



**CENTRO DE INVESTIGACIÓN Y DE ESTUDIOS
AVANZADOS DEL INSTITUTO POLITÉCNICO
NACIONAL**

UNIDAD ZACATENCO

Departamento de Ingeniería Eléctrica

Sección de Comunicaciones

**TÉCNICAS PARA MEJORAR EL DESEMPEÑO EN
REDES SOBREPUESTAS *PEER-TO-PEER*.**

TESIS

Que presenta:

Carlos Tonatiuh Juárez Campos

Para obtener el grado de:

MAESTRO EN CIENCIAS

Directores de Tesis:

Dr. Domingo Lara Rodríguez

Dr. Raúl García Ruiz

México, Ciudad de México

Febrero 2016

A toda mi familia, especialmente a

Ofelia y Dulce.

Agradecimientos

A mi madre Ofelia por todo tu esfuerzo y dedicación, no puedo imaginarme un apoyo más grande que el que me has brindado siempre. A mi hermana Dulce por todas las experiencias compartidas a lo largo de mi vida, tu optimismo siempre me impulsa a seguir adelante.

Al resto de mi familia, especialmente a mi tía Rayo y a mis tíos Paco y Daniel por apoyarme toda mi vida y motivarme a seguir creciendo. A mi tío Juan Carlos por haberme brindado un hogar sin esperar nada a cambio, siempre lo tendré presente.

A Isis por brindarme su amor y comprensión todos y cada uno de los días de la maestría y a su familia por todo el apoyo brindado siempre.

A mis asesores Dr. Domingo Lara y Dr. Raúl García por haber creído en mi desde el primer momento y brindarme la oportunidad de trabajar y aprender de ustedes.

Al Centro de Investigación y Estudios Avanzados del I.P.N. por permitirme haber sido parte de tan respetable institución. A los investigadores de la Sección de Comunicaciones porque con sus clases enriquecieron no solamente mi formación académica, sino mi formación personal. Al personal de la Sección de Comunicaciones por las facilidades brindadas a lo largo de mi estancia.

Al Consejo Nacional de Ciencia y Tecnología (CONACyT) por apoyarme económicamente durante el transcurso de la maestría.

A mis compañeros de maestría: Alavid, Ángel, Felipe, Gerardo y Víctor, porque en ustedes encontré a mi segunda familia. A mis amigos que siempre me apoyaron. A Rebeca por su compañía a lo largo de todo este camino, por no dejarme dar por vencido en ningún momento.

Resumen

En años recientes se ha presentado un crecimiento acelerado en el número de usuarios de Internet, por lo cual surge la necesidad de crear aplicaciones que conecten de manera simultánea a un gran número de usuarios. Los sistemas P2P (*Peer-to-Peer*) pueden cubrir este tipo de necesidades.

Para realizar la comunicación entre usuarios, el sistema P2P establece una red lógica, denominada *red sobrepuesta*, la cual se construye por encima de una red existente, denominada *red subyacente*. Debido al número de usuarios que puede soportar este tipo de sistemas, se requiere que el sistema cuente con un diseño de red sobrepuesta P2P con un buen desempeño en la entrega de mensajes a través de la red.

En este trabajo de tesis se presenta un estudio de diversas técnicas presentes en la literatura para mejorar el desempeño de redes sobrepuestas para su uso en sistemas P2P. Se identifican dos técnicas por medio de las cuales es posible mejorar el desempeño de redes sobrepuestas: la construcción de redes sobrepuestas conscientes de la topología subyacente y la construcción de redes sobrepuestas jerárquicas.

Posteriormente, se realiza la evaluación del desempeño de redes sobrepuestas conscientes de la topología subyacente, las cuales son adaptaciones del protocolo Chord propuestas en el presente trabajo que presentan una mejora en el desempeño de la red. Asimismo, se analiza el impacto de utilizar una elección de nodos marca propuesta en el desempeño de las construcciones realizadas. Finalmente, se evalúa una construcción de red sobrepuesta jerárquica consciente de la topología subyacente que utiliza la metodología propuesta.

Abstract

In recent years there have been an increased growth in the amount of users in Internet, which has made clear the necessity to develop large scale applications that face the challenge of connecting multiple users communicating in real time. Overlay networks, particularly Peer-to-Peer systems, fulfill this necessity.

In order to perform the communication between peers, the system establishes a logic topology denominated *overlay network*, which is built above the *underlying physical topology*. Because of the large number of users the system has to support, an efficient message delivery procedure has to be developed in the overlay network design.

In this work, a study of the different techniques that improve the performance of structured overlay networks is presented. From these techniques, two of them are reviewed: topology-aware overlay networks and hierarchical overlay networks.

Two topology-aware overlay networks are presented, which are Chord protocol adaptations proposed in this work. Additionally, the evaluation of the performance of these proposals is also reviewed. Moreover, the impact of the landmark nodes selection in the performance of the network is also presented. In particular, the hierarchical topology-aware overlay network performance is studied; this construction employs the proposed methodology.

Lista de Contenido

Resumen	7
Abstract	9
Lista de Contenido	11
Lista de Figuras	15
Introducción	19
Capítulo 1. Redes sobrepuestas P2P	23
1.1 Tecnología de redes sobrepuestas	23
1.1.1 Definición	24
1.1.2 Ventajas y limitaciones	27
1.1.3 Aplicaciones	28
1.2 Redes sobrepuestas P2P	30
1.2.1 Definición	30
1.2.2 Aplicaciones	32
1.2.3 Clasificación	35
1.3 Redes sobrepuestas P2P estructuradas	35
1.3.1 Tablas Hash Distribuidas	37
1.3.2 Ejemplos	40
1.3.3 Comparación	43
Referencias	47
Capítulo 2. Evaluación del protocolo Chord	51
2.1 Protocolo Chord	51
2.1.1 Topología de red	52

2.1.2	Búsqueda de una llave	54
2.1.3	Unión y abandono de la red	57
2.1.4	Longitud promedio de trayectoria	59
2.2	Simulación de la red subyacente	62
2.2.1	Generadores de topologías de red	62
2.2.2	Modelo Transit-Stub	64
2.3	Desempeño del protocolo Chord sobre una red subyacente	67
	Referencias	71
Capítulo 3. Redes sobrepuestas P2P conscientes de la topología subyacente		73
3.1	Introducción	73
3.2	Técnicas para generar información de proximidad en la red	75
3.3	Técnicas para introducir información de proximidad en la red	78
3.3.1	Enrutamiento basado en proximidad	78
3.3.2	Selección de nodos vecinos basados en proximidad	79
3.3.3	Disposición geográfica	79
3.4	Construcción de redes sobrepuestas conscientes de la topología subyacente	80
3.4.1	Construcción de red sobrepuesta consciente de la topología subyacente simple (RS-CTS-S)	81
3.4.2	Construcción de red sobrepuesta consciente de la topología subyacente mejorada (RS-CTS-M)	85
3.5	Elección de nodos marca	90
3.5.1	Optimización de la elección de nodos marca	92
3.5.2	Comparación de desempeño	93
	Referencias	96
Capítulo 4. Redes sobrepuestas P2P jerárquicas		99
4.1	Introducción	99
4.2	Clasificación	101
4.3	Construcción de una red sobrepuesta jerárquica (RSJ)	104
4.3.1	Análisis del número promedio de saltos	107

4.4	Construcción de una red sobrepuesta jerárquica consciente de la topología subyacente (RSJ-CTS)	110
4.5	Elección de los supernodos	114
	Referencias	116
	Conclusiones y Trabajo Futuro	119

Lista de Figuras

Figura 1.1. Ejemplo de una organización de los nodos en: A) Red sobrepuesta, B) Red subyacente. _____	25
Figura 1.2. Estructura de una red sobrepuesta organizada por capas en comparación con el modelo de Internet. _____	26
Figura 1.3. Estructura de anillo Chord. _____	41
Figura 1.4. Estructura de un espacio de 2 dimensiones en CAN. _____	42
Figura 2.1. Ejemplo de la asignación de llaves a nodos en el anillo Chord. _____	53
Figura 2.2. Ejemplo del enrutamiento de una búsqueda a través de los apuntadores a sucesores. _____	55
Figura 2.3. Ejemplo de la construcción de una tabla de dedos en Chord. _____	56
Figura 2.4. Ejemplo del enrutamiento de una búsqueda a través de la tabla de dedos. ____	57
Figura 2.5. Número de saltos promedio en Chord como función del tamaño de la red. _	61
Figura 2.6. Ejemplo de una estructura de red conformada por dominios aislados y dominios de tránsito. _____	65
Figura 2.7. Ejemplo de una topología de red de 100 nodos creada con el generador GT-ITM. _____	67
Figura 2.8. Retardo promedio de búsqueda de Chord sobre dos tipos de redes subyacentes. _____	69
Figura 2.9. Factor de estiramiento entre la topología de red sobrepuesta Chord y la topología de la red subyacente. _____	70
Figura 3.1. Ejemplo de agrupamiento de nodos utilizando nodos marca. _____	77

Figura 3.2. a) Ejemplo de la división de una topología de red en $M=4$ zonas. b) Ejemplo de la división del anillo Chord en $M=4$ arcos. _____	82
Figura 3.3. Comparación del desempeño del protocolo Chord tradicional y la construcción RS-CTS-S en base a: a) Número promedio de saltos por búsqueda, b) Tiempo promedio de búsqueda y c) Factor de estiramiento. _____	84
Figura 3.4. a) Ejemplo de la división de una topología de red en $M!=24$ zonas. b) Ejemplo de la división del anillo Chord en $M!=24$ arcos. _____	87
Figura 3.5. Comparación del desempeño del protocolo Chord tradicional, la construcción RS-CTS-S y la construcción RS-CTS-M en base a: a) Número promedio de saltos por búsqueda, b) Tiempo promedio de búsqueda y c) Factor de estiramiento. _____	89
Figura 3.6. Comparación del desempeño de la RS-CTS-M con tres diferentes conjuntos de marcas en base a: a) Número promedio de saltos por búsqueda, b) Tiempo promedio de búsqueda y c) Factor de estiramiento. _____	94
Figura 4.1. Arquitectura de red sobrepuesta jerárquica de dominio único. _____	102
Figura 4.2. Arquitectura de red sobrepuesta jerárquica de dominios múltiples. _____	103
Figura 4.3. Arquitectura de red sobrepuesta jerárquica. _____	104
Figura 4.4. Identificador jerárquico de dos niveles. _____	105
Figura 4.5. Comparación del desempeño de la construcción RSJ y su equivalente en arquitectura sin jerarquía en base a: a) Número promedio de saltos por búsqueda, b) Tiempo promedio de búsqueda y c) Factor de estiramiento. _____	106
Figura 4.6. Número promedio de saltos en la de red sobrepuesta jerárquica en función del número de dominios. _____	109
Figura 4.7. Arquitectura de red sobrepuesta jerárquica consciente de la topología subyacente. a) Ejemplo de la división de una topología de red en $M=4$ zonas. b) Ejemplo de la arquitectura de red jerárquica con $M=4$ dominios y un dominio de interconexión. ____	112
Figura 4.8. Comparación del desempeño de la construcción RSJ-CST, la construcción RSJ y su equivalente en arquitectura sin jerarquía en base a: a) Número promedio de saltos por búsqueda, b) Tiempo promedio de búsqueda y c) Factor de estiramiento. _____	113

Figura 4.9. Comparación del desempeño de la RSJ-CTS con tres diferentes conjuntos de marcas en base a: a) Número promedio de saltos por búsqueda, b) Tiempo promedio de búsqueda y c) Factor de estiramiento. _____ 115

Introducción

En años recientes se ha presentado un crecimiento acelerado en el número de usuarios de Internet y se estima que el porcentaje de usuarios a finales del año 2015 es de alrededor del 43% de la población mundial [1]. Con este crecimiento en el número de usuarios surge la necesidad de aplicaciones que conecten de manera simultánea a un gran número de usuarios permitiendo la compartición de recursos, la distribución masiva de información y la comunicación en tiempo real entre usuarios. Los sistemas de redes sobrepuestas, especialmente los sistemas P2P (*Peer-to-Peer*), son sistemas que pueden cubrir este tipo de necesidades y por lo tanto se han convertido en áreas activas de investigación y desarrollo.

Los sistemas de redes sobrepuestas P2P utilizan una arquitectura descentralizada [2], es decir, no se tiene una dependencia de servidores dedicados que se encuentren disponibles en la red en todo momento; en su lugar existen dispositivos terminales, denominados *pares*, los cuales se conectan de manera intermitente a la red y pueden comunicarse de manera directa entre ellos. Para realizar la comunicación entre pares, el sistema P2P establece una red lógica, denominada *red sobrepuesta*, la cual se encuentra por encima de una red existente, denominada *red subyacente*. De esta manera, los enlaces en la red sobrepuesta no son enlaces físicos, sino enlaces estrictamente virtuales entre los pares. Por lo tanto, un solo enlace en la red sobrepuesta puede utilizar varios enlaces físicos de la red subyacente.

Debido al gran número de usuarios que puede atender este tipo de sistemas, un diseño ineficiente de la red sobrepuesta puede provocar que el tráfico generado por los usuarios del sistema circule sobre un número redundante de saltos físicos en la red subyacente, ocasionando de esta manera un volumen grande de tráfico adicional en la red. Por lo tanto, se requiere que el sistema cuente con un diseño de red sobrepuesta P2P que presente un desempeño aceptable en la entrega de mensajes a través de la red y de esta manera evitar

afectar la operación de otras aplicaciones que se ejecuten en la infraestructura de la red subyacente.

El objetivo principal del presente trabajo de tesis es realizar un estudio de diversas técnicas presentes en la literatura para mejorar el desempeño en redes sobrepuestas para su uso en sistemas P2P, así como evaluar su desempeño en una red basada en el protocolo Chord.

La organización del trabajo es la siguiente:

- En el Capítulo 1 se introducen los conceptos básicos de las redes sobrepuestas P2P, haciendo énfasis en las redes sobrepuestas estructuradas basadas en una tabla *hash* distribuida. Asimismo, se presenta una comparación entre tres protocolos distintos de redes sobrepuestas: los protocolos Chord, CAN y Pastry.
- En el Capítulo 2 se presenta el protocolo Chord, y se estudian sus propiedades de enrutamiento de búsquedas en la red sobrepuesta. Posteriormente, se presenta la metodología para simular el protocolo Chord sobre una red subyacente y evaluar su desempeño por medio de las siguientes métricas: número promedio de saltos por búsqueda, tiempo promedio de búsqueda y factor de estiramiento.
- En el Capítulo 3 se estudian los conceptos básicos de la construcción de redes sobrepuestas conscientes de la topología subyacente. Con este fin, se presentan diversas técnicas para generar e introducir información de proximidad en la red. Posteriormente, se realiza la evaluación del desempeño de dos tipos de construcciones de redes sobrepuestas conscientes de la topología subyacente, las cuales son adaptaciones del protocolo Chord propuestas en el presente trabajo y presentan una mejora en el desempeño de la red en comparación con el protocolo Chord tradicional.
- En el Capítulo 4 se presentan los fundamentos básicos de redes sobrepuestas jerárquicas. Asimismo, se presenta una clasificación de este tipo de redes de acuerdo a su arquitectura de red. Para evaluar el desempeño de este tipo de redes se realiza la construcción de una red jerárquica que utiliza como base el protocolo Chord. Posteriormente, se realiza la

construcción de una red jerárquica consciente de la topología subyacente, la cual introduce la información de proximidad de la red subyacente utilizando las propuestas presentadas en el Capítulo 3.

- En el apartado “Conclusiones y Trabajo Futuro” se presentan las conclusiones del trabajo desarrollado en esta tesis y se presenta el posible trabajo futuro respecto al tema desarrollado.

Capítulo 1

1. Redes sobrepuestas P2P

En el presente capítulo se introducen los conceptos básicos de las redes sobrepuestas entre nodos con capacidades similares de procesamiento (P2P por *Peer-to-Peer*), haciendo énfasis en las redes sobrepuestas estructuradas basadas en una tabla *hash* distribuida. Asimismo, se presenta una comparación entre tres protocolos distintos de redes sobrepuestas.

1.1 Tecnología de redes sobrepuestas

El uso de redes sobrepuestas antecede a la aparición de aplicaciones P2P. Un ejemplo de una red sobrepuesta ampliamente utilizada para implementar servicios de red son los protocolos utilizados en servidores de correo electrónico en la Internet [3]. Sin embargo, estos primeros tipos de redes sobrepuestas especializadas utilizan protocolos específicamente diseñados, y están orientadas a interconectar servidores de infraestructura en lugar de dispositivos terminales. Con el paso del tiempo las redes sobrepuestas han evolucionado para satisfacer una variedad de necesidades presentes en las redes actuales. A continuación se presentan los conceptos básicos de las redes sobrepuestas.

1.1.1 Definición

Establecer una definición universal de una red sobrepuesta que englobe cada uno de sus aspectos es una tarea complicada debido a su constante evolución. Sin embargo, es posible identificar un conjunto de características mencionadas en la literatura [4] [5], las cuales conforman la estructura básica de una red sobrepuesta.

Una red sobrepuesta es una red que se construye por encima de una red existente, denominada red subyacente. De esta manera la red sobrepuesta utiliza los servicios disponibles de la red subyacente para funciones básicas de red (enrutamiento y reenvío). Por otra parte, la red sobrepuesta proporciona conectividad, enrutamiento e intercambio de mensajes entre dispositivos finales a nivel de capa de aplicación.

Los dispositivos terminales en la red sobrepuesta, comúnmente denominados *nodos*, se encuentran conectados entre sí por medio de enlaces lógicos, los cuales pueden extenderse a través de varios enlaces físicos. Por esta razón, un salto entre dos nodos pertenecientes a la red sobrepuesta puede necesitar varios saltos en la red subyacente. La Figura 1.1 muestra la organización de una red sobrepuesta en comparación con la ubicación de los nodos en la red subyacente, donde se puede observar que la red sobrepuesta crea enlaces lógicos que no se reflejan directamente en enlaces directos en la red subyacente.

Las redes sobrepuestas pueden ser utilizadas para superar algunas de las limitaciones de la red subyacente y permiten la introducción de funcionalidades de red que no se encuentran disponibles en las funciones de enrutamiento básicas del protocolo IP (*Internet Protocol*), el cual es el protocolo utilizado en la Internet. Algunas extensiones en las funcionalidades que se pueden proporcionar incluyen al manejo de la movilidad de los nodos, el enrutamiento personalizado, calidad de servicio, nuevos tipos de direccionamiento y mejoras en la seguridad.

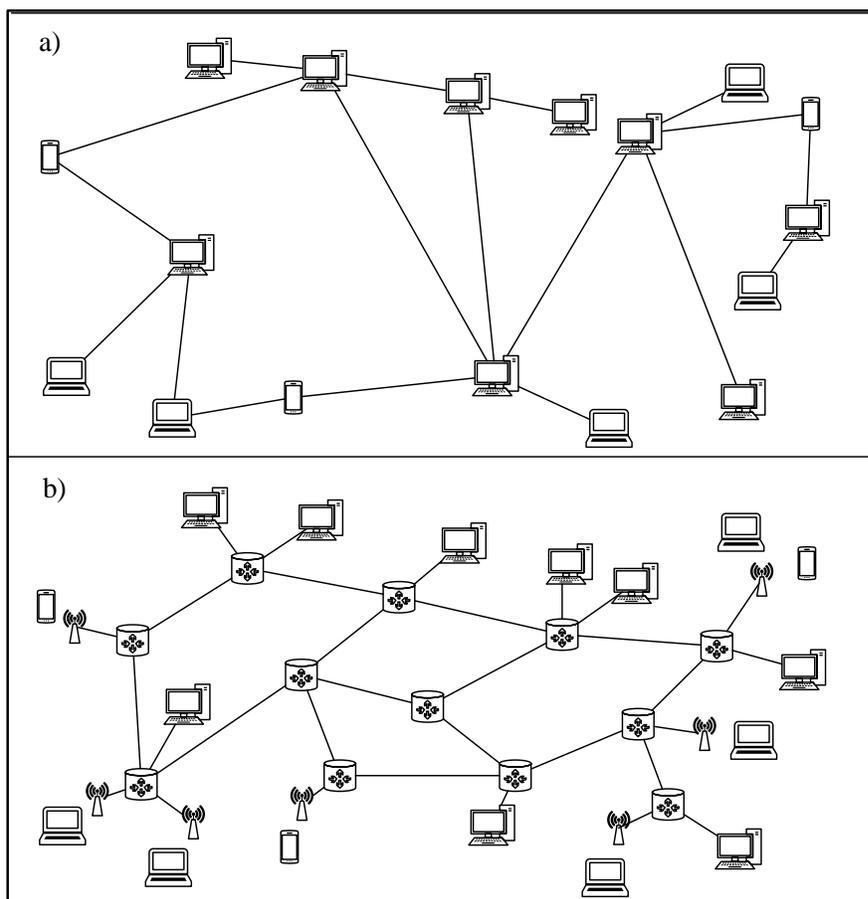


Figura 1.1. Ejemplo de una organización de los nodos en: A) Red sobrepuesta, B) Red subyacente.

Un ejemplo de este tipo de servicio implementado en redes sobrepuestas es el enrutamiento multidifusión [6]. El enrutamiento multidifusión permite a un mensaje enviado a una dirección multidifusión ser enviado a todos los miembros dentro de un grupo multidifusión. Esto es importante para reducir el tráfico en la red producido por aplicaciones como la difusión de video o videoconferencia.

La gran mayoría de los sistemas de redes sobrepuestas se construyen por encima de Internet. Así, los dispositivos terminales que conforman a la red sobrepuesta son dispositivos que se encuentran dispersos a lo largo de Internet, los cuales deben cumplir los siguientes requerimientos [7]:

- Capacidad de ejecución de una o más aplicaciones distribuidas proporcionando la infraestructura para ellas.
- Participar en las tareas de enrutamiento y reenvío a nivel de aplicación. La red sobrepuesta tiene como tarea proporcionar características de enrutamiento y reenvío diferentes a las que conforman los servicios básicos de Internet.
- Permitir un despliegue a lo largo de la Internet de manera tal que otras entidades puedan participar en la organización y operación de la red sobrepuesta.

En la Figura 1.2 se presenta la estructura de una red sobrepuesta organizada en forma de capas y se muestra como comparación el modelo de arquitectura de Internet.

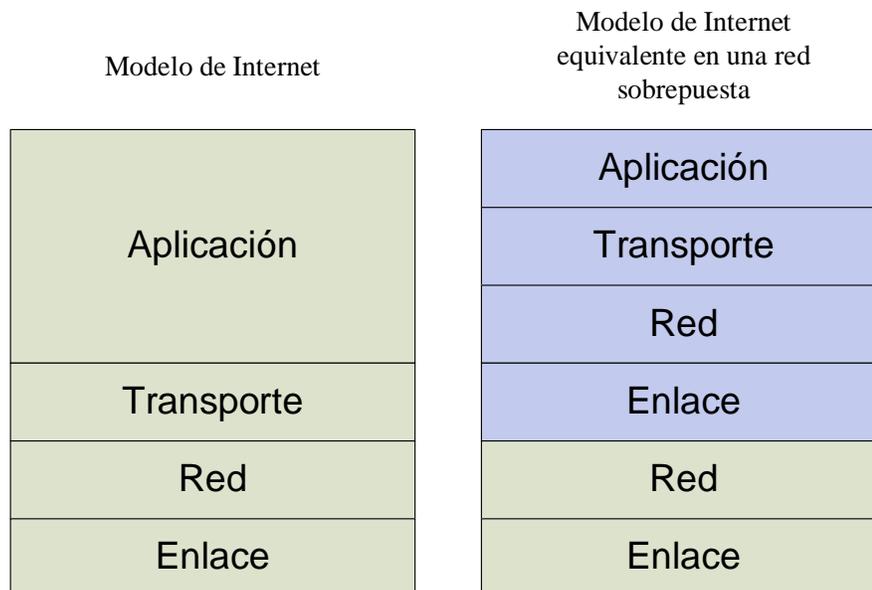


Figura 1.2. Estructura de una red sobrepuesta organizada por capas en comparación con el modelo de Internet.

La estructura de la red sobrepuesta comienza desde la parte inferior con la capa de enlace, la cual es responsable de ofrecer las primitivas básicas para enviar y recibir mensajes directamente entre los nodos de la red sobrepuesta. Como se puede apreciar esta capa corresponde con la capa de transporte del modelo de Internet. Los protocolos comúnmente

utilizados son los protocolos de transporte TCP (*Transmission Control Protocol*) y UDP (*User Datagram Protocol*).

Posteriormente, en la red sobrepuesta se tiene la capa de red, la cual se encarga de los servicios de enrutamiento y reenvío de mensajes entre nodos que conforman la red sobrepuesta. El enrutamiento es el proceso de construir y mantener las tablas de ruteo, mientras que el reenvío es el proceso de enviar el mensaje hacia su destino. Esta capa debe cumplir con sus funciones a pesar de los cambios en la red subyacente o en los dispositivos pertenecientes a la red sobrepuesta.

A continuación se encuentra la capa de transporte de la red sobrepuesta, la cual tiene como función la entrega de mensajes extremo a extremo entre nodos conectados a la red sobrepuesta. Esta capa proporciona dos servicios fundamentales: enviar y recibir mensajes directamente entre dispositivos conectados a la red sobrepuesta.

En la parte superior de la arquitectura de la red sobrepuesta está la capa de aplicación. Es en esta capa donde se ejecutan las aplicaciones que utilizan los servicios proporcionados por la red sobrepuesta.

1.1.2 Ventajas y limitaciones

Una red sobrepuesta ofrece ciertas ventajas en comparación a soluciones centralizadas o soluciones que introducen cambios significativos en los dispositivos que conforman la red existente [5]. Dentro de estas ventajas se encuentran:

- Las redes sobrepuestas no requieren cambios en los dispositivos que conforman a la red existente. Esto implica que una red sobrepuesta puede ser construida basada en protocolos estándar de red e interfaces existentes.
- La implementación de la red sobrepuesta puede realizarse añadiendo nodos de manera incremental, de esta manera es posible monitorear el desempeño de la red durante todo el proceso de su creación.

- El algoritmo de enrutamiento de la red sobrepuesta puede utilizar diversas métricas cuando realiza las decisiones de enrutamiento y reenvío. Por esta razón, la red sobrepuesta puede considerar aspectos específicos de la aplicación, tales como la latencia en la red, que son desconocidos por la infraestructura subyacente.
- Una red sobrepuesta es robusta a fallas presentes en la red subyacente y en los nodos que la conforman debido a su naturaleza adaptable. Con un número suficiente de nodos, es posible encontrar rutas alternas de un nodo hacia otro permitiendo así el enrutamiento en presencia de fallas.

Por otro lado, las redes sobrepuestas presentan ciertas limitaciones en su implementación, dentro de las cuales se pueden mencionar:

- Una red subyacente basada en el protocolo IP, como lo es la Internet, no proporciona conectividad universal extremo a extremo debido a la existencia de elementos en la red como *firewalls* y dispositivos traductores de direcciones (NAT por *network address traslation*). Esto quiere decir que la red sobrepuesta debe sobreponerse a estos obstáculos de accesibilidad. Una forma de conseguir esto es con el uso de protocolos específicos como el protocolo ICE (*Interactive Connectivity Establishment*).
- Una red sobrepuesta está conformada por dispositivos que se encuentran a lo largo de la Internet. Debido a esto, es complicado que la red sobrepuesta tenga la información adecuada acerca de la topología de la red subyacente para optimizar de manera adecuada el proceso de enrutamiento. Una manera de optimizar el intercambio de mensajes entre nodos de una red sobrepuesta es con el uso de redes sobrepuestas conscientes de la topología subyacente.

1.1.3 Aplicaciones

Existe una gran variedad de redes sobrepuestas que han surgido para satisfacer una serie de propósitos y necesidades, las cuales han sido propuestas tanto por la comunidad de investigación como por las compañías de Internet. Algunos ejemplos de casos donde es

posible hacer uso de redes sobrepuestas para cubrir una necesidad específica son los siguientes [7]:

- Redes P2P de compartición de archivos: Su propósito es la compartición de archivos distribuidos entre usuarios de la red superpuesta. Un sistema de este tipo es BitTorrent [8].
- Redes de entrega de contenido (CND por *content delivery networks*): Son redes que almacenan copias de cierto tipo de contenido en varios puntos de la red y permiten la distribución de datos en escala masiva con un menor retardo.
- Enrutamiento y reenvío: Su función es reducir el retardo y el costo en el enrutamiento. Aquí se pueden mencionar a las redes sobrepuestas resistentes (RON por *resilient overlay networks*), las cuales pretenden proporcionar el camino más corto en la red física para la entrega de un mensaje.
- Seguridad: Se utilizan para mejorar la seguridad de los usuarios terminales y brindar privacidad. Dentro de estos servicios se encuentran las redes privadas virtuales (VPN por *virtual private network*) y el enrutamiento cebolla (*onion routing*).
- Experimental: Ofrecen una maqueta de pruebas para la implementación y desarrollo de nuevas tecnologías en fase de experimentación. Un ejemplo de este tipo de red es PlanetLab [9], el cual es un conjunto grande de servidores distribuidos a lo largo de la Internet para la implementación de protocolos en fase de desarrollo.
- Comunicación: Sistemas de redes sobrepuestas que ofrecen una mejora en la comunicación entre usuarios de la red, como lo puede ser el correo electrónico, el enrutamiento multidifusión, servicio de voz sobre IP (VoIP por *Voice over IP*), entre otros.

Como se puede observar de los ejemplos anteriores, los servicios proporcionados por las redes sobrepuestas abarcan desde sistemas de propósito específico que proporcionan servicios avanzados de enrutamiento, hasta redes experimentales como PlanetLab que

existen como una plataforma para el despliegue de redes sobrepuestas múltiples para el estudio de su comportamiento.

1.2 Redes sobrepuestas P2P

Como se mencionó anteriormente, existen varios ejemplos de sistemas P2P de gran escala que se han implementado de manera exitosa sobre la Internet. El núcleo de estos sistemas son las redes sobrepuestas P2P, las cuales son capaces de conectar millones de usuarios y permitir descubrir de manera dinámica los datos almacenados en cualquier nodo de la red [4]. A continuación se presentan los conceptos básicos de las redes sobrepuestas P2P.

1.2.1 Definición

Considerando algunas de las particularidades que aparecen en la literatura [10] [3], es posible determinar un conjunto de características básicas que poseen las redes sobrepuestas P2P.

Una red sobrepuesta P2P es una arquitectura de red distribuida sin ningún tipo de control centralizado, la cual se encuentra conformada por dispositivos finales autónomos, denominados pares. Los pares forman una topología de red sobrepuesta de tal manera que asumen roles simétricos en la operación de la red. De esta manera, si dos pares no se encuentran directamente conectados, pueden hacer uso de los mecanismos de enrutamiento de la red sobrepuesta para el envío de mensajes de manera indirecta a través de otros pares. Asimismo, los pares comparten los recursos propios necesarios para proporcionar el servicio y contenido ofrecidos por la red.

Una característica de cualquier red sobrepuesta P2P es la simetría, es decir, cada nodo cumple con operaciones similares en la operación del sistema. Es decir, además de obtener servicios de la red, un par debe participar en el enrutamiento de búsquedas y recuperación

de recursos hacia otros pares de la red. Sin embargo, esta suposición no es realista ya que varios sistemas P2P aplican un enfoque jerárquico para lidiar con la heterogeneidad de los pares en la red [11].

Otra característica inherente en la mayoría los sistemas de redes sobrepuestas P2P es la propiedad conocida como *churn* [10]. Esta propiedad representa el dinamismo de la llegada y partida de los pares a la red sobrepuesta, lo cual tiene repercusiones en la topología de ésta. Por esta razón el mantenimiento de la red sobrepuesta P2P se convierte en un problema central. El mantenimiento de la red es la operación de reparar y estabilizar el estado del enrutamiento en la red sobrepuesta bajo la presencia de este tipo de cambios en la topología.

En comparación con otros sistemas descentralizados, los participantes en P2P poseen solamente información parcial acerca del estado entero de la red. Esto quiere decir que los recursos destinados por cada par para la operación de la red sobrepuesta crecen como función del tamaño de la red sobrepuesta, y normalmente es menor a una función lineal. Esto permite que las redes sobrepuestas P2P puedan operar con millones de nodos de manera simultánea de manera escalable.

Los protocolos de varias aplicaciones P2P tienen algunas características en común. Estos protocolos son construidos en la capa de aplicación de la pila de protocolos de la red. Adicionalmente, en la mayoría de los diseños, los pares poseen un identificador único, el cual es el identificador del par o la dirección del par en la red sobrepuesta.

La popularidad de los sistemas de redes sobrepuestas P2P se debe a los beneficios que ofrece a los usuarios terminales de la red. En comparación con la arquitectura cliente-servidor, los sistemas P2P presentan ciertas ventajas, entre las que se encuentran las siguientes [4]:

- Los usuarios terminales de la red sobrepuesta P2P pueden compartir o recuperar recursos de manera directa con otros usuarios de la red, sin necesidad de almacenarlos en un servidor centralizado.

- Las aplicaciones basadas en la arquitectura P2P son más robustas en comparación con las basadas en la arquitectura cliente-servidor ya que al no depender de un servidor centralizado, la falla en un nodo no impide las funciones de la red sobrepuesta.
- Debido a que las interacciones en la red se realizan entre pares, se eliminan los cuellos de botella de los servidores centralizados que se crean cuando muchos usuarios se comunican de manera simultánea con un solo servidor.

1.2.2 Aplicaciones

La popularización de los sistemas de redes sobrepuestas P2P comenzó con las redes P2P de compartición de archivos con aplicaciones como BitTorrent y KaZaA. En estas aplicaciones las búsquedas y transferencias de archivos se distribuyen entre los nodos participantes de la red sin necesidad de un directorio centralizado. El surgimiento de este tipo de sistemas así como su impacto en el desempeño de la Internet motivaron el interés de la comunidad de investigación para desarrollar soluciones a las deficiencias de este tipo de sistemas.

Recientemente, la distribución de video en tiempo real (*streaming*) conocida como P2PTV se ha convertido en una aplicación importante en P2P. A diferencia de los sistemas de compartición de archivos en los cuales los medios se descargan en la computadora del usuario para ser reproducidos de manera local, las aplicaciones de distribución de video en tiempo real proporcionan una tasa de transferencia del flujo de datos a cada par que en conjunto es igual a la tasa de reproducción del contenido. En este tipo de aplicaciones se utilizan varios servicios de redes sobrepuestas, incluyendo el enrutamiento multidifusión a nivel de aplicación y redes de entrega de contenidos híbridas. Ejemplos de este tipo de sistemas son PPStream y PPLive.

Por otro lado, los servicios de telefonía basados en voz sobre IP han presentado un aumento en su utilización en los últimos años. Estos servicios ofrecen llamadas entre usuarios por medio de computadoras e incluso permiten establecer llamadas a bajo costo a través de la red telefónica pública conmutada (PSTN por *Public Switched Telephone Network*). Dentro

de este tipo de sistemas el sistema con mayor penetración en el mercado es Skype [12]. Estimaciones indican que en el año 2013 cerca del 40% de llamadas internacionales en el mundo se realizaron a través de Skype [13].

El sistema Skype utiliza un protocolo propietario, donde los paquetes de datos y control se encuentran cifrados. Por esta razón es difícil determinar la forma exacta en la que el sistema opera. Sin embargo, existen estudios que muestran que Skype utiliza un modelo de red sobrepuesta P2P jerárquica, donde la información de los usuarios se encuentra distribuida entre los nodos P2P. Por otro lado, existen dos clases de nodos dependiendo de sus características, donde a los nodos con mayor jerarquía se les denomina *superpares*. Los superpares permiten conectar a pares ordinarios a través de dispositivos NAT [14].

Al utilizar sistemas basados en protocolos propietarios existen complicaciones de interoperabilidad entre sistemas realizados por distintos desarrolladores que no existen cuando se utilizan protocolos estandarizados. Un caso especial para efectuar la estandarización de servicios VoIP en redes sobrepuestas P2P son el conjunto de protocolos P2PSIP (*Peer-to-Peer SIP*).

- **P2PSIP**

El Protocolo de Inicio de Sesión (SIP por *Session Initiation Protocol*) [15] es un protocolo de señalización a nivel de capa de aplicación para el establecimiento, modificación y terminación de sesiones entre dos o más participantes. El término *sesión* puede ser utilizado para cualquier tipo de intercambio de datos, incluyendo VoIP, video, mensajería de texto, etc. El protocolo SIP ha sido utilizado ampliamente en aplicaciones que ofrecen servicios a través de la Internet. Sin embargo, la arquitectura SIP requiere el uso de servidores dedicados en la red para su funcionamiento, como el *Servidor de Registro* que se encarga del registro de los usuarios y el *Servidor de Ubicación* que realiza la conversión de una dirección SIP en la dirección IP del dispositivo.

El grupo de trabajo P2PSIP impulsado por el Grupo de Trabajo de Ingeniería de Internet (IETF por *Internet Engineering Task Force*) [16] tiene como objetivo desarrollar protocolos y mecanismos para el uso del protocolo SIP en ambientes donde no existe la dependencia de servidores centralizados. En otras palabras, el establecimiento y administración de sesiones es controlado por un conjunto de dispositivos terminales en lugar de servidores dedicados.

La arquitectura de P2PSIP se basa en un conjunto de nodos que constituyen una red sobrepuesta P2PSIP, en la cual se proporcionan mecanismos para la ubicación de recursos distribuidos entre los participantes de la red. De esta manera se elimina la necesidad de servidores centralizados al utilizar en su lugar los servicios proporcionados por la red sobrepuesta P2P.

El Protocolo de Ubicación de Recursos y Descubrimiento (RELOAD por *Resource Location and Discovery*) [17] es el protocolo propuesto por el grupo de trabajo P2PSIP como protocolo de señalización P2P para su uso en Internet. Este protocolo proporciona a sus clientes un almacenamiento abstracto y servicio de mensajes entre el conjunto de pares que forman la red sobrepuesta P2P. RELOAD fue diseñado para redes P2PSIP pero puede ser utilizado por otras aplicaciones con requerimiento similares al definir nuevos Usos (*usages*) que especifiquen el tipo de datos que necesitan ser almacenados por una aplicación particular.

El Uso SIP de RELOAD [18] es un uso en etapa de desarrollo que permite a agentes de usuario SIP establecer sesiones sin el requerimiento de Servidores de Registro o Servidores Proxy, es decir, un servicio de telefonía distribuido. En otras palabras, la red sobrepuesta RELOAD realiza por ella misma las operaciones de registro y localización del contacto (*rendezvous*). De esta manera se puede hacer la conversión de una dirección SIP a un nodo específico perteneciente a la red sobrepuesta.

1.2.3 Clasificación

Las redes sobrepuestas P2P pueden clasificarse de distintas maneras de acuerdo a la característica particular elegida para ello. Sin embargo, de manera general existen dos clases principales de acuerdo a la estructura de la red sobrepuesta: las redes sobrepuestas sin estructura y las redes sobrepuestas estructuradas [5] [4].

Las redes sobrepuestas sin estructura no mantienen una estructura en su topología de red. Este aspecto produce que un nodo en la red solamente se base en sus nodos adyacentes para la entrega de mensajes hacia otros nodos en la red sobrepuesta. Como ejemplos de estrategias de propagación de mensajes son *inundación* y *caminatas aleatorias*. Cada par que recibe una petición puede evaluarla de manera local utilizando su propio contenido, de esta manera los sistemas P2P sin estructura pueden soportar peticiones más complejas.

Las redes sobrepuestas estructuradas mantienen una estructura controlada y el contenido es distribuido entre los pares de manera que las peticiones pueden ser realizadas de manera eficiente. De esta manera, los nodos mantienen de manera cooperativa información de enrutamiento que permite alcanzar a todos los nodos de la red sobrepuesta. En comparación con las redes sobrepuestas sin estructura, las redes sobrepuestas estructuradas proporcionan un límite en el número de mensajes necesarios para encontrar cualquier objeto en la red sobrepuesta.

1.3 Redes sobrepuestas P2P estructuradas

Las redes sobrepuestas P2P estructuradas imponen restricciones en la colocación de los nodos dentro de la red sobrepuesta para obtener una estructura de datos distribuida que permita un comportamiento determinístico en la búsqueda de datos en la red. Por esta razón, las redes sobrepuestas P2P estructuradas utilizan una *topología de red* específica para ordenar a los nodos participantes en la red. Así, el desempeño de las búsquedas en la red sobrepuesta se encuentra directamente relacionado con la disposición de los nodos y la

manera en que la topología se mantiene cuando los nodos entran o salen de la red. Las topologías más utilizadas en las redes sobrepuestas P2P son *árboles*, *toroides* y *anillos* [5].

La topología de red define la estructura para organizar a los nodos en un espacio virtual de direcciones, conocido como el *espacio de identificadores*, donde cada nodo posee un identificador único dentro de este espacio conocido como *identificador de nodo*. Asimismo, se le asigna un identificador único a cada objeto de datos, denominado *llave o identificador de objeto*, el cual se encuentra dentro del mismo espacio de identificadores y por medio del cual se puede recuperar el objeto. Un objeto puede ser cualquier tipo de recurso. Posteriormente, las llaves se mapean a un nodo dentro de la red sobrepuesta, usualmente el nodo con el identificador más cercano dentro del espacio de identificadores, denominado comúnmente como *nodo responsable de la llave*. De esta manera, cada nodo en el sistema es responsable de un conjunto de llaves.

Cada nodo mantiene una *tabla de enrutamiento*, la cual se encuentra conformada con la información de contacto de algunos nodos en la red, usualmente sus nodos vecinos. De esta forma, las peticiones de objetos se enrutan a través de la red sobrepuesta utilizando la información contenida en las tablas de enrutamiento. Usualmente, cada paso en el enrutamiento reenvía la petición hacia un nodo más cercano a su destino, es decir, el nodo responsable de la llave. A este tipo de enrutamiento se le conoce como *enrutamiento basado en llave*.

Dependiendo de la topología particular que se utilice en la red sobrepuesta, existen distintos tipos de algoritmos de enrutamiento. El algoritmo de enrutamiento define la manera en que se localiza a un nodo o llave específica dentro de la red sobrepuesta y depende en gran manera de la topología de la red sobrepuesta y la información almacenada en cada nodo. La organización de la tabla de ruteo para realizar la selección del siguiente salto en la ruta de una petición es un aspecto importante en el desempeño del algoritmo.

Las búsquedas de objetos basadas en *tablas hash distribuidas* (DHT por *Distributed Hash Table*) es un esquema de búsqueda ampliamente adoptado en las redes sobrepuestas P2P

debido a su eficiencia. Utilizando este enfoque, cualquier nodo puede encontrar un objeto en particular utilizando su nombre. Asimismo, cualquier objeto existente en la red sobrepuesta puede ser encontrado por cualquier nodo participante. A continuación se presentan los aspectos generales de los sistemas basados en DHT.

1.3.1 Tablas Hash Distribuidas

Una Tabla Hash Distribuida (DHT) es un tipo de algoritmo distribuido descentralizado que ofrece un servicio de búsqueda. Para realizar esta función, la DHT almacena un objeto con su llave asociada formando una pareja (*llave, objeto*), de esta manera se puede realizar la búsqueda de un objeto asociado a una cierta llave. Las parejas (*llave, objeto*) se distribuyen en el sistema, y el sistema DHT debe asegurarse que los nodos tengan suficiente información del estado global para ser capaces de reenviar y procesar las peticiones de búsqueda de manera apropiada.

En una red sobrepuesta P2P estructurada basada en DHT la responsabilidad del conjunto de llaves se divide entre los pares activos en la red. El algoritmo DHT es responsable de la distribución de llaves, de tal forma que la recuperación de un objeto que corresponde a una llave se realice de manera eficiente. El servicio de búsqueda de un objeto se realiza tomando como parámetro de entrada la llave asociada al objeto, obteniendo de esta manera el identificador del nodo responsable de ella. Dado que los nodos pueden entrar o salir de la red, se requiere que este algoritmo sea capaz de lidiar con los cambios en el sistema distribuido.

En un sistema tradicional basado en una *tabla hash*, un cambio en el número de elementos del sistema origina que la mayoría de las llaves cambien de responsable y por lo tanto deban ser transferidos a su nuevo responsable, lo cual no resulta adecuado en un sistema con nodos dinámicos. Por esta razón los sistemas P2P basados en DHT utilizan el concepto de *hash consistente* [19] para dividir el espacio de identificadores entre un conjunto distribuido de nodos y con ello los recursos existentes en el sistema.

Una característica importante del hash consistente es la propiedad de *suavidad*. Esta propiedad implica que cambios suaves en el conjunto de nodos generan cambios suaves en la ubicación de los recursos. En otras palabras, cuando un nodo entra o sale del sistema, la fracción esperada de recursos que deben ser transferidos a un nuevo responsable es la mínima para mantener balanceada la carga a través de los nodos. De hecho, en esta técnica en promedio solamente el K/N de las llaves deben ser transferidas a un nuevo responsable, donde K es el número de llaves y N es el número de nodos en el sistema.

En los sistemas basados en DHT, los identificadores de los nodos y de los recursos provienen del mismo espacio de identificadores de m -bits. Es decir, el identificador de un nodo se obtiene al aplicar una *función hash* a un parámetro característico del nodo como puede ser su dirección IP o su llave pública. Asimismo, el identificador de un objeto se obtiene al aplicar la misma función hash a uno de sus metadatos.

Una función hash es una función que se utiliza para mapear una cadena de datos de longitud arbitraria a una de longitud fija, llamada comúnmente como *digesto* [20]. Usualmente el dominio de una función hash es mas grande que su rango y por lo tanto dos entradas pueden mapearse al mismo digesto, a esto se le conoce como *colisión*. Una función hash con buenas características de distribución es aquella que produce colisiones con muy baja probabilidad. Al hacer uso de una función hash con buenas características de distribución, los identificadores de nodos y las llaves se distribuyen de manera uniforme en el espacio de identificadores, consiguiendo de esta forma que cada nodo almacene un número similar de llaves.

Las arquitecturas de sistemas basadas en DHT presentan las siguientes propiedades [4]:

- La red sobrepuesta se estructura de manera que el número máximo de saltos necesarios para alcanzar a cualquier par en la red sobrepuesta se minimice. Esquemas DHT típicos tienen un límite en la longitud de la trayectoria de la búsqueda promedio de $O(\log N)$, donde N es el número de nodos en la red.

- Cada par mantiene solamente un número pequeño de conexiones activas hacia otros pares. De esta manera el almacenamiento en los nodos crece de manera logarítmica con el número de nodos en la DHT. Esto permite alcanzar una gran escalabilidad en el número de usuarios.
- Las búsquedas de recursos pueden resolverse con alta probabilidad aun ante la presencia de fallas y cambios en las condiciones de la red incluyendo particiones de la red y la llegada y partida de los pares a la red sobrepuesta de manera continua.

Por otro lado, el espacio de identificadores establecido por la red sobrepuesta es una construcción virtual y los identificadores de los nodos usualmente se asignan de manera aleatoria. Por esta razón los pares que son vecinos en la red sobrepuesta pueden tener una gran separación en la red subyacente. Esta particularidad mejora la tolerancia de la red sobrepuesta ante la presencia de fallas en la red subyacente, sobretodo si se trata de una falla en una zona geográfica particular. Sin embargo, se tiene una pérdida en el desempeño de la red ya que la trayectoria de un mensaje no es óptima. Es decir, la trayectoria en la red subyacente entre dos pares puede ser significativamente diferente a la trayectoria utilizada por la red sobrepuesta. Debido a esto la latencia de la búsqueda en las redes sobrepuestas P2P basadas en DHT pueden ser altas.

Para reducir el retardo en el enrutamiento de los mensajes se puede implementar una arquitectura de red sobrepuesta consciente de la topología subyacente, la cual considera la información de la localización de los nodos [11]. En estos sistemas se utilizan medidas de proximidad entre los pares en la red subyacente para crear vecinos en la red sobrepuesta. De esta manera, se reduce el número de saltos en la red subyacente que implican realizar un salto en la red sobrepuesta, reduciendo de esta manera la latencia en el enrutamiento.

1.3.2 Ejemplos

A continuación se presentan tres diferentes diseños de protocolos distribuidos basados en DHT para la localización de recursos y el enrutamiento utilizados en redes sobrepuestas P2P, los cuales se encuentran comúnmente referidos en la literatura [21].

- **Chord**

En Chord [22] cada nodo u objeto tiene un identificador de m -bits y se organizan de acuerdo a su identificador en un espacio en forma de anillo, el cual contiene 2^m direcciones diferentes. El identificador se genera al aplicar una función hash a la dirección IP del nodo y al nombre del archivo en el caso del objeto. La función hash utilizada en Chord es SHA-1, por lo que el espacio de identificadores es de tamaño 2^{128} . De esta manera, el primer nodo que sigue a una llave a lo largo del anillo en dirección de las manecillas del reloj es responsable de este objeto.

Cada nodo mantiene una tabla de enrutamiento de m entradas, donde cada entrada i apunta al siguiente nodo que sigue a lo largo del anillo a una distancia de al menos 2^i , donde $i = 0, \dots, m - 1$. La Figura 1.3 muestra un diagrama esquemático de la estructura de un anillo Chord con 2^6 direcciones y 7 nodos donde se muestra además la región de responsabilidad de cada nodo en el anillo. Adicionalmente, cada nodo almacena a su nodo predecesor en el anillo y a los r nodos sucesores en el anillo, donde r es un parámetro de diseño.

La estrategia de enrutamiento de Chord reenvía una petición en dirección de las manecillas del reloj de tal manera que un nodo intermedio elige al nodo precedente más cercano posible hacia el destino de acuerdo a su tabla de enrutamiento como siguiente salto. El número de nodos intermedios visitados a lo largo de la trayectoria de una búsqueda es $O(\log N)$ en una red Chord de N nodos.

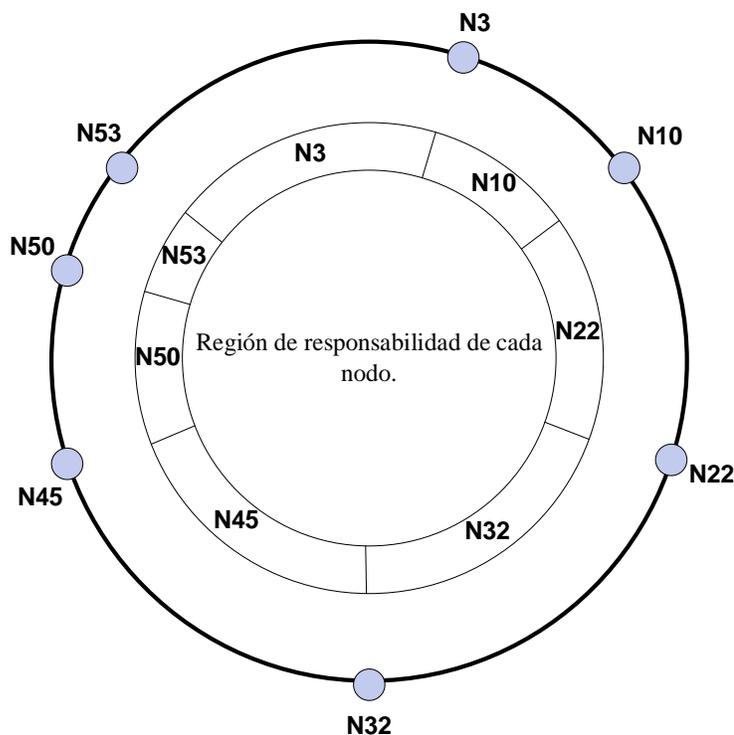


Figura 1.3. Estructura de anillo Chord.

- **CAN**

CAN (*Content-Addressable Network*) [23] es una DHT que utiliza un espacio coordinado de d dimensiones dividido en N zonas, donde cada zona es manejada por un nodo. Dos nodos son vecinos si sus zonas colindan a lo largo de una dimensión y se traslapan a lo largo de las otras $d - 1$ dimensiones, de esta manera cada nodo mantiene $2d$ vecinos. Cada identificador es una coordenada en d dimensiones en el espacio de identificadores La Figura 1.4 muestra un ejemplo de un espacio de 2 dimensiones en CAN con 22 zonas en donde los vecinos del nodo X es el conjunto de nodos $\{A, B, C, D\}$.

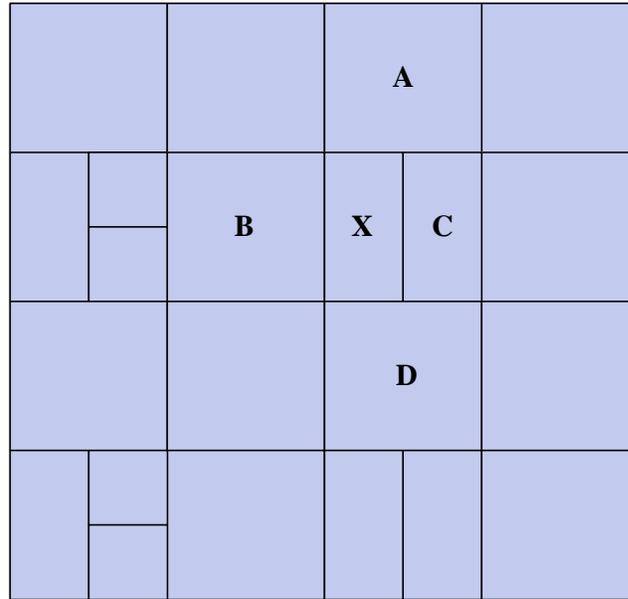


Figura 1.4. Estructura de un espacio de 2 dimensiones en CAN.

CAN utiliza una estrategia de enrutamiento donde cada nodo reenvía una petición hacia su vecino cuya zona se encuentra más cercana a la llave solicitada. La longitud promedio de una trayectoria de enrutamiento es $O(dN^{1/a})$ en una red CAN de N nodos.

- **Pastry**

De manera similar a Chord, Pastry [24] organiza a los nodos en una estructura de anillo, donde cada identificador especifica una posición a lo largo del anillo. El identificador del nodo se obtiene al aplicar una función hash a la dirección IP o a la llave pública del usuario, y la llave de un objeto se obtiene al aplicar una función hash a su nombre.

Cada identificador tiene un tamaño de 128 bits y consiste en una secuencia de dígitos, llamados *niveles*, en base $B = 2^b$, donde b es un parámetro configurable y usualmente $b = 4$. Cada nodo es responsable de las llaves más cercanas numéricamente en el anillo. Los nodos mantienen tablas de enrutamiento con información de otros nodos con los cuales comparten prefijos comunes de diferentes tamaños, así como los vecinos más cercanos en el espacio de identificadores y los vecinos más cercanos al nodo en distancia en la red

subyacente. Un prefijo son los primeros dígitos comenzando por la izquierda de un identificador.

En cada paso del enrutamiento un nodo intermedio elige un vecino que comparta un prefijo más grande con el destino como siguiente salto. Si dicho vecino no existe, la petición se envía a un nodo que comparta un prefijo de la misma longitud, pero se encuentre numéricamente más cercano a la llave. La longitud promedio de una trayectoria de enrutamiento es $O(\log_B N)$ en una red Pastry de N nodos.

1.3.3 Comparación

En esta sección se realiza una comparación de las redes sobrepuestas P2P estructuradas mencionadas en la sección anterior, donde cada una de ellas presenta algunas características únicas. Se presentarán sus similitudes y diferencias desde la perspectiva de la construcción de la red sobrepuesta, la métrica de distancia utilizada en cada una de ellas y el algoritmo de enrutamiento empleado para el envío de mensajes entre participantes.

- **Topología de red**

Chord construye un anillo unidimensional donde todos los identificadores de nodos y objetos se organizan en el sentido de las manecillas del reloj en orden ascendente. Cada nodo es responsable de los objetos que residen en rango comprendido entre su identificador y el identificador de su nodo precedente en el anillo. CAN construye un espacio cartesiano de dimensión d , el cual se divide en zonas que no se traslapan que corresponden a cada uno de los nodos. Cualquier objeto se mapea como un punto en una de las zonas del espacio de dimensión d . Pastry utiliza una estructura de anillo donde cada nodo es responsable de las llaves más cercanas numéricamente a él.

▪ Distancia

La métrica de *distancia*, $d(i, j)$, calcula la distancia entre dos nodos u objetos i y j en un sistema distribuido [25]. Esta métrica se utiliza como base para las operaciones de enrutamiento dentro de la red sobrepuesta. Por esta razón, la diferencia entre las métricas de distancia utilizadas en diferentes arquitecturas produce diferentes estrategias de enrutamiento.

En Chord, los identificadores de nodos y objetos se procesan de la misma manera. La distancia de i a j se define como $d(i, j) = (j - i) \bmod 2^m$, donde m es el número de bits del identificador del nodo u objeto. En CAN, la distancia es la distancia Euclidiana en el espacio de dimensión d . En Pastry la distancia entre identificadores se evalúa dígito a dígito con el algoritmo de Plaxton.

▪ Enrutamiento

Cada nodo perteneciente a una red sobrepuesta basada en DHT mantiene información de un conjunto de otros nodos para realizar el enrutamiento. Sin embargo, cada una de las estructuras mantiene diferentes conjuntos de nodos.

En Chord, cada nodo almacena tres conjuntos de nodos: un nodo predecesor, una tabla de dedos que consiste de m entradas y una lista de sucesores de r entradas. El número total de nodos que debe mantener cada nodo es $O(\log N)$. En CAN, cada nodo en el espacio de dimensión d mantiene una lista de $2d$ vecinos, la cual es independiente del número total de nodos en el sistema. En Pastry, cada nodo mantiene tres conjuntos de nodos: una tabla de enrutamiento que consiste de aproximadamente $(B - 1) \log_B N$ nodos, un conjunto de M nodos y un espacio de nombres de L nodos, donde los valores usuales de L y M son B y $2B$ respectivamente. Por lo tanto el número de nodos que cada nodo debe mantener es $O(\log_B N)$.

El tamaño promedio de una trayectoria de enrutamiento es un parámetro importante que refleja el desempeño del enrutamiento de cada sistema, mientras menor sea el número de saltos en una trayectoria, es mejor la eficiencia del enrutamiento.

En Chord, la longitud de las trayectorias de enrutamiento no es mayor a $O(\log N)$ saltos, como resultado de la construcción de la tabla de dedos de una manera exponencial a lo largo del anillo. En CAN, la trayectoria promedio tiene una longitud de $\frac{d}{4} N^{1/d}$ y cada nodo puede alcanzarse por cualquier otro nodo en $O(dN^{1/d})$ saltos. En Pastry la longitud promedio de una trayectoria de enrutamiento es $O(\log_B N)$. Cuando $B = 1$, Pastry y Chord tienen el mismo desempeño en la eficiencia de enrutamiento.

En la Tabla 1.1, se presenta una comparación entre las características mencionadas anteriormente. De la comparación realizada se puede observar que existe un compromiso entre el desempeño del enrutamiento de la red y el estado que debe almacenar, es decir, el número de nodos almacenados en su tabla de enrutamiento. En CAN es posible mejorar el desempeño del enrutamiento al aumentar el número de dimensiones. Sin embargo, este aumento en el número de dimensiones aumenta la complejidad de la arquitectura, así como el estado del enrutamiento. En Pastry es posible modificar el parámetro B para obtener una mejora en el desempeño del enrutamiento, sin embargo, el estado del sistema aumenta de manera lineal a este parámetro.

Chord presenta una topología más sencilla con respecto a CAN y el estado requerido es considerablemente menor al de Pastry sin disminuir en gran medida el desempeño del enrutamiento. Por estas características, Chord es apto para su uso en sistemas de gran escala como los estudiados posteriormente. Por otro lado, el protocolo Chord ha sido elegido como algoritmo DHT obligatorio en el protocolo RELOAD del conjunto de protocolos de P2PSIP.

Característica	Chord	CAN	Pastry
Topología	Espacio circular unidireccional.	Espacio coordinado de identificadores de dimensión d .	Malla estilo Plaxton.
Distancia de i a j	$(j - i) \bmod 2^m$.	Distancia Euclidiana.	Evaluada dígito a dígito.
Algoritmo de búsqueda	Envía el mensaje hacia el nodo precedente más cercano posible.	Envía el mensaje hacia el nodo vecino más cercano al destino.	Envía el mensaje hacia el nodo que posea una coincidencia en su identificador con la llave.
Parámetros del sistema	Número de pares N .	Número de pares N , número de dimensiones d .	Número de pares N , base del identificador $B = 2^b$.
Desempeño del enrutamiento	$O(\log N)$.	$O(dN^{1/a})$.	$O(\log_B N)$.
Estado del enrutamiento	$O(\log N)$.	$2d$.	$O(B \log_B N)$.

Tabla 1.1. Comparación de algoritmos DHT.

Referencias

- [1] ITU, «ICT Facts and Figures – The world in 2015,» 2015.
- [2] J. F. Kurose y K. W. Ross, *Computer Networking: A Top-Down Approach Featuring the Internet*, Boston: Pearson, 2013.
- [3] J. F. Buford, H. H. Yu y E. K. Lua, *P2P networking and applications*, Amsterdam: Elsevier/Morgan Kaufmann, 2009.
- [4] X. Shen, *Handbook of Peer-to-Peer Networking*, New York: Springer, 2010.
- [5] S. Tarkoma, *Overlay Networks: Toward Information Networking*, Boca Raton Fla.: CRC Press, 2010.
- [6] R. Iorga, E. Borcoci, R. Miruta, A. Pinto, G. Carneiro y T. Calcada, «Management driven hybrid multicast framework for content aware networks,» *Communications Magazine, IEEE*, vol. 52, nº 1, pp. 158-165, 2014.
- [7] D. Clark, B. Lehr, S. Bauer, P. Faratin, R. Sami y J. Wroclawski, «Overlay Networks and Future of the Internet,» *In Journal of Communications and Strategies*, pp. 1-21, 2006.
- [8] «BitTorrent,» [En línea]. Available: <http://www.bittorrent.com/>. [Último acceso: Agosto 2015].
- [9] «PlanetLab,» [En línea]. Available: <https://www.planet-lab.org/>. [Último acceso: Agosto 2015].
- [10] D. Korzun y A. Gurtov, *Structured Peer-to-peer Systems: Fundamentals of Hierarchical Organization, Routing, Scaling, and Security*, New York: Springer, 2013.

- [11] H. Shen, G. Liu y L. Ward, «A Proximity-Aware Interest-Clustered P2P File Sharing System,» *Parallel and Distributed Systems, IEEE Transactions on*, vol. 26, nº 6, pp. 1509-1523, 2015.
- [12] «Skype,» [En línea]. Available: www.skype.com. [Último acceso: Agosto 2015].
- [13] TeleGeography, «Skype traffic continues to thrive,» [En línea]. Available: <https://www.telegeography.com/products/commsupdate/articles/2014/01/15/skype-traffic-continues-to-thrive/>. [Último acceso: Agosto 2015].
- [14] D. Bonfiglio, M. Mellia, M. Meo y D. Rossi, «Detailed Analysis of Skype Traffic,» *Multimedia, IEEE Transactions on*, vol. 11, nº 1, pp. 117-127, 2009.
- [15] J. Rosenberg, H. Schulzrinne, G. Camarillo, A. Johnston, J. Peterson, R. Sparks, M. Handley y E. Schooler, «SIP: Session Initiation Protocol,» *RFC 3261*, 2002.
- [16] Internet Engineering Task Force, «Peer-to-Peer Session Initiation Protocol (p2psip),» [En línea]. Available: <http://datatracker.ietf.org/wg/p2psip/charter/>. [Último acceso: Agosto 2015].
- [17] C. Jennings, B. Lowekamp, E. Rescorla, S. Baset y H. Schulzrinne, «REsource LOcation And Discovery (RELOAD) Base Protocol,» *RFC 6940*, 2014.
- [18] C. Jennings, B. Lowekamp, E. Rescorla, S. Baset, H. Schulzrinne y E. T. Schmidt, «A SIP Usage for RELOAD,» *draft-ietf-p2psip-sip-15, work in progress*, 2015.
- [19] D. Karger, E. Lehman, T. Leighton, R. Panigrahy, M. Levine y D. Lewin, «Consistent Hashing and Random Trees: Distributed Caching Protocols for Relieving Hot Spots on the World Wide Web,» *Proceedings of the Twenty-ninth Annual ACM Symposium on Theory of Computing*, pp. 654-663, 1997.

- [20] J. Katz y Y. Lindell, *Introduction to Modern Cryptography*, Boca Raton: CRC Press/Taylor & Francis, 2015.
- [21] P. Felber, P. Kropf, E. Schiller y S. Serbu, «Survey on Load Balancing in Peer-to-Peer Distributed Hash Tables,» *Communications Surveys & Tutorials, IEEE* , vol. 16, n° 1, pp. 473-492, 2014.
- [22] I. Stoica, R. Morris, D. Karger, M. F. Kaashoek y H. Balakrishnan, «Chord: a scalable peer-to-peer lookup protocol for Internet applications,» *Networking, IEEE/ACM Transactions on*, vol. 11, n° 1, pp. 17-32, 2003.
- [23] S. Ratnasamy, P. Francis, M. Handley, R. Karp y S. Shenker, «A Scalable Content-addressable Network,» *SIGCOMM Comput. Commun. Rev.*, vol. 31, n° 4, pp. 161-172, 2001.
- [24] A. Rowstron y P. Druschel, «Pastry: Scalable, Decentralized Object Location, and Routing for Large-Scale Peer-to-Peer Systems,» *Proceedings of the IFIP/ACM International Conference on Distributed Systems Platforms Heidelberg*, pp. 329-350, 2001.
- [25] H. Zhang, Y. Wen, H. Xie y N. Yu, *Distributed Hash Table: Theory, Platforms and Applications*, New York: Springer, 2013.

Capítulo 2

2. Evaluación del protocolo Chord

En el presente capítulo se presenta el protocolo Chord, y se estudian sus propiedades de enrutamiento de búsquedas en la red sobrepuesta. Posteriormente se presenta la metodología para simular el protocolo Chord sobre una red subyacente y evaluar su desempeño por medio de las siguientes métricas: número promedio de saltos por búsqueda, tiempo promedio de búsqueda y factor de estiramiento.

2.1 Protocolo Chord

El protocolo Chord [1] [2] es un protocolo de búsqueda distribuido basado en una tabla *hash* distribuida (DHT por *Distributed Hash Table*) que se encarga de localizar de manera eficiente a un nodo que almacena un cierto objeto en una red sobrepuesta P2P. Para lograr su objetivo, el protocolo Chord mapea una llave en particular a un nodo específico dentro de la red. Al asociar una llave con un objeto de datos y almacenar la pareja de (*llave, objeto*) en el nodo correspondiente, se puede recuperar el objeto almacenado en la red sobrepuesta.

Una característica particular de Chord es que utiliza hash consistente [3] para la asignación de llaves a los nodos. El uso de hash consistente tiende a balancear la carga en los nodos ya que cada nodo recibe aproximadamente la misma cantidad de llaves. Adicionalmente, se

requieren relativamente pocos movimientos de llaves cuando los nodos se unen o abandonan del sistema.

Por otro lado, un nodo Chord necesita información sobre un conjunto pequeño de otros nodos para realizar la operación de enrutamiento de manera eficiente. Sin embargo, solamente una pieza de información en cada nodo debe ser correcta para que se garantice el correcto enrutamiento de las peticiones.

2.1.1 Topología de red

En Chord se utiliza una función hash consistente para asignar a cada nodo y objeto de datos un identificador dentro de un espacio de identificadores unidimensional. El identificador de un nodo se genera al aplicar la función hash a la dirección IP del nodo. Asimismo, el identificador de una llave se genera al aplicar la función hash a la llave del objeto, la cual puede ser su nombre.

La longitud del identificador es de m bits, donde m es un parámetro del sistema predefinido, y se utiliza SHA-1 como función hash base. De esta manera el rango de un identificador ID se encuentra entre $0 \leq ID < 2^m$. La longitud del identificador m debe ser suficientemente grande para que la probabilidad que dos nodos o llaves obtengan el mismo identificador después de aplicar la función hash sea despreciable.

El protocolo Chord ordena los identificadores de nodos y llaves en un círculo de identificadores módulo 2^m , es decir, el espacio de identificadores se representa como un círculo de posiciones de 0 a $2^m - 1$. A este círculo de identificadores se le denomina *anillo Chord*.

El anillo de nodos se enlaza en dirección de las manecillas del reloj. El *nodo sucesor* de un nodo n , es el nodo inmediato en el anillo en dirección de las manecillas del reloj, y se denota como *sucesor*(n). De la misma manera, el *nodo predecesor* de un nodo n , es el siguiente nodo en el anillo en dirección contraria a las manecillas del reloj, y se denota como *predecesor*(n).

Una llave k se asigna al primer nodo cuyo identificador es igual o sigue al identificador de k en el espacio de identificadores. En otras palabras, el nodo responsable de una llave k es el nodo sucesor de la llave k en el anillo, y se denota como $\text{sucesor}(k)$. Esto quiere decir que si el responsable de una llave es el nodo n , entonces no existe otro nodo n' tal que $k \leq n' < n$. La Figura 2.1 presenta la estructura de un anillo Chord con $m = 6$, es decir, el anillo se conforma de 2^6 posiciones. En el anillo existen 9 nodos y se presenta la región de responsabilidad de cada nodo. Adicionalmente, existen en el sistema 7 llaves, las cuales se encuentran almacenadas en el nodo sucesor correspondiente como se puede observar en la tabla de almacenamiento de cada nodo.

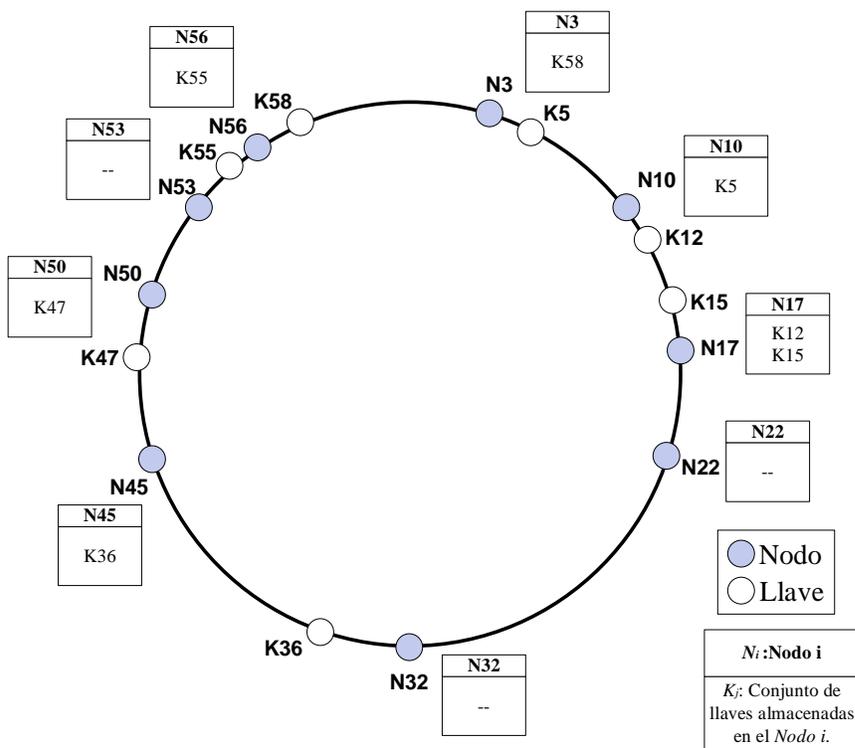


Figura 2.1. Ejemplo de la asignación de llaves a nodos en el anillo Chord.

Los resultados en [3] muestran que al hacer uso de una función hash consistente con buenas propiedades de distribución, existe una alta probabilidad de que los N nodos participantes y las K llaves almacenadas en el sistema se encuentren uniformemente distribuidos a lo largo

del espacio de identificadores, lo cual asegura una distribución equitativa de las llaves en cada nodo. Esto quiere decir que cada nodo es responsable de a lo más $(1 + \epsilon) K/N$ llaves, donde ϵ tiene una cota de $\epsilon = O(\log N)$.

De la misma manera, el uso de una función hash consistente permite a los nodos unirse o abandonar la red con la mínima perturbación. Cuando un nodo n se une a la red, ciertas llaves previamente asignadas al nodo $\text{sucesor}(n)$, se asignan al nodo n . Análogamente, cuando un nodo n abandona el sistema, todas las llaves asignadas se reasignan al nodo $\text{sucesor}(n)$. Esto quiere decir que cuando el $(N + 1)$ –ésimo nodo se une o abandona la red, solamente cambiará la responsabilidad de $O(K/N)$ llaves, sin que exista algún otro cambio en la asignación de llaves en el sistema. Esta característica es importante en el desempeño de la red sobrepuesta cuando existe la llegada y partida de nodos de manera continua.

2.1.2 Búsqueda de una llave

La búsqueda de una llave almacenada en una red sobrepuesta requiere que cada nodo almacene la información de contacto de otros nodos en la red. Particularmente en el protocolo Chord, cada nodo n necesita solamente conocer la información de contacto de su nodo sucesor en el anillo para garantizar el correcto funcionamiento de las operaciones de búsquedas de llaves en la red sobrepuesta.

De esta manera, la búsqueda de un cierto identificador puede ser reenviada alrededor del anillo a través de los enlaces de sucesores hasta que encuentre al nodo sucesor del identificador, es decir, el nodo al cual se mapea la búsqueda. Sin embargo, esta manera de realizar la búsqueda es ineficiente ya que se requiere atravesar cada uno de los nodos entre el nodo que inicia la búsqueda y el nodo responsable de la llave. La longitud de la trayectoria de búsqueda en este algoritmo es de $O(N)$ saltos, donde N es el número de nodos en el sistema. En la Figura 2.2 se muestra un ejemplo de la búsqueda de la llave con identificador 55 iniciada por el nodo con identificador 10, la cual se encuentra almacenada en el nodo 56

para la red mostrada en la Figura 2.1. La búsqueda se propaga en el anillo por medio de los apuntadores a sucesores de cada nodo.

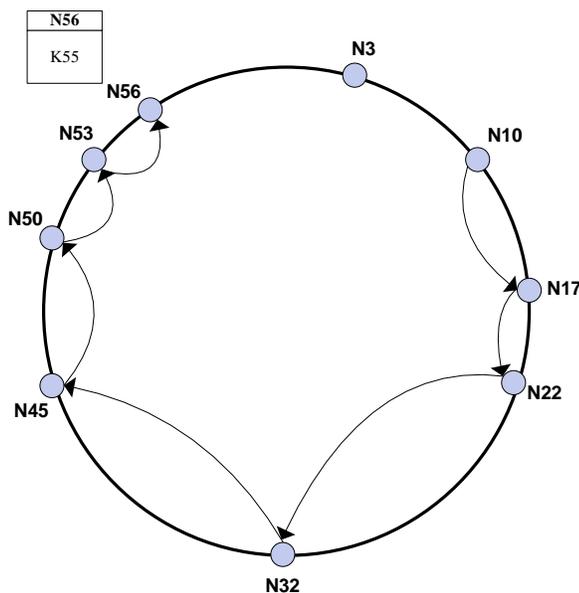


Figura 2.2. Ejemplo del enrutamiento de una búsqueda a través de los apuntadores a sucesores.

Para acelerar el proceso de una búsqueda, cada nodo almacena información de enrutamiento adicional, la cual no es esencial para el funcionamiento correcto de las operaciones de búsqueda ya que éste se consigue al mantener de manera correcta la información del sucesor. De esta manera, cada nodo mantiene una tabla de enrutamiento con hasta m entradas, denominada *tabla de dedos*, donde m es el número de bits de cada identificador. La i -ésima entrada de la tabla de dedos del nodo n contiene la información de contacto del siguiente nodo s que sucede al nodo a una distancia de al menos 2^{i-1} posiciones en el anillo, donde $1 \leq i \leq m$, es decir, $s = \text{sucesor}(n + 2^{i-1})$ donde todas las operaciones aritméticas se realizan en módulo 2^m . La Figura 2.3 presenta un ejemplo en el cual el nodo con identificador 3 construye su tabla de dedos, como se puede observar, las entradas en la tabla de dedos corresponden con el nodo sucesor al identificador correspondiente al cual apunta cada entrada.

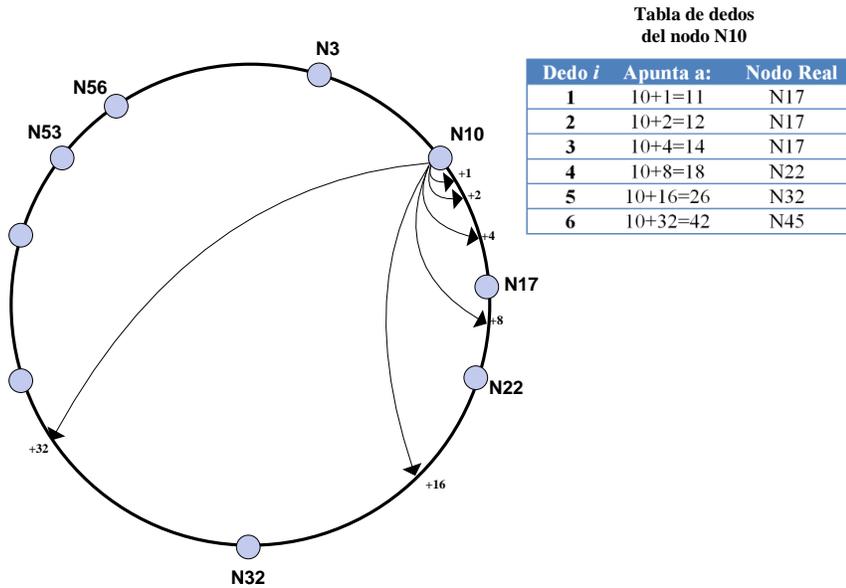


Figura 2.3. Ejemplo de la construcción de una tabla de dedos en Chord.

Como se puede observar de la Figura 2.3, la información contenida en la tabla de dedos generalmente no es suficiente para determinar de manera directa al sucesor de una llave aleatoria k . Por esta razón, el algoritmo de enrutamiento de Chord utiliza la información almacenada en la tabla de dedos de cada nodo para la propagación de una búsqueda a través del espacio de identificadores.

De esta forma, al recibir una petición de búsqueda, un nodo n comprueba si el identificador k solicitado se encuentra entre su propio identificador y el identificador de su sucesor, es decir, si $k \in (n, \text{sucesor}(n)]$. Si esta condición se cumple el nodo determina que el nodo sucesor del identificador es su propio sucesor. En el caso que la condición no se cumpla, el nodo n busca en su tabla de dedos al nodo n' cuyo identificador preceda de manera inmediata al identificador solicitado en el anillo. En otras palabras, el nodo busca en su tabla de dedos el nodo n' con el identificador más grande que satisfaga la condición $n' \in (n, k)$. Una vez identificado el nodo n' se ejecuta el mismo procedimiento ejecutado para el nodo n hasta obtener al nodo sucesor del identificador. La Figura 2.4 muestra un ejemplo de la llave con identificador 55 iniciada por el nodo con el identificador 10, la cual se encuentra almacenada

en el nodo 56. La búsqueda se propaga en el anillo por medio de los apuntadores en la tabla de dedos de cada nodo.

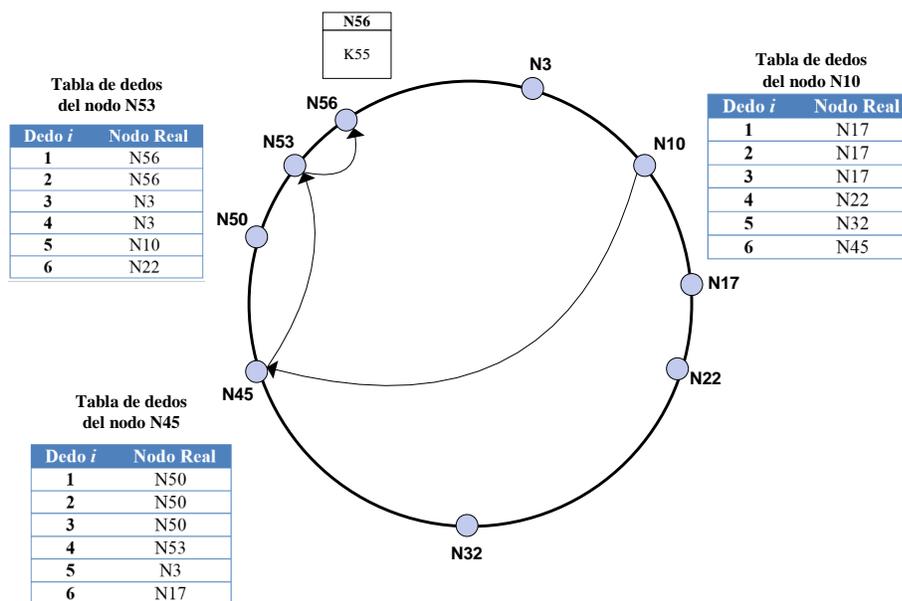


Figura 2.4. Ejemplo del enrutamiento de una búsqueda a través de la tabla de dedos.

Al hacer uso de la tabla de dedos, la longitud de la trayectoria de una búsqueda se reduce de manera significativa ya que cada nodo tiene entradas en su tabla de dedos en intervalos de potencia dos alrededor del círculo de identificadores. Por esta razón, cada nodo intermedio puede reenviar la búsqueda al menos la mitad de la distancia remanente entre el nodo y el identificador destino. De esta manera con una alta probabilidad, el número de nodos que deben ser contactados para encontrar al sucesor de un identificador en una red de N nodos es $O(\log N)$.

2.1.3 Unión y abandono de la red

Una red sobrepuesta P2P es una red dinámica, es decir, el sistema tiene que lidiar con la unión de nuevos nodos al sistema y la posible falla o abandono voluntario de algunos nodos en cualquier momento. Sin embargo, el sistema debe adaptarse a estos cambios y mantener la habilidad de localizar cualquier llave almacenada en la red.

Por esta razón, el protocolo Chord utiliza un protocolo de estabilización para mantener actualizado el apuntador a su sucesor y de esta manera garantizar que las búsquedas se ejecuten de manera correcta a pesar de que el conjunto de nodos en el sistema cambie. De esta manera, el protocolo de estabilización se ejecuta en cada nodo de manera periódica para actualizar el apuntador a sucesor y las entradas de la tabla de dedos.

Cuando un nodo n busca unirse a la red, necesita realizar las siguientes operaciones:

- Generar su identificador para posteriormente determinar su posición en el anillo y obtener las llaves que se encuentran bajo su responsabilidad.
- Inicializar su tabla de dedos.
- Actualizar la tabla de dedos de otros nodos para reflejar la presencia del nuevo nodo.

Para realizar la primera operación el nodo debe encontrar a su nodo sucesor en el anillo s , para lo cual contacta a un nodo *bootstrap* n' , el cual es un nodo existente en el anillo y puede efectuar la operación de búsqueda del nodo *sucesor*(n). Una vez encontrado el nodo s , el nodo n se coloca entre el nodo s y su nodo predecesor en el anillo, el nodo p , y obtiene las llaves del nodo s por las cuales ahora n es responsable y reemplaza a p como nodo precedente de s . En este punto, cuando el protocolo de estabilización se ejecuta en los nodos se actualiza al nodo n como el nuevo nodo sucesor de p , y al nodo p como nodo predecesor de n .

Cuando un nodo se ha colocado en su posición correspondiente dentro del anillo, puede inicializar su tabla de dedos al enviar peticiones de búsqueda para encontrar a los nodos sucesores de los valores específicos en las entradas de la tabla de dedos. Por otro lado, al ejecutarse el protocolo de estabilización en cada uno de los nodos, se actualizan las entradas de la tabla de dedos de los nodos que han sido afectadas debido a la unión de nuevos nodos al sistema.

Por otra parte, el posible fallo o el abandono voluntario de los nodos puede ocasionar que los punteros a sucesores no se encuentren actualizados y afectar el correcto funcionamiento de las búsquedas. Para aumentar la robustez del sistema, cada nodo Chord mantiene una

lista de sucesores de tamaño r , la cual contiene los primeros r nodos sucesores en el anillo. Los resultados en [2] muestran que al mantener una lista de sucesores de tamaño $r = O(\log N)$, las búsquedas se realizan de manera correcta con una alta probabilidad aun cuando cada nodo tenga una probabilidad de fallo de $1/2$. Este mecanismo puede ser utilizado para almacenar réplicas de los objetos asociados con cierta llave en algunos nodos pertenecientes a la lista de sucesores.

Para mejorar el desempeño del sistema, cuando un nodo desea abandonar el sistema de manera voluntaria, el nodo puede transferir las llaves de las cuales es responsable a su nodo sucesor antes de abandonar el sistema. Asimismo, el nodo puede informar a su nodo sucesor y predecesor de su abandono del sistema, permitiendo de esta manera la pronta reconfiguración del sistema.

2.1.4 Longitud promedio de trayectoria

Cuando una búsqueda se envía de un nodo al siguiente en su camino en la red sobrepuesta, se transmite un mensaje a nivel de la red subyacente entre los nodos. Cada paso en la transmisión del mensaje entre pares vecinos se denomina salto en la red sobrepuesta, el cual puede corresponder con varios saltos en la red subyacente dependiendo de la proximidad de los pares en la red subyacente. De esta manera, cuando el número de saltos que una búsqueda utiliza para encontrar un objeto aumenta, igualmente aumenta la latencia de la búsqueda. Por esta razón, el desempeño del protocolo Chord depende del número de nodos que deben visitarse para resolver una búsqueda.

Para realizar el análisis del número de saltos promedio para resolver una búsqueda, en [4] se considera un anillo Chord de tamaño 2^m con N pares, donde se supone que los pares se encuentran igualmente espaciados en el anillo, es decir, la distancia entre dos pares adyacentes es $2^m/N$.

De esta forma, dado un par inicial y una llave generada de manera aleatoria, la variable aleatoria H denota el número de saltos Chord necesarios para alcanzar al nodo responsable

de la llave. Asimismo, la variable aleatoria T denota la distancia medida en número de pares, en sentido de las manecillas del reloj, desde el par inicial hasta el par destino. Debido a que la elección de la llave se realiza de manera aleatoria, la variable T posee una distribución uniforme. Así:

$$E\{H\} = \sum_{n=0}^{N-1} P\{T = n\} E\{H|T = n\} = \frac{1}{N} \sum_{n=0}^{N-1} E\{H|T = n\} \quad (2.1)$$

Como puede observarse de la Ecuación (2.1), es suficiente calcular $E\{H|T = n\}$ para obtener $E\{H\}$. Debe notarse que $E\{H|T = 0\} = 0$ así como $E\{H|T = 1\} = 1$.

Por otro lado se define la variable j_n , la cual representa el número de entradas en la tabla de dedos que preceden al nodo destino, excluyendo al nodo sucesor, la cual se obtiene de la siguiente manera:

$$j_n = \max \left\{ j : 2^j \leq \frac{2^m n}{N} \right\}$$

Cuando una búsqueda realiza un salto al siguiente nodo en su trayectoria, la búsqueda avanza $\left\lceil \frac{2^{j_n}}{2^m/N} \right\rceil$ pares en el anillo, ya que este es el número de pares que se acerca la búsqueda al utilizar la entrada en la tabla de dedos más cercana al nodo destino. Por lo tanto, se obtiene la siguiente expresión:

$$E\{H|T = n\} = 1 + h \left(n - \left\lceil \frac{2^{j_n}}{2^m/N} \right\rceil \right) \quad (2.2)$$

Utilizando la Ecuación (2.2) se puede calcular todos los valores $E\{H|T = n\}$ de manera iterativa a partir de $E\{H|T = 0\} = 0$. Posteriormente se utiliza la Ecuación (2.1) para obtener el número de saltos promedio $E\{H\}$.

Para validar el análisis anterior, se realizó la simulación de una red que utiliza el protocolo Chord con $N = 2^k$ nodos. Se varió el valor de k desde 3 hasta 14 y se realizaron

experimentos separados para cada valor de k . En un experimento dado, cada nodo elige de manera aleatoria un conjunto de llaves del sistema para realizar su búsqueda y se realiza la medición de la longitud promedio de la búsqueda, es decir, del número de saltos promedio realizados para resolver la búsqueda. La Figura 2.5 muestra los resultados obtenidos por medio del análisis analítico y mediante la simulación del sistema.

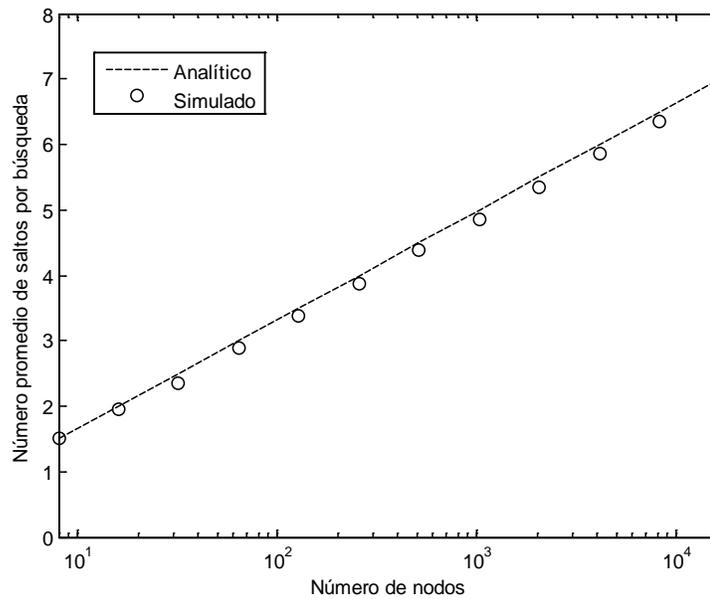


Figura 2.5. Número de saltos promedio en Chord como función del tamaño de la red.

De la Figura 2.5 se puede observar que la longitud promedio de búsqueda aumenta de manera logarítmica con el número de nodos, lo cual se debe al uso de la tabla de nodos en el algoritmo de enrutamiento. Asimismo, se puede obtener que el valor promedio de la longitud promedio de búsqueda es de $\frac{1}{2} \log_2 N$.

2.2 Simulación de la red subyacente

El desempeño de una red sobrepuesta se encuentra directamente relacionado con la manera en que los nodos se encuentran comunicados, de manera lógica en la red sobrepuesta y de manera física en la red subyacente. De esta manera, la estructura de la red subyacente influye en algunas propiedades de la red sobrepuesta, tales como la escalabilidad, el desempeño y la capacidad de recuperación.

La *topología de red* describe la manera en la cual los elementos de la red se encuentran interconectados entre ellos y por lo tanto define la manera en que fluye la información dentro de la red [5]. De este modo, la topología de la red subyacente tiene un impacto en el comportamiento del enrutamiento de los mensajes dentro de la red sobrepuesta. Por esta razón, las características de la topología de red subyacente deben ser consideradas para el diseño y evaluación de la red sobrepuesta. A continuación se presentan los conceptos generales de los generadores de topologías de red.

2.2.1 Generadores de topologías de red

Los *generadores de topologías de red* son herramientas utilizadas para generar topologías realistas con propósitos de investigación de las características de desempeño de una red [6]. La manera en la cual se construye la topología de red se denomina modelo de topología y define las propiedades estructurales de la red. Estos generadores no aspiran a producir réplicas exactas de una red existente, como lo puede ser Internet. En cambio, estos generadores intentan producir topologías de red que incorporen características fundamentales de las redes reales.

Los generadores de topologías de red modelan usualmente la estructura de la red por medio de un grafo, donde los vértices del grafo representan los elementos de la red y las aristas representan conexiones directas entre éstos [7]. De esta forma, inicialmente se genera un archivo de topología de red representada por un grafo para un número predefinido de nodos.

Posteriormente se utiliza alguna estrategia dependiente del escenario específico para generar el retardo extremo a extremo de los mensajes intercambiados dentro de la red.

Un grafo $G(V, E)$ consiste en un conjunto no vacío V de vértices conectado por un conjunto posiblemente vacío E de aristas, donde cada arista es un conjunto de dos vértices de V . El número de vértices y aristas se denota por $|V|$ y $|E|$, respectivamente. En el contexto de redes de comunicaciones, los vértices y aristas se denominan nodos y enlaces respectivamente [8].

Un grafo se denomina grafo ponderado si cada enlace tiene un peso asignado. De esta manera, un enlace entre el nodo i y el nodo j se caracteriza por un peso de enlace $w(i \rightarrow j)$, el cual usualmente es un número real positivo que representa alguna característica del enlace como distancia física, capacidad disponible, etc.

Por otro lado, el intercambio de mensajes entre dos nodos de la red que no se encuentran conectados directamente se realiza a través de una trayectoria. Una trayectoria del nodo A al nodo B con $k - 1$ saltos es la lista de nodos $P_{A \rightarrow B} = n_1 \rightarrow n_2 \cdots \rightarrow n_{k-1} \rightarrow n_k$ donde $n_1 = A$, $n_k = B$ y $n_i \neq n_j$ para cada índice i y j . Dos trayectorias entre los nodos A y B se denominan trayectorias disjuntas si no poseen nodos en común.

El modelo básico de grafo aleatorio es el denominado grafo aleatorio de Erdős-Rényi [9], el cual asigna una probabilidad uniforme a la creación de un enlace entre cualquier par de nodos. De esta manera, dado un número fijo de nodos y una probabilidad p , se crea de manera independiente un enlace entre dos nodos con probabilidad p . Debido a que las conexiones entre nodos se distribuyen con igual probabilidad sobre los pares de nodos, este modelo no manifiesta las propiedades de jerarquía en la red, por lo que este modelo no es apto para implementaciones realistas de redes.

Una variación del modelo anterior es el grafo aleatorio geométrico, el cual considera la distancia entre dos nodos y por lo tanto introduce aspectos de preferencia. Un modelo de este tipo es el modelo Waxman [10], el cual se utiliza en uno de los primeros generadores

de topologías ampliamente utilizado en simulaciones de protocolos. En este modelo los nodos se posicionan de manera aleatoria en un espacio coordinado cartesiano. Posteriormente, la probabilidad P de que exista un enlace que conecte a dos nodos u y v depende de la distancia Euclidiana entre ellos, $d(u, v)$, de la siguiente forma:

$$P(u, v) = \alpha e^{-d(u,v)/L\beta}, 0 < \alpha, \beta \leq 1$$

donde $0 < \alpha, \beta \leq 1$ y L es la máxima distancia entre dos nodos. El grafo de Waxman es un modelo apropiado para redes de tamaño pequeño que incluyen aspectos de localidad.

Sin embargo, las redes de computadoras de gran escala presentan una estructura jerárquica en lugar de una estructura aleatoria [11], es decir, la estructura de la red consiste usualmente de varios niveles. De esta forma, los generadores de topologías de red que utilizan modelos jerárquicos, usualmente conocidos como generadores estructurales, se enfocan en representar la estructura jerárquica de niveles múltiples que presentan las redes reales. El modelo jerárquico usualmente utilizado en el contexto de redes de computadoras es el modelo *Transit-Stub* [12], el cual genera un grafo ponderado jerárquico de dos niveles.

Posteriormente, se encontró que las distribuciones de grado de los grafos en la Internet son del tipo conocido como ley potenciales [13], lo cual generó un nuevo tipo de modelo de topología. Así, los generadores basados en grado utilizan leyes potenciales para generar grafos con las propiedades de la Internet. Estos generadores incorporan la suposición de que es más importante hacer coincidir una cierta propiedad local, como lo es la distribución del grado en cada nodo, que capturar la estructura jerárquica de la Internet [11].

2.2.2 Modelo *Transit-Stub*

La estructura de Internet puede observarse como una colección de dominios de enrutamiento o sistemas autónomos (SA) interconectados, donde cada SA consiste en un grupo de nodos que se encuentran bajo un mismo control administrativo [7]. Los nodos que se encuentran dentro de un mismo SA ejecutan el mismo algoritmo de enrutamiento y comparten información de enrutamiento entre ellos.

En el modelo *Transit-Stub* [12], cada dominio de enrutamiento en la Internet puede clasificarse como un dominio aislado (*stub domain*) o como un dominio de tránsito (*transit domain*). Un dominio aislado solamente transporta tráfico que se origina o termina dentro de ese dominio, es decir, la trayectoria que conecta dos nodos en la red u y v pasa a través de un dominio aislado solamente si u o v se encuentran dentro de este dominio. Por otro lado, un dominio de tránsito no tiene esta restricción. El propósito de los dominios de tránsito es interconectar dominios aislados de manera eficiente. La Figura 2.6 presenta un ejemplo de una estructura de red conformada por 2 dominios de tránsito y 4 dominios aislados.

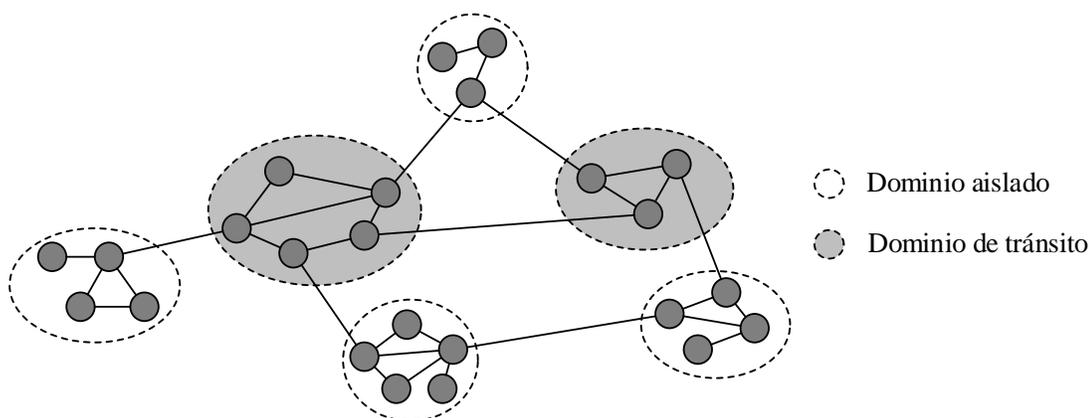


Figura 2.6. Ejemplo de una estructura de red conformada por dominios aislados y dominios de tránsito.

De esta manera, las características de enrutamiento en la Internet pueden resumirse por los siguientes principios generales relativos a la ruta entre dos nodos arbitrarios u y v [12]:

- Si u y v se encuentran dentro del mismo dominio, la trayectoria entre ellos permanece enteramente dentro de ese dominio. El algoritmo de enrutamiento que se ejecuta dentro de un mismo dominio se denomina protocolo de enrutamiento intra-dominio.
- Si u se encuentra en el dominio U y v se encuentra en el dominio V , la trayectoria de u a v inicia en U y atraviesa cero o más dominios de tránsito hasta V . El algoritmo de enrutamiento que se utiliza para conectar diferentes dominios se denomina protocolo de enrutamiento inter-dominio.

Un generador de topología de red ampliamente utilizado que implementa el modelo *Transit-Stub* es el generador GT-ITM (por *Georgia Tech Internetwork Topology Models*) [14], el cual es una paquetería para generar y analizar modelos de grafos de red. Este generador también puede generar grafos aleatorios de Erdős-Rényi y grafos aleatorios geométricos.

Para generar una topología con el modelo *Transit-Stub*, el generador GT-ITM realiza lo siguiente:

- Se construye un grafo aleatorio completamente conectado, es decir, todos los nodos del grafo pueden ser alcanzados a partir de cualquier nodo. Cada nodo de este primer grafo representa un dominio de tránsito.
- Posteriormente, cada nodo de este grafo se reemplaza con otro grafo aleatorio completamente conectado que representa la topología del dominio de tránsito.
- A continuación, por cada nodo en cada dominio de tránsito se genera un número de grafos aleatorios que representan los dominios aislados conectados a cada nodo.
- Finalmente, se agrega un número de enlaces adicionales entre pares de nodos, entre un nodo en un dominio de tránsito y un nodo en un dominio aislado, o dos nodos en diferentes dominios aislados.

Una vez completada la construcción del grafo, se asignan pesos a los enlaces de tal manera que la trayectoria más corta calculada considerando los pesos de los enlaces cumpla las restricciones presentadas del enrutamiento basado en dominio presentadas anteriormente. Es decir, la trayectoria más corta calculada atravesará dominios de tránsito si y sólo si los dos nodos se encuentran en diferentes dominios.

En la Figura 2.7 se muestra un ejemplo de topología de red creada por medio del generador GT-ITM, la cual se encuentra compuesta por 100 nodos. Para la visualización del grafo generado se utiliza la plataforma Gephi [15], la cual es una herramienta de código abierto para el análisis de grafos y redes.

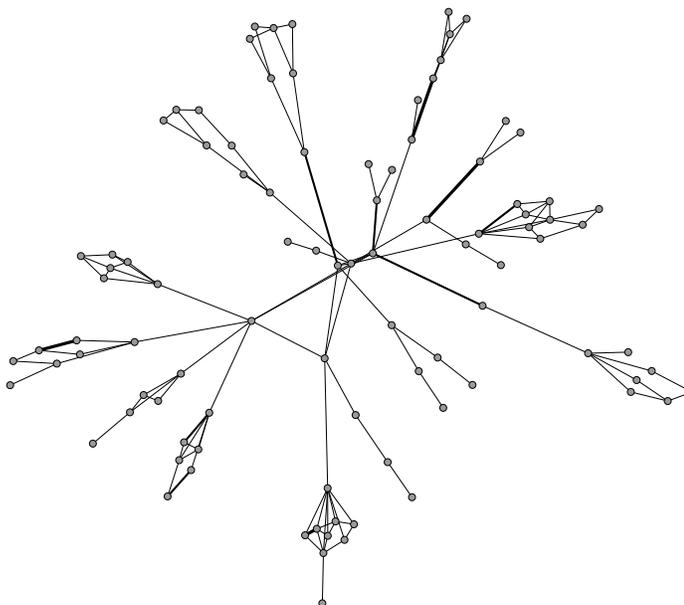


Figura 2.7. Ejemplo de una topología de red de 100 nodos creada con el generador GT-ITM.

2.3 Desempeño del protocolo Chord sobre una red subyacente

Para obtener los resultados del desempeño del protocolo Chord al ejecutarse sobre una red subyacente, inicialmente se debe generar una topología de red sobre la cual se conforma la red sobrepuesta. La topología de red utilizada como topología de red subyacente es generada con el modelo *Transit-Stub* utilizando el generador GT-ITM. Se generaron dos topologías distintas con 4525 y 5400 denotadas como Topología A y Topología B con los siguientes parámetros:

- La Topología A es una topología de red donde la mayoría de los nodos son nodos aislados. Cada nodo de tránsito tiene en promedio 3 dominios aislados y existen 5 dominios de tránsito. Los dominios de tránsito tienen en promedio 5 nodos. En cada dominio aislado existen en promedio 60 nodos.
- La Topología B es una topología de red donde la mayoría de nodos son nodos de tránsito. Cada nodo tránsito tiene en promedio 4 dominios aislados y existen 120 dominios de

tránsito. Los dominios de tránsito tienen en promedio 5 nodos. En cada dominio aislado existen 2 nodos.

La razón de seleccionar estas dos topologías de red es porque se desea evaluar el comportamiento del protocolo Chord en diferentes tipos de redes subyacentes y cada una de ellas representa un caso extremo del comportamiento de una red. La trayectoria recorrida por un mensaje transmitido entre dos nodos cualesquiera de la red subyacente se calcula por medio del algoritmo de Dijkstra [16].

Posteriormente, se realizó la simulación de una red que utiliza el protocolo Chord con $N = 2^k$ nodos, donde cada uno de los nodos pertenecientes a la red sobrepuesta es un nodo seleccionado de manera aleatoria de un dominio aislado de la red subyacente. Asimismo, la asignación del identificador del nodo se realiza de manera aleatoria al entrar a la red. Se varió el valor de k desde 8 hasta 12 y se realizaron experimentos separados para cada valor de k . En un experimento dado, cada nodo elige de manera aleatoria un conjunto de llaves del sistema para realizar su búsqueda.

El retardo o *latencia* de la comunicación en la red es una métrica utilizada como indicador del desempeño de la red experimentado por los nodos, ya que puede ser medida de una manera no intrusiva, práctica y extremo a extremo. De esta manera el *retardo extremo a extremo* se define como el periodo de tiempo que tarda un mensaje en viajar desde el nodo origen hasta el nodo destino. Por lo tanto, el retardo extremo a extremo es la suma de los retardos experimentados por el mensaje en cada uno de los saltos en el camino hacia su destino [5]. En la Figura 2.8 se muestra el retardo promedio de búsqueda del protocolo Chord al ejecutarse sobre cada una de las redes subyacentes descritas previamente.

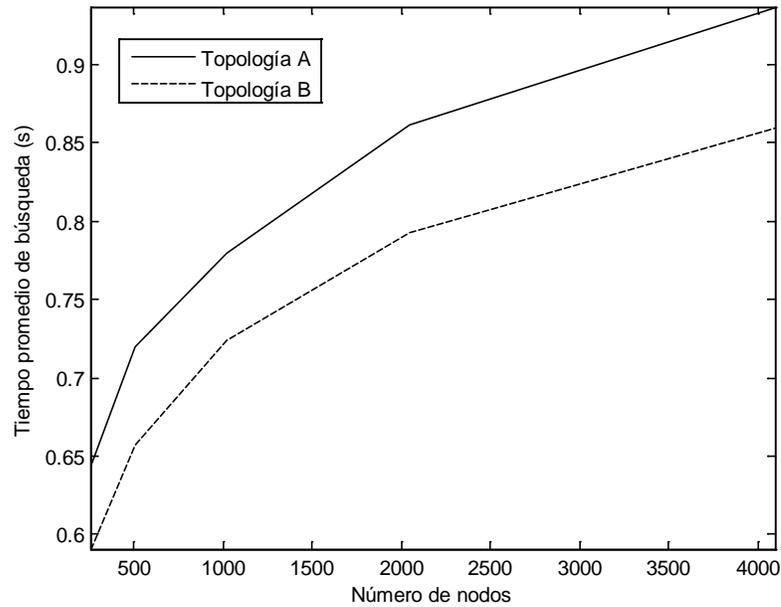


Figura 2.8. Retardo promedio de búsqueda de Chord sobre dos tipos de redes subyacentes.

En la Figura 2.8 se puede observar que el protocolo Chord tiene un mejor desempeño en la Topología B en base al tiempo promedio de búsqueda en la red sobrepuesta. Esto se debe a que la Topología B posee un mayor número de nodos de tránsito por lo que existen más conexiones entre nodos de tránsito y el retardo de los mensajes disminuye.

Por otro lado, se define el *factor de estiramiento* (*stretch factor*) como la razón entre la latencia promedio del enlace lógico sobre la latencia promedio del enlace físico, donde el enlace lógico se refiere al enlace virtual entre nodos en la red sobrepuesta. Es decir, si la función de la latencia promedio del enlace lógico y la latencia promedio del enlace físico entre dos nodos u y v , se denotan como $\Gamma(u, v)$ y $\gamma(u, v)$ respectivamente, entonces el factor de estiramiento, $F.E.$, se calcula de la siguiente manera para cada par de nodos $u, v \in G$:

$$F.E. = \frac{\Gamma(u, v)}{\gamma(u, v)}$$

El factor de estiramiento cuantifica el grado de correspondencia entre la topología de la red sobrepuesta y la topología de la red subyacente. Una manera alternativa de cuantificar el

grado de correspondencia entre estas dos redes se presenta en [17]. En la Figura 2.9 se muestra el factor de estiramiento que existe entre la topología de la red sobrepuesta Chord y la topología de la red subyacente.

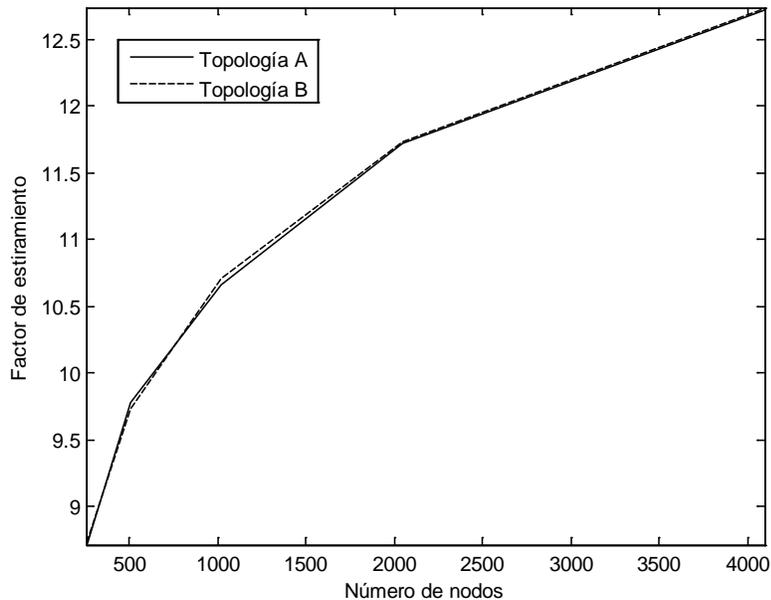


Figura 2.9. Factor de estiramiento entre la topología de red sobrepuesta Chord y la topología de la red subyacente.

De la Figura 2.9 se puede observar que ambas topologías presentan valores similares del factor de estiramiento, esto debido a que la selección de nodos y la asignación de identificadores se realizan de manera aleatoria. Asimismo se observa que incluso para valores pequeños del número de nodos, el valor del factor de estiramiento es alto, es decir, existe una gran disparidad entre la estructura de la red sobrepuesta y la red subyacente. Esto se debe a que la asignación de los identificadores se realiza de manera aleatoria y por lo tanto dos nodos cercanos en la red sobrepuesta pueden encontrarse a una gran distancia en la red subyacente. Una forma de reducir esta disparidad y de esta manera reducir la latencia promedio de las búsquedas, es con la construcción de una topología consciente de la topología subyacente, las cuales se estudiarán posteriormente en el Capítulo 3.

Referencias

- [1] I. Stoica, R. Morris, D. Karger, M. F. Kaashoek y H. Balakrishnan, «Chord: A Scalable Peer-to-peer Lookup Service for Internet Applications,» *SIGCOMM Comput. Commun. Rev.*, vol. 31, n° 4, pp. 149-160, 2001.
- [2] I. Stoica, R. Morris, D. Karger, M. F. Kaashoek y H. Balakrishnan, «Chord: a scalable peer-to-peer lookup protocol for Internet applications,» *Networking, IEEE/ACM Transactions on*, vol. 11, n° 1, pp. 17-32, 2003.
- [3] D. Karger, E. Lehman, T. Leighton, R. Panigrahy, M. Levine y D. Lewin, «Consistent Hashing and Random Trees: Distributed Caching Protocols for Relieving Hot Spots on the World Wide Web,» *Proceedings of the Twenty-ninth Annual ACM Symposium on Theory of Computing*, pp. 654-663, 1997.
- [4] L. Garces-Erice, E. W. Biersack, K. W. Ross, P. A. Felber y G. Urvoy-Keller, «Hierarchical Peer-to-Peer systems,» *Proc. ACM/IFIP International Conference on Parallel and Distributed Computing (Euro-Par)*, 2003.
- [5] K. Wehrle, M. Güneş y J. Gross, *Modeling and Tools for Network Simulation*, Berlin: Springer, 2010.
- [6] H. Haddadi, M. Rio, G. Iannaccone, A. Moore y R. Mortier, «Network topologies: inference, modeling, and generation,» *Communications Surveys & Tutorials, IEEE*, vol. 10, n° 2, pp. 48-69, 2008.
- [7] J. F. Kurose y K. W. Ross, *Computer Networking: A Top-Down Approach Featuring the Internet*, Boston: Pearson, 2013.
- [8] P. Van Mieghem, *Performance Analysis of Complex Networks and Systems*, Cambridge: Cambridge University, 2014.

- [9] P. Erdős y A. Rényi, «On Random Graphs I,» *Publicationes Mathematicae (Debrecen)*, vol. 6, pp. 290-297, 1959.
- [10] B. M. Waxman, «Routing of Multipoint Connections,» *IEEE Journal on Selected Areas in Comm.*, vol. 6, n° 9, p. 1617–1622, 1988.
- [11] H. Tangmunarunkit, R. Govindan, S. Jamin, S. Shenker y W. Willinger, «Network Topology Generators: Degree-based vs. Structural,» *SIGCOMM Comput. Commun. Rev.*, vol. 32, n° 4, pp. 147-159, 2002.
- [12] E. W. Zegura, K. L. Calvert y M. J. Donahoo, «A Quantitative Comparison of Graph-based Models for Internet Topology,» *IEEE/ACM Transactions on Networking*, vol. 5, n° 6, pp. 770-783, 1997.
- [13] M. Faloutsos, P. Faloutsos y C. Faloutsos, «On Power-law Relationships of the Internet Topology,» *SIGCOMM Comput. Commun. Rev.*, vol. 29, n° 4, pp. 251-262, 1999.
- [14] «Georgia Tech Internetwork Topology Models. GT-ITM,» [En línea]. Available: <http://www.cc.gatech.edu/projects/gtitm/>. [Último acceso: Septiembre 2015].
- [15] M. Bastian, S. Heymann y M. Jacomy, «Gephi: an open source software for exploring and manipulating networks.,» *International AAAI Conference on Weblogs and Social Media*, vol. 8, pp. 361-362, 2009.
- [16] A. Drozdek, *Data Structures and Algorithms in C++*, Boston, MA: Cengage Learning, 2013.
- [17] H. Rostami y J. Habibi, «Topology awareness of overlay P2P networks,» *Concurrency and Computation: Practice and Experience*, vol. 19, n° 7, pp. 999-1021, 2007.

Capítulo 3

3. Redes sobrepuestas P2P conscientes de la topología subyacente

En el presente capítulo se estudian los conceptos básicos de la construcción de redes sobrepuestas conscientes de la topología subyacente. De esta manera, se presentan diversas técnicas para generar información de proximidad en la red subyacente e introducirla en la red sobrepuesta. Posteriormente, se realiza la evaluación del desempeño de dos tipos de construcciones de redes sobrepuestas conscientes de la topología subyacente, las cuales son adaptaciones del protocolo Chord propuestas en el presente trabajo y presentan una mejora en el desempeño de la red en comparación con el protocolo Chord tradicional.

3.1 Introducción

Una red sobrepuesta inconsciente de la topología subyacente es capaz de controlar la secuencia de pares por la cual un mensaje se propaga antes de alcanzar su destino. Sin embargo, ignora completamente la manera en la cual los paquetes se conmutan a través de la infraestructura subyacente debido a la trayectoria impuesta por la red sobrepuesta [1]. Por lo tanto, una red sobrepuesta estructurada basada en una tabla hash distribuida es una red inconsciente de la topología subyacente en su diseño, ya que los identificadores de los nodos se eligen de manera aleatoria y cada nodo crea enlaces lógicos con sus nodos vecinos dentro del espacio de identificadores, sin tomar en cuenta la topología de red subyacente.

En la mayoría de las topologías de red sobrepuesta P2P, cuando un nodo se une a la red no es posicionado en el espacio de identificadores de manera óptima respecto a la red subyacente debido a que los identificadores de los nodos se asignan de manera aleatoria sin tomar en cuenta la topología subyacente. La asignación aleatoria de identificadores resulta en excelentes propiedades de balance de carga en los nodos, escalabilidad y robustez de la red sobrepuesta, como se mencionó en el Capítulo 1.

Sin embargo, la asignación aleatoria de identificadores provoca un *desajuste de topologías* entre la red sobrepuesta y la red subyacente. Esto significa que una distancia corta en el espacio de identificadores utilizado por la red sobrepuesta puede resultar en una distancia larga durante el enrutamiento de la red subyacente. Por esta razón, incluso cuando el nodo destino puede ser alcanzado en un número logarítmico de saltos en la red sobrepuesta, la distancia atravesada en la red subyacente durante el proceso de enrutamiento puede ser lejana de la óptima, lo que ocasiona un efecto negativo en el desempeño del enrutamiento de la red sobrepuesta [2].

Por otro lado, un enlace lógico entre dos pares de la red sobrepuesta comúnmente involucra múltiples enlaces en la red subyacente. Si las redes presentan desajuste de topologías entre ellas, un mensaje pueda atravesar el mismo enlace físico múltiples ocasiones durante su trayectoria, causando tráfico redundante y retardo adicional en la entrega de mensajes en la red sobrepuesta [3].

Por lo tanto, un aspecto fundamental al diseñar redes sobrepuestas de gran escala es incorporar información de la topología subyacente en su construcción para mejorar el desempeño en la entrega de mensajes a través de ella y evitar afectar la operación de otras aplicaciones que se ejecuten en la infraestructura de la red subyacente [4].

El proceso para utilizar de manera efectiva la información de la red subyacente en la construcción de la red sobrepuesta es el siguiente [5]:

- Obtener información de proximidad entre nodos de la red subyacente a través de alguna técnica de descubrimiento.

- Aplicar una técnica que introduzca la información de proximidad obtenida en la construcción de la red sobrepuesta, en el enrutamiento de mensajes entre nodos de la red o en los procesos de enrutamiento de búsqueda y recuperación de la información almacenada en la red.

A continuación se presentan las técnicas comúnmente utilizadas en la literatura para generar información de proximidad en la red y para introducirla en el diseño de una red sobrepuesta consciente de la topología subyacente.

3.2 Técnicas para generar información de proximidad en la red

La latencia en la red es la métrica de red utilizada para permitir a las redes sobrepuestas P2P tomar decisiones para optimizar el desempeño de aplicaciones mediante el empleo de la proximidad y la información topológica de los nodos en la red subyacente [2].

Existen varias métricas para estimar la distancia entre dos nodos en Internet en términos de la latencia en la red, por ejemplo, la longitud de la trayectoria IP, la longitud de la trayectoria a nivel de sistemas autónomos, la distancia geográfica real y mediciones del retardo ida y vuelta (RTT por *Round-Trip delay Time*) [6]. Sin embargo, las mediciones del RTT son comúnmente empleadas debido a que éstas son un indicador directo del desempeño visto entre nodos finales y pueden ser realizadas de manera sencilla, extremo a extremo y de una manera no intrusiva [4].

En la literatura se identifican tres técnicas para generar información de proximidad entre nodos dentro de una red [5]: búsqueda por anillo de expansión, agrupamiento utilizando nodos marca y sistemas de coordenadas globales.

- **Búsqueda por anillo de expansión**

Consiste en una técnica de inundación para medir el RTT entre un nodo y todos los otros nodos dentro de un radio definido en término de saltos en la red. Esta técnica se utiliza en la

construcción de una red sobrepuesta consciente de la topología subyacente en [7]. A pesar de que esta técnica utiliza un algoritmo simple para su implementación, puede introducir una carga excesiva de tráfico adicional en la red.

- **Sistemas de coordenadas globales**

Existen sistemas de coordenadas que incorporan a los nodos a un espacio cartesiano. Como en la técnica siguiente, se necesita el apoyo de nodos marca adicionales en la red. En primer lugar los nodos marca miden el RTT entre ellos mismos y utilizan esta información para calcular unas coordenadas en el espacio cartesiano para cada uno de ellos. Estas coordenadas se distribuyen entre los nodos pertenecientes a la red sobrepuesta, los cuales miden el RTT hacia estos nodos marca y calculan por sí mismos sus propias coordenadas en el espacio cartesiano en base al RTT y las coordenadas de los nodos marca, las cuales caracterizan la posición del nodo en la red. De esta manera, la proximidad en la red puede ser estimada utilizando la distancia euclidiana entre nodos en el espacio cartesiano. En [8] y en [9] se utiliza este enfoque de coordenadas globales para la construcción de una red sobrepuesta consciente de la topología subyacente.

- **Agrupamiento utilizando nodos marca**

En este esquema se asume la existencia de un conjunto de nodos bien conocido que se encuentran distribuidos a lo largo de la red subyacente y actúan como puntos de referencia en la red, los cuales son conocidos como *nodos marca*. El agrupamiento utilizando nodos marca asume que los nodos cercanos entre sí pueden tener latencias de red similares hacia algunos puntos de referencia distribuidos a través de la red. De esta forma, los nodos cercanos en la red subyacente se agrupan dentro de un vecindario del espacio de identificadores utilizado por la red sobrepuesta durante su formación.

Al utilizar este esquema, un nodo mide el RTT hacia cada uno de los nodos marca y los ordena de manera creciente respecto al valor de RTT. Por lo tanto, en base a los retardos medidos hacia los diferentes nodos marca, cada nodo tiene asociado un ordenamiento de

nodos marca. De esta manera, los nodos que comparten el mismo ordenamiento se encuentran cercanos unos a otro en términos de latencia en la red. La Figura 3.1 muestra un ejemplo de la manera en la que este esquema agrupa los nodos en la red de acuerdo al ordenamiento de nodos marca. Uno de los primeros trabajos que utiliza este esquema para la construcción de una red sobrepuesta consciente de la topología subyacente es [4]. De manera alternativa, se puede realizar el agrupamiento de los nodos de acuerdo al nodo marca más cercano.

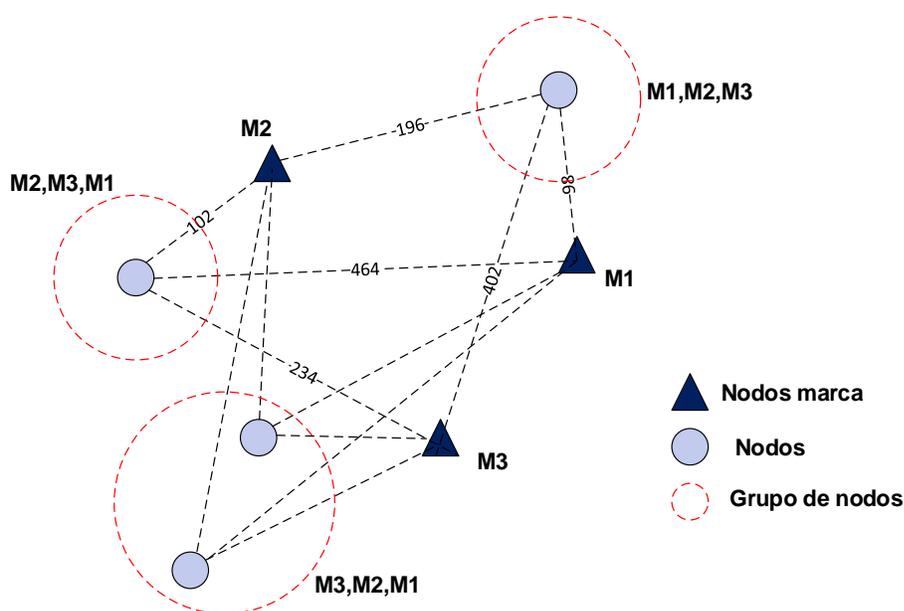


Figura 3.1. Ejemplo de agrupamiento de nodos utilizando nodos marca.

Este esquema de agrupamiento distribuido es escalable debido a que los nodos pueden descubrir de manera independiente el grupo al que pertenecen sin comunicación o coordinación con otros nodos en la red. Un nodo solamente necesita descubrir la distancia entre sí mismo y los nodos marca, lo cual puede realizar por sí mismo.

3.3 Técnicas para introducir información de proximidad en la red

De manera general, las técnicas utilizadas en la literatura para introducir en la red sobrepuesta cierto grado de consciencia acerca de la topología subyacente pueden dividirse en tres categorías: Enrutamiento basado en proximidad, Selección de nodos vecinos basados en proximidad y Disposición geográfica[10] [5]. A continuación se introducen cada uno de estos enfoques.

3.3.1 Enrutamiento basado en proximidad

En este enfoque la red sobrepuesta se construye de manera independiente de la topología subyacente. Sin embargo, el algoritmo de enrutamiento utiliza la información de proximidad en la red para decidir el siguiente salto en la trayectoria, es decir, en cada salto el nodo tiene oportunidad de elegir entre varios próximos saltos potenciales dentro de su tabla de enrutamiento. De esta manera, el nodo puede elegir como siguiente salto el nodo que se encuentre más cercano en la red subyacente o el nodo que representa el mejor compromiso entre proximidad y progreso en el espacio de identificadores. Esta última opción se puede realizar ponderando el progreso en el espacio de identificadores contra el costo de proximidad en la red [11].

La principal limitante de este enfoque es que su desempeño depende del número de opciones que posee el algoritmo de enrutamiento, el cual es proporcional al tamaño de la tabla de enrutamiento almacenada. Asimismo, dependiendo del algoritmo de enrutamiento, al seleccionar solamente a los candidatos con mayor proximidad, el número total de saltos en la trayectoria del mensaje puede incrementarse y no ser óptima.

A pesar de que el enrutamiento basado en proximidad presenta mejoras con respecto al enrutamiento inconsciente de la topología subyacente, es menos efectivo que utilizar el enfoque de disposición geográfica cuando se aplica en implementaciones del protocolo Chord y CAN [1].

3.3.2 Selección de nodos vecinos basados en proximidad

En este enfoque se construye una topología de red sobrepuesta consciente de la red subyacente debido a que el criterio de proximidad en la red se aplica al seleccionar los nodos vecinos en la red sobrepuesta y no solamente en el momento de la elección del próximo salto como en el enfoque anterior. De este modo, se eligen como entradas en la tabla de enrutamiento los nodos que se encuentran más cercanos en la red subyacente de entre aquellos con identificadores en la porción deseada del espacio de identificadores.

El éxito de esta técnica depende del grado de libertad con el que cuenta el protocolo de red sobrepuesta en la construcción de la tabla de enrutamiento sin afectar el número promedio de saltos en la trayectoria. Esta técnica es utilizada en el protocolo Pastry [12] cuando se construye el conjunto de nodos vecinos.

Una limitante de esta técnica es que no funciona de manera correcta con protocolos de red sobrepuesta como CAN [11] y Chord [13], los cuales requieren que las entradas en la tabla de enrutamiento se encuentren en puntos específicos del espacio de identificadores. Adicionalmente, el uso de esta técnica puede alterar la distribución uniforme de los nodos en el espacio de identificadores.

3.3.3 Disposición geográfica

Este enfoque también es conocido como *Asignación de Identificadores basado en la Topología Subyacente* debido a que intenta mapear el espacio de identificadores en la topología de red subyacente. De este modo, la asignación de los identificadores de los nodos se realiza de tal manera que los nodos que se encuentran cercanos en la topología subyacente se encuentren cercanos en el espacio de identificadores de la red sobrepuesta.

Algunos algoritmos [4] agrupan a nodos cercanos para reducir la latencia promedio de la red sobrepuesta. El agrupamiento de nodos utilizando nodos marca se puede utilizar para agrupar pares en la proximidad física.

Las implementaciones basadas en la disposición geográfica mejoran la eficiencia de búsqueda del sistema en general al reducir la latencia en las trayectorias y son especialmente efectivos cuando se aplican en implementaciones del protocolo CAN [1]. Sin embargo, el uso de esta técnica altera la distribución de nodos en el espacio de identificadores, lo cual puede resultar en problemas de balanceo de carga en la red sobrepuesta.

De la revisión anterior se puede concluir que el enfoque de disposición geográfica es el más adecuado para utilizar con el protocolo Chord debido a que no aumenta el número de entradas en la tabla de enrutamiento y puede ser utilizado sin realizar alguna modificación considerable al protocolo, con lo cual mantiene sus propiedades originales.

3.4 Construcción de redes sobrepuestas conscientes de la topología subyacente

En esta sección, se presenta la evaluación del desempeño de dos tipos de construcciones de redes sobrepuestas conscientes de la topología subyacente basadas en el protocolo Chord, las cuales presentan una mejora en su desempeño en comparación con el protocolo Chord tradicional de acuerdo a las métricas presentadas en el Capítulo 2. En estas construcciones se aplica la técnica de agrupamiento utilizando nodos marca para introducir la información de proximidad en la topología subyacente, donde M es el número de nodos marca utilizado.

Este enfoque de construcción ha sido empleado previamente en [4] para la construcción de redes sobrepuestas conscientes de la topología subyacente en una arquitectura de red sin jerarquía, sin embargo, en este trabajo se utiliza como base el protocolo CAN. El protocolo CAN es un protocolo que utiliza un espacio n -dimensional, lo cual aumenta la complejidad de una posible implementación. Por esta razón, en el presente trabajo se desarrollan dos tipos de formación de la red sobrepuesta que emplean esta técnica en una red que utiliza el protocolo Chord, el cual presenta buenas características de desempeño sin elevar la complejidad del sistema.

En las construcciones propuestas se introduce la consciencia de la topología subyacente al dividir el espacio de identificadores del protocolo Chord, es decir el anillo Chord, en un cierto número de arcos. Cada uno de estos arcos representa una porción de la red que contiene cierto número de nodos, es decir, los nodos pertenecientes a un mismo arco se encuentran dentro de la misma porción de la red.

3.4.1 Construcción de red sobrepuesta consciente de la topología subyacente simple (RS-CTS-S)

- **Arquitectura de red.**

En esta construcción, cuando un nodo ingresa a la red sobrepuesta, éste mide las latencias hacia cada uno de los nodos marca. Posteriormente ordena de manera ascendente las latencias obtenidas y elige unirse al arco correspondiente al nodo marca más cercano. La asignación del identificador dentro del arco correspondiente se realiza de manera aleatoria.

La Figura 3.2a ejemplifica la manera en la cual una topología de red compuesta de 100 nodos se divide en M zonas de acuerdo a la posición de los nodos en la red subyacente utilizando el procedimiento descrito anteriormente, donde se utilizan $M = 4$ nodos marca. De manera análoga, el espacio de identificadores, es decir el anillo Chord, se divide en $M = 4$ arcos como se muestra en la Figura 3.2b. De esta forma, los nodos que se encuentran dentro de la zona i en la red subyacente obtienen un identificador que se encuentra dentro del arco i en el anillo Chord.

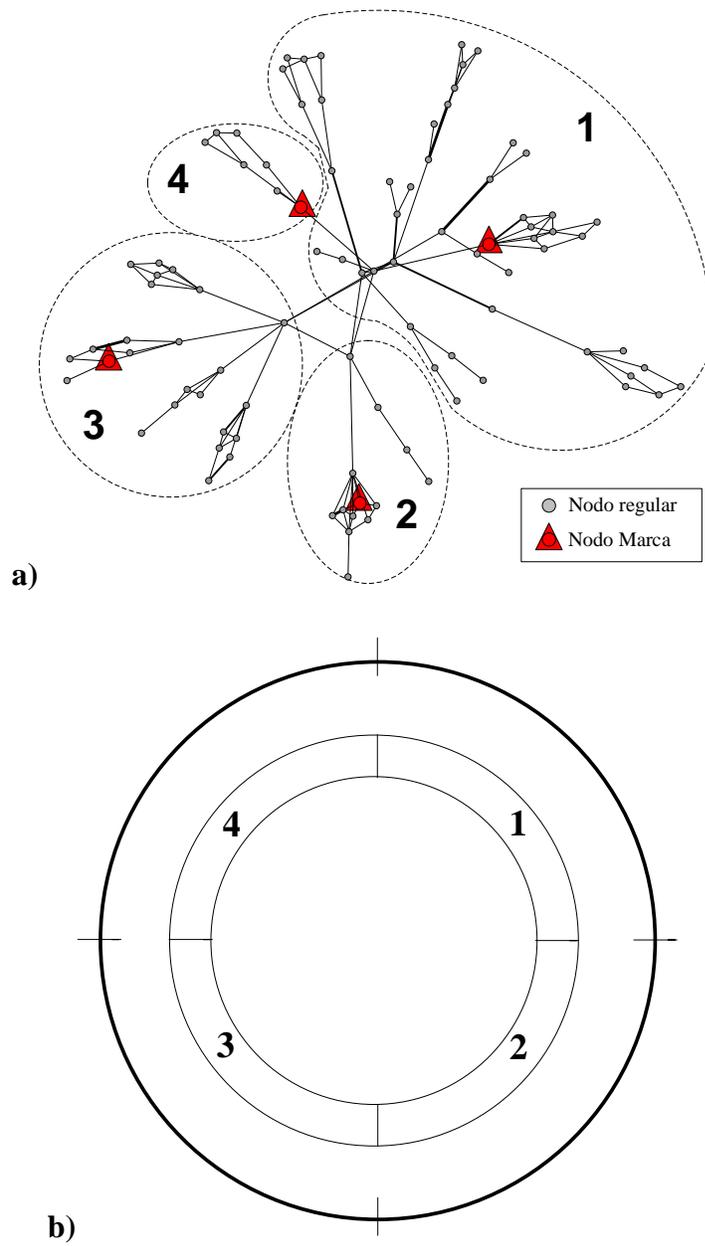


Figura 3.2. a) Ejemplo de la división de una topología de red en $M=4$ zonas. b) Ejemplo de la división del anillo Chord en $M=4$ arcos.

- **Metodología de simulación.**

Con el fin de evaluar el comportamiento y desempeño de la construcción RS-CTS-S, se desarrollaron experimentos por medio de simulación de esta construcción y del protocolo Chord tradicional, el cual es inconsciente de la topología subyacente, para $N = 2^k$ nodos. El simulador se desarrolló en lenguaje C, el cual cumple con las características del protocolo Chord como se mostró en la sección 2.1.4. Asimismo, la topología de red subyacente se genera con la herramienta GT-ITM [14]. Para esta simulación se ocupó la Topología A presentada en la sección 2.3 compuesta por 4525 nodos. Durante la simulación, se varió el valor de k desde 8 hasta 12 y se realizaron experimentos separados para cada valor de k utilizando 8 nodos marca. En un experimento específico, cada nodo elige de manera aleatoria un conjunto de llaves del sistema para realizar su búsqueda.

- **Resultados de simulación.**

A continuación, en la Figura 3.3 se muestra la comparación del desempeño del protocolo Chord tradicional y la construcción RS-CTS-S de acuerdo a las siguientes métricas: número promedio de saltos por búsqueda, tiempo promedio de búsqueda y factor de estiramiento.

En la Figura 3.3a se puede observar que el número de saltos en RS-CTS-S es ligeramente menor al de Chord tradicional. Esto se debe a que al realizar la construcción consciente de la topología subyacente existe un desbalance de carga en la red, es decir, existen arcos con mayor número de nodos en ellos lo cual provoca que se altere la distribución de los nodos en el espacio de identificadores. Por otro lado, en la Figura 3.3b se observa que el tiempo de búsqueda promedio disminuyó cerca del 30% debido a la construcción consciente. Esto puede confirmarse observando la Figura 3.3c, la cual muestra que el factor de estiramiento también se redujo aproximadamente en un porcentaje similar.

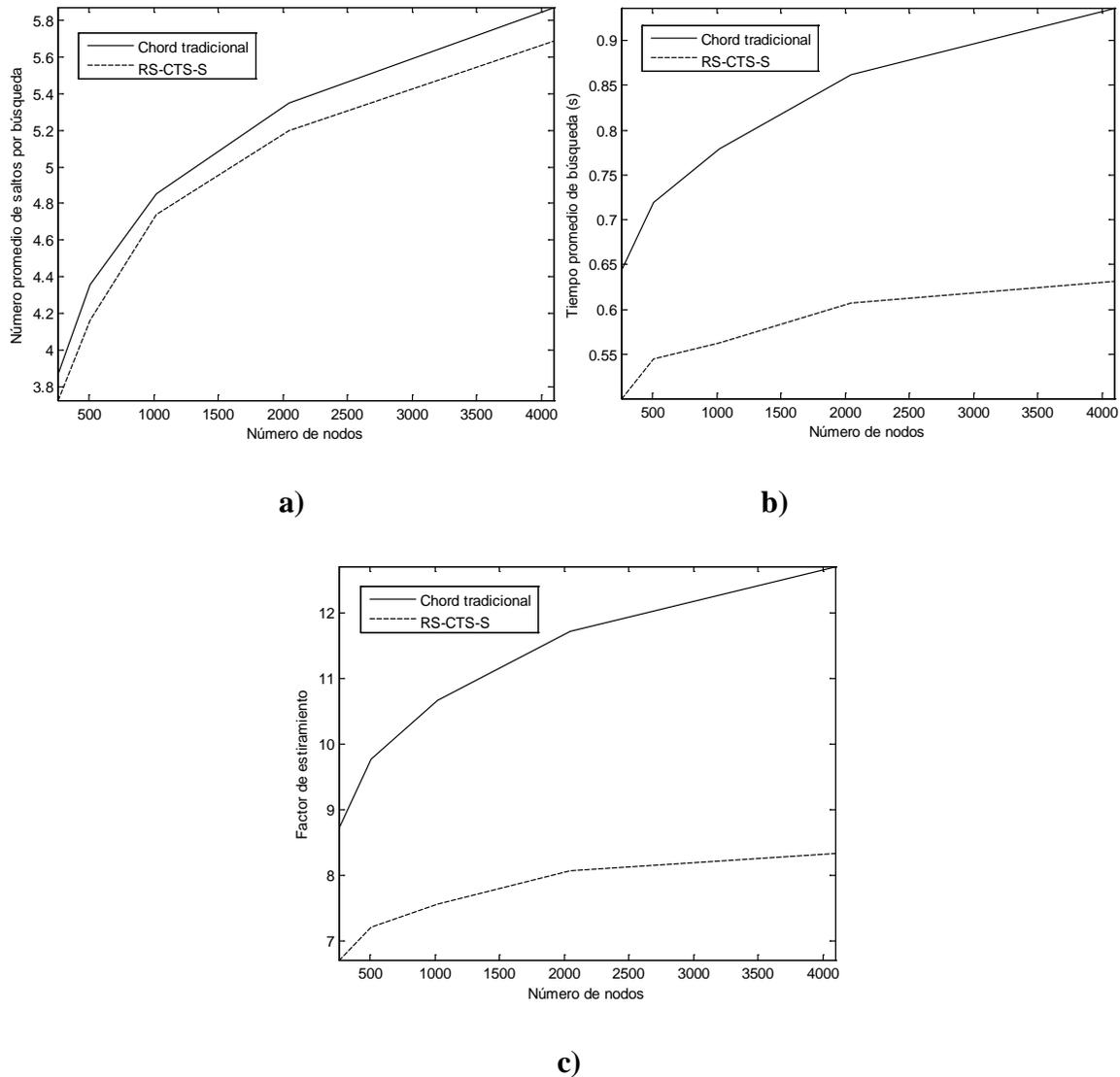


Figura 3.3. Comparación del desempeño del protocolo Chord tradicional y la construcción RS-CTS-S en base a: **a)** Número promedio de saltos por búsqueda, **b)** Tiempo promedio de búsqueda y **c)** Factor de estiramiento.

Como se puede observar de la comparación de desempeño entre la construcción RS-CTS-S y el protocolo Chord tradicional, la construcción RS-CTS-S presenta una mejora en el desempeño del sistema de acuerdo a las métricas presentadas. Sin embargo, el anillo se divide en un número pequeño de arcos y la asignación de los identificadores dentro de estos arcos es aleatoria. Esto quiere decir que las zonas en las cuales se divide la red pueden ser

muy grandes y nodos que se encuentran en los extremos de estas zonas pueden ser colocados como nodos contiguos en la red sobrepuesta, por lo cual el grado de consciencia de la red no es suficiente.

Bajo este enfoque es posible mejorar el desempeño del sistema al aumentar el número de marcas y así aumentar el número de arcos. Sin embargo, el aumento en el número de marcas representa un aumento en la complejidad del sistema, así como un aumento en el volumen de tráfico adicional en la red debido a las mediciones del RTT en la formación de la red sobrepuesta.

Por esta razón, en el presente trabajo se plantea una alternativa para aumentar el número de arcos en el anillo Chord manteniendo el mismo número de marcas. Esta propuesta de formación de la red sobrepuesta se evalúa por medio de la construcción RS-CTS-M.

3.4.2 Construcción de red sobrepuesta consciente de la topología subyacente mejorada (RS-CTS-M)

- **Arquitectura de red.**

En esta construcción, cuando un nodo ingresa a la red sobrepuesta, éste mide las latencias hacia cada uno de los nodos marca. Posteriormente ordena de manera ascendente las latencias obtenidas obteniendo un ordenamiento de marcas. Para obtener el arco al que pertenece, el nodo utiliza el Algoritmo 1, el cual utiliza como entrada el ordenamiento de marcas. La asignación del identificador dentro del arco correspondiente se realiza de manera aleatoria. El Algoritmo 1 es un algoritmo distribuido ya que cualquier nodo puede ejecutarlo de manera independiente con la información propia, sin realizar algún intercambio adicional de mensajes.

<p>Algoritmo 1 Obtención del número de arco correspondiente</p> <p>Entrada: M (Número de nodos marca), $P = \{p_0, \dots, p_{M-1}\}$ (Vector de ordenamiento de marcas).</p> <p>Salida: $NArco$ (Número de arco al que pertenece el nodo).</p> <p>Inicializar $NArco = 0$ y los vectores auxiliares $A = \{a_0 \leftarrow 0, a_1 \leftarrow 1, a_2 \leftarrow 2, \dots, a_{M-1} \leftarrow M - 1\}$ y $F = \{f_0 \leftarrow 0, f_1 \leftarrow 0, f_2 \leftarrow 0, \dots, f_{M-1} \leftarrow 0\}$.</p> <p>Para $i = 0, 1, \dots, M - 2$ hacer</p> <p style="padding-left: 2em;">$contador \leftarrow 0$ y $j \leftarrow 0$</p> <p style="padding-left: 2em;">Mientras p_i diferente de a_j hacer</p> <p style="padding-left: 4em;">Si f_j es igual a 0 entonces</p> <p style="padding-left: 6em;">$contador \leftarrow contador + 1$</p> <p style="padding-left: 4em;">Fin Si</p> <p style="padding-left: 2em;">$j \leftarrow j + 1$</p> <p style="padding-left: 2em;">Fin Mientras</p> <p style="padding-left: 2em;">$f_j \leftarrow 1$ y $NArco \leftarrow NArco + (contador)(M - 1 - i)!$</p> <p>Fin Para</p> <p>Regresa $NArco$</p>
--

Algoritmo 1. Algoritmo para la obtención del arco correspondiente en la construcción RS-CTS-M.

Al utilizar el Algoritmo 1 para obtener el arco al que pertenece cada nodo en la red, la red se divide en $M!$ zonas, donde los nodos pertenecientes a la misma zona comparten el mismo ordenamiento de marcas. De la misma manera, el espacio de identificadores se divide en $M!$ arcos.

La Figura 3.4a ejemplifica la manera en la cual una topología de red compuesta de 100 nodos se divide en $M! = 24$ zonas de acuerdo a la posición de los nodos en la red subyacente, donde se utilizan $M = 4$ nodos marca. De manera análoga, el espacio de identificadores, es decir el anillo Chord, se divide en $M! = 24$ arcos como se muestra en la Figura 3.4b. De

esta forma, los nodos que se encuentran dentro de la zona i , obtienen un identificador que se encuentra dentro del arco i .

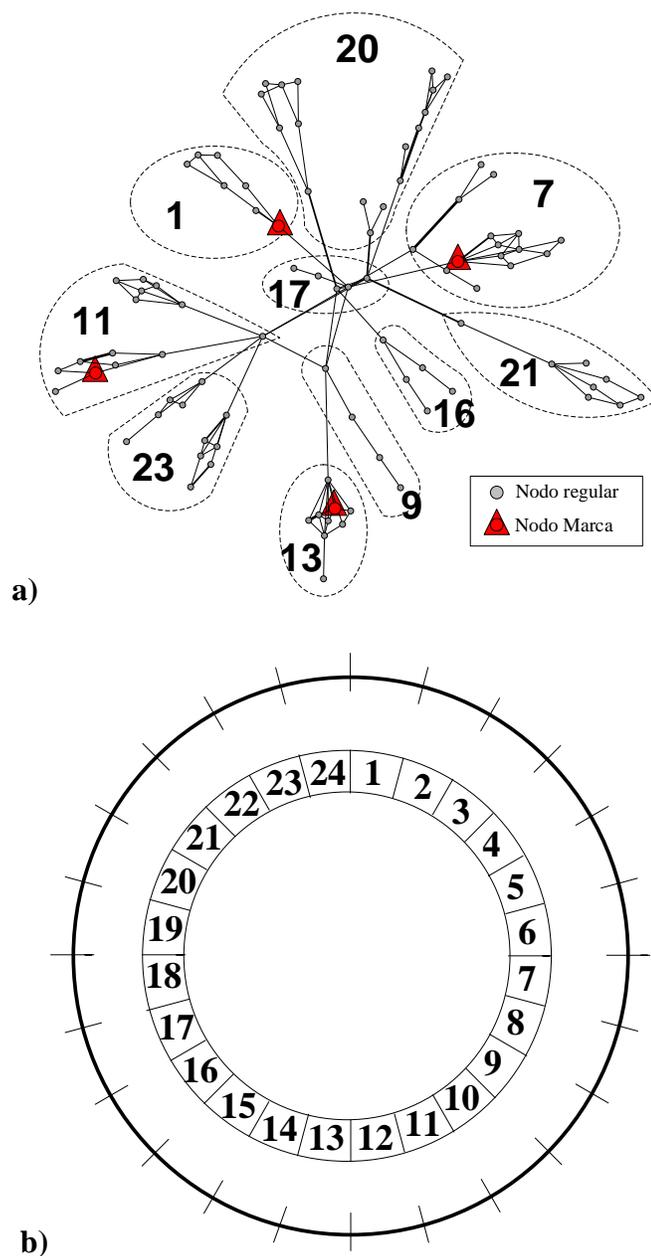


Figura 3.4. a) Ejemplo de la división de una topología de red en $M!=24$ zonas. b) Ejemplo de la división del anillo Chord en $M!=24$ arcos.

En el ejemplo que presenta la Figura 3.4 se puede apreciar que existen arcos en el anillo Chord que se mantienen vacíos porque no existen nodos en las zonas asociadas a estos arcos, es decir, no existen nodos en la red subyacente que poseen un ordenamiento específico de los nodos marca. Una consecuencia de esta situación es que los nodos no se encuentran distribuidos de manera uniforme en el espacio de identificadores, lo cual puede alterar el número promedio de saltos por búsqueda y el balance de la carga de los nodos.

- **Metodología de simulación.**

Con el fin de evaluar el comportamiento y desempeño de la construcción RS-CTS-M, se desarrollaron experimentos por medio de simulación de esta construcción, la construcción RS-CTS-M y del protocolo Chord tradicional, el cual es inconsciente de la topología subyacente, para $N = 2^k$ nodos. Para esta simulación se ocupó la Topología A presentada en la sección 2.3 compuesta por 4525 nodos. Durante la simulación, se varió el valor de k desde 8 hasta 12 y se realizaron experimentos separados para cada valor de k utilizando 8 nodos marca. En un experimento específico, cada nodo elige de manera aleatoria un conjunto de llaves del sistema para realizar su búsqueda.

- **Resultados de simulación.**

En la Figura 3.5 se muestra la comparación del desempeño del protocolo Chord tradicional, la construcción RS-CTS-S y la construcción RS-CTS-M de acuerdo a las siguientes métricas: número promedio de saltos por búsqueda, tiempo promedio de búsqueda y factor de estiramiento.

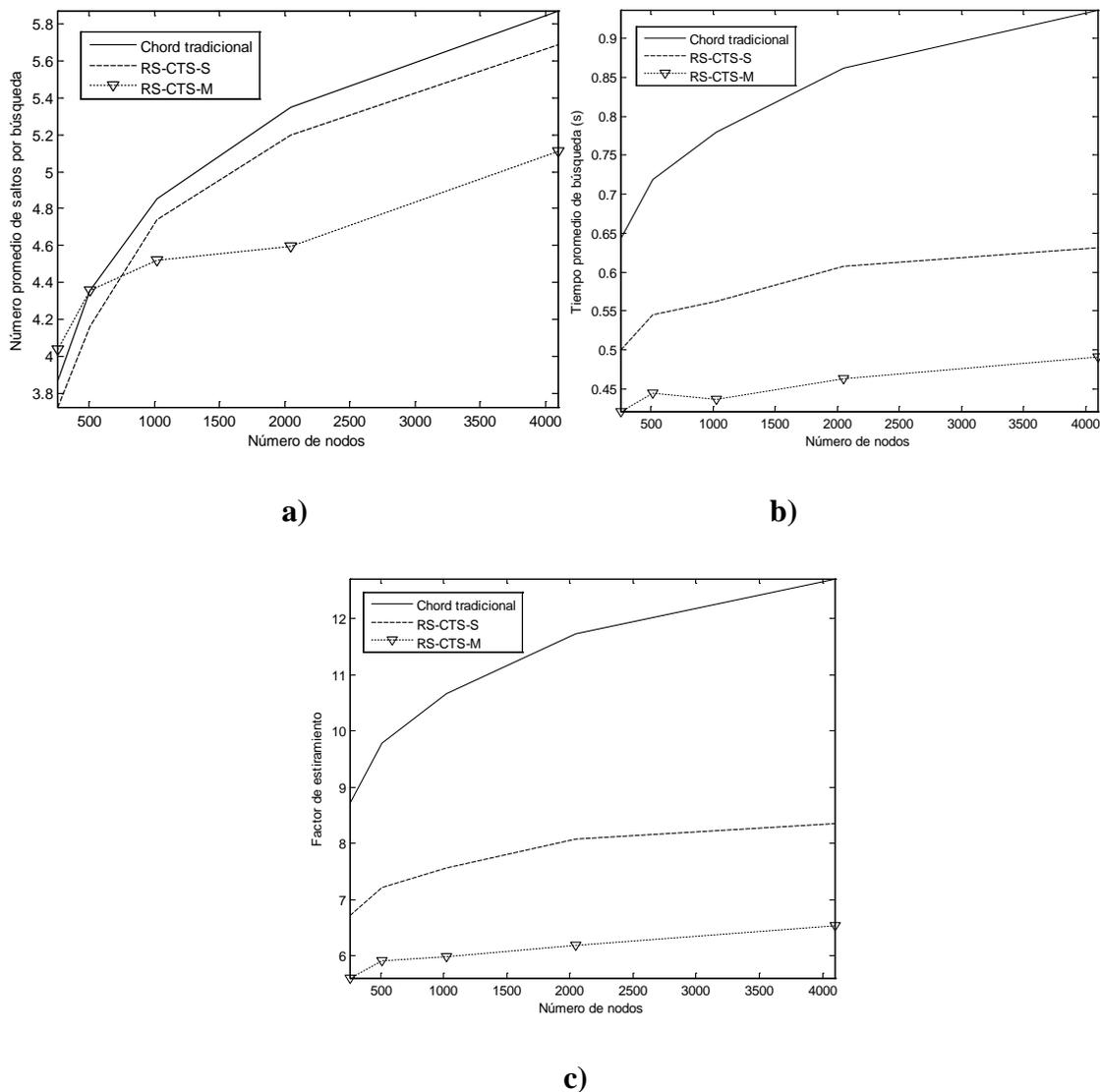


Figura 3.5. Comparación del desempeño del protocolo Chord tradicional, la construcción RS-CTS-S y la construcción RS-CTS-M en base a: a) Número promedio de saltos por búsqueda, b) Tiempo promedio de búsqueda y c) Factor de estiramiento.

Como se puede observar de la Figura 3.5a, el número de saltos de RS-CTS-M es todavía menor al de RS-CTS-S debido a que existe mayor desbalance de carga en los nodos, ya que al existir un gran número de arcos muchos de ellos pueden quedar vacíos y otros contener un gran número de nodos. Por otro lado, en la Figura 3.5b se observa que el tiempo de búsqueda promedio en RS-CTS-M disminuyó cerca de 20% más en comparación a RS-CTS-

S y cerca del 50% en comparación de Chord tradicional. Esto puede confirmarse observando la Figura 3.5c, la cual muestra que el factor de estiramiento se redujo en un porcentaje similar. Esto quiere decir que el grado de correspondencia entre la topología de red sobrepuesta y la topología de red subyacente es superior en la construcción RS-CTS-M.

Como se puede observar de la comparación de desempeño entre la RS-CTS-M, la construcción RS-CTS-S y el protocolo Chord tradicional, la construcción RS-CTS-M presenta un mejor desempeño con respecto al tiempo promedio de búsqueda y el factor de estiramiento. Sin embargo, también se pudo observar que el desbalance de la carga en los nodos también aumentó, por lo que existe un compromiso entre el aumento del grado de consciencia de la red subyacente y el balance de carga en los nodos de la red.

3.5 Elección de nodos marca

Los resultados generados en la sección anterior se obtuvieron al seleccionar de manera aleatoria a los nodos marca en la red. Sin embargo, se observó que el desempeño del agrupamiento utilizando nodos marca se encuentra relacionado con el posicionamiento de los nodos marca a lo largo de la red subyacente. El enfoque de elección de marcas de manera aleatoria es conveniente si los nodos se encontraran distribuidos de manera uniforme a lo largo de la red subyacente. Sin embargo, como se mencionó en la Sección 2.2, esta propiedad no se presenta en redes como Internet, por lo cual este enfoque de elección de nodos marca no es el adecuado.

A continuación se analizan diferentes enfoques para la elección de los nodos marca y se evalúa su impacto en el desempeño de las redes sobrepuestas conscientes de la topología subyacente analizadas en la sección anterior. Se analizan tres diferentes elecciones de nodos marca: Elección de nodos marca de manera aleatoria, Elección de nodos marca separados por un retardo determinado y la Elección de nodos marca en base al grado del nodo.

- **Elección de nodos marca de manera aleatoria.**

Esta elección se utilizó para generar los resultados presentados en la Sección 3.4 y consiste en elegir a los M nodos marca de manera aleatoria de entre todos nodos de la topología subyacente. De esta manera, los nodos marca se encuentran distribuidos de manera uniforme a lo largo de la red subyacente.

Esta elección puede implementarse mediante el uso de marcas definidas por su identificador como las presentadas en [15]. En este caso, dentro de la red existen identificadores bien conocidos, los cuales convierten al nodo responsable del identificador en ese momento en un nodo marca.

- **Elección de nodos marca separados por un retardo determinado**

Uno de los primeros trabajos en utilizar el enfoque de nodos marca para realizar la construcción de una red sobrepuesta consciente de la topología subyacente se presenta en [4], en el cual se propone elegir a los nodos marca requeridos de manera aleatoria con la única condición de que se encuentren separados entre ellos por un cierto número de saltos en la red subyacente. Sin embargo, debido a que la métrica utilizada para estimar la distancia entre dos nodos es la latencia en la red subyacente y no el número de saltos en la red subyacente, en el presente trabajo se propone que los nodos marca se encuentren separados entre ellos por un determinado valor de latencia en la red a diferencia del enfoque utilizado en [4].

En esta elección se eligen como nodos marca M nodos de la red subyacente, donde cada uno de los nodos marca se encuentra separado de los demás al menos un determinado valor de latencia.

- **Elección de nodos marca en base al grado del nodo.**

Debido a que los nodos no se encuentran distribuidos de manera uniforme en la red subyacente y en su lugar existen concentraciones de nodos, en el presente trabajo se

determinó que al utilizar el grado del nodo como parámetro de elección de los nodos marca se obtiene una mejora en el desempeño de la red. El grado del nodo se define como el número de conexiones que posee en la red subyacente. En esta elección de nodos marca, se eligen como nodos marca M nodos de la red subyacente que cumplen con la condición de poseer un cierto valor de grado.

3.5.1 Optimización de la elección de nodos marca

La elección de nodos marca es similar al problema de agrupamiento que ha sido estudiado anteriormente [16]. Un algoritmo de agrupamiento conocido por alcanzar buenos resultados para agrupar conjuntos grandes de datos en una variedad de aplicaciones es el algoritmo *k-means* [17].

El objetivo del algoritmo *k-means* es dividir a todos los nodos de la red en K grupos, agrupando a los nodos cercanos en la red sobrepuesta. Adoptando la nomenclatura para describir un grafo utilizada en la Sección 2.2.1, esto es $V = \{V_i | i = 1, 2, \dots, K\}$, donde $\bigcup_{i=1}^K V_i = V$ y $\bigcap_{i=1}^K V_i = \emptyset$. En este algoritmo el número de grupos K se mantiene fijo e inician de manera aleatoria. Cada nodo en la red se asigna al grupo de su nodo marca más cercano en base a la medida particular de proximidad elegida $d(u, v)$, la cual en este caso es el retardo en la red entre dos nodos u y v . Una vez que los grupos se han formado, los nodos marca se actualizan para cada grupo. El algoritmo repite de manera iterativa estos dos pasos mencionados hasta que las nuevos nodos marca calculados no cambien.

A continuación, en el Algoritmo 2, se presenta el pseudocódigo del algoritmo propuesto para la optimización de la elección de nodos marca, el cual es una adaptación del algoritmo *k-means*.

Un factor que puede impactar el desempeño del Algoritmo 2, así como el del algoritmo *k-means*, es la elección inicial de los nodos marca, ya que el algoritmo converge a un mínimo local. Por lo tanto, se introducen como entrada las marcas elegidas en base a su nodo como método de selección inicial de nodos marca, ya que es el que presenta mejor desempeño

Algoritmo 2 Optimización de nodos marca
Entrada: K (Número de grupos), $C = \{c_1, \dots, c_k\}$ (Conjunto de nodos marca iniciales).
Salida: $M = \{m_1, \dots, m_k\}$ (Conjunto de nodos marca finales).
<p>Paso 1: Para cada $v_i \in V - C$:</p> <p style="padding-left: 40px;">Medir el retardo entre v_i y c_j, es decir:</p> $d(v_i, c_j), j = 1, \dots, K$ <p style="padding-left: 40px;">Añadir v_i al conjunto V_S si satisface:</p> $d(v_i, c_S) = \min\{d(v_i, c_j) j = 1, \dots, K\}$ <p>Paso 2: Actualizar los nodos marca.</p> <p>Paso 3: Si $c_j \neq c_j'$ Regresar al Paso 1.</p> <p>Paso 4: $M \leftarrow C$, es decir, $m_i = c_i$ para $i = 1, \dots, K$.</p> <p>Paso 5: Salida $M = \{m_1, \dots, m_k\}$.</p>

Algoritmo 2. Algoritmo de optimización de nodos marca.

.Como se puede observar, el Algoritmo 2 requiere del conocimiento global de las latencias entre todos los nodos de la red y claramente no es práctico para su implementación en una red como Internet. Sin embargo, sirve como un límite inferior potencial para evaluar las técnicas de elección de nodos marca presentadas anteriormente.

3.5.2 Comparación de desempeño

A continuación en la Figura 3.6 se muestra una comparación del impacto que tiene la elección de los nodos marca presentadas anteriormente en el desempeño de una red sobrepuesta consciente de la topología subyacente como la construcción RS-CTS-M de acuerdo a las siguientes métricas: número promedio de saltos por búsqueda, tiempo promedio de búsqueda y factor de estiramiento. En esta comparación se muestra el desempeño de los siguientes métodos de elección de nodos marca: Elección de nodos marca de manera aleatoria (EA), Elección de nodos marca separados por un retardo determinado

(ER), Elección de nodos marca en base al grado del nodo (EG) y el resultado del Algoritmo de optimización de elección de nodos marca (EG-OPT).

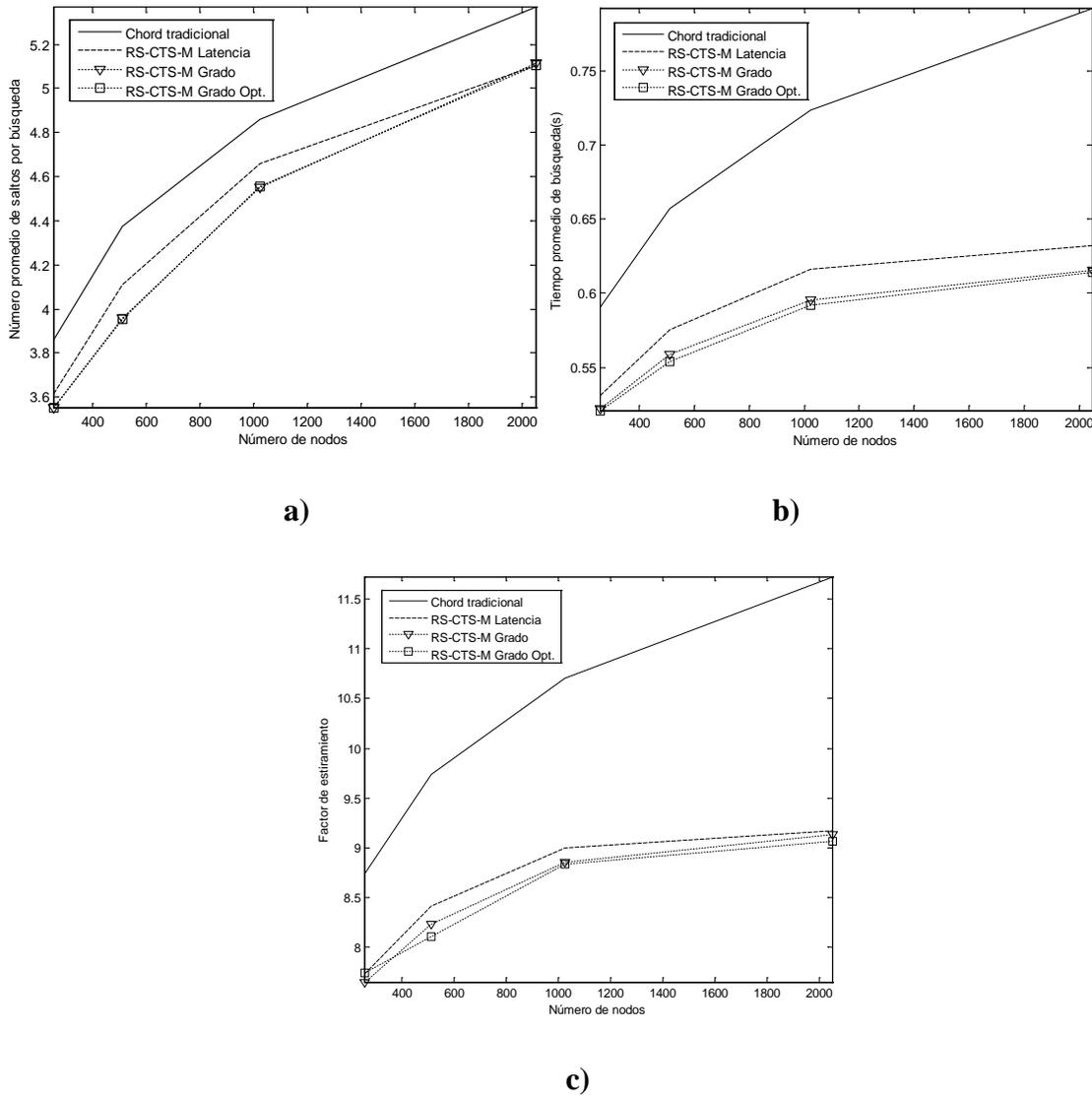


Figura 3.6. Comparación del desempeño de la RS-CTS-M con tres diferentes conjuntos de marcas en base a: a) Número promedio de saltos por búsqueda, b) Tiempo promedio de búsqueda y c) Factor de estiramiento.

Como se puede observar en la Figura 3.6 el desempeño mostrado por la construcción con la elección de nodos marca en base al grado del nodo es superior al presentado por los el resto

de los enfoques. Esto se debe a que existen concentraciones de nodos en la red subyacente, lo cual no se considera en los enfoques anteriores, por lo cual al elegir a los nodos marca que tienen un mayor número de conexiones, la agrupación de los nodos refleja de mejor manera las conexiones presentes en la red subyacente.

Por otro lado, los resultados obtenidos en la construcción utilizando el algoritmo de optimización son muy similares a los obtenidos por medio de la elección de nodos marca basados en el grado del nodo a pesar de que el algoritmo de optimización este último requiere del conocimiento global de las latencias entre todos los nodos de la red. Por lo que la elección de marcas propuesta presenta buenas características de desempeño debido a que se acerca al límite impuesto por el algoritmo de optimización.

Debido a que las construcciones evaluadas en este capítulo presentan diseños de redes sobrepuestas sin jerarquía, poseen ciertas limitaciones. Por ejemplo, al modificar el número de marcas en el sistema debe existir una reasignación total de los identificadores en la red sobrepuesta, lo cual generaría un excesivo volumen de tráfico adicional en una implementación real. Este tipo de limitaciones pueden superarse mediante el uso de una red sobrepuesta jerárquica como las que se analizan en el Capítulo 4.

Referencias

- [1] V. Moustakas, H. Akcan, M. Roussopoulos y A. Delis, «Alleviating the topology mismatch problem in distributed overlay networks: A survey,» *Journal of Systems and Software*, vol. 113, pp. 216-245, 2016.
- [2] J. F. Buford, H. H. Yu y E. K. Lua, P2P networking and applications, Amsterdam: Elsevier/Morgan Kaufmann, 2009.
- [3] H. Rostami y J. Habibi, «Topology awareness of overlay P2P networks,» *Concurrency and Computation: Practice and Experience*, vol. 19, n° 7, pp. 999-1021, 2007.
- [4] S. Ratnasamy, M. Handley, R. Karp y S. Shenker, «Topologically-aware overlay construction and server selection,» *Twenty-First Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings. IEEE*, vol. 3, pp. 1190-1199, 2002.
- [5] X. Shen, Handbook of Peer-to-Peer Networking, New York: Springer, 2010.
- [6] B. Huffaker, M. Fomenkov, D. Plummer, D. Moore y K. Claffy, «Distance metrics in the Internet,» *Proc. of IEEE international telecommunications symposium (ITS)*, vol. 19, n° 7, pp. 999-1021, 2002.
- [7] S. Ren, L. Guo, S. Jiang y X. Zhang, «SAT-Match: a self-adaptive topology matching method to achieve low lookup latency in structured P2P overlay networks,» *Parallel and Distributed Processing Symposium, 2004. Proceedings. 18th International*, pp. 83-91, 2004.
- [8] Z. Xu, C. Tang y Z. Zhang, «Building topology-aware overlays using global soft-state,» *Distributed Computing Systems, 2003. Proceedings. 23rd International Conference on*, pp. 500--508, 2003.

-
- [9] H. Shen, G. Liu y L. Ward, «A Proximity-Aware Interest-Clustered P2P File Sharing System,» *Parallel and Distributed Systems, IEEE Transactions on*, vol. 26, n° 6, pp. 1509-1523, 2015.
- [10] M. Castro, P. Druschel, Y. C. Hu y A. Rowstron, «Topology-aware routing in structured peer-to-peer overlay networks,» *Future directions in distributed computing*, pp. 103-107, 2003.
- [11] S. Ratnasamy, P. Francis, M. Handley, R. Karp y S. Shenker, «A Scalable Content-addressable Network,» *SIGCOMM Comput. Commun. Rev.*, vol. 31, n° 4, pp. 161-172, 2001.
- [12] A. Rowstron y P. Druschel, «Pastry: Scalable, Decentralized Object Location, and Routing for Large-Scale Peer-to-Peer Systems,» *Proceedings of the IFIP/ACM International Conference on Distributed Systems Platforms Heidelberg*, pp. 329-350, 2001.
- [13] I. Stoica, R. Morris, D. Karger, M. F. Kaashoek y H. Balakrishnan, «Chord: a scalable peer-to-peer lookup protocol for Internet applications,» *Networking, IEEE/ACM Transactions on*, vol. 11, n° 1, pp. 17-32, 2003.
- [14] «Georgia Tech Internetwork Topology Models. GT-ITM,» [En línea]. Available: <http://www.cc.gatech.edu/projects/gtitm/>. [Último acceso: Septiembre 2015].
- [15] Y. Yun-Shuai, M. Yu-Ben y S. Ce-Kuen, «Improving the Lookup Performance of Chord Network by Hashing Landmark Clusters,» *Networks, 2006. ICON '06. 14th IEEE International Conference on*, vol. 2, pp. 1-4, 2006.
- [16] J. Hong-Jun y D. Li-Juan, «Nodes Clustering Method in Large-Scale Network,» *Wireless Communications, Networking and Mobile Computing (WiCOM), 2012 8th International Conference on*, 2012.

- [17] C. C. Aggarwal, *Data Clustering: Algorithms and Applications*, Hoboken: CRC Press, 2013.

Capítulo 4

4. Redes sobrepuestas P2P jerárquicas

En el presente capítulo se presentan los fundamentos básicos de redes sobrepuestas jerárquicas. Asimismo, se presenta una clasificación de este tipo de redes de acuerdo a su arquitectura de red. Para evaluar el desempeño de este tipo de redes se realiza la construcción de una red jerárquica que utiliza como base el protocolo Chord. Posteriormente, se realiza la construcción de una red jerárquica consciente de la topología subyacente, la cual introduce la información de proximidad de la red subyacente utilizando las propuestas presentadas en el Capítulo 3.

4.1 Introducción

Las redes sobrepuestas estructuradas basadas en una tabla hash distribuida (DHT) como Chord [1], CAN [2] y Pastry [3], las cuales se analizaron en el Capítulo 1, se denominan *redes sobrepuestas sin jerarquía (flat overlay network)* debido a que emplean el mismo espacio para los identificadores de nodo y los identificadores de llave del sistema, además de que asumen que la distribución de los nodos y las llaves en el espacio de identificadores es uniforme, lo que proporciona escalabilidad y balance de carga en la red sobrepuesta [4].

Sin embargo, en una implementación de red sobrepuesta pueden existir nodos con capacidades diferentes. Por un lado, existen nodos cuyos tiempos de sesión son cortos y no

contribuyen de manera significativa a la operación de la red, denominados *nodos regulares*. Por otro lado, existen nodos estables y con más capacidad de procesamiento y almacenamiento, los cuales son capaces de tener más responsabilidad en la operación de la red, denominados *supernodos*.

De esta manera, las redes sobrepuestas sin jerarquía consideran a todos los participantes iguales en funcionalidad, es decir, todos los pares asumen roles iguales y tienen la misma responsabilidad en las operaciones de la red independientemente de sus capacidades. Si la asignación de tareas es igualitaria entre estos dos tipos de nodos, a medida que el sistema crece los nodos regulares pueden comprometer de manera seria la operación de la red entera, así como su escalabilidad [5].

En otras palabras, las redes sobrepuestas basadas en DHT sin jerarquía no toman en cuenta la heterogeneidad en los nodos y asignan responsabilidades iguales a nodos heterogéneos, lo cual es ineficiente. Por lo tanto, introducir jerarquía en la red sobrepuesta es un mecanismo común para hacer frente a la complejidad, escalabilidad, y heterogeneidad en este tipo de sistemas distribuidos [6].

Una jerarquía de red simple se consigue al organizar a los nodos participantes en la red sobrepuesta en dos capas, donde la capa superior está conformada por supernodos y la capa inferior por nodos regulares de la red. Cada supernodo se asigna a un conjunto de nodos regulares, los cuales forman una pequeña red entre ellos. Asimismo, el supernodo conecta a esta red de nodos regulares con el resto de la red al realizar búsquedas en su nombre. En esta red los supernodos actúan como dos nodos virtuales, uno por cada nivel [6].

De esta forma, una arquitectura de red sobrepuesta jerárquica ofrece algunas ventajas importantes en comparación a una red sin jerarquía [7]:

- Emplea de manera útil la heterogeneidad de los pares. Al designar como *supernodos* a los nodos con mejores características de procesamiento y almacenamiento, la red sobrepuesta de mayor nivel será más estable que su equivalente en red sin jerarquía. De este modo, un diseño de red jerárquica adopta la heterogeneidad mediante la

organización de los nodos de tal manera que nodos con distintas capacidades posean responsabilidades diferentes.

- Reduce el tiempo promedio de búsqueda. Debido a que el número de grupos es menor que el número total de pares en su equivalente de red sin jerarquía, los mensajes de búsqueda realizan menos saltos en la red sobrepuesta, por lo que el tiempo promedio de búsqueda disminuye. La reducción en el tiempo promedio es mayor cuando los pares pertenecientes al mismo grupo se encuentran cercanos en la topología [8]. Asimismo, un número menor de saltos por búsqueda significa que existe un menor número de mensajes de sobrecarga en la red.
- Transparencia. Los cambios en la responsabilidad de las llaves dentro de un mismo grupo debido a la entrada y salida de los nodos en la red (*churn*) es transparente a los demás grupos. De la misma manera, si existe un cambio en el algoritmo de búsqueda dentro de un grupo de un nivel inferior, el cambio es transparente a otros grupos a los algoritmos de niveles superiores.
- Aislamiento de fallas. La falla de un nodo regular dentro de un grupo G_i en un nivel inferior solamente tiene repercusiones en los nodos pertenecientes a G_i , es decir, las tablas de enrutamiento de los nodos pertenecientes a otros grupos no son afectadas.
- Facilita la interoperabilidad entre sistemas. Debido a que el enrutamiento dentro de un grupo es transparente a la jerarquía de mayor nivel, cada grupo de nodos puede elegir su propio protocolo de red sobrepuesta. De esta manera, al proporcionar autonomía administrativa dentro de cada grupo, se facilita la implementación de sistemas a gran escala y la interoperabilidad entre sistemas existentes.

4.2 Clasificación

Las redes sobrepuestas jerárquicas pueden clasificarse de acuerdo a su arquitectura de jerarquía. Una arquitectura de jerarquía define la manera en la cual se forman los grupos de

nodos en el sistema P2P y la manera en la cual se interconectan entre ellos [4]. En la literatura se pueden identificar dos tipos de arquitecturas de redes sobrepuestas jerárquicas: redes jerárquicas de dominio único y redes jerárquicas de dominios múltiples [5].

▪ Redes jerárquicas de dominio único

En esta clasificación existen las *redes sobrepuestas jerárquicas de un solo nivel* basadas en el modelo de jerarquía simple de dos niveles. En esta red, los supernodos forman una sola red sobrepuesta y los nodos ordinarios se conectan de manera directa a ellos como se muestra en la Figura 4.1. Una red jerárquica de este estilo se utiliza en el protocolo RELOAD [9], donde los nodos con poca capacidad de procesamiento, denominados clientes, se unen a un nodo participante de la red sobrepuesta. De esta manera, un cliente puede hacer uso de los servicios de la red sobrepuesta pero no participa en el mantenimiento ni el enrutamiento dentro de ésta.

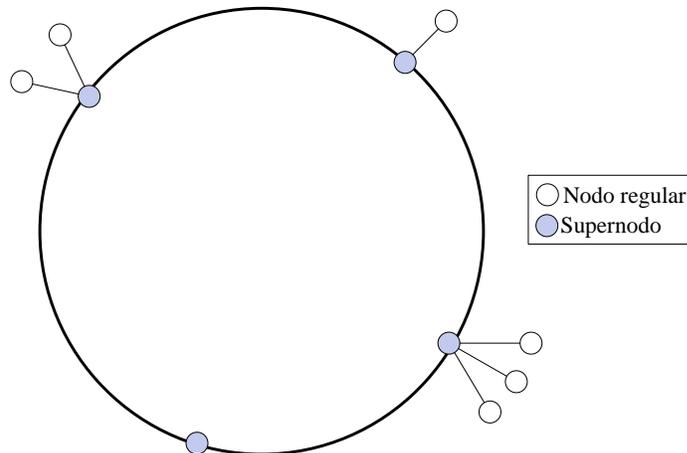


Figura 4.1. Arquitectura de red sobrepuesta jerárquica de dominio único.

▪ Redes jerárquicas de dominios múltiples

La propiedad característica de las redes jerárquicas de dominios múltiples es que se encuentran conformadas por distintos dominios interconectados por un dominio principal, es decir, cada dominio es una red sobrepuesta en sí misma. Debido a esta característica, cada

dominio puede tener su propio protocolo de red sobrepuesta como se muestra en la Figura 4.2. Cada dominio particular mantiene un supernodo dentro del dominio de interconexión, el cual se utiliza como puerta de enlace por los nodos regulares para mensajes entre dominios distintos. Una propuesta de red jerárquica de dominios múltiples para su uso en servicio de telefonía IP bajo el protocolo P2P-SIP [10] se presenta en [11].

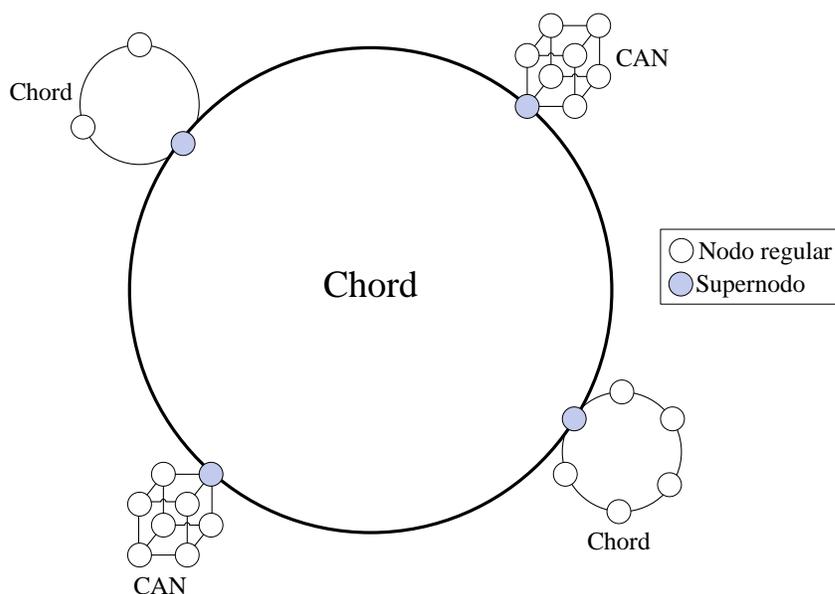


Figura 4.2. Arquitectura de red sobrepuesta jerárquica de dominios múltiples.

En esta arquitectura, un supernodo debe mantener una tabla de enrutamiento distinta por cada dominio al que pertenece, lo que puede generar una sobrecarga en su tabla de enrutamiento. Durante el enrutamiento, el supernodo decide la red sobrepuesta a utilizar en la búsqueda.

En [5] se presenta un análisis de costos para realizar la comparación en términos del número total de mensajes generados por las búsquedas entre los dos tipos de redes presentadas para un número de nodos determinado. La red jerárquica de dominios múltiples presenta un mejor desempeño en comparación de la red jerárquica de dominio único.

4.3 Construcción de una red sobrepuesta jerárquica (RSJ)

En esta sección se presenta la evaluación del desempeño de una construcción de red sobrepuesta jerárquica con arquitectura de red jerárquica de dominios múltiples, la cual se conforma de dos niveles y emplea al protocolo Chord en los diferentes dominios.

▪ Arquitectura de red.

En esta construcción, los N nodos pertenecientes en la red se dividen en K dominios. Cuando un nodo se une a la red sobrepuesta, éste decide unirse a un dominio existente de manera aleatoria, por lo que cada dominio tiene en promedio el mismo número de nodos, N/K . Asimismo, existe un dominio de interconexión, en el cual se introducen los K supernodos responsables de cada dominio, como se muestra en la Figura 4.3. La asignación del identificador de cada nodo dentro del dominio correspondiente se realiza de manera aleatoria.

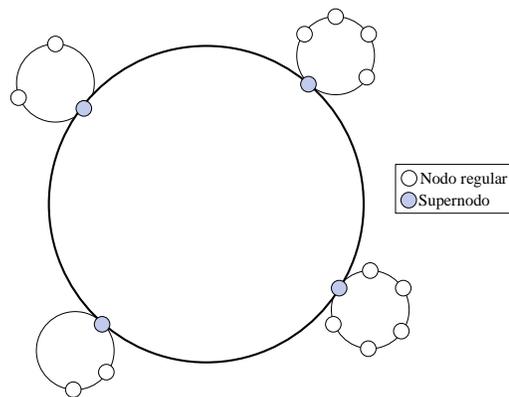


Figura 4.3. Arquitectura de red sobrepuesta jerárquica.

Para realizar el enrutamiento dentro de la red jerárquica se utiliza un identificador jerárquico como el propuesto en [11], el cual se encuentra compuesto de dos partes como se muestra en la Figura 4.4. De esta manera, los nodos regulares dentro de un mismo dominio utilizan el sufijo del identificador para el enrutamiento dentro del dominio. Por su parte, los

supernodos utilizan el prefijo del identificador para el enrutamiento dentro del dominio de interconexión.

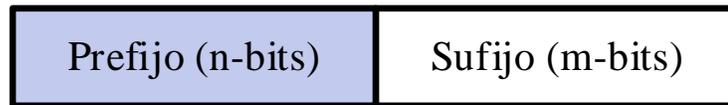


Figura 4.4. Identificador jerárquico de dos niveles.

En el caso en el cual un nodo regular requiera encontrar un nodo con prefijo de identificador diferente al utilizado dentro de su dominio origen, la búsqueda se envía al supernodo correspondiente para que reenvíe el mensaje de búsqueda a través del dominio de interconexión al supernodo responsable del dominio destino de la búsqueda. Posteriormente, el supernodo responsable del dominio destino de la búsqueda puede reenviar el mensaje dentro del dominio destino para encontrar al nodo destino.

- **Metodología de simulación.**

Con el fin de evaluar el comportamiento y desempeño de la construcción RSJ, se desarrollaron experimentos por medio de simulación de esta construcción y de su equivalente en arquitectura sin jerarquía, para $N = 2^k$ nodos. El simulador se desarrolló en lenguaje C, el cual cumple con las características del protocolo Chord como se mostró en la sección 2.1.4. Asimismo, la topología de red subyacente es generada con la herramienta GT-ITM [12], Para esta simulación se ocupó la Topología A presentada en la sección 2.3 compuesta por 4525 nodos. Durante la simulación, se varió el valor de k desde 8 hasta 12 y se realizaron experimentos separados para cada valor de k utilizando 8 supernodos, es decir, existen 8 dominios y un dominio de interconexión. En un experimento específico, cada nodo elige de manera aleatoria un conjunto de llaves del sistema para realizar su búsqueda.

▪ **Resultados de simulación.**

En la Figura 4.5 muestra la comparación del desempeño de la construcción RSJ y su equivalente en arquitectura sin jerarquía de acuerdo a las siguientes métricas: número promedio de saltos por búsqueda, tiempo promedio de búsqueda y factor de estiramiento.

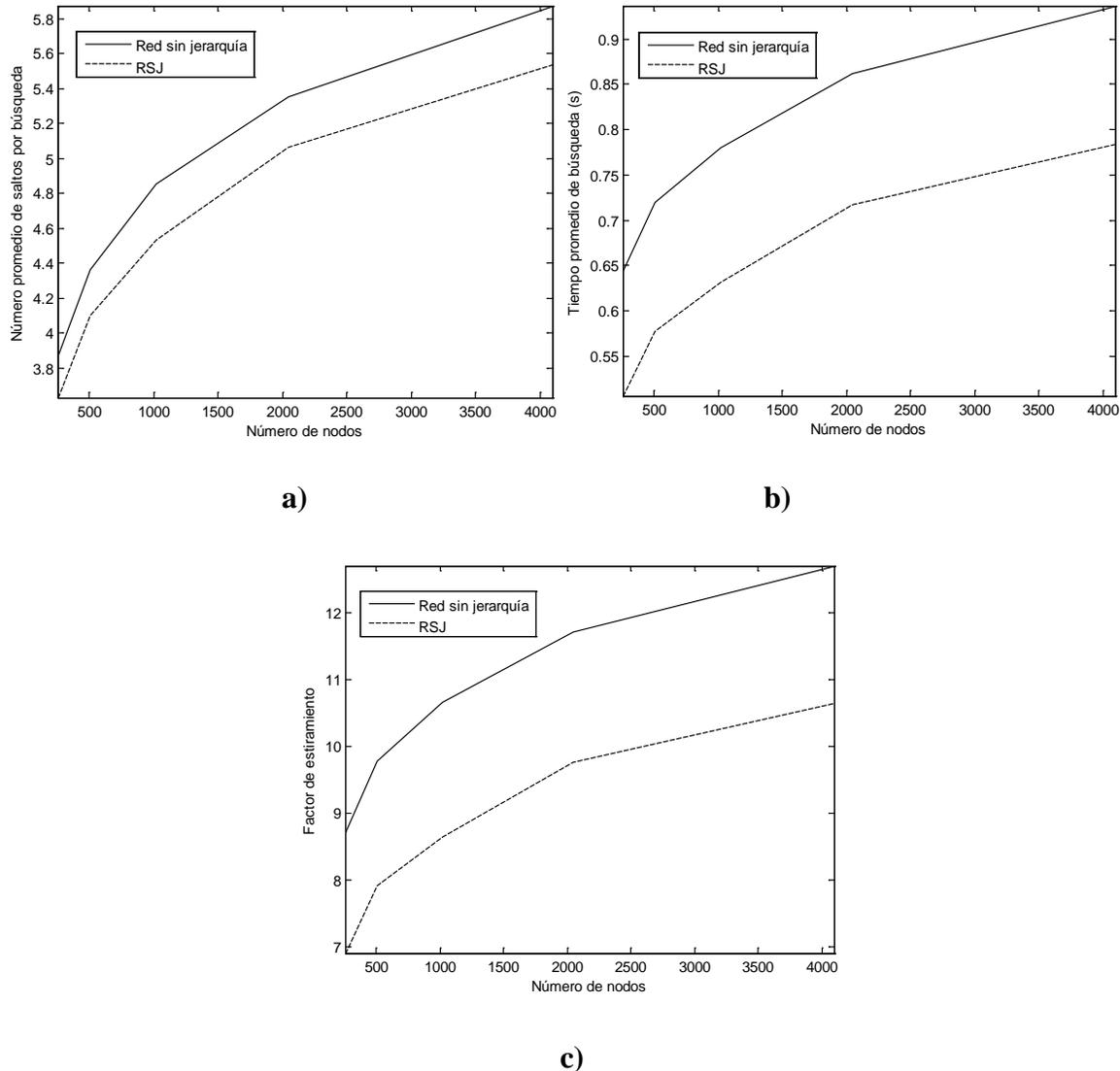


Figura 4.5. Comparación del desempeño de la construcción RSJ y su equivalente en arquitectura sin jerarquía en base a: **a)** Número promedio de saltos por búsqueda, **b)** Tiempo promedio de búsqueda y **c)** Factor de estiramiento.

Como se puede observar de la Figura 4.5a, el número promedio de saltos en RS-J disminuyó en comparación con su equivalente sin jerarquía. Eso se debe a que el número de grupos de nodos es más pequeño que el número total de nodos en la red, por lo cual la longitud de las trayectorias de los mensajes disminuye. Por otro lado, en la Figura 4.5b se observa que el tiempo promedio de búsqueda disminuyó aproximadamente un 20% debido a la implementación de esta construcción jerárquica. Esto puede confirmarse al observar la Figura 4.5c, en la cual se observa que el factor de estiramiento en la construcción RS-J también disminuyó en un porcentaje similar.

Al realizar la comparación de desempeño entre la construcción RS-J y su equivalente en red sin jerarquía, se puede observar que la construcción RS-J presenta un mejor desempeño en cuanto al tiempo promedio de búsqueda al disminuir el número de saltos en la red sobrepuesta. Sin embargo, la relación entre el número de dominios y la reducción en el número promedio de saltos en la red no se puede obtener de manera inmediata. A continuación, se presenta un análisis analítico para obtener el número promedio de saltos de un mensaje de búsqueda para la construcción presentada, El análisis se basa en el análisis propuesto en [11], con la adaptación para la construcción presentada, la cual utiliza el protocolo Chord en lugar del protocolo CAN.

4.3.1 Análisis del número promedio de saltos

Cuando un nodo se une a la red sobrepuesta, éste decide unirse a un dominio de manera aleatoria, por lo que el número promedio de nodos en el dominio i , M_i , es igual a N/K . De esta manera, el número promedio de saltos, H , se encuentra en función del número de saltos promedio en el dominio i , H_i , de la siguiente manera:

$$H = \frac{1}{N} \sum_{i=1}^K M_i H_i = \frac{1}{N} \sum_{i=1}^K \frac{N}{K} H_i = \frac{1}{K} \sum_{i=1}^K H_i \quad (4.1)$$

Donde H_i puede ser calculado con la Ecuación 4.2. Donde ρ_{ij} es la probabilidad de que una búsqueda tenga como origen el dominio i y como destino el dominio j . $C(K)$ es el costo de

encontrar un nodo dentro del dominio de interconexión en función del número de nodos en el dominio de interconexión K . $D_i(M_i)$ es el costo de encontrar un nodo dentro del dominio i en función del número de nodos en el dominio i , M_i .

$$H_i = \rho_{ii}D_i(M_i) + \sum_{j=1, j \neq i}^K \rho_{ij} [1 + D_j(M_j) + C(K)] \quad (4.2)$$

Como se puede observar de la Ecuación 4.2 el primer término es el costo de buscar un par dentro del dominio origen de la búsqueda, mientras que el segundo término es el costo de buscar un par que se encuentra en otro dominio. Cabe mencionar que en este análisis se supone que cada nodo regular conoce al supernodo correspondiente dentro de su dominio.

Sin embargo, como se mencionó anteriormente se ocupa el mismo protocolo de red sobrepuesta en cada uno de los dominios por lo que $D_j = D \forall j$. De la misma manera, el número de nodos en cada dominio es el mismo $M_i = M \forall i$. Asimismo, la probabilidad de buscar un par en otro dominio puede estar distribuida de manera equitativa entre todos los dominios externos, es decir, $\rho_{ij} = cte \forall j$. De esta manera se tiene que:

$$\sum_{j=1}^K \rho_{ij} = 1 \Rightarrow \rho_{ii} + \sum_{j=1, j \neq i}^K \rho_{ij} = 1 \Rightarrow \rho_{ii} + (K-1)\rho_{ij} = 1$$

$$\rho_{ij} = \frac{1 - \rho_{ii}}{(K-1)} \quad (4.3)$$

De esta manera utilizando la Ecuación 4.3 en la Ecuación 4.2 se obtiene lo siguiente:

$$H_i = \rho_{ii}D(M) + \frac{1 - \rho_{ii}}{(K-1)} \sum_{j=1, j \neq i}^K [1 + D(M) + C(K)] \quad (4.4)$$

Desarrollando la Ecuación 4.4 y utilizando el resultado obtenido en la Ecuación 4.1 y suponiendo que la probabilidad de buscar un nodo dentro del dominio origen es el mismo para todos los dominios, $\rho_{ii} = cte \forall i$, se obtiene la Ecuación 4.5, la cual denota el número promedio de saltos en la red jerárquica construida.

$$H = D(M) + (1 - \rho_{ii}) [1 + C(K)] \quad (4.5)$$

Debido a que el protocolo utilizado es el protocolo Chord, el costo de encontrar un nodo en el dominio origen o en el dominio de interconexión es el siguiente, como se demostró en el Capítulo 2:

$$C(x) = D(x) = \frac{1}{2} \log(x)$$

Por lo que el número promedio de saltos en la red jerárquica construida es:

$$H = \frac{1}{2} \log\left(\frac{N}{K}\right) + (1 - \rho_{ii}) \left[1 + \frac{1}{2} \log(K)\right] \quad (4.6)$$

A continuación, en la Figura 4.6 se muestra una comparación del resultado analítico obtenido en la Ecuación 4.6 y los resultados obtenidos por medio de simulación en función del número de dominios para un tamaño de red sobrepuesta de $N = 1200$ y probabilidad de una búsqueda dentro del dominio origen de $\rho_{ii} = 0.5$.

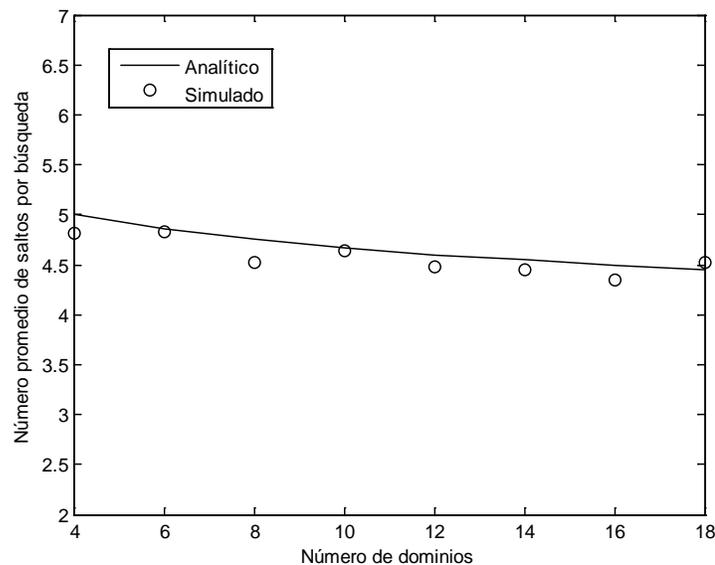


Figura 4.6. Número promedio de saltos en la de red sobrepuesta jerárquica en función del número de dominios.

En la Figura 4.6 se puede observar que el análisis presentado es una aproximación adecuada al compararse con resultados de simulación. Asimismo, se puede inferir que el número promedio de saltos en la red jerárquica puede disminuirse aumentando el número de dominios. Sin embargo, al aumentar el número de dominios, aumenta la complejidad del sistema y puede resultar en una generación excesiva de tráfico en la red.

Al realizar la construcción RS-J se disminuye el tiempo promedio de búsqueda al disminuir el número de saltos. Sin embargo, la elección aleatoria del dominio al que pertenece cada nodo, produce dominios cuyos nodos se encuentran distribuidos a lo largo de toda la red subyacente. Por esta razón, es posible mejorar el desempeño de esta construcción al introducir información de proximidad en la red, permitiendo que los nodos dentro de un mismo dominio se encuentren cercanos en la red subyacente. Esta alternativa de formación de la red sobrepuesta jerárquica se evalúa por medio de la construcción RSJ-CTS.

4.4 Construcción de una red sobrepuesta jerárquica consciente de la topología subyacente (RSJ-CTS)

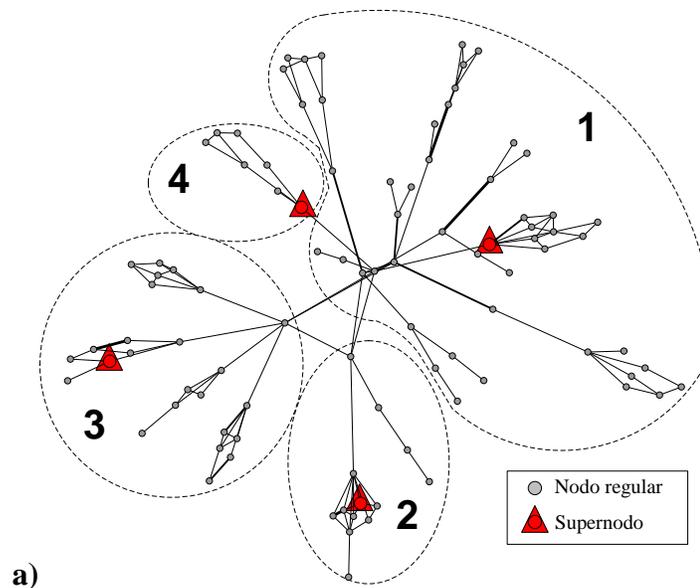
En esta sección se presenta la evaluación del desempeño de una construcción de red sobrepuesta jerárquica consciente de la topología subyacente con arquitectura de red jerárquica de dominios múltiples, la cual posee dos niveles y emplea al protocolo Chord en los diferentes dominios. En esta construcción se aplica la técnica de agrupamiento utilizando nodos marca para introducir la información de proximidad en la topología subyacente, donde M es el número de supernodos utilizado.

- **Arquitectura de red.**

En esta construcción, cuando un nodo ingresa a la red sobrepuesta, éste mide las latencias hacia cada uno de los nodos marca, los cuales también cumplen con la función de supernodos responsables de un dominio. Posteriormente, ordena de manera ascendente las latencias obtenidas y elige unirse al dominio correspondiente al supernodo más cercano. La

asignación del identificador dentro del dominio correspondiente se realiza de manera aleatoria.

La Figura 4.7a ejemplifica la manera en la cual una topología de red compuesta de 100 nodos se divide en M zonas de acuerdo a la posición de los nodos en la red subyacente, donde se utilizan $M = 4$ supernodos. De esta forma, los nodos que se encuentran dentro de la zona i ingresan al dominio i , el cual tiene como nodo responsable al supernodo marca más cercano. En la Figura 4.7b se ejemplifica la arquitectura de la red jerárquica construida cuando el número de supernodos es igual a $M = 4$, de esta manera, existen M dominios y un dominio de interconexión donde se encuentran los supernodos.



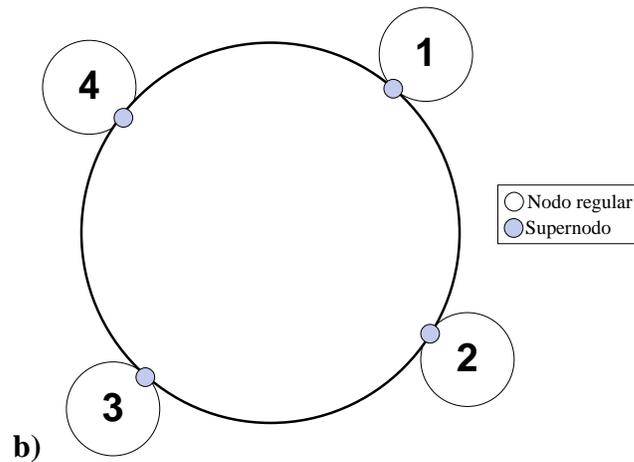


Figura 4.7. Arquitectura de red sobrepuesta jerárquica consciente de la topología subyacente. **a)** Ejemplo de la división de una topología de red en $M=4$ zonas. **b)** Ejemplo de la arquitectura de red jerárquica con $M=4$ dominios y un dominio de interconexión.

▪ Metodología de simulación.

Con el fin de evaluar el comportamiento y desempeño de la construcción RSJ-CTS se desarrollaron experimentos por medio de simulación de esta construcción, la construcción RSJ y de su equivalente en arquitectura sin jerarquía, para $N = 2^k$ nodos. Para esta simulación se ocupó la Topología A presentada en la sección 2.3 compuesta por 4525 nodos. Durante la simulación, se varió el valor de k desde 8 hasta 12 y se realizaron experimentos separados para cada valor de k . En un experimento específico, cada nodo elige de manera aleatoria un conjunto de llaves del sistema para realizar su búsqueda.

▪ Resultados de simulación.

En la Figura 4.8 se muestra la comparación del desempeño de la construcción RSJ-CTS, la construcción RSJ y su equivalente en red sin jerarquía de acuerdo a las siguientes métricas: número promedio de saltos por búsqueda, tiempo promedio de búsqueda y factor de estiramiento.

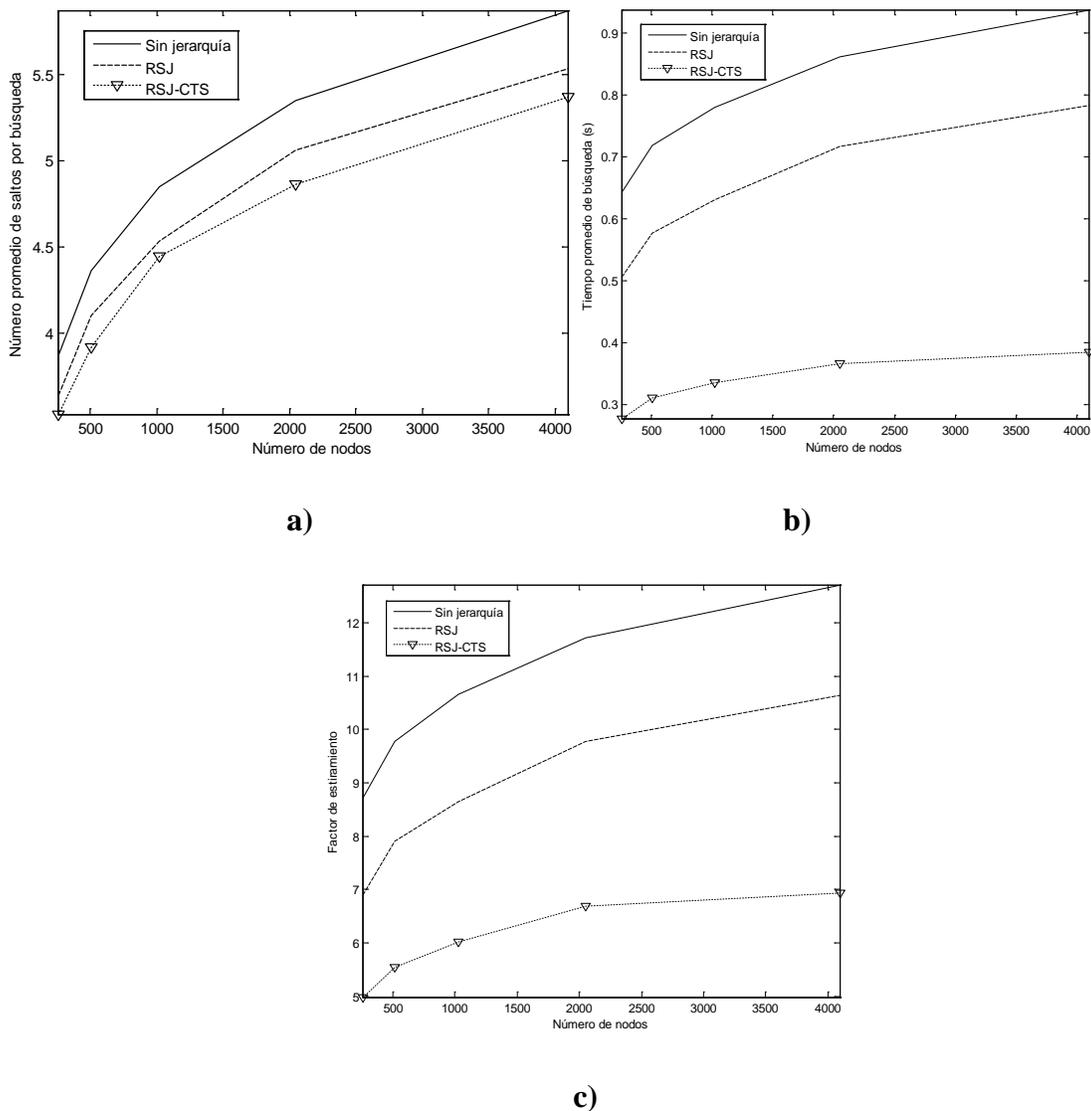


Figura 4.8. Comparación del desempeño de la construcción RSJ-CST, la construcción RSJ y su equivalente en arquitectura sin jerarquía en base a: **a)** Número promedio de saltos por búsqueda, **b)** Tiempo promedio de búsqueda y **c)** Factor de estiramiento.

Como se puede observar en la Figura 4.8a, el número de saltos de la construcción RSJ-CTS es menor a las otras construcciones dos debido a que al introducir información de la red subyacente, se altera la distribución de los nodos en los dominios de la arquitectura. Por otro lado, en la Figura 4.8b se observa que el tiempo promedio de búsqueda en la construcción RSJ-CTS disminuyó cerca del 50% en comparación de la construcción RSJ y cerca del 60%

en comparación a su equivalente sin jerarquía. Este efecto puede observarse también en la Figura 4.8c, donde el factor de estiramiento de la construcción RSJ-CTS disminuyó un porcentaje similar.

En base a los resultados obtenidos en la sección anterior, se puede observar que la construcción RSJ-CTS presenta un mejor desempeño respecto al tiempo promedio de búsqueda en comparación con la construcción RSJ y su equivalente de red sin jerarquía.

4.5 Elección de los supernodos

Los resultados generados en la sección anterior se obtuvieron al seleccionar de manera aleatoria a los supernodos de entre los nodos pertenecientes a la red. Sin embargo, como se mencionó en la Sección 3.5, el desempeño del agrupamiento utilizando nodos marca se encuentra relacionado con el posicionamiento de los nodos marca a lo largo de la red subyacente. Al utilizar a los supernodos como nodos marca, la elección de los supernodos impacta en el desempeño del sistema.

A continuación en la Figura 4.9 se muestra una comparación del impacto en el desempeño de una red sobrepuesta jerárquica consciente de la topología subyacente como la construcción RSJ-CTJ que presentan las elecciones de los supernodos presentadas en la Sección 3.5: Elección de supernodos de manera aleatoria (EA), Elección de supernodos separados por un retardo determinado (ER), Elección de supernodos en base al grado del nodo (EG) y el resultado del Algoritmo de optimización de elección de supernodos (EG-OPT). La comparación se realiza de acuerdo a las siguientes métricas: número promedio de saltos por búsqueda, tiempo promedio de búsqueda y factor de estiramiento.

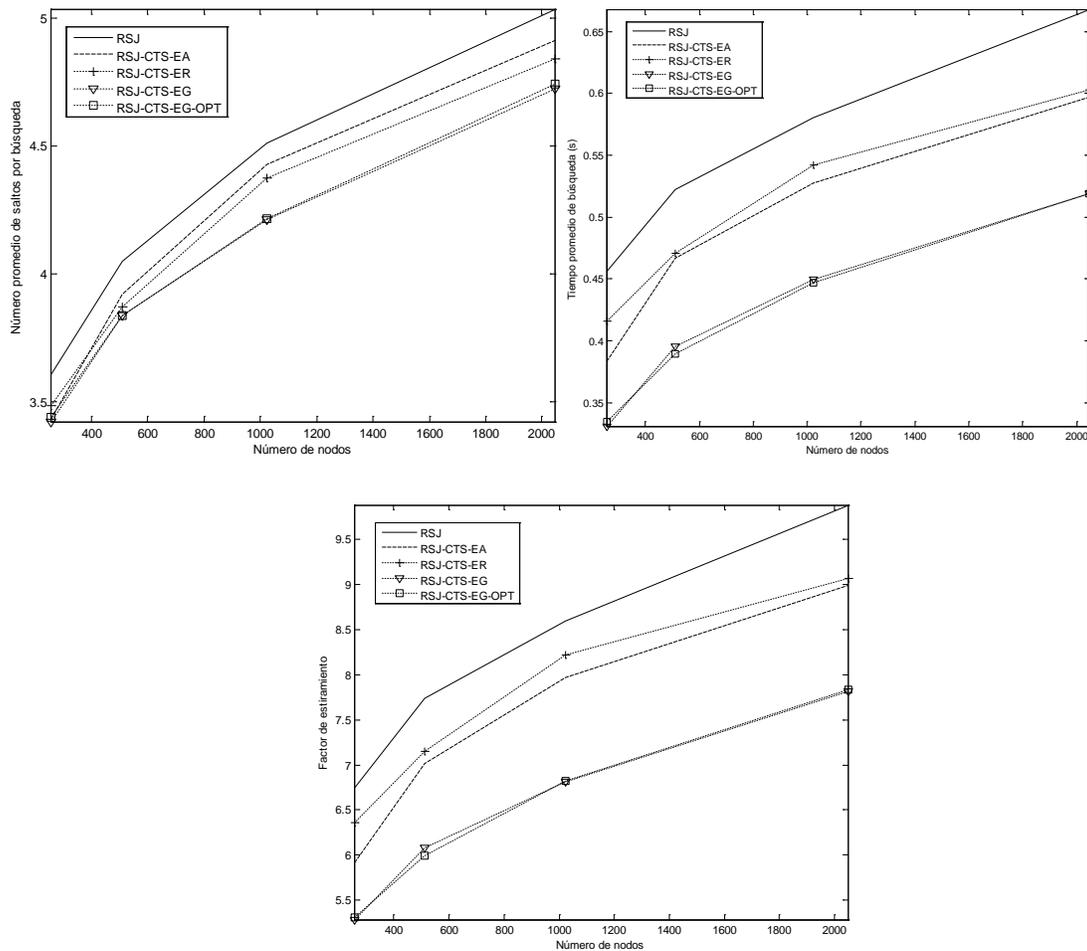


Figura 4.9. Comparación del desempeño de la RSJ-CTS con tres diferentes conjuntos de marcas en base a: **a)** Número promedio de saltos por búsqueda, **b)** Tiempo promedio de búsqueda y **c)** Factor de estiramiento.

En la Figura 4.9 se puede observar que el desempeño mostrado por la construcción que utiliza la elección de supernodos en base al grado del nodo es superior al presentado por el resto de los enfoques. Esto se debe a que existen concentraciones de nodos en la red subyacente, por lo cual al elegir como supernodos a los nodos que tienen un mayor número de conexiones se agrupa de mejor manera a los nodos en la red.

Referencias

- [1] I. Stoica, R. Morris, D. Karger, M. F. Kaashoek y H. Balakrishnan, «Chord: a scalable peer-to-peer lookup protocol for Internet applications,» *Networking, IEEE/ACM Transactions on*, vol. 11, nº 1, pp. 17-32, 2003.
- [2] S. Ratnasamy, P. Francis, M. Handley, R. Karp y S. Shenker, «A Scalable Content-addressable Network,» *SIGCOMM Comput. Commun. Rev.*, vol. 31, nº 4, pp. 161-172, 2001.
- [3] A. Rowstron y P. Druschel, «Pastry: Scalable, Decentralized Object Location, and Routing for Large-Scale Peer-to-Peer Systems,» *Proceedings of the IFIP/ACM International Conference on Distributed Systems Platforms Heidelberg*, pp. 329-350, 2001.
- [4] D. Korzun y A. Gurtov, *Structured Peer-to-peer Systems: Fundamentals of Hierarchical Organization, Routing, Scaling, and Security*, New York: Springer, 2013.
- [5] A. Diatta, I. Niang y M. Ba, «An Overview and Comparison of Hierarchical P2P-SIP Networks,» *International Journal of Computer Science and Information Security*, vol. 13, nº 3, pp. 47-58, 2015.
- [6] D. Korzun y A. Gurtov, «Hierarchical architectures in structured peer-to-peer overlay networks,» *Peer-to-Peer Networking and Applications*, vol. 7, nº 4, pp. 359-395, 2014.
- [7] L. Garces-Erice, E. W. Biersack, K. W. Ross, P. A. Felber y G. Urvoy-Keller, «Hierarchical peer-to-peer systems,» *Parallel Processing Letters*, vol. 13, nº 4, pp. 643-657, 2003.
- [8] W. Lv, Q. Liao, J. Zhao y Y. Xiao, «B_Chord: An Improved Routing Algorithm to Chord Based on Topology-Aware and Bi-Dimensional Lookup Method,» *Wireless*

Communications, Networking and Mobile Computing, 2009. WiCom '09. 5th International Conference on, pp. 1-4, 2009.

- [9] C. Jennings, B. Lowekamp, E. Rescorla, S. Baset y H. Schulzrinne, «REsource LOcation And Discovery (RELOAD) Base Protocol,» *RFC 6940*, 2014.
- [10] Internet Engineering Task Force, «Peer-to-Peer Session Initiation Protocol (p2psip),» [En línea]. Available: <http://datatracker.ietf.org/wg/p2psip/charter/>. [Último acceso: Agosto 2015].
- [11] I. Martinez-Yelmo, C. Guerrero, R. Cuevas y A. Mauthe, «A hierarchical P2PSIP architecture to support skype-like services,» *Parallel, Distributed and Network-based Processing, 2009 17th Euromicro International Conference on. IEEE*, pp. 316-322, 2009.
- [12] «Georgia Tech Internetwork Topology Models. GT-ITM,» [En línea]. Available: <http://www.cc.gatech.edu/projects/gtitm/>. [Último acceso: Septiembre 2015].

Conclusiones y Trabajo Futuro

Los sistemas de redes sobrepuestas P2P proporcionan una alternativa de solución para aplicaciones que necesitan mantener la conexión entre un gran número de usuarios de manera simultánea. Estos sistemas utilizan una arquitectura descentralizada, es decir, el sistema se encuentra conformado por dispositivos denominados *pares*, los cuales pueden comunicarse de manera directa entre ellos. Para su funcionamiento, el sistema establece una red lógica, denominada *red sobrepuesta*, la cual se construye por encima de una red existente, denominada *red subyacente*. De esta manera, los enlaces construidos en la red sobrepuesta son enlaces estrictamente virtuales, los cuales se encuentran compuestos de varios enlaces físicos en la red subyacente.

Debido al número de usuarios que puede soportar este tipo de sistemas, un diseño ineficiente de la red sobrepuesta, puede provocar que el tráfico generado por los usuarios del sistema circule sobre un número redundante de saltos físicos en la red subyacente ocasionando un volumen grande de tráfico adicional en la red. Por lo tanto, se requiere que el sistema cuente con un diseño de red sobrepuesta P2P con un buen desempeño en la entrega de mensajes a través de la red y de esta manera evitar afectar la operación de otras aplicaciones que se ejecuten en la infraestructura de la red subyacente.

En el presente trabajo de tesis se presenta un estudio de diversas técnicas presentes en la literatura para mejorar el desempeño de redes sobrepuestas para su uso en sistemas P2P. Se identifican dos técnicas por medio de las cuales es posible mejorar el desempeño de redes sobrepuestas de manera independiente: la construcción de redes sobrepuestas conscientes de la topología subyacente y la construcción de redes sobrepuestas jerárquicas. Cada una de estas técnicas consigue disminuir el tiempo promedio de búsqueda dentro de la red sobrepuesta al modificar un aspecto particular de su diseño.

En el Capítulo 1 se presentaron los elementos básicos de los sistemas de redes sobrepuestas P2P, haciendo énfasis en las redes sobrepuestas estructuradas basadas en una tabla hash distribuida (DHT). Asimismo, se realizó una comparación entre tres protocolos distintos de redes sobrepuestas: Chord, CAN y Pastry. En base a la comparación realizada se concluyó que el protocolo Chord presenta las mejores características de simplicidad, escalabilidad y balance de carga en los nodos, lo cual lo hacen apto para su uso en sistemas de gran escala como los estudiados en los capítulos posteriores.

En el Capítulo 2 se estudió de manera detallada el funcionamiento del protocolo Chord, este estudio sentó las bases para su posterior implementación. Asimismo, se presentó un análisis analítico del número de saltos promedio de una búsqueda en la red, el cual se validó con los resultados obtenidos por medio de simulación. Posteriormente, se introdujeron las métricas de evaluación utilizadas a lo largo de este trabajo: el número promedio de saltos por búsqueda, tiempo promedio de búsqueda y factor de estiramiento. En base a las métricas presentadas, se realizó la evaluación del desempeño del protocolo Chord al ejecutarse sobre una red subyacente, la cual se toma como punto de comparación para las construcciones de redes sobrepuestas implementadas posteriormente. Para la generación de la latencia en la red, se estudiaron diversas metodologías que realizan la simulación de una red subyacente, seleccionando un generador de topología de red que implementa el modelo *Transit-Stub* debido a que representa de mejor manera la topología de Internet a gran escala.

En el Capítulo 3 se presenta un estudio de diversos aspectos de las redes sobrepuestas conscientes de la topología subyacente (RS-CTS). Para ello, se presentaron las diversas técnicas para generar información de proximidad en la red, seleccionando al agrupamiento de nodos utilizando nodos marca como método para la construcción de RS-CTS debido a sus características de escalabilidad. Asimismo, se presentaron los tres enfoques comúnmente mencionados en la literatura para introducir la información de proximidad en la red y se concluyó que el enfoque de disposición geográfica es el más adecuado para utilizar con el protocolo Chord debido a que no aumenta el número de entradas en la tabla de enrutamiento y puede ser utilizado sin realizar alguna modificación considerable al protocolo.

Posteriormente, se presentó la evaluación del desempeño de dos construcciones de redes sobrepuestas conscientes de la topología subyacente: la construcción RS-CTS-S y la construcción RS-CTS-M. Estas construcciones son adaptaciones del protocolo Chord propuestas en el presente trabajo, las cuales introducen la información de la proximidad en la red subyacente por medio del agrupamiento de nodos utilizando nodos marca.

La construcción RS-CTS-S presenta una disminución en el tiempo promedio de búsqueda de aproximadamente 30% con respecto al protocolo Chord tradicional. Sin embargo, es posible mejorar estos resultados al aumentar el grado de consciencia de la red subyacente, lo cual se implementa por medio de la construcción RS-CTS-M.

La construcción RS-CTS-M presenta una disminución de aproximadamente la mitad del tiempo promedio de búsqueda con respecto al protocolo Chord tradicional. Esto se debe a que la red se divide en zonas pequeñas permitiendo que los nodos cercanos en la red subyacente se encuentren cercanos en el espacio de identificadores. Sin embargo, esta construcción presenta un mayor desbalance de carga en los nodos.

Asimismo, se realizó la evaluación del impacto en la elección de los nodos marca en el desempeño de la red. Al realizar la comparación del desempeño de distintas elecciones de nodos marca, la elección propuesta en el presente trabajo, es decir la elección de nodos marca en base al grado del nodo, presentó mejor desempeño de acuerdo a las métricas utilizadas. Esto se debe a que los nodos no se encuentran distribuidos de manera uniforme en la red subyacente y en su lugar existen concentraciones de nodos, lo cual no se consideraba en el resto de los enfoques analizados.

Para obtener un límite inferior en el desempeño del sistema se propuso un algoritmo de optimización de la elección de los nodos marca basado en el algoritmo *k-means*. Al realizar la comparación del impacto de los distintos tipos de elecciones de nodos marca, se observó que los resultados obtenidos mediante la elección de nodos marca en base al grado del nodo son muy similares a los obtenidos por medio del algoritmo de optimización propuesto, a pesar de que este último requiere del conocimiento global de las latencias entre todos los

nodos de la red. Por lo que la elección de marcas propuesta presenta buenas características de desempeño debido a que se acerca al límite impuesto por el algoritmo de optimización.

Por otro lado, debido a que las construcciones evaluadas en este capítulo presentan diseños de redes sobrepuestas sin jerarquía, poseen ciertas limitaciones. Por ejemplo, al modificar el número de marcas en el sistema debe existir una reasignación total de los identificadores en la red sobrepuesta, lo cual generaría un excesivo volumen de tráfico adicional en una implementación real. Este tipo de limitaciones pueden superarse mediante el uso de una red sobrepuesta jerárquica.

En el Capítulo 4 se presentaron los fundamentos de las redes sobrepuestas jerárquicas (RSJ) con las cuales se pueden resolver las limitaciones mencionadas anteriormente. Se presentó una clasificación de las RSJ de acuerdo a su arquitectura de red y se eligió la arquitectura de red jerárquica de dominios múltiples debido a que presenta un mejor desempeño en cuanto al número total de mensajes producidos.

Adicionalmente, se presenta la evaluación del desempeño de la construcción de redes sobrepuestas jerárquicas. Una construcción de jerarquía de dos niveles con arquitectura de red jerárquica de dominios múltiples que utiliza el protocolo Chord en ambos niveles, como la construcción RSJ, presenta una disminución del tiempo promedio de búsqueda en comparación a su equivalente sin jerarquía, ya que disminuye el tamaño de la trayectoria de un mensaje de búsqueda. Asimismo, se realizó el análisis del número promedio de saltos de un mensaje de búsqueda en la construcción RSJ, el cual se corroboró con los resultados obtenidos por medio de simulación. Sin embargo, la inclusión de consciencia de la topología subyacente puede mejorar el desempeño como se mostró con la construcción RSJ-CTS.

La construcción RSJ-CTS introduce la información de la proximidad en la red subyacente por medio del agrupamiento de nodos utilizando nodos marca. Esta construcción presenta una disminución en el tiempo promedio de búsqueda de cerca del 50% en comparación de la construcción RSJ y cerca del 60% en comparación a su equivalente sin jerarquía. Esto se

debe a que los nodos pertenecientes a un dominio en particular son nodos que se encuentran cercanos en la red subyacente.

Posteriormente, se realizó la evaluación del impacto en la elección de los supernodos en el desempeño de la red. Al realizar la comparación del desempeño de distintas elecciones de supernodos, la elección propuesta en el presente trabajo en la Sección 3.5, es decir la elección de supernodos en base al grado del nodo, presenta un mejor desempeño de acuerdo a las métricas utilizadas. Esto se debe a que los nodos no se encuentran distribuidos de manera uniforme en la red subyacente y en su lugar existen concentraciones de nodos.

Finalmente, al utilizar una arquitectura con jerarquía como la construcción RSJ-CTS, se obtienen beneficios en comparación de una arquitectura sin jerarquía como la construcción RS-CTS-M presentada en la Sección 3.4.2. Por ejemplo, al añadirse un nuevo nodo marca al sistema, no existe una reasignación total de identificadores como en la construcción sin jerarquía. Esto se debe a que al utilizar una estructura jerárquica, solamente los nodos que pertenecen al dominio del nodo marca adicional deben cambiar su posición en la red sobrepuesta. De la misma manera, la falla de un nodo dentro de un dominio solamente repercute en los nodos que se encuentran en el mismo dominio, es decir, las tablas de enrutamiento de los nodos pertenecientes a otros dominios no sufren cambios.

En base a los resultados obtenidos en el trabajo de tesis se puede identificar el siguiente trabajo a futuro a realizar:

- Analizar el uso de técnicas alternativas de agrupación de nodos en conjunto con el criterio de proximidad en la red subyacente, como por ejemplo la agrupación en base a la semántica de la búsqueda,
- Realizar un análisis de estabilidad de las construcciones presentadas en presencia de la entrada y salida de nodos en la red sobrepuesta.
- Evaluar el nivel de seguridad de las construcciones presentadas frente a posibles ataques en la red.