TESIS DEFENDIDA POR

Juan Manuel Oyoqui Sánchez

Y aprobada por el siguiente comité:

Dr. José Antonio García Macías
Director del Comité

Dr. Roberto Conte Galván
Miembro del Comité

Dr. Jesús Favela Vara
Miembro del Comité

Dr. Jesús Favela Vara
Jefe del Departamento de Ciencias de la Computación

Dr. Luis Alberto Delgado Argote
Director de Estudios de Posgrado

20 de Agosto del 2003

CENTRO DE INVESTIGACIÓN CIENTÍFICA Y DE EDUCACIÓN
SUPERIOR DE ENSENADA

DIVISIÓN DE FÍSICA APLICADA
DEPARTAMENTO DE CIENCIAS DE LA COMPUTACIÓN

Transferencia del Contexto en Redes Locales Inalámbricas

TESIS

que para cubrir parcialmente los requisitos necesarios para obtener el grado de

MAESTRO EN CIENCIAS

Presenta:

Juan Manuel Oyoqui Sánchez

Ensenada, Baja California a Agosto del 2003.

RESUMEN de la tesis de **Juan Manuel Oyoqui Sánchez**, presentada como requisito parcial para obtener el grado de MAESTRO EN CIENCIAS en CIENCIAS DE LA COMPUTACIÓN. Ensenada, Baja California, México. Agosto del 2003.

Transferencia del Contexto en Redes Locales Inalámbricas

Resumen aprobado por:

Dr. José Antonio García Macías

Director de Tesis

Hoy en día las redes inalámbricas y la tecnología de cómputo móvil están teniendo un profundo impacto en redes basadas en el protocolo IP. Sin embargo, debido a que los protocolos del conjunto IP fueron diseñados sin asumir la movilidad de los nodos, fue necesaria la creación de nuevos protocolos de movilidad que dieran solución a este problema. Entre dichos protocolos sin duda el más sobresaliente es *Mobile IP*. No obstante, este protocolo sufre de varios problemas cuando la movilidad es muy dinámica y se lleva a cabo al interior de un mismo dominio, por lo cual, surgió la necesidad de crear protocolos de micromovilidad cuya función principal es manejar de manera eficiente la movilidad al interior de dominios.

Sin embargo, los protocolos de micromovilidad no resuelven varios problemas importantes requeridos para lograr una movilidad transparente. Lo anterior se debe a que estos protocolos asumen que el nodo móvil ó el actual enrutador de acceso ya han seleccionado el nuevo enrutador con el cual se realizará un *handoff*. De la misma manera, cuando se desarrolla un *handoff* sería conveniente (si no es que obligatorio) que el MN (Mobile Node) encontrara en el nuevo enrutador el mismo contexto que tenía en su anterior enrutador de acceso.

Es por ello que en este trabajo se propone una arquitectura para manejar la movilidad transparente. Esta arquitectura incluye el proceso para el descubrimiento de los candidatos a enrutadores de acceso (CAR, Candidate Access Router) y la selección del mejor de ellos, así como de un protocolo para realizar la transferencia de contexto entre los enrutadores de acceso. Durante este trabajo, también se presentan consideraciones de diseño para evitar introducir latencia al proceso de la movilidad transparente, así como para evitar el rompimiento de las conexiones en las aplicaciones de los nodos móviles durante un *handoff* debido a la pérdida de paquetes.

Palabras clave: transferencia de contexto, descubrimiento de CARs, micromovilidad.

ABSTRACT of the thesis presented by **Juan Manuel Oyoqui Sánchez**, as a partial requirement to obtain the MASTER IN SCIENCES degree in COMPUTER SCIENCES. Ensenada, Baja California, México. August 2003.

Transfer Context in Wireless Local Networks

Abstract approved by:

Dr. José Antonio García Macías

Thesis director

Nowadays wireless networks and mobile computing technologies are having a profound impact on IP-based networks. However, due to the fact that the protocols of the IP suite were designed without assuming mobility of the network nodes, it was necessary to create new mobility protocols that address this problem. Among these protocols, undoubtedly the most prominent is Mobile IP. Nevertheless, this protocol has several drawbacks when the mobility is very dynamic and it is performed inside a single domain. Therefore it became necessary to create micro-mobility protocols whose main function is to handle the intra-domain mobility efficiently.

However, micro-mobility protocols do not address important issues required for seamless mobility. This is due to the fact that these protocols assume that the mobile node, or the current access router, already have selected the new router which will be used to carry out a handoff. In the same way, when performing a handoff it would be beneficial (if not mandatory) for the MN (Mobile Node) to find at the new access router the same context it had at its previous access router.

For this reason, this work proposes an architecture to handle seamless mobility. This architecture includes the process of candidate access router (CAR) discovery and the selection of the best of them, as well as a protocol to carry out the context transfer among the access routers. Also, design considerations are presented that aim to avoid introducing latency to the seamless mobility process, as well as to avoid breaking the connections in the applications of the mobile host during a handoff due to the loss of packages.

Keywords: transfer context, CARs discovery, micromobility.

Dedicatorias

A mis padres: Jorge y Margarita,
por brindarme su apoyo y amor incondicional.

A mis hermanas: Ana Lilia y Margarita,
a quienes quiero con toda el alma.

A mi novia Lupita,
por su amor y apoyo en este difícil camino.

Agradecimientos

A Dios,
por darme vida y fuerza para alcanzar esta meta.

Un gran agradecimiento a mi asesor el Dr. José Antonio García Macías,
por su paciencia y valiosa ayuda.

Al Dr. Jesús Favela Vara y al Dr. Roberto Conte Galván,
por sus aportaciones en favor de la mejora de este trabajo.

A mis amigos de la generación 2001-2003:
Lupita, Mike, Daniel, Oscar, Yobani, Domitilo, Chema, Alex Estrella, Alex Peña,
Rodrigo, Mirna, Pedro, Miguel Riesgo y Everardo.
Gracias por su amistad y por hacer más fáciles las desveladas.

A los estudiantes de doctorado:
Gabriel, Juan Contreras, Doris, Marcela y Juan Tapia,
por brindarme siempre su valiosa ayuda.

A todos mis maestros,
por enseñarme el camino para llegar hasta aquí.

Al Centro de Investigación Científica y de Educación Superior de Ensenada (CICESE),
por enseñarme el verdadero significado de estudiar.

Al Consejo Nacional de Ciencia y Tecnología (Conacyt),
por su apoyo económico sin el cual no hubiera sido posible este trabajo de
investigación.

Ensenada, México
20 de Agosto de 2003.

Juan Manuel Oyoqui Sánchez

Tabla de Contenido

Sección Página

Tabla de Contenido (Cont.)

Tabla de Contenido (Cont.)

Lista de Figuras

Lista de Figuras (Cont.)

Lista de Tablas

Capítulo I

Introducción

I.1 Motivación

En la actualidad las tecnologías inalámbricas y móviles están impactando fuertemente el entorno de las redes basadas sobre los protocolos de Internet (IP). Los dispositivos de comunicación móvil tales como computadoras portátiles, organizadores personales (PDAs), teléfonos celulares de 3ra generación, etc. tienen capacidad de utilizar los protocolos de Internet, pero usualmente estarán utilizando tecnologías inalámbricas de diferentes características. Así pues, nos enfrentamos a escenarios de ambientes heterogéneos donde diferentes tecnologías inalámbricas (estándar 802.11, Bluetooth, 3G, etc.) son utilizadas para comunicar dispositivos utilizando IP. Este tipo de redes tiene la propiedad de proveer movilidad a los nodos, o sea, que el nodo móvil puede seguir conectado a la red mientras se desplaza. Sin embargo, esta movilidad no puede ser resuelta por IP [Bhagwat *et al.*, 1996], dado que la dirección IP de un nodo móvil es utilizada para determinar la localización del mismo, y si este se mueve dentro de la red de acceso es necesario cambiar de dirección IP para ser propiamente localizado; esto conllevaría a romper las conexiones establecidas por las aplicaciones que se ejecutan en el dispositivo móvil. Para resolver este problema fue necesario la creación de nuevos protocolos para el manejo de la movilidad, de entre los cuales sobresale Mobile IP [Perkins, 1996a]. Sin embargo, este protocolo tiene grandes desventajas cuando la movilidad es llevada a cabo dentro de un mismo dominio [Fladenmuller y Silva, 1999], y dado que es ahí donde se llevará gran parte de la movilidad [Campbell *et al.*, 2002], es necesario contar con protocolos de micromovilidad. Así pues, usualmente un protocolo

de micromovilidad se encargará de manejar la movilidad dentro de un dominio mientras que Mobile IP se encargará de manejar la movilidad entre dominios. No obstante, los protocolos de micromovilidad aún tienen problemas por resolver para lograr una movilidad transparente, por lo tanto, la resolución de estos problemas es la principal motivación de esta tesis.

I.2 Planteamiento del problema

En redes IP que soportan la movilidad de los nodos, los caminos de enrutamiento entre el nodo y la red pueden cambiar de forma rápida y frecuente. Este problema puede ser resuelto eficientemente por algún protocolo de micromovilidad, sin embargo estos protocolos aun tienen varios problemas a resolver en beneficio de lograr una movilidad transparente. Entre los problemas que los protocolos de micromovilidad aún no resuelven se encuentran: el descubrimiento de los candidatos a enrutadores de acceso y la selección del mejor de entre ellos, así como la transferencia del estado - contexto - del nodo móvil hacia su nuevo enrutador de acceso, de tal forma que el nodo móvil pueda seguir contando con los mismos servicios que le proveía su antiguo enrutador.

Existen algunos casos en que el nodo móvil puede establecer ciertos servicios candidatos para la transferencia del contexto en subredes que son dejadas atrás mientras el nodo se mueve realizando transferencias de conexión (*handoffs*). Algunos de estos servicios son AAA (Authentication, Authorization, and Accounting), compresión de cabeceras, calidad de servicio (QoS), etc. Por lo tanto, si el nodo móvil quisiera restablecer estos servicios en el nuevo enrutador a través del mismo proceso que utilizó inicialmente, el retardo del tráfico en tiempo real puede ser seriamente impactado. No obstante, una alternativa para resolver este problema es transferir suficiente información del estado actual de los servicios candidatos para la transferencia de contexto a la nueva subred, de tal manera que los servicios puedan ser restablecidos rápidamente, en lugar

de comenzar las señalizaciones desde el principio [Levkowetz *et al.*, 2002]. Sin embargo, llevar a cabo la transferencia de contexto involucra la resolución de dos problemas. El primer problema a resolver es el descubrimiento de los candidatos a enrutadores de acceso y en base a las preferencias del nodo móvil, seleccionar un enrutador que pueda satisfacer estas necesidades [Krishnamurthi, 2002]. Una vez seleccionado un enrutador de acceso, el segundo problema consiste en transferir la información del contexto del nodo móvil desde su actual enrutador hacia el nuevo enrutador de acceso [Syed *et al.*, 2003]. Durante el desarrollo de este trabajo se presentará una solución a los problemas previamente descritos, para lo cual se propondrá una arquitectura y protocolos para dar solución al problema de la movilidad transparente.

I.3 Objetivos

El objetivo general de este trabajo de investigación es diseñar una arquitectura que permita la movilidad transparente, auxiliándose para ello del descubrimiento de los candidatos a enrutadores de acceso y la selección del mejor entre ellos así como de la transferencia de contexto. Para lograr este objetivo es necesario dividir esta tarea en objetivos específicos, los cuales son:

- Identificar escenarios de utilización de dispositivos móviles realizando handoff en redes locales inalámbricas.

- Proponer una arquitectura para realizar handoffs guardando el contexto de los dispositivos durante su desplazamiento. Esta arquitectura debe definir cuales son los pasos a seguir para llevar a cabo la transferencia de contexto, así como también especificar el diseño de los protocolos encargados de llevar a cabo esta tarea

- Implementar un prototipo funcional que maneje el descubrimiento de los candidatos a enrutadores de acceso y la selección del mejor entre ellos. El tiempo

de ejecución de este algoritmo no debe afectar el desempeño del proceso de la movilidad transparente.

I.4 Organización del documento

Este documento está compuesto por seis capítulos y dos apéndices. El primer capítulo ofrece una visión general de los problemas a atacar durante el desarrollo de la tesis. Los capítulos II, III y IV muestran el estado del arte del problema de la movilidad transparente; las soluciones propuestas se describen en el capítulo V y por último, las conclusiones se encuentran en el capítulo VI. Los apéndices incluyen información que ayudará a la mejor comprensión de los conceptos tratados en este trabajo. En la figura 1 se muestra una representación gráfica de la organización del documento. A continuación se presenta una descripción mas detallada de los temas que se abordan en cada capítulo.

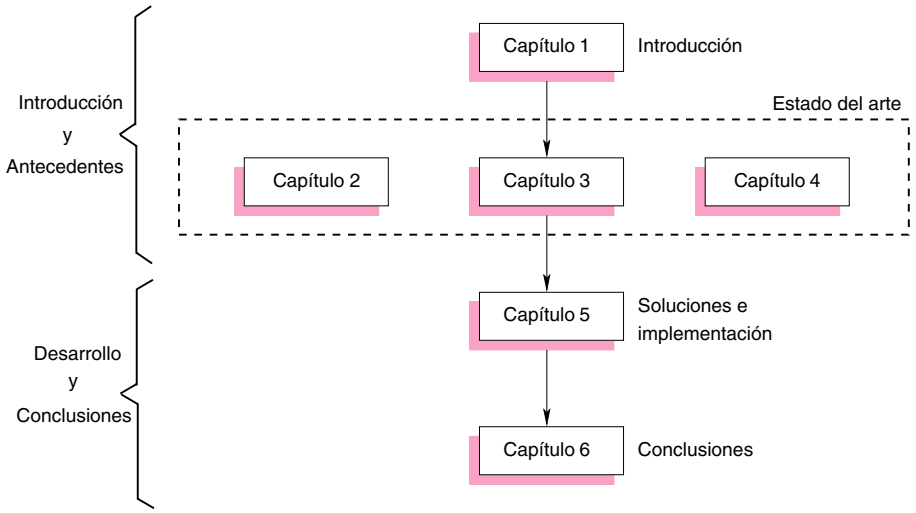

Figura 1. Organización de la tesis.

En el capítulo II se tratan los temas relacionados a las redes locales inalámbricas, por ejemplo, las tecnologías utilizadas para el desarrollo de las mismas, los estándares actuales y las características de cada uno de ellos.

Los problemas relacionados a la movilidad sobre IP son tratados en el capítulo III. Por lo tanto, en este capítulo se explica cual es el problema de la movilidad y cuales son los protocolos que han surgido para dar solución a este problema. También se explican los dos enfoques de la movilidad: macromovilidad y micromovilidad.

En el capítulo IV se describen los problemas relacionados a la movilidad transparente, entre los cuales encontramos el descubrimiento de los candidatos a enrutadores de acceso y la selección del mejor de ellos, así como el problema de la transferencia del contexto de los dispositivos móviles de su actual enrutador hacia su nuevo enrutador de acceso.

Las soluciones propuestas se describen en el capítulo V. Este capítulo contiene la descripción de la arquitectura propuesta, por lo tanto, se presenta una descripción y evaluación del proceso del descubrimiento de los candidatos a enrutadores de acceso y la selección del mejor entre ellos, así como el diseño del protocolo para la transferencia del contexto. Este capítulo también presenta consideraciones de diseño para evitar el rompimiento de las conexiones en las aplicaciones de los nodos móviles durante un handoff debido a la pérdida de paquetes.

Finalmente en el capítulo VI, se presentan las conclusiones de este trabajo de tesis así como las propuestas de trabajo futuro.

Capítulo II

Redes inalámbricas de área local

II.1 Introducción

En los últimos años se ha producido un crecimiento espectacular en el desarrollo y aceptación de las redes inalámbricas, y en concreto, de las redes inalámbricas de área local (Wireless LANs). La función principal de este tipo de redes es proporcionar conectividad y acceso a las tradicionales redes cableadas (Ethernet, Token Ring, etc), como si se tratara de una extensión de estas, pero con la flexibilidad y movilidad que ofrecen las comunicaciones inalámbricas. En países como los EUA, las WLANs son ampliamente usadas en ambientes como hospitales y universidades, donde los usuarios se encuentran en constante movimiento y sus requerimientos de ancho de banda no son muy grandes.

Debido a la creciente proliferación de WLANs, surgió la necesidad de estándares que permitieran interoperatividad a un creciente número de usuarios móviles. Entre los estándares mas importantes se encuentran, IEEE 802.11, HiperLAN y Bluetooth. De estos el más popular es IEEE 802.11. Tal popularidad es puesta en evidencia debido al número de dispositivos comerciales que están basados en este estándar.

II.2 Reseña histórica

Las redes de área local inalámbricas funcionan desde hace más de veinte años en entornos industriales y de investigación. Este tipo de redes se implementó por primera vez en el año de 1979, cuando la casa IBM Suiza utilizó enlaces infrarrojos creando una

red de área local en una fábrica [Gfeller y Bapst, 1979]. Posteriormente se utilizaron implementaciones basadas en tecnologías de microondas según los esquemas de transmisión de espectro ensanchado (Spread Spectrum) [Pahlavan, 1985]. En marzo de 1985 la Comisión Federal de Comunicaciones, FCC, organismo encargado de la regulación de telecomunicaciones en Estados Unidos, asignó a los sistemas WLAN las bandas frecuenciales 902-928 MHZ, 2.400-2.4835 GHz y 5.725-5.850 GHz también conocidas como ISM (Industrial, Científica y Médica) [Pahlavan *et al.*, 1995]. Esta asignación de una localización frecuencial fija propició una mayor actividad industrial. En este punto las redes de área local inalámbrica dejaron de ser meramente experimentales para empezar a introducirse en el mercado. Entre los años 1985 y 1990 se trabajó en el desarrollo de productos WLAN y finalmente, en mayo de 1991, se publicaron algunos trabajos de/sobre redes inalámbricas que superaban la velocidad de transferencia de 1Mbps velocidad mínima a partir de la cual el comité IEEE considera que una red es de área local [Freeburg, 1991; Tuch, 1991].

Hasta este momento las WLAN habían tenido una aceptación marginal en el mercado por dos razones fundamentales: falta de un estándar y precios elevados de la solución inalámbrica. En estos últimos años se ha producido un crecimiento en el mercado de hasta un 100% anual. Este hecho es atribuible a dos razones principales:

- El desarrollo de los equipos portátiles y de las comunicaciones móviles que han permitido que los usuarios puedan estar en continuo movimiento manteniendo comunicación constante con otras terminales y elementos de la red. En este sentido, las comunicaciones inalámbricas ofrecen una prestación no disponible en las redes cableadas: movilidad y acceso simultáneo a los recursos de la red.

- La conclusión de la definición del estándar IEEE 802.11 para redes de área local inalámbricas en junio de 1997 que estableció un punto de referencia y mejoró

muchos de los aspectos de estas redes.

II.3 Tecnologías usadas en las WLAN

Existen varias tecnologías utilizadas en redes inalámbricas. El empleo de cada una de ellas depende mucho de la aplicación dado que cada tecnología tiene sus ventajas y desventajas. Las tecnologías más importantes son:

- **Infrarrojo:** Los sistemas de comunicación por infrarrojo utilizan muy altas frecuencias, justo abajo del espectro de la luz visible para transportar datos. Como la luz, el infrarrojo no puede penetrar objetos opacos, ya sea directamente (línea de vista) o indirectamente (tecnología difundida/reflectiva). El alto desempeño del infrarrojo directo es impráctico para usuarios móviles pero su uso si resulta práctico para conectar dos redes fijas. La tecnología reflectiva no requiere línea de vista pero está limitada a cuartos individuales en zonas relativamente cercanas [Pahlavan *et al.*, 1995].

- **Banda angosta:** Un sistema de radio de banda angosta transmite y recibe información en una radio frecuencia específica. La banda angosta mantiene la frecuencia de la señal de radio lo más angosta posible, solo para pasar la información. El cruzamiento (crosstalk) no deseado entre canales es evitado al coordinar cuidadosamente diferentes usuarios en diferentes canales de frecuencia. En un sistema de radio, la privacidad y la no-interferencia se incrementa por el uso de frecuencias separadas de radio. El radio receptor filtra todas aquellas frecuencias que no son de su competencia. La desventaja de esta tecnología es el amplio uso de frecuencias, una para cada usuario, lo cual es impráctico cuando se cuenta con muchos de estos.

- **Espectro extendido:** La gran mayoría de los sistemas inalámbricos emplean la

tecnología de Espectro Extendido (Spread Spectrum), una tecnología de banda amplia desarrollada por los militares estadounidenses que provee comunicaciones seguras, confiables y de misión crítica. La tecnología de espectro extendido está diseñada para intercambiar eficiencia en ancho de banda por confiabilidad, integridad y seguridad. Es decir, más ancho de banda es consumido con respecto al caso de la transmisión en banda angosta, pero el "trueque" (ancho de banda/potencia) produce una señal que es en efecto mas fuerte y más fácil de detectar por el receptor, que conoce los parámetros de la señal de espectro extendido que está siendo difundida. Si el receptor no está sintonizado a la frecuencia correcta, una señal de espectro extendido se interpretaría como ruido de fondo. Otra característica del espectro extendido es la reducción de interferencia entre la señal procesada y otras señales no esenciales o ajenas al sistema de comunicación. Existen dos tipos de señales de radio de espectro extendido: salto en frecuencia y secuencia directa.

- **Espectro extendido con salto en frecuencia (FHSS):** FHSS[1] utiliza una portadora de banda angosta que cambia su frecuencia central siguiendo un patrón conocido tanto por el transmisor como por el receptor. Tanto transmisor como receptor están debidamente sincronizados comunicándose por un canal que está cambiando a cada momento en frecuencia [Pahlavan *et al.*, 1995]. FHSS es utilizado para distancia cortas, por lo general en aplicaciones punto a multipunto, donde se tiene una cantidad de receptores diseminados en un área relativamente cercana al punto de acceso.

- **Espectro extendido en secuencia directa (DSSS):** DSSS[2] genera un patrón de bits redundante para cada bit que sea transmitido. Este patrón de bits es llamado chip (o código chipping). Entre mas grande sea este chip, es mas grande

[1]Frequency-Hopping Spread Spectrum
[2]Direct-Sequence Spread Spectrum

la probabilidad de que los datos originales puedan ser recuperados, aunque se requerirá mas ancho de banda. Sin embargo, si uno o mas bits son dañados durante la transmisión, técnicas estadísticas embebidas dentro del radio transmisor podrán recuperar la señal original sin necesidad de retransmisión. DSSS se utiliza comúnmente en aplicaciones punto a punto y multipunto.

II.4 Estándar IEEE 802.11

El IEEE (Institute of Electrical and Electronic Engineers) ratificó la especificación original de 802.11 en 1997 [IEEE, 1996] como un estándar para las LANS Inalámbricas (WLAN). La versión 802.11[3] provee tasas de transferencia de bits de 1 Mbps y 2 Mbps así como un conjunto de métodos de señalización y otros servicios. La desventaja con el estándar original 802.11 son las bajas tasas de transferencia de bits que son insuficientes para soportar la mayoría de los requerimientos generales de los negocios. Reconociendo la necesidad de soportar tasas de transmisión más altas, el IEEE ratificó el estándar 802.11b (también conocido como WiFi) para transmisiones de hasta 11Mbps. Con el estándar 802.11b las redes inalámbricas son capaces de lograr un rendimiento y ancho de banda comparable al de las redes alambradas Ethernet de 10 Mbps. Su variante 802.11a ofrece velocidades de hasta 54 Mbps, pero trabaja en la frecuencia de los 5GHz, por lo que productos basados en este estándar no son compatibles con los basados en 802.11b [Nobel, 2000]. Debido a este inconveniente se desarrolló el estándar 802.11g, en donde se extiende la velocidad y el intervalo de frecuencias del 802.11, volviéndose más compatible con los estándares 802.11b y 802.11a. Varios grupos de trabajo están trabajando en diversos desarrollos del estándar 802.11 como se muestra en la tabla I.

Como todos los estándares 802.x, 802.11 se enfocó en las dos capas inferiores del

[3]http://standards.ieee.org/getieee802/802.11.html

Tabla I. Subdivisiones del estándar IEEE 802.11

Grupo	Características
802.11a	Opera entre 5 GHz y 6 GHz con modulación OFDM (Orthogonal Frequency Division Multiplexing). Su inconveniente es que por trabajar a altas frecuencias es mas propicia a interferencias.
802.11b	Llamado también WiFi, con velocidades de 11 Mbps en la banda de los 2.4 GHz. Es de bajo costo y amplia utilización.
802.11d	Nuevos dominios regulatorios para la interoperabilidad de los puntos de acceso.
802.11e	Mejoras al control de acceso al medio (MAC): multimedia, QoS y perfeccionamiento de la seguridad.
802.11f	Protocolo para la interacción del punto de acceso.
802.11g	Trabaja a frecuencias de 2.4 GHz con velocidades mayores a los 22 Mbps. Es compatible con 802.11b y 802.11a.
802.11h	Extensión en soporte para la banda de 5 GHz en Europa.

modelo de referencia OSI: la capa física y la capa de enlace de datos. Este estándar propone tres mutuamente incompatibles implementaciones para la capa física: infrarrojo, o señalización en radio frecuencia (RF, Radio Frequency) usando ya sea secuencia directa (DS, Direct Sequence) de espectro extendido, o salto en frecuencia (FH, Frequency Hopping) de espectro extendido. Una única capa MAC (Medium-Access Control) soporta las tres implementaciones para la capa física, como se muestra en la figura 2.

II.4.1 Arquitectura del estándar 802.11

Cada computadora ya sea móvil, portátil o fija es conocida como una estación en 802.11. Las estaciones móviles acceden a la LAN mientras éstas se mueven. El estándar 802.11 soporta dos tipos de modos: modo infraestructura y modo ad hoc. En el modo infraestructura (ver figura 3) la red inalámbrica consiste de al menos un punto de acceso (AP, Access Point) conectado a la infraestructura de la red alambrada y a un conjunto de estaciones inalámbricas. Esta configuración es llamada BSS (Basic Service Set). Un ESS (Extended Service Set) es un conjunto de dos o mas BSSs formando una subred simple. Dos o más BSS están interconectados usando un sistema de distribución (DS,

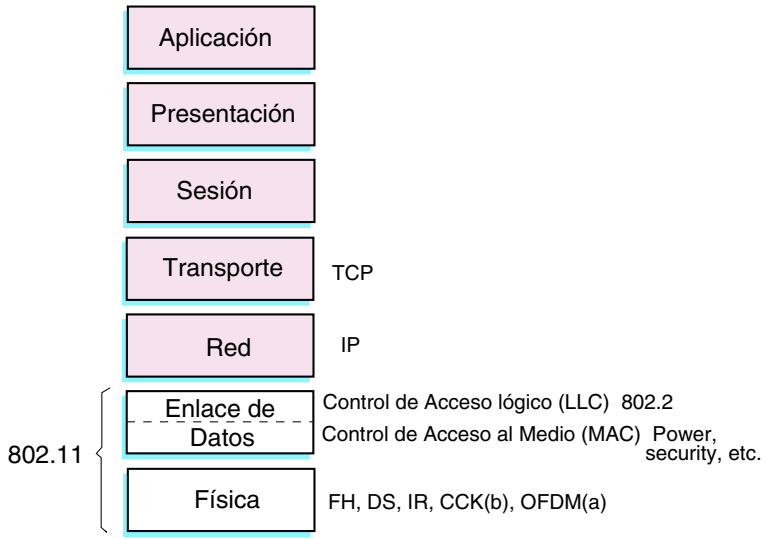

Figura 2. El estándar 802.11 y el modelo OSI

Distribution System). En un ESS la red entera simula ser un BSS independiente a la capa de control de enlace lógico (LLC, Logical Link Control). Esto significa que las estaciones dentro de un ESS pueden comunicarse o moverse entre las BSSs de manera transparente al LLC. Un ESS puede interconectar una WLAN con una LAN por medio de unos dispositivos conocidos como portales. Un portal es una identidad lógica que especifica el punto de integración en el DS donde la red IEEE 802.11 se integra con una red diferente a IEEE 802.11. Si la red es una 802.x, el portal incorpora funciones que son análogas a las de un puente (bridge), esto es, provee un rango de extensión y traducción entre diferentes formatos de marcos (frames) [Crow *et al.*, 1997]. Esta topología es útil para proveer cobertura inalámbrica a un edificio o a un campus.

El modo ad hoc (también conocido como igual-a-igual ó IBSS, Independent Basic Service Set) es básicamente un conjunto de estaciones 802.11 que se comunican una con otra de manera directa sin utilizar un punto de acceso o cualquier conexión a una red alambrada (ver figura 4). En una red ad hoc no existe una estación base y por lo tanto no es necesario un permiso para comunicarse. Esta topología es útil para aplicaciones

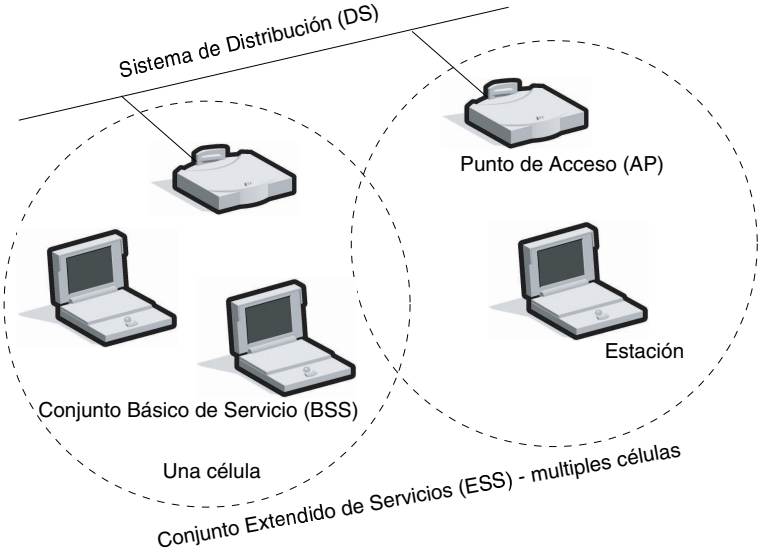

Figura 3. Modo Infraestructura

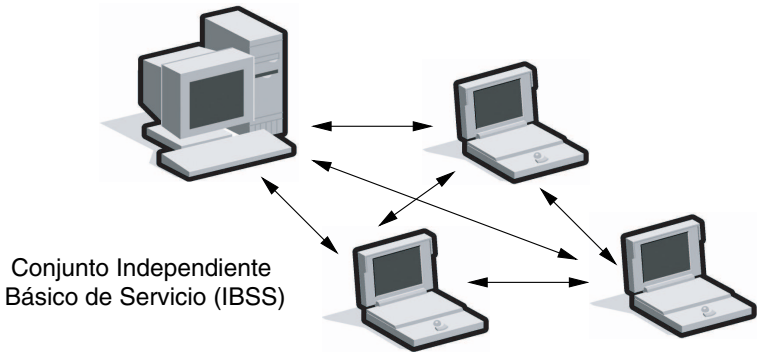

Figura 4. Modo Ad Hoc

tales como transferencia de archivos en lugares donde se forman redes espontáneas, por ejemplo, en conferencias [Crow *et al.*, 1997].

II.4.2 Capa Física

El estándar 802.11 provee tres tipos diferentes de capas físicas, dos de las cuales son de radio de espectro extendido y una especificación de infrarrojo [IEEE, 1996]. Los estándares basados en radio operan dentro de los 2.4 GHz en la banda ISM. El estándar 802.11 define tasas de transferencias de 1 Mbps y 2 Mbps por medio de ondas de radio usando salto en frecuencia (FH, Frequency Hopping) de espectro extendido o secuencia directa (DS, Direct Sequence) de espectro extendido. Es importante tomar en cuenta que FH y DS son fundamentalmente mecanismos de transmisión diferentes y no interoperarán el uno con el otro [LaMaire *et al.*, 1996]. Usando la técnica de salto en frecuencia, la banda de los 2.4 GHz es dividida en 75 subcanales de 1 MHz cada uno. El emisor y receptor acuerdan un patrón de salto en frecuencia, y los datos son enviados sobre una secuencia en los subcanales. Cada conversación dentro de la red 802.11 ocurre sobre un diferente patrón de salto, debido a que estos patrones son diseñados para minimizar la probabilidad de que dos emisores usen el mismo subcanal de manera simultánea. Las técnicas de FHSS permiten un diseño de radio relativamente simple, pero estos están limitados a velocidades no mayores de 2 Mbps. Estas limitaciones son manejadas principalmente por las regulaciones de la FCC (Federal Communications Commission) en Estados Unidos, que restringen el ancho de banda del subcanal a 1 MHz. Por lo tanto, estas regulaciones obligan a los sistemas FHSS a extender su uso a través de toda la banda de los 2.4 GHz, lo que significa que deben de brincar a menudo de frecuencia, lo que conduce a un alto número (*overhead*) de brincos. En contraste, la técnica de señalización de secuencia directa divide la banda de los 2.4 GHz en 14 canales de 22 MHz donde los canales adyacentes se traslapan parcialmente. Los datos

son enviados a través de estos canales de 22 MHz sin la necesidad de brincar a otro canal. Para compensar el ruido de un canal, se usa una técnica llamada “chipping” donde cada bit de los datos del usuario es convertido en una serie de patrones redundantes de bits llamados “chips”. La redundancia inherente de cada chip, combinado con la extensión de la señal a través del canal de 22 MHz, permite hacer posible una forma de verificación y corrección de errores aun si parte de la señal se encuentra dañada, ya que ésta puede ser recobrada en muchos casos minimizando la necesidad de retransmisiones.

II.4.3 Capa de enlace de datos

La capa de enlace de datos del estándar 802.11 consiste de dos subcapas: control de enlace lógico (LLC, Logical Link Control) y control de acceso al medio (MAC, Media Access Control). 802.11 usa el mismo LLC del 802.2 y 48 bits de direccionamiento como otras LANs 802, permitiendo una manera simple de cruzar (bridging) de redes inalámbricas a redes alambradas, aunque la capa MAC [IEEE, 1996] es única a las redes WLAN. La MAC 802.11 es muy similar en concepto a la del 802.3, que fue diseñada para soportar múltiples usuarios en un medio compartido, donde el emisor censa el medio antes de utilizarlo. En redes Ethernet el protocolo CSMA/CD (Carrier Sense Multiple Access with Collision Detection) regula la manera como las estaciones Ethernet establecen su acceso a la red alambrada, así como la manera en que pueden detectar y manejar las colisiones que ocurren cuando dos o mas dispositivos tratan de comunicarse simultáneamente en la red. En una WLAN 802.11 la detección de colisiones no es posible debido al problema conocido como cercano/lejano (near/far), donde para detectar una colisión, la estación debe de poder transmitir y escuchar al mismo tiempo, pero en los sistemas de radio la transmisión no permite a la estación tener la habilidad de escuchar una colisión. Para solucionar este problema 802.11 usa un protocolo ligeramente modificado conocido como CSMA/CA (Carrier Sense Multiple Access

with Collision Avoidance) o la función de coordinación distribuida (DCF, Distributed Coordination Function) [LaMaire *et al.*, 1996]. CSMA/CA intenta evitar colisiones usando un paquete de reconocimiento (ACK, acknowledgment), lo que significa que la estación receptora envía un paquete ACK para confirmar que el paquete de datos llegó de manera intacta. CSMA/CA trabaja de la siguiente manera: cuando una estación desea transmitir, ésta censa el medio, y si no se detecta actividad, la estación espera un periodo de tiempo adicional seleccionado de manera aleatoria y entonces transmite si el medio todavía se encuentra libre. Si el paquete es recibido de manera intacta, la estación receptora envía un ACK y una vez que éste es recibido por el emisor se completa el proceso. Si el ACK no es recibido por la estación emisora, ya sea porque el paquete de datos no fue recibido de manera intacta ó porque el ACK no fue recibido de manera intacta, entonces se asume que hubo una colisión y el paquete de datos es nuevamente transmitido después de esperar una cantidad aleatoria de tiempo. El mecanismo explícito de usar un ACK también maneja de manera efectiva la interferencia así como otros problemas relacionados con las señales de radio, sin embargo permite que las redes 802.11 sufran de un *overhead* que las redes 802.3 no tienen, por lo tanto, las redes 802.11 siempre tendrán un rendimiento inferior al de las redes Ethernet. Otro problema de la capa MAC que es específico de las redes inalámbricas es el del "nodo oculto" (ver figura 5), en el cual dos estaciones en lados opuestos de un AP pueden oír la actividad del AP, pero no pueden oírse entre ellas, probablemente debido a la distancia o a una obstrucción. Para resolver este problema 802.11 especifica un protocolo opcional conocido como RTS/CTS (Request to Send/Clear to Send). Cuando esta opción está en uso, una estación emisora transmite un RTS y espera a que el AP le conteste con un CTS. Debido a que todas las estaciones en la red pueden oír el AP, el CTS causa un retardo en cualquier intento de transmisión, permitiendo que la estación emisora transmita y reciba el paquete ACK sin tener ninguna posibilidad de colisión.

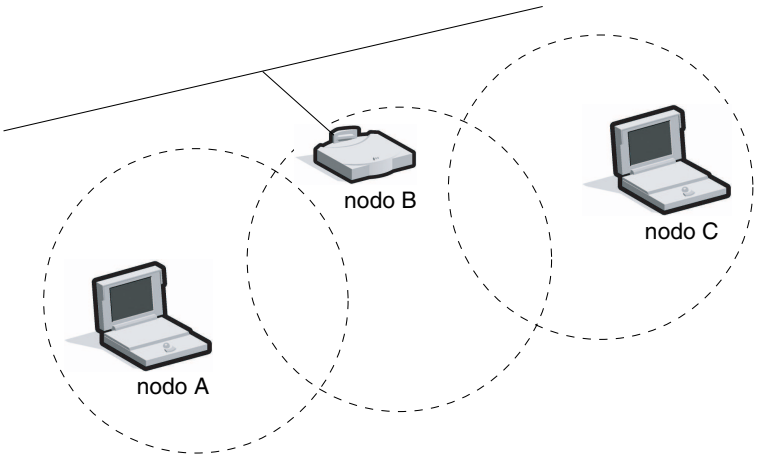

Figura 5. Problema del nodo oculto en redes inalámbricas

Debido a que RTS/CTS agrega *overhead* a la red, dado a que se está reservando temporalmente el medio, esta opción es típicamente usada solo en los paquetes de gran tamaño para los cuales la retransmisión puede ser costosa desde un punto de vista de ancho de banda.

II.5 Otros estándares

Existen otros estándares para WLANs además del 802.11. En esta sección se mostrarán dos de los mas prominentes: HiperLAN y Bluethooth, aunque es importante notar que hasta ahora el mercado de las redes inalámbricas ha sido dominado por el estándar 802.11. Actualmente han comenzado a aparecer varios dispositivos Bluethooth en el mercado, aunque estos están orientados hacia las WPAN (Wireless Personal Area Networks). Mientras tanto se ha demostrado que algunos prototipos de HiperLAN 2 todavía no son productos comerciales [García-Macías y Toumi, 2003].

II.5.1 HiperLAN

Entre 1990 y 1992 el ETSI (European Telecommunications Standards Institute), notó una tendencia que se encaminaba hacia redes inalámbricas más rápidas y mejores, y comenzó el desarrollo de estándares para este tipo de redes. Dentro de esta estructura el proyecto BRAN (Broadband Radio Access Networks)[4] está trabajando en un estándar llamado HiperLAN (High Performance Radio Local Area Network). Este proyecto fue separado en cuatro tipos diferentes de HiperLAN, debido a que se querían tener diferentes tipos de servicios y aplicaciones para WLANs. Estos cuatro tipos son:

- **HiperLAN 1.** Este estándar fue la primera fase en el desarrollo de la familia de estándares HiperLAN. Es una red inalámbrica que es compatible con el estándar ISO 8802 con transferencia de hasta 19 Mbps. Este estándar opera en la banda de los 5.2 GHz con un espectro de 100 MHz y ofrece comunicaciones uno-a-uno así como uno-a-muchos a través de broadcast. El canal provee auto configuración y flexibilidad de uso gracias al canal de acceso distribuido EY-NPMA (Elimination-Yield No-Pre-Emptive Priority Multiple Access). Aunque no es necesario una infraestructura, las estaciones HiperLAN pueden ser usadas para reenviar paquetes con la finalidad de aumentar el área de cobertura. HiperLAN usa la técnica CSMA/CA para resolver colisiones, el esquema reparte la capacidad de radio disponible entre los usuarios activos quienes intentan transmitir datos durante un intervalo de tiempo en común.

- **HiperLAN 2:** Especifica una red de acceso por medio de radio, que puede ser usada con una variedad de redes dorsales (core networks), como por ejemplo IP, ATM, UMTS [Khun-Jush *et al.*, 2000]. HiperLAN 2[5] opera en la banda de los 5.2

[4] http://www.etsi.org/bran
[5] http://www.hiperlan2.com

GHz de frecuencia con un espectro de 100MHz, y se tienen velocidades de hasta 54 Mbps.

- **HiperAccess:** Este estándar es el siguiente paso de HiperLAN 2 y provee acceso inalámbrico externo (outdoor). Tiene un alcance de hasta 5 Km de cobertura entre puntos de acceso inalámbricos y terminales inalámbricas, por ello es propuesto para aplicaciones estacionarias y semi-estacionarias. La frecuencia original de operación era de 5 GHz, pero actualmente se encuentra en el rango de 11 GHz a 42 GHz.

- **HiperLink:** Este estándar esta enfocado a proveer servicios de interconexión a recursos que necesiten altas velocidades de datos como las redes (ejemplo, HiperLANs). Por lo tanto, HiperLink provee interconexiones punto-a-punto a muy altas velocidades de datos que van hasta los 155 Mbps en distancias de hasta 150 metros. La frecuencia de operación se encuentra en la banda de los 17 GHz con un espectro de 200 MHz hasta el momento.

El estándar HiperLAN 1 fue terminado en 1996, aunque la última enmienda para el estándar fue hecha en el año de 1998. Los estándares HiperLANs 2 y 4 fueron diseñados para solamente soportar redes ATM. Aunque actualmente HiperLAN 2 también soporta accesos a redes IP y UMTS. Los nombres para los estándares HiperLAN 3 y 4 fueron cambiados a HiperAccess y HiperLink respectivamente. La figura 6 muestra un panorama general de los estándares HiperLAN.

II.5.2 Bluetooth

Bluetooth es una tecnología de bajo costo, utilizada para conectividad inalámbrica de corto alcance entre dispositivos tales como Asistentes Digitales Personales (PDAs, Personal Digital Assistants), teléfonos celulares, teclados, monitores, impresoras, máquinas

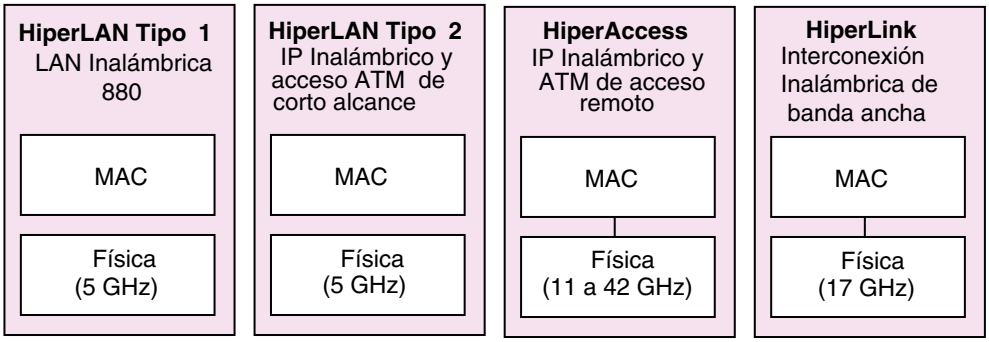

Figura 6. Panorama general de los estándares HiperLAN

de fax, computadoras de escritorio y portátiles, modems, proyectores, etc. El principal mercado es la transferencia de datos y voz entre dispositivos y computadoras personales. Muchas veces este estándar se confunde con el estándar IEEE 802.11, otra tecnología de RF de corto alcance. IEEE 802.11 ofrece más caudal eficaz pero necesita más potencia de transmisión y ofrece menos opciones de conectividad que Bluetooth para el caso de aplicaciones de voz. Las distancias de conexión en Bluetooth pueden ser de hasta 10 metros, aunque actualmente se discute si se aumenta su alcance a 100 metros incrementando la potencia del transmisor a 100 mW.

La versión 1.2 de la especificación Bluetooth acaba de ser liberada en este año, pero el desarrollo de esta tecnología empezó realmente 9 años atrás, en 1994, cuando la compañía Ericsson[6] empezó a estudiar alternativas para comunicar los teléfonos celulares con otros dispositivos. Actualmente es dirigido por el Bluetooth SIG[7] que incluye miembros como Nokia, IBM, Toshiba, Intel, 3Com, Motorola, Lucent Technologies y Microsoft. El origen del nombre de esta tecnología proviene de un Vikingo de origen danés llamado Harald Blatand (Bluetooth) quien en el siglo décimo unificó Dinamarca y Noruega. El nombre fue adoptado por Ericsson, quien espera que Bluetooth unifique

[6]http://www.ericsson.com/
[7]http://www.bluetooth.com

las telecomunicaciones y la industria del cómputo.

Bluetooth, es una tecnología de radiofrecuencia (RF) que opera en la banda de los 2.4 GHz bajo la tecnología de radio conocida como espectro disperso y opera en la banda de libre licencia ISM. La banda de operación está dividida en canales de 1 MHz, y a 1 megasímbolo por segundo puede obtenerse el ancho de banda máximo por canal. Con el esquema de modulación empleado, GFSK (Gaussian Frequency Shift Keying), esto equivale a 1 Mbps [Haartsen, 1998]. Utilizando GFSK, un 1 binario representa una desviación posititiva de la portadora nominal de la frecuencia, mientras que un 0 representa una desviación negativa. Después de cada paquete, ambos dispositivos re-sintonizan su radio transmisor a una frecuencia diferente, saltando de un canal a otro canal de radio de manera seudoaleatoria conocida como espectro disperso con salto en frecuencia (FHSS). De esta manera, los dispositivos Bluetooth utilizan toda la banda de 2.4 GHz y si una transmisión se interfiere sobre un canal, una retransmisión siempre ocurrirá sobre un canal diferente con la esperanza de que este canal esté libre. Cada ranura de tiempo tiene una duración de 625 microsegundos y generalmente los dispositivos saltan una vez por paquete, o sea, saltan cada ranura, cada 3 ranuras o cada 5 ranuras. Como Bluetooth fue diseñado para aplicaciones móviles de poca potencia, la potencia del radio transmisor debe ser minimizada. Tres diferentes clases de niveles de potencias están definidas, las cuales proveen rangos de operación de aproximadamente 10, 20 y 100 metros. Además de las distancias cortas de conexión de Bluetooth, en materia de ancho de banda solo soporta hasta 780 Kbps, los cuales pueden ser utilizados para transferir unidireccionalmente 721 Kbps y 57.6 Kbps en la dirección de retorno o hasta 432.6 Kbps de manera simétrica en ambas direcciones.

Los dispositivos Bluetooth pueden operar de dos maneras: como un maestro o como un esclavo. El maestro es el encargado de fijar la secuencia de los saltos en frecuencia. Los esclavos se sincronizan con el maestro en tiempo y frecuencia para seguir la secuencia

de los saltos del maestro. Además de lo anterior, para controlar la secuencia de los saltos de frecuencia el maestro es el que controla cuando los esclavos pueden transmitir. El maestro permite la transmisión de los esclavos colocando ranuras de tráfico de voz y tráfico de datos. En las ranuras de tráfico de datos los esclavos solo pueden transmitir cuando están contestando una transmisión hecha por el maestro. En las ranuras de tráfico de voz los esclavos pueden transmitir regularmente en ranuras reservadas, aun cuando estén o no contestando al maestro. El maestro controla la manera en que el ancho de banda es dividido entre los esclavos, decidiendo cuando y qué tan a menudo se comunica con cada esclavo. El número de ranuras que un dispositivo obtiene en un tiempo depende de sus requerimientos. El sistema para dividir las ranuras de tiempo entre los diferentes dispositivos es llamado Multicanalización por División de Tiempo (TDM, Time Division Multiplexing).

En Bluetooth una colección de dispositivos esclavos operando de manera conjunta con un solo maestro en común es llamada piconet. La especificación de Bluetooth limita el número de esclavos de una piconet a siete, donde cada esclavo solo puede comunicarse con un maestro compartido. Sin embargo, para obtener un área de mayor cobertura o un mayor número de miembros en la red, pueden enlazarse varias piconets para formar una red esparcida (scatternet), donde algunos dispositivos pueden ser miembros de mas de una piconet (ver figura 7). Cuando un dispositivo está presente en mas de una piconet, debe de compartir su tiempo gastando algunas ranuras en una piconet y algunas otras en la otra piconet. Un dispositivo puede ser esclavo en una piconet y maestro en la otra, o por el otro lado, puede ser esclavo en ambas piconets. No es posible que un mismo dispositivo sea maestro en ambas piconets, debido a que todos los esclavos en una piconet están sincronizados con la secuencia de salto de frecuencia del maestro [Johansson *et al.*, 2001].

Bluetooth permite comunicaciones de tiempo crítico, como aquellas requeridas para

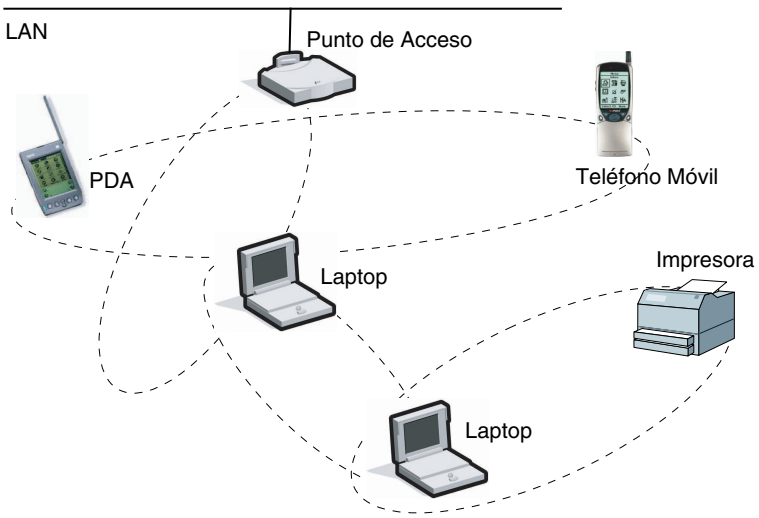

Figura 7. Scatternet Bluetooth de cuatro piconets.

voz y audio. Para llevar este tipo de tráfico, se definen dos tipos de enlaces entre cualquiera de los dos dispositivos. Para comunicación de voz: enlaces Orientados a Conexión Síncrona (SCO, Synchronous Connection Oriented), y para enlaces de datos: enlaces sin Conexión Asíncrona (ACL, Asynchronous Connectionless). El protocolo Bluetooth usa una combinación de circuito y switcheo de paquetes, ya que el canal es ranurado y cada ranura puede ser reservada para paquetes síncronos. La pila de protocolos (ver figura 8) puede soportar un enlace ACL para datos y tres enlaces simultáneos SCO para voz, o una combinación de datos asíncronos y voz síncrona (tipo de paquete DV). Cada canal de voz soporta 64 Kb/s en cada dirección. La pila contiene primeramente un protocolo de nivel físico (banda base) y un protocolo a nivel enlace (LMP) con una capa de adaptación (L2CAP) para los protocolos de capas superiores que interactúan con los de capas inferiores.

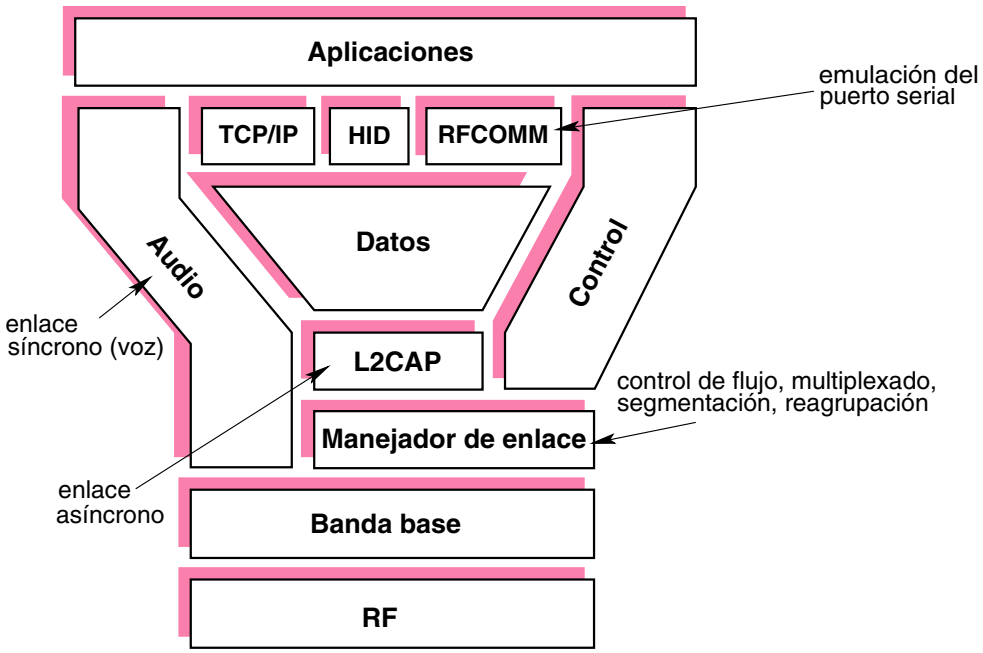

Figura 8. Pila de protocolos de Bluetooth.

II.6 Conclusiones

Debido a su fácil instalación, precios decrecientes y sus velocidades cada vez mayores, las WLANs están gradualmente complementando muchas redes alambradas y se han convertido en las redes de opción para realizar tareas comunes como acceso a Internet. Sin embargo, debemos de tener en cuenta que existen dispositivos que pueden ocasionar interferencia en esta banda regulada, por ejemplo: los hornos de microondas [Kamerman, 1997] y algunos teléfonos celulares de nueva generación. Para dar solución a los problemas relacionados con las diferentes tecnologías inalámbricas, el grupo de trabajo WPAN 802.15 está desarrollando mecanismos y prácticas que faciliten la coexistencia entre las WLANs (como 802.11) y las WPANs (como Bluetooth) aunque la versión 1.2 de Bluetooth ya soluciona el problema de la interferencia con el estándar 802.11. Debemos de tener en cuenta que las WLANs popularizaron el concepto de "movilidad".

Esta movilidad ha producido varios problemas de investigación dentro de las redes, incluyendo el de la transferencia del contexto el cual es el tema principal de esta tesis.

Capítulo III

Movilidad sobre IP

Como se observó en el capítulo anterior, las redes inalámbricas proveen a los nodos de movilidad, la cual sin embargo, no se da mas allá de la cobertura que ofrece su AP. Este problema tiene sus antecedentes desde el inicio de Internet. Este capítulo se enfocará a describir el problema de la movilidad sobre IP, así como también se hablará de Mobile IP, la principal propuesta de solución a este problema. Para entender el problema de la movilidad, es necesario comprender cómo funciona el enrutamiento de paquetes en redes IP y cómo este enrutamiento se ve afectado cuando existe movilidad. Por lo tanto el problema de la movilidad es el siguiente.

Internet es una gran colección de redes que comparten el mismo espacio de direcciones e interoperan utilizando un conjunto de protocolos en común conocido como Protocolo de Control de Transferencia/Protocolo de Internet (TCP/IP). Un concepto fundamental de la arquitectura de Internet es que cada nodo (host) tiene una dirección de red única, la cual utilizan otros nodos para comunicarse con él. Los datos que se transmiten en estas redes son transportados en forma de paquetes, los cuales contienen una dirección fuente y una dirección destino. Para que un nodo pueda comunicarse con cualquier otro nodo, el nodo fuente solo necesita conocer la dirección destino, y los enrutadores de Internet se encargarán de llevar los paquetes desde el nodo fuente hacia el nodo destino.

Dentro de Internet, los enrutadores mantienen un panorama general de la topología de la red en forma de tablas de enrutamiento. Estas tablas son consultadas cuando se toman decisiones de enrutamiento de los paquetes. El proceso de enrutamiento de

paquetes involucra inspeccionar la dirección destino contenida en el paquete, y basados en el contenido de la tabla de enrutamiento, determinar el próximo enrutador al cual se enviará el paquete. Cada enrutador a través de la ruta de un nodo fuente a un nodo destino, repite este proceso hasta que el paquete es finalmente entregado a su nodo destino. Si las direcciones de los nodos son tratadas como identificadores, entonces cada enrutador debería de mantener la información de enrutamiento de cada nodo conectado a Internet, lo cual no es factible debido al gran número de nodos conectados a esta red. Una solución a este problema es imponer a las direcciones una estructura jerárquica. Este direccionamiento jerárquico es esencial si se desea que la arquitectura de enrutamiento sea escalable.

A cada nodo se le asigna una dirección Internet de 32 bits (también conocida como dirección IP) la cual consiste de dos partes: identificador de red (network-id) e identificador de nodo (host-id). Aunque el direccionamiento jerárquico hace el enrutamiento más simple y manejable, conlleva también ciertas limitaciones en su uso. Una dirección jerárquica solo puede ser usada dentro del dominio de su definición. Debido a esto, una dirección IP tendrá algún significado solo si el nodo que la usa permanece conectado a la red denotada por el network-id de la dirección IP. Cuando el nodo se mueva a una nueva red, se le asignará una nueva dirección IP, la cual estará derivada del espacio de direcciones de la nueva red. Por lo tanto, para que el enrutamiento de Internet trabaje, el nodo móvil siempre debe ser asociado con una nueva red cuando se mueva.

Esto se debe a que desde los orígenes del desarrollo del protocolo TCP/IP se asumió que los nodos deberían ser estacionarios, y si durante una conexión activa un nodo se mueve, la conexión se rompe, desestabilizando todos los servicios que se encuentran en la cima de TCP/IP. Para resolver este problema existen dos enfoques:

1. El rediseño completo (ó rehacer el diseño) de los protocolos de Internet con el objetivo de soportar la movilidad de los nodos, o

2. proveer servicios adicionales a la capa de red que hagan posible la movilidad de los nodos en Internet.

Aunque el primer enfoque es sin duda interesante - desde el punto de vista de investigación -, su aplicación práctica no es fácil debido a que requiere cambios radicales en el actual desarrollo de la arquitectura de Internet. Por lo tanto, las actuales soluciones a la movilidad de los nodos descritas posteriormente están basadas en el segundo enfoque.

Para un mejor entendimiento del problema de la movilidad es necesario comprender la diferencia entre los conceptos de nombre y dirección. Un nombre es un identificador de nodo el cual es independiente de la localización del nodo. Una dirección, por otro lado, refleja el punto de conexión del nodo en la red. Los nodos que permanecen estáticos durante todo su tiempo de vida, pueden utilizar su nombre ó su dirección de manera intercambiable. Pero para un nodo móvil, una dirección IP no puede ser usada como único identificador debido a que ésta debe cambiar con la localización del nodo. El nombre es el único identificador independiente de la localización que puede ser usado para localizar a los nodos móviles. Sin embargo, cuando los nombres de los nodos fueron originalmente desarrollados, se asumió que el enlace nombre-a-dirección permanecería estático. Por lo tanto, en vez de referirse a un nodo por su nombre, los protocolos de Internet fueron desarrollados para identificar a los nodos a través de su dirección IP. Un ejemplo de esto, es una conexión TCP, la cual es identificada por una cuádrupla:

(dirección IP fuente, puerto TCP fuente, dirección IP destino, puerto TCP destino)

Si ninguno de los dos nodos se mueve, los componentes de la conexión permanecerán fijos, lo que permitirá que la sesión TCP pueda mantenerse entre los dos nodos, pero si cualquiera de ellos se mueve se tendrá alguno de los siguientes problemas:

- El nodo móvil adquirirá una nueva dirección IP, con lo que sus identificadores de conexión TCP también cambiarán. Esto causará que todas las conexiones TCP en las que se encuentre involucrado el nodo móvil se rompan.

- Si el nodo móvil retiene su dirección IP, entonces el sistema de enrutamiento no podrá enviar los paquetes a su nueva localización.

El problema fundamental reside en que en la arquitectura de Internet, la dirección IP sirve para dos propósitos. Desde la perspectiva de las capas de transporte y aplicación, sirve como un identificador de punto final (endpoint), y para la capa de red la misma dirección IP es usada como una directiva de enrutamiento. Por lo tanto, debido a que el objetivo es asegurar que las conexiones de red no se rompan mientras el host se mueve, o mejor dicho, para mantener las sesiones en la capa de transporte, la dirección IP del nodo móvil debe ser preservada independientemente de su punto de conexión en la red [Bhagwat *et al.*, 1996].

Estudios de investigación en movilidad sobre IP han sugerido que la movilidad es esencialmente un problema de traducción de direcciones, el cual es mejor resuelto en la capa de red [Bhagwat *et al.*, 1996]. Como se muestra en la figura 9, un nodo móvil (MH, Mobile Host) puede moverse fuera de su red hogar y conectarse a Internet a través de una red foránea. Mientras el MH se encuentra en la red foránea, este obtiene una dirección de reenvío (forwarding) derivada del espacio de direcciones de la red foránea. Sin embargo, si un nodo S intenta enviar paquetes al MH, este lo hará usando la dirección hogar del MH como destino. Debido a que el MH se encuentra en una red foránea y no en su red hogar, es necesario el uso de un agente de traducción de direcciones (ATA, Address Translation Agent) en la red hogar, y un agente de reenvío (FA, Forwarding Agent) en la red foránea. Estos agentes realizan las funciones f y g respectivamente, cuya definición es la siguiente:

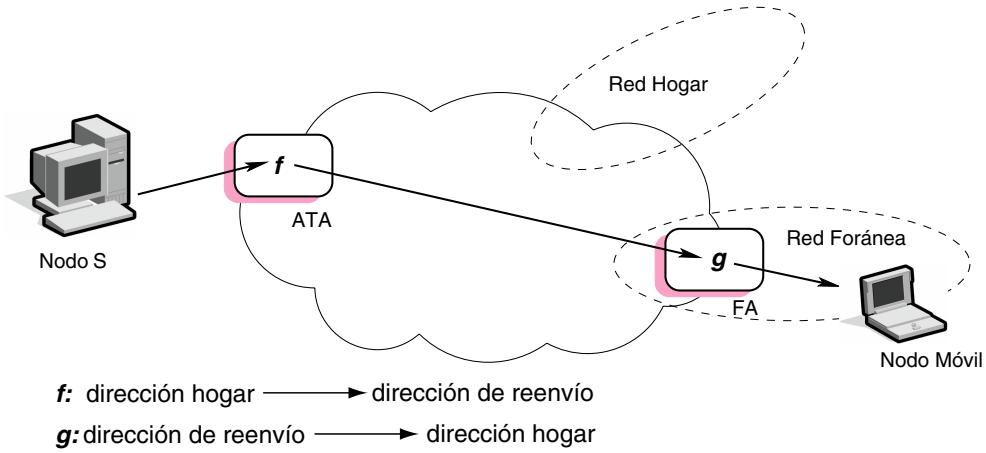

Figura 9. Movilidad como un problema de traducción de direcciones.

- f: dirección hogar → dirección de reenvío.
- g: dirección de reenvío → dirección hogar.

De esta manera, cuando S envía un paquete al MH, el paquete primero pasa a través del ATA. Este agente realiza un mapeo f para enviar los paquetes hacia la dirección que el MH adquirió en la red foránea. En la red foránea, el FA intercepta todos los paquetes que contienen la dirección de reenvío del MH, aplica la función g para mapear la dirección de reenvío a la dirección hogar del MH y entrega de manera efectiva los paquetes.

Se han identificado algunos otros requerimientos para los nodos móviles basados en IP [Myles y Skellern, 1993], entre los que se encuentran: Transparencia operacional, el cual es un requerimiento esencial, y significa que el usuario no debe de desarrollar ninguna otra operación adicional, como configuración manual antes, durante y después de la migración (movilidad) del nodo. Esto solo puede ser logrado usando mecanismos para la detección de la migración y realizando las acciones apropiadas para asegurar que se continúen con los servicios de red de todos los nodos hacia la nueva localización del

nodo móvil. Un requerimiento adicional es el de la transparencia del rendimiento, lo que significa que el uso de un protocolo para nodos móviles no debe de afectar el rendimiento de las aplicaciones del nodo móvil, donde tal rendimiento debe ser comparable con el de las mismas aplicaciones corriendo en nodos fijos. Los factores que aseguran un buen desempeño de la transparencia incluyen:

- Un óptimo enrutamiento de los paquetes hacia y desde los nodos móviles.
- Procedimientos de migración eficientes y robustos.
- Un eficiente uso de los recursos de la red como transmisión y ancho de banda.

Para resolver el problema de la movilidad han surgido diversos protocolos, de los cuales el más destacado es sin duda Mobile IP.

III.1 Mobile IP

Mobile IP es probablemente la propuesta más usada para manejar la movilidad. El estándar Mobile IP especifica una arquitectura para la administración de la movilidad en Internet. Para comprender el funcionamiento de Mobile IP se definirán algunos términos descritos en el RFC 2002 [Perkins, 1996a].

- **Care Of Address (COA).-** COA es el punto de terminación de un túnel hacia el nodo móvil, al cual llegan los datagramas enviados al nodo móvil cuando éste se encuentra fuera de su red hogar. El protocolo puede utilizar dos tipos diferentes de COA: un "foreign agent care of address" que es la dirección de un agente foráneo con el que el nodo móvil está registrado, y un "co-located care of address" el cual es una dirección local que el nodo móvil ha asociado con una de sus interfaces de red.

- **Nodo móvil (Mobile Node, MN).-** Un MN es un dispositivo o enrutador que puede cambiar su punto de conexión a Internet utilizando Mobile IP. El MN mantiene su dirección IP y puede comunicarse continuamente con otro sistema en el Internet mientras mantenga su conectividad por medio de la capa de enlace.

- **Red foránea (Foreign Network, FN).-** Cualquier otra red que no sea la red hogar del nodo móvil.

- **Red hogar (Home Network, HN).-** Una red, posiblemente virtual, que tiene un prefijo que concuerda con la dirección hogar del nodo móvil. Mobile IP no es necesitado dentro de la red hogar.

- **Agente foráneo (Foreign Agent, FA).-** El FA puede proveer varios servicios al MN durante su visita a la red foránea. El FA puede tener el COA que actúa como punto final del túnel y entrega paquetes al MN. Además el FA puede ser el enrutador por omisión del MN.

- **Agente en el hogar (Home Agent, HA).-** El HA, localizado en la red hogar, provee de varios servicios al MN. Primero, es el comienzo del túnel para enviar los paquetes al MN, además de mantener un registro de la localización del MN, esto es, el MN le informa su actual COA.

IP enruta paquetes desde un origen hacia un destino, permitiendo a los enrutadores enviar estos paquetes por medio de tablas de enrutamiento. Las tablas de enrutamiento típicamente mantienen información del siguiente salto para cada destino IP, de acuerdo al número de redes a las que la dirección IP esté conectada. Para que se puedan mantener las conexiones de la capa de transporte, es necesario que el MN mantenga su misma dirección IP, como se mostró al inicio del capítulo.

Mobile IP ha sido diseñado para resolver este problema, permitiéndole al nodo móvil utilizar dos direcciones IP: la *dirección hogar* y la *care-of-address* (COA). La dirección hogar es estática y es usada para identificar las conexiones TCP. La care-of-address cambia en cada nuevo punto de conexión, indica el número de la red e identifica el punto de conexión del nodo móvil con respecto a la topología de la red. La dirección hogar hace parecer que el nodo móvil siempre está preparado para recibir datos de su red hogar, donde Mobile IP requiere de la existencia de un nodo de red conocido como agente en el hogar (HA). Cuando el nodo móvil no esté conectado a su red hogar (y por lo tanto está conectado a una red foránea, FN), el agente en el hogar recibirá todos los paquetes del nodo móvil y los enviará al actual punto de conexión del nodo móvil.

Cada vez que el MN se mueve, debe de registrar su nuevo COA con su HA. Para entregar un paquete al MN desde su HN, el HA envía el paquete desde la HN hasta el COA del MN. Para lograr esto se requiere que el paquete sea modificado de tal manera que la COA aparezca como la dirección IP destino. Esta modificación puede ser entendida como una transformación del paquete, o más específicamente, una redirección. Cuando el paquete llega al COA, la transformación en reversa es aplicada, de esta manera una vez mas el paquete tiene la dirección del MN como dirección IP destino. Cuando el paquete llega al MN, este lo procesará adecuadamente por TCP o cualquier otro protocolo de alto nivel. Ver figura 10.

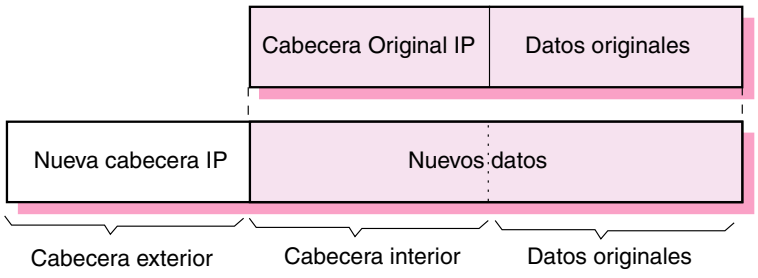

Figura 10. Encapsulamiento IP

En la figura 11 se muestra la operación básica de Mobile IP. Un MN es conectado normalmente a su HN usando una dirección hogar estática. Cuando el MN se mueve a una FN, debe de registrarse con el FA. Una vez registrado el MN se comunica con el HA que se encuentra en su HN y le envía su COA para que pueda identificar el lugar donde se encuentra el FA. Normalmente los enrutadores de la red desarrollarán los roles de FA y HA. Cuando datagramas IP son intercambiados sobre una conexión entre el MN A y el host correspondiente B, ocurren las siguientes operaciones:

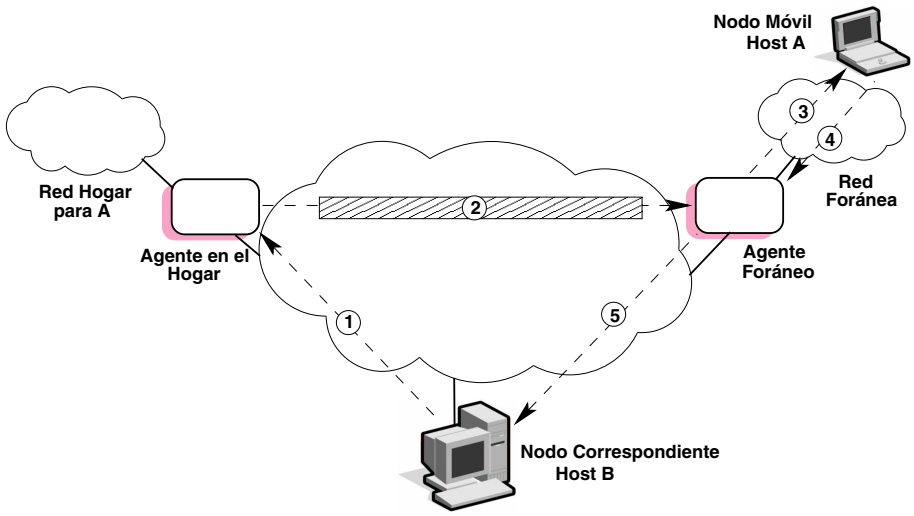

Figura 11. Escenario básico de Mobile IP

1. El host B transmite un datagrama IP destinado al MN A, con la dirección hogar de A en la cabecera IP. El datagrama es dirigido a la red hogar de A.

2. En la red hogar, el datagrama IP es interceptado por el HA. El HA encapsula el datagrama dentro de un nuevo datagrama IP, el cual tiene el COA de A como destino en la cabecera, y es retransmitido.

3. El FA recibe el datagrama, lo desencapsula y lo entrega al MN por medio de la red foránea (por ejemplo, a través de un marco Ethernet).

4. Cuando A envía tráfico IP a B, usa la dirección IP de B. Cada datagrama IP es enviado por A a un enrutador en la red foránea para que dirija los datagramas.

5. El datagrama IP de A a X viaja directamente a través del Internet hacia B, usando la dirección IP de B.

En Mobile IP, el HA redirecciona los paquetes de la HN hacia el COA construyendo una nueva cabecera IP que contiene al COA como dirección IP destino. Esta nueva cabecera protege o encapsula el paquete original, causando que la dirección hogar del MN no tenga efecto en el enrutamiento del paquete encapsulado hasta que éste llegue al COA. Este encapsulamiento es llamado tunelaje (*tunneling*) [Perkins, 1996a].

Mobile IP es mejor entendido como la cooperación de tres mecanismos separables:

- Descubrir el care-of-address.
- Registrar el care-of-address.
- Tunelaje al care-of -address.

III.1.1 Descubrir el care-of-address

El proceso de descubrimiento en Mobile IP ha sido construido usando el protocolo *Router Advertisement*, especificado en el RFC 1256 [Deering, 1991]. El descubrimiento en Mobile IP no modifica los anuncios (advertisement) originales de los enrutadores existentes, sino que simplemente lo extiende para asociarle funciones de movilidad. Un anuncio de enrutador puede llevar información acerca de los enrutadores por omisión, como antes, pero en suma lleva información acerca de uno o más care-of-address. Cuando un anuncio de enrutador es extendido para contener el COA, es conocido como anuncio de agentes (agent advertisements). El HA y el FA típicamente

difunden su anuncio de agentes en intervalos de tiempo, por ejemplo, cada segundo o cada pocos segundos. Si un MN necesita obtener un COA y no desea esperar el periodo de anuncio, el MN puede difundir (broadcast o multicast) una petición que será respondida por el FA o HA que la reciba. HA usa el anuncio de agentes para hacerse conocer, aún si ellos no ofrecen ningún COA.

El anuncio de agentes desarrolla las siguientes funciones:

- Permite la detección de agentes de movilidad.
- Lista una o más COA disponibles.
- Informa al MN acerca de aspectos especiales provistos por los FA, por ejemplo, técnicas de encapsulamiento alternativas.
- Permite a los MN determinar el número de la red y el estado de sus enlaces hacia el Internet.
- Permite a los MN conocer si el agente es un HA, FA o ambos, y por consiguiente si se está en una HN o FN.

Los MN usan solicitudes de enrutador (como se define en el RFC 1256 [Deering, 1991]) para detectar cualquier cambio en el grupo de agentes móviles disponibles en el actual punto de conexión. En Mobile IP esto es llamado solicitud de agente (agent solicitation). Si no son muy lejanos los anuncios detectados de los FA, los cuales previamente habían ofrecido una COA al MN, entonces el MN debe esperar que el FA no se encuentre lejos del rango de la interfaz de red del MN. En esta situación, el MN debe de empezar a buscar su COA, o posiblemente use el COA conocido que se encuentra en los anuncios que él está recibiendo. EL MN puede escoger entre esperar otro anuncio (si es que no ha recibido ningún anuncio reciente del COA) o mandar una solicitud de agente.

III.1.2 Registrar el care-of-address

Una vez que un MN tiene un care-of-address, su HA debe encontrarlo. El proceso comienza cuando el MN, posiblemente con la ayuda de un FA, envía una petición de registro con la información de su COA. Cuando el HA recibe esta petición, agrega la información necesaria a su tabla de enrutamiento, aprueba la petición y envía una respuesta del registro al MN (ver figura 12).

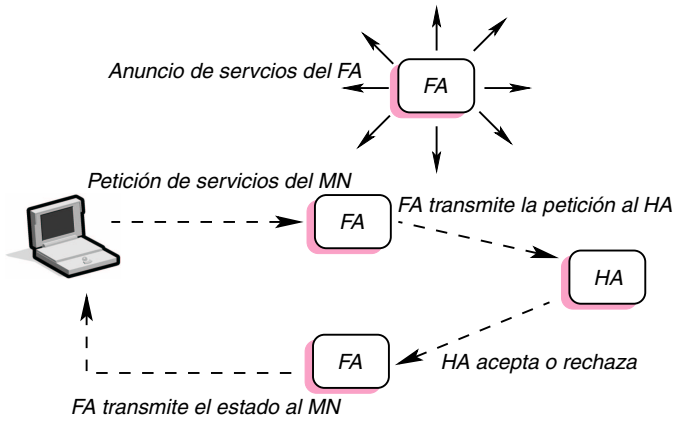

Figura 12. Operación de registro en Mobile IP

Autentificación: Las peticiones de registro contienen parámetros y banderas que caracterizan al tunel por medio del cual el HA entrega los paquetes al COA. Cuando un HA acepta la petición, éste comienza a asociar la dirección hogar del MN con su COA y mantiene esta asociación hasta que el tiempo de vida del registro expira. A la tripleta que contiene la dirección hogar, el COA y el tiempo de vida del registro se le llama *binding* de un MN. Una petición de registro enviada por un HA puede ser considerada también como un *binding update*.

Un *binding update* es un ejemplo de una redirección remota, dado a que esta es enviada remotamente al HA para afectar su tabla de enrutamiento. De aquí se ve muy clara la necesidad de la autentificación. El HA debe estar seguro que dicho registro

lo hizo el MN y no otro nodo malicioso que pretende ser el MN. Un nodo malicioso puede causar que el MN altere su tabla de enrutamiento con información errónea del COA, y de esta manera puede hacer que el MN permanezca inalcanzable para cualquier comunicación desde el Internet.

La necesidad de autentificar la información de registro ha jugado un papel importante en determinar los parámetros aceptables de diseño para Mobile IP. Cada nodo móvil y su HA deben compartir una conexión segura y deben ser capaces de usar el sistema de encriptación *Message Digest 5* (MD5) (RFC 1321 [Rivest, 1992]) con una llave de 128 bits para crear una firma digital segura para la petición de registro.

Para hacer seguras las peticiones de registro cada petición debe contener datos únicos, de tal forma que dos registros diferentes en términos prácticos nunca tendrán el mismo *MD5 hash*. De otra manera el protocolo debe de ser susceptible a ataques repetidos (replay attacks), en que nodos maliciosos puedan guardar registros válidos para posteriormente transmitirlos, y de esta manera romper el enlace del HA con el COA del MN. Para asegurar que esto no pase, Mobile IP incluye dentro de su mensaje de registro un campo especial de identificación que cambia con cada nuevo registro. La semántica exacta del campo de identificación depende de varios detalles, los cuales están descritos en el RFC 2002 [Perkins, 1996a]. Existen dos maneras para hacer el campo de identificación único. Una es el uso de una etiqueta de tiempo (timestamp), de esta manera cada registro tendrá un posterior timestamp y este diferirá de registros previos. La otra manera es hacer que la identificación sea un número aleatorio con suficientes bits de aleatoriedad, los cuales dan una alta probabilidad de que dos valores escogidos para el campo de identificación sean diferentes.

El campo de identificación también es usado por el FA para comparar peticiones de registros pendientes con las contestaciones de registro cuando ellas llegan al HA, y subsecuentemente estará preparado para transmitir la contestación al MN. El FA

también almacena otra información para registros pendientes, incluyendo la dirección hogar del MN, la dirección MAC (Media Access Layer) del MN, el número de puerto para la petición de registro del MN, el tiempo de vida del registro propuesto por el MN y la dirección del HA. El FA puede limitar el tiempo de vida del registro por medio de un valor configurable que coloca dentro de los anuncios de agentes (agent advertisements). El HA puede reducir el tiempo de vida del registro, el cual lo incluye dentro de la contestación del registro, pero jamás podrá incrementarlo [Perkins, 1996a].

III.1.3 Tunelaje al care-of-address

El mecanismo por omisión utilizado para la encapsulación que debe ser implementado por todos los agentes de movilidad que usan Mobile IP es IP-dentro de-IP. Utilizando este método, el HA (inicio del túnel), inserta una nueva cabecera IP o *tunnel header* al principio de la cabecera IP de cualquier datagrama direccionado a la dirección hogar del MN. La nueva cabecera del túnel utiliza la dirección del COA como su IP destino o *tunnel destination*. La dirección IP origen del túnel es el HA, y la cabecera del túnel usa el protocolo de alto nivel número 4, indicando que la siguiente cabecera del protocolo es otra cabecera IP (ver figura 13). En IP-dentro de-IP la cabecera original es preservada como la primera parte de la carga útil de la cabecera del túnel. Por consiguiente, al recobrar el paquete original, el FA elimina la cabecera del túnel y entrega el resto al MN [Perkins, 1996a].

En la figura 13 se muestra como algunas veces la cabecera del túnel usa el protocolo número 55 en la cabecera interna. Esto sucede cuando el HA usa encapsulamiento mínimo [Perkins, 1996b] en lugar de IP-dentro de-IP. Realizar este tipo de encapsulamiento es un poco más complicado que IP-dentro de-IP, debido a que alguna de la información de la cabecera del túnel es combinada con la información de la cabecera interna, que tiene encapsulación mínima para reconstituir la cabecera original IP. Por

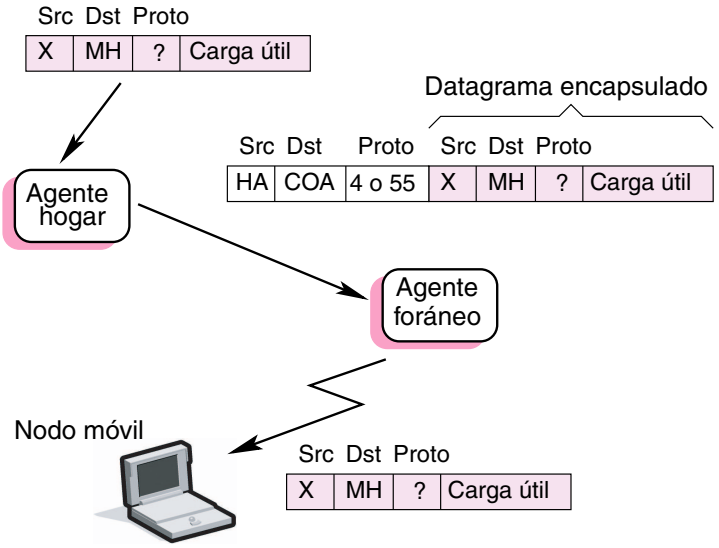

Figura 13. Operación tunneling en Mobile IP

otro lado el *overhead* es reducido.

III.2 La micromovilidad IP

Mobile IP sufre de diferentes debilidades bien conocidas [Campbell *et al.*, 2002; Campbell y Gomez-Castellanos, 2000; Fladenmuller y Silva, 1999] que han llevado a la definición de un enfoque de macro/micro movilidad. En esta sección revisaremos algunas de estas debilidades para mostrar las ventajas de este paradigma.

- Latencia de *handover* y control de tráfico: en Mobile IP, los procedimientos de manejo de movilidad básica están compuestos de dos partes: la detección de movimiento del MN y el registro al HA. Cada vez que el MN cambia su IPPOA[1], estos dos pasos deben ser realizados para que el MN pueda recibir paquetes. Sin embargo es el MN quien inicia este proceso de mandar una petición de registro

[1]IP Point Of Attachment

una vez que ha detectado que se ha movido de una red a otra y que ha obtenido su nuevo COA. Esto introduce dos causas de latencia:

1. Latencia de detección de movimiento: es el tiempo requerido por el MN para detectar que ha cambiado de IPPOA. Este puede ser largo, ya que los mecanismos de detección de movimiento en Mobile IP son basados, ya sea en la expiración del tiempo de vida del aviso del FA o en la comparación de los prefijos de las direcciones de dos o más avisos de agentes.
2. Latencia de registro: como el HA puede estar localizado en cualquier parte del Internet, este proceso puede tomar un tiempo muy grande y algunas veces puede ser imposible completarlo. Es en esta etapa donde se tiene la mayor parte de la latencia del *handover* total.

En el caso de que el MN se mueva velozmente cambiando de red de manera rápida, el proceso de registro será muy ineficiente. Más aún, este mecanismo produce mucho tráfico dentro del dominio local y a través del Internet.

- Calidad de servicio (QoS): Debido a que frecuentemente se cambia de punto de conexión y de COA es difícil otorgar a los usuarios móviles servicios de QoS. Con RSVP (Resource Reservation Protocol) [Braden *et al.*, 1997], por ejemplo, las reservaciones pueden ser hechas una y otra vez, cada vez que el MN cambia de COA, a lo largo de la ruta entera, aún si la parte más larga de la ruta permanece sin cambio.

Por lo anteriormente expuesto, normalmente se utilizarán dos diferentes protocolos para manejar la movilidad:

- Mobile IP maneja los movimientos del MN entre dominios inalámbricos distantes y a través del Internet.

- Otro protocolo maneja el movimiento de los MN dentro del dominio inalámbrico.

Ambos protocolos interactúan de la siguiente manera: El MN obtiene un COA local cuando se conecta a un dominio. El COA permanece válido mientras el MN se encuentre en ese dominio y el móvil realizará su registro con el HA mientras se conecta con el dominio. Los movimientos del usuario dentro del dominio son manejados por un protocolo de micromovilidad, esto es transparente al HA y al resto de Internet. De hecho, para el HA cada dominio inalámbrico se convertirá en una subred Mobile IP. La latencia y el control de tráfico a través de la red serán extremadamente reducidos. Ésta es la principal razón para adoptar el enfoque de micromovilidad, dado que cada propuesta de micromovilidad apunta a reducir la latencia de detección de movimiento y a optimizar el manejo del *handover* dentro de un dominio.

En el caso de calidad de servicio, como la red no está consciente de los movimientos de un usuario dentro de un dominio, las reservaciones serán hechas de nuevo sólo cuando el MN cambie de dominio. Esto reduce el control de tráfico y el retardo en cuanto al proceso de registro. Esto solo es posible, si el protocolo de micro-movilidad soporta el uso de RSVP [Braden *et al.*, 1997] u otros mecanismos de calidad de servicio.

Los protocolos de micromovilidad fueron diseñados para utilizarse en ambientes donde los MN cambian de manera frecuente su punto de conexión en la red, provocando que el mecanismo de tunelaje de Mobile IP introduzca *overhead* a la red en términos de un mayor retardo (delay), pérdida de paquetes y señalización. El objetivo de los protocolos de micromovilidad es manejar la movilidad local (dentro de un dominio) de los MNs sin la interacción con Mobile IP. Una característica importante de estos protocolos es su habilidad para reducir el *overhead* de señalización causado por las frecuentes migraciones de los MN, tomando en cuenta el modo de operación de estos (activo ó inactivo). Los protocolos de micromovilidad pueden ser ubicados dentro de una de las siguientes categorías: movilidad jerárquica, tunelaje jerárquico y

enrutamiento específico a la movilidad, las cuales se describen a continuación:

Movilidad jerárquica

Este enfoque reduce el impacto de la ejecución de la movilidad por medio del manejo local de las migraciones locales, además de ocultar estas migraciones a los HAs. Esto significa que la dirección IP conocida por el HA no refleja el punto de conexión exacto del MN, sino que representa la dirección IP de un gateway que es común a un gran número de puntos de acceso en la red. Cuando un MN se mueve de un AP a otro (los cuales son accesibles a través del mismo gateway), el HA no necesita ser informado del movimiento del MN. Por lo tanto, el rol de los protocolos de micromovilidad es asegurar que los paquetes que lleguen al gateway sean enviados al AP apropiado. Para enrutar estos paquetes se mantiene una base de datos de localización, la cual mapea el identificador del MN a la localización de este. Los protocolos de micromovilidad requieren que los host que participan en el enrutamiento de paquetes mantenga una lista de entradas (una entrada por cada MN) y busquen en ésta por cada paquete que se desea enrutar. Para enrutar un paquete, cada host debe leer la dirección IP destino del paquete, buscar una entrada correspondiente en la lista y reenviarlo hacia el siguiente host. Existen dos tipos de movilidad jerárquica: tunelaje jerárquico y enrutamiento específico a la movilidad [Campbell y Gomez-Castellanos, 2000].

Tunelaje jerárquico (Hierarchical Tunneling)

En este enfoque, la base de datos es mantenida de manera distribuida por un conjunto de agentes foráneos que se encuentran en la red acceso. Cada agente foráneo lee la dirección IP destino del paquete y busca en su lista de visitantes una entrada que corresponda a esa dirección IP. Si la entrada existe, ésta contiene la dirección IP del siguiente agente foráneo. La secuencia de estas entradas constituye la ruta de los

paquetes y la localización del MN. Esta propuesta se apoya en una estructura arbórea de agentes foráneos. El agente foráneo raíz recibe tráfico encapsulado del HA. Cada agente foráneo en la red desencapsula y reencapsula los paquetes conforme los enrutan en la red. Como los MNs se mueven de un AP a otro, las actualizaciones de la rutas se hacen en el punto mas óptimo del árbol, desviando el tráfico hacia el nuevo AP. Hierarchical Mobile IP usa este tipo de movilidad.

Enrutamiento específico a la movilidad (Mobile-Specific Routing)

Este enfoque evita el *overhead* que presenta el esquema anterior al desencapsular y reencapsular los paquetes. Esta propuesta se basa en la actualización de rutas para enrutar los paquetes hacia el MN. La actualización se hace a través de señalización implícita (basada en el espionaje de paquetes) o explícita, o estando consciente de que un protocolo de enrutamiento está en uso. El gateway desencapsula los paquetes y los envía hacia la estación base. Dentro de la red de acceso los host son identificados por su dirección IP y los paquetes son enrutados usando mobile-specific routing sin necesidad de tunelaje o conversión de direcciones. Este protocolo de enrutamiento asegura que los paquetes sean entregados al MN correspondiente. Cellular IP y HAWAII utilizan este enfoque de micromovilidad.

III.2.1 Hierarchical Mobile IP

Hierarchical Mobile IP es una extensión natural a Mobile IP para dar un soporte eficiente a la micromovilidad. El principio es que el MN que se conecte por primera vez al dominio se registre solamente una vez con su HA, con la dirección del *Gateway Foreign Agent* (GFA) como COA [Gustafsson *et al.*, 2001]. Cualquier movimiento futuro del móvil dentro del dominio será transparente al HA. Dentro del dominio, el MN solamente podrá realizar registros regionales. Este tipo de registro es mandado por el MN

al GFA cada vez que este cambia de FA (o sea de IPPOA). El registro contiene el nuevo COA "local" del MN: la dirección que puede ser usada por el GFA para alcanzar al MN mientras éste se encuentra conectado al mismo FA. Esta dirección puede ser de dos formas, una dirección co-localizada, o la dirección del FA. El enrutamiento con Hierarchical Mobile IP es muy simple, un paquete que se envía al MN es en principio interceptado por el HA y por medio del túnel (tunneling) enviado a el GFA. El GFA desencapsula el paquete y es reenviado por medio del túnel hacia el actual COA local del MN.

Hierarchical Mobile IP soporta también una jerarquía multi-nivel de FAs entre los niveles IPPOA y el GFA. Cada FA en la jerarquía debe de mantener un enlace en su lista de visitantes para cada MN conectado a un IPPOA de menor nivel en la jerarquía. Estos enlaces son establecidos y refrescados por las peticiones regulares de registro y contestados por los MN que cambiaron su posición en la red. En este caso, los registros regionales enviados por un MN son solamente enviados al primer FA que ya tiene un enlace para el MN. Los niveles superiores de la jerarquía no están conscientes de los detalles de los movimientos de los MNs ya que ellos no tienen que cambiar sus enlaces. De esta manera la administración del *handoff* está limitada a un número muy pequeño de máquinas [Gustafsson *et al.*, 2001]. En la figura 14(A) y 14(B) se muestran modelos básicos de estas redes, en donde se observa, que la figura 14(B) contiene un número mayor de niveles jerárquicos. Hierarchical Mobile IP también ha incursionado en el mundo de IPv6. Esto se debe a que IPv6 fue desarrollado para manejar la macromovilidad pero no la micromovilidad, por lo tanto en [Castelluccia, 1998] se muestra el funcionamiento de este protocolo bajo IPv6.

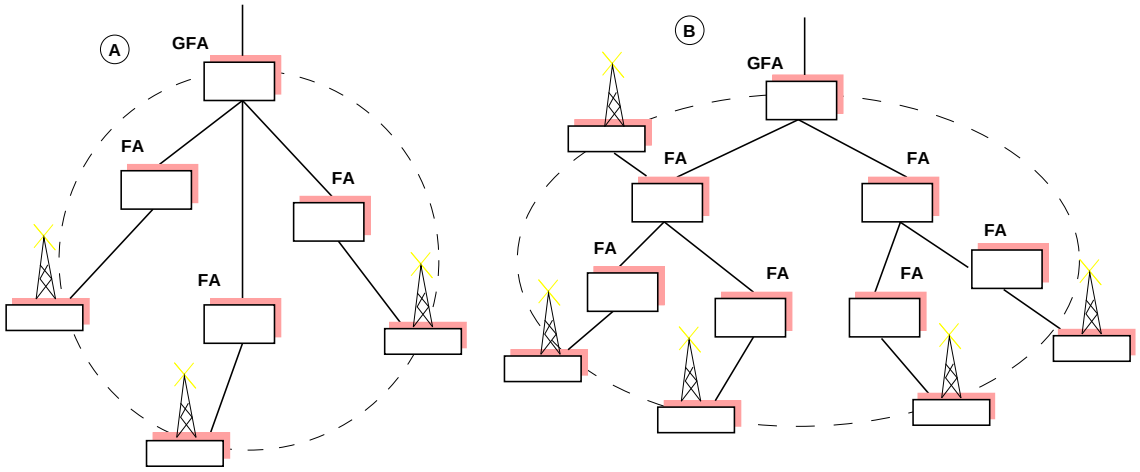

Figura 14. Arquitectura básica en Hierarchical Mobile IP

III.2.2 Cellular IP

El objetivo de Cellular IP es el de integrar los principios de la tecnología celular dentro de las redes IP. Esta propuesta posee grandes cambios, los cuales son diferencias fundamentales entre las arquitecturas de redes celulares y redes IP. Un dominio Cellular IP está compuesto de Agentes de Movilidad (MA, Mobility Agent) y uno de ellos actúa como un gateway hacia Internet y como un FA de Mobile IP para macromovilidad. Un nodo Cellular IP constituye el componente universal de una red Cellular IP, ya que éste sirve como un punto de acceso inalámbrico pero al mismo tiempo enruta paquetes IP e integra funciones de control celular, las cuales son tradicionalmente encontradas en Mobile Switching Centres (MSC) y Base Station Controllers (BSC). Los nodos Cellular IP son nodos IP modificados donde el enrutamiento estándar es modificado por el propio enrutamiento de Cellular IP y por sus funciones de administración de localidad.

La movilidad entre los gateways es manejada por Mobile IP, mientras que la movilidad dentro de las redes de acceso es manejada por Cellular IP. Los host móviles conectados a la red utilizan la dirección IP del gateway como su Mobile IP COA. En

la figura 15 se muestra una red de acceso Cellular IP. Los paquetes primeramente son enrutados hacia el HA y de ahí son enviados por medio de un túnel (tunneling) hacia el gateway. El gateway desencapsula los paquetes y los envía hacia la estación base. Dentro de la redes Cellular IP, los host móviles son identificados por su dirección hogar y los paquetes son enrutados sin el uso de un túnel o conversión de direcciones [Valko, 1999].

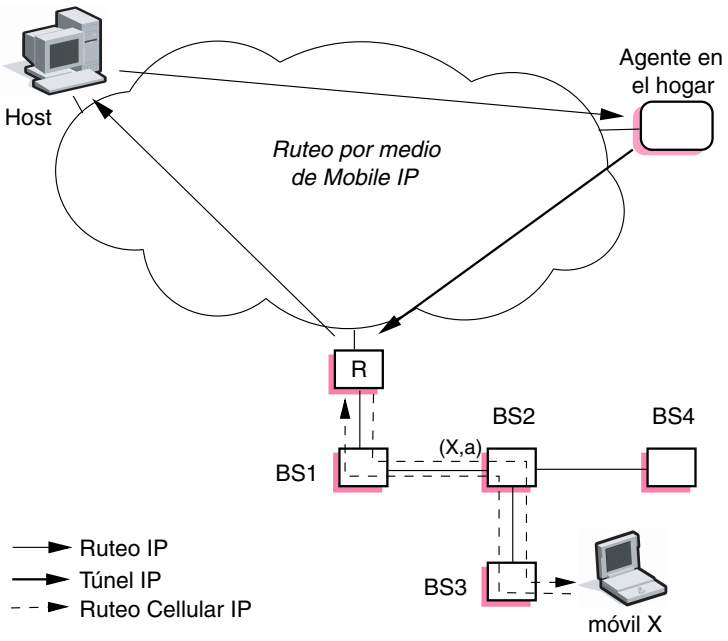

Figura 15. Red de acceso Cellular IP

Como se mencionó anteriormente, los nodos Cellular IP integran administración de localidad y funciones de enrutamiento, por lo tanto lo que sigue es un panorama general de estas funcionalidades, las cuales incluyen enrutamiento, paginación (paging) y *handoff*.

Enrutamiento (Routing)

Por simplicidad y estabilidad, ninguno de los nodos de Cellular IP conoce la localización exacta de un host móvil. Los paquetes dirigidos a un host móvil son enrutados salto por salto a su actual estación base, donde cada nodo sólo necesita saber cuál de sus puertos de salida necesita usar para enviar los paquetes. El gateway periódicamente difunde (broadcast) una bandera y la usa para enrutar paquetes hacia el gateway. Todos los paquetes transmitidos por un host móvil, sin importar su dirección de destino, son enrutados hacia el gateway usando estas rutas.

La manera en que se enruta en Cellular IP puede ser ejemplificada utilizando la figura 15. Los paquetes que son transmitidos por un host móvil con dirección IP X entran a *BS2* por medio de su interfaz a. En el cache de enrutamiento de *BS2* esto es indicado por un mapeo *(X, a)*. Este mapeo permanece válido por una cierta cantidad de tiempo, la cual es el tiempo válido de la ruta (route-timeout) y su validez es renovada por cada paquete de datos que cruza la misma interfaz de entrada para el mismo host móvil. Para que el host móvil esté enviando regularmente paquetes de datos, las estaciones base que se encuentran a lo largo de la ruta de la actual localidad del host y el gateway, mantienen entradas válidas en su antememoria (cache) de enrutamiento para formar una ruta de estado suave (soft-state) entre el host móvil y el gateway. Los paquetes enviados al mismo host móvil son enrutados salto por salto usando el cache de enrutamiento establecido. Si el host móvil no está transmitiendo paquetes de datos regularmente, el puede mantener su cache de enrutamiento enviando paquetes de datos vacíos, en intervalos regulares de tiempo. Estos paquetes son llamados paquetes de actualización de ruta *(route update packets)*.

Handoff

Todos los *handoff* en Cellular IP son iniciados por el host móvil. El algoritmo de

handoff duro es basado en un enfoque simple de administración de movilidad que soporta *handoffs* rápidos y simples al precio de que algunos paquetes se pierdan. Los host móviles escuchan las banderas transmitidas por las estaciones base e inician el *handoff* basado en la medida de la fuerza de la señal. Cuando un host móvil se acerca a una nueva estación base, éste redirecciona sus paquetes de datos de la antigua estación base a la nueva. El primero de los paquetes redireccionados automáticamente configurará un camino de enrutamiento en el cache del host, esta vez para la nueva estación base. La latencia del *handoff* es el tiempo que se tarda entre el *handoff* y la llegada del primer paquete por medio de la nueva ruta. Para mejorar este mecanismo, el protocolo también define el también llamado *handoff* semi-suave. Basado en la recepción de un SHRT (Strong Handoff Radio Trigger) por el host móvil antes de la ocurrencia de un *handoff*, el móvil puede enviar un paquete especial para establecer un bicasting del tráfico a la estación base antigua y a la nueva, con lo cual se reduce significativamente las pérdidas de paquetes. El *handoff* duro no provee garantías, mientras que el *handoff* semi-suave asegura que la pérdida de paquetes sea significativamente reducida [Valko, 1999].

Paging

Los host móviles inactivos, periódicamente generan pequeños paquetes de control llamados *paging update packets*, los cuales son enviados a las estaciones base más cercanas. Estos paquetes al igual que los de datos son enrutados salto por salto hacia el gateway. Las estaciones base pueden mantener un *cache paging*, el cual es actualizado por cualquier paquete enviado por el host móvil, incluyendo los paging update packets. De esta manera podemos decir que paging es un mecanismo que maneja la administración de la localidad de los host móviles inactivos.

III.2.3 HAWAII

HAWAII fue sometido para su consideración en el IETF (Internet Engineering Task Force), pero no avanzó más allá de un borrador (Draft) de Internet [Ramjee *et al.*, 1999]. Bajo su enfoque, Mobile IP es usado como la base para el manejo de la movilidad en redes inalámbricas de área local,pero han sido desarrollados nuevos métodos para manejar la movilidad dentro de dominios administrativos. Un punto importante a recalcar es que los host móviles retienen su dirección de red mientras se mueven dentro de un dominio, de esta manera el HA (si usan Mobile IP) y cualquier otro host, no están conscientes de que el host ha desarrollado un movimiento dentro del dominio.

En forma similar a Cellular IP, HAWAII es un protocolo de micromovilidad que se apoya en Mobile IP para la macromovilidad. A diferencia de Cellular IP, HAWAII no reemplaza a IP sino que trabaja encima de él. Cada estación dentro de la red no solamente actúa como un enrutador IP, sino que también soporta funciones específicas de movilidad. El principio básico de HAWAII es similar al principio de Cellular IP, donde cada estación mantiene un cache de enrutamiento para manejar la movilidad y las transmisiones salto por salto de paquetes especiales en la red activan a las estaciones a actualizar su cache de enrutamiento. Como en Cellular IP, se supone que la red está organizada como un árbol jerárquico y un único gateway está localizado en la raíz de éste.

Para tener una mejor comprensión del funcionamiento de HAWAII, se definirá "enrutador de intersección" como el enrutador que se encuentra más próximo al MN, y que es la intersección de dos rutas: la primera ruta entre el enrutador raíz de dominio (gateway) y la antigua estación base, y la segunda ruta entre la antigua estación base y la nueva estación base. Cuando un MN realiza un *handoff* en HAWAII, todas las tablas de reenvío involucradas son modificadas para reencaminar los paquetes hacia la nueva

localización del MN. Estos cambios son hechos a través de uno de los cuatro posibles esquemas de configuración de rutas que determinan cuándo, cómo y qué enrutadores son actualizados. Estos esquemas son: MSF (Multiple Stream Forwarding), SSF (Single Stream Forwarding), UNF (Unicast NonForwarding) y MNF (Multicast NonForwarding), los cuales se encuentran agrupados en esquemas de reenvío y no reenvío:

- **Reenvío (forwarding).-** Este tipo de esquemas son independientes del enlace inalámbrico y se apoyan en la red alambrada para almacenar (buffer) paquetes y después enviarlos a la nueva estación base. Los paquetes son primeramente enviados de la antigua estación base a la nueva estación base, antes de que sean desviados al enrutador de intersección.

- **No reenvío (nonforwarding).-** Este esquema toma ventaja de ciertas propiedades de los enlaces inalámbricos, donde la antigua y la nueva estación base pueden mantener conectividad con el MN para realizar una entrega transparente de paquetes durante un *handoff*. En este esquema, los mensajes de configuración de ruta viajan de la nueva estación base hacia la antigua estación base, por lo que los paquetes son desviados en el enrutador de intercepción hacia la nueva estación base, lo que resulta en un no reenvío de paquetes desde la antigua estación base.

Esta clasificación se debe a la manera en que se entregan los paquetes al MN durante un *handoff*.

Esquemas forwarding: MSF y SSF

Debido al orden en que las tablas de reenvío son modificadas, bajo el esquema MSF (Multiple Stream Forwarding) se pueden originar ciclos transitorios de enrutamiento. También existe la posibilidad de que se originen flujos de paquetes desordenados, lo que se debe al método de reenvío empleado. Para superar los problemas de MSF se

propuso otro método de forwarding llamado SSF (Single Stream Forwarding), el cual no solamente toma en cuenta la dirección IP y la interfaz de salida, sino también la interfaz (ó interfaces) de entrada. SSF posee un mayor grado de complejidad, pero ayuda a resolver los ciclos transitorios de enrutamiento así como los flujos de paquetes desordenados que presenta MSF. Los autores de HAWAII muestran que la complejidad agregada a SSF no mejora de manera significativa el establecimiento de las nuevas rutas, debido a que los *handoff* típicos involucran a enrutadores que se encuentran a uno o dos saltos de distancia.

Esquemas nonforwarding: UNF y MNF

Los dos tipos de esquemas nonforwarding, fueron motivados por los dos tipos de redes inalámbricas existentes. El esquema UNF (Unicast NonForwarding) está optimizado para redes donde el MN es capaz de escuchar/transmitir a dos o más estaciones base de manera simultánea durante un corto periodo de tiempo, como es el caso de las redes CDMA[2]. El esquema MNF (Multicast NonForwarding) está optimizado para redes donde el MN es capaz de escuchar/transmitir solamente con una estación base, como es el caso de las redes TDMA[3]. Gracias a la redirección utilizada tanto en UNF como en MNF, se asegura que no existan pérdidas de paquetes o flujos de paquetes desordenados, así como también evitá los ciclos transitorios de enrutamiento. Bajo MNF existe un corto periodo de tiempo donde el enrutador de intersección envía datos hacia ambas interfaces de salida. Los esquemas nonforwarding son menos complejos y es más probable que puedan ser implementados[4]. En [Ramjee *et al.*, 2002] se muestra la manera en que trabajan estos esquemas.

[2]Code Division Multiple Access

[3]Time Division Multiple Access

[4]Las soluciones propuestas para administrar la movilidad en HAWAII solamente han sido validadas por medio de simulaciones y hasta ahora no se conoce ninguna implementación.

Como se muestra en la figura 16(a), cuando en UNF la nueva estación base recibe el mensaje de configuración de ruta, ingresa una dirección de reenvío para la dirección IP del MN, cuya interfaz de salida es la interfaz a través de la cual recibió el mensaje. En seguida revisa su tabla de enrutamiento para determinar el próximo enrutador, Enrutador 2. Este enrutador realiza una acción similar y envía el mensaje hacia el enrutador 0, el enrutador intersección en este caso. El enrutador 0 ingresa una nueva entrada de reenvío, con el objetivo de que los nuevos paquetes sean desviados directamente hacia el MN a través de la nueva estación base. Eventualmente, el mensaje 5 alcanzará la antigua estación base, la cual cambiará su entrada de reenvío y enviará un mensaje de aceptación (acknowledgment) - mensaje 6 - hacia el MN. El esquema MNF es muy similar al esquema UNF y la principal diferencia es que el enrutador de intersección, Enrutador 0, realiza un multicast de paquetes de datos durante un periodo corto de tiempo. En la figura 16(b), después de que el enrutador 0 recibe el mensaje 3, éste envía paquetes hacia ambas estaciones base, la antigua y la nueva, hasta que este recibe el mensaje 6. Esto ayuda a limitar el número de paquetes perdidos en la red, donde el MN solo puede escuchar a una sola estación base.

Al igual que Cellular IP, HAWAII soporta conectividad pasiva con un mecanismo de paginación (paging). Las áreas de paginación están compuestas de estaciones pertenecientes al mismo grupo IP multicast. Las peticiones de paginación, deben alcanzar todas las estaciones en un área y son transmitidas por medio del grupo multicast correspondiente a esa área. Para soportar eficientemente la calidad de servicio, HAWAII define una integración nativa de *RSVP* adaptada a la movilidad del usuario [Ramjee *et al.*, 2002].

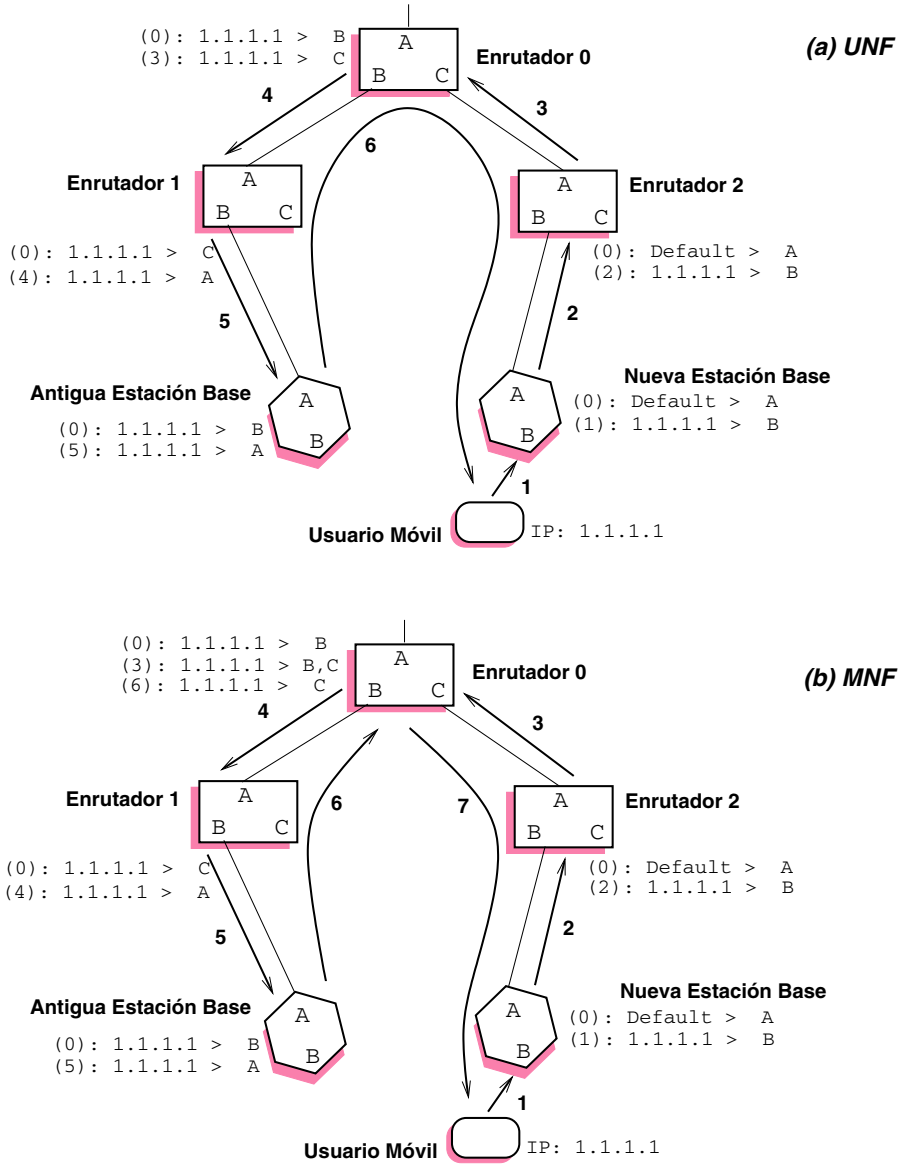

Figura 16. Esquema nonforwarding para configuración de rutas.

III.3 Conclusiones

Debido a que el Internet nació en una era donde no existían dispositivos móviles, todos los protocolos básicos fueron diseñados con la idea de que los dispositivos deberían permanecer en un lugar fijo. Conforme fueron apareciendo dispositivos portátiles, surgió la necesidad de que estos dispositivos pudieran permanecer conectados mientras se desplazaban, por lo cual fue necesario crear nuevos protocolos que se encargaran de resolver este problema. Mobile IP puede verse como una culminación de muchas de estas propuestas, aunque también se debe hacer hincapié en sus debilidades cuando es necesario realizar varios *handoffs* dentro de un mismo dominio. Por lo tanto, surgió la necesidad de realizar un protocolo de micromovilidad que se encargara de realizar la movilidad dentro de un dominio, mientras que Mobile IP se encargaría de manejar la movilidad entre dominios. Entre los protocolos de micromovilidad más importantes encontramos a Hierarchical Mobile IP, Cellular IP y HAWAII.

Estos protocolos de micromovilidad no resuelven varios problemas importantes para lograr una movilidad transparente, como es el caso del descubrimiento de los posibles candidatos a enrutadores de acceso dentro de la red y la selección del mejor de entre ellos, así como tampoco realizan la transferencia del contexto del MN entre los enrutadores de acceso. Estos temas son tratados en el siguiente capítulo.

Capítulo IV

Movilidad transparente

Cuando un nodo móvil se desplaza y requiere realizar un *handoff*, los protocolos de micromovilidad asumen que el nodo móvil ya seleccionó el nuevo enrutador con el cual se realizará el *handoff*. Por lo tanto, estos protocolos no realizan el descubrimiento de los posibles enrutadores de acceso, ni verifican sus capacidades para seleccionar el enrutador que mejor cumpla con las preferencias del nodo móvil. Asimismo, otro de los problemas asociados con los protocolos de micromovilidad es que estos no transfieren el estado -contexto- del nodo móvil hacia su nuevo enrutador, de tal forma que el nodo móvil pueda seguir contando con los mismos servicios que le proveía su antiguo enrutador. Estos problemas ocasionan que no se pueda otorgar una movilidad transparente a los nodos móviles. En este capítulo se tratarán los problemas asociados a la movilidad transparente, la cual incluye el descubrimiento de los candidatos a enrutadores de acceso y la selección del mejor entre ellos así como la transferencia del contexto.

IV.1 Descubrimiento de candidatos a enrutadores de acceso

Un *handoff* con protocolos de movilidad IP involucra el movimiento de un MN de un enrutador de acceso a otro por medio de la capa 3, después de que el MN tenga establecida una conexión por medio de la capa 2 con el nuevo enrutador de acceso. Existen varias maneras de lograr esto, las cuales varían dependiendo de la movilidad IP y de los protocolos para *handoff* suaves, sin embargo, un requisito en común para realizar un *handoff* a nivel IP es descubrir los candidatos a enrutadores de acceso con

los cuales el MN puede realizar un *handoff*. Descubrir el candidato a enrutador de acceso involucra identificar su dirección IP así como sus capacidades en las cuales el MN pueda estar interesado. El grupo de trabajo Seamoby[1] especifica los principales problemas asociados con el descubrimiento de los candidatos a enrutadores de acceso [Trossen *et al.*, 2002], así como los prerequisitos que debe cumplir el protocolo que realice esta función [Krishnamurthi, 2002]. Las siguientes definiciones pueden ayudar a comprender mejor los problemas mencionados anteriormente:

- **Punto de acceso (AP, Access Point).-** Es un emisor/receptor de radio por medio del cual un MN obtiene conectividad con la red alambrada a través de la capa 2.

- **Enrutador de acceso (AR, Access Router).-** Un enrutador IP residiendo en la red de acceso y conectado a uno o mas AP. Un AR ofrece conectividad IP al MN.

- **AR Candidato (CAR, Candidate AR).-** Un AR con el cual el MN tiene una oportunidad de desarrollar un *handoff* a nivel IP. Esto significa que el MN tiene la interfaz de radio correcta para conectarse a un AP que es servido por este AR, cuando la cobertura de este AR se traslape con la del AR al cual está actualmente adjunto el MN.

- **AR Objetivo (TAR, Target AR).-** Un AR con el que son iniciados los procedimientos para realizar un *handoff* a nivel IP. El TAR es seleccionado después de correr un algoritmo de selección de TAR que toma en cuenta las capacidades del CAR, preferencias del MN y algunas políticas locales.

[1] http://www.ietf.org/html.charters/seamoby-charter.html

IV.1.1 Problema de descubrimiento del CAR

Existen dos problemas básicos asociados con el descubrimiento del CAR:

1. Mapear el identificador de capa 2 de un AP a la dirección IP del CAR.

2. Identificar las capacidades del CAR

Los dos problemas están relacionados en tanto que ambos se preocupan por obtener información a nivel IP acerca de un CAR con el propósito de determinar un enrutador de acceso con el cual realizar un *handoff*. Estos dos problemas son discutidos a continuación:

Descubrimiento de las direcciones IP de los CAR

Los protocolos para *handoff* suave requieren una cierta cantidad de señalización IP entre el actual AR del MN y el nuevo AR con el cual el MN desarrollará un *handoff* o desarrolló el *handoff*. Esta señalización es requerida para reconfigurar el enrutamiento de los paquetes durante un *handoff* en Mobile IP cuando el enlace del MN se mueve al AR objetivo. La señalización también puede ser usada para que el actual AR pueda transferir el contexto de los servicios IP al AR objetivo. El contexto de los servicios IP puede incluir: estado de la QoS, el estado AAA, etc. El poder establecer rápidamente el contexto del servicio IP en el AR objetivo es importante, porque determina qué tan rápido el MN puede recibir el mismo nivel de servicio IP en el AR objetivo como lo recibía en su antiguo AR. Para que la señalización a nivel IP pueda ocurrir, el AR actual requiere la dirección IP del TAR. El identificador de capa 2 del TAR puede ser obtenido cuando el MN tiene contacto con la bandera (beacon) del TAR a través del AP conectado a él. Por lo tanto, mapear el identificador de capa 2 del AP a la dirección IP del CAR al cual se encuentra conectado el AP, es tarea de un protocolo de descubrimiento de CARs. Este problema no es distinto al de la determinación de

una dirección en reversa (reverse address resolution [Finlayson *et al.*, 1984]) o al uso de DHCP (Dynamic Host Configuration Protocol) [Droms, 1993] para obtener la dirección IP de un nodo basado en su dirección MAC. Sin embargo, en este caso se debe obtener la dirección IP del CAR en base a la dirección MAC del AP conectado a él, dado a que el MN puede escuchar al AP pero no al CAR. Cualquier solución a este problema debe proveer una autoconfiguaración dinámica en la determinación de la dirección IP, de tal forma que los ARs y APs que sean ingresados o removidos sean rápidamente descubiertos sin requerir mucha o ninguna intervención humana [Trossen *et al.*, 2002].

En la figura 17 se observa que los ARs emiten sus beacons, los cuales contienen sus direcciones MAC, estas direcciones serán almacenadas en una lista que será enviada al actual AR cuando el host móvil desee realizar un *handoff*.

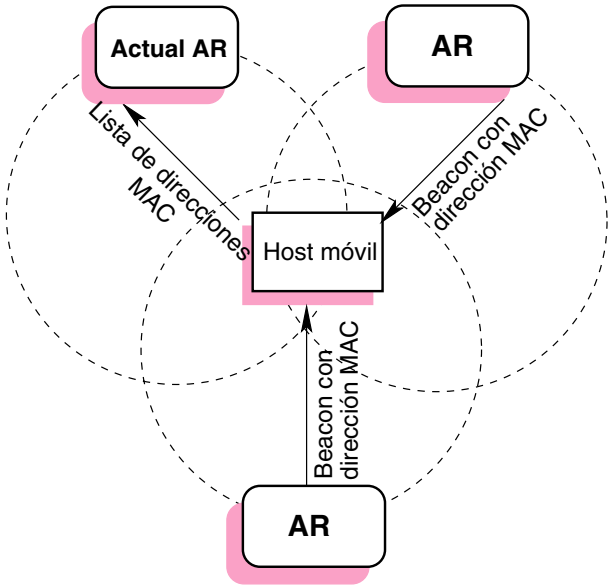

Figura 17. Descubrimiento de direcciones CAR.

Identificación de las capacidades del CAR

Aunque no es común ahora, las futuras generaciones de redes móviles pueden consistir de AR que ofrecen cobertura en la misma área geográfica pero son heterogéneos en capacidades. La función básica de todos los enrutadores IP es el envío de paquetes, y en ese aspecto todos los enrutadores son iguales. Sin embargo, la heterogeneidad puede surgir entre los diferentes AR debido a factores tales como funciones adicionales soportadas por estos AR (soporte para *handoff* suaves, funciones de seguridad, funciones que mejoran el desarrollo inalámbrico, etc.), aspectos administrativos y comerciales para proporcionar servicio al MN (proveedor del servicio, costo de acceso, etc.), disponibilidad de ciertos tipos de recursos con el AR (disponibilidad de QoS), etc. Debido a esta heterogeneidad se necesita una solución que permita al MN enterarse de las capacidades de los CARs en las cuales él esté interesado. La implementación de un protocolo para el descubrimiento de CARs debe permitir esto.

En la figura 18 se muestra el escenario donde el actual AR le pregunta a los ARs (que el MN puede escuchar) sus capacidades, las cuales se utilizarán como entrada en el algoritmo para la selección del TAR.

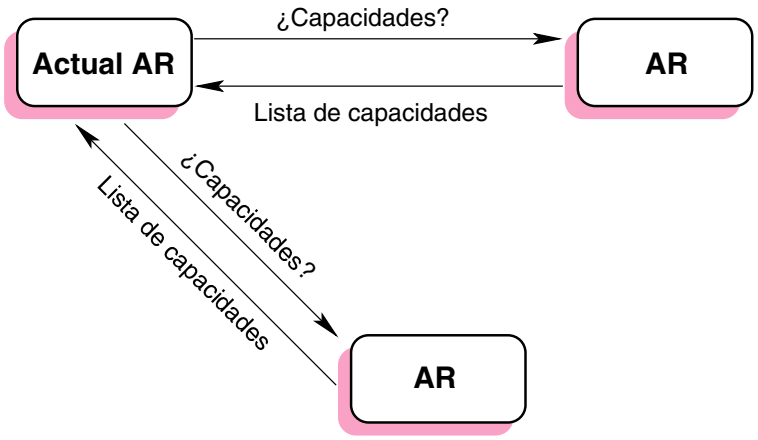

Figura 18. Identificación de las capacidades de los CARs.

IV.1.2 Consideraciones de seguridad

El protocolo para el descubrimiento de los CARs debe permitir a otros nodos adquirir información acerca de un AR, incluyendo su dirección IP y sus capacidades. Nodos maliciosos pueden utilizar este tipo de información para lanzar ataques tipo Denial of Service (DoS) y/o de secuestro (hijacking) de servicios. Por consiguiente, los siguientes puntos deben ser cubiertos por cualquier solución desarrollada para el descubrimiento de CARs:

- Autentificación de los nodos.
- Asociación de seguridad entre nodos.
- Encriptación del mensaje/carga útil.

IV.1.3 Protocolos propuestos para la selección del CAR

Actualmente, el grupo de trabajo SeaMoby ha liberado una especificación del protocolo CAR Discovery [Liebsch *et al.*, 2003], sin embargo, esta especificación se encuentra en sus etapas iniciales (draft) y todavía tiene muchos problemas por resolver, como por ejemplo: seguridad, la aceptación o rechazo de un servidor CARD[2], etc. El protocolo CAR Discovery se basa en la implementación de un servidor CARD, en el cual se encuentren habilitados todos los AP. De esta manera cuando el host móvil requiera realizar un *handoff* enviará a su AR actual las direcciones MAC de los AP que pueda escuchar. El AR actual enviará esta lista de direcciones al servidor CARD, el cual se encargará de mapear la dirección MAC de cada AP a la dirección IP del AR al cual se encuentra conectado cada AP. En base a estas direcciones IP, el AR actual se encargará de preguntar a cada AR candidato sus capacidades y en base a estas capacidades realizar el algoritmo de selección del TAR.

[2]Candidate Access Router Discovery

Así mismo, se sometió la especificación del protocolo DyCARD[3] [Trossen *et al.*, 2003] al grupo de trabajo SeaMoby. Sin embargo, este trabajo se consideró una aportación personal para resolver el problema del CARD y no avanzó mas allá de un borrador (draft) de Internet. El protocolo DyCARD se basa en anticipar la selección del nuevo AR. Esto es posible manteniendo en el cache de los AR sus AP mas cercanos de manera que el AR actual sabe cuales son los posibles AP con los que el MN puede realizar un *handoff*. Cuando el MN envía la lista de AP que puede escuchar, el AR actual puede mapear la dirección MAC del AP a la dirección IP del AR al cual se encuentra conectado el AP. Este protocolo obtiene las capacidades de los AR a través de mensajes diseñados para este propósito. Cabe señalar que ninguna de las dos propuestas especifican como representar las capacidades de los AR, así como tampoco describen el algoritmo para la selección del TAR.

IV.2 Transferencia de contexto

En redes IP que soportan la movilidad de los host, los caminos de enrutamiento entre el host y la red pueden cambiar de manera rápida y frecuente. En algunos casos, el host puede establecer ciertos servicios candidatos para la transferencia del contexto en su localización actual en la red, y cuando el host se mueve realizando *handoffs*, desea beneficiarse de los mismos servicios en su nueva localización. Ejemplos de tales servicios son: calidad de servicio, compresión de cabeceras, seguridad, etc. Sin embargo, para que el host obtenga estos mismos servicios en su nueva localización, el host debe realizar los flujos de señalización necesarios para configurar estos servicios desde el principio. En algunos casos, este proceso puede ser considerablemente lento, debido a que se requiere de cierta cantidad de tiempo para configurar el estado de los servicios así como para realizar los flujos de información necesarios entre los protocolos involucrados. Por lo tanto,

[3]Dynamic Candidate Access-Router Discovery

si el host quisiera restablecer estos servicios por el mismo proceso que utilizó inicialmente, el retardo en tiempo real puede sería fuertemente impactado. Una alternativa es transferir suficiente información del estado actual de los servicios candidatos para la transferencia del contexto - o su contexto - hacia su nueva localización en la red, de tal manera que los servicios puedan ser reestablecidos rápidamente a través de un proceso llamado transferencia de contexto. La transferencia del contexto de los servicios puede ser ventajosa, debido a que minimiza el impacto de la movilidad de los MN en aspectos tales como QoS, políticas locales, servicios como PPP, compresión de cabeceras, etc. La transferencia de contexto puede ser usada como mínimo para enviar la información de configuración necesaria para establecer los servicios y protocolos respectivos. Los siguientes términos descritos por el grupo de trabajo SeaMoby ayudarán a comprender mejor el problema de la transferencia de contexto [Levkowetz *et al.*, 2002]:

- **Contexto.-** Es la información requerida del estado actual de un servicio para restablecer ese mismo servicio en una nueva subred, sin tener que desarrollar desde el principio todo el protocolo de intercambio con el host móvil.

- **Transferencia de contexto.-** Es el movimiento del contexto de un enrutador a otro, o de una red a otra, como un medio para restablecer servicios específicos en una nueva subred o colección de subredes.

- **Servicio candidato para la transferencia de contexto.-** Un servicio que es candidato para la transferencia de contexto. Los Servicios involucrados en el trato que se otorga a los paquetes durante su enrutamiento tales como QoS y seguridad, o aquellos involucrados en otorgar o denegar acceso a la red al MN tal como AAA, son considerados servicios candidatos para la transferencia de contexto.

- **Handoffs.-** Es el proceso de re-encaminar las rutas de conexión de un nodo móvil conforme este se mueve desde un punto de acceso a otro, de manera tal que la

conexión sea preservada de manera transparente.

En la figura 19 podemos observar la manera en que se llevaría a cabo la transferencia de contexto[4], donde el pAR (previous Access Router) hace una petición de transferencia de contexto (CT, Context Transfer) al nAR. Si este acepta, se comenzará la transferencia de contexto durante la cual el nAR estará avisando al pAR sobre el estado de la transferencia. Una vez que el contexto se encuentre en el nAR, el host móvil pedirá que le sea entregado su contexto, para lo cual tendrá que identificarse como una medida de seguridad y el actual AR procederá a entregarselo una vez que esté seguro que ese contexto le pertenece al host móvil .

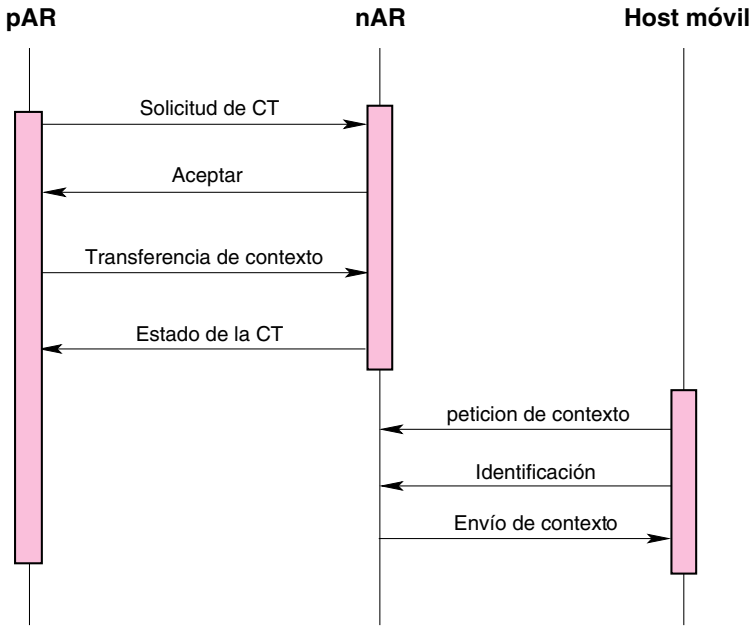

Figura 19. Transferencia de contexto.

IV.2.1 Motivaciones para la transferencia de contexto

Existen dos motivaciones básicas para la transferencia del contexto:

[4]Este diagrama no es una solución general y solo será visto como un medio para comprender mejor el problema de la transferencia de contexto.

1. La necesidad del rápido restablecimiento de servicios candidatos para la transferencia del contexto, sin que el host móvil tenga que desarrollar de nuevo todo el protocolo.

2. Proveer una solución interoperable, que trabaje para cualquier tecnología de acceso a la capa 2 por medio de radio.

Estos puntos son discutidos en mas detalle en las siguientes subsecciones.

IV.2.1.1 Rápido restablecimiento de servicios candidatos para la transferencia de contexto

Existe una variedad de servicios candidatos para la transferencia del contexto que pueden utilizar una solución de transferencia de contexto. En esta sección examinaremos tres servicios representativos:

- **Autentificación, autorización y contabilidad.-** Una de las aplicaciones urgentes de transferencia de contexto, es facilitar la re-autentificación del host móvil y restablecer la autorización de este mismo en una nueva subred, transfiriendo el contexto AAA (Authentication, Authorization and Accounting) del antiguo servidor AAA del host móvil al nuevo servidor AAA. Esto permite que el host móvil continúe el acceso a la nueva subred sin tener que rehacer un intercambio AAA con el nuevo servidor AAA de la subred. Naturalmente, es necesario mantener una asociación segura entre los servidores AAA para que la información de autentificación del host móvil pueda ser transferida de manera segura. Accounting es la función de registrar información, por ejemplo: el estado de una conexión, la cantidad de los datos a transferir, la duración de la conexión, etc.

- **Compresión de encabezado.-** La compresión es a menudo deseable porque reduce los requerimientos del ancho de banda y porque puede ser un aspecto

muy importante en medios inalámbricos de baja velocidad. Sin embargo, es problemático usar la compresión en la capa de enlace en algunas circunstancias, debido a que la mejor compresión se logra casi siempre en los protocolos de alto nivel, especialmente en la capa de aplicación. Desafortunadamente, intentar la compresión a dos diferentes niveles de protocolo es menos eficiente que desarrollarla a un solo nivel, ya que se corre el riesgo de que se incremente la cantidad de los datos que van a ser transmitidos.

- **Calidad de servicio (QoS).-** La calidad de servicio es un concepto que existe incluso en sistemas no computarizados para el envío de información, por ejemplo el caso del correo postal que ofrece servicios de entrega estándar y exprés. Este mismo concepto puede ser aplicado a redes de computadoras. Obviamente, los clientes pagan diferentes precios de acuerdo a la calidad del servicio que se les proporcione, tomando en cuenta que no todas las aplicaciones tienen los mismos requerimientos, y es necesario hacer un uso eficiente de los recursos disponibles con el objetivo de obtener la calidad de servicio que más se acople a sus necesidades. Los dos enfoques actuales de calidad de servicio son: servicios diferenciados (DiffServ, Differentiated Services) [Nichols *et al.*, 1998; Blake *et al.*, 1998; Black, 2000] y servicios integrados (IntServ, Integrated Services) [Baker *et al.*, 1996; Shenker *et al.*, 1997; Baker *et al.*, 1997]. Establecer niveles de QoS ayuda a las aplicaciones del MN a tener un mejor desempeño, es por ello que es necesario que estos niveles de QoS puedan ser transferidos hacia la nueva localización del MN en la red.

IV.2.1.2 Interoperatividad

Un tema de interés en la realización de *handoffs* suaves es que diferentes protocolos de radio de capa 2 pueden definir sus propias soluciones de transferencia de contexto. Estas soluciones son desarrolladas primeramente para permitir la transferencia de contexto

relacionado a la capa 2, el cual se transmite entre dos redes inalámbricas o dos puntos de acceso a través de una red alambrada IP. Si se adoptan los protocolos de capa 2 como una optimización a la solución de la transferencia de contexto de capa 3, entonces la movilidad transparente de los MNs con interfaces de red de capa 2 que soportan múltiples protocolos de radio puede ser difícil de lograr. Esto se debe a que se necesitaría un gateway o un traductor de protocolos de radio de capa 2, ya que de otra forma el MN requerirá de desarrollar la reinicialización de los servicios candidatos para la transferencia de contexto en su nueva subred, dado a que el flujo de información se realizará entre dos tecnologías diferentes . Por lo tanto, una solución general de capa 3 a la transferencia de contexto puede ser de utilidad a los protocolos de capa 2 que no definan su propia solución a la transferencia de contexto [Levkowetz *et al.*, 2002]. La consideración de este problema está fuera del alcance de esta tesis.

IV.2.2 Limitaciones en la transferencia del contexto

La transferencia del contexto no siempre será la mejor solución para restablecer los servicios candidatos para la transferencia de contexto en una nueva subred. Existen ciertas limitaciones que si llegan a ser eliminadas, harán útil la transferencia del contexto. Tales limitaciones son [Levkowetz *et al.*, 2002]:

- **Compatibilidad de enrutadores.-** La transferencia de contexto entre dos enrutadores es posible solo si el enrutador receptor soporta los mismos servicios candidatos de transferencia de contexto que el enrutador emisor. Esto no significa que los dos enrutadores tengan que ser idénticos en su implementación, ni tampoco implica que deban de tener las mismas capacidades. Un enrutador que no pueda manejar el contexto recibido debe rechazar la transferencia.

- **Requerimientos para reinicializar el servicio desde el principio.-** La principal motivación para el desarrollo de la transferencia del contexto asume que es

deseable el rápido restablecimiento (al mismo nivel) de los servicios candidatos en la nueva subred. Aun así, puede haber situaciones donde ambos, el dispositivo o la red de acceso, prefieran reestablecer o renegociar el nivel del servicio. Por ejemplo, cuando un host móvil ingresa a un dominio administrativo donde las políticas operacionales son diferentes, o cuando la negociación de un diferente nivel de servicio puede ser requerido.

- **Conveniencia para un servicio particular.-** Se asume que es más rápido establecer el servicio por medio de transferencia de contexto en lugar de comenzar todo de nuevo. Esto no es cierto para ciertos tipos de servicios, como por ejemplo el multicast.

IV.2.3 Consideraciones de desarrollo

El propósito de la transferencia de contexto es sustentar los servicios candidatos para la transferencia del contexto que son provistos al tráfico de un host móvil durante un *handoff*. Esto es esencialmente una mejora a la movilidad IP, lo cual lleva a una mejoría también en el funcionamiento del *handover*. Una solución de transferencia de contexto debe de proveer un desempeño que sea igual o mejor que reinicializar los servicios candidatos para la transferencia del contexto entre el host móvil y la red desde el principio. Los requerimientos generales para realizar una solución a la transferencia de contexto están descritos en [Syed *et al.*, 2003] dentro del grupo de trabajo de SeaMoby.

IV.2.4 Consideraciones de seguridad

Cualquier transferencia de contexto estándar debe proveer mecanismos de seguridad en la transferencia del contexto. Algunas consideraciones generales para la seguridad en la transferencia del contexto incluyen:

- **Privacidad de información:** El contexto puede contener información que el

usuario final o el operador de la red prefiera mantener oculta para usuarios no autorizados.

- **Legitimidad de transferencia:** Una transferencia de contexto falsa ó corrompida puede tener un severo impacto sobre la operación del enrutador receptor, y por consiguiente puede afectar la operación misma de la red. Las amenazas potenciales incluyen ataques de: denegación del servicio y robo de servicios.

- **Preservación de seguridad:** Parte de la transferencia del contexto puede incluir información pertinente a un contrato de seguridad establecido entre el nodo móvil y otra entidad en la red. Para que esta asociación de seguridad sea preservada durante un *handover*, la transferencia del contexto de seguridad debe incluir medidas apropiadas de seguridad.

Se espera que las medidas de seguridad usadas durante el transporte de la información entre los puntos en una red IP puedan ser suficientes para la transferencia del contexto. Sin embargo, y dadas las consideraciones mostradas anteriormente, existen razones para proveer medidas de seguridad adicionales que vayan mas allá de las soluciones propuestas por el IETF. Debido a que la transferencia de contexto requiere una asociación segura entre las entidades de la red, la seguridad de toda la red se puede ver comprometida si tan solo una de las entidades no cumple con este compromiso de seguridad. Cuando un MN se mueve de una entidad segura a otra insegura en presencia de una transferencia de contexto, la información incluída en la transferencia de contexto puede ser usada para descubrir los códigos de seguridad del canal de comunicaciones. Cuando un MN se mueve hacia una entidad de red insegura en ausencia de transferencia de contexto, la seguridad puede ser re-establecida en la nueva entidad. Sin embargo, lograr extender la seguridad a la transferencia de contexto y lograr que el contexto solo transite a través de los AR, dependerá de la seguridad con que cuenten los enrutadores.

Por lo tanto, la transferencia de contexto requiere de medidas novedosas de seguridad que excedan las capacidades de las soluciones existentes en el IETF [Levkowetz *et al.*, 2002].

IV.2.5 Protocolos propuestos para la transferencia de contexto

El grupo de trabajo de SeaMoby, especifica un protocolo general para la transferencia del contexto [Loughney *et al.*, 2003], cuyo funcionamiento está ilustrado en la figura 20.

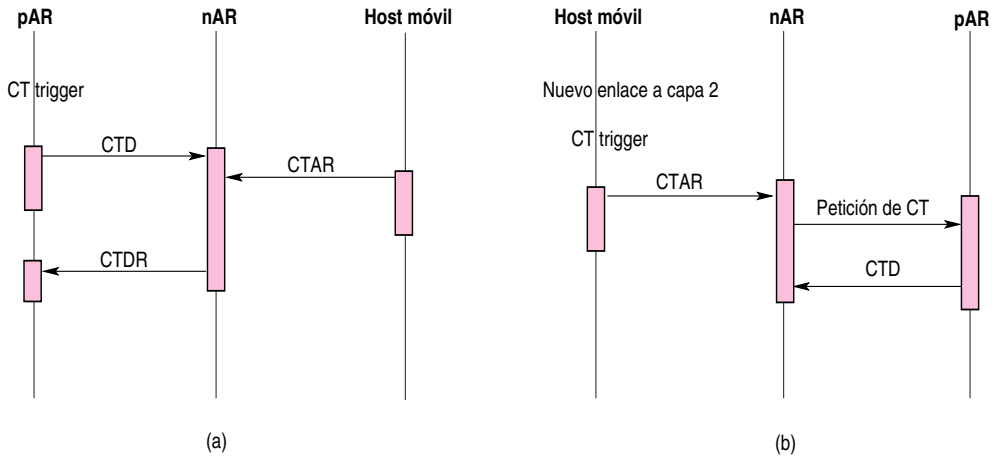

Figura 20. Protocolo para la transferencia de contexto

Este protocolo puede ser comenzado por una petición del MN (mobile controlled) o por una petición del nuevo o del previo enrutador de acceso (network controlled). En ambos casos el proceso de la transferencia de contexto comienza cuando un evento toma lugar, por ejemplo un *handover*. En la figura 20(a) se observa como un evento (trigger) es disparado en el pAR y se comienza el proceso de la transferencia del contexto. El pAR envia un mensaje CTD (Context Transfer Data) al nAR, este mensaje contiene la información referente al contexto de los servicios. Si el nAR acepta el mensaje significa que acepta la transferencia y aceptará el mensaje CTAR del MN, este mensaje contiene la dirección IP del nAR, la dirección IP de MN en el pAR, una lista del contexto a ser

transferido, y una clave con la cual el MN puede autentificarse y recibir su contexto. Por último el nAR envía un mensaje CTDR (CTD Request) al pAR para informarle del estado de la transferencia.

En la figura 20(b) se muestra el funcionamiento del protocolo cuando este es iniciado por el MN. En este caso el MN establece un enlace de capa 2 con el nAR, lo cual generará un trigger y hará que el MN envíe el mensaje CTAR al nAR. Cuando el nAR recibe la petición para comenzar el *handoff*, este envía una petición de transferencia de contexto al pAR, el cual por último enviará el contexto del MN en el mensaje CTD. Cabe señalar que este protocolo no ha definido medidas de seguridad, ni la manera en que el contexto será representado. Este protocolo simplemente define mensajes para pedir, iniciar y llevar el control de la transferencia del contexto.

Existen otras propuestas de protocolos para la transferencia de contexto, las cuales se han sometido como trabajos individuales al grupo de trabajo SeaMoby. Entre estos se encuentran: un protocolo para la transferencia de contexto como una extensión a Cellular IP [Georgiades *et al.*, 2003], y el protocolo TEXT [5] [Nakhjiri, 2003] el cual hace uso de agentes de movilidad, por lo que este protocolo no es compatible con HAWAII o Cellular IP. Ambos protocolos también tienen problemas de seguridad y ninguno de ellos especifica la representación del contexto a ser transferido. Es importante decir que las tres propuestas de protocolos para la transferencia de contexto presentadas en este capítulo se encuentran en nivel de borrador (draft) y no existe ninguna implementación de ellos.

IV.3 Conclusiones

Mobile IP permite a los nodos móviles realizar *handoffs* entre los ARs, los cuales actúan como puntos de conexión a la red IP. Sin embargo, en muchos escenarios el retardo

[5] Time Efficient context Transfer

y la pérdida de paquetes ocasionada por Mobile IP es muy alta. Por lo tanto, fue necesario crear protocolos de micromovilidad que se encargarán de realizar *handoffs* transparentes (poca latencia y poca pérdida de paquetes) entre los enrutadores. Sin embargo, estos protocolos asumen que el MN o el AR tienen un conocimiento previo del nuevo enrutador con el cual se realizará el *handoff*. De la misma manera, una vez que se realiza el *handoff*, los protocolos de micromovilidad no transmiten el estado de los servicios que el MN tenía en su antiguo punto de conexión, para que estos mismos servicios le puedan ser entregados al MN en su nuevo punto de conexión. Es por ello que en este capítulo se presentan las soluciones propuestas por el grupo de trabajo SeaMoby a estos dos problemas. Sin embargo, hasta ahora los protocolos propuestos para el descubrimiento de los CARs no poseen un algoritmo para la selección del TAR. En el caso de los protocolos para la transferencia de contexto, estos tampoco describen una manera de representar el contexto, además de que todavía se encuentra en discusión el protocolo (TCP/UDP) en el que se apoyará la implementación del protocolo para la transferencia del contexto. Otros aspectos importantes a señalar de estas dos propuestas es que tienen medidas pobres de seguridad, y aun no han sido implementadas. Así también ninguno de estos protocolos ha avanzado mas allá de un borrador (draft) de Internet.

Capítulo V

Una arquitectura para manejar la movilidad transparente

Como se ha observado hasta ahora, la movilidad dentro de un dominio es manejada por un protocolo de micromovilidad que permite al MN desarrollar *handoffs* entre las células de una red inalámbrica. Aunque varios protocolos de micromovilidad han sido propuestos y probados (algunos de ellos discutidos anteriormente), todos ellos asumen que un TAR ha sido seleccionado de entre un número de CARs. De la misma manera, cuando se desarrolla un *handoff*, sería conveniente (quizá obligatorio) que el MN encontrara en el TAR el mismo contexto que tenía en su anterior enrutador de acceso. Los problemas del descubrimiento de enrutadores de acceso y selección del TAR así como la transferencia de contexto, son problemas abiertos y están siendo abordados dentro del IETF por el grupo de trabajo SeaMoby. Este grupo ha liberado propuestas preliminares para la solución de estos problemas, pero hasta el momento estos no han avanzado mas allá de un borrador (draft) de Internet. En este capítulo se presentará una arquitectura para solucionar el problema de la movilidad transparente, la cual hace uso del proceso para el descubrimiento de los candidatos a enrutadores de acceso y la selección del mejor de entre ellos, así como de un protocolo encargado de realizar la transferencia de contexto entre los enrutadores de acceso.

V.1 Descripción general de la arquitectura propuesta

Una arquitectura de red es aquella que contiene un conjunto organizado de capas y protocolos. La especificación de esta arquitectura debe contener suficiente información que permita a un desarrollador escribir el código o construir el hardware para cada capa, de tal manera que se obedezca correctamente el protocolo apropiado [Tanenbaum, 1996]. Los elementos que conforman la arquitectura que aquí se propone son los siguientes: un MN el cual realizará *handoffs* entre las células de cobertura inalámbrica, enrutadores de acceso de entre los cuales se seleccionará un TAR, un repositorio en el cual se encontrarán almacenadas las políticas locales y la información descriptiva de las capacidades de los ARs, un proceso para el descubrimiento de los CARs y selección de un TAR así como un protocolo para la transferencia de contexto.

Los elementos de la arquitectura propuesta tienen diversas interacciones entre sí, las cuales se llevan a cabo cuando un MN desarrolla un *handoff* en el que se involucra el descubrimiento de los CARs y la transferencia de contexto (ver figura 21):

1. Un evento de capa 2 dispara la necesidad de un *handoff*. Algunos ejemplos de estos eventos pueden ser: establecer un enlace inalámbrico de capa 2 con el nuevo AR, la degradación de la calidad de la señal que el MN recibe de su AR actual o el MN recibiendo un anuncio (beacon) de un CAR. Se debe de tener en cuenta que hasta este punto el MN no puede establecer una comunicación de capa 3 con ninguno de los CARs, debido a que no conoce su dirección IP.

2. El proceso de descubrimiento de los CARs inicia y se selecciona un TAR. Si el MN recibe señales de capa 2 de diferentes ARs, este problema consistirá en determinar la dirección IP de los ARs y descubrir sus capacidades. Este proceso es conocido como descubrimiento de candidatos a enrutadores de acceso (CARD, Candidate

Access Router Discovery). Después de descubrir los posibles CARs un TAR debe ser seleccionado, lo que se logra a través de un algoritmo que seleccione el CAR que pueda cumplir con las preferencias del MN.

3. Inicia un *handoff* con el TAR. El firmware de la tarjeta de red realiza un *handoff* de capa 2 asociando al MN con el nuevo punto de acceso. Análogamente, el protocolo de micromovilidad tendrá que desarrollar un *handoff* de capa 3 con el TAR al cual se encuentra conectado el nuevo punto de acceso.

4. La transferencia de contexto es iniciada paralelamente con el desarrollo del *handoff*, con lo que se espera optimizar el rendimiento del *handoff*. En casos donde las capacidades de los CARs sean inadecuadas, la transferencia de contexto será inapropiada o incluso imposible. Sin embargo, si se selecciona un TAR entre los CARs y sus capacidades permiten la transferencia de contexto, esta se llevará a cabo.

5. El *handoff* es completado. El protocolo de micromovilidad debe terminar el *handoff* con o sin la optimización del rendimiento que provee la transferencia de contexto.

En las siguientes secciones del capítulo sólo se explicarán detalladamente los pasos 2 y 4 de la arquitectura. Los pasos 1, 3 y 5 serán llevados a cabo por un protocolo de micromovilidad, y sólo se explicarán brevemente.

En la figura 21 se ejemplifican las interacciones de los diferentes elementos de la arquitectura propuesta, las cuales pueden resumirse en los siguientes 3 pasos:

- En el primer paso, el MN almacena en una lista los beacons de los AR que pueda escuchar, la cual será enviada después a su AR actual.

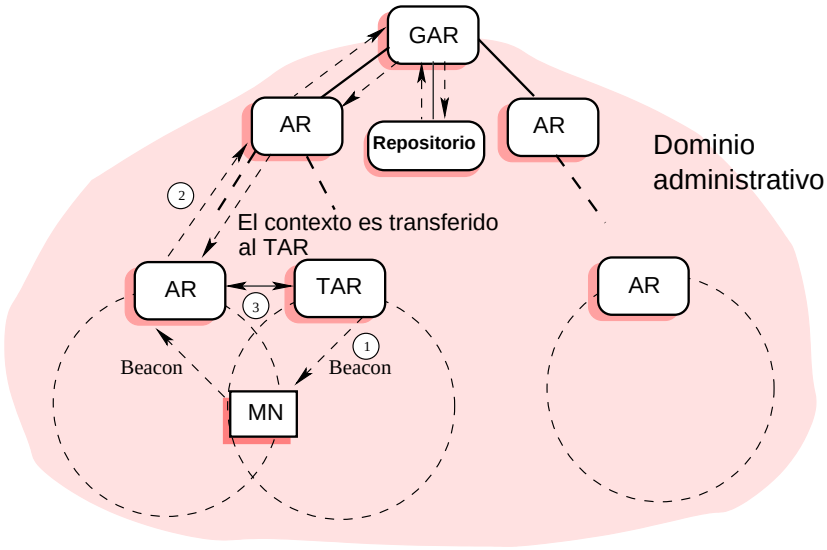

Figura 21. Movilidad con descubrimiento de CARs y transferencia de contexto.

- En el segundo paso, el AR actual se encargará de mapear la dirección MAC de los enrutadores a su dirección IP, así como de descubrir sus capacidades y seleccionar el TAR. El AR actual debe enviar la dirección IP del TAR al MN para que comience un *handoff* con éste.

- Por último, paralelo al desarrollo del *handoff* se realizará la transferencia del contexto hacia el TAR y se finaliza el *handoff*.

V.2 Descubrimiento del CAR y selección del TAR

El protocolo para el descubrimiento de los CAR (CARD) y selección del TAR debe contemplar las funcionalidades de: traducción de direcciones en reversa (reverse address translation) y el descubrimiento de las capacidades de los CAR, como se especificó en la sección IV.1.1. El descubrimiento de las capacidades de los ARs puede ser de gran ayuda al momento de realizar las decisiones de *handoff*, por lo tanto, para poder representar de manera adecuada estas capacidades, aquí se propone el uso de conceptos

relacionados a PBNM (Policy Based Network Management, ver apéndice A) [Verma, 2002]. Particularmente se apoya en el trabajo del IETF Policy Framework Working Group[1], el cual propone un modelo orientado a objetos [Moore *et al.*, 2001] para representar la información de las políticas. Este modelo define una jerarquía de dos niveles para representar las clases de objetos, las cuales son:

1. Clases estructurales, encargadas de representar la información y control de las políticas, y

2. Clases de asociación, encargadas de indicar cuantas instancias de las clases estructurales se encuentran interrelacionadas entre ellas.

Este grupo de trabajo (policy) sugiere el uso de LDAP [Hodges y Morgan, 2002] como el protocolo para acceder a las bases de datos de las políticas [Strassner *et al.*, 2002; DMTF, 2002] y también define un conjunto de clases y atributos genéricos para la representación de políticas. Aquí se retoman dichos conceptos de manera que las políticas se almacenarán en un repositorio LDAP, el cual también almacenará la información descriptiva de los diferentes dispositivos (ARs, APs, MNs) de red; esta información estará representada a través de esquemas LDAP. Así pues, el siguiente constituye un ejemplo de un esquema que representa a un AR:

```
dn: cn=router1, ou=Recursos, dc=cicese, dc=mx
ipHostNumber:158.97.22.201
linkQuality: 7
channelNumber: 11
essid: WLAN
troughput: 456
ipNetmaskNumber: 255.255.252.0
objectClass: top
objectClass: pcimPolicy
objectClass: ipNetwork
```

[1]http://www.ietf.org/html.charters/policy-charter.html

```
objectClass: ipHost
objectClass: device
objectClass: ieee802Device
headerCompresion: 0
numberNodes: 2
cn: router1
description: Este es el enrutador 1
LinkEncap: ethernet
macAddress: 00:20:E0:6E:DC:CD
ipNetworkNumber:158.97.22.201
typeQoS: 111
QoS: DiffServ
```

Para poder representar este esquema en LDAP es necesario apoyarse en los diferentes esquemas que este incluye, por ejemplo: los tipos de datos ipHostNumber, ipNetmaskNumber, ipNetworkNumber, etc. están descritos en el esquema *nis* (network information service) descrito en el RFC 2307 [Howard, 1998]. Estos datos ayudarán a identificar los atributos de red de los dispositivos (en este caso un AR). Debido a que no existe un esquema para representar las capacidades de los ARs, fue necesaria la creación de un esquema propio, al cual se llamó "PolicyCorewin" en el cual se encuentran las variables que representa las capacidades de los ARs. Estas variables son: troughput, linkQuality, numberNodes, headerCompresion, QoS y typeQoS. El esquema PolicyCorewin también define otras variables, las cuales se tratarán mas adelante.

De esta manera se puede observar que el esquema descrito arriba puede representar los atributos comunes del AR (dirección IP, dirección MAC, etc.), así como también puede representar las capacidades de éste (ancho de banda, calidad de la señal, número de nodos, compresión de cabeceras y calidad de servicio). Debe aclararse que en el desarrollo de esta arquitectura se asumirán dos tipos de QoS: IntServ y DiffServ [Baker *et al.*, 1996; Cisco, 2001]. Cada tipo de QoS tiene sus propias divisiones, las cuales estarán representadas en el vector binario typeQoS. Los tres valores para DiffServ son:

```
typeQoS[0] = servicio garantizado.
typeQoS[1] = servicio asegurado.
typeQoS[2] = mejor esfuerzo.
```

Y los dos valores de IntServ son:

```
typeQoS[0] = servicio garantizado.
typeQoS[1]= servicio de carga controlada.
```

De manera similar, los MNs pueden ser representados en un esquema LDAP, donde además de describir sus atributos de red (ipHostNumber, ipNetworkNumber, etc.), también se representan las variables utilizadas para conectarse a los APs (channelNumber y essid) así como las variables que representan sus preferencias (selCriteria, TypeOfService, QoS y typeQoS). Estos dos últimos grupos de tipos de variables también fueron definidos en el esquema PolicyCorewin debido a que tampoco existe un esquema para representar las preferencias de los MNs. El uso de selCriteria y TypeOfService será descrito posteriormente. Un ejemplo del esquema de un MN es el siguiente:

```
dn: cn=mncomp2, ou=Recursos, dc=cicese, dc=mx
linkQuality: 7
ipHostNumber: 158.97.22.202
channelNumber: 8
essid: WLAN
troughput: 500
objectClass: pcimPolicy
objectClass: ipNetwork
objectClass: ipHost
objectClass: top
objectClass: device
cn: mn2
description: nodo movil 1
typeQoS: 01
ipNetworkNumber: 158.97.22.221
TypeOfService: gold
QoS: IntServ
selCriteria: 11110
```

Por último, un ejemplo del esquema de un AP es el siguiente:

```
dn: cn=accesspoint3, ou=Recursos, dc=cicese, dc=mx
ipHostNumber: 158.97.22.202
linkQuality: 8
channelNumber: 3
essid: WLAN
troughput: 800
ipNetmaskNumber: 255.255.252.0
objectClass: top
objectClass: pcimPolicy
objectClass: ipNetwork
objectClass: ipHost
objectClass: device
objectClass: ieee802Device
numberNodes: 3
cn: punto de acceso 3
macAddress: 00:20:E0:6E:AC:EF
description: Este es el punto de acceso 3
ipNetworkNumber: 158.97.22.232
```

En el esquema de un AP se utilizarán las variables macAddress e ipHostNumber (definidas en *nis*) para mapear a cual AR se encuentra conectado el AP. Las variables ipNetworkNumber, essid (Extended Service Set Identifier) y channelNumber son utilizadas para que el MN conozca los elementos necesarios para conectarse con el nuevo AP. El procedimiento a través del cual el MN obtiene estos valores será descrito posteriormente.

La manera en que se utilizarán estos esquemas se representa en la figura 22, donde cada AR, AP y MN que sea ingresado ó dado de alta en la red deberá de enviar su esquema al repositorio que se encuentra en el GAR (Gateway Access Router).

Además, los ARs serán los encargados de informar al repositorio sobre el estado actual de la célula que ellos controlan, y dado que esta actualización se hará periodicamente[2], el repositorio siempre tendrá un panorama actual del estado de la red. La información del repositorio será difundida periódicamente en la jerarquía hacia los ARs

[2]El tiempo de las actualizaciones será configurado por el administrador de la red y será idealmente de cada pocos segundos.

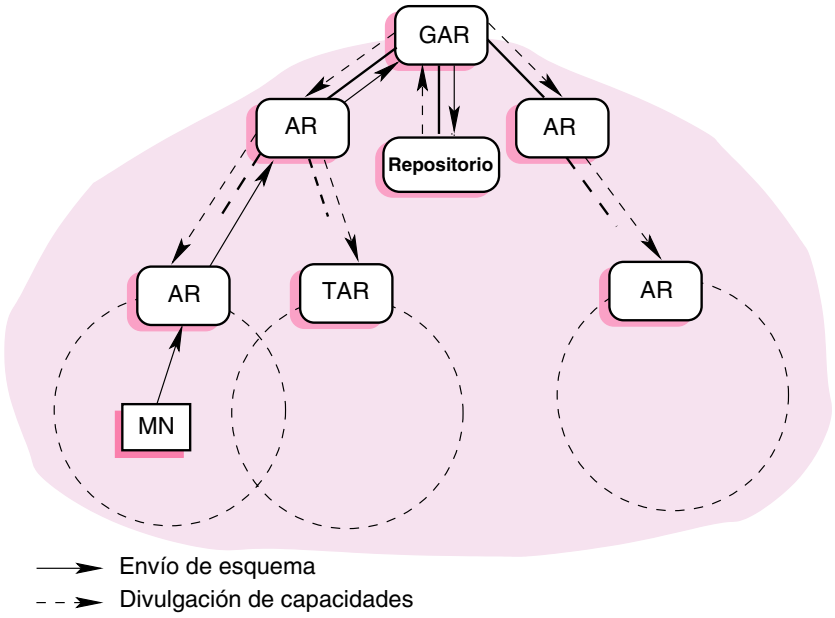

Figura 22. Ingreso de dispositivos y divulgación de recursos.

y será almacenada en el cache (ó repositorio local) de los ARs. Por lo tanto, cuando un TAR deba ser seleccionado, el antiguo AR utilizará la información almacenada en su cache para descubrir las capacidades de los ARs. Sin embargo, si el tiempo de vida asociado a estas capacidades es muy corto, el antiguo AR solicitará esta información al repositorio y con base en la información actualizada ejecutará un algoritmo para la selección del TAR. Este algoritmo debe estar basado en las capacidades de los ARs, las preferencias del MN y una política local. Un ejemplo de una de estas políticas es el siguiente:

```
policyRule selectTAR{
     selCriteria:11001
     TypeOfService: silver
     policyAction: setTar
}
policyAction setTar{
     connect: Allowed
}
```

Estas políticas deben ser creadas por el administrador de la red y almacenadas en el repositorio localizado en el GAR. El funcionamiento de las políticas es parecido a un enunciado IF donde, si la sección policyRule se cumple, entonces deberá ejecutarse lo que se encuentra en la sección policyAction. La política anterior indica que si un CAR es encontrado y satisface los valores descritos en la sección policyRule, entonces el CAR debe ser seleccionado como TAR, tal y como lo indica la sección policyAction. Por consiguiente, el MN desarrollará un *handoff* con el TAR para que éste se convierta en su nuevo AR. Esta política hace uso de un campo para la selección de los criterios (selCriteria, también encontrado en el esquema del MN), el cual servirá para determinar cual de los CARs es el apropiado para ser seleccionado como TAR.

Este criterio (selCriteria) es un vector binario donde cada bit representa uno de los criterios a considerar en la selección del TAR. Si el bit es 1, entonces ese criterio será tomado en consideración en el algoritmo; de lo contrario, significa que el MN no tiene interés en encontrar ese servicio en su nuevo AR. El orden de los servicios (en el vector selCriteria) en los que el MN puede establecer sus preferencias son:

```
selCriteria[0] = Ancho de banda.
selCriteria[1] = Calidad de la señal.
selCriteria[2] = Número de nodos en la célula.
selCriteria[3] = Compresión de cabeceras.
selCriteria[4] = Calidad de servicio.
```

En el ejemplo de la política mostrada anteriormente, puede observarse que en este caso el algoritmo no debe tomar en cuenta el número de nodos en la célula ni la compresión de cabeceras, ya que los atributos de ancho de banda, calidad de la señal y QoS serán los que determinen cuál de los CARs se convertirá en el TAR.

El valor de los primeros tres parámetros (ancho de banda, calidad de la señal y número de nodos) dependerá del valor asignado al campo TypeOfService. En este caso es “plata” (silver), lo que para el valor de ancho de banda representaría un valor que

se encuentra dentro del rango de 500 Kbps a 1Mbps. Debe tenerse en cuenta que los rangos de los valores del campo TypeOfService son configurados por el administrador de la red, por lo que éste puede especificar el funcionamiento de la red de acuerdo a sus necesidades. Los tres valores disponibles para TypeOfService, los cuales fueron definidos aquí son: oro (gold), plata (silver) y bronce (bronze). Por consiguiente, se puede decir que esta política es usada como un parámetro de entrada en el algoritmo que realiza la selección del TAR. Para ejemplificar más detalladamente la manera en que se utiliza el campo TypeOfService para especificar el valor de las capacidades de los ARs, la tabla II muestra los valores que fueron definidos al momento de realizar las pruebas del algoritmo del CARD, las cuales serán descritas posteriormente.

Tabla II. Valores de las capacidades asignados por TypeOfService.

Atributo	**Bronze**	**Silver**	**Gold**
Troughput	> 0	> 500	> 1000
LinkQuality	> 0	> 5	> 8
NumberNodes	≤ 7	≤ 4	≤ 2

Debe tenerse en cuenta que el nuevo AR puede no ser el que mejores capacidades tenga, sino aquel cuyas capacidades satisfagan la política establecida y por ende las preferencias del MN. Si ninguno de los CARs puede satisfacer los requerimientos del MN, entonces se enviará un mensaje al MN indicando que la transferencia de contexto no puede ser llevada a cabo. Otra posibilidad puede ser el negociar con uno o más CARs para saber si alguno de ellos puede satisfacer algunos de los requerimientos del MN, pero este proceso puede necesitar una gran cantidad de tiempo y por consiguiente, la transferencia de contexto perdería su mayor ventaja que es la realización de un *handoff* mas eficiente. Es por ello que en este caso sería mejor restablecer los servicios desde el principio con el nuevo AR.

En la figura 23 se observa un escenario en el que se muestra la interacción de las entidades involucradas en la realización de un *handoff* con descubrimiento de CARs y transferencia de contexto, tal y como se mostró en la arquitectura propuesta.

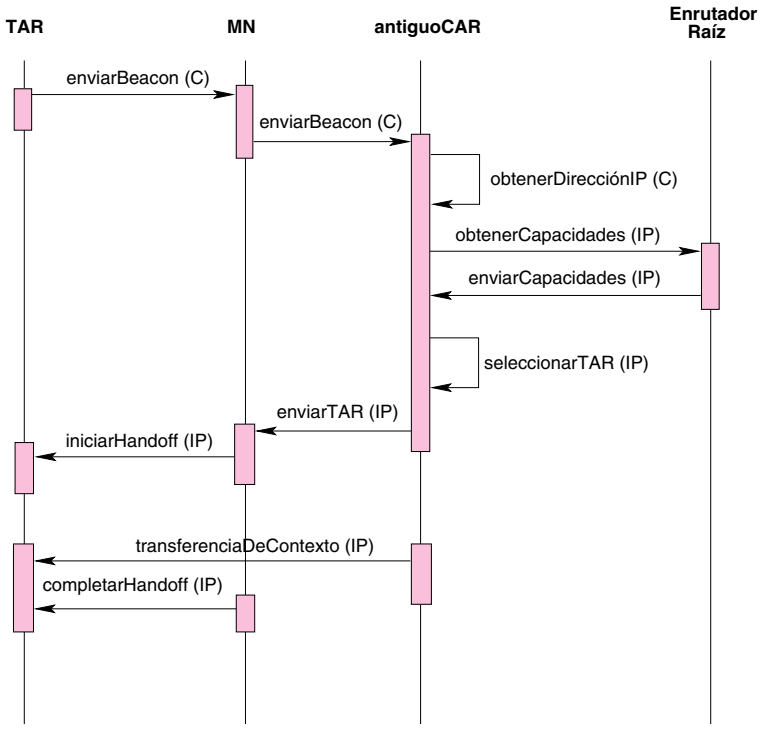

Figura 23. Interacciones en la movilidad transparente.

Cuando el MN se mueve hacia las periferias de su actual célula inalámbrica, un evento hace que el MN comience a escuchar y guardar en una lista los anuncios (beacons) de los demás APs geográficamente adyacentes a él. Estos anuncios contienen la dirección (MAC) de capa 2 de los AP con los que el MN puede conectarse. El MN enviará a su AR actual sus preferencias junto con la lista de beacons solicitándole que encuentre, por cada AP, la dirección IP del AR asociado con el AP, además de tener que descubrir la capacidades de estos ARs.

Estas capacidades pueden ser obtenidas del cache del antiguo AR, aunque estas capacidades tienen asignado un tiempo de vida y si éste es demasiado corto se tendrá que

hacer una solicitud al repositorio localizado en el GAR para que envíe las capacidades actuales de los CARs que el MN puede escuchar. Con base en estas capacidades, las preferencias del MN y una política local, el antiguo AR realizará la selección del TAR. Este proceso se hará de la siguiente manera: la política local será la encargada de verificar cual de los CARs (con base en las capacidades de éste) puede satisfacer las preferencias del MN. Si un TAR es seleccionado, entonces el AR actual debe enviar al MN la información necesaria del AP conectado al TAR seleccionado para que éste pueda desarrollar un *handoff* con él. Esta información es: dirección IP, channelNumber y essid (Extended Service Set Identifier). Paralelo a la realización del *handoff* se realizará la transferencia del contexto, de esta manera el MN encontrará en el TAR los servicios que le proveía su antiguo AR. Por último, el protocolo de micromovilidad se encargará de terminar el *handoff*, y de esta manera establecer las comunicaciones entre el MN y el TAR. Las eventuales implicaciones de seguridad que pudieran surgir durante el desarrollo de este proceso están fuera del alcance de este trabajo.

V.2.1 Desempeño del algoritmo

Para probar el desempeño del algoritmo para el descubrimiento de los CARs y selección de un TAR propuesto aquí, fue necesario tomar su tiempo de ejecución bajo diferentes escenarios. Primeramente se describirán las condiciones de la prueba: una computadora Laptop Pentium III a 1GHz operando con RedHat 8.0 Linux jugó el rol del MN, mientras que el repositorio se encontraba en una computadora de escritorio Pentium II a 300 MHz operando con Mandrake 9.0 Linux, la cual también jugaba el rol del GAR. Las condiciones de la red fueron las siguientes: 2/3 de la red estaba configurada a 100 Mbps mientras que el 1/3 restante estaba configurado a 10 Mbps (ver figura 24). El punto de acceso utilizado es marca Belkin con tecnología 802.11b a 11 Mbps, mientras que el MN tenía una tarjeta inalámbrica D-Link modelo DWL-650H con tecnología 802.11b a

11Mbps.

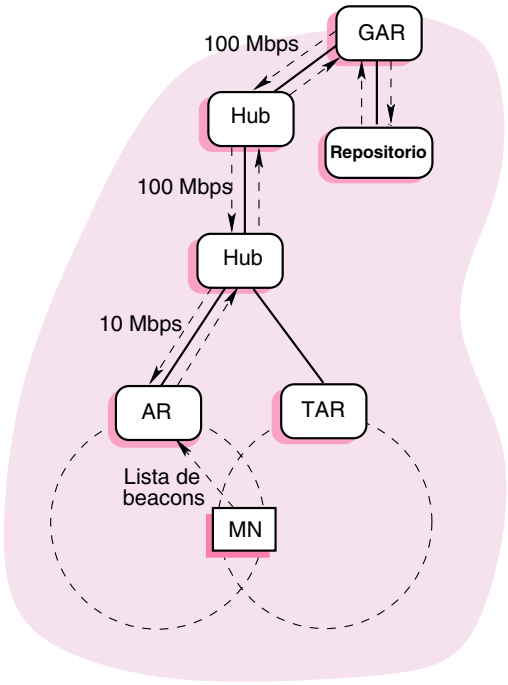

Figura 24. Condiciones de la red en la prueba del CARD.

El funcionamiento del programa encargado del descubrimiento de los CARs y selección del TAR es el siguiente: el MN debe enviar a su AR actual una lista de las direcciones MAC de los APs que puede escuchar; el AR actual se encargará de mapear la dirección MAC de los APs a la dirección IP de los enrutadores a los cuales se encuentran conectados. Una vez identificada la dirección IP de los ARs se analizarán sus capacidades almacenadas en el cache del AR actual y con base en ellas se ejecuta el algoritmo para la selección del TAR. Este proceso es realizado por el AR actual, dado que es ahí donde reside este programa. El peor caso del algoritmo se llevará a cabo cuando las capacidades de los CARs almacenadas en el cache del AR actual tenga un tiempo de vida demasiado corto, por lo cual será necesario solicitar estas capacidades

al repositorio localizado en el GAR (ver figura 24). Lo anterior introducirá mas retardo al desempeño del algoritmo debido a que se requiere de cierto tiempo para que el GAR procese la petición, busque las capacidades de los CARs y las envíe al AR actual. Además debe tomarse en cuenta la congestión de la red, lo cual también puede ocasionar que el desempeño del algoritmo se vea afectado.

El tiempo de ejecución del programa fue tomado con el comando *time*, el cual se encuentra incluido en los sistemas UNIX y derivados. En las pruebas se midió el tiempo de ejecución para el peor de los casos, el cual incluye las siguientes tres acciones: el AR actual recibe la lista de los APs que el MN puede escuchar y mapea la dirección MAC de los APs a la dirección IP de los ARs asociados a estos. Una vez identificados los CARs verifica sus capacidades, sin embargo, debido a que el tiempo de vida asociado a éstas es muy corto solicita estas capacidades al repositorio LDAP. Con las capacidades de los CARs actualizadas, el AR actual realiza el algoritmo de la selección del TAR.

Este tiempo es de 0.1 segundos, por lo tanto se espera que el proceso de descubrimiento de CARs y selección del TAR no introduzca retardo extra al desempeño de un *handoff*. Las pruebas descritas arriba también fueron realizadas conectando el MN a la red a través de un cable (conexión no inalámbrica), esto con el fin de comparar el desempeño del algoritmo en el peor caso en una red inalámbrica contra el de una red alambrada. El tiempo arrojado en la prueba de la red alambrada fue de 0.09 segundos, con lo cual se observa que las tecnologías inalámbricas no introducen mucho retardo extra al desempeño del algoritmo.

V.3 Transferencia de contexto

Una vez que se ha seleccionado el TAR hacia el cual se realizará el *handoff*, es necesario especificar cómo es que se va a llevar a cabo la transferencia de contexto (CT, Context Transfer). Aquí se propone un protocolo para la transferencia de contexto

(CTP, Context Transfer Protocol) cuyos objetivos serán la reducción de latencia, minimización de pérdida de paquetes y evitar reiniciar la señalización para establecer los servicios desde el principio. El restablecer los servicios del MN en el nuevo AR es un proceso necesario para que el MN pueda contar con tales servicios después de un *hand-off*, sin embargo, este proceso puede consumir mucho tiempo y puede ocasionar que la latencia sea percibida por el usuario. Si la transferencia de contexto se realiza de manera correcta, el retardo asociado con el restablecimiento de los servicios no afectará el periodo de disrupción del servicio. Por lo tanto, la transferencia de contexto ofrece un mejor soporte a los nodos basados en movilidad debido a que las aplicaciones ejecutándose en ellos pueden operar con mínima interrupción. Cuando un MN se mueve a un AR diferente, el reacomodo de su contexto provee varios beneficios importantes, tales como:

- **Operación transparente en el flujo de las aplicaciones**, debido a que el MN no necesita restablecer su contexto en el nuevo AR.
- **Ahorro de ancho de banda**, debido a que se puede evitar restablecer múltiples contextos a través de un enlace de baja velocidad (inalámbrico) reacomodando el contexto a través de un enlace de alta velocidad (red alambrada).
- **Reducir la susceptibilidad a errores**, dado que la mayor parte del protocolo opera sobre redes confiables (alambradas), reemplazando las operaciones a través de enlaces propensos a errores (inalámbricos).

En el RFC3374 [Levkowetz *et al.*, 2002] se muestra una descripción detallada de la motivación, necesidades y beneficios de la transferencia de contexto. Cualquier solución a este problema debe incluir lo siguiente:

- Representación de las características del contexto.

- Mensajes para iniciar y autorizar la transferencia de contexto.
- Mensajes para transferir el contexto antes, durante y después de un *handover*.

El protocolo que aquí se propone deberá trabajar con otros protocolos[3] en beneficio de lograr una movilidad transparente. Por lo tanto, para lograr una movilidad transparente es necesario evitar las interrupciones (disruptions) en el desempeño de las aplicaciones. Para resolver este problema algunos protocolos de micromovilidad utilizan buffers para bifurcar el flujo de paquetes del MN hacia los dos ARs, el antiguo y el nuevo (ver figura 25a), por lo que el MN puede seguir recibiendo su flujo de paquetes.

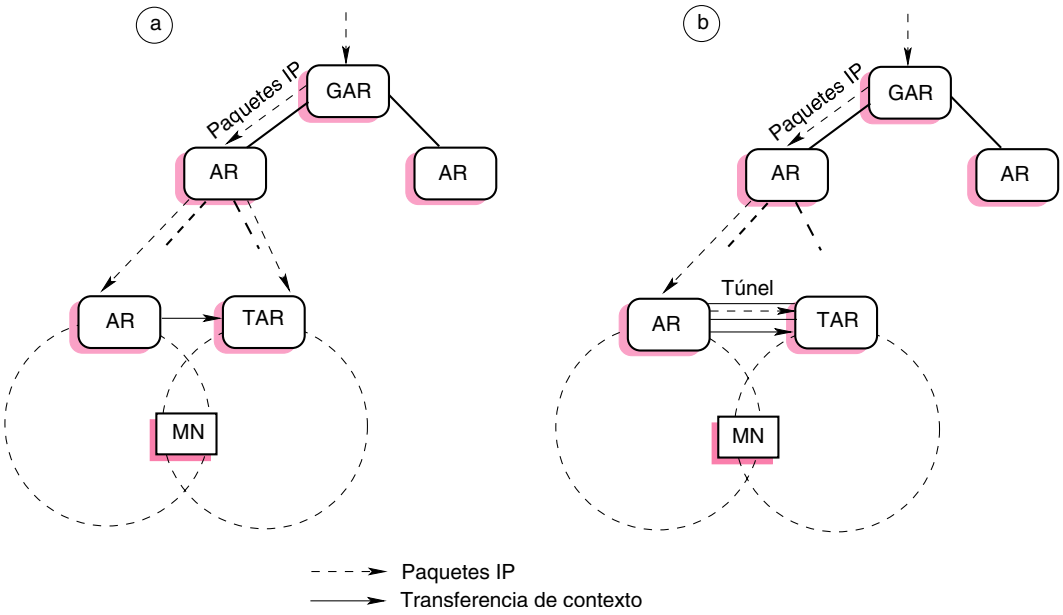

Figura 25. Túnel para evitar interrupciones durante un *handoff*.

Sin embargo, si el protocolo de micromovilidad no realiza esta función, proponemos que el antiguo AR establezca un túnel con el nuevo AR y que el tráfico del MN fluya

[3]IP, TCP, UDP, protocolo de micromovilidad, etc.

a través de este túnel, con lo cual se evitarían las posibles interrupciones en las aplicaciones debido a la pérdida de paquetes (ver figura 25b). La transferencia de contexto no necesita pasar a través de este túnel, no obstante, este túnel debe permanecer hasta que se complete la transferencia de contexto.

Durante el establecimiento del túnel el nuevo AR reenvía el tráfico del MN sin tener un conocimiento detallado de los servicios del MN o de su información de enrutamiento; el registro del MN con el nuevo AR es pospuesto hasta que el MN lo crea más conveniente. Sin embargo, esta solución solo ayuda a remover la latencia involucrada en establecer un enrutamiento en capa 3 de los paquetes fuera de la ruta crítica del *handover*, y cualquier retardo asociado con el restablecimiento de las características de los servicios podrá ocasionar interrupciones en los servicios. Durante el reenvío de los paquetes a través del túnel, el AR antiguo se encargará de procesar la información asociada al contexto del MN, o dicho de otra forma, mientras el túnel esté establecido, el AR antiguo se encargará de manejar el contexto del MN. De esta manera, el MN no sólo recibe sus paquetes, sino que también sus servicios son procesados por el AR antiguo sin interrupción. Una vez que la transferencia de contexto se considere completa, el AR nuevo simplemente tomará el control del procesamiento de los servicios del MN.

Mantener el contexto en el AR actual durante un *handoff* tiene sus ventajas, como por ejemplo: si un MN se desplaza y comienza a realizar un *handoff* con un nuevo AR, pero antes de terminar la transferencia de contexto el MN continúa moviéndose, y ahora empieza a realizar un *handoff* con un segundo AR (ver figura 26). Así el AR antiguo termina las comunicaciones con el primer AR y comenzará el proceso del restablecimiento del contexto del MN con el segundo AR. De esta manera, el MN puede seguir recibiendo el tráfico dirigido hacia él, y además se tiene la ventaja de que este tráfico sigue siendo procesado por el AR antiguo debido a que no ha desechado el contexto del MN. En otro escenario se puede considerar que en vez de que el MN busque otro AR,

éste regrese a su AR antiguo, donde su AR antiguo seguirá manteniendo el contexto del MN, por lo que no se requerirá de ningún otro proceso.

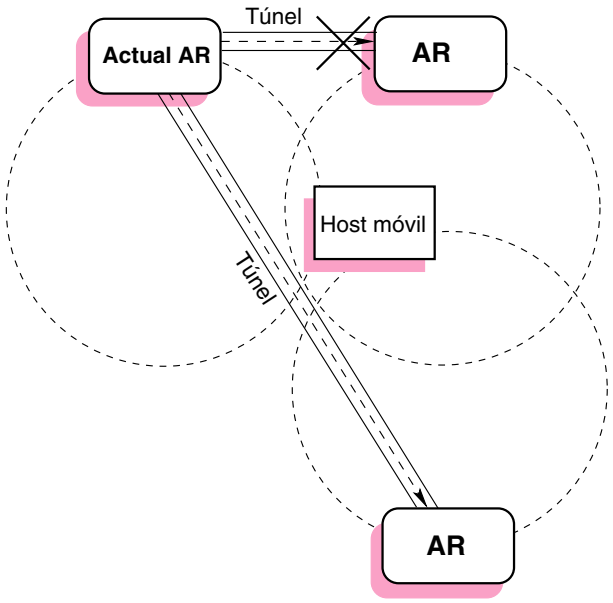

Figura 26. Establecimiento de un túnel con un tercer AR.

Por último, se debe tomar en cuenta que una vez que el MN ha obtenido un nuevo AR, el AR antiguo debe desechar el contexto asociado a ese MN. Sin embargo, para evitar el problema mencionado anteriormente proponemos la especificación de un umbral. De esta manera cuando el MN traspase este umbral y su contexto haya sido transferido a un nuevo AR, el AR antiguo deberá desechar el contexto del MN (ver figura 27).

En la figura 28 se puede observar como se lleva a cabo el intercambio de paquetes durante un *handoff*, donde primeramente el AR antiguo envía una petición al nuevo AR para comenzar la transferencia de contexto, si esta petición es aceptada la transferencia será llevada a cabo. Durante el proceso del *handoff* los caminos de enrutamiento serán actualizados para que los paquetes del MN sean dirigidos hacia el TAR, así también

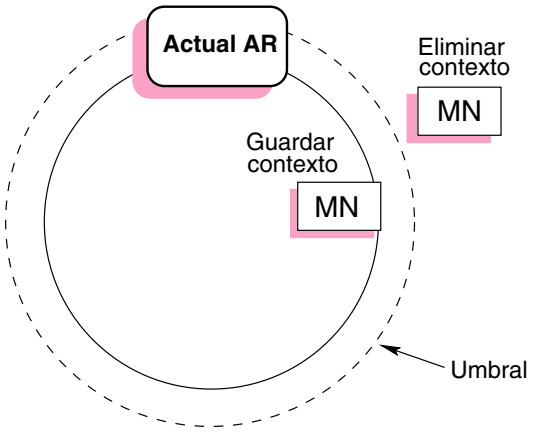

Figura 27. Umbral para eliminar el contexto de un MN.

el MN establecerá comunicaciones con el TAR. Una vez que el MN pueda tener comunicación a capa 3 con el nuevo AR, este tendrá que autentificarse para poder recibir su contexto. Esta autentificación sirve como un medio de seguridad para evitar que nodos maliciosos soliciten el contexto de algún otro MN. Una vez terminado el *handoff* el nuevo AR se encargará de procesar el contexto del MN.

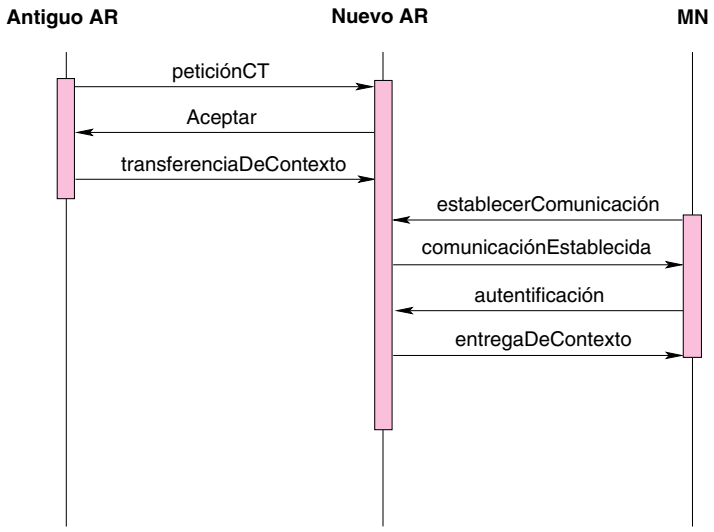

Figura 28. Diagrama de secuencia del protocolo CTP.

V.3.1 Formato de los mensajes del protocolo CTP

El formato general de los paquetes propuestos aquí consistirá de una cabecera IP, una cabecera del protocolo CTP, la carga útil del mensaje y un campo de autentificación. En la figura 29 podemos observar el formato genérico de los paquetes[4] del protocolo CTP propuesto en este trabajo.

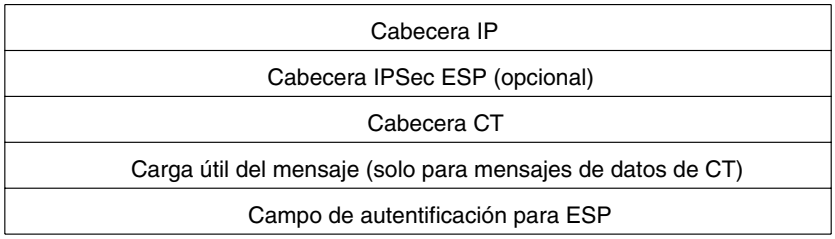

Figura 29. Formato genérico de los paquetes para la transferencia de contexto.

El tamaño total del paquete incluyendo las cabeceras del protocolo de transporte debe ser menor que el MTU (Maximum Transmisión Unit) para evitar la fragmentación de los paquetes. La descripción de cada uno de los campos del paquete genérico CTP es la siguiente:

Cabecera IP.- Los datos del contexto son enviados usando paquetes IP. En caso que el tráfico en la red pueda ser manejado con esquemas de calidad de servicio, los paquetes IP que llevan el contexto deben ser marcados para que sean tratados con más que "mejor esfuerzo", para que de esta manera se tenga mayor robustez en la transferencia del contexto.

Cabecera IPSec ESP (opcional).- En caso de que sea necesaria la encriptación y/o

[4]Las figuras de los paquetes solo proveen el esquema del formato y no los tamaños de los campos.

autentificación de la carga útil (payload), se puede agregar una cabecera ESP (Encapsulating Security Payload) para agregar seguridad a la transferencia del contexto.

Cabecera CT (Transfer Context).- Esta cabecera consiste en la información común a los mensajes de transferencia de contexto, ya sea de datos o de señalización (ver figura 30). El propósito de cada uno de los campos de esta cabecera es el siguiente:

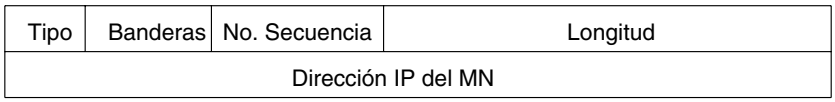

Figura 30. Cabecera CT.

Tipo.- Es el tipo de mensaje en la transferencia de contexto y es definido de la siguiente manera:

0 Mensaje de petición de inicio de CT.

1 - 2 Reservado.

3 Mensaje de datos de CT.

4 Reservado.

5 Mensaje PNACK CT (opcional).

6 Mensaje FNACK CT (opcional).

7 Mensaje abortar CT.

8 Mensaje abortar característica de CT (opcional).

9 Reservado.

Banderas.- Este espacio puede ser utilizado para establecer banderas, ya sea de confiabilidad en la transmisión o para indicar si el paquete es una retransmisión, actualización, etc.

Número de Secuencia.- En caso de que los paquetes del contexto sean transmitidos en varios paquetes o fases, este campo puede ayudar al nuevo AR a ordenar los paquetes y/o a detectar pérdida de paquetes. El número de secuencia para el primer paquete es 0. Retransmisiones o actualizaciones de los datos del contexto deben usar el mismo número de secuencia.

Longitud.- Longitud de la carga útil del mensaje en bytes.

Dirección IP del MN.- Este campo es usado para identificar el dueño del contexto, de esta manera el nuevo AR entrega el contexto al MN que tenía anteriormente esa dirección IP.

Carga útil del mensaje.- Este mensaje utilizará la cabecera mostrada en la figura 30 cuyo campo tipo será igual a 3 para mensaje de datos. La carga útil del paquete contiene datos del contexto en forma de opciones (uno por cada servicio) como se muestra en la figura 31. Cada una de las opciones de los datos es construida como se muestra en la figura 32, y la descripción de cada uno de éstos es la siguiente:

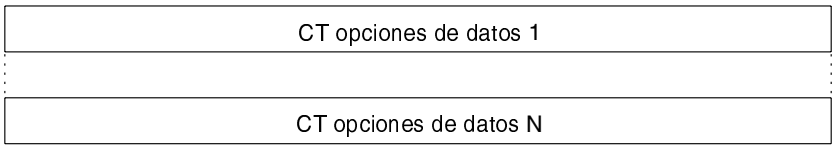

Figura 31. Mensajes de datos en CT.

Identificador de microflujo.- Este campo debe contener el identificador de un microflujo dado. Un microflujo [Blake *et al.*, 1998] es una instancia

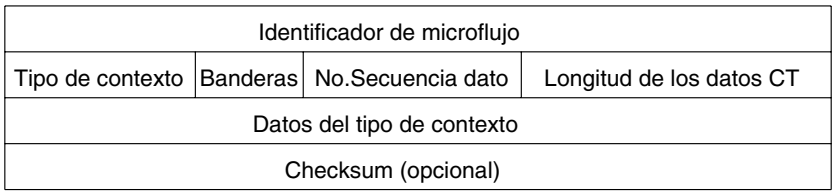

Figura 32. Aspectos de los datos CT.

de un flujo de una aplicación-a-aplicación, el cual es identificado por una dirección fuente, un puerto fuente, una dirección destino, un puerto destino y un identificador de protocolo.

Tipo de contexto.- Identifica el servicio al cual corresponde ese contexto, por ejemplo: compresión de cabeceras, QoS, etc.

Banderas.- Este campo es utilizado para establecer banderas, las cuales pueden identificar si se requiere el uso de confiabilidad, si existen otros aspectos del contexto en el paquete o si ya es el último, actualización de contexto, etc.

No. Secuencia dato.- Es usado en caso de que el contexto sea enviado en varios datagramas. Este número de secuencia es iniciado en 0 para cada uno de los servicios transferidos. Las actualizaciones o retransmisiones deben utilizar el mismo número de secuencia del paquete que se va a actualizar.

Longitud de los datos CT.- Indica la longitud de los datos del contexto.

Datos del tipo de contexto.- Contienen la información del estado actual del servicio a ser transmitido.

Checksum (opcional).- Este campo es utilizado para verificar que los datos no sean corrompidos durante la transmisión.

Campo de autentificación para ESP.- Este campo contendrá la llave de seguridad ESP para que las comunicaciones puedan ser establecidas de manera segura.

Los siguientes mensajes son definidos para ayudar a verificar si existen errores en la transferencia o si se desea cancelar la misma. El protocolo CTP reconocerá estos mensajes debido a que ya se les asignó un número que debe ir en el campo tipo de la cabecera CTP.

Mensaje PNACK CT.- Este mensaje indica uno o más paquetes de datos CT faltantes. Este mensaje puede ser usado cuando existan criterios de confiabilidad (reliability).

Mensaje FNACK CT.- Este mensaje indica que se ha recibido un paquete de datos CT con aspectos de contexto faltantes. Este tipo de mensaje también puede ser usado en casos de confiabilidad.

Mensaje abortar CT.- La transferencia de contexto debe ser abortada en dos casos:

1. El MN se mueve a un tercer AR, mientras se está realizando una transferencia de contexto.

2. El contexto es innecesario, debido a que el nuevo AR determina que no necesita el contexto que le provee el AR actual (por ejemplo: en caso que el MN comience a establecer las negociaciones de contexto directamente con el nuevo AR).

En ambos casos, el AR nuevo debe enviar este mensaje al AR actual para abortar la transferencia del contexto. Este mensaje es formado asignando al campo tipo de la cabecera CT el valor de 7.

Mensaje abortar característica de CT.- Este mensaje es usado en caso que el MN comience a establecer características del contexto directamente con el nuevo AR.

Este grupo de mensajes se utilizarán en el desarrollo de un prototipo del protocolo para la transferencia del contexto CTP. El desarrollo de este prototipo está fuera del alcance de este trabajo.

V.3.2 Consideraciones de seguridad

El protocolo para la transferencia de contexto (CTP) transfiere estados entre los ARs. Si los MN no son autentificados y autorizados antes de moverse en la red, existe la posibilidad de ataques del estilo DoS (Denial of Service), intercambio de estados entre los ARs, etc., causando interrupciones en la red. Para evitar introducir latencia a la transferencia de contexto debido a la necesidad del establecimiento de canales seguros entre los dos puntos (ARs), los dos ARs deben establecer tales canales anticipadamente. Si el protocolo IPSec [Kent y Atkinson, 1998a; Kent y Atkinson, 1998b; Kent y Atkinson, 1998c] es usado, los dos ARs necesitan comprometerse (anticipadamente a cualquier transferencia del contexto) con mecanismos para el intercambio de llaves, como por ejemplo: establecer asociaciones seguras IPSec, definición de las llaves IPSec, algoritmos y protocolos IPSec (como ESP). Esto puede minimizar el tiempo necesitado para transmitir el contexto de un MN durante un *handoff* a través de canales seguros. Además, uno o ambos ARs deben autentificar al MN y autorizar sus credenciales antes de autorizar la transferencia de contexto o la entrega del contexto al MN. Otra consideración importante es que el MN reclame su propio contexto y no el de algún otro MN. Un método posible para lograr esto es a través de una "cookie" de autentificación, la cual será incluida en la transferencia de contexto y la cual debe ser confirmada por el MN antes de que el AR nuevo pueda entregarle su contexto.

V.3.3 Consideraciones de diseño

Para que el protocolo CTP sea eficiente, debe ser capaz de enviar mensajes a través de la señalización IP del *handover*, por lo que se propone como trabajo futuro que se explore el uso de la señalización en banda para este propósito[5]. Como otra medida para optimizar el tiempo de ejecución, proponemos que CTP permita enviar mensajes de señalización independientemente del flujo IP, permitiendo el uso de un flujo independiente CTP. Una manera de lograr esto es a través del uso de mensajes ICMP. Para seleccionar el protocolo de transporte en el cual se apoyará el desarrollo de CTP, algunas consideraciones de red deben de ser tomadas en cuenta, como por ejemplo el control de congestión de la red o el control de errores. El control de congestión de la red puede afectar el desempeño de CTP, pero ayudará a mantener el buen estado de la red, es por ello que se propone analizar el costo-beneficio que proporcionan los protocolos que la proveen, así como los que no la proveen. ICMP y UDP no proveen control de congestión, mientras que TCP y SCTP si la proveen. Este problema debe de ser muy bien analizado, debido a que existe contexto que debe ser restablecido rápidamente, y el retardo del control de la congestión puede disminuir el rendimiento del MN. Otro problema relacionado es la corrección de errores, debido a que de esta manera se evita el reenvío innecesario de los paquetes CTP. Esta corrección de errores es provista por TCP y SCTP a través de cheksums, mientras que UDP solo provee un checksum opcional. Se propone como trabajo futuro evaluar las ventajas y desventajas proporcionadas por estos protocolos, para poder decidir que protocolo es el más adecuado en la transferencia de contexto.

[5]Con este enfoque, la información de la señalización puede ser codificada dentro de los campos IP no utilizados: opciones IP

V.4 Conclusiones

En este capítulo se propuso una arquitectura para manejar la movilidad transparente, para lo cual se propuso el uso de conceptos relacionados a PBNM, así como el uso de un repositorio de políticas localizado en el GAR. Para resolver el problema del descubrimiento de los candidatos a enrutadores y selección del TAR se propuso el proceso CARD y se mostró que este proceso no introduce *overhead* significativo al desarrollo de un *handoff*. Para resolver el problema de la transferencia de contexto se propuso el protocolo CTP, el cual establece una serie de mensajes que permiten a los ARs establecer el contexto de los MNs en su nuevo enrutador de acceso. De igual manera propusimos algunas consideraciones de diseño y de seguridad de este protocolo, las cuales serán cruciales en el desarrollo de un prototipo de este protocolo. Cabe señalar que aunque se realizaron algunas consideraciones de seguridad, este tema está fuera del alcance de este trabajo. Para evitar interrupciones en las aplicaciones que residen en los MN se hicieron varias propuestas para mejorar el desempeño y evitar la pérdida de paquetes durante los *handoffs*. Entre las consideraciones propuestas se encuentran: el establecimiento de un túnel entre el AR actual y el nuevo, así como el mantener el contexto en el AR actual hasta que el MN haya traspasado un umbral. Por último, debemos señalar que este trabajo fue enfocado a la movilidad transparente intradominios, pero las consideraciones propuestas de los protocolos CARD y CTP pueden ser extendidas para ser aplicadas a una movilidad transparente interdominios, lo cual está fuera del alcance de este trabajo.

Capítulo VI

Conclusiones

Actualmente las redes inalámbricas de área local se han vuelto cada vez más populares, esto se puede observar en la actual tendencia de instalar este tipo de redes en hoteles, aeropuertos, restaurantes, salones de conferencias, etc. Sin embargo, aunque existe una gran tendencia hacia el uso de redes inalámbricas y computo móvil, muy poco se está haciendo para ofrecer a los usuarios una movilidad transparente cuando se desplazan realizando *handoffs* entre las células de cobertura inalámbrica. Es por ello que en esta tesis se propone una arquitectura que permita a los nodos móviles desplazarse entre las células inalámbricas de manera transparente al usuario. Esta arquitectura describe la manera en que deben llevarse a cabo las interacciones entre los diferentes dispositivos de la red, así como también agrega un elemento a la red (repositorio) como se observa en la sección V.1. Además propone el proceso de descubrimiento de los CARs y selección del TAR (ver sección V.2) así también, propone un protocolo encargado de realizar la transferencia de contexto del MN entre los enrutadores de acceso (ver sección V.3).

Para poder realizar el proceso del descubrimiento de los CARs y selección del TAR, se representó la información descriptiva de los dispositivos de red en esquemas de objetos de LDAP, para lo cual además de apoyarse en los esquemas predefinidos de LDAP, se creó un esquema propio (policyCorewin) en el cual se describen los objetos que representan las capacidades de los ARs así como las preferencias de los MNs. Para almacenar los esquemas LDAP de los diferentes dispositivos se propuso el uso de un repositorio cuya ubicación se sugiere sea el GAR. Debido a que este proceso debe realizar la selección de un TAR, se creó un algoritmo encargado de realizar este proceso, el

cual toma en consideración las siguientes variables: las capacidades de los CARs, las preferencias del MN y una política local. Debe señalarse que en este trabajo se propuso el uso de políticas como un medio para seleccionar el TAR. Para probar el desempeño del algoritmo propuesto se realizaron pruebas en diferentes escenarios, y con base en los resultados arrojados se puede decir que este proceso no introduce *overhead* al desempeño del proceso de la movilidad transparente.

Por otro lado, para realizar la transferencia de contexto entre el AR actual y el AR nuevo se realizó el diseño del protocolo CTP (Context Transfer Protocol), en el cual se describen las interacciones así como el conjunto de paquetes de datos que serán intercambiados entre los dos enrutadores para llevar a cabo la transferencia de contexto del MN. Debido a que la transferencia de contexto se lleva al mismo tiempo que un *handoff*, se propuso el establecimiento de un túnel de manera previa a cualquier transferencia de contexto entre el AR actual y el TAR, para evitar disrupciones en las aplicaciones de los MN debido a la pérdida de paquetes. De igual manera se realizaron otras consideraciones de diseño para evaluar en qué momento el AR actual debe desechar el contexto del MN, así como la manera en que debe ser tratada la transferencia de contexto en un escenario donde se involucra un tercer AR con el cual el MN desea realizar un *handoff* (ver sección V.3).

VI.1 Aportaciones

Durante el desarrollo de este trabajo de investigación se realizaron las siguientes aportaciones:

- Se desarrolló una arquitectura para manejar la transferencia de contexto cuando un dispositivo móvil realiza un *handoff* entre las células de cobertura inalámbrica.
- Se diseñó e implementó un prototipo del protocolo para el descubrimiento de los

CARs y selección de un TAR. Además, se hicieron pruebas de desempeño de este prototipo y se observó que su tiempo de ejecución es muy corto como para introducir *overhead* al proceso de la movilidad transparente.

- Se desarrolló un esquema de objetos LDAP (policyCorewin) para representar la información descriptiva de las capacidades de los ARs, las preferencias de los MN y la descripción de las políticas locales.

- Se propuso el uso de políticas como un medio para seleccionar un TAR entre los posibles CARs. Para ello se utilizó LDAP como un medio para almacenar y acceder a tales políticas.

- Se propuso un algoritmo para realizar la selección del TAR, el cual está basado en las capacidades de los CARs, las preferencias del MN y una política local.

- Se propuso el protocolo para la transferencia de contexto CTP, el cual muestra las interacciones y los paquetes a ser transmitidos entre los dos enrutadores de acceso (antiguo y nuevo) para llevar a cabo la transferencia de contexto.

- Se propuso el establecimiento de un túnel entre el AR antiguo y el TAR, el cual debe establecerse de manera previa a cualquier transferencia de contexto para evitar disrupciones en las aplicaciones del MN en un *handoff* debido a la pérdida de paquetes.

Es importante notar que actualmente se están realizando otros trabajos enfocados a dar solución al problema de la movilidad transparente. Tal es el caso del grupo de trabajo SeaMoby, el cual ha propuesto algunas soluciones a estos problemas [Liebsch *et al.*, 2003; Loughney *et al.*, 2003]. Sin embargo, hasta ahora estas soluciones no han abordado la representación de las capacidades de los CARs, ni definido el algoritmo

encargado de realizar la selección del TAR, así como tampoco han realizado ninguna implementación para resolver el problema de la movilidad transparente.

Estos problemas son atacados en este trabajo de investigación, por lo tanto, las aportaciones mencionadas anteriormente ayudarán a complementar el desarrollo del proceso de la movilidad transparente.

VI.2 Trabajo futuro

Durante el desarrollo de este trabajo de investigación se cumplieron los objetivos propuestos. Sin embargo, durante el transcurso de su realización surgieron nuevas inquietudes que pueden ser de gran ayuda para fortalecer la arquitectura propuesta. Por lo tanto, como trabajo futuro se propone lo siguiente:

Debido a que los temas relacionados con la seguridad estuvieron fuera del alcance de este trabajo se propone que en un futuro se realice más investigación en estos temas. Además, se propone que con base en la descripción del protocolo CTP (ver sección V.3) y las consideraciones de diseño (ver sección V.3.3) propuestas aquí, se realice un prototipo del mismo. Por último, se propone que los prototipos de los protocolos CARD y CTP sean integrados a algún protocolo de micromovilidad y una vez realizado esto se realicen pruebas de desempeño de la plataforma.

Bibliografía

Baker, F., Guerin, R., y Kandlur, D. 1996. "Specification of Commited Rate Quality of Service". `Internet Draft. draft-ietf-intserv-commit-rate-svc-00.txt`, 1996.

Baker, F., Krawczyk, J., y Sastry, A. 1997. "Integrated Services Management Information Base using SMIv2". Internet RFC, RFC 2213.

Bhagwat, P., Perkins, C., y Tripathi, S. 1996. "Network Layer Mobility: an Architecture and Survey". PERSCOMM, 3(3):54-64 p.

Bhatia, R., Lobo, J., y Kohli, M. 2000. "Policy Evaluation for Network Management". En "INFOCOM (3)". Madhur KohliBell Labs. 600 Mountain Ave., Murray Hill, NJ 07975, USA., 1107-1116 pp.

Black, D. 2000. "Differentiated Services and Tunnels". Internet RFC, RFC 2983.

Blake, S., Black, D., Carlson, M., Davies, E., Wang, Z., y Weiss, W. 1998. "An Architecture for Differentiated Services". Internet RFC, RFC 2475.

Braden, R., Zhang, L., Berson, S., Herzog, S., y Jamin, S. 1997. "Resource ReSerVation Protocol (RSVP)". Internet RFC, RFC 2205.

Campbell, A. T., Gomez, J., Kim, S., Wan, C.-Y., Turanyi, Z. R., y Valko, A. G. 2002. "Comparison of IP micromobility protocols". IEEE Wireless Communications, 9(1):72 - 82 p.

Campbell, A. T. y Gomez-Castellanos, J. 2000. "IP micro-mobility protocols". ACM SIGMOBILE Mobile Computing and Communications Review, 4(4):45-53 p.

Castelluccia, C. 1998. “A Hierarchical Mobile IPv6 Proposal”. Technical Report INRIA RT-0226. Unit´e de recherche INRIA Rhˆone-Alpes 655, avenue de l’Europe, 38330 MONTBONNOT ST MARTIN (France). 25 pp.

Cisco 2001. “DiffServ —The Scalable End-to-End QoS Model”. (Search online at:) `http://www.cisco.com/`, whitepaper, 2001.

Crow, B., Widjaja, J., Kim, J., y Sakai, P. 1997. “IEEE 802.11 Wireless Local Area Networks”. IEEE Communications Magazine, 35(9):116 - 126 p.

Deering, S. 1991. “ICMP Router Discovery Messages”. Internet RFC, RFC 1256.

DMTF 2002. “CIM Core Model V2.5 LDAP Mapping Specification”. URL: http://www.dmtf.org/standards/documents/DEN/DSP0123.pdf.

Droms, R. 1993. “Dynamic Host Configuration Protocol”. RFC 1531, Octubre 1993.

Finlayson, R., Mann, T., Mogul, J., y Theimer, M. 1984. “A Reverse Address Resolution Protocol”. RFC 903, Junio 1984.

Fladenmuller, A. y Silva, R. D. 1999. “The effect of mobile IP handoffs on the performance of TCP”. ACM Mobile Networks and Applications, 4(2):131-135 p.

Freeburg, T. 1991. “A New Technology for High Speed Wireless Local Area Networks”. IEEE Workshop on Wireless LANs. Motorola Inc., Arlington Heights, IL. 127-139p.

García-Macías, J. y Toumi, L. 2003. “Wireless Local Access to the Mobile Internet”. En Furht, B. y Ilyas, M., editors, “Wireless Internet Handbook”. Auerbach Publications. Boca Raton, FL : CRC Press. 227-244 p.

Georgiades, M., Politis, C., Tafazolli, R., y D.Gatzounas 2003. “Context Transfer Extension to Cellular-IP”. draft-georgiades-seamoby-ctecip-00.txt, work in progress.

Gfeller, F. R. y Bapst, U. 1979. "Wireless in-house data communication via diffuse infrared radiation". Proceedings of the IEEE, 67(11):1474-1486 p.

Gustafsson, E., Jonsson, A., y Perkins, C. 2001. "Mobile IP Regional Registration". draft-ietf-mobileip-reg-tunnel-04.txt, work in progress.

Haartsen, J. 1998. "The Bluetooth radio system". IEEE Personal Communications, 7(1):28 - 36 p.

Hodges, J. y Morgan, R. 2002. "Lightweight Directory Access Protocol (v3): Technical Specification". Internet RFC, RFC 3377.

Howard, L. 1998. "An Approach for Using LDAP as a Network Information Service". Internet RFC, RFC 2307.

IEEE 1996. "IEEE Standards for Information Technology – Telecommunications and Information Exchange between Systems – Local and Metropolitan Area Network – Specific Requirements – Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications".

Johansson, P., Kazantzidis, M., Kapoor, R., y Gerla, M. 2001. "Bluetooth: An enabler for personal area networking". IEEE Network, 15(5):28 - 37 p.

Kamerman, A. 1997. "Spread Spectrum Schemes for Microwave-frequency WLANs". Microwave Journal, 72(4):80-90 p.

Kent, S. y Atkinson, R. 1998a. "IP Authentication Header". Internet RFC, RFC 2402.

Kent, S. y Atkinson, R. 1998b. "IP Encapsulating Security Payload (ESP)". Internet RFC, RFC 2406.

Kent, S. y Atkinson, R. 1998c. "Security Architecture for the Internet Protocol". Internet RFC, RFC 2401.

Khun-Jush, J., Malmgrem, G., Schramm, P., y Torsner, J. 2000. "Hiperlan type 2 for broadband wireless communication". ER, (2):108-119 p.

Krishnamurthi, G. 2002. "Requeriments for CAR discovery protocols". draft-ietf-seamoby-card-requirements-02.txt, work in progress.

LaMaire, R., Krishna, A., Bhagwat, P., y Panian, J. 1996. "Wireless LANs and Mobile Networking: Standards and Future Directions". IEEE Communications Magazine, 34(8):86 - 94 p.

Levkowetz, O. H., Calhoun, P. R., Kempf, J., Kenward, G., Syed, H., Manner, J., Nakhjiri, M., y Krishnamurthi, G. 2002. "Problem description: reasons for performing context transfer between nodes in an IP access network". Internet RFC, RFC 3374.

Liebsch, M., Singh, A., Chaskar, H., y Funato, D. 2003. "Candidate Access Router Discovery". draft-ietf-seamoby-card-protocol-01.txt, work in progress.

Loughney, J., Nakhjiri, M., Perkins, C., y Koodli, R. 2003. "Context Transfer Protocol". draft-ietf-seamoby-ctp-02.txt, work in progress.

Lupu, E. C. y Sloman, M. S. 1997. "Conflict analysis for management policies". En "Proceedings of the 5th IFIP/IEEE International Symposium on Integrated Network management IM'97, San Diego, CA".

Moore, B., Ellesson, E., Strassner, J., y Westerinen, A. 2001. "Policy Core Information Model – Version 1 Specification". Internet RFC, RFC 3060.

Myles, A. y Skellern, D. 1993. "Comparing Four IP Based Mobile Host Protocols". Computer Networks and ISDN Systems, 26(3):349-355 p.

Nakhjiri, M. 2003. "Time Efficient context Transfer (TEXT)". draft-nakhjiri-seamoby-text-ct-02.txt, work in progress.

Nichols, K., Blake, S., Baker, F., y Black, D. 1998. "Definition of the Differentiated Services Field (DS Field) in the IPv4 and IPv6 Headers". Internet RFC, RFC 2474.

Nobel, C. 2000. "Making 802.11 standards work together". eWeek, july 19, 2000.

Pahlavan, K. 1985. "Wireless communications for office information networks". IEEE Communications Magazine, 23(6):19-27 p.

Pahlavan, K., Probert, T., y Chase, M. 1995. "Trends in Local Wireless Networks". IEEE Communications Magazine, 33(3):88-95 p.

Perkins, C. 1996a. "IP Mobility Support". IETF RFC 2002.

Perkins, C. 1996b. "Minimal Encapsulation Within IP". Internet RFC 2004, 1996.

Ramjee, R., Porta, T. F. L., Thuel, S., Varadhan, K., y Wang, S. Y. 2002. "HAWAII: A Domain-Based Approach for Supporting Mobility in Wide-Area Wireless Networks". IEEE/ACM Transactions on Networking, 10(3):396-410 p.

Ramjee, R., Porta, T. L., Thuel, S., Varadhan, K., y Salgarelli, L. 1999. "IP micro-mobility support using HAWAII". draft-ietf-mobileip-hawaii-00.txt, work in progress.

Reinbold, P. y Bonaventure, O. 2003. "Ip micro-mobility protocols". IEEE Communications Surveys and Tutorials.

Rivest, R. 1992. "The MD5 Message-Digest Algorithm". IETF RFC 1321.

Shenker, S., Partridge, C., y Guerin, R. 1997. "Specification of Guaranteed Quality of Service". Internet RFC, RFC 2212.

Strassner, J., Moore, B., Moats, R., y Ellesson, E. 2002. "Policy Core LDAP Schema". draft-ietf-policy-core-schema-16.txt, work in progress.

Syed, H., Kenward, G., Calhoun, P., Nakhjiri, M., Koodli, R., Atwal, K., Smith, M., y Krishnamurthi, G. 2003. "General Requirements for Context Transfer". draft-ietf-seamoby-ct-reqs-05.txt, work in progress.

Tanenbaum, A. S. 1996. "Computer Networks". Prentice Hall PTR. ISBN 0-13-349945-b.

Trossen, D., Krishnamurthi, G., Chascar, H., y Kempf, J. 2002. "Issues in candidate access router discovery for seamless IP-level handoffs". draft-ietf-seamoby-cardiscovery-issues-03.txt, work in progress.

Trossen, D., Krishnamurthi, G., Chaskar, H., Chalmers, R. C., y Shim, E. 2003. "A Dynamic Protocol for Candidate Access-Router Discovery". draft-trossen-seamoby-dycard-01.txt, work in progress.

Tuch, B. 1991. "An ISM band spread spectrum local area network: WaveLAN". IEEE Workshop on Wireless LANs. Worcester Polytechnic Institute. 140-145 p.

Valko, A. 1999. "Cellular IP - A New Approach to Internet Host Mobility". ACM Computer Communication Review, 29(1):50-65 p.

Verma, D. C. 2002. "Simplifying Network Administration using Policy based Management". IEEE Network, 16(2):20 - 26 p.

Verma, D. C., Calo, S., y Amiri, K. 2002. "Policy-based management of content distribution networks". IEEE Network, 16(2):34-39 p.

Apéndice A

Administración de red basada en políticas

Actualmente, la gestión de redes en empresas o instituciones es una tarea muy compleja. Esto se debe a que nos encontramos en una era donde la complejidad técnica aumenta día con día, por lo tanto, cada vez es más difícil contar con personal capacitado que pueda manejar las nuevas capacidades con las que cuentan los servidores, enrutadores y switches. Sin embargo, PBNM (Policy Based Network Management) proporciona los medios, a través de los cuales el proceso de la administración puede ser simplificado y ampliamente automatizado. Para poder desarrollar sistemas de PBNM se requiere de un proceso de estandarización que asegure la interoperabilidad entre equipos de diferentes proveedores y entre sistemas PBNM de diferentes desarrolladores. Para solucionar esto el IETF (Internet Engineering Task Force)[1] y el DMTF (Distributed Management Task Force)[2] están trabajando actualmente en la definición de estándares para PBNM.

El DMTF está enfocado principalmente en la representación de políticas y la especificación de un esquema y modelo de información correspondiente. El IETF está trabajando en ese campo en cooperación con el DMTF, mientras tanto también está tratando de definir una estructura general para un sistema PBNM, así como un protocolo que pueda ser usado para implementar un sistema PBNM. En comparación con los tradicionales enfoques de administración de redes, PBNM ofrece una solución más

[1] http://www.ietf.org
[2] http://www.dmtf.org

flexible y de completa administración, esto permite a cada router/switch ser configurado por una aplicación específica hecha a la medida del cliente. Esto lleva un costo, por lo cual, se tiene que hacer un cambio en la semántica de las políticas, en el modelo actual de PBNM [Verma, 2002].

Las políticas que pueden ser definidas son limitadas por el actual modelo de información, el cual incluye solamente clases para la representación de condiciones de políticas basadas únicamente en tiempo y en las cabeceras de los paquetes. No es posible definir nuevos tipos de condiciones (condiciones basadas en el estado de un nodo o de la red) sin extender el núcleo del modelo de información. Además la actual arquitectura de PBNM solo puede manejar un bastante limitado dominio de problemas, como los que pueden ser trasladados a escenarios fijos de configuración. Mas aun, de acuerdo a las estructuras de las políticas, estas son definidas o modificadas por una herramienta administrativa y la intervención del administrador siempre es requerida.

A.1 Cuerpos de estándares relevantes e iniciativas.

El IETF distingue los siguientes componentes para una arquitectura PBNM: Punto de Decisión de Políticas (PDP, Policy Decision Point), Punto de Ejecución de Políticas (PEP, Policy Enforcement Point), Repositorio de políticas y una herramienta de administración de políticas (ver figura 33) [Verma *et al.*, 2002]. Un administrador usa la herramienta para la administración de políticas, así como para definir las políticas que serán implementadas dentro de la red [Lupu y Sloman, 1997; Bhatia *et al.*, 2000]. Un dispositivo que pueda aplicar y ejecutar las diferentes políticas es conocido como PEP. El repositorio de políticas es usado para almacenar las políticas generadas por la herramienta de administración. Para lograr la interoperabilidad entre los productos de diferentes vendedores, la información almacenada en el repositorio debe corresponder a un modelo de información especificado por el grupo de trabajo del IETF llamado

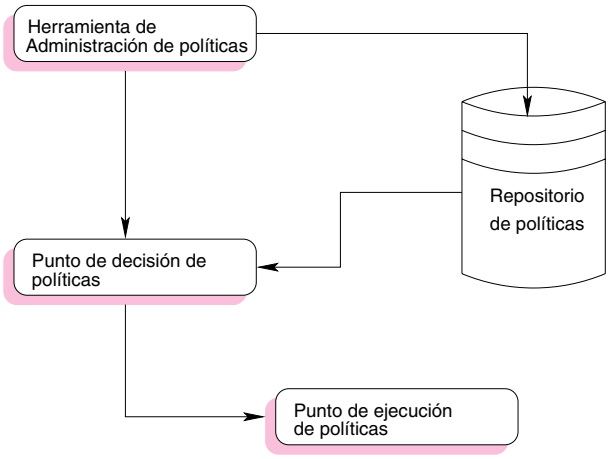

Figura 33. Componentes de una arquitectura PBNM

Policy (Policy Framework Working Group)[3]. El PEP usa un intermediario conocido como PDP para comunicarse con el repositorio. El PDP es responsable de interpretar las políticas almacenadas en el repositorio y comunicárselas al PEP. El PEP o PDP pueden estar en un simple dispositivo o en diferentes dispositivos físicos. Un repositorio puede ser un directorio de un servidor en la red, el cual es accesado usando LDAP [Hodges y Morgan, 2002].

Para adaptar la dirección de la red hacia al cliente, es necesario hacer factible la arquitectura PBNM a las aplicaciones para establecer sus propias políticas de dirección. De esta manera, es posible colocar lo mejor del conocimiento del cliente en sus aplicaciones, de tal manera, que el usuario final tenga un pequeño control sobre los recursos disponibles. Comenzando desde estas suposiciones, un mecanismo posible puede consistir en insertar datos de las políticas en paquetes activos de administración, para que el manejo del ambiente de ejecución sea capaz de extraer y almacenar las políticas cuando estas llegan al nodo, procediendo después a su ejecución.

Como se menciono anteriormente, las políticas serán almacenadas en un repositorio

[3]http://www.ietf. org/html.charters/policy-charter.html

LDAP, por lo tanto en la siguiente sección se mostrará el uso de esta tecnología.

A.2 Lightweight Directory Access Protocol (LDAP)

Lightweight Directory Access Protocol (LDAP) es un conjunto de protocolos usados para acceder a información centralizada en una red, este protocolo está basado en el estándar X.500 usado para compartir directorios. Sin embargo, es más ligero, y por esta razón algunas veces es llamado X.500 lite. LDAP organiza la información de manera jerárquica usando directorios. Estos directorios pueden almacenar una variedad de información e incluso pueden ser usados de una manera similar a NIS (Network Information Service), permitiendo a cualquier usuario acceder su cuenta desde cualquier máquina en la red.

En muchos casos, LDAP es usado simplemente como un directorio telefónico, permitiendo a los usuarios encontrar la información de otros usuarios de una manera simple. Pero LDAP va mas allá de un directorio telefónico tradicional, debido a que es capaz de propagar sus directorios a otros servidores LDAP a través del mundo, proveyendo acceso global a la información. Actualmente LDAP es mas comúnmente usado en ambientes individuales como universidades, departamentos de gobierno y compañías privadas.

Este protocolo es un sistema cliente-servidor. El servidor puede usar una variedad de bases de datos almacenadas en un directorio, donde cada una de ellas es optimizada para realizar rápidas operaciones de lectura de información. Cuando un cliente LDAP se conecta a un servidor LDAP, el puede consultar la información de un directorio o ingresar información al directorio. En el evento de una consulta, el servidor puede contestar la petición o si el no puede contestar localmente, el servidor puede enviar la consulta a los servidores LDAP localizados más arriba en la jerarquía, los cuales pueden tener la respuesta a la consulta. Si el cliente LDAP intenta ingresar información al servidor LDAP, el servidor verificará que el usuario tenga permiso de hacer los cambios

y después ingresará o actualizará la información.

El principal beneficio de LDAP es que la información de una organización puede ser consolidada dentro de un repositorio central. De esta manera, en vez de usar una lista de usuarios por cada grupo en la organización, puede usarse un directorio central LDAP el cual es accesible desde cualquier parte de la red. Otro de los beneficios de LDAP es que soporta SSL (Secure Sockets Layer) y TLS (Transport Layer Security) para que de esta manera, los datos puedan viajar de manera segura. Sin embargo, un aspecto negativo de LDAP es que su configuración no es una tarea trivial.

A.2.1 Comandos de LDAP

Los clientes LDAP son utilizados para ejecutar los comandos de: ingresar, modificar y borrar entradas en el directorio LDAP[4]. El formato de estos comandos es:

- `Ldapadd`.- ingresa una entrada a un directorio LDAP, la cual debe encontrase en un formato estándar LDIF (posteriormente descrito). Este comando es un enlace fuerte a `ldapmodify -a`.
- `Ldapmodify`.- modifica una entrada en el directorio LDAP.
- `Ldapsearch`.- busca una entrada en el directorio LDAP.
- `Ldapdelete`.- elimina una entrada de un directorio LDAP.

Con excepción de `ldapsearch`, cada uno de estos comandos puede ser utilizado de manera mas simple a través de un archivo que contiene las acciones a realizar, en vez de ejecutar cada una de esas operaciones a través de una ventana de línea de comandos. Posteriormente se mostrará un ejemplo de estos comandos.

[4]Este directorio se encuentra en /usr/bin en los sistemas UNIX

A.2.2 Terminología LDAP

En un directorio LDAP cada entrada es una unidad, la cual es identificada por un nombre distintivo único (DN, Distinguished Name). Cada una de las entradas contiene atributos, los cuales son piezas de información asociadas directamente con la entrada. Por ejemplo: una organización puede ser una entrada LDAP y sus atributos asociados puede ser su número de fax, su dirección, etc. Las personas también pueden ser entradas de un directorio LDAP, donde sus atributos incluyen número de teléfono, dirección de correo, etc.

Algunos de estos atributos son requeridos, mientras que otros son opcionales. Una definición de clase de objetos (objectclass) determina cuales atributos son requeridos y cuales no. Estos `objectclass` pueden ser encontrados en varios archivos llamados esquemas[5].

El formato para el intercambio de datos en LDAP (LDIF, LDAP Data Interchange Format) es un formato de texto ASCII, el cual es utilizado para representar las entradas LDAP. Los archivos importados o exportados de una base de datos LDAP deben de estar en formato LDIF. Una entrada LDIF es similar al siguiente ejemplo:

```
[<id>]
dn: <nombre distintivo>
<atributo>: <valor>
<atributo>: <valor>
<atributo>: <valor>
```

Una entrada puede contener el número de pares de $< atributo >:< valor >$ que ella necesite. Una línea en blanco indica el final de un archivo. Sin embargo, para que estos pares puedan ser utilizados, deben de ser previamente definidos en un esquema.

[5]En los sistemas UNIX estos archivos se encuentran en /etc/openldap/schema/

Para comprender mejor como están formados estos archivos LDIF, a continuación se muestra un ejemplo de la representación de una persona en uno de estos archivos:

```
dn: cn=oyoqui,ou=Personas, dc=cicese,dc=mx
businessCategory: Alumno
roomNumber: 211
employeeType: estudiante
preferredLanguage: es
homePhone: 1875200
givenName: Juan Manuel
givenName: Manuel
objectClass: top
objectClass: person
objectClass: organizationalPerson
objectClass: inetOrgPerson
objectClass: ipHost
objectClass: posixAccount
objectClass: ieee802Device
userPassword:: Y2VyZWJyYXRl
uid: 132
mail: joyoqui@cicese.mx
uidNumber: 876
cn: Oyoqui
cn: juanma
initials: JMOS
macAddress: 00:30:BD:60:00:11
telephoneNumber: 1890707
carLicense: 12334556
street: Av. Ruiz
ipHostNumber: 158.97.22.218
gidNumber: 987
st: BC
l: ensenada
description: Administrador
homeDirectory: /home/oyoqui
sn: Juan
```

A.2.3 Uso de LDAP en nuestro trabajo

En la sección V.2 se describió el papel que jugó LDAP en la arquitectura propuesta, sin embargo en esta sección se examinará la forma en que se utiliza LDAP.

Configuración de los archivos LDAP

Para poder hacer uso de LDAP, primero se tiene que editar su archivo de configuración[6], dentro del cual se encontrará una configuración como la siguiente:

```
database bdb
suffix "dc=<MI-DOMINIO>,dc=<COM>"
rootdn "cn=Manejador,dc=<MI-DOMINIO>,dc=<COM>"
rootpw secret
directory /usr/local/var/openldap-data
```

Esta configuración debe ser reemplazada con los valores que reflejen el nombre de dominio donde se utilizará la base de datos LDAP, como por ejemplo cicese.mx. También deben cambiarse los valores del manejador de la base de datos y su contraseña. Un ejemplo de esta configuración puede ser:

```
database bdb
suffix "dc=cicese,dc=mx"
rootdn "cn=juanmanuel,dc=cicese,dc=mx"
rootpw cicese2003
directory /usr/local/var/openldap-data
```

Para poder utilizar los esquemas de LDAP, en ese mismo archivo *slapd.conf* deben incluirse las referencias a los archivos que contienen los esquemas[7], esto se hace de la siguiente manera:

[6] En sistemas UNIX este archivo usualmente se encontrará en /usr/local/etc/openldap/slapd.conf

[7] En los sistemas UNIX estos archivos usualmente se encontrarán en el directorio /usr/local/etc/openldap/schema/.

```
# include schema
include /usr/local/etc/openldap/schema/core.schema
include /usr/local/etc/openldap/schema/cosine.schema
include /usr/local/etc/openldap/schema/inetorgperson.schema
include /usr/local/etc/openldap/schema/policyCorewin.schema
```

De igual manera si se desea crear un esquema propio, éste debe referenciarse en este archivo de configuración. Para crear un archivo de configuración puede verse http://www.openldap.org/doc/admin21/schema.html.

Por último, para levantar la base de datos LDAP, debe ejecutarse el siguiente comando:

```
/etc/rc.d/init.d/ldap start
```

Debe tenerse en cuenta que para realizar las configuraciones, así como para levantar el servidor LDAP, deben tenerse permisos de administrador (root).

Uso de los comandos LDAP

En esta sección se muestra como se utilizan los comandos LDAP para administrar la base de datos. Primeramente se mostrará el uso de `ldapadd`, cuya estructura es la siguiente:

```
ldapadd -x -D "cn=juanmanuel,dc=cicese,dc=mx" -W -f router.ldif
```

Donde el parámetro x especifica el identificador de autorización y el parámetro D especifica el enlace a la base de datos, estos dos parámetros están especificados en la cadena que se encuentra entre comillas (“ ”). El parámetro W es usado para indicar que se requiere una contraseña para ingresar la entrada, y el parámetro f indica que la entrada se encuentra en un archivo, en este caso *router.ldif*. En la figura 34 se muestra el uso de este comando en una consola.

Ahora, si se quisiera buscar la entrada para comprobar que fue ingresada correctamente, se haría de la siguiente manera:

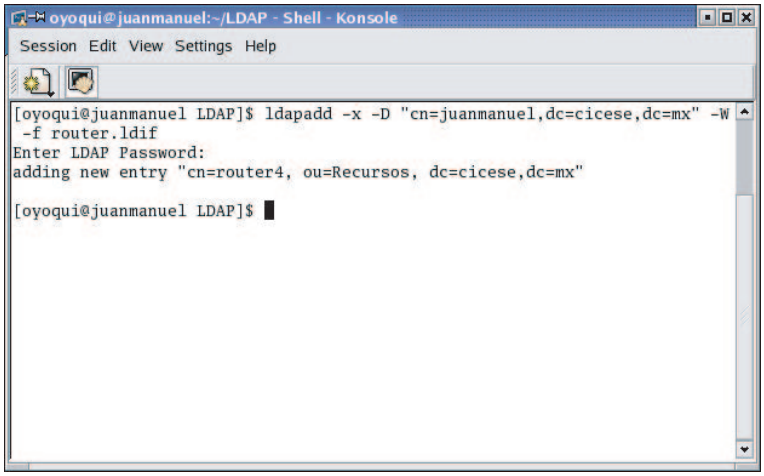

Figura 34. Ingreso de una entrada en LDAP

```
ldapsearch -x -b 'dc=cicese,dc=mx' "cn=router4"
```

Donde el parámetro *b* indica la base del directorio sobre la cual se hará la búsqueda. La base se encuentra en la cadena encerrada entre apóstrofes (' '), y la entrada que se busca se encuentra en la cadena encerrada entre comillas (" "). En la figura 35 se muestra el uso de este comando en una consola.

Para modificar cualquier entrada, solo se necesita editar el archivo *.ldif* que contiene la información de tal entrada y se utiliza la misma estructura del comando `ldapadd`, solo que con el comando `ldapmodify`, por ejemplo:

```
ldapmodify -x -D "cn=juanmanuel,dc=cicese,dc=mx" -W -f router.ldif
```

Por último, si se quisiera borrar cualquier entrada en el directorio LDAP, esto se haría de la siguiente manera:

```
ldapdelete -x -D "cn=juanmanuel,dc=cicese,dc=mx" -W "cn=router4
,ou=recursos, dc=cicese, dc=mx"
```

Donde la primera cadena especifica el usuario y la base de datos a la que se enlazará

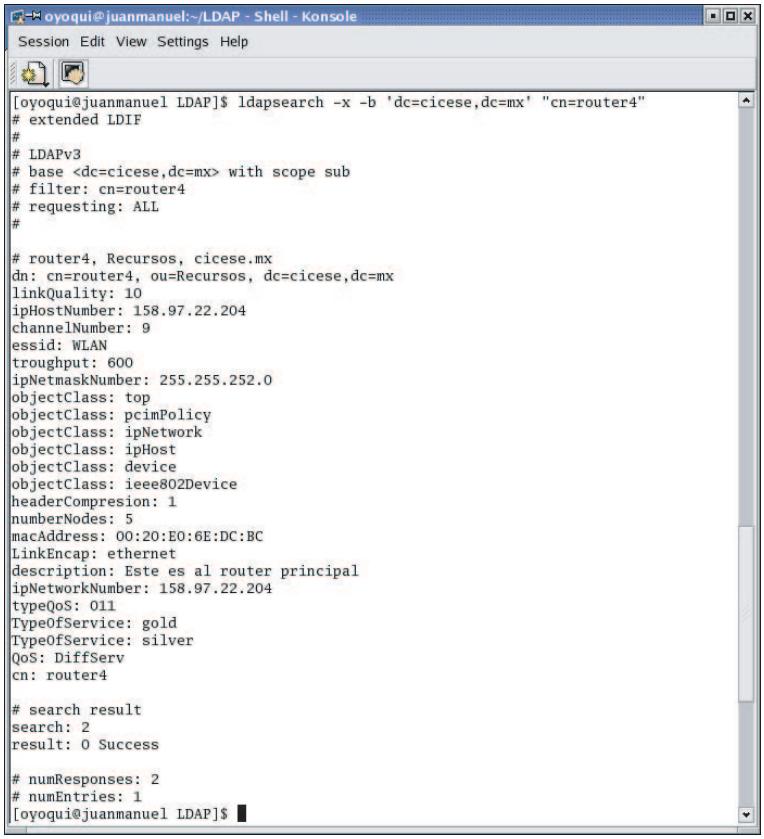

Figura 35. Búsqueda de una entrada en LDAP

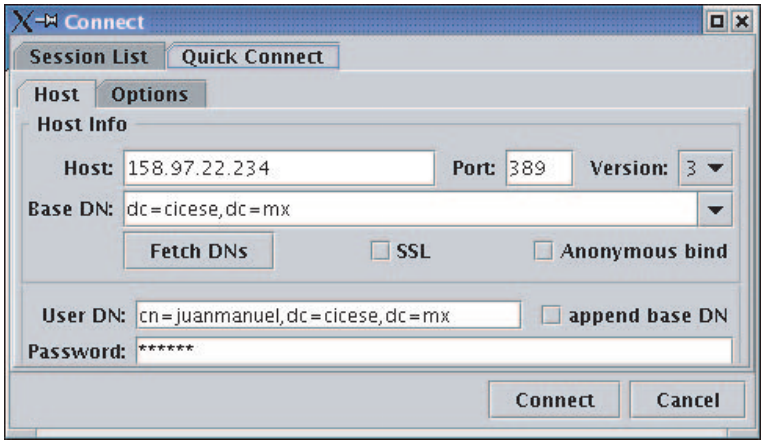

Figura 36. Configuración de un cliente LDAP

la petición, y la segunda cadena especifica la entrada que se borrará[8].

Uso de un cliente LDAP

LDAP puede ser mas fácilmente usado y administrado a través de un cliente gráfico. Durante la realización de este trabajo se utilizó la herramienta LDAP Browser[9]. La configuración de este cliente puede hacerse a través del menú FILE→connect, y enseguida aparecerá la ventana donde se especifican las opciones para conectarse. En la pestaña Quick Connect se encuentra la forma para conectarse al servidor LDAP, como se muestra en la figura 36. La ventaja de utilizar un cliente gráfico, es que se puede modificar una entrada sin necesidad de editar un archivo, al igual que se pueden exportar o importar las entradas de la base de datos de manera sencilla.

Una vez conectados, la base de datos será mostrada en el programa cliente como se muestra en la figura 37. Un inconveniente de los clientes gráficos, es que cuando se genera un error, estos no especifican adecuadamente el tipo de error, haciendo un poco

[8]En http://www.openldap.org/doc/admin21/ puede obtenerse mas información sobre el uso de estos comandos.

[9]Esta herramienta puede ser descargada de http://www.iit.edu/ gawojar/ldap/

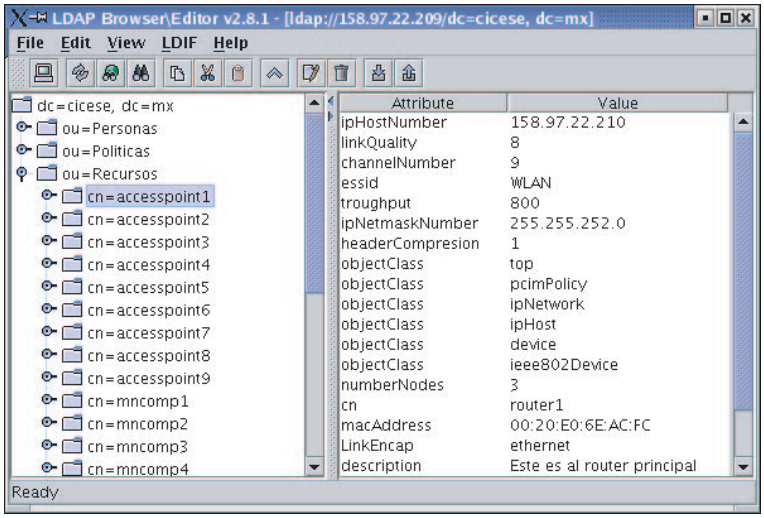

Figura 37. Cliente LDAP

mas compleja la tarea del administrador.

Apéndice B

Glosario de términos y acrónimos

Para una mejor comprensión de los temas tratados durante este trabajo de investigación, en este apéndice se describen los términos y acrónimos utilizados, los cuales son:

ACK.- Notificación enviada por un dispositivo de la red a otro para comunicar que se produjo un evento determinado (por ejemplo, la recepción de un mensaje). También llamado acuse de recibo.

Actual AR.- El enrutador de acceso que ofrece conectividad al nodo móvil después de la realización de un handoff.

Agente foráneo (Foreign Agent, FA).- El FA puede proveer de varios servicios al MN durante su visita a la red foránea. El FA puede tener el COA que actúa como punto final del túnel y entrega paquetes al MN. Además el FA puede ser el enrutador por omisión del MN.

Agente en el hogar (Home Agent, HA).- El HA, localizado en la red hogar, provee de varios servicios al MN. Primero, es el comienzo del túnel para enviar los paquetes al MN. Además el HA mantiene un registro de la localización del MN, esto es, el MN le informa su actual COA.

Antiguo AR.- Un enrutador de acceso que ofreció conectividad a un nodo móvil antes de la realización de un handoff.

AR Candidato (CAR, Candidate AR).- Un AR con el cual el MN tiene una oportunidad de desarrollar un handoff a nivel IP. Esto significa que el MN tiene la interfaz de radio correcta para conectarse a un AP que es servido por este AR, cuando

la cobertura de este AR se traslape con la del AR al cual está actualmente adjunto el MN.

AR Objetivo (TAR, Target AR).- Un AR con el que son iniciados los procedimientos para realizar un handoff a nivel IP. El TAR es seleccionado después de correr un algoritmo de selección de TAR que toma en cuenta las capacidades del CAR, preferencias del MN y algunas políticas locales.

Aspectos del contexto.- Es la colección de información que representa el contexto de un aspecto dado. El contexto completo asociado con un nodo móvil es la colección de uno o más aspectos del contexto.

ATM.- Modo de transferencia asíncrona. Estándar internacional para relevo (relay) de celdas en el que múltiples tipos de servicios (como por ejemplo, voz, vídeo o datos) se transmiten en celdas de longitud fija (53 bytes). Las celdas de longitud fija permiten que el procesamiento de las celdas se produzca en el hardware, reduciendo así los retrasos de tránsito. ATM se encuentra diseñado para aprovechar los medios de transmisión de alta velocidad como E3, SONET y T3.

Bandera (Beacon).- Un mensaje de control difundido (broadcast) por un nodo (especialmente, una estación base) para informar a todos los otros nodos en la red de la presencia continua de un nodo de difusión. Este mensaje también puede contener información del estado del nodo así como información de configuración.

Broadcast.- Paquete de datos enviado a todos los nodos de una red. Los broadcasts se identifican mediante una dirección de broadcast.

Buffer.- Área de almacenamiento utilizada para manejar datos en tránsito. Los búferes se usan en la interconexión de redes (internetworking) para compensar las diferencias en velocidad de procesamiento entre dispositivos de red. Se pueden almacenar ráfagas de datos en los búferes hasta que los dispositivos de procesamiento más lentos las puedan manejar.

Capacidades del AR.- Son las características de los servicios ofrecidos por un AR, que pueden ser de interés a un nodo móvil, cuando el AR es considerado un candidato para realizar un handoff.

Care Of Address (COA).- COA es el punto de terminación de un túnel hacia el nodo móvil, al cual llegan los datagramas enviados al nodo móvil cuando éste se encuentra fuera de su red hogar. El protocolo puede utilizar dos tipos diferentes de COA: un "*foreign agent care of address*" que es la dirección de un agente foráneo con la que está registrado el nodo móvil, y un "*co-located care of address*" la cual es una dirección local que el nodo móvil ha asociado con una de sus interfaces de red.

Contexto.- Es la información requerida del estado actual de un servicio para restablecer ese mismo servicio en una nueva subred, sin tener que desarrollar desde el principio todo el protocolo de intercambio con el host móvil.

Datagrama.- Agrupamiento lógico de información enviada como unidad de capa de red a través de un medio de transmisión sin establecer previamente un circuito virtual. Los datagramas IP son las unidades principales de información de la Internet.

Dirección hogar (Home Address, HA).- Una dirección IP asignada a un nodo móvil, la cual es usada como su dirección permanente. El prefijo de red de esta dirección IP indica la red hogar del nodo móvil.

Dirección IP.- Dirección de 32 bits asignada a los hosts que usan TCP/IP. Cada dirección consta de un número de red, un número opcional de subred, y un número de host. Los números de red y de subred se utilizan conjuntamente para el enrutamiento, mientras que el número de host se utiliza para el direccionamiento a un host individual dentro de la red o de la subred.

Dirección MAC.- Dirección de capa de enlace de datos estandarizada, necesaria para cada puerto o dispositivo que se conecta a una LAN. Otros dispositivos de la red usan estas direcciones para ubicar puertos específicos en la red y para crear y actualizar

las tablas de enrutamiento y las estructuras de los datos. Las direcciones MAC tienen una longitud de 6 bytes y son controladas por el IEEE. También denominada dirección de hardware, dirección de capa MAC o dirección física.

Dominio administrativo.- Una colección de redes bajo el mismo control administrativo y agrupados conjuntamente para propósitos administrativos.

DoS (Denial of Service).- Es un incidente en el que un usuario o una organización son privados de los servicios de un recurso que normalmente esperan tener.

Encapsulamiento.- Es el mecanismo de tomar un paquete compuesto de una cabecera y datos, y colocarlo completo dentro de la parte de datos de un nuevo paquete con una nueva cabecera.

Enrutador de acceso (AR, Access Router).- Un enrutador IP residiendo en la red de acceso y conectado a uno o más AP. Un AR ofrece conectividad IP al MN.

Gateway.- En la comunidad IP, término antiguo que se refiere a un dispositivo de enrutamiento. Actualmente, el término enrutador se utiliza para describir nodos que desempeñan esta función y gateway se refiere a un dispositivo especial que realiza una conversión de capa de aplicación de la información de una pila de protocolos a otro.

Handoff (o Handover).- Es el proceso de re-encaminar las rutas de conexión de un nodo móvil conforme este se mueve desde un punto de acceso a otro, de manera tal que la conexión sea preservada de manera transparente.

Handoff duro.- Un handoff cuyo principal objetivo es minimizar el retardo, sin el interés explícito de pérdida de paquetes.

Handoff suave.- Un handoff cuyo principal objetivo es minimizar la pérdida de paquetes, sin el interés explícito de retardos adicionales en el reenvío de paquetes.

Handoff transparente.- Un handoff en el que no hay cambio en las capacidades de servicio, seguridad o calidad. En la práctica, se espera alguna degradación en el servicio. La definición de handoff transparente en un caso práctico es que otros protocolos,

aplicaciones o usuarios finales no detecten ningún cambio en la capacidad del servicio, seguridad o calidad.

Hub.- Dispositivo de hardware o software que contiene múltiples módulos independientes pero que están conectados a los equipos de red y de internetwork. Los hubs pueden ser activos cuando repiten señales enviadas a través de ellos, o pasivos cuando no repiten las señales sino simplemente dividen las señales enviadas a través de ellos.

Identificador de microflujo.- Este campo debe contener el identificador de un microflujo dado. Un microflujo [Blake *et al.*, 1998] es una instancia de un flujo de una aplicación-a-aplicación, el cual es identificado por una dirección fuente, un puerto fuente, una dirección destino, un puerto destino y un identificador de protocolo.

IEEE.- Instituto de Ingeniería Eléctrica y Electrónica. Organización profesional cuyas actividades incluyen el desarrollo de estándares de comunicaciones y redes. Los estándares de LAN de IEEE son los estándares que predominan en las LAN de la actualidad.

IETF.- Fuerza de Tareas de Investigación de Internet. Fuerza de tareas compuesta por más de 80 grupos de trabajo responsables por el desarrollo de estándares de Internet.

Internet.- Término utilizado para referirse a la interconexión de redes (internetwork) más grande del mundo, que conecta decenas de miles de redes de todo el mundo, y con una cultura que se concentra en la investigación y estandarización basada en el uso real.

IP (Internet Protocol).- Protocolo de capa de red de la pila TCP/IP que ofrece un servicio de interconexión de redes (internetwork) no orientada a la conexión. El IP brinda funciones de direccionamiento, especificación del tipo de servicio, fragmentación y reensamblaje, y seguridad.

IPPOA (IP Point Of Attachment).- Es la entidad mas pequeña en una red. Desde esta perspectiva el conjunto de puntos de conexión IP forman una red, la cual está

conectada a una dorsal IP con un gateway hacia el Internet [Reinbold y Bonaventure, 2003].

ISO (International Organization for Standardization).- Organización Internacional para la Normalización. Organización internacional que tiene a su cargo una amplia gama de estándares, incluidos aquellos referidos a la networking. ISO desarrolló el modelo de referencia OSI, un popular modelo de referencia de conexión de redes (networking).

LAN (Local Area Network).- Red de datos de alta velocidad y bajo nivel de error que cubre un área geográfica relativamente pequeña (hasta unos pocos miles de metros). Las LAN conectan estaciones de trabajo, periféricos, terminales y otros dispositivos en un solo edificio u otra área geográficamente limitada.

LDAP.- Es un conjunto de protocolos usados para acceder a información centralizada en una red.

MAC (Media Access Control).- Control de acceso al medio. Capa inferior de las dos subcapas de la capa de enlace de datos, según la define el IEEE. La subcapa MAC maneja el acceso a los medios compartidos, por ejemplo, si se utilizara la transmisión o la contención de estafetas (tokens).

Macromovilidad.- Movilidad sobre un área grande. Esto incluye soporte de movilidad y procedimientos asociados al registro de direcciones que son necesitados cuando un nodo móvil se mueve entre dominios IP. Mobile IP puede ser visto como un medio para proveer macromovilidad.

MD5 (Message Digest 5).- Algoritmo utilizado para la autentificación de mensajes en SNMP v.2. MD5 verifica la integridad de la comunicación, autentifica el origen y verifica la puntualidad.

Micromovilidad.- Movilidad sobre un área pequeña. Usualmente esto significa movilidad dentro de un dominio IP con énfasis en el soporte de nodos móviles activos,

aunque también puede incluir procedimientos para nodos móviles inactivos.

MTU.- Unidad máxima de transmisión. Tamaño máximo de paquete, en bytes, que puede manejar una interfaz en particular.

Nodo móvil (Mobile Node, MN).- Un MN es un dispositivo o enrutador que puede cambiar su punto de conexión a Internet utilizando Mobile IP. El MN mantiene su dirección IP y puede comunicarse continuamente con otro sistema en el Internet mientras mantenga su conectividad por medio de la capa de enlace.

Host móvil.- Un nodo móvil que es un host final (end host) y no un enrutador.

Paquete.- Agrupación lógica de información que incluye un encabezado que contiene la información de control y (generalmente) los datos del usuario. El término "paquete" se usa con mayor frecuencia para referirse a las unidades de datos de la capa de red.

Paging.- Es el proceso de dividir la red en distintas áreas geográficas llamadas áreas paging. Este proceso es utilizado para conocer la localización aproximada de los nodos móviles cuando éstos se encuentran en estado inactivo [Reinbold y Bonaventure, 2003].

PBNM (Policy Based Network Management).- Es un medio por el cual el proceso de administración de la red puede ser simplificado y ampliamente automatizado.

Políticas.- Es un conjunto de reglas para administrar, manejar y controlar el acceso a los recursos de la red.

Protocolo.- Descripción formal de un conjunto de reglas y convenciones que rigen la forma en la que los dispositivos de una red intercambian información.

Punto de acceso (AP, Access Point).- Es un emisor/receptor de radio por medio del cual un MN obtiene conectividad con la red alambrada a través de la capa 2.

QoS (Quality Of Service).- Calidad de servicio. Medida de desempeño para un sistema de transmisión que refleja su calidad de transmisión y disponibilidad de servicio.

Red foránea (Foreign Network, FN).- Cualquier otra red que no sea la red hogar del nodo móvil.

Red hogar (Home Network, HN).- Una red, posiblemente virtual, que tiene un prefijo que concuerda con la dirección hogar del nodo móvil. Mobile IP no es necesario dentro de la red hogar.

Servicio candidato para la transferencia de contexto.- Un servicio que es candidato para la transferencia de contexto. Servicios involucrados en el trato que se otorga a los paquetes durante su enrutamiento como QoS y seguridad, o aquellos involucrados en otorgar o denegar acceso a la red al MN como AAA, son considerados servicios candidatos para la transferencia de contexto.

Tabla de enrutamiento.- Es una tabla almacenada en un enrutador o en algún otro dispositivo de interconexión de red (internetworking) que realiza un seguimiento de las rutas hacia destinos de red específicos y, en algunos casos, las métricas asociadas con esas rutas.

TCP (Transfer Control Protocol).- Protocolo para el control de la transmisión. Protocolo de la capa de transporte orientado a conexión que proporciona una transmisión confiable de datos bidireccional (full duplex). TCP es parte de la pila de protocolos TCP/IP.

Transferencia de contexto.- Es el movimiento del contexto de un enrutador a otro, o de una red a otra como un medio para restablecer servicios específicos en una nueva subred o colección de subredes.

Trigger.- Es la información de capa 2 que informa a la capa 3 de los eventos detallados involucrados en la secuencia de un handover a capa 2.

Túnel.- El proceso de encapsular y desencapsular datagramas es conocido como tunelaje de datagramas, donde el encapsulador es el punto inicial del túnel y el desencapsulador es el punto final del túnel.