# MOBILITY AND APPLIANCE NETWORKS – A NEW MOBILITY MODEL

Bob Askwith, Madjid Merabti and Anirach Mingkhwan
*School of Computing and Mathematical Sciences, Liverpool John Moores University, Byrom St., Liverpool, L3 3AF.*

Abstract: In this paper we propose a novel network architecture, called Node-Centric Networking, that challenges current thinking about mobility in IP-based networks. We are motivated by the increasing proliferation of small personal mobile devices such as PDAs and mobile phones that people carry and the desire to support the development of novel applications for these devices, including the potential for interaction with a new class of device; Networked Appliances. Although IP host mobility now has a long research history to draw on, we are not convinced current solutions meet our vision for personal mobility services that will better enable new applications to be developed. Our challenge to present a new architectural model is based on creating networks of services that are centred on the user and the tasks they wish to perform. It is a form of Ad Hoc networking, though our approach is novel and we openly desire integration with, though not reliance on, fixed Internet infrastructure. Our key innovation is the unlinking of device mobility and service mobility, using IP as a separator. The paper first presents the overall vision for personal mobility services and explains some of the networking and application challenges and the work we have done to solve them.

Key words: Internet; IP; mobility; network architecture; ad hoc network; personal area network; networked appliance.

## 1. INTRODUCTION

In the space of around a decade personal computing has moved from being a static desktop activity to one dominated by connectivity to the Internet. Increasingly this connectivity is being utilised via wireless links and via devices far removed from the humble desktop PC. The initial technical driver for popularising the Internet was the invention of the World

Wide Web, an application that allowed easy to use access to online information across the globe. In parallel we became used to the idea of wireless and mobile communications through the introduction of $2^{nd}$ generation telecommunications systems, in Europe known as GSM.

In the last five years the lines have begun to blur between these previously separate worlds. GSM began a service expansion as part of the evolution towards $3^{rd}$ generation (3G) systems, a phase called '2.5G'. The first significant development was the deployment of the Wireless Application Protocol (WAP) (Erlandson and Ocklind, 1998), allowing a 'micro-browser' in a terminal to access modified Internet content, WML rather than HTML, through a service provider gateway. Two other significant 2.5G activities have been the roll out of a packet service called GRPS (Kalden et al., 2000) that improves Internet performance for phones, and the sudden take up of camera phones and picture messaging via Multimedia Messaging Service (MMS) (Novak and Svensson, 2001).

This process of convergence has been added to by a number of wireless communications standards for computing devices, most notably the Wireless LAN standard IEEE802.11 (IEEE802.11), essentially Ethernet without cables. WLAN has proven particularly helpful in disconnecting users from the desktop PC and enabling them to perform the same tasks from a laptop/notebook PC. The final piece of the current picture is the explosion of consumer electronics 'gadgets' that have appeared on the market, many with connectivity of some kind. Popular gadgets today include digital cameras, portable MP3 music players, and personal organisers (PDA), all of which need connectivity to transfer information, most commonly to the desktop PC which is still the principle access point to the wider Internet. The important observation we are making is that connectivity and computing are becoming generic activities, not associated with one particular device, application, or location.

The only apparent access restriction the architecture of the Internet puts on devices is that they must be able to run the Internet Protocol (Postel, 1981). Having observed the trend above it is reasonable to suggest that the range of devices owned by a user can simply connect to the Internet to give the required connectivity. One problem with this vision is the mobility of terminals, since IP assumes a static location, and yet the purpose of many new devices is their very portability. This problem is not new; indeed there is over 10 years of research to look back on. In this paper we are not concerned with the problem of wider Internet mobility. Instead, we are looking at another problem – the architectural assumptions of IP as it applies to the Internet (Carpenter, 1996) and how this affects the building of mobile applications based on IP.

The remainder of the paper is structured as follows; section 2 discusses the requirements for consumer mobile appliances that use IP-based applications and services. In Section 3 we present a new architecture in which these requirements are achievable. A discussion of the architecture focusing on the mobility issues is given in section 4, while section 5 examines remaining work and draws some conclusions.

## 2.       REQUIREMENTS FOR APPLIANCE MOBILITY

The Internet, to the ordinary person, is really no more than a very big computer program, albeit a very useful one. The bedrock of that application is the network protocol called IP (version 4). A user who wishes to send or retrieve data over the Internet does so by transmitting and receiving IP packets. Although user applications need IP to use the Internet they are decoupled from being a part of the Internet. This distinction is important and it means that the same application can run on any network that deploys IP. For example Intranets are an IP-based network commonly found inside an enterprise. Applications inside an Intranet may use the same application software as they would do on the Internet, give or take some reconfiguration.

The common properties to these traditional applications of IP-based internetworking are as follows:
- All nodes on the network are assigned a single IP address that is unique to the network; regardless whether they are likely to ever communicate. Node IP addresses must be unique in order to guarantee that packets can be delivered to the correct node.
- IP is used to create a network of networks, with users at the edge and a core that supports the connection of these networks. Edges are often 'private' (e.g. enterprise LANs) and the core is often called 'public' as it is shared between the private areas. The core is made up almost entirely of routers whose main task is to transport packets from edge to edge.
- Users do not manage or administer the network; this is left to third parties. The third party may belong to the same organisation in the case of edge networks and Intranets, or may be an external third party in the case of the core networks.
- No matter what application a user is running, the network looks the same due to the common use of IP address for each task – the IP address is per node, not per application. Activity in the network is based on connectivity rather than application need. This, combined with the other properties creates a network-centric architecture.

In the two examples of IP-based networking mentioned above, Internet and Intranet, the network-centric architecture is geared towards a rather

static sharing of varied resources between many people and some semicentralised control and administration. This model of 'connecting to the network' is clearly not effective for connecting user appliances. First, the networks are smaller, made up of mostly local resources required to perform a particular task. Second, the administrator is the user of the application, a user who is often likely to lack the competence to administer an IP-based network. Third, the very nature of appliance computing implies that the networks formed in the course of an activity are likely to be temporary (i.e. for the duration of the task) and dynamic if the user moves or requires a new resource during execution.

The Internet architecture works well for distributed applications that use geographically separated resources. What is really needed is an architectural model that more closely supports appliance computing, still using IP as a basis. This model should have the following properties:

- An IP network need only exist for the duration of an application execution. Each node within this network must have different IP address, unique within the network. This is normally called Ad Hoc Networking, but we will see later why our vision represents a special case.
- The Internet should be accessible as an application resource. This requirement requires some kind of gateway translation function to hide the new IP network from the Internet.
- The user application should represent the core of the network with the associated resources being at the edge. This allows the management and configuration aspects to be controlled within one node. This node-centric architecture enables resources to be connected as and when the application needs them.
- Nodes can be part of more than one application network. This results from the previous requirement that an application is the core of its own network. Since a user may perform many tasks the implication is that many networks may be created. Conversely, the resources on the same node may become part of another network on another node. A special case of this problem is when two or more users create a peer-to-peer (P2P) network. In the P2P case the network is created co-operatively.

In summary we require an architectural model that is centred on the needs of the user application rather than an edge node sharing a larger shared infrastructure. The model needs to be dynamic such that only the required resources are part of the network and the node is responsible for managing this dynamic situation. The Internet is seen as a separate resource within the model accessible through a gateway function. Nodes can manage simultaneous application networks and can co-operate with other nodes to provide service to another application network. We call such a vision Node-

Centric Networking. The next section presents the initial results of our work towards this vision.

## 3.     A MODEL FOR APPLIANCE MOBILITY

The fundamental principle of our proposed Node-Centric Networking model is to allow a single device or application to draw together all the resources it requires into one single-hop IP network. The significance of single-hop IP is that routing is removed from the functionality of layer 3, where needed multi-hop communications is handled at layer 2. This allows the IP layer to be more concerned with service discovery than with internetwork packet delivery. These IP networks operate in an ad hoc manner except that one node is always in control of configuration and management. We will call this controlling node the Application Control Node (ACN) and all other node Application Resource Nodes (ARN). The ACN must handle a number of operations at different layers of the stack in order to function properly, we shall the layers from bottom to top.

### 3.1     Layer 1/2

We eschew providing multi-hop IP networking at layer 3 in favour of a lightweight layer 3 that provides one logical L3 hop between nodes. At L2 multi-hop communications may or may not be available depending on the link layer being used. In the case of Wi-Fi we have developed a way of connecting networks together up to 5 hops away, yet appearing as one hop to IP (Mingkhwan et al., 2003a). Bluetooth may similarly provide a multi-hop network by forming a scatternet topology (Haartsen, 2000). We are also looking at wider problems of L2 standards interoperability, and in particular ad hoc bridging (Mingkhwan et al., 2003b), to support more flexible 1-hop L3 networks. This bridging and multi-hop functionality is combined in the architecture into a network-connectivity sub-layer between L2 and L3 (i.e. below IP but using the services of each respective L2).

This hiding of the physical structure from L3 makes sense in an environment where resources are likely to be used in close proximity to each other (e.g. a printer will usually be required to be near the user who is printing the document). Here IP is simply used as a convenient glue to connect different resources together and to provide a compatible protocol for applications to use. Nodes that are mobile have their basic connectivity managed at this layer. Failed connectivity can be managed at higher layers, which may then choose to re-map resources or whatever form of adaptation is most appropriate. We will return to these ideas later.

## 3.2    Layer 3

At the IP layer the ACN is responsible for allocating IP addresses to all the ARN it wishes to communicate with as part of the applications. This allocation is performed using non-Internet routable address space, so-called 'private' addresses. A part of this function is to make sure that there are no address conflicts with other nodes in the area, including those in different application networks. If there is conflict then a resolution phase takes place; this also takes care of hidden terminals.

An IP network is first created when a node is switched on. The node selects an IP address at random and broadcasts using a modified ARP to determine if this IP is already in use. If there is no response the address is good, otherwise a node will inform the new node that the IP is already in use. This will trigger a repeat of the process until a good IP is chosen. This node becomes the ACN of a new network.

When nodes are mobile their IP address could conflict with other devices that come within range. As mentioned in the previous section this problem is dealt with at L2, but layer 3 must be able to provide transparency to higher layers that require a constant IP address, e.g. TCP. To solve this we propose a per node address translation service to work at the IP layer. The translation service simply makes sure that on one hand the correct address is being used in communications, and on the other it remains constant during the application execution. When a node providing a service moves or stops providing its service, and another service is remapped from another device then the per-node translation will hand over the old IP to the new node/service. In this way higher layer protocols are protected from address changes.

## 3.3    Transport layer

Since the user application maintains a static IP address for the duration of the application execution the transport layers operate as normal, though optimised versions of TCP may be employed to suit the environment (Pentikousis, 2000). Session and Presentation layers may be present unaltered, e.g. if devices are part of WAP. The case of WAP is an interesting one as this allows device profiles to be shared via UAProf (Arnby et al., 2001). Where re-mapping of services occurs then this may prove useful.

## 3.4    Application layer

Application layer protocols such as HTTP can run unaltered in our architectural model. Resources that require access out to the wider Internet

will be redirected via the gateway function; we have chosen NAT (Egevang and Francis, 1994) to perform this task. The presence of a NAT gateway to the Internet might typically occur in a local access point, e.g. an 802.11 node. An ACN can access this point directly or it could be redirected via an ARN that is advertising Internet access as a service, e.g. a desktop PC belonging to the user.

More interestingly at the application layer we have been able to develop a service composition framework for appliances using the lower layer features of the architecture to build distributed applications (Mingkhwan et al., 2004a). Devices are able to advertise their services to their environment using a web services description language, DAML-S (Ankolekar et al., 2002). Applications can then compile and adapt their required functionality from amongst the services on offer. Users can configure an application with profile policies so that the best alternatives can be found. For example, a home theatre application would involve some combination of playback device, a display device, and an audio device. The profile policy represents the semantics of the application, and is therefore able to use the results of service discovery to attempt to build that application. Although we have not yet extended the prototype to cope with mobility, the architecture should support the profile policy being able to re-map functions to different devices.

## 4.      DISCUSSION

This section provides some discussion on how the architecture presented challenges IP mobility. In the introduction to the paper we noted that current IP-based applications are usually seen in the context of the Internet, i.e. an architecture involving edge subscriber nodes (users) and core access 'infrastructure' networks as the model of access. The main purpose of IP in the Internet is to provide unique addresses to each node, and to provide a routing infrastructure to join nodes together.

We argue that for networking personal appliances there is no need for an IP-level routing infrastructure, or indeed unique (or static) addresses for each node, but rather for IP to provide a common interface for applications and services, one particular service being the Internet. In this discussion we consider solutions that lie between the two mobility models, i.e. traditional static IP and our node-centric solution and argue that our proposal best suits the domain of personal mobility services.

## 4.1      Internet mobility

The first efforts at creating mobility in the Internet resulted in the Mobile IP protocol (Perkins, 1998). In Mobile IP packets destined for a permanent IP address are redirected at the Mobile IP-aware network gateway (home agent) towards a temporary (care-of) address at a Mobile IP-aware foreign network (foreign agent). IP version 6 (Stallings, 1996) contains a more streamlined version of this mobility model (Johnson and Perkins, 2003), although standardisation is still taking place. The other significant form of Internet mobility support is called Micro-mobility (Campbell et al., 2002), and is concerned with managing movement within a LAN segment, and in particular aimed at handling multimedia connection handover. These two solutions are extensions to the infrastructure model of the traditional Internet, allowing mobility at the edges of the infrastructure. Their purpose is to maintain IP-level routing across the network. We would advocate the use of Internet Mobility as part of the Internet Service, but it is inappropriate for our model since routing is not required at the IP layer.

Another mobility solution being pursued is the IETF NEMO work (IETF, 2004). This takes the notion of node mobility a step further and allows the access point or router to be mobile, thereby allowing an entire network to be come mobile. Whilst this may have some application to our vision, it is still a network-oriented view, rather than a user/service-oriented view.

## 4.2      Ad hoc networking

In recent years, spurred on by the activity in mobility research, interest has grown in infrastructure-less networks, also called Ad Hoc Networks (Chlamtac et al., 2003). Our model utilises the idea of ad hoc networking, but we do so in a novel way. In current solutions researchers are tackling ad hoc routing problems to enable connectivity between cooperating nodes. It is our argument that in appliance networking the control of the network should belong as close to the user as possible, since they are operating the application network, and the user application should form the network based on the services it needs to use. Similarly, physical network connectivity need not be provided by the IP layer, but instead handled at the link layer. Initial simulation results suggest that this model compares favourably in performance with IP routing connectivity (Mingkhwan, 2004b).

So, our model clearly advocates an ad hoc approach to networking, but the flexibility comes from being able to use IP just as a service-handle to applications. This means that the underlying model becomes transparent to applications, which simply request service from the network. Device mobility, and resulting network connectivity, is handled between the link

and network layers, in a bridging/multi-hop sub-layer, independently from the application activities. From the application point of view service mobility becomes a separate issue (previously service and device mobility were IP issues), handled either in the application or in a sub-layer below. This separation gives us great control flexibility, since we can focus control over services independently of the connectivity issues. Services mobility can involve both the device and the service moving from place to place, yet a constant IP address can be maintained in a relatively simple way.

## 5. CONCLUSIONS AND FURTHER WORK

In this paper we have challenged the orthodoxy of IP as used on the Internet. We propose that in many personal networking scenarios a user is controlling an application from one device and using the resources of other local devices. In these scenarios a single-hop IP network that is user-centric makes the most sense. In such a model IP addresses are treated as a resource under the control of the central node (i.e. where the user is controlling the application). We called such a vision Node-Centric Computing.

Different applications that have created different IP networks around them may overlap with each other. In these cases the networks are able to ensure that conflicts of addresses are solved. Mobility is handled at both the link layer and at the IP layer, though for different purposes. Link layer mobility allows a device to travel outside of the range of a direct link by configuring a multi-hop link layer network. Mobility at the IP layer uses a translation service within the central node that makes the movement of services and nodes transparent to the application above.

We have produced many results in this work so far, but there is still a long way to go. Our achievements so far include layer 2 mechanisms for multi-hop wireless using Wi-Fi, Bluetooth interoperability, IP address configuration for use in the architecture and an application semantics framework that allows services to be composed according to service descriptions and policy profiles. This allows applications and services to be mobile within the architecture, whilst all the time maintaining use of IP.

Amongst the further work remaining, apart from completing the prototype work of the ideas presented here, include the development of a security model to fit the personal mobility environment. Security models are bound up in the environment they are applied to, and this environment is very different to the Internet. We are currently studying a component composition model that preserves the security of an overall system (Llewellyn-Jones et al., 2004). We need also to examine how certain applications can cope with multimedia and mobility. The architecture we

have described is relatively lightweight, but the prospect of adding in full QoS support appears daunting, yet is likely to be important.

## REFERENCES

Ankolekar, A., et al., 2002, DAML-S: Web service description for the semantic Web", in: *Proceedings of 1st International Semantic Web Conference (ISCW)*, Sardinia, Italy.

Arnby, P., Hjelm, J., and Stark, P., 2001, WAP 2.X architecture – features, services and functions, *Ericsson Review* 5(4):178-183.

Campbell, A. T., Gomez, J., Kim, S., Wan, C.-Y., Turanyi, Z. R., and Valko, A. G., 2002, Comparison of IP micromobility protocols, *IEEE Wireless Communications*, 11(1):72-82.

Carpenter, B., 1996, Architectural principles of the Internet, IETF RFC 1958; http://www.ietf.org/rfc/rfc1958.txt

Chlamtac, I., Conti, M., and Liu, J. J-N., 2003, Mobile ad-hoc networking: imperatives and challenges." *Ad Hoc Networks*, 1(1):13-64.

Egevang, K., and Francis, P., 1994, The IP Network Address Translator, IETF RFC 1631; http://www.ietf.org/rfc/rfc1958.txt

Erlandson, C., and Ocklind, P., 1998, WAP – the wireless application protocol, *Ericsson Review* 2(4):150-153.

Haartsen, J. C., 2000, The Bluetooth radio system, *IEEE Personal Communications* 7(1):28-36.

IEEE 802.11 Wireless LAN Standard; http://grouper.ieee.org/groups/802/11/main.html

IETF, 2004, Internet Engineering Task Force Working Group on Network Mobility (IETF NEMO); http://www.ietf.org/html.charters/nemo-charter.html

Johnson, D. B. and Perkins, C., 2003, Mobility support in IPv6, IETF Internet Draft; http://www.ietf.org/internet-drafts/draft-ietf-mobileip-ipv6-24.txt

Kalden, R., Meirick, I., and Meyer, M., 2000, Wireless Internet access based on GPRS, *IEEE Personal Communications* 7(2):8-18.

Llewellyn-Jones, D., Merabti, M., Shi, Q., and Askwith, B., 2004, Utilising Component Composition for Secure Ubiquitous Computing, in: 2nd UK-UbiNet Workshop - Security, Trust, Privacy and Theory for Ubiquitous Computing, Cambridge, UK.

Mingkhwan, A., Merabti, M., Askwith, B., and Hanneghan, M., 2003a, Global Wireless Framework, in: *Proceedings of European Personal Mobile Communications Conference (EPMCC'03)*, Glasgow, Scotland, UK.

Mingkhwan, A., Abuelma'atti, O., Merabti, M., and Askwith, B., 2003b, Multihop communication in global wireless framework, in: *Proceedings of IEEE Wireless Communications and Networking Conference (WCNC'03)*, New Orleans, USA.

Mingkhwan, A., Fergus, P., Abuelma'atti, O., Merabti, M., 2004a, Implicit Functionality: Dynamic Services Composition for Home Networked Appliances, in: *Proceedings of IEEE International Communications Conference (ICC 2004)*, Paris, France.

Mingkhwan, A., 2004b, An Integrated Personal Mobility Services Architecture, *PhD Thesis*, Liverpool John Moores University, Liverpool, UK.

Novak, L. and Svensson, M., 2001, MMS – building on the success of SMS, *Ericsson Review* 5(3):102-109.

Pentikousis, K., 2000, TCP in wired-cum-wireless environments, *IEEE Communications Surveys & Tutorials* 3(4) only available online; http://www.comsoc.org/livepubs/surveys

Perkins, C. E., 1998, Mobile Networking Through Mobile IP, *IEEE Internet Computing*, 2(1):58-69.

Postel, J., 1981, Internet Protocol, IETF RFC 791; http://www.ietf.org/rfc/rfc0791.txt
Stallings, W., 1996, IPv6: the new Internet Protocol, *IEEE Communications Magazine*,
   **29**(7):96-108.