

MEJORAMIENTO DEL TIEMPO DE RESPUESTA EN UNA NUEVA ARQUITECTURA DE MULTIDIFUSIÓN PARA AUDIOCONFERENCIAS BASADA EN EXTENSIONES SIP Y GESTIÓN DE GRUPOS

CARLOS MORENO¹ Y MARIBEL ALVAREZ²

¹Universidad Central de Venezuela. Escuela de Ingeniería Eléctrica. e-mail:carlos.moreno@ucv.ve

²Universidad Simón Bolívar. Decanato de Estudios Tecnológicos. e-mail:maribelalvarez@usb.ve

Recibido: octubre 2011

Recibido en forma final revisado: abril 2012

RESUMEN

Los servicios de grupo como la audioconferencia requieren de un nivel mínimo de calidad de servicio para la transmisión adecuada de paquetes multimedia, así como del mejoramiento de los procesos de señalización para el establecimiento, el mantenimiento y la liberación de sesiones de grupos de multidifusión. Tradicionalmente, los servidores SIP (Session Initiation Protocol) han trabajado con señalización de unidifusión, y en forma complementaria soportan multidifusión IP (Internet Protocol). En la carga útil con SDP (Session Description Protocol) se describen las pautas de distribución de la sesión de multidifusión IP. Como consecuencia de ello, los mismos inconvenientes de la multidifusión IP afectan el rendimiento de las audioconferencias basadas en solicitudes con protocolo SIP. El presente trabajo propone una nueva arquitectura basada en extensiones SIP y en un sistema de gestión de grupos. Tal arquitectura soporta multidifusión de sobrealineación tanto en el plano de datos, como en el plano de control. Las funciones de gestión de grupos son ejecutadas en un gestor de conferencias. Se montaron dos grupos de experimentos para medir el tiempo de adherencia y el tiempo de desincorporación tanto con multidifusión IP como con la propuesta de multidifusión SIP para varios niveles de tráfico extra de datos desde un potencial miembro de grupo hasta la fuente de tráfico multimedia. Los resultados demuestran que la diferencia entre la desviación estándar mínima y máxima del tiempo de adherencia y del tiempo de desincorporación para la arquitectura propuesta con multidifusión SIP es menor que en la multidifusión IP para los escenarios estudiados.

Palabras clave: Audioconferencia, Multidifusión, Señalización, SIP, Respuesta.

IMPROVEMENT OF THE RESPONSE TIME IN A NEW MULTICAST ARCHITECTURE FOR AUDIOCONFERENCES BASED ON SIP EXTENSIONS AND GROUP MANAGEMENT

ABSTRACT

Group services like the audioconference require a minimum level of quality of service for the adequate transmission of multimedia packets, and the improvement of signaling processes for the establishment, maintenance and release of multicast group sessions. Traditionally, the SIP (Session Initiation Protocol) servers have dealt with unicast signaling, and complementary support IP (Internet Protocol) multicast. The distribution patterns of the session are described on the SDP (Session Description Protocol) payload. As a consequence, the same existing drawbacks of IP multicast affect the performance of SIP-based requested audioconferences. This work proposes and improves a new multicast architecture based on SIP extensions and a group management system. Such architecture supports overlay multicast in both data and control plane. The group management functions are executed inside a conference manager. Two testbeds were mounted to measure the joining and leaving time for both IP multicast and SIP multicast with several levels of extra load traffic from a potential group member to the source of multimedia traffic. The results show that the difference between the minimum and maximum standard deviation value for both joining and leaving time on the architecture proposed with SIP multicast is lower than with IP multicast for the studied scenarios.

Keywords: Audioconference, Multicast, Signaling, SIP, Response.

INTRODUCCIÓN

Hoy en día el tráfico de voz, datos y video está siendo integrado en la misma plataforma de red. El diseño de una arquitectura debe tomar en cuenta las aplicaciones a ser desplegadas con un enfoque sin olvidar las capas altas del modelo TCP / IP (Winter, 2008). La audioconferencia entre grupos es una de las aplicaciones más demandadas, y su mejoramiento en los procesos de señalización se está realizando principalmente en la capa de aplicación y no en la capa IP. Un nivel adecuado en la calidad de servicio se necesita para el uso de esta aplicación, y ello conlleva a la disminución del tiempo de respuesta para adherir o desincorporar a un miembro de grupo además de incorporar la multidifusión (Williamson, 2001; Seifert, 2001). A diferencia de las implementaciones de arquitecturas de multidifusión basadas en IP (Minoli, 2008), aquéllas que se basan en la sobrealineación (Jannotti et al. 2000) se pueden desarrollar y mejorar sin las restricciones relacionadas con los protocolos de enrutamiento de multidifusión en capa IP. Esta es una de las razones por las cuales los servicios de televisión IPTV (Harte, 2008) no se han expandido con rapidez.

Una arquitectura de audioconferencias con gran cantidad de usuarios se puede implementar con software libre como SER (IPTEL, 2011), OPENSIPS (OPENSIPS, 2011) o ASTERISK® (Gómez et al. 2008); o software propietario como 3CX (3CX, 2011), en forma centralizada o descentralizada.

El protocolo SIP (Perea, 2008; Camarillo, 2002) es uno de los más utilizados para arquitecturas de audioconferencias basadas en unidifusión, y sus definiciones base (Schulzrinne et al. 2002) hacen referencia a la multidifusión IP (Hersent et al. 2004). Algunas extensiones que han surgido en IETF como (Rosenberg, 2006) permiten desarrollar soluciones para realizar audioconferencias con servidores y multidifusión IP. La nueva arquitectura que se propone se basa en una extensión del protocolo SIP hacia la multidifusión de sobrealineación (Moreno et al. 2009; Gao et al. 2006), introduciendo un mejoramiento en el tiempo de respuesta en la adherencia y desincorporación de miembros de grupo en relación con la predictibilidad del comportamiento. A gran escala, una mejor predicción del tiempo de respuesta resultará en innovaciones posteriores para calidad de servicio.

Algunos trabajos relacionados bibliográficamente con la presente investigación, se señalarán a continuación:

Chang et al. 2004

En la Universidad Nacional Changhwa de Taiwan, los investigadores sugieren la presencia de agentes especiales que conviertan mensajes del protocolo IGMP (Internet Group Management Protocol) en otros del protocolo SIP en un ambiente de red celular, y que se use el protocolo RSVP (Resource Reservation Protocol) para reservaciones. Desafortunadamente, el sistema se hace muy complejo con la movilidad (Chang et al. 2004).

Sbata et al. 2005

En Telecom SudParis, una arquitectura denominada mPnt (Multipoint Proxy Network) fue desarrollada para integrar SIP con la multidifusión IP en un servidor de audioconferencias. Soporta multidifusión de sobrealineación para la red de transporte únicamente con implementación propietaria y no en el acceso. Se compararon topologías en árbol expandidas, hipercubo y mixtas. SIP se puede usar sin extensiones nativamente, y el tiempo de respuesta depende de la topología interna así como de la implementación de la multidifusión IP (Sbata et al. 2005).

Chua & Pheanis (2006)

En la Arizona State University, se diseñó un sistema de teleconferencia para VoIP (Voice over IP) basado en un servidor con multidifusión IP. Esta solución sólo trabaja con “menrutadores” y con protocolos de gestión dinámica de grupos (Chua & Pheanis, 2006). No soporta el protocolo SIP.

EL PROTOCOLO SIP

El protocolo SIP permite el establecimiento, mantenimiento, liberación y modificación de las características de una sesión de voz, datos o video (PEREA, 2008). Se encuentra en la capa aplicativa del modelo TCP / IP y se basa en una interacción cliente-servidor entre agentes de usuario clientes y servidores. Se definen dos tipos de mensajes: las solicitudes y las respuestas. Una sesión se compone de un diálogo, el cual está a su vez conformado por transacciones. Cada transacción contiene dos o más mensajes. Es un protocolo orientado punto a punto principalmente y se han creado especificaciones adicionales acerca de mensajería instantánea, confiabilidad de respuestas provisionales, actualizaciones de descripción de la sesión y precondiciones para calidad de servicio.

LA MULTIDIFUSIÓN

La multidifusión es una tecnología de transmisión en redes de conmutación de paquetes en la que un flujo específico es enviado a un grupo de terminales (MINOLI, 2008). Esto es posible debido a que los puertos físicos de los equipos activos intermedios o los terminales finales están asociados con dichos grupos en forma dinámica o estática. Hay 3 tipos de multidifusión: 1) Multidifusión IP, 2) Multidifusión en capa aplicativa, y 3) Multidifusión de sobrealineación. En el primer tipo, un huésped se desincorpora o incorpora a un grupo enviando un mensaje de tipo IGMP (Internet Group Management Protocol) hacia el enrutador designado.

Este, en conjunto con otros enrutadores, forman árboles de distribución, intercambiando mensajes de acuerdo a un protocolo de enrutamiento de multidifusión.

En el segundo tipo, la membresía de grupos, la construcción de los árboles de multidifusión y el reenvío de los mensajes de datos es controlado únicamente por los terminales. El tercer tipo de multidifusión define la existencia de proxys o nodos de servicio, los cuales conforman la columna vertebral de sobrealineación y trabajan en conjunto con los terminales para construir la infraestructura y establecer los árboles de multidifusión.

SIP Y MULTIDIFUSIÓN

Las especificaciones del protocolo SIP establecidas inicialmente (Schulzrinne et al. 2002) han sido complementadas con otros RFCs acerca de mensajería instantánea (Campbell et al. 2002), confiabilidad de respuestas provisionales (Rosenberg & Schulzrinne, 2002), actualizaciones de descripción de sesión (Rosenberg, 2002), y precondiciones para calidad de servicio (Camarillo & Kyzivat, 2005) entre otros. Las definiciones de multifusión IP con SIP presentan fuertes limitaciones en el documento original de este último. La multidifusión IP puede ser usada sólo para un servicio de descubrimiento de un simple huésped, en el que se envía una solicitud a un grupo de servidores homogéneos, y sólo se procesa la respuesta de uno de ellos (Schulzrinne et al. 2002). Esta funcionalidad es mayormente usada para procesos de registro con múltiples servidores (Figura 1).

Basado en las reglas de procesamiento de transacciones, el cliente de este proceso aceptará la primera respuesta, y verá las otras como retransmisiones ya que todas contienen al mismo identificador Via (Camarillo, 2002). Otra limitación es que el cliente que envía la solicitud necesita formar parte del grupo de multidifusión al que envía la solicitud (Figura 1).

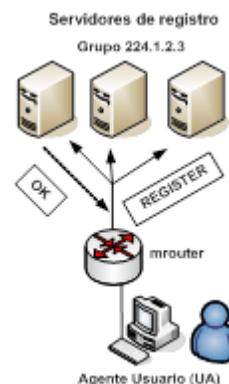


Figura 1. Registro con múltiples servidores

Algunos esfuerzos se han realizado desde el año 1998 para unir sesiones de audioconferencias con multidifusión IP y unidifusión basadas en SIP. En ese año, un artículo publicado por Schulzrinne y Rosenberg (Schulzrinne & Rosenberg, 1998) muestra cómo las audioconferencias podían establecerse. Estos autores sugieren que un nuevo participante de multidifusión IP tenga como primer paso la obtención de una dirección IP clase D. Luego debe enviar un mensaje INVITE a cada miembro de grupo de la audioconferencia indicando en la descripción SDP (Handley & Jacobson, 1998; Rosenberg & Schulzrinne, 2002) que desea enviar o recibir tráfico multimedia en el grupo de multidifusión. Otros terminales que soportan multidifusión IP responderán con un mensaje de clase 200, otros con un mensaje de clase 600. Esto permitiría que el nuevo miembro sepa de cuáles usuarios puede recibir datos de multidifusión y de cuáles a través de unidifusión.

Este enfoque presenta inconvenientes que derivan de la multidifusión IP en una LAN o WAN. Es necesario que el administrador de red instale o haga una migración a enrutadores especiales o “menrutadores” (Williamson, 2001) y a conmutadores que soporten protocolos dinámicos de gestión de grupos como IGMP snooping o CGMP (Minoli, 2008).

En los casos de audioconferencias centralizadas, un simple agente de usuario referido como “Focus” (Figura 2) mantiene un diálogo con cada participante y coordina procesos de señalización (Rosenberg, 2006).

El Focus (Figura 2) se encarga de manejar la señalización SIP (trabaja en el plano de control) y el Mixer coordina la transmisión de la información (trabaja en el plano de datos).

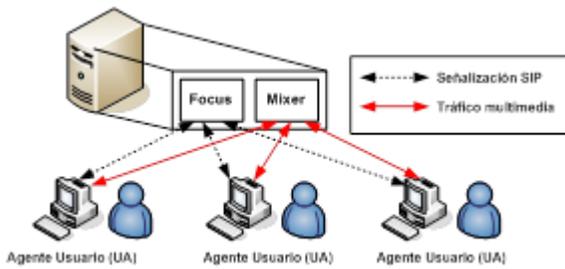


Figura 2. Focus en servidor de conferencias

ARQUITECTURA PROPUESTA

Se propone una nueva arquitectura para redes fijas basada en el protocolo SIP, con multidifusión de sobrealineación y gestión de grupos (Figura 3) que puede extenderse a comunicaciones con multidifusión inalámbrica en el futuro. Tal arquitectura contiene: (1) Extensor de multidifusión, (2) Un módulo especial denominado Agente Pasarela de Multidifusión (MGA - Multicast Gateway Agent) que interpreta las extensiones SIP para multidifusión y (3) Un gestor de multidifusión.

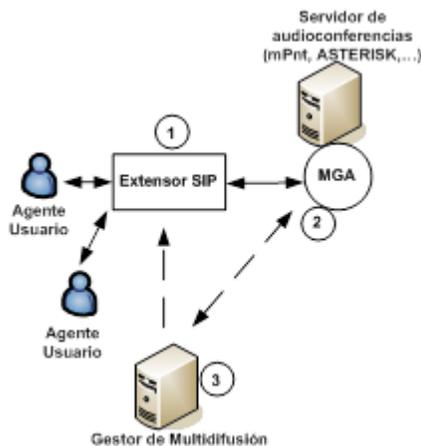


Figura 3. Arquitectura propuesta

Extensor SIP

Este componente permite trabajar con señalización SIP extendida normalmente hacia el servidor de audioconferencias, luego de añadir un encabezado SIP adicional que indica el uso de multidifusión de sobrealineación. También convierte mensajes extendidos con multidifusión SIP en mensajes convencionales hacia el agente usuario. Intercambia mensajes SIP extendidos de multidifusión con el módulo MGA, el cual se encuentra embebido en el servidor de audioconferencias.

Agente Pasarela de Multidifusión (MGA Multicast Gateway Agent)

Interpreta los mensajes extendidos de multidifusión SIP (Moreno et al. 2007) y responde al agente usuario. Llena y lee información de una tabla de multidifusión aplicativa que contiene información acerca de los agentes de usuario asociados con sus respectivos extensores SIP o de uno compartido entre ellos. El URI (Identificador Uniforme de Recurso) respectivo y una dirección IP identifican a un agente usuario; el ID de la audioconferencia es el URI que se encuentra en un nuevo campo de encabezado añadido que posee además el status registrado, incluido o excluido.

Gestor de multidifusión

Consiste en un programa que trabaja únicamente en el plano de control. Un operador humano puede decidir si un agente usuario específico solicitará su adherencia o desincorporación a una audioconferencia si se hace una configuración estática. Para eso, activa al módulo extensor SIP a través de un mecanismo de conexión con TCP. En modalidad automática puede obtener información de todos los servidores de audioconferencias que juegan un rol en la sesión utilizando una conexión bajo TCP con cada módulo MGA. Si hay varios servidores en la red WAN, puede calcular la topología más óptima. Cada período, el gestor de multidifusión interroga a los módulos MGA de los servidores de audioconferencias acerca del status de sus miembros de grupos (Figura 4). Mantiene su propia tabla con la mejor topología, la cual es determinada con el manejo de grafos y algoritmos genéticos.

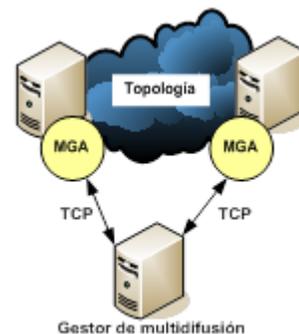


Figura 4. Interacción entre el Gestor de multidifusión y los MGAs

EXPERIMENTACIÓN

Escenarios planteados

Previamente a la experimentación, se visualizaron 2 escenarios.

En el primero, un huésped de multidifusión IP envía un mensaje al enrutador designado de tipo IGMP “Join” (si desea adherirse a un grupo) o IGMP “Leave” (si desea salirse de un grupo). El enrutador designado a su vez intercambia mensajes del protocolo PIM-SM (Protocol Independent Multicast – Sparse Mode) con la finalidad de que se cree un árbol, para distribuir el flujo multimedia que proviene de una fuente, que no pertenece al grupo. Se definió la raíz previamente cuando se configuraron los equipos.

En el segundo, un cliente extendido de multidifusión SIP envía un mensaje al MGA de tipo INVITE extendido (si desea adherirse a un grupo) o BYE extendido (si desea salirse de un grupo). El gestor de multidifusión da instrucciones al MGA (ambos contenidos en el mismo PC en este escenario) para que el tráfico multimedia que proviene de una fuente, sea o deje de ser reenviado al cliente extendido de multidifusión SIP. El gestor de multidifusión SIP posee información de la localización del cliente.

Realización de experimentos

Se realizaron 2 montajes. El primero contiene una arquitectura basada en multidifusión IP con el uso del programa cliente VideoLAN y el segundo posee un subconjunto de la arquitectura propuesta para su evaluación usando el programa cliente SJphone®. Ambos montajes fueron implementados en el laboratorio Percevale de Telecom SudParis (PERCEVALE, 2011) para realizar una comparación y tener una evaluación preliminar de la propuesta. Se realizaron 2 experimentos para cada montaje: en el primer montaje se configuró la multidifusión IP y se midió en el experimento #1 el tiempo de adherencia y en experimento #2 el tiempo de desincorporación; en el segundo montaje se implementó un subconjunto de la arquitectura nueva y se midieron en el experimento #3 el tiempo de adherencia y en el experimento #4 el tiempo de desincorporación.

Montaje #1: medición del tiempo de adherencia y desincorporación en escenarios con multidifusión IP

El montaje de la Figura 5 contiene 3 enrutadores modelo CISCO 3640, 2 conmutadores LAN modelo ENTERASYS 3820 y 2 computadores de escritorio.

Los enrutadores fueron configurados para enrutamiento de

unidifusión con protocolo RIPv2 (Routing Internet Protocol versión 2) y enrutamiento de multidifusión con protocolo PIM-SM (Protocol Independent Multicast – Sparse Mode) e IGMPv3 (Internet Group Management Protocol versión 3). El enrutador R2 fue seleccionado como el RP (Rendezvous Point).

El computador PC1 contiene la aplicación de software libre VideoLAN (lado cliente) y un generador de tráfico denominado TrafGen (TrafGen,2011) creado en Telecom SudParis. El PC2 contiene la aplicación VideoLAN (lado servidor) y TrafGen.

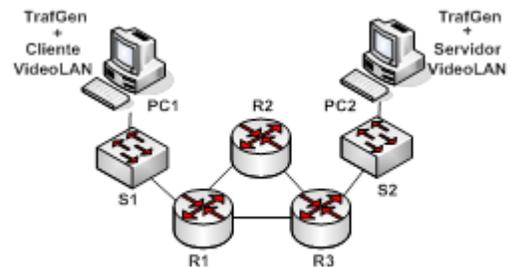


Figura 5. Componentes de montaje #1

Experimento #1: determinación del tiempo de adherencia en montaje #1

En este experimento, el software VideoLAN (lado servidor) en modo de multidifusión IP comienza a transmitir un flujo de tráfico de audio contenido en un archivo con formato MPEG1 hacia el grupo 239.255.12.42 y puerto 1234 de forma continua. Tal flujo es originado de PC2 (la fuente no pertenece al grupo 239.255.12.42). El software VideoLAN (lado cliente) recibe el flujo de multidifusión escuchando al puerto 1234 para el tráfico enviado a la dirección 239.255.12.42. PC1 envía un mensaje de adherencia IGMP (Join).

El tiempo de adherencia (T_{IN}) ha sido medido restando dos instantes: el primero (T₁) es cuando el primer mensaje (Join) es enviado por el PC1 al grupo 239.255.12.42, y el segundo (T₂) es cuando el primer paquete multimedia (UDP / RTP) es recibido por el huésped de multidifusión IP (PC1).

$$T_{IN} = T_2 - T_1 \quad (1)$$

Un tráfico de carga en la dirección del mensaje IGMP inicial es incrementado progresivamente con períodos de intersalida T con los siguientes valores :

Sin tráfico extra, $T=28$ ms , $T=21$ ms , $T=14$ ms y $T=7$ ms

Este tráfico extra pasa por los saltos R_1 y R_3 antes de alcanzar PC2, debido a que este flujo es de unidifusión, y por tanto se utiliza la tabla de enrutamiento de RIP.

Para producir el tráfico extra, se utilizó el software TrafGen. Se estableció un flujo CBR (Constant Bit Rate) de paquetes IP con carga útil 1024 bytes. Se realizó en sentido de receptor a emisor, para evaluar el efecto de la congestión en el sentido de la solicitud IGMP.

Experimento #2: determinación del tiempo de desincorporación en montaje #1

En este experimento, el software VideloLAN (lado servidor) transmite un flujo de audio con multidifusión de un archivo MPEG1 al grupo 239.255.12.42 en el puerto 1234 en forma continua. VideoLAN (lado cliente) recibe el flujo de multidifusión escuchando al puerto 1234 para dicho grupo. Luego hace una solicitud de desincorporación al grupo enviando un mensaje IGMP (leave).

El tiempo de desincorporación (T_{OUT}) se midió como la substracción de dos instantes: el primero (T_1) es cuando el primer mensaje IGMP (leave) es enviado de PC1 al grupo 239.255.12.42 y el segundo (T_2) es cuando el último paquete multimedia (UDP / RTP) es recibido por el cliente de multidifusión IP.

$$T_{OUT} = T_2 - T_1 \quad (2)$$

Al igual que en el experimento #1, se inyectó un tráfico de carga extra de paquetes IP con carga útil 1024 bytes con las mismas características. Se realizó en sentido de receptor a emisor, para evaluar el efecto de la congestión en el sentido de la solicitud IGMP.

Montaje #2: medición del tiempo de adherencia y desincorporación en escenarios con multidifusión SIP

El montaje de la Figura 6 contiene 3 enrutadores del modelo CISCO 3640, 3 conmutadores LAN modelo ENTERASYS y 3 computadores personales. Los enrutadores fueron configurados con protocolo de enrutamiento de unidifusión RIPv2 solamente. En la Figura 6, el agente de usuario (UA – User Agent) soporta multidifusión SIP (el módulo “SIP extensor está embebido”). Este manifiesta su intención de formar parte de un grupo enviando un mensaje SIP extendido al MGA. Varios servidores de audioconferencias con sus respectivos módulos MGA forman una topología que puede ser optimizada por algún algoritmo que se ejecuta

en el gestor de multidifusión. En este escenario preliminar, el MGA y el gestor de multidifusión están localizados en el mismo computador PC3.

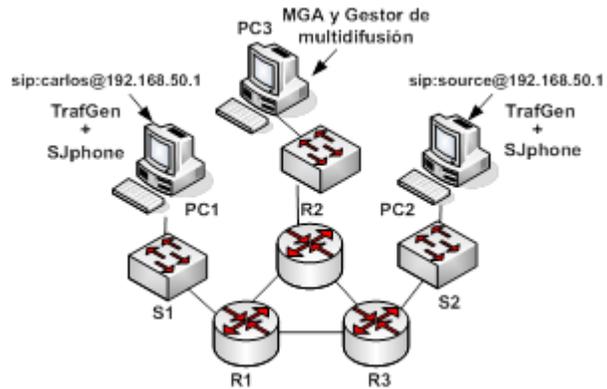


Figura 6. Componentes de montaje #2

El agente de usuario que se encuentra en el computador PC1 contiene la aplicación SJphone® (lado de cliente) en conjunto con un extensor SIP, y el software TrafGen que se utilizó en el montaje #1. El computador PC2 contiene también la aplicación SJphone® (lado de fuente) y TrafGen. Similarmente en el montaje #2, un tráfico de carga extra en la dirección del mensaje extendido SIP inicial es enviado con las mismas características. Este se origina en PC1 y termina en PC2 atravesando el camino R_1 y R_3 .

La carga se eleva progresivamente con periodos interlegadas T:

Sin tráfico extra, $T=28$ ms , $T=21$ ms , $T=14$ ms y $T=7$ ms.

Se seleccionaron los valores de T ya que en estas condiciones se produce un alto nivel de congestión en el sentido de una transacción INVITE o BYE extendida para multidifusión SIP, lo cual fue necesario para crear las condiciones en el estudio. Se evaluaron escenarios en los cuales las retransmisiones de los mensajes SIP con protocolo UDP fuesen posibles ya que SIP trabaja con éste comúnmente en la conexión agente usuario-proxy (Camarillo et al. 2003). Los mensajes INVITE son retransmitidos luego de un tiempo T_A (500 ms por defecto) si ninguna respuesta provisional o final llega en dicho intervalo de $64 \times T_1$ preferiblemente (Schulzrinne et al. 2002). Para las otras transacciones diferentes que INVITE, los agente usuarios retransmiten continuamente hasta que se recibe una respuesta final hasta un tiempo T_B . Particularmente, la retransmisión del mensaje ACK para terminar una transacción INVITE depende de la subcapa SIP núcleo y

no de la subcapa SIP transaccional del modelo de capas SIP si la respuesta final es de tipo OK. Las condiciones de congestión afectarán a los mensajes INVITE y ACK extendidos, al igual que los mensajes IGMP “join” para el tiempo de adherencia del montaje #1.

Igualmente, se afectará al tiempo de envío del mensaje BYE extendido así como ocurre con el mensaje IGMP “leave” del montaje #1.

Experimento #3: determinación del tiempo de adherencia en montaje #2

El tiempo de adherencia (T_{IN}) se midió como la diferencia entre 2 instantes: el primero (T_1) es cuando el mensaje SIP INVITE extendido es enviado por el computador PC1 al módulo MGA en PC3 y el segundo (T_2) es cuando el primer paquete multimedia (UDP /RTP) es recibido por el programa SJphone® extendido en PC1 (Figura 7).

$$T_{IN} = T_2 - T_1 \quad (3)$$

Se utilizó wireshark para medir los instantes T_1 y T_2 .

Un MGA toma la información que proporciona el gestor de multidifusión, como resultado de calcular la mejor topología para un grupo de MGAs. En este caso de estudio, la transmisión de la información entre MGA y el gestor de multidifusión ocurre dentro del computador PC3 por comunicación entre procesos.

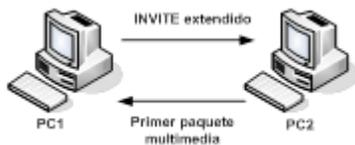


Figura 7. Medición de tiempo de adherencia en montaje #2

Experimento #4: determinación del tiempo de desincorporación en montaje #2

El tiempo de desincorporación (T_{OUT}) se midió como la diferencia entre 2 instantes: el primero (T_1) es cuando el mensaje SIP BYE extendido es enviado por el computador PC1 al módulo MGA en PC3, y el segundo (T_2) es cuando el último paquete multimedia (UDP /RTP) es recibido por el programa SJphone® extendido en PC1.

$$T_{OUT} = T_2 - T_1 \quad (4)$$

Al igual que en el experimento # 3, la transmisión de la información entre el MGA y el gestor de multidifusión ocurre dentro del computador PC3 por comunicación entre procesos.

RESULTADOS

Experimento #1: determinación del tiempo de adherencia en montaje #1

Se realizó diez veces este experimento para cada valor de tiempo intersalida T en la carga extra de tráfico. Los resultados mostrados en la Figura 8 indican que los valores del tiempo de adherencia T_{IN} están situados generalmente por debajo de los 100 ms. Hay dos excepciones, ambas en el caso de $T=14$ ms en el cual supera los 400 ms (442 ms y 546 ms). Esto es la consecuencia del comportamiento del tráfico IGMP hacia el enrutador designado R_1 (mensaje IGMP join) y de los mensajes del protocolo PIM intercambiados entre R_1 y el enrutador RP (Rendezvous Point) R_2 de la columna vertebral de la red de transporte (PIM join) con un nivel de calidad de servicio de mejor esfuerzo. No hay correlación entre el resultado de una muestra con un valor de $T=7$ ms y el resultado con una muestra equivalente para un valor diferente de T . Esto se valida por el comportamiento de los protocolos IGMP y PIM-SM (Minoli, 2008).

El objetivo del método que se usó para presentar los resultados en la Figura 8, fue para obtener información sobre la distribución de la data.

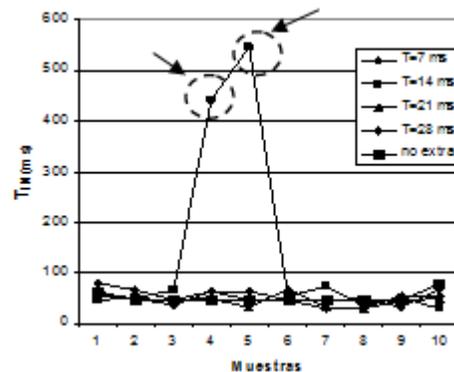


Figura 8. Tiempo de adherencia T_{IN} por muestra en montaje #1

En la Tabla 1 se muestra el promedio y la desviación estándar del tiempo de adherencia T_{IN} para cada valor de T .

Tabla 1. Promedio y desviación estándar de T_{IN}

| T (ms) | Promedio (ms) | Desviación estándar (ms) |
|----------|---------------|--------------------------|
| 7 | 47 | 11 |
| 14 | 139 | 189 |
| 21 | 45 | 12 |
| 28 | 56 | 14 |
| No extra | 51 | 11 |

Se puede observar que los 2 picos del tiempo de adherencia T_{IN} en la Figura 8 no ocurren en el caso de mayor carga de tráfico extra ($T=7$ ms), sino en un valor de menor carga ($T=14$ ms). Por lo tanto, no se puede decir que los picos se deben necesariamente a un alto nivel de carga extra. Sin embargo, se observa que la desviación estándar para $T=14$ ms en la Tabla 1 tiene el mayor valor (189 ms) y además del valor promedio mayor (139 ms).

Experimento #2: determinación del tiempo de desincorporación en montaje #1

Se realizó diez veces este experimento para cada valor de tiempo intersalida T en la carga extra de tráfico. Los resultados mostrados en la Figura 9 indican que los valores del tiempo de desincorporación T_{OUT} están situados alrededor de los 2000 ms. Sólo en un caso ($T=28$ ms de carga extra) el valor de T_{OUT} es mucho mayor (2499 ms) o mucho menor (1000 ms). Esto es la consecuencia del comportamiento del tráfico IGMP hacia el enrutador designado R_1 (mensaje IGMP Leave), el de mensajes del protocolo PIM (Leave) intercambiados entre R_1 y el enrutador RP (Rendezvous Point) R_2 de la columna vertebral de la red de transporte con un nivel de calidad de servicio de mejor esfuerzo y de la redundancia del tráfico multimedia. No hay correlación entre el resultado de una muestra con un valor de $T=7$ ms y el resultado con una muestra equivalente para un valor diferente de T al igual que en el experimento #1 y se valida por el comportamiento de los protocolos IGMP y PIM-SM (Minoli, 2008).

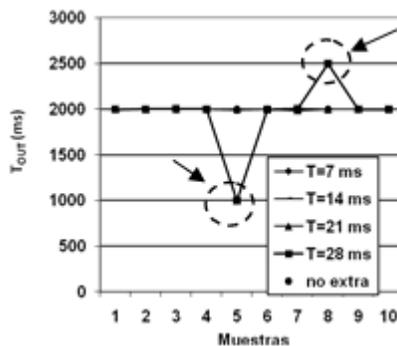


Figura 9. Tiempo de desincorporación T_{OUT} por muestra en montaje #1

En la tabla 2 se muestra el promedio y la desviación estándar del tiempo de desincorporación T_{OUT} para cada valor de T .

Tabla 2. Promedio y desviación estándar de T_{OUT}

| T (ms) | Promedio (ms) | Desviación estándar (ms) |
|----------|---------------|--------------------------|
| 7 | 1999 | 4 |
| 14 | 2000 | 2 |
| 21 | 1998 | 5 |
| 28 | 1949 | 369 |
| No extra | 2000 | 1 |

Se puede observar en la Figura 9 que el pico del tiempo de desincorporación T_{OUT} no ocurre en el caso de mayor carga de tráfico extra ($T=7$ ms), sino en un valor de poca carga extra ($T=28$ ms). Por lo tanto, no se puede decir que el pico se debe necesariamente a un alto nivel de carga extra. Adicionalmente, en la Tabla 2 se observa que los promedios para $T=14$ ms y en ausencia de carga extra tienen el mayor valor (2000 ms) y la desviación estándar para $T=28$ ms es la mayor (369 ms). La configuración de los enrutadores con los parámetros de intervalo de cuenta para última consulta en $n=2$ e intervalo de última consulta en 1 s para el enrutador designado R_1 y la redundancia del tráfico multimedia justifican el comportamiento. Durante 2 s el enrutador R_1 espera por la respuesta del último huésped (un mensaje adicional de grupo específico es enviado antes). El primer mensaje IGMP (leave) ha sido enviado al comienzo, y el último paquete multimedia ha sido recibido poco antes o después de 2 s en promedio. Hay 4 instantes que permiten describir el tiempo de desincorporación T_{OUT} en este escenario: el tiempo que tarda un mensaje IGMP (leave) para llegar al enrutador designado R_1 (T_{L1}), el tiempo que tarda el mensaje específico de grupo en transmitirse de R_1 al último huésped (T_{L2}) del grupo, el tiempo que tarda la confirmación final del último huésped en llegar al enrutador designado R_1 (T_{L3}) y el tiempo en que tarda el último mensaje multimedia en llegar de R_1 al huésped (T_{M4}) a partir de una referencia establecida. Si el flujo multimedia se transmite de forma continua desde la fuente PC2, entonces R_1 deja de enrutar paquetes multimedia una vez que pasa poco más de 2 s si sólo un huésped formaba parte del grupo de multidifusión. En el escenario real, había un huésped, y el flujo multimedia no es continuo debido a que posee redundancia que se aprovecha con la compresión. En consecuencia hay 2 subescenarios: el primero ocurrió si el último paquete multimedia fue enrutado por R_1 en el segundo intervalo, antes de 2 s, y por lo tanto el tiempo de respuesta sería:

$$T_{OUT} = T_{L1} + 1s + T_{M4} \quad (5)$$

Suponiendo que en el segundo intervalo de 1 s llegó el último paquete multimedia, T_{M4} tiene como referencia el instante $T_{L1} + 1$ s. No se considera T_{L3} debido a que sólo había un huésped, ni T_{L2} porque es despreciable a diferencia de T_{L1} que debe aumentar con una mayor carga de tráfico extra (gestión en un sentido).

En el segundo escenario se venció el intervalo de 2 s, y el tiempo adicional es T_{L1} , de modo que el tiempo de respuesta sería:

$$T_{OUT} = T_{L1} + 2s \quad (6)$$

Experimento #3: determinación del tiempo de adherencia en montaje #2

Se realizó el experimento diez veces para cada valor de T.

Los resultados mostrados en la Figura 10 muestran que el tiempo de adherencia T_{IN} varía entre 200 y 500 ms. Hay sólo dos excepciones; una cuando $T=21$ ms cuyo valor es 771 ms y otra sin tráfico extra cuyo valor es 664 ms. No se puede concluir que los picos ocurren necesariamente debido a la carga de tráfico extra. Sin embargo, el documento SIP base indica que al enviarse un mensaje SIP INVITE sobre protocolo de transporte UDP, pueden realizarse retransmisiones si durante un tiempo de T_A (500 ms por defecto) no hay una respuesta transitoria ni final. Se realizan hasta un máximo de seis retransmisiones en las cuales se duplica el tiempo de espera. En ambos picos los valores de T_{IN} están por debajo del máximo por defecto para una sola retransmisión (1500 ms) lo cual indica que posiblemente hubo al menos una retransmisión. Igualmente, puede haber retransmisiones del mensaje ACK para que el servidor se asegure de que la respuesta final OK llegó al agente usuario.

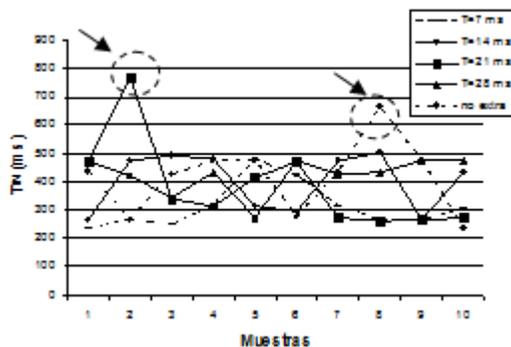


Figura 10. Tiempo de adherencia T_{IN} por muestra en montaje #2

Estas características en las retransmisiones se conservan para el manejo de mensajes de multidifusión SIP.

No hay correlación entre el resultado de una muestra de T_{IN} con un valor de $T=7$ ms y el resultado con una muestra equivalente para un valor diferente de T al igual que en los experimentos #1 y #2.

En la Tabla 3 se muestra el promedio y la desviación estándar del tiempo de adherencia T_{IN} para cada valor de T.

Tabla 3. Promedio y desviación estándar de T_{IN}

| T (ms) | Promedio (ms) | Desviación estándar (ms) |
|----------|---------------|--------------------------|
| 7 | 416 | 128 |
| 14 | 311 | 79 |
| 21 | 400 | 102 |
| 28 | 386 | 159 |
| No extra | 422 | 66 |

El máximo valor del promedio en T_{IN} ocurre en ausencia de tráfico extra, y el mayor valor de desviación estándar ocurre para $T=28$ ms.

El tiempo de adherencia está definido principalmente en este escenario por la interacción entre el cliente para multidifusión SIP extendida y el MGA asociado, debido a que este último se encuentra en el mismo computador del gestor de multidifusión.

Experimento #4: determinación del tiempo de desincorporación en montaje #2

Se realizó cinco veces el experimento para cada valor de T. Los resultados de la Figura 11 muestran que en algunos casos el valor del tiempo de desincorporación T_{OUT} posee valores negativos. Esto significa que el último paquete multimedia llegó antes del envío del mensaje SIP extendido. Esto es posible si el tráfico multimedia tiene gran redundancia y por unos instantes no se envía.

En dos casos se presentan picos de T_{OUT} : para $T=7$ ms y para $T=21$ ms para los cuales el valor de T_{OUT} es de 248 ms y 379 ms respectivamente. No se puede generalizar que a mayor carga de tráfico extra, existen picos. Sin embargo, el documento base de SIP indica que para los mensajes diferentes que INVITE, las retransmisiones se realizan hasta que los mensajes lleguen a su destino. En la arquitectura de multidifusión SIP propuesta, se conservan las propiedades de las retransmisiones con SIP nativo.

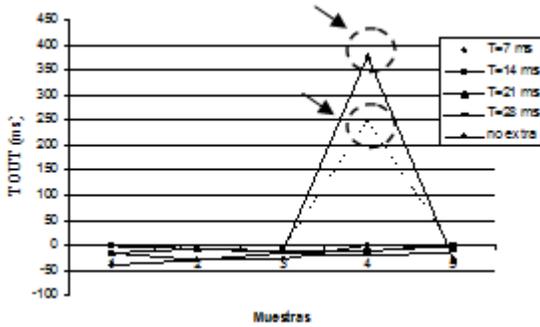


Figura 11. Tiempo de desincorporación T_{OUT} por muestra en montaje #2

En la Tabla 4 se muestra el promedio y la desviación estándar del tiempo de desincorporación T_{OUT} para cada valor de T .

Tabla 4. Promedio y desviación estándar de T_{OUT}

| T (ms) | Promedio (ms) | Desviación estándar (ms) |
|----------|---------------|--------------------------|
| 7 | 45 | 113 |
| 14 | -6 | 6 |
| 21 | 65 | 176 |
| 28 | -18 | 7 |
| No extra | -19 | 17 |

El máximo valor del promedio en T_{OUT} , y el mayor valor de desviación estándar ocurren para $T=21$ ms lo cual no constituye una condición de máxima carga extra. Sin embargo, para ambos valores de T se presenta el mayor valor de la desviación estándar de T_{OUT} .

El tiempo de desincorporación T_{OUT} está definido principalmente en este escenario por la interacción entre el cliente para multidifusión SIP extendida y el MGA asociado, debido a que este último se encuentra en el mismo computador del gestor de multidifusión. Esto de manera similar al experimento #3.

DISCUSIÓN

Al comparar el tiempo de adherencia T_{IN} en multidifusión IP vs multidifusión SIP de la Tabla 5 se puede observar que para la arquitectura propuesta con el escenario planteado, no hay una disminución. Inclusive los picos en multidifusión SIP poseen mayores valores.

Tabla 5. Validación para T_{IN} (ms)

| <i>Multidifusión IP</i> | <i>Multidifusión SIP</i> |
|------------------------------|---|
| Max: 81 | Max: 492 |
| Min: 30 | Min: 235 |
| Picos: 442 y 546 @ $T=14$ ms | Picos: 771 y 664 @ $T=21$ ms y no extra |

Particularmente, para el valor de $T=14$ ms existen 2 picos y la desviación estándar de T_{IN} es la mayor en el experimento #1. La probabilidad de que haya una retransmisión de un mensaje IGMP Join en un instante específico para multidifusión IP, tiene relación con el comportamiento de los instantes adyacentes para un mismo nivel de carga extra de tráfico en el escenario estudiado. En el caso de multidifusión SIP al haber una mayor carga de tráfico, no necesariamente hay un menor tiempo de respuesta; sin embargo, los picos se producen con un nivel de carga extra cuya desviación estándar en T_{IN} es 102 ms por encima de los valores de 66 y 79 ms correspondientes a ausencia de tráfico extra y de $T=14$ ms.

Al comparar el tiempo de desincorporación T_{OUT} en multidifusión IP vs multidifusión SIP de la Tabla 6 se puede observar que para la arquitectura propuesta con el escenario planteado, hay un menor tiempo de respuesta. Esto es posible con la configuración realizada en el enrutador designado R_1 .

Tabla 6. Validación para T_{OUT} (ms)

| <i>Multidifusión IP</i> | <i>Multidifusión SIP</i> |
|-------------------------|--------------------------|
| Max: 2003 | Max: 2 |

Adicionalmente, se observó anteriormente que el pico T_{OUT} ocurre para el nivel de carga extra en el que la desviación estándar correspondiente es mayor.

La mayor contribución de la arquitectura propuesta, con los escenarios estudiados está contemplada en la estabilidad del tiempo de respuesta al inyectarse la carga de tráfico extra. Si se consideran todas las muestras recolectadas para calcular la desviación estándar del tiempo de adherencia T_{IN} en la Tabla 7, se puede observar que la diferencia entre el mínimo y el máximo valor en multidifusión SIP es menor que la diferencia en multidifusión IP.

Tabla 7. Desviación estándar para T_{IN} (ms)

| <i>Multidifusión IP</i> | <i>Multidifusión SIP</i> |
|-------------------------|--------------------------|
| Max: 189 | Max: 159 |
| Min: 11 | Min: 66 |
| $\Delta=178$ | $\Delta=93$ |

Igualmente, si se consideran todas las muestras recolectadas para calcular la desviación estándar del tiempo de desincorporación T_{OUT} en la Tabla 8, se puede observar que la diferencia entre la mínima y la máxima en multidifusión SIP es menor que la diferencia en multidifusión IP.

Tabla 8. Desviación estándar para T_{OUT} (ms)

| <i>Multidifusión IP</i> | <i>Multidifusión SIP</i> |
|-------------------------|--------------------------|
| Max: 388 | Max: 176 |
| Min: 0,8 | Min: 6 |
| $\Delta=367$ | $\Delta=170$ |

En las Tablas 7 y 8 se demuestra que el comportamiento de la arquitectura con multidifusión SIP propuesta en los escenarios estudiados es más predecible que la equivalente con multidifusión IP, en relación con el tiempo de respuesta cuando se inyecta un tráfico de carga extra como el que se introdujo. Esto implica una mayor certeza a la hora de hacer un diseño de una red con multidifusión SIP para una empresa proveedora de servicios de VoIP considerando tales escenarios. Además, se demuestra que la arquitectura con multidifusión SIP presenta menos retransmisiones y mayor estabilidad que una arquitectura con multidifusión IP en los escenarios estudiados.

CONCLUSIONES

El protocolo SIP no soporta características nativas para un comportamiento de multidifusión, pero se apoya en una infraestructura con multidifusión IP.

La arquitectura propuesta permite el reuso de los métodos básicos del protocolo SIP para soportar señalización nativa de multidifusión, permitiendo el mejoramiento del tiempo de respuesta en algunos escenarios.

Tal arquitectura contiene 3 módulos: un extensor SIP que modifica las solicitudes y respuestas básicas, un MGA (Agente Pasarela de Multidifusión / Multicast Gateway Agent) que interpreta los mensajes extendidos SIP y procesa

los requerimientos, y un gestor de multidifusión para controlar y calcular la mejor topología de sobrealineación.

El gestor de multidifusión, como parte de la arquitectura de multidifusión SIP propuesta, permite optimizar la topología compuesta por los MGAs. También puede ser usada por un operador para definir una topología estática.

En los escenarios estudiados, la desviación estándar del tiempo de respuesta es directamente proporcional a la probabilidad de ocurrencia de retransmisiones en los escenarios estudiados.

La arquitectura con multidifusión SIP estudiada en los escenarios planteados permite predecir más fácilmente el comportamiento de los tiempos de respuesta, si se inyecta un tráfico de carga extra adicional, desde el receptor hasta la fuente, con incrementos preestablecidos.

Los tiempos de respuesta tal como fueron definidos tienen relación directa con los algoritmos de compresión del tráfico multimedia que emite la fuente.

El trabajo futuro se enfocará a estudiar diferentes escenarios con presencia de cajas intermedias (NAT y cortafuegos) y con 2 o más clientes de multidifusión SIP. Adicionalmente, se realizarán simulaciones a gran escala de la arquitectura.

AGRADECIMIENTOS

Los autores agradecen al Consejo de Desarrollo Científico y Humanístico (CDCH) de la Universidad Central de Venezuela por el financiamiento otorgado para la primera etapa del proyecto: “Mejoramiento de arquitectura para servicios multimedia basada en multidifusión SIP” bajo el número PI-08-7820-2009/1.

REFERENCIAS

- CAMARILLO, G., KANTOLA., R., SCHULZRINNE, H. (2003). Evaluation of Transport Protocols for the Session Initiation Protocol. Journal IEEE Network.17(5); p. 40-46.
- CAMARILLO, G. & KYZIVAT, P. (2005) Update to the Session Initiation Protocol (SIP) Preconditions Framework. IETF RFC 4032. Recuperado el 29 de Septiembre de 2011, de <http://www.ietf.org/rfc/rfc4032.txt>
- CAMARILLO, G. (2002). SIP Demystified, USA: Editorial: McGraw Hill; pp. 116-220.

- CAMPBELL, B., ROSENBERG, J., SCHULZRINNE, H., HUTTEMA., C., GURLE, D. (2002). Session Initiation Protocol (SIP) Extension for Instant Messaging. IETF RFC 3428. Recuperado el 29 de Septiembre de 2011, de <http://www.ietf.org/rfc/rfc3428.txt>
- CHANG, I., MAI, Y., CHANG, L. (2004). Hierarchical Multicast SIP Protocol for Streaming Multimedia in Wireless Networks. Proc. of IEEE International Conference on Networking, Sensing and Control Vol (1). Taipei, Taiwan; pp. 361-366.
- CHUA, T & PHEANIS, D.(2006). Innovative Server-Based Multicast VoIP Teleconference System. Proc. of 4th IASTED International Conference on Communications, Internet, and Information Technology. USA Islas Virgenes. pp. 533-060.
- GAO, H., PAPKA, N., STEVENS, R. (2006). Extending Multicast Communications by Hybrid Overlay Network. Proc. of 2006 IEEE International Conference on Communications, 2 (1); pp. 820-828.
- GÓMEZ, J. & MONTOYA, F. (2008). VoIP y ASTERISK. Redescubriendo la telefonía. USA: Editorial Ra-Ma; pp. 1-350.
- HANDLEY, M. & JACOBSON, V. (1998). Session Description Protocol. IETF RFC 2327. Recuperado el 29 de Septiembre de 2011, de <http://www.ietf.org/rfc/rfc2327.txt>
- HARTE, L. (2008). IPTV Testing, USA: Editorial Althos; pp.77-79
- HERSENT, D., GURLE, D., PETIT, J.P. (2004). La Voix sur IP. Codecs, H.323, SIP, MGCP, deployment et dimensionnement. Paris: Editorial DUNOD. pp-137-235.
- JANNOTTI, J., GIFFORD, D., JOHNSON, K., KAASHOEK, M., O'TOOLE J. (2000). Overcast: Reliable Multicasting with an Overlay Network," Proceedings of. 4th Symp. Operating Systems Design and Implementation (OSDI); pp-201-234.
- MINOLI, D. (2008). IP Multicast with Applications to IPTV and Mobile DVB-H. USA: Editorial Willey-Interscience; pp. 1-124.
- MORENO, C., BECKER, M., VINCENT, P. (2007). An Improved Multicast Network Architecture for Multimedia Services with SIP Extensions. Proc. of the Second International Conference on the Latest Advances in Networks ICLAN2007. Paris, Francia; pp. 60-64.
- MORENO, C., BECKER, M., VINCENT, P. (2009). Performance of a new Multicast Architecture based on SIP Extensions and Group Management. Proc. IEEE/IFIP 6th Latin American Network Operations and Management Symposium LANOMS2009, Punta del Este, Uruguay. pp. 1-10.
- PEREA, R. (2008). Internet Multimedia Communications Using SIP. USA: Editorial Morgannn Kauffma; pp. 1-550.
- ROSENBERG, J. & SCHZULRINNE, H. (2002). An Offer/Answer Model with the Session Description Protocol (SDP). IETF RFC 3264. Recuperado el 29 de Septiembre de 2011, de <http://www.ietf.org/rfc/rfc3264.txt>
- ROSENBERG, J. & SCHZULRINNE, H. (2002). Reliability of Provisional Responses in the Session Initiation Protocol (SIP). IETF RFC 3262. Recuperado el 29 de Septiembre de 2011, de <http://www.ietf.org/rfc/rfc3262.txt>
- ROSENBERG, J. (2002). The Session Initiation Protocol (SIP) UPDATE Method, IETF RFC 3311. Recuperado el 29 de Septiembre de 2011, de <http://www.ietf.org/rfc/rfc3311.txt>
- ROSENBERG, J. (2006). A Framework for Conferencing with the Session Initiation Protocol (SIP). IETF RFC 4353. Recuperado el 29 de Septiembre de 2011, de <http://www.ietf.org/rfc/rfc4353.txt>
- SBATA, K., BENAINI, R., VINCENT, P. (2005). Comparative study of MPNT Topological Models. Proc. on Systems Communications; 98-103.
- SCHZULRINNE, H. & ROSENBERG, J. (1998). Signaling for Internet Telephony. Proc. of IEEE Sixth International Conference on Network Protocols. USA, Austin. pp-298-307.
- SCHZULRINNE, H., CAMARILLO, G., JOHNSTON, A., PETERSON, J., SPARKS, R., HANDLEY, M., SCHOOLER, E. (2002). SIP: Session Initiation Protocol. IETF RFC3261. Recuperado el 29 de Septiembre de 2011, de <http://www.ietf.org/rfc/rfc3261.txt>
- Seifert, R. (2001). The Complete Guide to LAN Switching Technology, USA: Editorial Willey; pp 201-353.

WILLIAMSON, B. (2001). Developing IP Multicast Networks. Volume I. Indianapolis: Editorial CISCOPRESS; pp. 121-230.

WINTER, R. (2008). Towards a Unified Internet Control Architecture(2008). Electronics Proceedings of EUROVIEW2008 (Würzburg, Alemania). Recuperado el 29 de Septiembre de 2011, de <http://euroview2008.com/data/slides/25-presentation-winter.pdf>

REFERENCIAS DE ARQUITECTURAS

IPTTEL (2011). Official Site of SER Project, Fraunhofer FOKUS. Recuperado el 29 de Septiembre de 2011, de <http://www.iptel.org/ser/>

OPENSIPS (2011). The Scientifical Community, Official Site of Project OPENSIPS Project (Former OPEN SER). Recuperado el 29 de Septiembre de 2011, de <http://www.opensips.org/index.php?n=Development.Development>

PERCEVALE (2011). Plate-forme d'Expérimentation Recherche en Communication d'Évry Val d'Essonne en Telecom SudParis. Recuperado el 29 de Septiembre de 2011, de <http://percevale.telecom-sudparis.eu/en/>

TRAFGEN (2011). Generador de tráfico UDP TCP diseñado en Departamento LOR de Telecom SudParis. Recuperado el 29 de Septiembre de 2011, de <http://www-lor.int-evry.fr/~vincent/java/trafGen/trafGenEn.htm>

3CX (2011). 3CX Software based PBX for Windows. Recuperado el 29 de Septiembre de 2011, de <http://www.3cx.com/>

