

## Introduction

---

The degree of interest in IEEE 802.11 based wireless local area networks (WLANs) has been rapidly increasing over the past several years. This is evidenced in any number of distinct areas: the number of refereed journals with papers on the matter, the number of new start-ups selling Wi-Fi equipment, the number of installed Wi-Fi networks, the number of businesses engaged in aggregating Wi-Fi networks, the number of detracting articles published by Qualcomm and others interested in public wireless networks, and so on.

In point of fact, the detractors are not completely without their points. WLANs do have particular issues as one considers how best to scale them to form interconnected networks that cover metropolitan or national service areas. The rather formidable issues that arise in such a context include identifying the physical and logical location of a mobile Wi-Fi device (that is, managing its mobility), authenticating the device properly, and billing for service rendered. Generally, these issues can be categorized as mobility management and operations, administration, maintenance, and provisioning (MM and OAM&P).

It is not too surprising that IEEE 802.11 networks should lack support for MM and OAM&P. After all, the variety of 802.11 protocols pertain to the physical and medium access control layers of TCP/IP based local area networks. They were never envisaged as providing these functions, certainly not in the way in which, for example, they are integrated into public wireless services like the GSM.

A number of ways to address these issues have been proposed by those well-disposed to the growth of Wi-Fi. While the author is not aware of everyone's thoughts on the matter, it appears that there are two broad themes here: address the issues strictly within the context of a

wireless Internet or leave the isolated WLANs as they stand and allow for dual-mode devices that can flexibly and gracefully interwork via TCP/IP with either a public 3G network or a WLAN.

These two themes are each associated with certain advantages and disadvantages. We shall develop those in more detail in what follows. The aim of this paper is to present an alternative to either of these themes that tightly integrates WLANs with ReFLEX™ NPCS networks to provide the *missing link*; namely, a common control channel. We assume that the interested reader will typically be reasonably familiar with IEEE 802.11, TCP/IP, and the protocols of the various public wireless networks (called 2G, 2.5G, 3G); but that they will lack familiarity with ReFLEX NPCS. With that in mind, this paper will include some background material on the history and operations of ReFLEX NPCS networks before showing how they can provide much, if not all, of the missing component protocols that would easily allow Wi-Fi systems to scale nationally or internationally.

## The issues

---

### Mobility Management

---

When a Wi-Fi capable computer becomes active on a WLAN, it is typically assigned an IP address that is consistent with the gateway or router that manages the physical WLAN segment. Most often, this done by the Dynamic Host Configuration Protocol (DHCP). If the device moves to a new WLAN, with a new router, DHCP will assign a new, appropriate IP address from the range assigned to the new router.

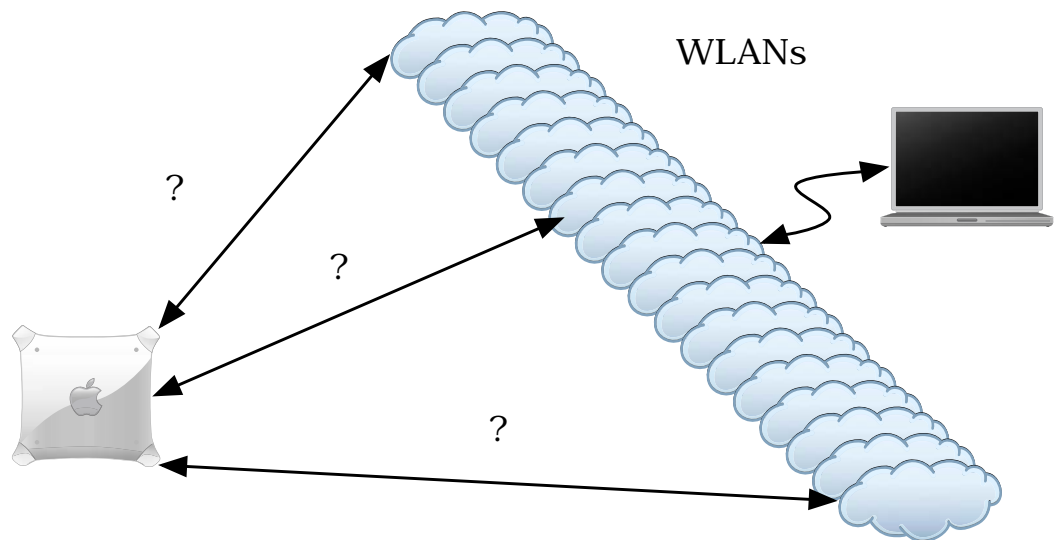
Within this model, it is straightforward for the mobile computer to act as a client to Internet-based hosts. The obvious use-case is a "road warrior" who carries his laptop from place to place, achieves local access to the

Internet via Wi-Fi, and can access the Web from a browser, can download email from a remote mail server, can interconnect to any one of a number of IM servers, can establish voice over IP (VoIP) calls, and can establish a virtual private network connection to his company's network through an IPsec router at the home base. While this is a great degree of utility, and is in no small measure responsible for the interest in Wi-Fi services, there is one thing that cannot be achieved in this model: it is impossible to push information to the mobile computer either when it is inactive and has no IP address anywhere or when its current IP address is not known to the computer that has information to push. This situation is shown graphically in Exhibit 1, where a host computer cannot identify the one WLAN in which the mobile computer is currently active.

In other words, as it stands, Wi-Fi does not support either permanently active hosts or clients. A permanently active host would typically require a name entry in one or more Domain Name Servers (DNSs) with mappings to permanent public IP addresses. A permanently active client would require either a permanent public IP address or a permanently-connected proxy box. The latter kind of functionality is commonly associated with secure communications to the public Internet from client computers behind a corporate firewall that uses both DHCP and Network Address Translation (NAT) to hide the actual private IP addresses of client computers from the ever-present array of malicious parties that threaten them.

This limitation prevents one of the simplest use cases imaginable from being achieved; namely, terminating a "call" to the mobile computer of any sort to deliver, or even indicate the presence of, important informa-

**Exhibit 1. Find me if you can**



tion. Instead, the user must be called via cell-phone or paged... a disagreeable state of affairs.

## OAM&P

It may be generally assumed that among those interested in scaling up Wi-Fi to a regional or national level will be businesses whose aim would be to achieve a profit on operations. We take this to be both natural and laudible. Assuming for the time being that the MM issue could be addressed (pun intended), a service provider must find ways to provision services, bill for them as they are consumed, prevent fraudulent use or mis-use, prevent passive or active attacks against legitimate and authorized user communications, and deal with customer complaints concerning coverage (or its absence), billing errors, denial of service, and so on.

Occasionally, one finds the acronym 'FCAPS' associated with the concepts of OAM&P, where FCAPS expands out to fault, configuration, accounting, provisioning, and security.

These OAM&P functions can be broken down into two classes in terms of their relationship to the time at which the user is active on a billable WLAN channel. Technical sales and planning, provisioning, call set-up, billing, network or service management by the customer,

repair, and technical support either must occur, or at least, can occur when the user is not on a billable channel. FCAPS functions that occur while the customer is active on the WLAN can occur on the same WLAN segment as the billable traffic in association with it.

In the context of a pure TCP/IP solution, what “non-billable channel” means is fairly fungible. It could be an Internet connection by some other means than the service provider’s WLAN. It could be a pathway forced to a specified server and port, no matter what host or protocol the mobile computer had attempted to access. At present, Wi-Fi service providers regularly employ both methods. As will be seen later, in the context of public wireless networks, the usual mechanism is signaling over a *common control* channel.

## The existing solutions

### An instructive example

In one approach to dealing with the short-comings of existing Wi-Fi, the remedies are sought within the context of extensions to TCP/IP and the protocols that operate over them. [This is not to exclude UDP, ICMP, etc.; it is simply a common turn of speech to refer to Internet protocols as “TCP/IP”.] In the other case, WLANs are “married” in some way with public 2.5G or 3G wireless networks.

To draw out some of the elements of feasible solutions, let us go back to the road warrior with a VPN connection to his corporate network via a WLAN. At the WLAN, he has likely been assigned an IP address by DHCP from the IP address space managed by the Wi-Fi service provider and consistent with his gateway router’s configuration. This IP address may be in public address space or private address space. For some degree of simplicity in what follows, let us assume that it is a public IP address. Next, there is an IPsec tunnel opened to a router with a public IP address at the corporate network’s interface to the Internet. For the sake of argument, let’s say this IPsec router is named as *secure.corp1.com*. At a layer above this router, and within the IPsec tunnel, another DHCP agent will have

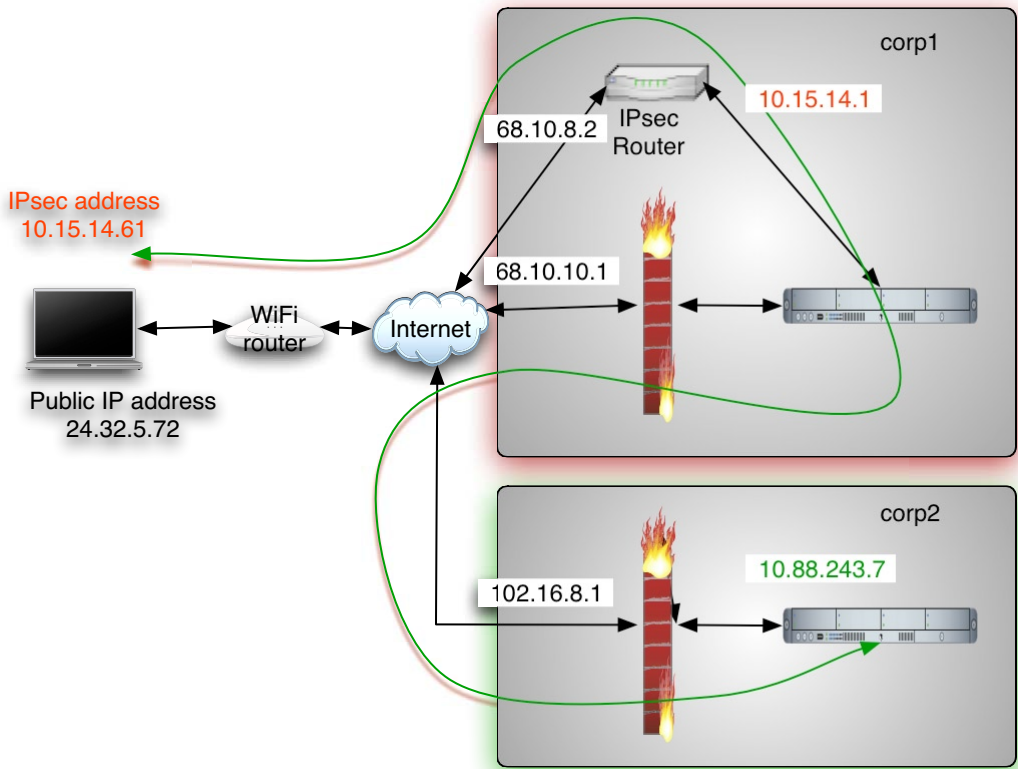
assigned the mobile computer an IP address from within the corporation’s private address range and consistent with the configuration of the IPsec router.

Now assume that the road warrior uses a web browser to open a connection to *http://www.corp2.com*, the URI for another company’s web site. His traffic is routed via the WLAN to his corporate IPsec router where it is decoded into the http, port 80 connection request that is implied to the named host at *www.corp2.com*. More than likely, this connection is achieved via a firewall and NAT router combination of some sort that proxies the internal TCP connection request from the private IP address of the mobile computer and a high order port onto the public IP address for *corp2.com* at some other high order port [or some similar legerdemain.]

A sample is shown in Exhibit 2. In this case, the public IP address of the mobile computer is 24.32.5.72 [chosen purely at random by the author, for the sake of this example.] It has a VPN path open to the IPsec router at corp1’s Internet presence at 68.10.8.2. Inside this VPN tunnel, the mobile has been assigned the private IP address of 10.15.33.61. Packets from it are injected into the corporate network from the IPsec router’s internal port at 10.15.14.1. In this example, these packets are to be routed to a web server at corp2. The flow is shown proceeding through the corp1 firewall [at 68.10.10.1] and NAT router out through the Internet and across to the firewall at 102.16.8.1, which is a component of corp2’s Internet presence. Presumably, DNS associates this address with the name *www.corp2.com*. Because of network address translation, IP packets on this segment apparently flow between 102.16.8.1 and 68.10.10.1. Finally, in this example, the web server is internally located at corp2 at the private IP address, 10.88.243.7. This server behaves as if it were connected to 68.10.10.1. At the other side, the mobile computer’s browser behaves as if it were connected to 102.16.8.1 within the VPN tunnel. An observer inspecting packets between the WLAN and corp1 would see a connection between 24.32.5.72 and 68.10.8.2.

If the uninitiated reader finds this example flow rather complex, it may be pointed out that this kind of thing occurs millions, if not billions, of times a second on the

**Exhibit 2. Routing for Wi-Fi VPN connection**



Internet. It draws out a contrast between the Internet and other networks like the PSTN; namely, that IP is extremely effective at masquerading the identity of the nodes engaged in communication.

In any case, from the point of view of the corp2 web server, the mobile computer appears to be anchored at the Internet presence of corp1. This anchor point is invariant to whatever WLAN, or other temporary Internet attachment, that the mobile computer has from time to time. This apparent permanence is an effect of the VPN tunnel into corp1 and NAT out to the Internet from that location.

This arrangement typically doesn't allow content to be pushed at the mobile computer from locations outside the corp1 network or for the mobile to act as a host. The private IP address of the mobile is usually dynamically assigned by DHCP for each VPN connection. As a consequence, the mobile's identity is not present in

the corp1 external DNS, even if the corp1 security staff would allow such an entry to be made within common security policies. Since even the internal private IP address assigned to the mobile changes dynamically, a permanent DNS entry would not be feasible. While providing a more permanent IP address to a specific mobile computer may not be a significant technical issue, dynamic assignment is presently the usual rule.

Whatever else this functional example may demonstrate, it embodies two essential components of a feasible solution: first, a reference point for the mobile device on the fixed public Internet; and second, a mobile reference

point that can be mapped onto the fixed reference point as the need arises.

### TCP/IP-based MM solutions

In a pure TCP/IP-based solution, the mobile device is associated with at least two IP addresses. The first represents its "home" IP address; the second represents its local or "temporary" IP address. To allow TCP connections to survive gracefully through switch-overs between independent WLANs (complete with independent temporary IP address assignments) the mobile computer must be capable of uniquely and globally identifying itself to the home network. In this way, the home network can consistently re-assign the mobile the same IP address and maintain TCP connectivity throughout.

To consider one possible implementation, if in the "road warrior" example just given, the mobile device

had a permanent IP address assignment from within the corp1 private space, it would be at least feasible for a live TCP connection to survive the loss and re-establishment of a VPN tunnel as the mobile transitioned from one WLAN to another.

Naturally, it would be quite another thing to make such a transition without any loss of data over the TCP link, or without having to reset the link (with new SYN/ACK exchanges to resynchronize transmissions).

Work on this form of solution addresses how to manage both mobile and home IP addresses consistently without resort to VPNs, how to co-ordinate graceful IP layer switch-overs between independent WLANs, how to maintain TCP connectivity between peer end points where one or both are mobile, and so on. While the example we've been referring to is based upon IPv4, much of the new work is in the context of IPv6 with its much larger IP address space. Where IPv4 allows for 32-bit addresses, IPv6 has 128-bit addresses.

Also, IPv6 has more sophisticated mechanisms for managing the functions of DHCP than IPv4. In addition, Apple Computer has presented to the IETF a package of protocols, called *Rendezvous*, for the automatic construction of networks of computers that includes Zero Configuration networking [ZC], automatic link-local addressing [v4LL RFC 2462], and multicast DNS [mDNS]. *Rendezvous* would potentially allow mobile computers to dynamically join or re-join a home network after a brief outage between remote WLANs, to update a DNS, and to maintain TCP connections in process.

To summarize, the pure TCP/IP approaches to dealing with the MM issues of Wi-Fi networks bind the mobile computer to a permanent home on the Internet, at which point it has an accessible IP address and perhaps, name. This fixed anchor point is bound dynamically to the current location of the mobile computer in whatever WLAN (or fixed LAN) it happens to be present.

## Session Initiation Protocol

The Session Initiation Protocol is perhaps the most sophisticated and complete version of a solution to the problem of terminating calls on arbitrary mobile Inter-

## SIP & SDP

The Session Initiation Protocol and the Session Description Protocol are structurally similar to the combination of HTTP and HTML. First of all, SIP carries SDP, just as HTTP carries HTML.

Like HTML describes the format and content of pages to be displayed by a browser, SDP describes communication sessions that are requested or accepted. SDP consists of alphanumeric strings that enumerate attributes of the session and their values. For example, the time for a session to start may be indicated by

```
t = 18000789 18001000
```

where the numbers indicate a start and stop time in seconds in Universal Co-ordinated Time (UTC).

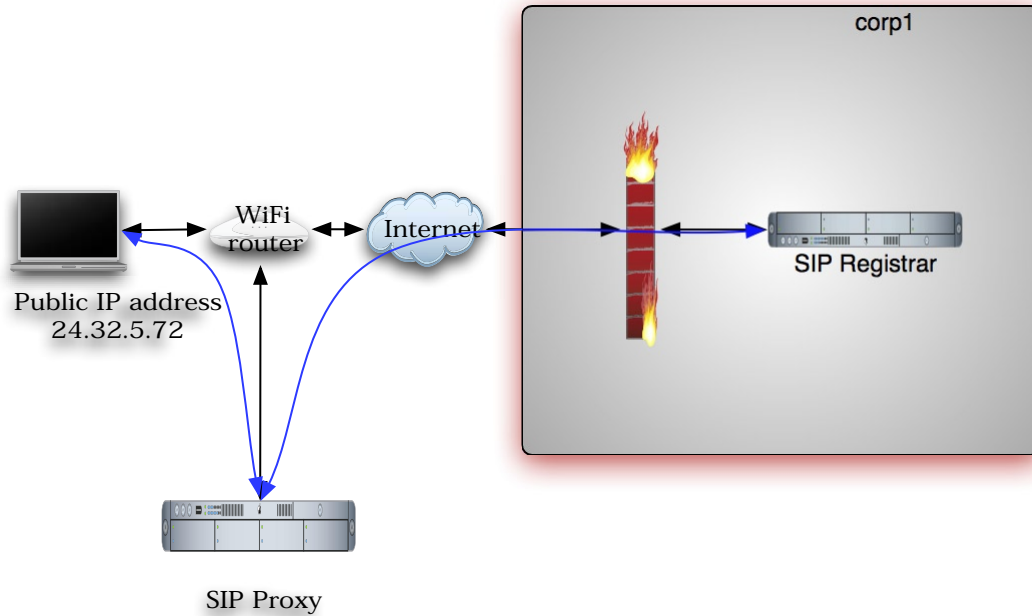
Like HTTP, SIP involves a server and a client. Most SIP-enabled devices will actively run both a server and a client simultaneously. Once a SIP client locates the address of a host that it wants to invite to a session, it sends an SDP packet to the SIP server running on the host. In response, the target host's SIP server will reply with SDP that describes its response. This may be a description of how it will set up the session, or a redirection to another point, or a decline.

The manner in which SIP supports the capability for clients to find servers that roam around the Internet in random ways is given in the body of this paper.

net hosts. The panel above tells some of the characteristics of how SIP operates to setup a connection from one host to another, given that the target host can be found. SIP also supports the location of roaming hosts.

SIP's roaming function is shown in Exhibit 3. The roaming device interacts with a local SIP Proxy server, which relays location information to a SIP Registrar at the roamer's home location. In effect, this constitutes a kind of dynamic DNS mechanism. Instead of a Domain Name Server, there is a Registrar at the home domain. Using SIP, this Registrar can be updated by proper Proxy agents with the current Internet location of the roam-

**Exhibit 3. SIP location registration**



ing host. The Proxy is essential because the roaming host may not actually know its real Internet location. For example, in any case where the roamer’s path to the public Internet is via Network Address Translation (NAT), the IP address that the roamer knows will *not* be the address needed to route traffic to it from the Internet. The Proxy must have access to this information, of course.

SIP and SDP are of use in initiating a connection that terminates on a roamer; however, they add little value directly in mobility management for active connections. For this, we must consider switch-over (or hand-over, or hand-off.) Switch-over occurs when a mobile device moves from one RF network serving point to another. This may be from one cell to another, from one network provider to another, one channel to another, and so on. The combination of SIP and SDP would allow either party at the end of a dropped connection to re-establish a new connection, but this is not quite the same as transparent switch-over.

**SDP**

SDP is given in IETF RFC 2327 and is updated in RFC 3266 for IPv6 functionality. The text payload of SDP should be no greater than 1kB. SDP is based upon strictly text encoding of lines of the form <type>=<value>. The values are all single character fields. For example, “v=” is protocol version, “o=” is owner, “s=” is session name, “e=” is email address, “t=” is time, and so on. There are components for session description, for time, and for media description.

Text encoding was chosen specifically so that a variety of transport protocols could be used to send SDP, including SIP, SMTP, and HTTP. We will later consider SDP over WCTP.

**Switch-over**

There are three basic models associated with switch-overs from one “cell” to another: network-directed, mobile-assisted, and mobile-directed. These may be briefly described as follows.

In network-directed switch-over, the decision to allocate a mobile to one access point or another is made wholly by elements within the network. This assumes that these elements can communicate across the boundaries of independent networks to support inter-system switch-over. This also assumes that the various networks have excess resources that can be allocated to assess the relative value of potential switch-overs before they are attempted. Since the switch-over begins in one network where the mobile is a “trusted” client, it is assumed that this originating network will vouch for it to the new serving system. In other words, “trust” is transitive. The

original AMPS network used this model.

In mobile-assisted switch-over, the decision as when and where to switch-over is still maintained at the network; but the information on relative signal quality is obtained by the mobile and delivered back to the network over a special signaling channel. This minimizes the requirement for network resources while maintaining network control and security.

In mobile-directed switch-over, the decision to attempt a switch-over lies completely with the mobile. The mobile must continuously assess the relative strength of various links that may be available to it, and to gain access rights to any new networks that may be essential to its access before a switch-over to them is needed. It must also be able to provide the new network with any information necessary for it to take over any live connections at the time the switch-over is affected. Since there is little, if any, inter-network communication in this model, the onus is on the new serving network to assess the credentials of the incoming mobile rapidly in order to prevent fraud while still providing an acceptable grade of service to roamers.

It goes without saying that the standard Internet does not include any network elements that perform network-directed or mobile-assisted switch-over. Nor does it incorporate any protocols that would support mobile-directed switch-over, although there are security protocols that can be employed to authenticate mobile roamers.

To say that the resolution of these issues is important to the question of scaling Wi-Fi is almost tautological (or maybe a *Kantian synthetic a priori*). Even if the mobile computer's home network has a flat rate billing model, it is almost certain that the operators of networks that mobiles roam onto will desire some form of variable cost recovery that depends upon traffic. At a minimum, there is likely to be a distinction in billing plans that include roaming and those that don't, with the automatic requirement for other networks to distinguish the roaming rights of incoming mobiles in real time.

Therefore, in addition to assessing the access rights of mobile roamers, networks need to be able (i) to measure

the consumption of roamers in some agreed, and objective way, (ii) to accumulate those measures and communicate between themselves for billing purposes, and (iii) to allow for auditing of the underlying records in the event of a dispute.

In the case of inter-system switch-over, there is a question of how to route the traffic. In an inter-system switch-over, a connection is created when the roamer is in the domain of one service provider, but the roamer moves into the domain of a new service provider during the connection. In the approach commonly used in cellular systems, the initial network remains the "anchor point" for the connection. The anchor network arranges for the switch-over to the new network (and validates the roamer's identity and credit), and the forwarding of traffic is transparent to the far side. This model can be extended almost indefinitely, if the roamer continues to hold up the connection across several new networks, although some practical limits can be designed in. It is common in the cellular service for the anchor system to receive all of the revenue for such a call. Such a model avoids the complexity of detailed call records that accumulate all of the parameters of inter-system roaming on a single connection. Customers tend not to want such detail. This business model works where there is reciprocity in the volumes of inter-system roaming traffic, but it can be abused.

In another possible approach to inter-system switch-over, the traffic does not thread through the original system but rather is re-routed to the new network. The issues with this approach are many and varied. For one, if the roaming rates of the new network are different than the original, should the user be informed, should the call be dropped, or should some other action be taken? What agent in the network should assess these rates in any case? Also, security issues are numerous. How is authentication among the various parties to be accomplished in real-time? The new roamer must be authenticated to the new network, either by its home network or by the original anchor network or both. The new identity of the roamer must be authenticated to the other end point of the connection. Any firewalls or other security points along the route must find the re-

direction of the connection acceptable. And all of these node updates must occur in negligible time as far as the connection is concerned; if any fail, the call is dropped.

Considerations such as these tend to motivate the use of the other, “anchor point” model.

While work is being done in many locations on these issues, there is no final resolution of them at present. Many remain, as they say, “interesting research projects.”

## **TCP/IP-based OAM&P solutions**

---

In one respect, OAM&P looks relatively easy within the context of TCP/IP. Much of the work is already embodied in the Simple Network Management Protocol (SNMP), or so it seems. Since the provisioning function must inter-work with the inter-system roaming and billing functions that we just outlined, the apparent ease is seen to be rather superficial as soon as the mobile device is accounted for. Even if the mobile were to embody an SNMP Management Information Base (MIB), when an error arises, where should the trap be pushed? to the home fault management system or the visited one? Again, if an SNMP daemon is running on the mobile, how can it securely communicate its current location and IP address to the home management system?

In short, the OAM&P question is implicitly tied up with the mobility management question. If Internet protocols are to be used to perform OAM&P functions, then creating and maintaining connections for these protocols across precisely those remote physical links that are a component, “untrustworthy” part of the problem is critical to the solution. For a simple example, how does a remote computer that cannot reliably obtain a DHCP lease for a local IP address communicate this to its home network from a roaming location, when having an IP address is essential to the communication process?

## **Public 2.5G, 3G solutions**

---

Let us turn our attention now to the other theme; namely, solutions that integrate WLANs with public 2.5G/3G solutions. For brevity in what follows, we shall simply refer to these as “3G” even if a GPRS or similar implementation may be feasible. Our points will be

broader than quibbles about that level of detail.

The general picture here is quite straight-forward. At an over-arching level, a mobile accesses a public 3G back-bone network. Geographically embedded within this network may be one or more isolated WLANs. Even in the presence of one of these WLANs, the mobile always has access to the public network. Because of its evolution from the 1G and 2G technologies, the public network has support for MM and OAM&P.

Furthermore, in the 3G picture, the mobile can obtain a local IP address easily; and the home network maintains an anchored Internet presence for the mobile, wherever it may roam to within the 3G system.

In this model, the independent WLANs do not really need to be inter-connected with one another directly; rather, the public network acts as a sort of “glue” that binds them together, at least from the point of view of the roaming mobile computer. Thus, mobiles do not so much roam between WLANs directly as they move into and out of them from the public back-bone network. This back-bone provides the home anchor point for the mobile from the point of view of any other computer that connects to the mobile.

In most discussions of 3G networks, the IP back-bone is based on v6 with its large public address space for all of the mobile nodes that are assumed to be inter-connected. The extent to which the public 3G service providers will offer additional security services such as firewalling, spam filtering, intrusion detection, and so on, is a moot point at present. Considering the rather limited computing resources of many mobile devices, architectures for mobile device security are an important component of any OAM&P model. However, our primary attention in this discussion is focussed on mobile computers with the horsepower to run anti-virus, firewall, anti-vandal, intrusion detection, buffer overflow protection, virtual private network, and similar security software. The physical configuration of such a laptop would include one or more modules to access both 802.11x and 3G networks, while its networking software would allow for the use of local temporary IP addresses as well as the permanent IP address of its home, anchor point. We merely observe in passing that users of hand-

sets or PDAs given a public IP address by their providers should practice the principle of *caveat emptor*.

To return to the theme of 3G networks over-laying WLANs, we can observe that the public wireless networks have completely developed mobility management and OAM&P solutions from within the context of their networks. However, the implementations of these protocol suites are natively based upon the ISO/OSI protocol stack, and upon SS7 and related ITU protocols (e.g., CMIP) in particular, rather than TCP/IP. Even in implementations where SS7 may be transported over IP, the communicating nodes must have SS7 addresses in order to participate in the OAM&P protocols. Equally importantly, the mobile end equipment is not assumed to be “trusted” or managed network equipment; rather, it is assumed to be customer owned and maintained (COAM) equipment. If the mobile fails, the sole responsibility of the service provider is to remove or disable the offending mobile in order to protect the network and other users.

In this context, a mobile computer operating within a WLAN is roughly in the same position as the user of a desk-set on a private branch exchange (PBX). The desk-set apparently has a unique, publically accessible phone number even though all its calls are routed to and from the address of the PBX. Maintenance of the desk-set is not the responsibility of the local exchange carrier (in the absence of any contractual business relationship).

---

## SIP & 3G

The 3G Partnership Program (3GPP) also supports SIP. For a 3G mobile device, it might be natural to anchor its home to the domain of the 3G carrier. Then, if the mobile roams onto the coverage of a WLAN, or even if it is connected to a wired network, it can use SIP to update the home Registrar with its current location. Under the broad assumption that the mobile unit will spend the greatest part of its time on the 3G network, this model works well.

Of course, it is also feasible “to home” the mobile device onto a wired corporate network or onto a WLAN and use SIP to note when the device roams onto a 3G

system. Presumably, the best “home” is the form of network where the mobile would spend the greatest amount of time.

So, if WLAN access constitutes a relatively small amount of the useful service time for the mobile, then the use of SIP and 3G could be a very effective mechanism for locating the roamer when it is not “at home.” On the other hand, if the mobile host is to spend the best part of its service on various heterogeneous WLANs, SIP and the 3G back-bone may not add more value than SIP alone.

---

## 3G MM & OAM&P

It goes without saying that 3G networks will be completely provided with solutions for mobility management and OAM&P *within the scope of the networks*. The extent to which these protocols and services will migrate outside the network, to a user’s corporate network for example, is questionable. In short, phone companies trust themselves to manage the devices on their networks; but they are not so sure about anybody else. Hence, automatic roaming between WLANs and 3G networks is not likely to be supported outside of the functions of SIP that have already been pointed to.

---

## An Alternative: NPCS and IEEE 802 WLANs

Our proposal is simplicity itself: use the mechanism of SIP and SDP over the existing NPCS networks as a common control channel for terminating calls on mobile hosts that are capable of IEEE 802 WLAN connectivity. In order to explain this proposal, it behooves us to provide some information on the structure and function of a NPCS network.

NPCS is a natural extension of traditional paging into fully bi-directional data communication. Hence, NPCS devices have many of the same characteristics of traditional paging units: low cost, excellent in-building penetration, broad coverage, and true “push” for content delivery. In addition to these starting points, NPCS adds the ability to exchange binary content with arbitrary



Internet connected hosts by name or IP address using an address formalism that subsumes both the notations of (RFC 822) email and Uniform Resource Locators (URLs). In short, it is feasible to send arbitrary binary or alphanumeric content to a NPCS mobile at a specified port or path and to declare that content to be of a specified class.

NPCS networks transparently manage the mobility of devices and present their identities to the Internet in terms of account identifiers, which may be user names (“john.doe”) or phone numbers (“NPA NXX XXX”) or PINs (“XXX XXXX”). In this way, it is quite possible to address traffic to a common gateway to the Internet via HTTP push, formatted in accordance to an open extensible markup language document type definition (XML DTD), and deliver virtually any class of content whatsoever to be forwarded to any port or path at the mobile host.

In particular, the delivery of an SDP description for an Internet session is almost immediately feasible. Under the assumption that the destination unit has the capability to use any form of return Internet path whatsoever, it can respond to the calling party within the prescribed time-frame to complete the call.

So, a typical call setup proceeds in this manner: the calling party addresses an SDP message to the called party through the NPCS network. The called party has or obtains an IP presence (of any form whatsoever) and replies to the calling party via the Internet. The calling party authenticates the called party, and the call proceeds. In the absence of normal call completion, redirection to voice mail, call forwarding, or other services, may easily be implemented per the mechanisms of SIP/SDP.

## Identifying the calling party

There may be a number of circumstances under which the calling party does not know its return address. These are similar to the situations in which the called party does not know its address; for example, its IP address is within some section of private space and translated onto a public address at a corporate point of presence.

In other words, the caller’s host is behind a firewall and it is essentially impossible to open a session directly to such a host. In this case, although the calling party could put their IP address in the SDP that they send to the called party, it would be of little or no use in delivering a response.

Within the context of SIP, this situation is solved by having the calling party work through a local proxy. The proxy handles the calling party’s SIP invitation to the called party; the proxy acts as the caller’s agent in setting up the request. The proxy uses its IP address for the reply path and will interpret the routing information necessary to forward a response from the called party to the calling party. For example, if the calling party is identified at the SIP layer as “< sip:allan@weblink-wireless.com >”, the return path would first route to the SIP proxy at weblinkwireless.com and from there to the user “allan”.

In a case like this, it would be essential to allow the NPCS network to inter-work with the calling party’s SIP proxy. To do so, the receiving NPCS gateway would most likely present itself as another SIP proxy. The destination address for the called party would then be presented at the SIP layer as; for example, “< sip:NPANXXXXXX@npcs.net >”. The opaque payload of the SIP INVITE would remain SDP describing the session. The FLEXsuite Uniform Addressing and Routing protocol would support the delivery of the pure SIP alphanumeric content with the SDP payload to an appropriate port at the called device.

Of course, the same could be achieved if the calling party has and knows its public IP address. That is, the calling party could send SIP directly to an NPCS SIP Proxy also. The main point here is that there are two potential ways to deliver SDP via NPCS: first, as an opaque payload via WCTP (XML over HTTP) to the called party; and second, via a SIP Proxy for the called party.

## Extensions to traditional paging

While ReFLEX is broadly deployed throughout the United States, Canada and Mexico, it is “thin on the

ground” elsewhere. However, more traditional FLEX and POCSAG paging systems are extensively available in most parts of the world. Both of these systems can also support the forward channel delivery of alphanumeric content of the sort represented by SDP. This would enable a variation of the full 2way NPCS service to be accomplished where only 1way paging was available.

The comments of the previous section concerning the delivery of SIP and SDP together apply in common to both FLEX and ReFLEX devices, since both support the FLEXsuite protocol stack.

In such a case, the natural return path would be the Internet. Also, an international roaming service could be constructed by borrowing part of the mechanism of SIP. The use case would run as follows: A roaming subscriber from North America visits, say, Taiwan, where FLEX, but not ReFLEX is available. Using an Internet connection from the airport, hotel, or dialup (it really doesn’t matter), the customer manually registers their current location. A return Internet message reconfigures the paging receiver to manage the local FLEX channel. Now, when Internet SDP arrives at the home carrier’s gateway for this subscriber, it is simply forwarded to the Taiwanese paging service provider, who transmits it by FLEX paging to the customer. The customer responds to the Internet SDP invitation by means of Internet pathways.

## **Advantages to this approach**

---

This approach has a number of advantages over any of the alternatives, which are as follows.

*1. No need for SIP Proxies or Registrars.* In this proposal, there is no requirement for the broad deployment or management of SIP Proxies or Registrars. In SIP as presently envisaged, if the local network does not have a SIP Proxy, roaming is not supported. In this proposal, the local service does not need to have deployed SIP for SIP to work. Likewise, the home network does not have to maintain a Registrar for its mobiles to be reached.

*2. Improved security.* In the SIP service, virtually any asking party can obtain the current IP location of a roamer from the Registrar. In the current proposal, the user’s

gateway and identity are invariant to roaming. Calling parties are identified to the called party, who can choose to answer directly or not. In any case, a user’s IP address, location, or presence is not available merely for the asking.

*3. No need for an always-on Internet connection.* In the SIP model, only hosts with an active Internet connection can be reached. In the present proposal, the called party can obtain an Internet connection and IP address after the SDP session invitation is received. For situations in which IP connectivity is expensive; for example, billed by the minute, the economic value of this should be apparent. If the user cannot immediately establish an Internet connection, but wishes to acknowledge the session invitation, or negotiate an alternative time, a response message can be sent via the NPCS network. In fact, neither party needs to have an Internet connection at the time of call negotiation, if both parties have also got NPCS capabilities.

*4. Battery saving.* Since NPCS technology is extremely battery-efficient, and since there is no requirement to support an always-on Internet connection in order to manage Internet call termination, the battery consumption of a mobile can be reduced to a minimum.

*5. Security (Part 2).* NPCS standards support an elliptic curve public key infrastructure. By adding a challenge-response component to authenticate the called party, the calling party can perform its own authentication of the replying party. This authentication can occur not only at the initial call setup, but also during the call, and in particular, at intersystem switch-overs where the source IP address of the mobile may change.

*6. Remedial action for link loss or no coverage.* In the event that a call in progress is lost because the mobile cannot restore IP network service, either party can employ the NPCS network to send SDP to remediate the loss of the call. Naturally, in cases where the exchange of simple text messages can complete the call, the NPCS network can support the information exchange. Likewise, if the mobile unit has no Internet access at all, there may be cases where a simple information exchange suitable for the NPCS channels can suffice, or can direct the mobile user where information can be recovered later. For

example, if the calling party intended to send a file to the called party, the calling party can instead deposit the file on some mutually accessible server where the called party can recover it later. If the server were publicly accessible, the sending party could also supply access keys or tokens (for example, Kerberos tokens) to enable only the proper party to gain access to the file. Such exchanges would easily be accomplished via NPCS.

7. *Authentication, MM, & OAM&P.* As mentioned, NPCS networks support an ECC PKI for devices. By making the user's account on the NPCS network a clearing point for the billing of roaming-charges, users can obtain, in effect, instant credit for WLAN or other roaming. By billing-back verified roaming service to their home NPCS account, users gain the advantage of an instantaneous introduction to a broad range of potential service providers with whom they would otherwise have no direct business contact or credit arrangements.

Consider a mobile-directed switch-over model. If the roaming host, during a call, scans for alternative system opportunities and presents its credentials for roaming to a new target system before any switch-over is necessary, that target system can verify those credentials against NPCS key-servers, which would hold signed public keys of the users with NPCS accounts. Later, if the mobile deems a switch-over to be necessary, it can obtain a new IP address in the new system and forward a signed authentication token to the far end host to validate the IP address "take-over." Admittedly, there is a present generation of stateful firewalls that will have problems with such a scenario; however, these are not used every-

where. There is also no reason that a stateful firewall itself could not verify the credentials of a mobile to complete a connection take-over; this is simply not done at present.

8. *Billing & clearing.* In the current generation of cellular networks, call detail records (CDRs) are kept for visiting mobiles. These are later delivered to a clearing house, which tallies the net payables and receivables of the participating carriers on the basis of the aggregate roaming traffic to and from each of the networks. To avoid fraud, the visited network can authenticate the roamer by presenting it with a time-stamped "ticket" for service. The roamer validates this ticket by signing and returning it. Signed service tickets can be forwarded to the clearing house by visited systems as part of the verification that their payables claims are accurate.

---

## Summary

This white paper presents a modest proposal for the use of NPCS networks as a common control channel for a general model of roaming Internet devices, including specifically IEEE 802.11x units (but in no way limited to them.)

The proposal depends upon the use of a session description mechanism like SDP over SIP to carry session "invitations" to target roamers. As well, the target roamers would need have dual or multi-mode connection capabilities; for example, NPCS and an Internet connection method. In practice, the advantages of the proposal over the obvious alternatives are numerous

and include better battery life, reduced infrastructure, improved security, mobility management, and OAM&P functions.

## Bibliography

1. GSM 03.60 *Digital cellular telecommunications (Phase 2+) General Packet Radio Service (GPRS) Service description; Stage 2, version 7.4.1*, Figure 2. ETSI 1998.
2. GSM 03.60 *Digital cellular telecommunications (Phase 2+) General Packet Radio Service (GPRS) Service description; Stage 2, version 7.4.1*, Figure 49. ETSI 1998.
3. GSM 04.60 *Digital cellular telecommunications (Phase 2+) General Packet Radio Service (GPRS); Mobile Station (MS) - Base station System (BSS) interface; Radio Link Access Control (RLC/MAC) protocol* Figure 1, ETSI 1999.
4. ETSI TS 101 350 *Digital cellular telecommunications system (Phase 2+); General Packet Radio Service (GPRS); Overall description of the GPRS radio interface; Stage 2 (GSM 03.64 version 8.5.0 Release 1999)*, Chapter 6.
5. 3GPP TS 22.174 V6.2.0 *Push Service; Service aspects; Stage 1*, Release 6, 2003-03, <http://www.3gpp.org/ftp/Specs/html-info/22174.htm>
6. *Session Description Protocol*, IETF RFCs 2327 & 3266, <http://www.rfc-editor.org>.
7. *Session Initiation Protocol* IETF RFCs 3261 through 3265, <http://www.rfc-editor.org>.
8. SIP Forum web site, <http://www.sipforum.org>.
9. SIP Tutorial web site, <http://www.iptel.org/sip/>
10. IEEE Wireless Standards, <http://standards.ieee.org/wireless>
11. *Simple Network Management Protocol v3*, IETF RFCs 3410 through 3418, <http://www.rfc-editor.org>.
12. EDS clearing house site, [http://www.edsch.com/services/cont\\_serv.html](http://www.edsch.com/services/cont_serv.html)
13. Cibernet clearing house site, <http://www.cibernet.com/>
14. *FLEXsuite™ of Enabling Protocols*, v2.1.1, Chapter 4.0, UAR, Motorola, 1997.



Allan Angus  
TVP Architecture  
WebLink Wireless

For more information, contact:

Allan Angus, PE,  
TVP Architecture  
E-mail: [allan.angus@weblinkwireless.com](mailto:allan.angus@weblinkwireless.com)  
Phone: (214) 765-3470  
2-Way: (800) 970-9325