camino más corto entre dos puntos de una red, en base a las direcciones de los dispositivos de origen y destino.
Rotación Faraday	El efecto Faraday causa una rotación en la polarización de una señal con respecto a la orientación del plano de polarización.
Router	Un dispositivo que conecta dos segmentos de LAN, que utilizan arquitecturas similares o diferentes en la capa de red OSI, la capa 3. El *router* determina la ruta más eficaz para transmitir datos por medio del Internet. Los paquetes que contienen una dirección de red diferente a la dirección del PC de origen se remiten a una red contigua.
Routing	Enrutamiento. El proceso de selección de la trayectoria del circuito para un mensaje.

SAPI Service Access Point Identifier. El SAPI identifica un punto lógico en el que los servicios de capa de enlace son provistos por una entidad de capa de enlace de datos a una entidad de capa 3 (capa de red).

Satélite Plataforma receptora de microondas, que funciona como repetidor y/o regenerador de señales que circunda en órbita sobre la tierra.

Señal portadora Una forma de onda continua (normalmente eléctrica), cuyas propiedades son susceptibles de ser modulada con una segunda señal portadora de información.

Símbolo Estado eléctrico reconocible el cual está asociado con un elemento de señal que es una señal eléctrica dentro de un período de tiempo definido.

Síncrono Sistema que utiliza una tasa de reloj para la transmisión de datos.

SMDS Switched Multimegabit Data Service. Una conexión de alta velocidad de datos de servicios de transmisión destinada a la aplicación en una red de área metropolitana (MAN).

Socket Una abstracción del sistema operativo que proporciona la capacidad a los programas de aplicación para acceder de forma automática los

	protocolos de comunicación. Desarrollado como parte de los primeros trabajos sobre TCP/IP.
SONET	Synchronous Optical NETwork. Una familia de tasas de transmisión por fibra óptica de 51,84 Mbps a 39,812 Gbps creada para proporcionar la flexibilidad necesaria para el transporte de muchas señales digitales con capacidades diferentes.
Spread spectrum	También llamado salto de frecuencia, es una técnica de modulación utilizado en los sistemas inalámbricos. Los datos a ser transmitidos son paquetizados y esparcidos sobre una amplia gama de ancho de banda que la demanda por el contenido del flujo de la información original.
SS7	Signaling System, #7. Conjunto de protocolos de señalización telefónica, utilizados en la mayoría de redes de telefonía.
Stack	Lista de instrucciones internas que se ejecutan en un sistema.
STM-1	Synchronous Transport Module 1. Estándar SDH para la transmisión sobre fibra óptica OC-3 a 155,52 Mbps.
STS	Synchronous Transport Signal. El equivalente eléctrico de SONET OC-nivel. La señal eléctrica se

convierte en óptica antes de la presentación al medio de fibra óptica.

STS-3C Synchronous Transport Signal level 2. Implementación de la capa física ATM que soporta 155 Mbps.

Subred Una subred es una parte de la red, que puede ser una red físicamente independiente, que comparte una dirección de red con otras partes de la red y se distingue por un número de subred.

Switch Dispositivo mecánico, eléctrico o electrónico que abre o cierra los circuitos, completa o interrumpe una trayectoria eléctrica, o selecciona vías o circuitos. Los *switches* trabajan en las capas 1 (física) y 2 (enlace de datos) del modelo de referencia OSI, haciendo énfasis en la capa 2. Un *switch* verifica los datos entrantes (datos o voz) para determinar la dirección de destino. Con base en esa dirección, una conexión se establece a través de la matriz de conmutación entre los puertos físicos de entrada y salida del dispositivo.

Switching Establecimiento de una ruta de transmisión de una entrada en particular a una salida en específico de un grupo de entradas y salidas.

Tandem	En el contexto de las telecomunicaciones, el término se refiere a los *switches*, circuitos, u otros elementos de red, que permiten que otros elementos de red trabajen en conjunto.
Tasa de errores de bit	Porcentaje de bits recibidos con error en comparación con el número total de bits recibidos.
TCP	Transmission Control Protocol. TCP es un protocolo de capa de transporte, orientado a la conexión. Proporciona una entrega confiable, ordenada, y no duplicada de bytes enviados a un usuario remoto o local.
TDM	Time Division Multiplex. Una técnica de multiplexación para la transmisión de una serie de señales separadas de datos, voz y/o video simultáneamente a través de un medio de comunicación intercalando una pieza de cada señal una después de la otra.
Telnet	Un programa que le permite conectarse a otros ordenadores en internet.
Terminal	Un dispositivo de entrada/salida para comunicarse con computadoras.
Throughput	La cantidad real de información útil y no redundante que es transmitida o procesada.

Time to live Mecanismo utilizado por el protocolo IP para controlar la vigencia de un paquete y evitar que éste se mantenga en la red por siempre.

Timeout Es la cantidad de tiempo que el hardware o el software espera por un evento antes de tomar acciones correctivas para terminar la conexión o sesión.

Token En las redes, una combinación única de bits utilizados para otorgar el acceso a una computadora a transmitir a una en una red de área local. También proporciona información importante para el enrutamiento de mensajes a través de la red, tales como direcciones de origen y destino, información de control de acceso, información de control de ruta.

Topología Se refiere a cómo los dispositivos se cablean. La topología lógica se refiere a cómo los nodos en realidad interactúan.

Trail Como término ATM es una entidad que transfiere la información proporcionada por la capa de cliente entre los puntos de acceso en una red de capa de servidor.

Transponders Un *transponder* es un equipo de radio a bordo de un satélite de comunicaciones. Recibe una señal, la

amplifica, cambia su frecuencia y la reenvía de regreso a la tierra.

UBR Unspecified bit rate. Una clase de servicio ATM. UBR no ofrece ninguna garantía de servicio, por lo que se tendría que utilizar para datos de texto, transferencia de imágenes, mensajería y distribución de información no crítica, en el que no tiene que tener un tiempo de respuesta o garantía de servicio.

UDP User Datagram Protocol. Provee el intercambio de datagramas sin acuses de recibo ni entrega garantizada.

UMTS Universal Mobile Telecommunications System. Miembro europeo de la IMT-2000 (Telecomunicaciones móviles internacionales para el año 2000), familia 3G (tercera generación) de estándares inalámbricos. UMTS está diseñado para soportar velocidades de transferencia de datos de 144 Kbps para el tráfico vehicular, 384 Kbps para el tráfico peatonal, y hasta 2 Mbps en apoyo de los servicios dentro de un edificio.

UNI User Network Interface. Especificaciones para los procedimientos y protocolos entre el equipo del usuario o bien una red ATM. La UNI es el punto físico, eléctrico y funcional de demarcación entre el usuario y el proveedor de servicios de la red pública.

Uplink	En los satélites es el enlace de la estación de tierra hasta el satélite.
V.35	Estándar para la interfaz troncal entre un dispositivo de acceso a la red y una red de paquetes que define la señalización de las velocidades de datos superiores a 19,2 Kbps.
VoIP	Voice over IP. Una tecnología para el transporte de voz digital integrada, video, y datos sobre redes IP.
VPN	Virutal Private Network. Red de paquetes de datos que ofrece servicios de red privada.
VSAT	Very Small Aperture Terminal. Pequeña antena satelital, por lo general 1,5 a 3,0 metros de diámetro, que se utiliza para aplicaciones de comunicaciones de datos punto a multipunto.
WAN	Wide Area Network. Es una red pública que cubre un área metropolitana, que puede extenderse más allá de los límites de la ciudad.
WWW	World Wide Web. Es el universo de información accesible disponible en muchos equipos extendidos por el mundo y unidos a la red de computadoras gigantescas que se llama Internet. La Web tiene un cuerpo de software, un conjunto de protocolos y un

conjunto de convenciones definidas para llegar a la información en la *web*.

X.25 Un estándar que define la interfaz entre el DTE y DCE para equipos que operan en el modo de paquetes en redes públicas de datos. También se define un protocolo de control de enlace.

RESUMEN

El presente trabajo de graduación tiene como objetivo reunir en un documento la información necesaria para dar a conocer el mundo convergente de redes de telefonía de próxima generación. Día a día la rama de las comunicaciones celulares se adapta y acopla a la tendencia global de convergencia, cambiando sus procedimientos, controles y lógica de negocios para realizar la transición a infraestructura sobre IP; por lo que se hace necesario realizar una descripción y análisis de las redes y arquitecturas de la siguiente generación y presentar una discusión de las características, funciones y propiedades de los protocolos que se fusionan en esta nueva generación de red convergente.

Al revisar todo este proceso se podrá ver más claramente los beneficios que el servicio de telefonía y de internet brinda a una sociedad que demanda tecnología. Entre los beneficios inmediatos que se encuentra son, la reducción de costos en la infraestructura y un manejo mejorado de las mismas y entre los beneficios a largo plazo se obtiene la habilidad de implementar nuevos servicios en corto tiempo.

En el capítulo 1 se presenta todo el trasfondo histórico, desde los inicios de la señalización telefónica, pasando después por la digitalización de sus redes, obteniendo posteriormente un método de canal-común denominado SS7, hasta llegar a la creación de un protocolo de red robusto y potente denominado protocolo de internet (IP), el que desde sus inicios fue diseñado pensando en la interconectividad de redes.

En el capítulo 2 se muestra la convergencia que se está dando entre SS7 e internet, que son dos redes independientes, las cuales proveen servicios distintos y llevan a cabo diferentes tareas; pero debido a la demanda mundial entre los servicios de telefonía y el acceso de internet se está dando una fusión entre los distintos servicios que ofrecen, a fin de dar respuesta a las necesidades actuales.

En el capítulo 3 se expone el diseño de "S*imple Control Transport Protocol*" (SCTP), que va desde control de protección, hasta enrutamiento IP. SCTP fue creado para ser un protocolo de transporte para la señalización telefónica realizando esta función como una capa en completa unión con el protocolo de internet (IP) dando el paso para la convergencia de ambas redes.

En el capítulo 4 se presenta *Signaling Transport* (SIGTRAN), que es un nuevo ser de protocolos para el transporte de mensajes de señalización SS7 sobre IP. Estos protocolos son el primer paso para la fusión de redes.

OBJETIVOS

General

Proporcionar una descripción y análisis de los protocolos SIGTRAN y su papel en la siguiente generación de redes celulares.

Específicos

1. Analizar el entorno histórico que gobierna y genera la necesidad de globalización y adaptación de las redes hacia la unificación de estándares.

2. Describir y analizar los protocolos actualmente involucrados en las comunicaciones y su fiable convergencia hacia una red incluyente de todos los servicios.

3. Describir y analizar los protocolos emergentes en SIGTRAN para transporte de señalización y su acople a las tecnologías actuales.

INTRODUCCIÓN

Con la creciente popularidad de las redes basadas en IP se hizo obvio el objetivo de integrar las redes existentes de señalización con las redes IP. Las telecomunicaciones modernas como las redes 3G móviles están desarrollando y encontrando su camino hacia una arquitectura todo-IP. Este tipo de arquitectura significa que todos los nodos que conforman la red se comunican unos con otros utilizando tecnología de IP.

Tradicionalmente las redes telefónicas han empleado el protocolo SS7 para la señalización entre *switches* de telefonía. La utilización de protocolo IP permita el acople de ambas tecnologías permitiendo que los nodos de telefonía utilicen señalización SS7, pero ésta será transportada sobre una red IP. Por ejemplo, nodos como HLR, VLR/MSC se comunicarán unos con otros usando SIGTRAN.

SIGTRAN es un grupo de protocolos estandarizados por IETF (Internet engineering task force), todos estos protocolos son independientes unos de otros. Cada uno de ellos ha sido desarrollado para emular las funciones correspondientes a los protocolos SS7 sobre una red IP. Por ejemplo, M3UA es un protocolo SIGTRAN realizando las funciones en IP, similares a las que el protocolo MTP-3 realiza en una red SS7.

Por las razones anteriores se considera necesario realizar la tesis en este tema, debido a la importancia primordial que en la actualidad tiene no solo la telefonía celular, sino los avances que constantemente deben tomar en cuenta para mantenerse con lo más reciente en tecnología.

1. TRASFONDO HISTÓRICO

El mundo convergente de redes de próxima generación (NGN, por sus siglas en inglés) hace su aparición con promesas de obtener voz, video y data sobre una misma red de alto ancho de banda. Esta transición de las tradicionales redes basadas en circuitos, hacia redes basadas en paquetes ha estado presente ya por varios años.

1.1. Señalización en telefonía

Alexander Graham Bell patentó el teléfono en 1876 e inmediatamente hubo una gran demanda por el nuevo invento. Inicialmente el uso del teléfono era muy simple y no existía ninguna empresa telefónica, sino que los teléfonos eran vendidos en pares y el comprador era el encargado de establecer la conexión al conectar un solo cable entre ellos, (la superficie del planeta funcionaba como tierra por lo que solamente era necesario utilizar un cable).

Los teléfonos no tenían un tono de aviso y la manera de establecer una llamada se hacía gritando al micrófono y esperando que la otra persona estuviese lo suficientemente cerca de su teléfono para escuchar al otro "llamando". Esto le dio al dueño del teléfono la posibilidad de hablar con otro cliente. Uno debía tener la misma cantidad de pares de teléfonos como la cantidad de diferentes personas con las que se quería comunicar. La figura 1 muestra esta situación, si nueve personas deseaban estar conectadas entre ellas.

Figura 1. **Evolución en redes de telefonía**

Fuente: elaboración propia, con Microsoft Office Visio 2007.

Un año después de esto las ciudades se encontraban cubiertas de cables pasando sobre las casas y los árboles, haciéndose obvio que este modelo de conexión no funcionaría. Tomando ventaja de esto, Bell creo la compañía telefónica (Bell Telephone Company) y abrió la primera oficina de conmutación en 1878.

La compañía establecía un cable hacia cada cliente u oficina. Cuando ellos deseaban utilizar el teléfono debían primero levantar el receptor, permitiendo que corriente DC fluyera a través del teléfono y de regreso a través del retorno del circuito, encendiéndose una lámpara en el tablero de conmutación del operador, quien luego lo conectaba con la persona que deseaba contactar utilizando un puente (*jumper*).

De esta manera, un cliente podía comunicarse con todos los demás usuarios conectados a la misma oficina, solamente teniendo un aparato telefónico (ahora equipados con tono de aviso) y un solo cable (ahora balanceados, aislados y cables en pares). Este modelo se ilustra en la figura 1.

Al tiempo que la popularidad del teléfono se incrementó, las personas deseaban realizar llamadas de larga distancia entre ciudades, así que las oficinas de las compañías fueron interconectadas. Pero luego, el mismo problema de interconectar todas las compañías salió a luz nuevamente, y entonces se creó un nuevo nivel de oficinas, como se muestra en la figura 1, eventualmente la jerarquía creció a cinco niveles.

La corriente DC y el tono de aviso fueron el primer tipo de señalización de telefonía que se utilizó para establecer y terminar llamadas telefónicas, a pesar que se realizaba de forma manual por el operador. Sin embargo, la señalización evolucionó, incluyendo hoy en día mucha más información de lo que los métodos anteriores podían hacer y reduciendo la intervención a su mínimo.

La señalización telefónica se encontraba limitada en sus inicios por el hecho que sobre el mismo circuito se transportaba la voz y la señalización, a este método se le denomina *in-band signaling*. Además, la señalización telefónica era análoga y tenía una pequeña cantidad de posibles estados y muy poca información podía ser manejada, haciendo necesaria la intervención del operador la mayoría de las veces. Para empeorar las cosas, el método señalización en banda (*in-band signaling*) provocaba que el circuito utilizado para la llamada telefónica estuviera ocupado, desde el momento que la persona marcaba hasta que se encontraba conectado con el otro extremo.

Por lo tanto, las compañías telefónicas se encontraron rápidamente sin circuitos para atender la demanda, ya que los clientes comenzaron a incrementarse a la cantidad de millones y se creó una gran cantidad de tráfico. Por un lado, las compañías de telefonía necesitaban una manera de administrar las llamadas que les ahorrara en las inversiones que debían realizar al agregar

nuevas instalaciones, pero por el otro lado, necesitaban encontrar métodos para soportar los servicios que los suscriptores estaban demandando.

A inicios de los años 60, las compañías de telefonía europeas empezaron a digitalizar sus redes. Uno de los primeros pasos tomados fue el separar la señalización de los circuitos de voz y se creó una red independiente para el uso de ésta, conocida como señalización de canal común (Common Channel Signaling, CCS).

Este nuevo método mostró beneficios, por ejemplo los procesos de establecer y desmontar llamadas eran realizados de manera más rápida y era menos propensa a errores. La digitalización de las líneas telefónicas no solo mejoró la calidad de las llamadas (especialmente las de larga distancia), sino que también permitió que los equipos utilizados disminuyeran de precio. CCS aún se utiliza ampliamente, es el protocolo SS7, el modelo y arquitectura que mayormente se utiliza en la actualidad y en las redes más nuevas. Sin embargo, en la historia de la señalización telefónica se han utilizado muchos otros métodos:

- DC signaling: éste fue el primer tipo de señalización utilizado. Cuando un suscriptor levantaba el auricular, una corriente DC corría desde la oficina de la compañía hasta el teléfono y de regreso. Un detector de corriente DC proveía de un tono de llamada y el suscriptor marcaba el número utilizando un marcador de rotación, que a su vez utilizaba un *relay* para interrumpir la corriente, creando una ráfaga de pulsos (10 pulsos por segundo).

  La oficina central determinaba el número marcado y establecía el circuito hacia el otro extremo. El suscriptor destinatario era alertado por el timbre

en su teléfono, mientras el suscriptor original continuaba recibiendo el tono de llamada. Cuando el destinatario contestaba, el tono era interrumpido y el circuito entonces transportaba la voz. El circuito era liberado cuando cualquiera de los dos suscriptores colgaba.

- *In-band signaling*: este método de señalización se basa en el uso de tonos a ciertas frecuencias en lugar de utilizar una corriente DC. Los tonos son transmitidos sobre el mismo circuito que la voz, y por ende, estos tonos deben estar en el mismo ancho de banda que la voz (0 a 4kHz aprox.). Los tonos están diseñados para minimizar la posibilidad de duplicar los tonos de voz, pero no son 100 % libres de fallas.

  Los tonos pueden ser a frecuencia simple (*single frequency*, SF), los cuales aún se utilizan en algunas partes de las redes telefónicas; frecuencia múltiple (*multi-frequency*, MF) o frecuencia múltiple de tono dual (*dual tone multi-frequency*, DTMF), mayormente utilizados para enviar información de los dígitos marcados hacia la oficina central de la compañía. Más allá de la posibilidad de malinterpretar los tonos de señalización como voz, este método requiere detectores de tonos bastante costosos y aun es limitado en los distintos valores que puede manejar.

- *Out-of-band signaling*: método muy similar al *in-band signaling*, con la diferencia que la portadora de voz está limitada a 3,5 kHz y la banda entre 3,5 kHz y 4 kHz está asignada a los tonos de señalización. Esto libra del error de malinterpretación pero las demás fallas no son solucionadas.

- Digital *signaling*: una de las técnicas usadas en la señalización cuando las redes telefónicas se digitalizaron, fue la utilización de ciertos *bits* de las tramas de la voz para la señalización. Esta práctica no deterioraba la calidad de la voz. Este método aunque significa menor costo, aún tiene la limitante del tipo de señalización que provee al no ser basada en mensajería.

- *Common channel signaling*: método también digital, pero su mayor diferencia es que coloca la información de señalización en un canal o *time slot* separado al de la voz, así que las tramas de voz y de data son utilizadas solamente para transportar la voz o la data del usuario.

  Este método es utilizado en SS7 y las redes ISDN (Integrated Services Digital Networks). Es capaz de enviar y recibir mensajes (data) que pueden tener gran cantidad de valores y por lo tanto puede extenderse a soportar nuevas funcionalidades. No solamente puede controlar el estado de las llamadas telefónicas, sino que también puede realizar solicitudes a base de datos remotos para soportar servicios especiales.

Entre todos estos métodos de señalización, el más importante de ellos es CCS, y se discutirá del mismo más adelante. En la siguiente sección se introducirá SS7 y se discutirá sobre su funcionalidad y arquitectura.

1.2. Redes SS7

CCS es mucho más flexible y poderoso que la señalización sobre el mismo canal, está bien adaptado para soportar los requerimientos de las redes digitales integradas. La culminación de la transición sobre las redes telefónicas de un método en el canal, a un método de canal-común, es SS7, emitido por

primera vez por el Comité Consultivo Internacional Telegráfico y Telefónico (Comité Consultatif International Télégraphique et Téléphonique, CCITT) en 1980, con revisiones cada cuatro años.

SS7 fue diseñado para ser un estándar de señalización de canal-común que puede ser utilizado por la gran variedad de redes de circuitos-conmutados. El propósito general de SS7 es proveer un estándar internacional de propósito general con las siguientes características principales:

- Optimizado para el uso en redes de telecomunicaciones digitales relacionadas con el intercambio de programas controlados digitalmente, usando canales digitales de 64 kbps. Sin embargo, está también adaptado para ser utilizado sobre canales análogos a la misma velocidad de 64 kbps.

- Diseñado para cumplir con los requisitos de transferencia de información actuales y futuros sobre control de llamada, control remoto, administración y mantenimiento.

- Provee un medio confiable para la transferencia de información en secuencia, sin pérdida o duplicación.

- Adaptada para el uso de conexiones punto a punto, terrestre o satelital.

La aplicación de SS7 es vasta, ya que cubre todos los aspectos de control de señalización para las redes digitales complejas. Sin embargo, el primer uso de SS7 no fue para el control de llamadas, sino para obtener accesos remotos a bases de datos.

En 1980, algunas compañías telefónicas empezaron a ofrecer un nuevo servicio denominado servicios telefónicos de amplia área (Wide Area Telephone Service, WATS), que usaba un código de área en común (800) sin importar donde fuese el destino de la llamada. Pero todo el equipo de conmutación en ese entonces tomaba decisiones de enrutamiento basados en el código de área a través de la red PSTN (Public Switched Telephone Network).

El problema fue resuelto al asignar un número “normal” secundario a cada número que iniciaba con 800, el cual contenía un código de área real y por lo tanto podía ser usado en el enrutamiento. Sin embargo, la cantidad de números con 800 creció rápidamente y fue necesario el almacenarlos todos en una base de datos central que pudiera ser consultada por las oficinas centrales. Por lo tanto, la red SS7 se inició al ser utilizada para obtener información de tarificación y rutas de la base de datos central, al realizar consultas dentro de los paquetes de mensajes.

Luego los servicios de SS7 fueron expandidos para incluir otros servicios incluyendo el control de inicio y fin de llamada. Otra función de las redes de telefonía obtenida a través de SS7 es LNP (Local Number Portability), la cual permite que los suscriptores puedan cambiar de compañía telefónica y mantener el número telefónico que utilizaban con anterioridad. LNP también puede evitar que el número cambie una vez ha sido actualizado el servicio de POTS (Plain Old Telephone Service) a ISDN (Integrated Services Digital Network). Este servicio requiere el uso de una base de datos que no difiere mucho de la utilizada para los números 800.

SS7 puede proveer mucho más que información de enrutamiento y tarificación. Provee los medios para que distintos equipos remotos de conmutación puedan comunicarse. Por ejemplo, si el número al que se llama se

encuentra ocupado, el suscriptor que llama puede utilizar un servicio que se conoce como *automatic callback*. Entonces, cuando el número al que se llamó se encuentra disponible, la red realiza una llamada al suscriptor que originó la llamada. Tan pronto el suscriptor conteste, se realizará automáticamente la llamada al número que deseaba contactar.

Esta característica es posible gracias a la capacidad de SS7 de enviar mensajes desde una central hacia otra, permitiendo que cada sistema pueda utilizar los servicios sin necesidad de establecer una conexión entre ambos.

Roaming es un servicio de las telefonías celulares que descansa sobre el protocolo SS7. Los proveedores de telefonía almacenan la información de sus suscriptores en bases de datos que se denominan HLR (Home Location Register), y comparten dicha información con otros proveedores de telefonía con los que tienen acuerdos. De este modo, el cliente no debe registrarse con otro proveedor cuando viaja**,** ya que la nueva red es seleccionada automáticamente.

Hoy en día SS7 es utilizada por la mayoría de compañías telefónicas independientes. Todas estas subredes propiedad de las compañías telefónicas, proveedores de servicios celulares y *carriers* internacionales, están interconectadas entre sí gracias al protocolo SS7. Esto hace de SS7 la red de comunicaciones más larga del mundo. En las siguientes secciones se desarrollará la estructura interna de esta red.

1.2.1. Funcionalidad SS7

Con la señalización de canal-común, los mensajes de control son enrutados a través de la red para realizar la administración de las llamadas

(establecimiento, mantenimiento y finalización) y las funciones de administración de la red.

Dichos mensajes de control son pequeños paquetes que deben ser enrutados a través de toda la red hasta su destino final. Por lo tanto, aun si la red que está siendo controlada es una red de conmutación de circuitos, la señalización de control es implementada utilizando conmutación de paquetes. En efecto, una red de conmutación de paquetes es sobrepuesta sobre la red de conmutación de circuitos, para operar y controlar la red de conmutación de circuitos.

SS7 define las funciones a ser realizadas en la red de conmutación de paquetes pero no dicta ninguna implementación de hardware particular. Por ejemplo, todas las funcionalidades de SS7 pueden ser implementadas en los nodos de las redes de conmutación de circuitos como funcionalidades extras; este método es denominado modo de señalización asociado (*associated signaling mode*).

Por otro lado, pueden existir distintos puntos de conmutación que transporten únicamente los paquetes de control, a este modo se le conoce como modo de señalización no asociado (*non- associate signaling mode*).

Aun en el segundo escenario, los nodos basados en conmutación de circuitos deben implementar partes de SS7 para que puedan recibir los paquetes de control. Hoy en día, los *switches* telefónicos utilizados en muchas de las operadoras realizan funciones de señalización. Esto es realizado al implementar un ordenador adjunto que se conecta a la red a través de un *link* digital. Dichos ordenadores son denominados SP (*Signaling Points*). Estos están a cargo de conmutar mensajes a través de la red, utilizando puntos de

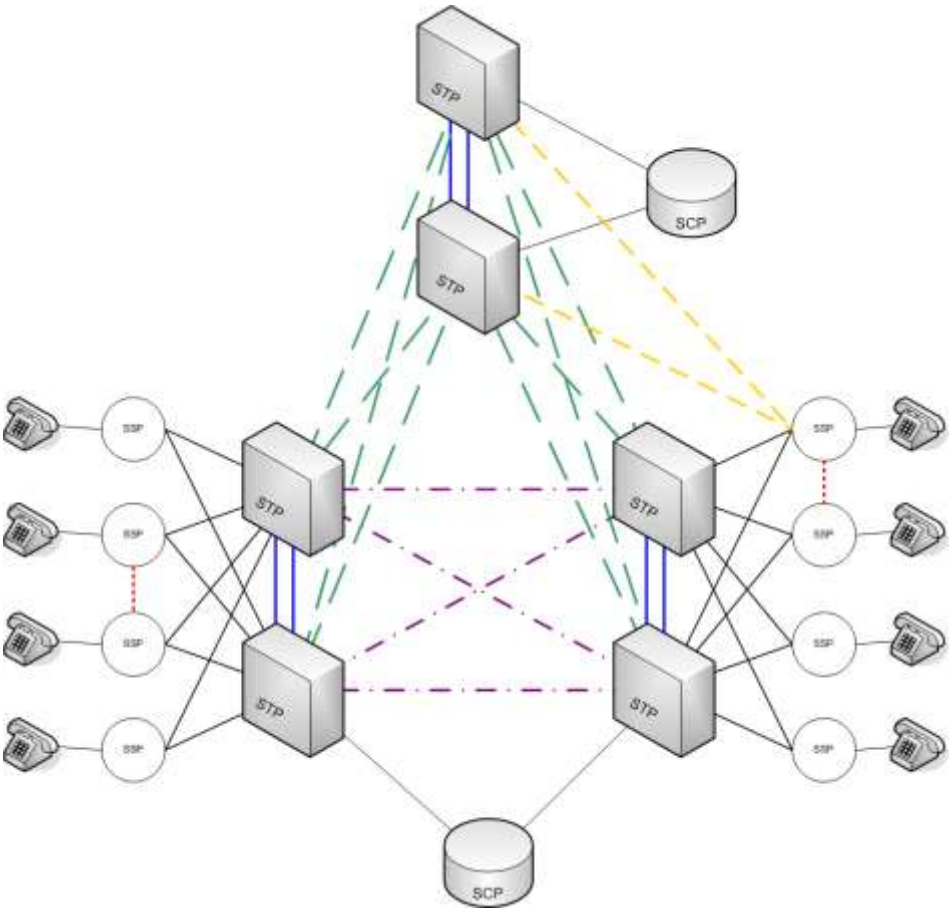
transferencia para enrutar los mensajes desde un punto a otro y también proveen acceso a bases de datos.

Todos los nodos en una red SS7 son llamados puntos de señalización. Un punto de señalización tiene la habilidad de realizar discriminación de mensajes (*Message Discrimination*), es decir son capaces de leer la dirección y determinar si el mensaje es para ese nodo, también pueden enrutar mensajes SS7 hacia otro SP. Cuando se utiliza SS7 para soportar la red inteligente (IN, Intelligent Network) se puede encontrar tres tipos distintos de SP:

- Service Switching Point (SSP)
- Signal Transfer Point (STP)
- Service Control Point (SCP)

SP proveen acceso a la red SS7, proveen acceso a las bases de datos utilizadas por los *switches* dentro y fuera de la red SS7 y transfieren mensajes SS7 a otros SP dentro de la red. Estos están interconectados gracias a los *links* de señalización que proveen la velocidad necesaria para la entrega de mensajes SS7. La arquitectura se muestra en la figura 2.

Figura 2. **Arquitectura funcional de SS7**



Fuente: elaboración propia, con Microsoft Office Visio 2007.

Tanto SP como *links* de señalización son siempre implementados en pares para redundancia. SS7 se asegura que la red se encuentre siempre operacional proveyendo rutas alternas en el caso de fallas. Esto asegura que los mensajes puedan siempre alcanzar su destino. La red es implementada en dos niveles o planos distintos. Hay un nivel internacional, que utiliza el estándar ITU-T del protocolo SS7, y hay un nivel nacional.

El nivel nacional se rige bajo cualquier estándar que exista en el país en el que está siendo implementado. Sin embargo, todos los países son capaces de comunicarse unos con otros a través de puertas de acceso (*gateways*) que convierten la versión nacional de SS7 hacia la versión internacional. Esto asegura que todas las naciones puedan trabajar con el resto, mientras aun atienden los requerimientos de su propia red.

1.2.1.1. Service Switching Point (SSP)

SSP es la central local de una red de telefonía. Un SSP puede ser una combinación de un *switch* de voz y uno de SS7, o un ordenador adjunto conectado al *switch* de voz de la central local. SSP debe convertir la señalización del *switch* de la voz hacia mensajes de señalización SS7, los cuales pueden ser enviados a otras centrales a través de la red SS7. La central enviará mensajes relacionados con sus circuitos de voz hacia las demás centrales con conexión directa a ella.

La función de los SSP es la de utilizar la información provista por la parte que llama (como los dígitos marcados), y determinar cómo conectar la llamada utilizando tablas de enrutamiento. Enviará un mensaje SS7 a la central adyacente correspondiente, solicitando una conexión. La otra central reconoce la solicitud y otorga el permiso para conectar la troncal. El mismo procedimiento es repetido, conectando troncales entre varias centrales adyacentes hasta que se alcance el destino final.

Muchas funciones de SSP son conseguidas al adjuntar ordenadores a los *switches* ya existentes. Esta computadora recibe señales del *switch* de voz que son utilizados para disparar mensajes específicos de SS7. Utilizar este método de ordenadores adjuntos permite a las compañías telefónicas, actualizar sus

puntos de señalización (SP) sin necesidad de remplazar costosos *switches* que realicen ambas tareas, proveyendo una construcción de red modular. Las actualizaciones se convierten entonces únicamente en cambios de software.

Un SSP debe tener la habilidad de enviar mensajes utilizando el protocolo ISUP (ISDN User Part) y el protocolo TCAP (Transaction Capabilities Application Part) los cuales se describirán más adelante.

1.2.1.2. Signal Transfer Point (STP)

Todos los paquetes SS7 viajan desde un SSP hacia otro, por al menos un STP. El STP actúa como un enrutador en la red SS7 y usualmente no origina ni termina los mensajes. El STP típicamente se encuentra adjunto a un *switch* de voz y raramente es un sistema construido solamente para la funcionalidad de STP. Existen tres niveles de STP.

- STP Nacional existe dentro de una red nacional y es capaz de transferir mensajes utilizando los estándares nacionales del protocolo. Los mensajes pueden ser transferidos a otro nivel de STP pero los STP nacionales no tienen capacidad de convertir mensajes a otra versión o formato.

- STP Internacional funciona de la misma manera que el anterior, pero éste opera sobre la red internacional. La red internacional provee de conectividad con las redes externas globales utilizando los estándares ITU-T. Todos los nodos que se conecten al STP Internacional deben cumplir con el protocolo estándar ITU-T.

- STP Gateway (puerta de enlace), provee la conversión de protocolo desde un estándar nacional al estándar ITU-T y vice-versa. Este STP es utilizado como un acceso a la red internacional.

El STP Gateway funciona como interfaz hacia otra red. Los proveedores de larga distancia o servicios internacionales, pueden tener acceso a información del suscriptor local a través de las bases de datos de la compañía local, o la compañía local puede necesitar información de la base de datos del proveedor internacional. En cualquier caso, esto se realiza a través del Gateway STP.

El Gateway STP utiliza funciones de detección para mantener la seguridad de la red. Tienen la capacidad de examinar todos los paquetes entrantes o salientes y permitir el paso solamente a aquellos autorizados. Cuando se considera el rendimiento de la red, un nivel de STP es recomendable. Sin embargo, al considerar fiabilidad y disponibilidad de la misma se proponen soluciones de más niveles. Las siguientes son sugerencias del estándar ITU-T:

- En una red de señalización jerárquica con un solo nivel de STP:
  - Cada SP que no cumpla funciones de STP debe estar conectado mínimo a dos STP.
  - El interconectado de STP debe ser tan completo como sea posible.

- En una red de señalización jerárquica con dos niveles de STP:
  - Cada SP que no sea también un STP debe estar conectado mínimo a dos STP de un nivel superior.
  - Cada STP en un nivel inferior está conectado mínimo a dos STP de nivel superior.

- Los STP de nivel superior deben estar completamente interconectados.

La figura 2 es un ejemplo de una red jerárquica con dos niveles de STP. Los cuatro STP en la parte baja de la figura podrían ser STP de la red nacional, mientras que los otros dos STP podrían ser *gateways* o STP internacionales en la red nacional SS7, según especificación de ITU-T. Más allá de las tareas de ruteo, el STP realiza tareas de medición.

Existen dos tipos básicos de medición: mediciones de tráfico y mediciones de uso.

Las mediciones de tráfico proveen datos estadísticos correspondientes al tipo de mensajes que entran y salen de la red. Los eventos de la red como duración de un *link* fuera de servicio, uso de procesador, etc., son también almacenados sobre todo para propósitos de mantenimiento. Debido a la velocidad de respuesta de SS7, las mediciones de tráfico son la mejor manera en que el personal de mantenimiento pueda mantener un monitoreo sobre lo que sucede en la red y prevenir fallas.

Las mediciones de uso mantienen un conteo y estadísticas de la cantidad de mensajes que entran y salen de la red (por tipo de mensaje). Estos conteos son almacenados en una cinta magnética. Esta cinta es utilizada para la creación de las facturas de los clientes.

En la red SS7 el STP recibe mensajes del SSP. Estos paquetes están relacionados a conexiones de llamadas o consultas a base de datos. El acceso a la base de datos está provisto por otro elemento de la red el SCP. Si el SSP no conoce la dirección de destino del SCP, el STP debe proveer dicha

dirección. En este caso, el SSP envía una consulta a la base de datos hacia el STP local. El STP obtendrá los dígitos marcados (denominados Global Title Digits) y determina a través de sus tablas, la dirección de la base de datos. Esto se denomina Global Title Translation. El STP es entonces el nodo más versátil en SS7, pues provee una gran variedad de servicios a los usuarios de la red.

1.2.1.3. Service Control Point (SCP)

El SCP funciona como una interfaz a las bases de datos de la compañía telefónica, no necesariamente almacena información, pero actúa como conexión hacia el sistema que mantiene la información. Estas bases de datos se utilizan para almacenar información sobre servicios a suscriptores, enrutamiento de numeración especial, validación de tarjetas para llamadas o protección contra fraude.

El SCP es usualmente una computadora que funciona como entrada al sistema de base de datos. Esta base de datos usualmente se conecta al SCP utilizando *links* X.25, pero en sistemas STP/SCP integrados, la base de datos reside en el SCP. El SCP puede entonces realizar conversiones de protocolo entre SS7 y X.25, o puede proveer una interfaz para acceder a la base de datos directamente. El protocolo utilizado para acceder e interactuar con la base de datos es TCAP. El tipo de base de datos depende de la red. Las más utilizadas en las redes son:

- Call management services database (CMSDB): provee instrucciones de enrutamiento para números de servicios especiales, así como información de facturación. También provee información de rutas para evitar nodos congestionados.

- Local number portability (LNP): esta base de datos contiene la información que permite a los suscriptores cambiar de compañías telefónicas sin necesidad de cambiar sus números telefónicos.

- Line information database (LIDB): provee información sobre los suscriptores, como clase de servicio, facturación a terceros y características de llamadas especiales como *speed dial* y *call forwarding*.

- Business service database (BSD): permite a los suscriptores almacenar instrucciones para el proceso de sus llamadas, también almacena procedimiento para la administración de la red.

- Home location register (HLR): este tipo de base de datos hace su aparición en redes de telefonía celulares. El HLR almacena información sobre facturación, servicios disponibles, así como la ubicación actual del teléfono celular.

- Visitor location register (VLR): almacena las ubicaciones actuales para los suscriptores visitantes, cuando estos se encuentren fuera de su red de origen.

Cada base de datos contiene información sobre aplicaciones específicas y posee una dirección, denominada Subsystem Number (SSN), éste es utilizado en consultas de rutas desde los SSP a través de la red SS7 hasta la base de datos destino.

1.2.1.4. *Links* de señalización

Los *links* son bidireccionales y *full-duplex* trabajan a velocidades que varían desde 4,8 Kbps hasta 1,536 Mbps, dependiendo del estándar nacional de la red SS7. Los *links* se implementan en grupos, denominados *linksets*. Todos los *links* dentro del grupo deben tener el mismo nodo adyacente. El equipo de conmutación alternará la transmisión entre todos los *links* dentro del *linkset* para asegurar igual uso de ellos. Se pueden asignar hasta 16 *links* a un *linkset*.

En el caso común de que un nodo tenga conexión con dos STP, los *links* son asignados a dos *linksets*, uno por cada STP. Ambos *linkset* pueden luego ser configurados como un *linkset* combinado. Los *linkset* combinados son utilizados para compartir carga, donde el SP que origina puede enviar mensajes a ambos STP dividiendo así la carga a través de los *links*. Dentro de la red SS7, *linksets* alternados son utilizados para proveer rutas alternas para los mensajes. Un *linkset* es utilizado cuando existe una congestión sobre el *linkset* primario, logrando así tomar ventaja de la diversidad de rutas existentes para evitar mayor congestión.

Los *links* deben estar siempre disponibles para el tráfico de SS7 con un mínimo de indisponibilidad, (se debe permitir un máximo de 10 minutos de baja por año para cualquier *linkset*). Cuando un *link* falla, los demás *links* dentro del *linkset* deben tomar el tráfico. De igual manera, si un nodo dentro de la red (por ejemplo un STP) falla es el nodo pareja el encargado de tomar la carga.

Esto quiere decir que en cualquier momento un *link* puede ser sobrecargado más allá de lo que pueda soportar. Por esta razón, los nodos de la red SS7 deben estar restringidos a no mandar más de 40 % de tráfico en

cada *link*, así en el caso de una falla, cualquier *link* del grupo es capaz de manejar un 80 % de carga en respaldo por el *link* fallido y aun tener un rango de guarda por cualquier tráfico extra.

Si la longitud promedio de un mensaje es de 40 *bytes*, ej. 320 *bits*, y se considera las especificaciones ANSI de SS7 con *links* a 56 Kbps, trabajando a un 40 % da una capacidad disponible de 22,4 Kbps. Cada *link* puede transportar hasta 70 mensajes por segundo. Esta fórmula es utilizada para el dimensionamiento de la red. Existen seis distintos tipos de *links* de acuerdo a su función:

- Access links (A): son utilizados entre el SSP y el STP, o SCP y STP, Proveen el acceso a la red y a las bases de datos a través del STP.

- Bridge links (B): son utilizados para conectar STP acoplados con otros STP acoplados del mismo nivel jerárquico. *Links* B son implementados de cuatro en cuatro como se muestra en la figura 2, se pueden utilizar un máximo de ocho *links* B entre STP acoplados.

- Cross links (C): son utilizados para conectar un STP a su par STP. Tráfico normal de SS7 no es enrutado a través de este *link*, a excepción que se esté experimentando congestión o se tenga aislamiento de un nodo y el único camino sea a través de los *links* C. Los únicos mensajes que viajan a través de los *links* C en condiciones normales son mensajes de administración de la red.

- Diagonal links (D): son utilizados para conectar un par de STP de un nivel jerárquico con otro par de STP de un distinto nivel jerárquico. Poseen las mismas características que los *links* C.

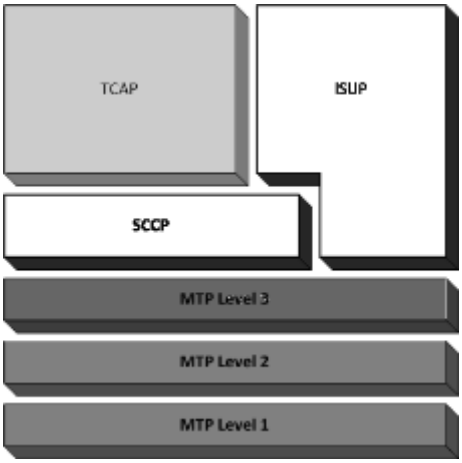
- Extended links (E): son utilizados para conectar SSP con STP remotos. Se utilizan como ruta alterna para mensajes de SS7 en caso de congestionamiento contra un par de STP.

- Fully associated links (F): son utilizados cuando se tiene gran cantidad de tráfico entre dos SSP, o cuando económicamente no sea conveniente tener una conexión directa entre SSP y STP. Solamente se cuenta con tráfico sobre procedimientos para establecer y desconectar llamadas a través de estos *links*.

1.2.2. Protocolo SS7

Así como el modelo OSI (Open System Interconnection), el estándar SS7 posee una arquitectura de capas. El término nivel en SS7 es utilizado bajo el mismo contexto como lo es el término capa en el modelo OSI. La figura 3 muestra la estructura de SS7.

Figura 3. **Arquitectura del protocolo SS7**



Fuente: elaboración propia, con Microsoft Office Visio 2007.

Los tres niveles más bajos de la arquitectura del protocolo de SS7 son referidos como MTP (Message Transfer Part), estos proveen un confiable enrutamiento de los mensajes a través de la red SS7.

En las versiones SS7 de 1984 en adelante, se agrega un módulo adicional, el cual reside en el nivel 4 de la arquitectura, dicho módulo se conoce como SCCP (Signaling Connection Control Part). En conjunto el SCCP y MTP se conocen como NSP (Network Service Part). SCCP define una variedad de distintos servicios de red para cumplir las demandas de los usuarios de NSP. El resto de módulos de SS7 se consideran de nivel 4 e incluye a los distintos usuarios de NSP, que se puede considerar como un sistema de entrega de mensajes, las demás partes tratan con el contenido de dichos mensajes.

ISUP (ISDN User Part) provee el necesario control de señalización en ISDN para tratar con las llamadas de los suscriptores y sus funciones relacionadas, la mayoría para establecer y terminar llamadas entre nodos terminantes. ISUP se deriva de TUP (Telephone User Part) el cual es el equivalente a ISUP en las especificaciones ITU-T. Además de las funcionales TUP, ISUP ofrece también el beneficio de soportar funciones de IN y servicios de ISDN. TCAP (Transaction Capabilities Application Part), se introdujo en 1988, provee los mecanismos para aplicaciones y funciones en una arquitectura orientada en transacciones.

Existen otros protocolos que al ser derivados de TUP forman parte de SS7, por ejemplo BISUP (Broadband ISDN User Part), el cual es utilizado para establecer y terminar circuitos de BISDN (Broadband ISDN) o como el protocolo DUP (Data User Part), diseñado para proveer capacidad de transmisión de data a través de redes en modo de circuitos. Sin embargo, DUP no está relacionado con ISDN como lo está ISUP por lo que su uso es ya casi obsoleto.

1.2.2.1. Message Transfer Part (MTP)

El Protocolo MTP se encuentra en los niveles bajos de la pila de SS7 y es un protocolo de transporte utilizado por los niveles superiores. Se divide en tres distintos niveles con las mismas funcionalidades de las capas uno, dos y tres del modelo OSI (Física, Datos, Red).

- MTP level 1 (MTP1) permite el uso de interfaces de tipo digital. Interfaces comunes en la mayoría de redes SS7 incluyen E1 (canales de 2,048; 64; 32 Kbps), DS1 (canales de 1,544 Mbps; 64; 24 Kbps), V.35 (64 Kbps), DS0 (64 Kbps), y DS0A (56 Kbps).

- MTP level 2 (MTP2) provee las funciones necesarias para detección y corrección de errores. Este nivel se enfoca en proporcionar una entrega confiable entre dos SP, no posee conocimiento del destino final del mensaje.

  Este nivel provee funcionalidades de control de flujo y numeración secuencial de las unidades de la señal enviadas a través de la conexión punto a punto.

  Otra funcionalidad es la corrección de errores. Existen dos procedimientos para la corrección de errores. Basic error correction es utilizado en *links* con retraso menor a los 15ms. Este utiliza retransmisión *go-back-n***,** donde un cuadro perdido o corrupto y todos los cuadros siguientes se vuelven a transmitir por el que origina.

  El esquema preventive cyclic retransmission (PCR) se utiliza en links con mayor rango de retraso. En PCR las señales son retransmitidas

automáticamente durante periodos libres (sin transmisión), hasta que se recibe un reconocido del receptor.

MTP2 también realiza dos funciones para monitorear la tasa de errores. Signal unit error rate, realiza un conteo de errores individuales utilizando un esquema conocido como *leaky bucket*: un contador es incrementado en uno, cada vez que se detecta un error y es reducido en uno, después de haber recibido exitosamente 256 unidades. Si el contador alcanza el valor de 64, una indicación es enviada a MTP3.

Alignment error rate es utilizado para asegurar la sincronía de las unidades transmitidas.

- MTP level 3 (MTP3) tiene la responsabilidad de transportar los mensajes entre SP. Existen dos mayores funcionalidades realizadas por este nivel administración de red (*network management*) y manejo de mensajes (*message handling*). *Network management* está a cargo de proveer configuración de la red de señalización en el caso que se tengan fallas de SP o *links*. También controla el tráfico en caso de congestión. Las funcionalidades de señalización en la administración de la red son:

    - Signaling link management: activa nuevos *links* y reinicializa o remueve de la operación los *links* fallidos (siguiendo indicaciones de MTP2). Informa solamente del problema al SP que se encuentra al otro extremo del *link* con fallas. Por lo tanto, es un proceso local. El re-enrutamiento de tráfico no es parte de las funciones de esta administración. Otra característica que algunos SSPs proveen que es responsabilidad de *signaling link management* es automatic allocation. Este consiste en remover

circuitos de voz para utilizarlos como *links* de señalización SS7 o viceversa.

- Signaling traffic management: realiza en cierta manera una tarea similar a *signaling link management*, ya que también realiza reemplazo de *links* de señalización. Sin embargo, trata únicamente con los *links* que han sufrido una falla completa. Los mensajes utilizados para deshabilitar el *link* que ha fallado son enviados por distintas rutas. Así que, la diferencia básica entre *signaling traffic management* y *signaling link management* reside en el mecanismo utilizado para informar al SP adyacente sobre la falla. Sera entonces *signaling link management* el encargado de deshabilitar el *link* fallido.

- Signaling Route Management: es utilizado para avisar a otros SP acerca de la incapacidad de un SP para alcanzar a otro SP. Por lo tanto, cuando un SP identifica que no puede comunicarse con un SP adyacente, éste informa a los demás SP que eviten enviar mensajes al SP aislado.

Dentro de la categoría del manejo de mensajes se pueden encontrar tres grandes funciones:

- Message Discrimation: determina si determinado mensaje MTP2 pertenece a un específico SP o a otro, basado en la etiqueta de ruta del mensaje. Si la etiqueta de ruta del mensaje contiene la dirección de un SP local, este mensaje es entregado para su distribución (*message distribution*). De lo contrario, es enviado a ruteo del mensaje (*message routing*).

- Message Distribution: si el mensaje pertenece a este SP, el mensaje es enviado al respectivo usuario MTP (ISUP o TCAP) o a su función MTP3.

- Message Routing: si el mensaje MTP2 recibido pertenece a otro SP o si es un mensaje originado en este SP, debe ser enviado a través de otro *link* de señalización, el cual se obtiene gracias a las tablas de ruteo.

Una gran parte de las especificaciones de la red y su funcionalidad involucran procesos para superar fallas de *links* y congestión. Estos procesos son especificados para determinar rápidamente cuando existe una falla de *link*, deshabilitarlo del servicio, reenrutar el tráfico y habilitar de nuevo el *link* luego de ser reparado. Existe entonces una preocupación por la fiabilidad de la red, el objetivo debe ser 99 998 % de disponibilidad. Este objetivo se cumple en SS7 gracias a su redundancia de equipo y el dinamismo de reconfiguración de la red y su capacidad de reenrutamiento.

1.2.2.2. Signaling Connection Control Part (SCCP)

MTP fue diseñado originalmente para cubrir las demandas en tiempo real de las redes de señalización telefónica y por esta razón, provee un servicio no orientado a conexión. Algunas aplicaciones sin embargo, requieren capacidad de transferencias orientadas a conexión y un área más completa para su direccionamiento que la que provee MTP.

MTP provee el OPC (origination point code) y el DPC (destination point code), con una longitud de 14 *bit*. En ambos casos, el *point code* es desde una perspectiva nodo a nodo. Además, MTP tiene una capacidad limitada de

distribución utilizando un indicador de 4 *bits* en el campo SIO (service indicator octet). Esta capacidad de direccionamiento es adecuada solamente para un número reducido de servicios.

Una de las más grandes mejoras provistas por SCCP es su campo de direccionamiento expandido. SCCP complementa el direccionamiento MTP al definir un campo adicional denominado SSN (subsystem number), que consiste en información local utilizada para identificar al usuario SCCP en cada nodo.

La combinación de OPC y SSN forman la dirección de la parte que se llama (calling party address), y la combinación DPC y SSN es la dirección de la parte a la que se llama (called party address). Otra mejora de SCCP es la capacidad de utilizar *global titles* como dirección. El *global title* es una dirección especial que no provee información para enrutamiento. SCCP es el protocolo que realiza el *global title transaltion*. SCCP se utiliza únicamente con TCAP, aunque el estándar indica que puede utilizarse con ISUP.

Esto en teoría, permitiría que mensajes ISUP se asociarán a una conexión ya establecida para ser enrutadas utilizando la conexión punto a punto, como se hace con TCAP. Sin embargo, esta funcionalidad no ha sido implementada en las redes SS7.

1.2.2.3. Transaction Capabilities Application Part (TCAP)

TCAP provee funcionalidades de operaciones remotas para SS7. Provee la capacidad para que una aplicación en un nodo, invoque la ejecución de cierta operación en otro nodo y que reciba el resultado de dicho nodo remoto. TCAP fue diseñado originalmente para soportar consultas a base de datos.

TCAP comprende dos subcapas denominadas TSL (transaction sublayer) y CSL (component sublayer). TSL es la subcapa más baja de TCAP y define como se realizará la transacción o el diálogo.

Existen dos tipos de diálogos: diálogo no estructurado, que es comunicación en un sentido en el cual el proceso remoto recibe el mensaje pero no envía ninguna respuesta de vuelta, y el dialogo estructurado, en el cual las consultas producen una respuesta.

CSL es la subcapa superior de TCAP y define los mensajes reales llamados componentes, que están contenidos en los mensajes TSL. Existen cuatro tipos de componentes CSL: *invoke* (solicita una operación remota), *return result* (contiene la respuesta a la operación solicitada), *return error* (indica cualquier tipo de error), y *reject* (indica algún error de sintaxis).

Los servicios TCAP son provistos a una aplicación superior que se denomina ASE (application service element), responsable de proveer información que requieran las aplicaciones específicas, como obtener el número de facturación de una tarjeta telefónica.

1.2.2.4. ISDN user part (ISUP)

ISUP es un protocolo relacionado a circuitos utilizado para establecer, administrar y liberar troncales transportando llamadas de voz y data sobre PSTN. Es utilizado tanto para llamadas ISDN o No-ISDN y fue adoptado por SS7 para reemplazar a TUP, el cual no soportaba transmisión de data o circuitos digitales.

Sin embargo, ISUP no soporta tecnologías de banda ancha. Estas nuevas tecnologías se trabajarán bajo una nueva versión de ISUP denominada BISUP y que aún está bajo desarrollo de ITU-T.

ISUP puede utilizar servicios de transporte ya sea del nivel MTP o SCCP. Sin embargo, la interfaz entre ISUP y SCCP no ha sido implementada. Los servicios de MTP son utilizados para el transporte de los mensajes de señalización relacionados con la llamada. Mientras que SCCP podría ser utilizado para servicios de conexión adicionales, así como señalización punto a punto.

ISUP es compatible con el protocolo ISDN, el cual fue desarrollado como una extensión de SS7 para el suscriptor. El propósito de la compatibilidad con ISDN es el permitir a los suscriptores enviar información de señalización a los suscriptores remotos. Esto puede ser utilizado para soportar características como llamadas en conferencia o devolver llamada automáticamente.

1.3. El internet

El internet y los servicios que ofrece se han convertido en los últimos años en un fenómeno en masa, que cada vez gana más popularidad. Detrás de este crecimiento existe un protocolo denominado IP (internet protocol).

IP ha demostrado ser un protocolo de red muy robusto que puede soportar la implementación de nuevas tecnologías con mínimos cambios, siendo este protocolo aun válido luego de más de 30 años desde su diseño inicial. Esto es porque a diferencia de muchos otros protocolos de red, IP fue diseñado desde el comienzo pensando en interconectividad de redes.

El número de usuarios del internet ha estado siempre en crecimiento, sin embargo, no ha sido sino hasta años más recientes que la comunidad de usuarios de Internet se convirtió en un grupo más que significante. Esto sucedió gracias al desarrollo de nuevas aplicaciones que utilizan el internet y que hacen posible una nueva era de comunicación.

1.3.1. Historia de internet

A mediados de los años 60, a la altura de la guerra fría, el Departamento de Defensa DoD (Department of Defense) de los Estados Unidos de América, querían tener una red que sobreviviera una guerra nuclear. Las redes de conmutación de circuitos eran consideradas muy débiles debido a que la pérdida de un nodo o línea podía terminar con todas las comunicaciones que estuvieran utilizando esta red e incluso podía dividirla.

El Departamento de Defensa a través de su área de investigación ARPA (Advanced Research Project Agency), inició la investigación sobre una red utilizando la idea de conmutación de paquetes. Teniendo una subred de datagramas, si alguna línea o nodo se perdía, los mensajes podían ser entregados automáticamente por una ruta alternativa.

ARPA dio concesiones a ciertas universidades para que investigaran sobre este tema, y finalmente en diciembre de 1969, una red de conmutación de paquetes con 4 nodos nació, la red ARPANet.

ARPANet creció rápidamente, y unos años después, experimentos mostraron que los protocolos existentes de ARPANet no eran capaces de correr sobre múltiples redes. Esta observación incrementó la investigación sobre protocolos, terminando en la invención de TCP (Transmission control protocol),

y el modelo TCP/IP en 1974. TCP/IP fue específicamente diseñado para manejar la comunicación entre redes.

Para 1983 ARPANet era estable y exitosa, con más de 200 redes y cientos de *hosts*, con TCP/IP como el único protocolo estándar que se utilizaba. DNS (Domain Name System) fue creado durante los años 80 para organizar ordenadores dentro de dominios y mapear los nombres de los *hosts* dentro de direcciones IP. En los años 80 ARPANet estaba conectada a varios nodos fuera de los Estados Unidos, la mayoría en Europa y Japón, pero el crecimiento y evolución real de internet sucedía en Norte América.

Para 1990, ARPANet había sido superada por redes más recientes que ella misma había generado, por lo que fue cerrada y desmantelada. Para los años 70s la NSF (National Science Foundation) de los Estados Unidos de América, se dio cuenta del impacto que ARPANet tuvo en universidades y centros de investigación y su capacidad para compartir ideas y proyectos. El problema era que las universidades que desearan unirse a ARPANet debían tener un contrato de investigación proporcionado por su dueño, el Departamento de Defensa.

Así que, la NSF empezó a diseñar un sucesor de alta velocidad para ARPANet disponible a todas las universidades, y el resultado de dicha investigación fue NSFNet, fundada a mediados de los años 80's utilizando el mismo hardware que ARPANet. NSFNet fue un éxito instantáneo. Pocos años después de su lanzamiento ya contenía miles de *hosts* de las distintas universidades, laboratorios de investigación, bibliotecas y museos, incluyendo *hosts* conectados a ARPANet.

En los años siguientes, los *links* utilizados por la red debieron ser actualizados de 56 Kbps, a 1,5 Mbps en 1990. Sin embargo, esta actualización no era gratis, y se hizo obvio que el gobierno no estaba en la capacidad de financiar la red. Ese mismo año algunas compañías crearon una corporación sin fines de lucro denominada ANS (Advanced Networks and Services), este fue el primer paso para la comercialización de NSFNet. ANS se hizo cargo de la red NSF y actualizó sus *links* a 45 Mbps. Para este punto el internet tenía ya 200 000 ordenadores contenidos en aproximadamente 3 000 redes.

En 1991, el Congreso de los Estados Unidos aprobó la creación de NREN (National Research and Education Network), la red sucesora de NFSNet, corriendo a velocidades de *Gigabits*. Al inicio de los años 90 las compañías comerciales comenzaron a implementar sus propias redes basadas en IP, así que la infraestructura de NSFNet ya no era necesaria. NSFNet fue vendida a AOL (America On Line) en 1995, y desde entonces el internet ya no fue mantenido por los Estados Unidos o los gobiernos locales.

Hasta inicios de los años 90s los servicios tradicionales provistos por internet eran *e-mail* (la aplicación más popular desde ARPANet), *news* (foros internacionales sobre diferentes temas), *remote login* (normalmente usando el protocolo Telnet para el acceso remoto) y *file transfer* (para realizar copias de archivos utilizando FTP (File Transfer Protocol) o TFTP (Trivial File Transfer Protocol). Dichos servicios eran mayormente utilizados por investigaciones académicas, gubernamentales e industriales.

Pero en 1990, Tim Berners-Lee, un científico trabajando para El Consejo Europeo de Investigación Nuclear CERN (Conseil Européenne pour la Recherche Nucléaire), creo el Protocolo de Transferencia de Hipertexto HTTP (Hypertext Transfer Protocol), el lenguaje que utilizarían las computadoras para

comunicar documentos de hipertexto sobre el internet, él también diseñó un esquema para dar direcciones en el internet a dichos documentos, el URI (Uniform Resource Identifier).

A final de 1990 creó un servidor de documentos de hipertexto, y un programa cliente (*browser*) para obtener y ver dichos documentos. El nombró a esta aplicación WWW (World Wide Web). El siguiente año, en 1991 hizo público en el internet, a su servidor y clientes y lo que hoy se conoce como La Red, tuvo su inicio.

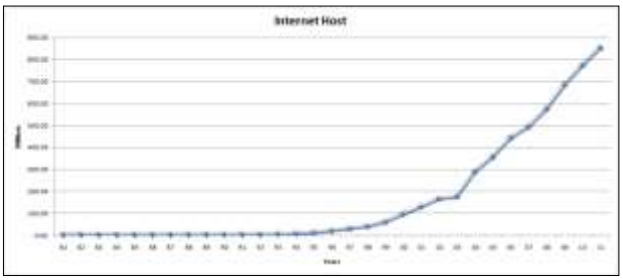
El navegador (*browser*) de Berners-Lee estaba específicamente diseñado para la computadora personal que él utilizaba, así que otros, mayormente estudiantes, empezaron a programar sus propios navegadores. El primer navegador con soporte para multimedia fue Mosaic, escrito por la NCSA (National Center for Supercomputer Applications) en 1993, y luego de éste, los desarrollos se dieron tan rápido que se volvió prácticamente imposible de seguir la pista a los avances.

La aparición de WWW fue la aplicación que atrajo a millones de nuevos usuarios no académicos a la red, y es la aplicación que ha hecho tan popular al internet. Su primer uso fue el hacer más simple el compartir documentos entre científicos e investigadores, pero hoy en día su uso es mayormente comercial. Virtualmente no existe ninguna compañía que no posea una página web ofreciendo productos, ni gobiernos que mantengan páginas web donde se pueda realizar algún trámite oficial.

La figura 4 muestra el crecimiento exponencial del internet desde sus inicios, publicada por ISC (Internet Software Consortium). La última medición de

la cantidad de *hosts* conectados a Internet fue tomada en junio de 2011 y para ese entonces se tenían aproximadamente 850 millones de *hosts*.

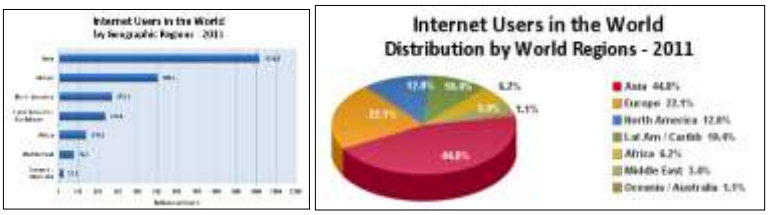
Figura 4. **Crecimiento de internet (1981-2011)**



Fuente: ISC, Internet Host History, https://www.isc.org/solutions/survey/history. Consulta: 16 de mayo de 2012.

El método utilizado para medir la cantidad de *hosts* conectados a internet son diferentes y sus resultados varían considerablemente al consultar otras fuentes. Por otra parte, el número de *hosts* conectados a internet no es la única variable que da una idea del éxito de éste. Como ejemplo se adjunta las estadísticas provistas por el sitio web internet *world stats*, el cual es un sitio que ofrece estadísticas actualizadas del uso de internet, estadísticas de población que lo utiliza y estadísticas de uso del Mercado para más de 233 países individuales y regiones del planeta. Dichos datos se observan en las figuras 5.

Figura 5. **Usuarios de internet por regiones del mundo**



Fuente: Internet World Stats, http://www.internetworldstats.com/stats.htm. Consulta: 16 de mayo de 2012.

Sin embargo IP aún no estaba listo para tan increíble éxito. Debido a las nuevas aplicaciones que hacen tan interesante el internet para el público en general, el número de usuarios en línea ha crecido exponencialmente desde mediados de la década de los años 90 y se espera que su crecimiento continúe aún más. Hay millones de usuarios utilizando dispositivos *wireless* para mantenerse conectados a un nodo central. Además, está la convergencia de computadoras, comunicaciones, la industria del entretenimiento, y ya se tienen desde teléfonos celulares, hasta televisores, que representan cada uno un nodo más conectado a internet.

Esto trajo dos problemas: por un lado, las direcciones IP son números de 32 *bits*, lo cual da un máximo teórico de 4 billones de direcciones para *hosts*. Pero la práctica de organizar las direcciones IP en distintas clases para mejorar el enrutamiento, ha desperdiciado millones de estas direcciones. Así que, con el crecimiento del internet las direcciones IP se han convertido en un tema prioritario. Por otro lado, el tener tal cantidad de *hosts* convierte a los algoritmos de enrutamiento en deficientes, haciendo que el enrutamiento sea más lento e incrementando el consumo de recursos.

Bajo estas circunstancias se hizo imperativo que IP debía evolucionar y hacerse más flexible. Así que en 1990 la IETF (Internet Engineering Task Force), el organismo internacional que produce los estándares con respecto a internet, comenzó a trabajar sobre una nueva versión de IP. La principal característica de esta nueva versión debe ser el uso de un espacio mayor para la dirección, a fin de nunca quedarse sin direcciones IP, pero al mismo tiempo debe solventar varios inconvenientes:

- Reducir el tamaño de las tablas de ruteo.

- Simplificar el protocolo permitiendo a los *routers* realizar su trabajo más rápido.

- Proveer mejoras en seguridad (autenticación y privacidad) que la versión actual de IP.

- Prestar mayor atención al tipo de servicio, particularmente para data en tiempo real.

- Hacer posible para un *host* que cambie de lugar sin cambiar su dirección.

- Hacer el protocolo lo suficientemente abierto para que pueda evolucionar en el futuro.

- Permitir que ambas versiones de IP coexistan por años.

IPv6 aún no está completamente desplegado, y prácticamente el único protocolo usado en internet es la antigua versión de IP, ahora llamado IPv4.

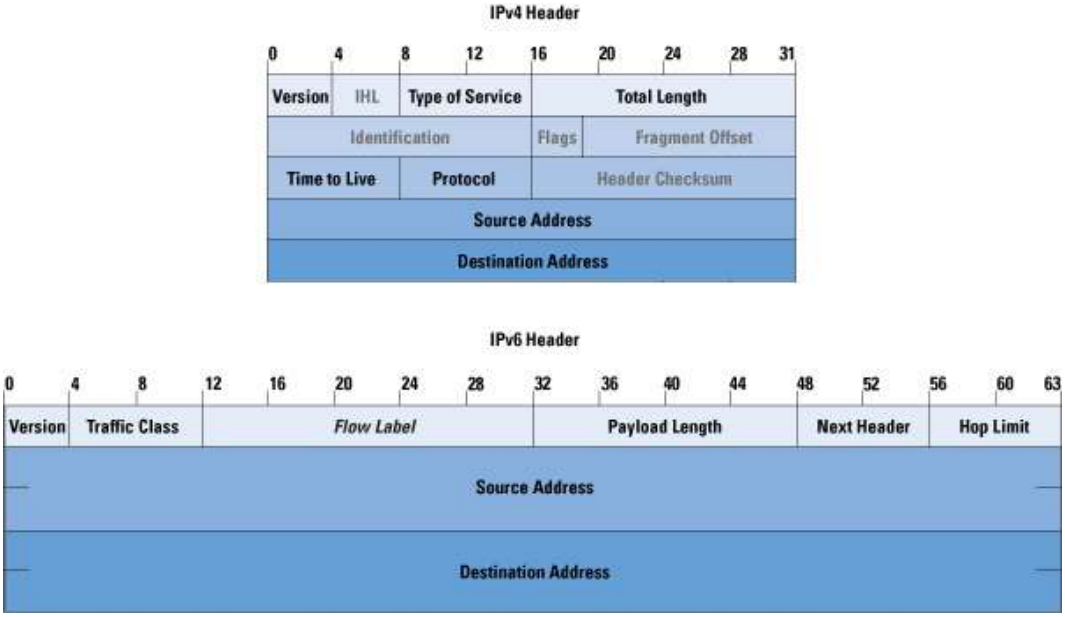
Pero, ya se tienen varias implementaciones IPv6 que trabajan en conjunto con IPv4.

1.3.2. Protocolo de internet (IP)

IP es quien mantiene a todas las redes juntas es el lenguaje que toda computadora conectada a Internet debe hablar y entender, para hacer posible la comunicación entre hosts. IP es un protocolo de red que provee la mejor manera para transportar piezas de información llamadas datagramas, de un ordenador a otro remoto, y viceversa, sin importar si estas máquinas se encuentran en la misma red o no.

La figura 6 muestra los encabezados de los mensajes. El primer campo tanto en IPv4 como en IPv6 es el campo de versión. Este mantiene información sobre la versión de protocolo al que el datagrama pertenece, haciendo posible la transición entre las versiones de IP.

Figura 6. **Encabezado IP**



Fuente: CISCO, The Internet Protocol Journal - Volume 9, Number 3. http://www.cisco.com/web/about/ac123/ac147/archived_issues/ipj_9-3/ipv6_internals.html. Consulta: 17 de mayo de 2012.

El siguiente campo en IPv4 es IHL (Internet Header Length), necesario solamente en IPv4, ya que la longitud de encabezado en IPv4 puede ser variable, mientras que en IPv6 el encabezado tiene un tamaño fijo. IPv4 puede tener *options* en su encabezado y es la razón de que exista el campo IHL. Este campo de opciones permite que el transmisor del datagrama especifique la ruta que éste debe seguir, permitiendo crear un *log* de las rutas transitadas en el trayecto.

Sin embargo, el espacio para este campo es insuficiente, solamente 40 bytes. Por otra parte, la mayoría de los *routers* ignoran estas opciones, así que la conclusión es que al realizar un chequeo de la longitud que posee el

encabezado, solamente hace más lento el procesamiento del datagrama. Esta es la razón por la que los encabezados en IPv6 son de un tamaño fijo de 40 *bytes*.

Sin embargo, para proveer posibilidades de extensión IPv6 posee el campo Next Header. El encabezado en IPv6 fue simplificado a su máximo, pero para proveer mayor capacidad al protocolo en caso de ser necesario, se agregan encabezados para extensión adicionales. Estos se localizan luego del encabezado principal y se identifican como Next Header. Se pueden agregar varias extensiones en el mismo datagrama, conectados en cadena. La última extensión se establece como el identificador de la carga de protocolo utilizado por el datagrama de IPv6, este identificador en IPv4 se localiza en el campo Protocol.

El *byte* utilizado por los campos DSCP (Differentiated Services Codepoint) y ECN (Explicit Congestion Notification), ha sido un *byte* inestable. Se definió en un inicio para IPv4 como campo de TOS (Type of Service). Los primeros tres *bits* eran llamados Precedence, el cual indicaba la importancia de la información que transportaba el datagrama en IPv4. Los siguientes tres *bits* eran las banderas**,** D (Delay), T (Throughput), R (Reliability), permitiendo que el *host* especificará por cuál de ellas era de mayor importancia para él. Los últimos dos *bits* eran establecidos a cero. Pero en la práctica, los *routers* ignoraban el campo TOS por completo.

Luego, el séptimo *bit* del campo TOS fue convertido a *bit* de bandera, el cual se utilizaba para indicar la preferencia por bajo costo monetario. Aun así, el campo TOS no era tomado en cuenta por los *routers*. Similarmente, en las primeras especificaciones de IPv6 los *bits* 4-7 del encabezado principal forman el campo denominado Priority, que tiene un uso similar al campo Precedence

en IPv4. En la última especificación de IPv6, este campo fue expandido hasta ocupar un *byte* completo denominado Traffic Class, pero su uso no estaba definido.

Aprovechando los constantes cambios, otra modificación fue realizada a este *byte* en ambas versiones de IP. Los primeros 6 *bits* se convirtieron el campo DSCP, y los últimos dos *bits* se volvieron el campo ECN.

Estos últimos dos *bits* son usados para evitar utilizar la pérdida de paquetes en los *routers*, como único indicador de congestión. En cambio, si un *router* tiene congestión utiliza este campo para indicar esto. El receptor de la información utiliza los acuse de recibido del protocolo de transporte, para indicar al transmisor sobre la situación de congestión, quien disminuirá su tasa de envío.

El campo Total Length en IPv4 cuenta la longitud total del datagrama. Es el equivalente al campo Payload Length en el encabezado de IPv6, el cual no incluye la longitud del encabezado principal. En teoría, los datagramas de mayor tamaño pueden alcanzar un tamaño de 64 *Kbytes*, pero en la práctica estos llegan a alcanzar alrededor de 1,5 *Kbytes* debido a las limitaciones de la capa de enlace de datos.

El campo Identification es utilizado cuando un datagrama es fragmentado en la red. Todos los fragmentos del datagrama transportan el mismo valor de identificación. Este campo, junto al campo Fragment Offset, indican la posición del fragmento dentro del datagrama original, y la bandera MF (More Fragments) marca el final de un datagrama fragmentado, haciendo posible su reconstrucción en el lado del receptor.

La bandera DF (Don't Fragment) representa un imperativo para los *routers* de no fragmentar el datagrama, porque el lado receptor es incapaz de volverlo a armar. En IPv6 no existe un campo en su encabezado principal que proporcione capacidades de fragmentación, simplemente porque los *routers* no están autorizados a fragmentar los datagramas.

En cambio, existe un encabezado de fragmentación que puede ser utilizado únicamente por las partes transmisoras y receptoras, pero no por los *routers* de la red. Los *routers* simplemente descartan cualquier datagrama que sea demasiado grande, devolviendo un código de error, lo cual simplifica el trabajo.

El campo Time to Live evita que se tengan datagramas sin entregarse por tablas de ruteo corruptas. Inicialmente fue diseñado para contar segundos, pero en la práctica este campo se disminuye en uno por cada *router*, así que únicamente cuenta saltos. En IPv6 el campo Hop Limit realiza el mismo trabajo.

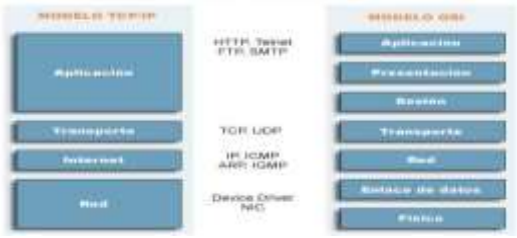
El campo Header Checksum protege únicamente al encabezado. Se asume que si el usuario desea proteger la información transportada por IP éste proveerá su propio esquema de detección de errores. Sin embargo, debido que el campo *Time to Live* varía en cada salto, el *checksum* debe ser recalculado en cada *router*. IPv6 no incluye ninguna verificación (*cheksum*) en su encabezado principal, pero existen dos encabezados de extensión disponibles para superar este problema. El campo AH (Authentication Header) provee integridad y autenticación de los datos de origen. El campo ESP (Encapsulating Security Payload) provee encriptación.

Los campos de Source Address y Destination Address presentes en ambas versiones de IP indican el origen y destino del datagrama. Mientras que

las direcciones de IPv4 son de longitud de 4 *bytes*, en IPv6 estas son de 16 *bytes*. La propuesta original de Deering contemplaba direcciones de longitud de 8 *bytes*, pero eso podría terminar en que IPv6 tuviera el mismo problema que IPv4 en unos años, por lo que la decisión final fue proveer direcciones de 16 *bytes*. Para tener una idea de la cantidad de direcciones, se puede calcular que si éstas fueran distribuidas equitativamente sobre la tierra, se tendrían aproximadamente 6,7 x 10^{23} direcciones por metro cuadrado. No obstante, parte de esos 16 *bytes* contiene información de rutas, permitiendo a los *routers* trabajar más eficientemente en una red de esas dimensiones.

No obstante, aun cuando IP es el protocolo más importante en el internet, debido a que hace posible la comunicación entre redes remotas, no es más que una parte del set de protocolos que utiliza internet. La figura 7 muestra algunos de estos protocolos y su relación con el modelo de referencia OSI.

Figura 7. **Algunos miembros del set de protocolo de internet**



Fuente: http://www.textoscientificos.com/redes/tcp-ip/comparacion-modelo-osi. Consulta: 18 de mayo de 2012.

A los niveles más bajos de protocolos, el internet puede utilizar cualquiera de las tecnologías disponibles, desde Token Ring o Token Bus, hasta redes de área local (LANs) de alto desempeño como FDDI. Sin embargo, una de las LAN de mayor uso es Ethernet. En la arquitectura de internet (modelo TCP/IP), tanto

la capa física como la de enlace de datos, están unidas a un mismo nivel, que es normalmente denominada como la capa de Host to Network.

A un nivel superior se tiene al protocolo IP intentando hacer una sola red de una gran cantidad de LANs. En la arquitectura de internet (modelo TCP/IP), IP funciona al nivel de internet, el cual es el equivalente a la capa de red del modelo OSI. Luego, sobre IP tenemos el nivel de *transport*, en el cual usualmente se tienen dos protocolos: UDP (User Data Protocol) y TCP (Transmission Control Protocol).

UDP es muy simple y es utilizado en aplicaciones no orientadas a conexión. TCP es un protocolo orientado a conexión que provee fiabilidad y control de congestión. Han sido los protocolos de transporte más utilizados de los últimos años, pero hoy en día también se cuenta con SCTP (Stream Control Transmission Protocol), muy cercano a TCP en la funcionalidad que provee, pero mejorado. SCTP se tratará más adelante en este documento.

Existen también protocolos de control, que actúan directamente sobre IP. Por ejemplo: está ICMP (Internet Control Message Protocol), el cual funciona para IPv4 e IPv6, es utilizado para reportar errores a nivel de IP, también protocolos de ruteo como OSPF (Open Shortest Path First), utilizado para calcular las tablas de ruteo.

En el modelo de internet, no se cuenta con las capas de sesión y presentación. En su lugar, directamente sobre la capa de Transporte se tiene la placa de aplicación (*Application*). Los protocolos mostrados en la figura 7 son solamente una porción del total.

Aparte de Telnet utilizado para conexión remota, FTP utilizado para transferencia de archivos, HTTP utilizado para la transferencia de archivos de hipertexto y SMTP (Simple Mail Transport Protocol) utilizado para el transporte de e-mail, los cuales se muestran en la figura, existen protocolos como NFS (Network File System), utilizado para proveer acceso transparente a archivos por parte de aplicaciones clientes, o SNMP (Simple Network Management Protocol) utilizado para manejar sistemas remotos a través de la red. Muchos de esos protocolos pueden utilizar tanto UDP como TCP, pero siempre uno de ellos es preferido.

Sobre todo, IP provee de una red de conmutación de paquetes muy flexible, que es capaz de transportar una multitud de protocolos diseñados para la más amplia gama de usos.

2. CONVERGENCIA: UNA DEMANDA ACTUAL

SS7 e internet han sido dos redes independientes que realizan distintas tareas y proveen servicios distintos. SS7 es utilizada para señalización de telefonía e internet y para transferencia de data en redes de conmutación de paquetes. Sin embargo, en los últimos años los servicios que ofrecen se están fusionando. Por un lado, los usuarios de telefonía demandan servicios que involucran el acceso a internet, y por el otro, los desarrollos en los servicios de IP ya pueden proveer ciertos niveles de calidad. Esto hace de IP más confiable para el transporte de información más delicada, como lo es voz y señalización telefónica.

2.1. Voz sobre IP

Durante los últimos años VoIP (Voice over IP) se ha convertido en un tema de moda. Su influencia en las llamadas telefónicas globales se ha incrementado, así como su popularidad. Esta popularidad está basada en que VoIP hace un mejor uso de los recursos utilizados para transmitir voz, y por ende, los precios que las compañías ofrecen por telefonía IP pueden ser menores, especialmente para llamadas de larga distancia. Sin embargo, su mayor problema es que no puede ofrecer los mismos niveles de calidad que PSTN.

Telefonía IP al contrario**,** no realiza ninguna reserva exclusiva de ningún recurso sino que maneja la llamada sobre la red como cualquier otra fuente de información. Normalmente, un usuario en telefonía IP digita un número gratis y se conecta a una puerta de enlace IP (IP Telephony Gateway), enviando

también la información necesaria como su número de cuenta y el número de destino.

Estas puertas de enlace conectan la red de telefonía pública y la red IP proveyendo el servicio y están también a cargo de recibir los datos, comprimir la voz y transportarla sobre la red IP pública o privada hasta la puerta de enlace del receptor. Este *Gateway* conectará la llamada a la red de telefonía local, descomprimiendo los paquetes IP y enviando el flujo de voz al suscriptor deseado.

La mayor diferencia entre estos dos esquemas es que el transporte de larga distancia es sustituido por una red IP. Así que, una llamada de larga distancia se convierte en dos llamadas locales más un transporte IP de larga distancia. Por lo tanto, un ITSP (IP Telephony Service Provider) puede ofrecer mejores precios a sus clientes.

Los costos de transportar la voz utilizando una red IP pueden ser más bajos que los de un transporte de llamada de larga distancia tradicional, ya que las instalaciones son compartidas entre todos los usuarios y no existen los canales dedicados. Si se tiene un circuito dedicado *full-duplex* para transmitir una conversación telefónica se hace un muy pobre uso del mismo, ya que la mayoría del tiempo por lo menos una de las partes estará en silencio y el canal permanece sin uso durante ese tiempo.

Inicialmente la telefonía sobre internet significaba la existencia de algún software que era capaz de establecer una llamada entre computadoras conectadas a internet**,** estas llamadas eran gratis**,** pero ofrecían muy baja calidad al inicio. El primero de dichos programas apareció en 1995, y desde entonces comenzaron los intentos de interconectar IP y PSTN. En 1997 fue

lanzado el primer servicio teléfono-a-teléfono. Hoy en día, muchos ITSP ofrecen llamadas de larga distancia.

Sin embargo, PSTN e internet tienen características muy distintas que hacen del transporte de voz sobre IP una tarea muy difícil. Algunas de estas diferencias se enlistan en la tabla I.

Tabla I. **Diferencia entre redes de telefonía e IP**

PSTN	VoIP
Líneas (canales dedicados).	Todos los canales son transportados sobre una conexión a internet.
Cada canal es de 64 Kbps (en cada dirección).	Compresión puede resultar en canales de 10 Kbps (en cada dirección).
Características como *call waiting*, *caller ID*, etc., usualmente disponibles bajo costos extras.	Características como *call waiting*, caller ID, etc., usualmente disponibles sin costo adicional.
Puede mejorarse o expandirse con nuevo equipo y canales.	Mejoras usualmente solo requieren mayor ancho de banda y mejoras de software.
Larga distancia se cobra usualmente por minuto o en algún paquete de suscripción.	Larga distancia es usualmente incluida en el precio mensual.

Fuente: The Difference between VoIP and PSTN Systems.
http://www.webopedia.com/DidYouKnow/internet/2008/voip_pots_difference_between.asp.
Consulta: 20 de mayo de 2012.

Una de las mayores diferencias que tienen es la QoS (Quality of Service) que proveen. Mientras que PSTN ha sido diseñada para ser una red altamente confiable**,** en la cual un paquete es raramente perdido o retrasado**;** internet por

el contrario es una red de mejor-esfuerzo**,** donde una mayor cantidad de paquetes se pierden**,** (de hecho TCP necesita la pérdida de paquetes como retroalimentación para proveer control de congestión) y en la que no se provee un límite para los retrasos, lo cual puede dañar severamente la calidad de la voz.

El QoS ofrecido por VoIP es altamente dependiente de la congestión que exista en la red, degradándose conforme el ancho de banda se disminuye. Este problema puede enfrentarse con el simple hecho de agregar ancho de banda, pero esto sería una solución temporal. Por lo que es necesario el desarrollo de mecanismos de red más apropiados para garantizar el ancho de banda necesario para servicios como VoIP**,** y ayudar a las compañías a minimizar sus costos mientras aun cumplen con niveles satisfactorios de QoS.

Sin embargo, el futuro en este aspecto se ve prometedor. La IETF ha desarrollado varias tecnologías para agregar QoS a las redes IP que pueden mejorar los problemas originados por Telefonía IP y su transporte en tiempo real. Entre estos esfuerzos se pueden resaltar los siguientes:

- RSVP (Resource Reservation Protocol) es utilizado por los ordenadores para solicitar un QoS específico de la red para su flujo de data. Los *routers* utilizan RSVP para comunicar solicitudes de QoS a todos los nodos en la ruta del flujo, y para establecer y mantener el nivel. Las solicitudes RSVP por lo regular significan la reserva de recursos en cada nodo a lo largo de la ruta a seguir por la data. El nivel deseado de QoS se consigue al reservar los recursos de antemano.

- RAP (Resource Allocation Protocol) es un protocolo utilizado por *routers* con capacidad de RSVP para comunicarse con servidores de políticas

que se encuentren en la red. Cuando no existan suficientes recursos para satisfacer todas las peticiones RSVP, los servidores de políticas son los que determinan a quien se le concederán los recursos de la red y que solicitudes tendrán prioridad.

- COPS (Common Open Policy Service) es el protocolo base para comunicar la información de las políticas entre los servidores de políticas y los *routers* que se encuentren en el marco de RAP.

- RTP (Real Time Protocol) es un protocolo especialmente diseñado para transportar información en tiempo real. Opera sobre UDP y puede ser utilizado en media *on-demand* o VoIP. Consiste de dos partes: la parte de data, denominada RTP Data Transfer Protocol, es un protocolo que provee reconstrucción de tiempos, detección de pérdidas, seguridad e identificación de contenido. La parte de control, denominada RTP Control Protocol (RTCP), chequea la calidad de la transmisión y controla el estado de los participantes. RTP por sí mismo no provee ningún mecanismo para asegurar la entrega a tiempo o proveer garantías de QoS, pero depende de servicios de capas inferiores como RSVP para lograrlo.

- RTSP (Real Time Streaming Protocol) es una extensión de control para RTP. Añade funciones como retroceder, adelantar y pausa.

- SIP (Session Initiation Protocol) es un protocolo de control de la capa de aplicación que se puede utilizar para establecer, manejar y terminar sesiones multimedia (incluyendo VoIP). SIP puede utilizarse para establecer sesiones multi-punto y las puertas de enlace en la telefonía de internet que se conectan con PSTN pueden usar SIP para establecer

llamadas entre ellos. Éste también define nuevos tipos de URL (Uniform Resource Locators) que ayudan a traducir números de teléfono a direcciones IP y viceversa.

- SDP (Session Description Protocol) definido con el propósito de describir las sesiones multimedia para que sea posible anunciar sesiones, invitar a sesiones y otros medios de iniciar sesiones multimedia. El anuncio de la sesión en sí**,** es realizado utilizando SAP (Session Announcement Protocol) el cual realiza un *multicast* del anuncio el cual contiene la descripción de la sesión.

- DiffServ (Differentiated Services) permite a los proveedores de servicio clasificar los paquetes con varias prioridades utilizando el campo DSCP del encabezado IP. Se espera que los *routers* existentes en la red reconozcan dichas prioridades y concedan a los paquetes ciertos privilegios de rendimiento de acuerdo a estas prioridades.

- MPLS (Multiprotocol Label Switching Architecture) en esencia**,** éste impone algún tipo de conmutación de circuitos a una red IP. Los paquetes pueden ser agrupados al etiquetarlos con una sentencia en común que permite el paso acelerado a través de la red. Las etiquetas no solo informan a los *routers* sobre el QoS que se debe cumplir**,** sino que sustituyen las decisiones de ruteo, la cual se debe realizar una sola vez y se aplica al grupo de paquetes etiquetados.

Sin embargo, una vez que se consigue una calidad razonable para el habla transportada sobre IP existen otras tareas que se deben solventar. Uno de los más grandes problemas relacionados con la telefonía IP han sido sus limitantes para operar con PSTN. Los *gateways* de VoIP son capaces de

proveer medios para el transporte de la trama de voz, pero muchos de los servicios provistos por PSTN provienen de la red de señalización que utiliza: la red SS7.

Las funcionalidades provistas por SS7 para los operadores incluyen una amplia gama de características, desde la identificación de la llamada hasta características más complicadas basadas en IN. Solo cuando una adecuada interoperabilidad entre redes IP y SS7 sea provista, los servicios VoIP serán altamente aceptados por los consumidores. Como un sencillo ejemplo, sin una completa interconexión con SS7, los ITSPs deben continuar con sus prácticas de marcación en varios estados (el suscriptor debe primero marcar al *gateway*, luego su ID de cliente y finalmente el número al que desea contactar).

Por otra parte, una verdadera integración de las redes de voz (SS7) y datos (IP) introduciría los beneficios de VoIP sobre multimedia y aplicaciones de varios servicios, algo con lo que la telefonía de hoy no puede competir. Más allá de reemplazar la conmutación de circuitos, VoIP tiene el potencial de hacer los servicios de telefonía tan flexibles y programables como el email y servicios web, acelerar la disponibilidad de las comunicaciones multimedia, así como integrar los servicios de telefonía con los servicios existentes de internet.

Hoy en día uno de los mayores intereses en VoIP proviene de sus bajos costos en llamadas de larga distancia. Sin embargo, en el futuro, los beneficios reales de VoIP vendrán de la posibilidad de ofrecer estos nuevos servicios. Una integración de voz y datos permitirá mayor estandarización y reducirá el total de equipos necesitados. También el tener solamente una red para voz y datos hace su administración más sencilla.

2.2. Telefonía de tercera generación

El primer servicio de telefonía móvil fue provisto en los USA en los años 40 y a inicios de los años 50 en Europa. Eran teléfonos análogos con restricciones en cuanto a cobertura y número de suscriptores. Eran pesados, caros, expuestos a interferencias, con un alto consumo de energía y pobre calidad de voz. A inicios de los años 1980s existía alrededor de un millón de suscriptores en todo el mundo.

A finales de los años 70 e inicios de 80**,** la introducción de los sistemas celulares era un salto enorme en las comunicaciones móviles. Gracias a los semiconductores y microprocesadores**,** nuevos teléfonos más livianos, pequeños y más sofisticados se hicieron realidad.

Estos primeros sistemas celulares que eran capaces de transmitir únicamente voz análoga fueron conocidos como telefonía móvil de primera generación (1G). Entre las más importantes se tiene AMPS (Advanced Mobile Phone System) en América, parte de Europa y Rusia, Australia y parte de Asia; NMT (Nordic Mobile Telephone) en los países nórdicos, y TACS (Total Access Communication System), en Inglaterra. Existían alrededor de 20 millones de suscriptores de 1G para 1990.

La telefonía móvil de segunda generación (2G) es la que mayormente se utiliza hoy en día. Es digital, por lo que provee nuevos servicios como fax, mensajes cortos y transmisión de datos, aun con posibilidad de encriptación. Por otra parte, provee servicios móviles avanzados (*roaming*), que hacen posible que los clientes tengan movilidad hacia áreas donde operan distintas compañías telefónicas manteniendo el servicio.

El sistema más exitoso en los estándar de telefonía 2G es GSM (Global System for Mobile Communications) alcanzando aproximadamente 170 países alrededor del mundo**,** otras versiones importantes de 2G son CDMA (Code Division Multiple Access), mayormente utilizada en la región de Asia del Pacífico; TDMA (Time Division Multiple Access), muy utilizada en USA; y PDC (Personal Digital Cellular), para clientes en Japón.

Sin embargo, las redes 2G aun cuentan con ciertas incompatibilidades. Existen varios estándar que no son compatibles haciendo a las terminales móviles inservibles en áreas o países con distinta tecnología; la tasa de transmisión de datos (9,6 Kbps en GSM) es demasiado lenta y la calidad de la voz es buena pero podría ser mejorada; y hoy en día el suscriptor demanda nuevos servicios como aplicaciones multimedia**,** los cuales no encuentran un lugar en la red 2G.

Mientras que se desarrollan teléfonos de nueva generación, algunas tecnologías se han agregado a las redes GSM, como HSCSD (High Speed Circuit Switched Data) que provee hasta 57,2 Kbps al abrir varios canales de comunicación. GPRS (General Packet Radio Service) que permite una transferencia de datos de hasta 171,2 Kbps utilizando paquetes IP. EDGE (Enhanced Data for GSM Evolution) con dos versiones basadas en HSCSD y GPRS respectivamente, que debe proveer hasta 384 Kbps; o AMR (Adaptive Multi Rate) para optimizar la calidad de la voz. GSM, junto a estas tecnologías es conocida como *2G+ o 2.5G*.

En este ambiente, la telefonía móvil de tercera generación (3G) está siendo especificada por el proyecto mundial 3GPP (3G Partnership Project), con el objetivo de lograr una red global móvil con completa compatibilidad.

Las redes 3G colectivamente conocidas como IMT-2000 (International Mobile Telephony 2000) y UMTS (Universal Mobile Telecommunication System)**,** empezó a ser diseñada a mediados de los años 1990s y la primera operación aparece aproximadamente en el 2002.

UMTS es retrocompatible con GSM y también utiliza SS7 para señalización. Sin embargo, UMTS pone mayor énfasis en la conmutación de paquetes que GSM, usándola no solo para señalización sino también para datos. UMTS ofrece hasta 1 920 Mbps para ser utilizados por aplicaciones multimedia.

En la primera edición de especificaciones de UMTS, ATM (Asynchronous Transfer Mode) fue la red de conmutación de paquetes escogida. Esto fue debido a que provee la seguridad de obtener el QoS necesario, y también porque su espacio para direccionamiento era lo suficientemente grande.

Sin embargo, los últimos avances sobre el QoS en IP y el desarrollo de IPv6 con un rango de direccionamiento más amplio están cambiando las cosas. Sobre todo, la mayoría de los servidores que contienen la información que será transmitida a los usuarios UMTS se espera que se localicen sobre internet. Por lo que hace sentido el uso de IP para transferir información a las terminales que se esperan sean IPv6 a través de la red UMTS. Además, las redes IP son de mucho menor costo.

Una vez que se tenga una red IP transportando la información se desea que la señalización sea basada en IP para poder utilizar la misma red para ambos propósitos. Ha habido dos emisiones más de UMTS, donde el papel de IP es mucho más protagónico. En la emisión 3GPP R5, también conocida como *All IP release*, la red de transporte utiliza IP. Éste y los protocolos adyacentes

serán utilizados en el control de la red de igual manera y el flujo de información del usuario se espera que sea todo sobre IP.

En otras palabras, las redes móviles implementadas bajo la especificación 3GPP R5 serán redes de conmutación de paquetes utilizando IP como protocolo de transporte**,** en lugar de SS7. Pero la red IP aún debe soportar conmutación de circuitos, y UMTS debe ser compatible con GSM. Esto significa que aún se necesita un medio para utilizar el protocolo SS7 en nuestra red IP. Como resultado de esto, las nuevas redes 3G necesitan un medio para transportar mensajes SS7 sobre IP.

2.3. La necesidad de un nuevo protocolo de transporte

Mientras que el grupo de trabajo de MMUSIC (Multiparty Multimedia Session Control) de la IETF estaban a cargo de proveer lo medios necesarios para mejorar la capacidad de QoS de las redes IP, un nuevo grupo de trabajo se fundó el 23 de noviembre de 1998, SIGTRAN (Signaling Transport), con la misión de direccionar el transporte de la señalización PSTN sobre redes IP. La manera en que se enfrentaba dicha misión era mantener tanto las pilas de SS7 y de IP y definir una interfaz que hiciera posible el transporte tanto de voz como de señalización SS7 a través de redes IP.

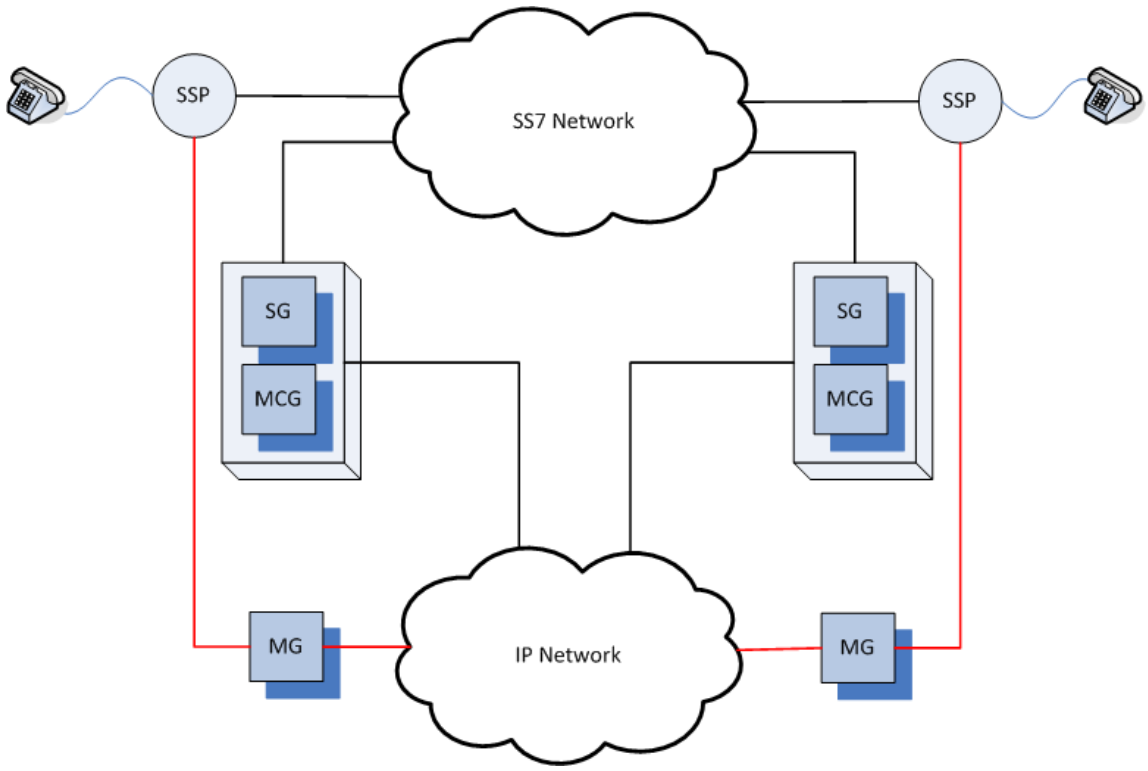
Los mensajes de señalización tienen requisitos sobre retraso y pérdida bastantes estrictos, así como seguridad y flexibilidad, debido a esto se diseñó una arquitectura como la que se muestra en la figura 8. En dicha figura se pueden identificar los *gateways* que conectan las redes SS7 e IP, las líneas color rojo representan canales de voz y las líneas color negro representan *links* de señalización, se pueden también identificar los siguientes tres elementos:

- MG (Media Gateway): termina el flujo de SS7, empaqueta la voz y entrega los paquetes a la red IP. Del lado del receptor realiza la función inversa.

- SG (Signaling Gateway): es el agente que recibe la señalización SS7, la traduce y envía a la red IP, y viceversa.

- MGC (Media Gateway Controller): maneja el registro y administración de los recursos utilizados por el MG, con la posibilidad de autorizar el uso de recursos en base a políticas locales.

Figura 8. **Modelo funcional de SIGTRAN**



Fuente: elaboración propia, con Microsoft Office Visio 2007.

Los *gateways* SS7-IP no solamente traducen y transportan SS7 a través de las redes IP, sino pueden también recibir mensajes de administración direccionados a ellos. Proveen transporte transparente sobre las redes IP para los protocolos de señalización basados en mensajes. De este modo, tanto datos como señalización pueden navegar la red IP y alcanzar su destino, proveyendo el mismo tipo de servicios que ofrece PSTN, haciendo al mismo tiempo un mejor uso de la red que transporta los datos de voz.

Antes que la arquitectura estuviera completamente definida, el personal de SIGTRAN inició determinando los protocolos que utilizarían para proveer la traducción de los mensajes SS7. Obviamente se necesitaba un protocolo de transporte**,** pero no existía ningún consenso sobre cual se debía de utilizar, así que se definió solamente como CTP (Common Transport Protocol). Existió un intento de no complicar el trabajo y utilizar TCP o UDP. La funcionalidad esperada de CTP era:

- Transporte de la variedad de protocolos SCN (Switched Circuit Network) como MTP3, ISUP, SCCP, TCAP, con la habilidad de proveer una manera de identificar el protocolo específico que se está transportando.

- Proveer un protocolo en común que defina formatos de encabezados, extensiones de seguridad y procedimientos para el transporte de señalización, y soportar extensiones para agregar protocolos SCN específicos si se necesitara.

- Junto con IP, proveer la funcionalidad esencial como está definida por la capa inferior SCN. Esta funcionalidad puede incluir:
    - Control de flujo.

- Entrega secuencial de los mensajes de señalización dentro de un mismo flujo de control.
- Detección de error.
- Recuperación de los componentes en una ruta en caso de falla.
- Retransmisión y otras correcciones de errores.
- Detección de nodos no disponibles.

- Soportar multiplexión de varias sesiones SCN de capas superiores dentro de una misma sesión de transporte. El protocolo debe incluir la capacidad de evitar el bloqueo de la entrega de mensajes de un flujo de control debido a errores secuenciales de otro flujo de control.

- Tener la capacidad de transportar mensajes de mayor longitud que los especificados por las limitaciones de segmentación/reconstrucción de SCN.

- Permitir esquemas robustos de seguridad para proteger la información de señalización que se transporta. El transporte de la señalización debe ser capaz de operar sobre sesiones *proxy* y sobre *firewall*.

- Proveer medios para evitar congestión y reacción en redes congestionadas.

Por lo tanto UDP no fue ni considerado, y en un inicio TCP fue el candidato para convertirse en CTP. Sin embargo, luego de una investigación se determinó que TCP poseía ciertas deficiencias al proveer transporte de señalización PSTN sobre IP, entre éstas, se pueden identificar:

- TCP es un protocolo de transporte que provee transferencia de data confiable y un estricto orden de envío de la información. Esto puede ser lo que normalmente se requiere, pero existen aplicaciones que necesitan la confiabilidad en la transferencia, pero no administración secuencial, o solamente un ordenamiento parcial de la información. Cualquier aplicación con estas necesidades sufriría del bloqueo HOL (Head Of Line) que produce TCP**,** este bloqueo es debido a que TCP maneja los mensajes como una misma cadena de bytes, así que**,** si se utiliza TCP para enviar mensajes sin relación, los entregará en el mismo orden que fueron enviados, y si un datagrama se pierde afectará los mensajes siguientes, los que serán retenidos hasta que arribe el mensaje faltante, esto se denomina bloqueo HOL, causando un retraso innecesario e indeseable.

- TCP es orientado a flujo, y esto también puede ser un inconveniente para algunas aplicaciones, las cuales usualmente incluyen sus propias marcas dentro del flujo para lograr identificar el inicio y fin del mensaje.

- TCP nunca fue diseñado para *multihomed* (nodo que por lo regular posee varias tarjetas de red y puede hacer uso de cualquier número de IPs en cualquier momento).

- TCP no realiza un buen escalamiento, ya que el número máximo de conexiones simultáneas de TCP depende de las limitaciones de *kernel*. Por esto es que generalmente TCP se implementa a nivel de sistema operativo.

- En TCP no existe la posibilidad de control de tiempo**,** pues no provee control de la aplicación respecto a tiempos de inicialización, terminación y retransmisión.

- TCP es bastante vulnerable a ataques de negación de servicio. Este tipo de ataques intentan hacer un servicio indisponible, comúnmente se realizan intentando agotar los recursos que utiliza el servicio.

El transporte de señalización PSTN a través de redes IP es el tipo de aplicación que hace que las limitaciones de TCP se hagan relevantes. Existió un intento inicial para modificar y mejorar TCP para que cumpliera dichos requisitos. Sin embargo, esto fue descartado debido a que ya existían estudios sobre la dificultad que conllevaba modificar TCP.

Por lo tanto, se decidió diseñar un nuevo protocolo de transporte que operara sobre UDP. Aparte de las funcionalidades que eran esperadas en CTP, se identificaron otras características deseadas:

- Habilidad para descubrir el MTU (Maximum Transfer Unit) de la ruta utilizada desde la IP de envío**,** hasta la IP de recepción y la posibilidad de fragmentar la información para cumplir con el MTU descubierto.

- Posibilidad de enviar información sobre múltiples flujos dentro de la misma asociación. Entrega secuencial de los mensajes enviados en el mismo flujo, y la posibilidad de entrega en orden de arribo de mensajes individuales.

- Posibilidad de unir múltiples mensajes en un mismo paquete.

Con estos objetivos, el personal de SIGTRAN inició el trabajo de diseño de un nuevo protocolo que superara los problemas encontrados en TCP.

2.4. Propuesta que no se podía rechazar

A finales de 1998 se presentaron varias propuestas de protocolos que cumplían los requisitos expuestos de una manera parcial o total. Uno de ellos se llamó RUDP (Reliable UDP), el cual soportaba acuse de recibido y retransmisión, pero no soportaba *multihoming*, ni evitaba congestionamiento.

Otra propuesta fue T/UDP (UDP for TCAP) que incluyo control de flujo y transferencia de datos confiable, pero igualmente fue descartado. Otro protocolo con características similares fue SSTP (Simple SCCP Tunneling Protocol), éste era capaz de funcionar sobre UDP o TCP, pero también fue descartado. El protocolo PURDET fue otra opción utilizando UDP y soporte secuencial, control de flujo, identificación de protocolo, retransmisión en caso de error y detección de pérdidas de *links*, fue también descartado.

Sin embargo, existió una propuesta denominada MDTP (Multi-Network Datagram Transmission Protocol), que atrajo la atención del grupo de trabajo de SIGTRAN. En 1997 se inició el diseño de MDTP de parte del trabajo del grupo SIGTRAN, como una solución para las debilidades de TCP.

En su diseño preliminar MTPD era un protocolo a nivel de aplicación trabajando sobre UDP**,** que además cumplía con la mayoría de los requisitos impuestos por SIGTRAN para CTP. Esta propuesta era la única que soportaba *multihoming* y que evitaba el bloqueo HOL. Estas eran buenas razones para escoger MTPD**,** durante los siguientes meses se trabajó en su desarrollo y se utilizó como base para SCTP (Simple Control Transport Protocol).

El cambio de MTPD a SCTP no solamente involucró un cambio de nombre, sino una profunda revisión de protocolo. Fue ahí cuando el encabezado y su estructura interna fueron modificados casi completamente convirtiéndose en altamente extensible; se adoptó el mecanismo de *cookie* en la iniciación para evitar los ataques de negación de servicios, se implementaron las características de control de congestión de TCP y se modificaron otras características como negociación de flujo, agrupación de mensajes y fragmentación de información.

Para el 2000 se introdujo otro gran cambio, el grupo de trabajo revisó la pila del protocolo e implementó que éste funcionara directamente sobre IP. Este cambio fue una noticia polémica**,** ya que implícitamente significaba que SCTP se debía implementar dentro del *kernel* del sistema operativo, a pesar que las implementaciones de SCTP no iniciarían hasta después de varios años, y teniendo que esperar a que los vendedores de sistemas operativos lo tuvieran disponible en sus productos.

Por otra parte, teniendo SCTP dentro del *kernel* haría mucho más difícil tener el control sobre valores de tiempos y parámetros que permitían adaptar SCTP a distintos ambientes. Sin embargo**,** los beneficios de ubicar SCTP en el lugar adecuado en la pila de IP se sobrepusieron a estos problemas. Moviendo a SCTP sobre IP y teniendo éste su propio puerto, abrió una vía para que se convirtiera en un protocolo de transporte importante, al mismo nivel que TCP, haciendo de SCTP de amplio uso para otras aplicaciones y no solamente para transporte de señalización telefónica.

SCTP fue diseñado para ser un protocolo de transporte para la señalización telefónica**,** no fue la idea original diseñar un protocolo que pudiera competir con TCP. En su largo período de diseño, muchas características

fueron agregadas a la idea original, muchas de ellas intentando resolver problemas que se habían identificado cuando se utilizaba TCP/IP, aun cuando no eran de vital importancia para el transporte de PSTN sobre IP.

SCTP puede por lo tanto ser utilizado como el principal protocolo de transporte en internet sustituyendo a TCP en el futuro. Es por ello que se puede pensar que SCTP es una versión renovada de TCP que incluye capacidades de extensión, y que utilizado en conjunto con IPv6, se espera que cambien en el futuro la manera es de enviar y recibir información sobre internet.

3. DISEÑO DE SCTP

Durante su período de diseño las características de SCTP se incrementaron, la especificación final de SCTP es el resultado de un trabajo conjunto de muchos especialistas de distintas aéreas, desde control de protección hasta enrutamiento IP.

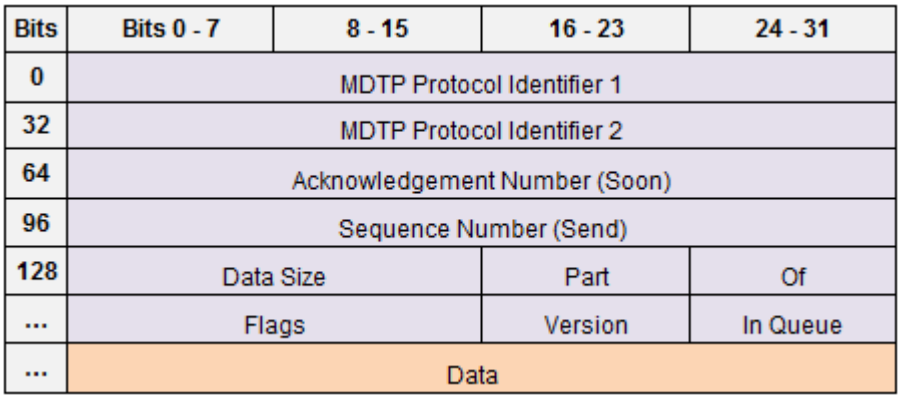
3.1. Datagrama SCTP

La estructura interna de los datagramas de SCTP ha cambiado por completo desde la primera versión de MDTP (Multi-Network Datagram Transmission Protocol). Sus características han sido mejoradas y muchos errores se han solventado. Sin embargo**,** las ideas básicas aún se mantienen y el diseño final interno de SCTP es mucho más parecido a MDTP que lo que se esperaba en un inicio. Ya que SCTP es una evolución de MDTP, se tratará en primer lugar los encabezados de MDTP y su estructura interna como lo fue en la primera versión. Luego se discutirá el diseño final de SCTP.

3.2. Encabezado y estructura interna de MDTP

Cuando se presentó la primera versión de MDTP en agosto de 1998, la estructura de su datagrama era muy similar a la de TCP la cual se puede observar en la figura 9.

Figura 9. **Estructura del datagrama MDTP**



Bits	Bits 0 - 7	8 - 15	16 - 23	24 - 31
0	MDTP Protocol Identifier 1			
32	MDTP Protocol Identifier 2			
64	Acknowledgement Number (Soon)			
96	Sequence Number (Send)			
128	Data Size		Part	Of
...	Flags		Version	In Queue
...	Data			

Fuente: elaboración propia, con Microsoft Office Excel 2007.

TCP es hoy en día el protocolo de transporte más exitoso utilizado en internet, y por lo tanto, era la mejor referencia sobre la cual basarse. Luego, con la evolución de MDTP y SCTP, las similitudes con TCP permanecieron más en su comportamiento interno que en su estructura externa:

Todo datagrama MDTP tiene un encabezado de 8 *bytes* que contiene el identificador propio de MDTP, mientras que TCP no posee algo parecido. La razón para esto es que MDTP fue diseñado como un protocolo de aplicación corriendo sobre UDP, y luego el encabezado IP que identifica el protocolo sería siempre establecido a 17, identificando el protocolo UDP.

Pero es importante para los *proxies*, *firewalls* y *routers* conocer que protocolo es transportado por los datagramas de IP, así pueden decidir mejor que hacer con ellos. Así que, estos 8 *bytes* eran un mal necesario.

Los campos MDTP Protocol Identifier 1 y MDTP Protocol Identifier 2 fueron establecidos originalmente a números hexadecimales F7873072 y

17074012 respectivamente, dichos valores fueron escogidos al azar y utilizados**,** ya que era altamente improbable que cualquier data transportada por UDP empezara con esos mismos 8 *bytes*.

Luego, se determinó que esta manera de identificar MDTP no era la apropiada**,** ya que era necesario adentrarse en el datagrama de UDP para conocer que se estaba transportando. Así que, se aceptó establecer ciertos puertos de UDP específicos para el transporte de MDTP y ayudar con su identificación. Sin embargo**,** para soportar la multiplexión de protocolos y que otros pudieran utilizar los mismos puertos, este identificador de 8 *bytes* no fue eliminado inmediatamente.

En primer lugar, el campo MDTP Protocol Identifier 2 desapareció en la séptima publicación de MDTP; y luego el campo MDTP Protocol Identifier 1 fue reducido de 32 a 28 *bits* en la siguiente versión. Luego, en la novena publicación de la versión, este campo fue establecido como opcional compartiendo ese campo con CRC (Cyclic Redundacy Code).

Finalmente, el identificador de MDTP fue completamente eliminado cuando fue publicada la primera versión de SCTP y su identificación residía en el uso de un puerto exclusivo en UDP. El problema fue finalmente eliminado cuando SCTP funciono sobre IP y se le asignó el número de protocolo 132 como identificador.

Los campos Acknowledgement Number y Sequence Number son completamente equivalentes con los campos en TCP que utilizan el mismo nombre.

El campo Data Size fue incluido a un inicio en el encabezado de MDTP. TCP tiene el campo Data Offset que incluye únicamente la longitud del encabezado. Esto es suficiente**,** ya que la capa de IP informa del tamaño completo del datagrama TCP.

Esto no funcionó con MDTP**,** ya que éste fue diseñado para agregar relleno**,** (*bytes* con valor de 0) al final del datagrama para lograr que su longitud sea un múltiplo de 4 *bytes*, y dichos *bytes* no son parte de la información. Este relleno fue establecido debido que la mayoría de ordenadores leen información en piezas de 32 *bits* o mayores. Así que, es más efectivo leer 4 *bytes* en la fila aunque solo una de ellas contenga información utilizable.

Este campo fue trasladado en la primera versión de SCTP**,** desde los encabezados hacia *chunks*, el cual es una unidad de información en los paquetes de SCTP que consiste en un encabezado y un contenido.

MDTP era un protocolo orientado a mensajes, lo cual es una gran diferencia con TCP que simplemente maneja un flujo de *bytes*, y es la aplicación la que delimita los diferentes mensajes transportados en el mismo flujo y recortados en las piezas específicas.

En MDTP los mensajes enviados son identificados por los campos *part* y o*f*. Si el datagrama contiene un mensaje completo, entonces el campo *part* se establece en 0 y el campo Of en 1. De otro modo, el campo *Of* informa al receptor el número de fragmentos del mensaje y el campo *part* indica el orden de los fragmentos**,** para que el mensaje puede ser ensamblado correctamente.

Esto provocó que la longitud del mensaje transmitido por MDTP pudiera tener una longitud de 255 veces el MTU de la red utilizada para transmitir

datagramas**,** menos los encabezados IP y MDTP. Si se está utilizando *ethernet*, cuyo MTU es de 1500 *bytes* e IPv4 como el protocolo de red, con un encabezado típico sin los 20 *bytes* de opciones, se obtiene 255*(1500-20-24) = 371 280 *bytes*. Aun cuando este valor debe ser más que suficiente para un mensaje, el mecanismo fue modificado más adelante en SCTP haciendo la longitud máxima del mensaje técnicamente infinita.

Existen 16 *bits* de banderas denominados *Flags* y *Mode*. Esto es similar a TCP, con la diferencia que las funcionalidades de MDTP necesitaban más *bits* para funcionar correctamente**,** por lo que solamente dos de estos *bits* se encontraban libres (RE1 y RE2). Dichos *bits* se utilizaban como negociadores en la fase de establecimiento para solicitar servicios opcionales, y también para informar al receptor cual era la estructura interna del campo d*ata*, el cual podía contener distintos tipos de información incluyendo data de usuario.

Estos dos *bits* los cuales ya se encontraban en uso desde la cuarta versión de MDTP, limitaron las posibilidades de extensión de MDTP. Por otra parte, tener esa gran cantidad de banderas era de cierta manera inapropiada y difícil de manejar, haciendo el procesamiento del datagrama ineficiente.

En la octava versión el datagrama MDTP fue transformado, incluyendo áreas denominadas *control parameter part* y *data part*. Las banderas fueron remplazadas por dos *bits* que únicamente indicaban si las áreas de Control y Data se encontraban presentes, además un identificador de 6 *bits* para el parámetro de control. Agregado a esto, se reservaron 8 *bits* para uso futuro.

El campo de *version* representaba un número de versión del protocolo. Este campo se redujo a 4 *bits* en la octava publicación de MDTP y finalmente eliminada en la sexta versión SCTP.

El campo In Queue contiene el número de mensajes que el transmisor del datagrama tenía en su cola de entrantes esperando a ser leídos por la aplicación. Éste se utilizaba para controles de flujo. Este campo es equivalente al campo window en TCP, con la diferencia que indica cantidad de mensajes sin leer y no *bytes*.

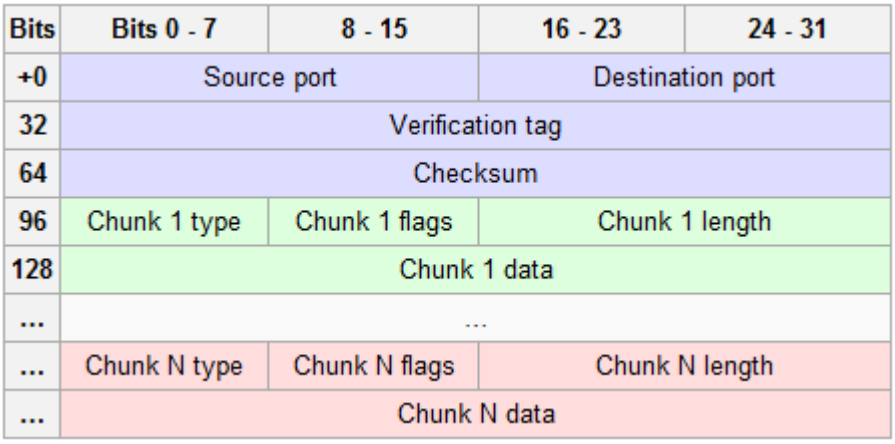
El campo In Queue fue útil para el transmisor en dar a conocer el estado del *buffer* del receptor. Sin embargo, sería mucho más útil si la información contenida en este campo expresara la cantidad en *bytes* y no en mensajes.

Así que**,** la decisión final fue que la información sobre el tamaño del *buffer* receptor sería intercambiado durante la fase de establecimiento y que un campo similar denominado Advertised Receiver Window sería utilizado, esta vez no en el encabezado**,** sino en las partes del reconocimiento (*acknowledgement*). Este cambio se dio en la primera versión de SCTP.

3.3. Encabezado y estructura interna de SCTP

Se revisó la estructura inicial de MDTP. La estructura de SCTP se parecía a la que se muestra en la figura 10. Como se puede observar, el encabezado de SCTP es completamente diferente que al de MDTP y mucho menos complejo. Esto permite que los datagramas de SCTP sean más fáciles de procesar. Sin embargo, la estructura interna es más elaborada ya que se puede observar que los datagramas de SCTP contienen varias estructuras en distintos niveles.

Figura 10. **Estructura del datagrama SCTP**



Bits	Bits 0 - 7	8 - 15	16 - 23	24 - 31
+0	Source port		Destination port	
32	Verification tag			
64	Checksum			
96	Chunk 1 type	Chunk 1 flags	Chunk 1 length	
128	Chunk 1 data			
...	...			
...	Chunk N type	Chunk N flags	Chunk N length	
...	Chunk N data			

Fuente: SCTP Packet Structure. http://en.wikipedia.org/wiki/SCTP_packet_structure. Consulta: 22 de mayo del 2012.

Cada datagrama de SCTP contiene un encabezado en común de 12 *bytes* seguido de una o más estructuras que se denominan *chunks*. El encabezado en común contiene los siguientes elementos:

- En enero del 2000 SCTP se convirtió en un protocolo de transporte sobre IP. Por lo tanto, la información transportada en el encabezado de UDP se movió hacia el encabezado de SCTP. Los parámetros Source Port Number y Destination Port Number debían aparecer en el encabezado de SCTP.

  Sin embargo, al tener a SCTP al mismo nivel que protocolos como TCP o UDP dio como resultado que SCTP debía tener interacción con otros protocolos que también corren sobre IP. Uno de dichos protocolos ICMP es el encargado de notificar a los usuarios IP sobre situaciones problemáticas**,** como falta de espacio de buffer en un *router*, una

dirección inalcanzable, o tablas de ruteo ineficientes. ICMP fue inicialmente definido para IPv4, pero también posee una versión IPv6.

Cuando existe alguna situación en la red IP que active la transmisión de un mensaje ICMP, dicho mensaje contiene el inicio del paquete IP que originó la anomalía. Los mensajes ICMPv6 incluyen tantos *bytes* del datagrama IP original como lo permiten las especificaciones (1280 *bytes*), mientras que el mensaje ICMP contiene solamente los primeros 8 *bytes* del datagrama IP. Debido a que los campos source port number y destination port number son vitales para identificar qué asociación SCTP originó el mensaje ICMP, dichos campos debieron ser ubicados en estos primeros 8 *bytes* del datagrama SCTP.

Cuando SCTP fue modificado para operar sobre IP, también se acordó que todos los puertos TCP utilizados por aplicaciones "bien conocidas"**,** serian automáticamente reservados en el espacio SCTP *port address*. Esto haría más sencilla la migración de TCP a SCTP.

- Verification tag, este campo es la evolución de los campos MDTP *acknowledgement number* y *sequence number*. Tiene el mismo rol que los campos en la fase de establecimiento, pero también provee protección contra ataques *blind*. Es un entero de 32 *bits*, seleccionado aleatoriamente e intercambiado en la fase establecimiento. Es utilizado sin modificaciones durante toda la duración de la asociación y valida que los datagramas entrantes sean realmente enviados por el extremo válido y no por un atacante.

- El campo Checksum ha evolucionado desde la aparición en el encabezado de CRC-16 (Cyclic Redundancy Check of 16 *bits*), esta

opción se mantuvo durante las primeras cinco versiones de SCTP. Luego, en la sexta versión de SCTP para obtener una verificación de 32 bits, Adler-32 c*hecksum*, siendo esta versión finalmente incluida en las especificaciones SCTP.

Sin embargo**,** luego se determinó que Adler-32 proveía baja detección de errores en pequeñas tramas y tomando en cuenta que en telefonía los mensajes usualmente ocupan menos de 128 *bytes*, se debería de usar otro tipo de verificación. Por lo tanto, el acuerdo se tuvo con la implementación de CRC-32 (Cyclic Redundancy Check of 32 *bits*).

Luego del encabezado común debe existir como mínimo un *chunk* (fragmento). Un *chunk* o fragmento es una estructura independiente con su identificador y significado específico**,** pero con una estructura TLV (Type-Length-Value) común. Son utilizados para enviar peticiones al punto final remoto y recibir respuestas de los mismos. Los fragmentos definidos en un inicio fueron:

- Initiation (init), Initiation Acknowledgement (init ack), State Cookie (cookie echo) y Cookie Acknowledgement (cookie ack), estos son utilizados en la fase de establecimiento.

- Payload Data (data) y Selective Acknowledgement (sack) son utilizados para la transferencia de información.

- Heartbeat Request (heartbeat) y Heartbeat Acknowledgement (heartbeat ack) son utilizados para monitorear el estado de las distintas interfaces involucradas en la asociación.

- Operation Error (error) es utilizado para reportar un error no fatal.

- Shutdown (shutdown), Shutdown Acknowledgement (shutdown ack) y Shutdown Complete (shutdown complete) son utilizados durante el término normal de una asociación.

- Abort (abort), el cual reporta un error fatal y termina una asociación.

El listado completo de *chunks* disponibles puede verificarse en la tabla II a continuación:

Tabla II. **Descripción completa de *chunks* disponibles**

Value	Abbreviation	Description
0	DATA	Payload data
1	INIT	Initiation
2	INIT ACK	initiation acknowledgement
3	SACK	Selective acknowledgement
4	HEARTBEAT	Heartbeat request
5	HEARTBEAT ACK	Heartbeat acknowledgement
6	ABORT	Abort
7	SHUTDOWN	Shutdown
8	SHUTDOWN ACK	Shutdown acknowledgement
9	ERROR	Operation error
10	COOKIE ECHO	State cookie
11	COOKIE ACK	Cookie acknowledgement
12	ECNE	Explicit congestion notification echo (reserved)
13	CWR	Congestion window reduced (reserved)
14	SHUTDOWN COMPLETE	Shutdown complete
15-62	N/A	Reserved by IETF
63		IETF-defined chunk extensions
64-126		Reserved by IETF
127		IETF-defined chunk extensions
128-190		Reserved by IETF

Continuación de la tabla II.

192-254		Reserved by IETF
255		IETF-defined chunk extensions

Fuente: SCTP Packet Structure, http://en.wikipedia.org/wiki/SCTP_packet_structure.
Consulta: 22 de mayo de 2012.

Esta estructura es la que diferencia a SCTP de TCP. Como se ha visto, el diseño inicial de MDTP era muy parecido a TCP, con muchas banderas y tamaños de campos definidos. La estructura de SCTP intenta evitar dos problemas presentes en TCP:

- TCP tiene un problema grande en sus posibilidades de extensión. TCP tiene únicamente 40 *bytes* de espacio para incluir opciones en su encabezado. Esto hace a TCP no extensible (y es justo ésta la razón por la que se necesitó un nuevo protocolo), así que los diseñadores de SCTP intentaron evitar enfrentarse con la misma situación en el futuro.

- Cuando TCP fue diseñado, una de las consignas principales en su diseño fue hacerlo tan eficiente como fuese posible en términos de la sobrecarga producida por el encabezado. Si algo hacia que el procesamiento del datagrama de TCP fuera difícil pero ahorraba un *byte* de su encabezado, entonces era una buena opción. Incluso existen estándares que definen métodos para comprimir el encabezado de TCP de su tamaño típico de 20 *bytes*, a un promedio de 3 *bytes*. Esto es posible debido a las similitudes entre encabezados TCP en segmentos que pertenecen a la misma conexión.

Estos esfuerzos pueden entenderse si se toma en cuenta que en los años 70 y 80, cuando se inicia el uso de TCP, la transmisión de la información podía ir a una velocidad tan alta como varias decenas de *kilobytes* por segundo en el mejor de los casos. Pero esto ya no es así. Hoy en día es muy común que un ordenador esté conectado a internet a través de una tarjeta Ethernet de 100 Mbps, 1000 veces más rápido que las conexiones utilizadas hace 30 años.

Así que, actualmente las características claves para un protocolo de transporte, no es si la sobrecarga aumenta o disminuye un poco, sino si el protocolo es simple, fácil de extender, y sus datagramas son rápidos de procesar.

Por lo tanto, la primera versión de SCTP fue en cierto sentido una revolución en este aspecto y fue profundamente modificado en la última versión de MDTP. Como se observa en la figura 10, todos los fragmentos siguen una línea básica en su definición. Estos contienen los siguientes campos:

- Campo Chunk Type, esto identifica un fragmento de otro y dice al receptor que hacer con los mismos. Originalmente, el valor 254 fue utilizado para fragmentos específicos del vendedor (*vendor-specific*), esto fue diseñado cuando SCTP corría sobre UDP y cuando cada compañía interesada en SCTP podría implementar su propia versión. Conforme que SCTP fue evolucionando, se situó sobre IP, y sobre todo existían planes para que su código se incluyera dentro del *kernel* del sistema operativo (ya realizado por UNIX / LINUX). Por lo que sería más factible que compañías interesadas en SCTP lo compraran directamente con su proveedor de sistema operativo.

Esto incrementaría el costo para una empresa que quisiera una extensión específica de SCTP, por lo tanto**,** el uso de fragmentos específicos para cada vendedor podía crear conflictos de interoperabilidad**,** ya que una extensión única del vendedor es funcional entre dos servidores que entiendan dicha extensión (normalmente del mismo vendedor), en cuyo caso el vendedor puede introducir su nueva funcionalidad en la manera que le plazca.

Debido a que los fragmentos (*chunks*) pueden ser creados para llenar las necesidades derivadas de cualquier nueva característica, el receptor podría tener fragmentos que puede o no entender. Algunas veces el transmisor del fragmento debe conocer si el receptor entendió o no, y en otras ocasiones, el procesamiento del fragmento puede ser de vital importancia para continuar con el procesamiento del resto del datagrama. Así que, los primeros dos *bits* del campo Chunk Type le dicen al receptor que hacer en caso no reconozca el fragmento. Dependiendo del valor, las acciones a tomar son:

- 00: el receptor debe descartar el datagrama por completo, sin procesar cualquier fragmento futuro con referencia al mismo.
- 01: igual comportamiento que 00 pero reportar al transmisor que no reconoció el fragmento.
- 10: el receptor debe descartar este fragmento, pero continuar con el procesamiento del resto.
- 11: igual comportamiento que 10 pero reportar al transmisor que no reconoció el fragmento.

- Campo Chunk Flag, usualmente existen campos del fragmento que se pueden expresar como valores booleanos y para estos existe una

estructura de banderas. El significado de las banderas depende del fragmento y existen únicamente tres fragmentos que poseen banderas: *data*, *shutdown complete* y *abort.*

- Campo Chunk Length, campo necesario ya que la longitud de los fragmentos puede ser variable.

- Campo Fixed Fields, no todo los fragmentos proveen información extra a la que se encuentra en c*hunk type*. Sin embargo, muchos contienen campos extras para proveer información extra.

- Si el fragmento contiene información opcional o una longitud variable puede contener información en el campo *parameters*.

- Finalmente, si el campo chunk length no es múltiple de 4 *bytes*, se agregan *bytes* de *padding* al final.

Los parámetros son similares a los fragmentos pero en un nivel más bajo. Fueron creados para contener información opcional o de longitud variable dentro de los fragmentos y proveer mayores posibilidades de extensión. Como se puede observar en la figura 10 contienen la siguiente estructura:

- Los parámetros fueron diseñados para hacer virtualmente ilimitadas las opciones de expansión de SCTP. Por lo tanto, poseen un campo Parameter Type de dos *bytes*. Como en el caso de los fragmentos, los primeros *bits* del campo le dicen al receptor que acción tomar en caso entiendan o no lo enviado. El comportamiento es básicamente el mismo que el descrito anteriormente, pero por estar a un nivel inferior cuando el

primer *bit* es 0**,** el receptor no debe descartar el datagrama sino el fragmento.

- Debido a que los parámetros contienen longitud variable existe el campo parameter length para proveer esta información al receptor.

- El campo Parameter Value es el que contiene la información (notar que este campo es opcional**)**, muchas veces es suficiente con el campo parameter type.

- Como en el caso de los fragmentos existen *bytes* de *padding* al final en los parámetros**,** si la longitud de estos no es múltiple de 4 *bytes*.

Existe otra estructura dentro de los datagramas de SCTP como se observa en la figura 10, denominada *error cause*. Básicamente igual a los parámetros. Tienen exactamente la misma estructura teniendo campos Cause Code, Cause Length y Cause Value. Las diferencias entre error causes y parameters son:

- *Error causes* informan sobre una situación problemática como la recepción de un fragmento o parámetro irreconocible, falta de recursos para abrir una nueva asociación, o tipo de dirección que no se puede administrar.

- Los primeros dos *bits* del campo *cause code* no tienen el mismo significado que en *parameter type*. En teoría, un servidor nunca podría recibir una causa de error desconocido (a menos que uno de los dos puntos relacionados involucre *bugs* en su implementación).

Como se aprecia, el diseño de SCTP es el que hace posible una de sus características principales: extensibilidad. SCTP fue creado para transportar distintos protocolos de señalización telefónica con distintos requisitos, y su amplio rango de uso hace que incluso sea posible considerarlo como un sustituto de TCP. Gracias a la estructura de fragmentos de un datagrama de SCTP, es muy sencillo diseñar nuevos fragmentos y parámetros que provean una nueva característica y definir causas de errores que informen sobre problemas al utilizar una nueva extensión.

3.4. Manejo de asociaciones SCTP

Como en TCP, los pasos requeridos para establecer y liberar una asociación se puede representar como una máquina de estados finita. En SCTP se tienen 8 estados en lugar de los 11 que se tienen en TCP, como se representa en la figura 11. En ésta se puede identificar los ocho diferentes estados como rectángulos redondeados**,** (notar que el estado *closed* aparece dos veces, y que el estado en la parte superior etiquetado como a*ny state* no es un estado distinto, sino significa que puede ser cualquiera de los 8 estados). Las dos representaciones de computadoras identifican aquellos servidores de donde o hacia donde se envían o reciben datagramas.

Como se observa en la leyenda de la figura en la parte inferior derecha hay tres distintos tipos de flechas que significan distintas cosas:

- Las flechas con muescas con texto en negrilla significan que el usuario superior realiza una llamada primitiva, denominada *associate* (para iniciar una asociación), *shutdown* (para terminar una asociación) y *abort* (para abortar una asociación).

- Las flechas con texto en itálicas representan los fragmentos de control enviados al punto receptor o recibido por él. Estas flechas van desde los rectángulos de estados hacia el servidor y viceversa.

- Finalmente, las flechas que no tienen ningún texto representan cambios internos de un estado. Éstas van desde un rectángulo de estado hacia otro.

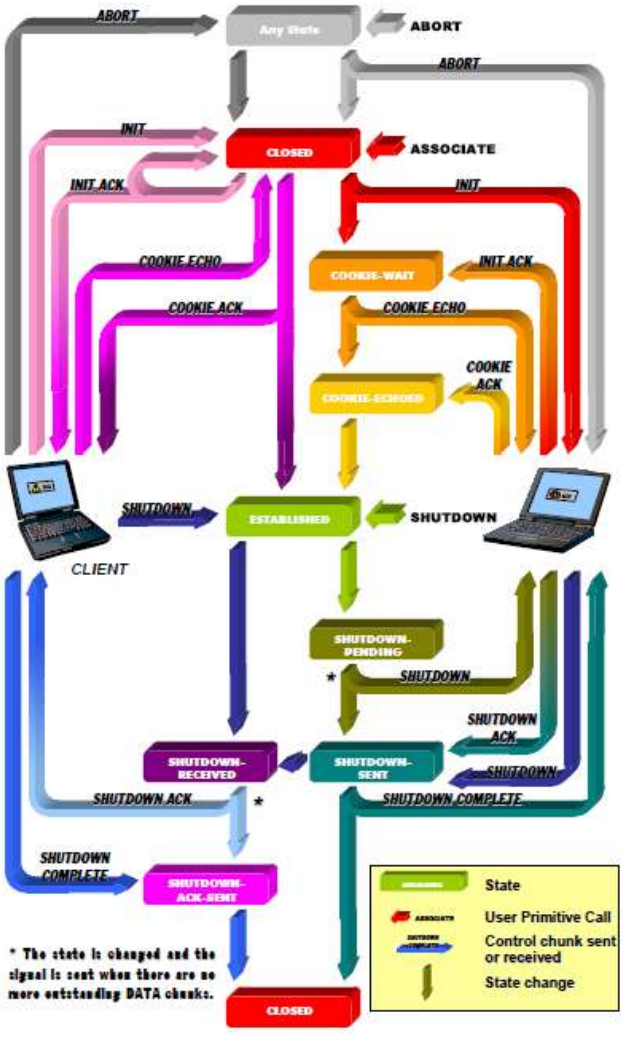
El diagrama es colorido y varias flechas tienen distinto color para su inicio y fin. Esto ayuda a entender de una mejor manera lo que se intenta transmitir con el diagrama.

Una llamada de un usuario superior o un fragmento entrante usualmente acciona un cambio de estado**,** y habitualmente otro fragmento es enviado, así que**,** la clave para identificar qué salida se relaciona con que entrada es a través del color. La respuesta a una llamada específica o un fragmento entrante es la flecha cuyo color de inicio es el mismo que el color final de la flecha que representa la llamada. Esto aplica no solamente al estado, sino también a las respuestas del otro servidor a los fragmentos de control salientes.

Como ejemplo, se observa que el fragmento *init* entrante en el estado *closed* aparece con un color rosa en la terminación de la flecha. Por lo tanto, se sigue la flecha saliente color rosa del estado *close* y se observa que la respuesta a ese fragmento es que se mantiene el mismo estado y se envía un fragmento *init ack* de regreso al servidor. Por otra parte, la flecha representando el *init ack* termina en color fucsia cuando alcanza al servidor, así que su respuesta al *init ack* es la flecha saliente del servidor que inicia con el mismo color, la cual en el ejemplo envía de vuelta un fragmento *cookie-echo.*

Existen dos cambios de estado marcados con el símbolo * los cuales no son producidos por ningún fragmento entrante o llamada primitiva representados, pero ocurren cuando los servidores reconocen todos los posibles datos pendientes que se puedan tener.

Figura 11. **Manejo de conexiones SCTP**



Fuente: Arias, Ivan, Stream Control Transmission Protocol. p. 49.

Otra manera para interpretar la figura es que la parte derecha representa las acciones tomadas cuando se es la parte activa en el establecimiento o terminación de una asociación**,** (por ejemplo cuando se actúa como el cliente**,** se conecta al servidor y luego termina la conexión), mientras que la parte izquierda es lo opuesto**,** (cuando se actúa como servidor esperando a que el cliente se conecte de regreso, y luego es el cliente quien libera la asociación). Esta regla se rompe solamente en un caso, que es en el fragmento *shutdown* enviado por el servidor de la derecha**,** (esto se hace para mantener la claridad del diagrama).

El diagrama de estados de SCTP es muy similar al de TCP, esto es debido a que TCP es uno de los antecesores de SCTP pero también porque la mayoría de protocolos de transporte manejan en cierta manera la misma representación de estados. Pero existen grandes diferencias entre ellos, siendo las más importantes:

- SCTP utiliza para la fase de establecimiento un intercambio de cuatro vías (*four-way handshake*) mientras que TCP utiliza uno de tres vías. Esto está ligado al mecanismo de *cookie* utilizado para evitar un ataque similar al ataque SYN en TCP.

- La terminación de una asociación es más simple. En SCTP no existen conexiones semiabiertas (*half-open connections*).

4. SIGTRAN

El grupo de trabajo SIGTRAN (Signaling Transport) que forma parte de IETF (Internet Engineering Task Force), ha diseñado un nuevo set de protocolos para el transporte de mensajes de señalización SS& sobre IP. La *suite* de protocolos consiste de un nuevo protocolo de transporte y varios protocolos de adaptación, y se convirtió en un estándar en los años 2000 y 2001. Utilizar protocolos SIGTRAN es el primer paso en fusionar redes SS7 con redes IP. La primera razón para la utilización de IP es descargar la red SS7 y hacer escalable esta red para el incremento de usuarios de telefonía.

La solución SIGTRAN también será utilizada para interconectar redes SS7 aisladas, que de otro modo requerirían una infraestructura de alto costo. Hoy en día, las compañías de telefonía están avanzando hacia una red todo IP (*all-IP*), donde IP reemplazará redes de telefonía tradicionales, sin embargo, es una tarea de largo plazo y el mayor enfoque actualmente es posibilitar que estos dos sistemas puedan coexistir y mejorar los servicios que proveen.

Para la entrega de mensajes sobre IP se tienen dos protocolos de transporte TCP (Transmission Control Protocol) y UDP (User Datagram Protocol), pero para la señalización en tiempo real implican ciertas limitantes, Sin embargo, las características deseadas en el transporte de señalización son:

- Transferencia ordenada y confiable
- Redundancia en caso de falla

- Poca pérdida y retraso

- Seguridad ante la negación de servicio (DoS, Denial of Service)

UDP y TCP no pueden soportar todos estos requisitos, por lo tanto, un nuevo protocolo de transporte fue diseñado por SIGTRAN, SCTP (Stream Control Transmission Protocol).

4.1. UDP

UDP es un protocolo de transporte no orientado a conexión, esto significa que intrínsecamente no utiliza un mensaje de ACK para garantizar el orden y confiabilidad del transporte. UDP es útil en situaciones cuando se necesitan altas tasas de transmisión, pero no tiene que cumplir con los demás requisitos de rendimiento de los mensajes de señalización.

4.2. TCP

TCP es un protocolo de transporte orientado a conexión**,** que provee un flujo de *bytes* y garantiza que su entrega sea en orden. Esto es ideal para transmitir gran cantidad de *data*, como archivos o correo electrónico, pero su estricta entrega-en-orden es lo que también lo hace inadecuado para mensajes de señalización. TCP es muy sensible a retrasos causados por la red o pérdida de paquetes que usualmente causan retransmisiones.

Cuando se espera por el reconocimiento de un paquete perdido todos los demás paquetes son retrasados, lo cual se denomina bloqueo de primero de fila (*head-of-line blocking*). Esto genera retrasos innecesarios para otros paquetes,

por lo tanto TCP es inadecuado para aplicaciones de tiempo real, como la señalización.

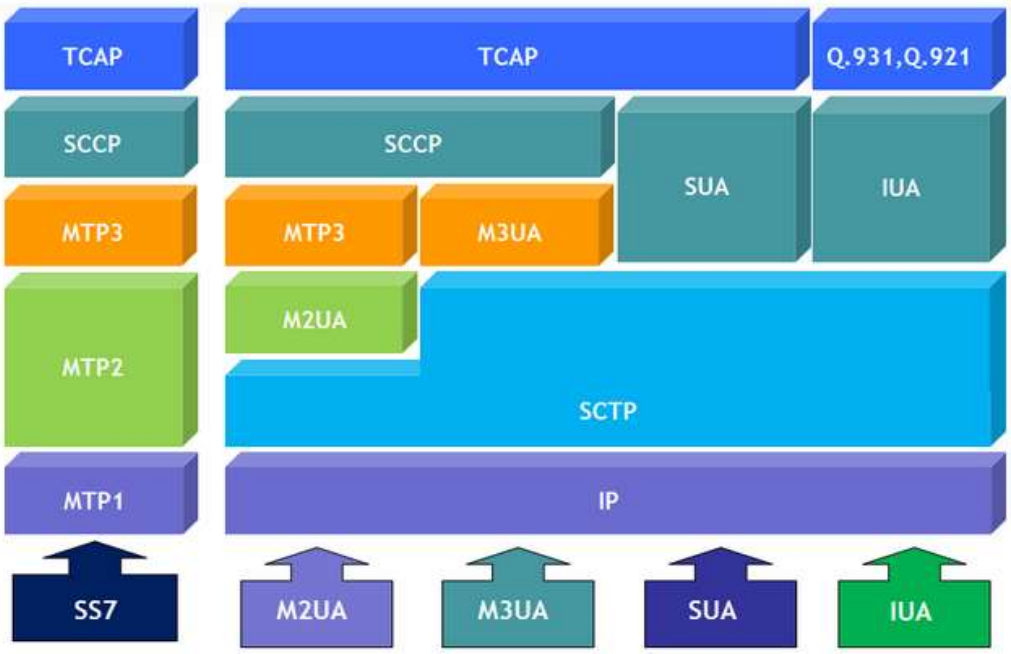
Otra desventaja de TCP es su vulnerabilidad a ataques DoS. Para establecer una conexión TCP el cliente debe enviar un mensaje SYN al servidor**,** el cual contestará con un mensaje SYN ACK. El servidor espera con el correspondiente ACK del cliente, que es el último paso en el intercambio para establecer una conexión TCP. Sin embargo, este procedimiento es susceptible a un tipo de ataque DoS denominado ataque SYN, causado por muchos mensajes SYN enviados al servidor**,** los cuales ocupan recursos de memoria y pueden derivar en un colapso del servidor, previniendo que clientes legítimos obtengan conexión. Esto no es aceptable en redes SS7**,** donde los servicios telefónicos deberían estar siempre disponibles.

4.3. Arquitectura

El protocolo de SIGTRAN incluye el protocolo de transporte SCTP junto con otras capas de adaptaciones de usuario (*UA- User Adaptation*)**,** que son necesarias para el transporte de mensajes SS7 sobre IP. La arquitectura de SIGTRAN consiste de tres niveles:

- Capa IP
- Capa de Transporte (SCTP)
- Capas de Adaptación de usuario (M2PA, M2UA, M3UA y SUA)

Figura 12. **Arquitectura SS7 vs. SIGTRAN**



Fuente: SS7 over IP. http://ict2.springnote.com/pages/9939166/attachments/5679946. Consulta: 25 de mayo del 2012.

En la figura 12 se puede observar las tres capas inferiores con los cambios presentados por SIGTRAN. Estos reemplazan las capas de SS7 (MTP1 y MTP2), permitiendo el transporte sobre IP. SCTP es un protocolo de transporte similar a TCP, con ciertos cambios que lo adaptan mejor a la señalización SS7. Los protocolos de adaptación de usuario permiten que los usuarios de SS7 (MTP3, SCCP, TCAP, ISUP etc.) no adviertan que las capas SS7 originales han sido remplazadas. Dependiendo de la red de telefonía, distintos protocolos de adaptación pueden ser utilizados.

4.4. SCTP

Para enviar información sobre una conexión IP usualmente son utilizados TCP o UDP. Sin embargo, como ya se mencionó, mensajes de señalización de SS7 poseen requisitos de pérdida y retraso muy exigentes, por lo tanto**,** TCP no es una elección adecuada ya que los retrasos son demasiado prolongados y UDP no provee la confiabilidad necesaria. El protocolo SCTP es similar a TCP**,** (ya que provee mecanismos para control de flujo y congestión)**,** pero posee ciertas diferencias denominadas multi-*homing* y multi-*streaming*.

4.4.1. Multi-*homing*

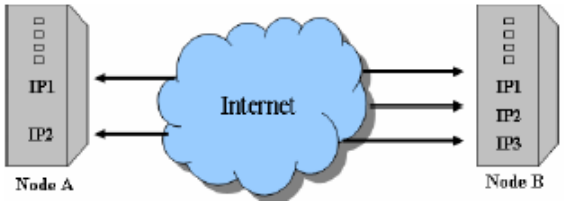
Un nodo en multi-*homed* es aquel con varias direcciones IP, donde cada par de direcciones IP entre dos nodos es denominado *path*. En la figura 13, el nodo A tiene tres *path* hacia el nodo B y el nodo B posee dos *path* hacia el nodo A.

En una conexión SCTP (en SCTP esto es denominado “asociación”), cada nodo elije un *path* primario. Si una falla ocurre en este *path*, las retransmisiones son enviadas a través del *path* alterno. Cada *path* está asociado con mensajes de *Heartbeat* los cuales indican si se encuentra en modo activo o inactivo. Luego de un número específico de retransmisiones, un *path* es considerado inactivo y un nuevo *path* es elegido, y si éste se encuentra activo se convierte en el nuevo *path* primario.

Esta característica de multi-*homing* permite a la red redireccionar tráfico a otras direcciones IP, dando a la red una mayor tolerancia a fallas físicas de *links*. En una red SS7 clásica existen siempre como mínimo dos *paths* físicos distintos por donde transmitir información. Ya que SIGTRAN debe proveer una

solución IP con todas las cualidades de una red SS7, la característica de multi-*homing* puede ser utilizada para proveer la misma capacidad de redundancia.

Figura 13. **Multi-*Homing***

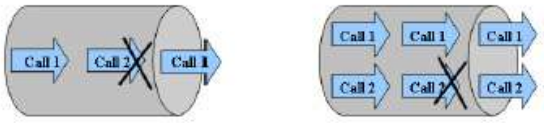


Fuente: Imonen, Mia. Signaling over IP. p. 10.

4.4.2. Multi-*streaming*

Multi-*streaming* es utilizado para evitar bloqueos *head-of-line* (primero de la línea), el cual es un fenómeno común en TCP, como se muestra en la figura 14. Cuando un paquete de señalización es perdido en una trama TCP, toda la conexión es bloqueada en espera de una retransmisión, resultando en un bloqueo *head-of-line*. El retraso para recuperar la información perdida puede ser de varios segundos, lo cual es inaceptable si se realiza una llamada telefónica

Figura 14. **Multi-*Streaming***



Fuente: Imonen, Mia. Signaling over IP. p. 11.

Por lo tanto en SCTP una asociación entre dos nodos puede tener varios flujos, cada uno asignado a una aplicación o recurso específico, y estos flujos no se bloquean unos a otros en el evento de una pérdida o retraso de un paquete. Crear varios flujos en TCP es también posible, pero implica utilizar varias conexiones TCP donde cada una actúa como un flujo.

Cada conexión introduce un TCB (Transport Control Block) en el lado del servidor, los cuales contienen toda la información importante sobre la conexión. Estos TCBs consumen memoria, y sus mediciones pueden ser significativas para un punto congestionado con varios clientes, por lo tanto, múltiples conexiones TCP no son una alternativa deseable. También al utilizar una conexión SCTP con varios flujos de información en lugar de varias conexiones TCP**,** permite evitar tiempos de *set up* innecesarios.

4.4.3. Otras características de SCTP

- Message Boundary Preservation: TCP es un protocolo orientado a *byte*, mientras que SCTP es un protocolo orientado a mensaje**,** que coloca uno o más mensajes de señalización completos dentro de un mensaje SCTP. Un mensaje SCTP está compuesto por un encabezado común y varios *chunks*, donde estos contienen la información del usuario con distintas longitudes.

- *Out of Order Transmission*: un nodo TCP siempre recibe paquetes en un orden secuencial. Con SCTP es posible el envío de paquetes en orden o fuera de orden, dependiendo como lo prefiera la aplicación. Cuando se refiere a señalización, el orden secuencial dentro de un mismo flujo es importante, pero no lo es entre flujos distintos.

- *Cookies*: ambos TCP y SCTP atraviesan un intercambio antes de establecer una conexión de punto a punto. Mientras TCP utiliza un intercambio de 3-vias, SCTP utiliza uno de 4-vias**,** el cual incluye *cookies* para proteger la conexión de ataques DoS. Un DoS ocurre cuando un atacante de un modo u otro, acapara el servicio de un usuario legítimo.

4.5. Capas de adaptación

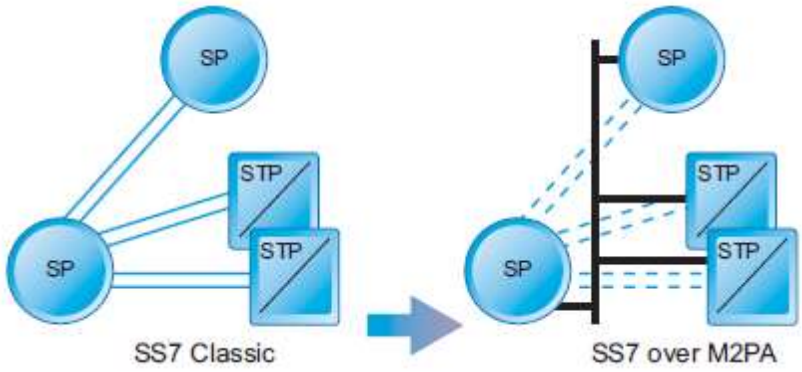
El protocolo SIGTRAN define cuatro capas de adaptación (AU – *Adaptation Layer*). Estas cuatro capas de adaptación sirven a un propósito en común: transportar mensajes de señalización desde el punto en que termina un mensaje SS7, sobre una red IP, hasta el punto de la capa superior.

4.5.1. M2PA

La capa de adaptación de punto a punto MTP2-User (M2PA por sus siglas en inglés)**,** es un protocolo SIGTRAN que transporta mensajes de señalización MTP de SS7 sobre IP utilizando SCTP. Es un protocolo de adaptación entre MTP3 y SCTP y funciona entre pares de nodos de señalización.

Utilizando M2PA hace posible mantener la topología original de una red SS7, por ejemplo**,** todos los elementos de red como STP (Signaling Transfer Points), *point codes*, etc. El único punto que cambia es que el transporte de la señalización ocurre sobre IP**,** en lugar de los *links* tradicionales de 64 Kbps, ver figura 15.

Figura 15. **Configuración M2PA**

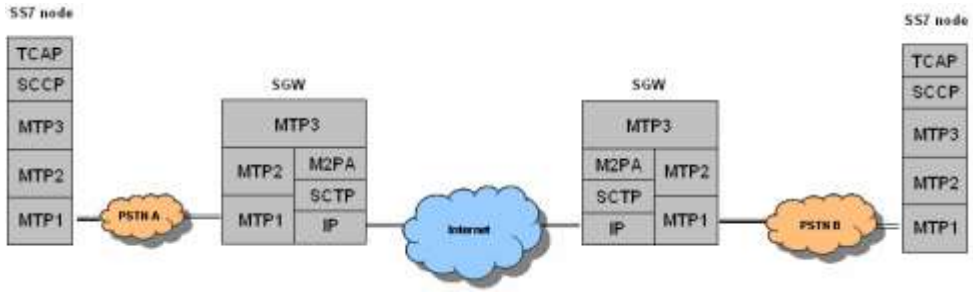


Fuente: Nothridge, Steve. Convergent SS7 Signaling. p. 3.

M2PA puede ser utilizado entre dos nodos de señalización IP, o entre un SGW (Signaling Gateway) y un nodo de señalización IP, pero es más común entre dos SGWs, por ejemplo**,** interconectar dos redes SS7 aisladas (PSTN A y PSTN B) a través de una red IP. La figura 16 muestra las dos redes SS7 distantes**,** que son combinadas a través de una red IP de mucho más bajo costo.

Debido a que los *links* SS7 son dedicados únicamente para tráfico de señalización existe una continua asignación de ancho de banda, por lo tanto**,** *links* que no se utilizan frecuentemente de igual manera tienen un ancho de banda asignado**,** aun cuando éste es un recurso valioso y escaso. La solución IP mezcla el tráfico de señalización con otro tráfico IP y por lo tanto reduce el costo de la señalización, ya que un *link* puede ser utilizado entre varios usuarios.

Figura 16. **Interconexión vía M2PA**



Fuente: Imonen, Mia. Signaling over IP. p. 13.

Ya que ambos SGW tienen capa MTP3 poseen también un *point code*, y un PC SS7 debe ser asignado a cada SGW. Debido a la capacidad punto a punto de M2PA, es posible para los puntos MTP3 comunicarse directamente. El usuario de M2PA es MTP3 en ambos nodos, tal como MTP3 es el usuario de M2PA en la pila de SS7. Esto significa que M2PA es en realidad un remplazo de MTP2 y por lo tanto tiene funciones similares a MTP2.

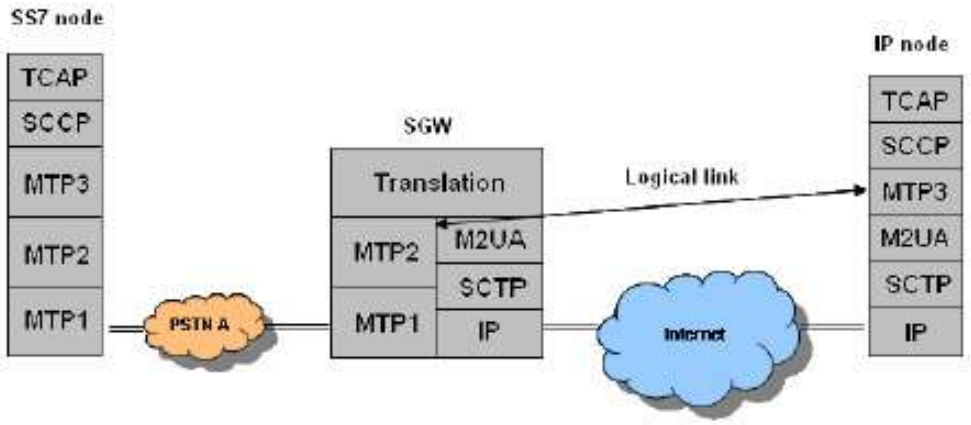
4.5.2. M2UA

La capa de adaptación MTP2-User (M2UA) también acopla MTP3 a SCTP, y es un protocolo para enviar mensajes de señalización entre capas MTP3 en un MGC (Media Gateway Controller), y la capa MTP2 de un SGW, por ejemplo: en una red VoIP. En lugar de ser un protocolo de punto a punto como M2PA, éste opera en configuración cliente-servidor, donde el MGC (nodo IP) es el cliente, y el SGW actúa como servidor.

De este modo, la capa MTP3 en el MGC es el usuario para la capa MTP2 del SGW, y ninguno de los dos está consciente que en realidad son nodos remotos. Este fenómeno donde los mensajes de señalización son transportados

vía IP**,** desde una capa SS7 superior hacia una inferior es denominado *backhauling***,** debido a que el SGW no tiene una capa MTP3, y solamente el MGC posee un *point code*, ver figura 17.

Figura 17. **Interconexión vía M2UA**



Fuente: Imonen, Mia. Signaling over IP. p. 14.

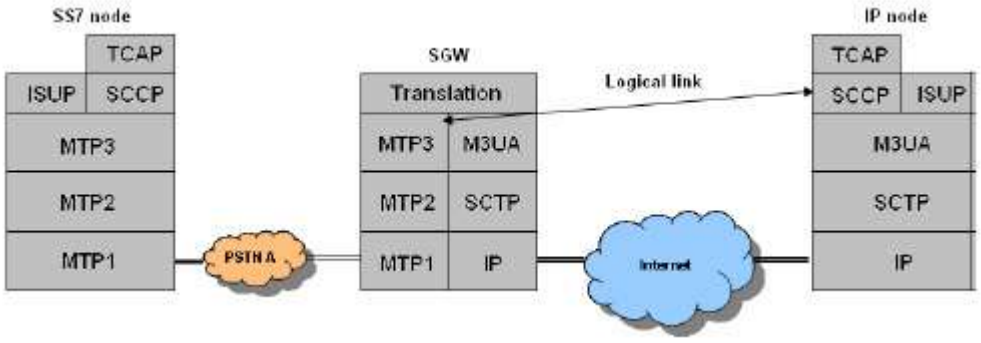
M2UA es frecuentemente usado cuando existe una baja densidad de *links* físicos SS7 en algún área particular de la red, o si los SGW están muy distanciados entre ellos. En este caso**,** *backhauling* puede conectar varios de estos nodos de señalización a un elemento de red central, permitiendo a estos nodos distantes compartir un mismo SGW. Ya que esto es realizado sobre IP, es mucho más barato que *links* SS7, por lo tanto M2UA es una alternativa de bajo costo.

Otra ventaja es el hecho que cada SGW que conecta un punto de señalización remoto a un MGC no posee un *point code*. El *point code* es asignado al MGC lo cual ahorra muchos PCs de SS7**,** que de otro modo serían requeridos por cada SGW (como cuando se utiliza M2PA).

4.5.3. M3UA

La capa de adaptación MTP3-User (M3UA) opera en un esquema cliente-servidor como M2UA**,** es utilizado para proveer conexión remota entre dos capas SS7**,** en un SGW y un MGC (nodo IP). Sin embargo, en este caso, el SGW posee una capa MTP3 (y un *point code*) que se comunica con la capa ISUP/SCCP del MGC, ver figura 18. Aun en este caso, los nodos no están conscientes el uno del otro; la capa MTP3 del SGW no sabe que su usuario (ISUP o SCCP) es remoto y de similar manera, la capa ISUP/SCCP del MGC desconoce que la capa MTP3 del SGW no es propia. Esto es otro ejemplo de *backhauling*.

Figura 18. **Interconexión vía M3UA**



Fuente: Imonen, Mia. Signaling over IP. p. 15.

Como en M2UA, M3UA no procesa ningún paquete de señalización, simplemente lo reenvía a su destino. Esto significa que la capa M3UA en los nodos IP**,** no poseen tablas de enrutamiento y no ejecuta ninguna otra operación de las correspondientes en la capa MTP3. Si se utiliza M3UA en una red completa IP sin nodos puros de SS7, éste reemplaza la capa MTP3 en

ambos nodos IP y opera en modo punto a punto que se conoce como comportamiento IPSP (IP Signaling Point).

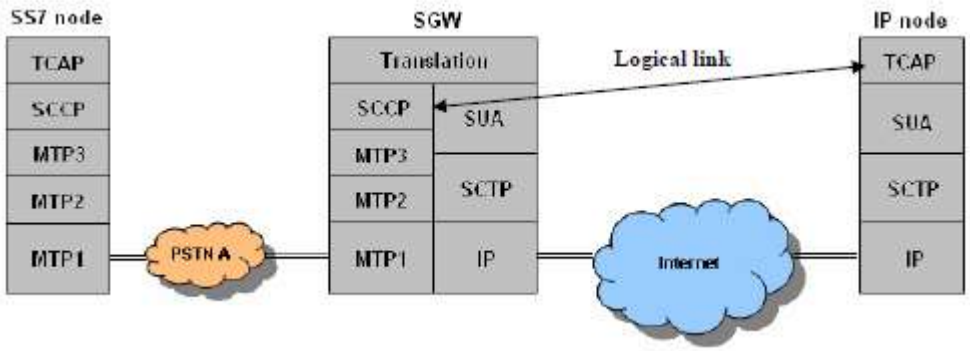
M3UA es una de las capas de adaptación que remueve la mayor cantidad de capas SS7 de los puntos de señalización y que cambia la topología de la red a una más parecida a IP. En una estructura basada en IP, M3UA no está restringida a los requisitos de SS7 del tamaño máximo del mensaje (272 *bytes*), pero puede usar un mayor ancho de banda el cual está disponible en una red IP. La flexibilidad de M3UA y su habilidad para hacer un mejor uso de redes IP y sus ventajas**,** la han llevado a ser el protocolo estándar en las redes UMTS.

4.5.4. SUA

Cuando se realiza una migración desde una red SS7, los operadores de la red IP deben mantener varias aplicaciones valiosas de las redes tradicionales como prepago y *roaming*.

El grupo de trabajo SIGTRAN hizo esto posible al definir la capa de adaptación SCCP-User (SUA), que no solamente provee estos servicios a la red IP, sino que elimina de la pila de SS7**,** que las otras capas de adaptación**;** ver figura 19**.** Además de utilizar rutas IP y ancho de banda de una manera más eficiente, los Nodos con SUA son más simples y por lo tanto más baratos que nodos con otras capas de adaptación.

Figura 19. **Interconexión vía SUA**



Fuente: Imonen, Mia. Signaling over IP. p. 16.

La tarea principal de la capa SUA es la de transferir información SCCP entre un SGW y un MGC (modelo cliente-servidor), y el mapear entre direcciones SCCP y direcciones IP en el SGW. Sin embargo, debido a la incapacidad de SUA para transportar mensajes ISUP, 3GPP ha escogido utilizar M3UA como el protocolo de señalización estándar en las partes centrales de las redes UMTS, y utilizar SUA como un complemento para nodos con base de datos, por ejemplo HLRs (Home Location Registers).

4.6. Seguridad

En la red de acceso en telefonía los protocolos de acceso son utilizados para señalización, y en la red *core* el protocolo SS7 es utilizado también para señalización. Las redes SS7 son por lo regular físicamente inaccesibles para el usuario final, así que se consideran protegidas de ataques, ya que el equipo de red se encuentra detrás de “puertas cerradas”. La red de acceso por otro lado, es utilizada para la señalización de los usuarios finales y aquí los puntos de seguridad son importantes. Los mayores peligros son atacantes pasivos, que

simplemente leen mensajes de la red, observan *passwords*, etc., junto con atacantes activos los cuales escriben, borran o modifican mensajes.

Algunos puntos de seguridad importantes son: autenticación de usuarios, integridad, confidencialidad de la información del usuario, evitar el uso inapropiado y no autorizado y DoS (Denial of Service). Todas las capas de adaptación de SIGTRAN utilizan SCTP para el transporte de información, lo cual provee algunas características de seguridad cómo su resistencia a ataques DoS (*flooding*, *masquerading* y monopolización de los servicios):

- *Cookies*: en el intercambio de 4-vias de SCTP las *cookies* son intercambiadas; esto previene que atacantes establezcan conexiones sin utilizarlas y de este modo bloquear que usuarios legítimos puedan establecer una conexión.

- *Verification Tag*: el encabezado SCTP contiene una etiqueta de verificación que indica si un paquete pertenece a cierta asociación. Si no pertenece a ninguna es descartado; esto protege al usuario de un ataque tipo *man-in-the-middle*.

Para proveer seguridad punto-a-punto entre dos redes remotas es recomendable que se utilice IPsec o TLS. Con IPsec se establece un túnel seguro de comunicación entre los dos puntos, lo cual es el equivalente a un link aislado como se utiliza en las redes SS7 tradicionales.

CONCLUSIONES

1. El estándar SIGTRAN fue originalmente desarrollado para el transporte de señalización sobre IP, pero a través de los años a encontrado nuevas áreas de aplicación.

2. SIGTRAN, al ser liberado del comportamiento de retransmisión extremadamente complejo de TCP, el protocolo SCTP puede confiadamente utilizarse para el transporte de Media sobre IP (MoIP).

3. El protocolo SIGTRAN es una nueva manera de transportar mensajes a través de IP. Ya que los servicios basados en IP se encuentran en constante expansión.

4. Las capacidades del protocolo SCTP y la versatilidad de SIGTRAN permiten tener una alternativa de convergencia entre las redes tradicionales de señalización SS7 y el protocolo IP.

5. Debido a la convergencia de dichas redes es posible para los implementadores (empresas, desarrolladores, etc.) optar a nuevas maneras de comunicación, hoy en día no solamente se tiene un protocolo que ofrece el soporte para el transporte de señalización sino que es capaz de competir contra un protocolo gigante del internet, TCP.

RECOMENDACIONES

1. Evaluar los sistemas de comunicación que se utilizan en la actualidad, su versatilidad y adaptación a los nuevos protocolos y estándares establecidos bajo SIGTRAN/SCTP.

2. Planificación del diseño de red, costos y beneficios que se obtendrá con la implementación de SIGTRAN.

3. Desarrollar nuevas áreas de negocios disponibles por la versatilidad del protocolo SCTP, incursionando en la generación, distribución y control de media sobre IP, generando de esta manera un cambio y ventaja en el mercado en los productos que se ofrecen y asegurando gracias a las características de SCTP e IP la calidad del servicio que se provee.

BIBLIOGRAFÍA

1. ALLMAN, M.; PAXSON, V.; STEVEN, W. R. *TCP congestion control,* RFC 2581. [en línea]. <http://www.ietf.org/rfc/rfc2851.txt>. [Consulta: 07 de junio 2012].

2. ARIAS, Ivan. *Stream control transmission protocol*. Helsinki, Finland: Helsinki University, 2002. 143 p.

3. BERNERS-LEE, T.; FIELDING, R. T.; FRYSTYK, H. *Hypertext transfer protocol – HTTP/1.0,* RFC 1945. [en línea]. <http://www.ietf.org/rfc/rfc1945.txt>. [Consulta: 07 de junio de 2012].

4. HINDEN, R. M.; DEERING, S. E. *IP Version 6 addressing architecture,* RFC2373. [en línea]. < http://www.ietf.org/rfc/rfc2373.txt >. [Consulta: 03 de julio de 2012].

5. IMMONEN, Mia. *Signaling over IP – a step closer to an all-IP network*. Stockholm, Sweden: IMIT/LCN, 2005. 38 p.

6. LOUGHNET, J. *Security considerations for signaling transport (SIGTRAN) Protocols*. [en línea]. <http://www.rfc-archive.org/getrfc.php?rfc=3788>. [Consulta: 12 de junio 2012].

7. NORTHRIDGE, Steve. *Convergent SS7 signaling for seamless service deployment.* [en línea]. <

http://www.ulticom.com/docs/SS7_Signaling_Convergence_White_Paper.pdf>. [Consulta: 15 de junio 2012].

8. TUXEN, Michael. SCTP/SIGTRAN & SS7 Overview. California, USA: Foothill Collegue, 2008. 26p.