Exploiting Network Diversity in MARS – A Mobile Access Router System

Rajiv Chakravorty† Ian Pratt† Pablo Rodriguez‡

†Systems Research Group, University of Cambridge Computer Laboratory
Firstname.Lastname@cl.cam.ac.uk *e.g. Rajiv.Chakravorty@cl.cam.ac.uk*

‡Microsoft Research, Cambridge, U.K.
pablo@microsoft.com

Abstract— **In this paper, we make a case for exploiting the inherent network diversity in wireless access available from different wireless and cellular networks, and of the operators that can be used to provide a sufficiently sustainable and reliable wireless communication link. Based on a real-world example of cellular networks that are currently deployed in a (small) town in the U.K., we show that there is a substantial overlap in terms of coverage being offered by most operators (e.g. Vodafone, Orange etc.), and also currently to some extent even across networks (e.g. GPRS, 3G and/or WLANs). We advocate that diversity in wireless access from cellular and other types of wireless access networks can be exploited to aggregate bandwidths from multiple network interfaces simultaneously in order to provide a fairly reliable communication link. We also give experimental evidence on the impact mobility can have in such environments, and discuss why bandwidth aggregation using simultaneous wireless access in this environment is useful. We exploit network diversity in MARS - a commuter Mobile Access Router System – that we are currently building for providing *on-the-move* mobile Internet access. We conclude the paper discussing research challenges that we anticipate in the design and implementation of MARS.**

I. Introduction

There is a strong growth in mobile Internet access, fuelled by the increasing popularity of *WiFi* (*i.e.* IEEE 802.11b-based WLANs), and the worldwide deployment of wide-area wireless networks such as 2.5G GPRS and third generation wireless (3G). Multi-mode devices (e.g. WLAN-GPRS cards) are becoming increasingly affordable, and a growing number of mobile devices such as laptops, PDAs and handhelds are equipped to connect to multiple networks.

With the proliferation and ever decreasing costs associated with wireless access devices, providers are increasingly looking towards the practical issues of service deployment and performance gaurantees. Mobility that involves handovers between Wi-Fi 'hotspots', 2.5G and 3G wireless data services continue to pose a significant challenge, as does the intelligent manipulation of channels and multiplexing/striping of data across available wireless links to achieve the best possible performance and access under variable and often unpredictable conditions.

In light of this, it is important for us to identify the challenges in building reliable wireless communication systems:

Wireless Link-related Problems. Cellular networks in the wide-area such as 2.5G General Packet Radio Service (GPRS) and 3G (e.g. UMTS, CDMA2000 etc.) promise users *always-on* connectivity in the wide-area. However, real experiments conducted over production networks (e.g., GPRS, CDMA2000) does indicate that such links are currently plagued with several problems such as high and variable round trip times, patterns of burst packet loss, frequent link outages, and significantly lower bandwidths than originally claimed (see, [8], [12]). In other words, it seems that there is currently no wireless technology in the wide-area, that can offer the level of reliability expected out of a communication channel. This means that we either have to adapt protocols and applications to work over such links (e.g., [8], [9]), or work for some other more practical alternative.

Spectrum Limitations. Wireless networks are spectrum constrained. Cellular network (and WLAN) operators are only allocated a limited amount of bandwidth. This fixed bandwidth enables them to only support a limited number of subscribers in each service area (cell). However, increasing the data rate for each subscriber is a trade-off against the number of subscribers the service area can support. This situation is particularly exacerbated when subscriber density and application/content size increase at the same time. Furthermore, licensing laws and competition in this space only adds to this problem [2].

Lack of Real Systems Exploting Diversity. Most current communication systems are single input single output (SISO) systems; such systems cannot afford to exploit enough diveristy because of the use of only one transmitter and receiver able to operate over the communication channel. As a result, the ability in SISO to provide sustainable data rates and reliable wireless channel poses a significant challenge. However, spatial domain solutions can result in considerable improvement in wireless system performance. Techniques that exploit spatial diversity (e.g., Tx/Rx diversity, beam forming, MIMO systems etc.) are proven techniques to improve wireless system (and link) performance. Unfortunately, current wireless systems have yet to exploit such techniques to improve performance in wireless [2].

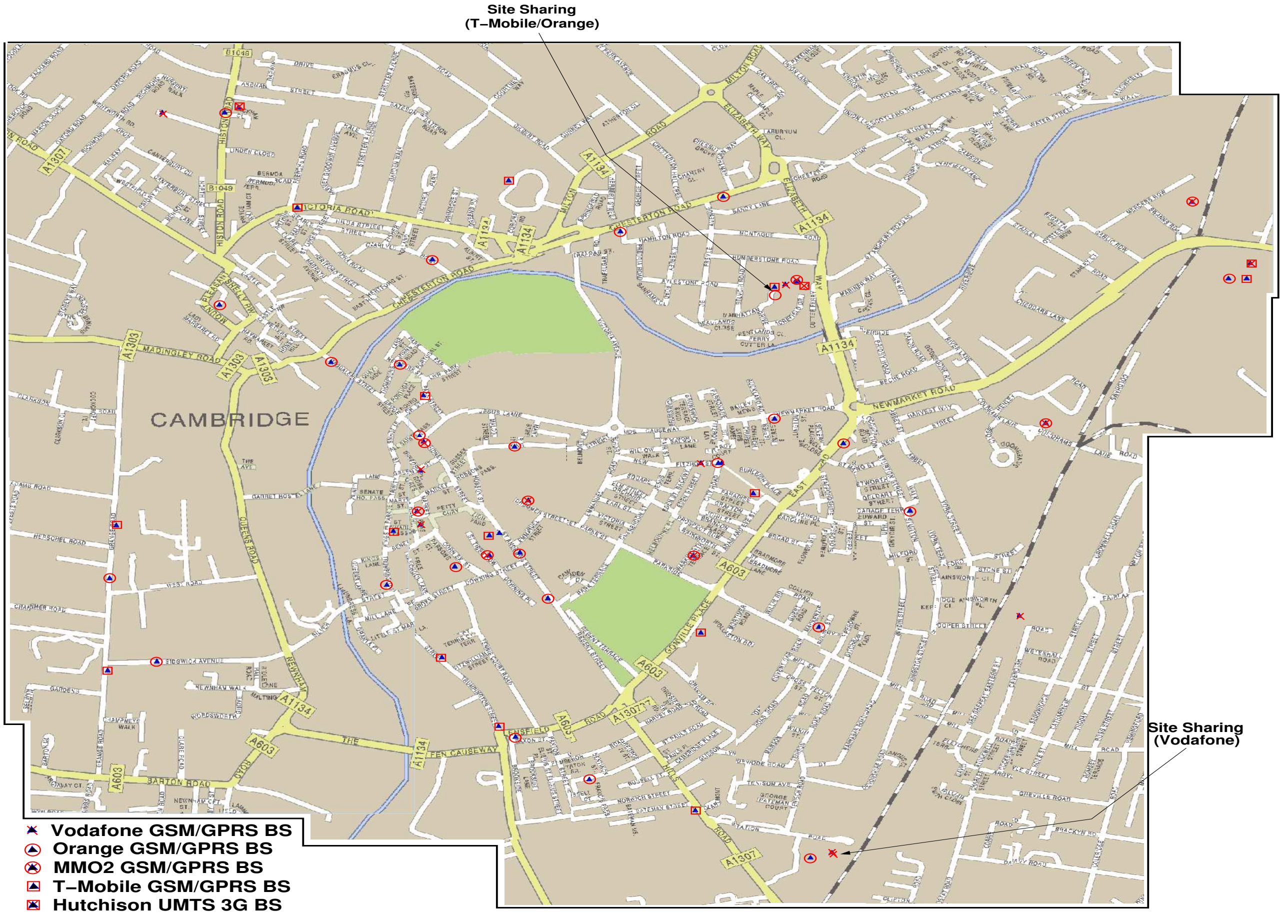

Fig. 1. Cellular Network Diversity in Cambridge, U.K.. The map gives the location of Base Stations (BS) of different cellular network operators in Cambridge. as of July 2003 (source: http://www.sitefinder.radio.gov.uk/). Due to substantial overlap in the coverage offered by the cellular operators in the U.K. and other European countries, ability to exploit the network diversity using simultaneous wireless access has the potential to provide highly-sustainable wireless communication channel in the wide-area, even when some links may suffer from link-outages and loss effects. MARS benefits from network diversity.

After identifying some of the challenges, we feel that it may be difficult (at least in the near terms) to realize a wireless communication system that uses a single air interface, and is still able to cater to the requirements of all the mobile applications. Instead, we argue for a more practical case in this paper, where we advocate use of multitude of air interfaces simultaneously, for providing a fairly reliable wireless communication channel (link). Such a link can provide sustainable data-rates and will be able to cover different mobile application requirements as well as mobility scenarios. To that end, we advocate exploiting network diversity from different wireless networks and operators to be able to aggregate bandwidths that can then be offered as a single large, stable pipe to the end users.

We exploit use of network diversity in MARS – a commuter Mobile Access Router System that we are implementing. MARS routers can be placed in a highly mobile (*on-the-move*) devices (e.g. car, bus etc.) able to perform bandwidth aggregation across network interfaces, and provide mobile Internet access to the MARS users. MARS exploits network diversity available from several wireless access networks simultaneously to provide a highly-reliable wireless link. We discuss more about MARS in section IV.

II. Exploiting Network Diversity in MARS

The ability to provide sustainable connectivity and data rates across multiple wireless access technologies available in the same mobile terminal is always a challenge. The challenge is made significantly more difficult because of the wide variety of environments (indoor, outdoor, moving, fixed, etc.) where a user can exist in and also move through.

Figure 1 shows wireless network diversity inherent in a cellular network infrastructure. The map shows the base stations of the cellular network operators located in the city of Cambridge, U.K. The cellular infrastructure in Cambridge comprises GSM/GPRS/UMTS base stations from 5 different cellular operators (Vodafone, Orange, MMO2, T-Mobile and Hutchison) that are located all over the city. An interesting point to note here is that number of base stations offered by the operators tend to have higher build-up near to the town centre. As such, coverage offered by the cellular operators in and around Cambridge is vitually pervasive, and hence, there is also a substantial overlap.

Altough this is the case for the wireless coverage made available in towns small or big, the same may not be the case while travelling in trains and/or along highways, or even in more rural locations (such as in the country-side). Of course, we do find

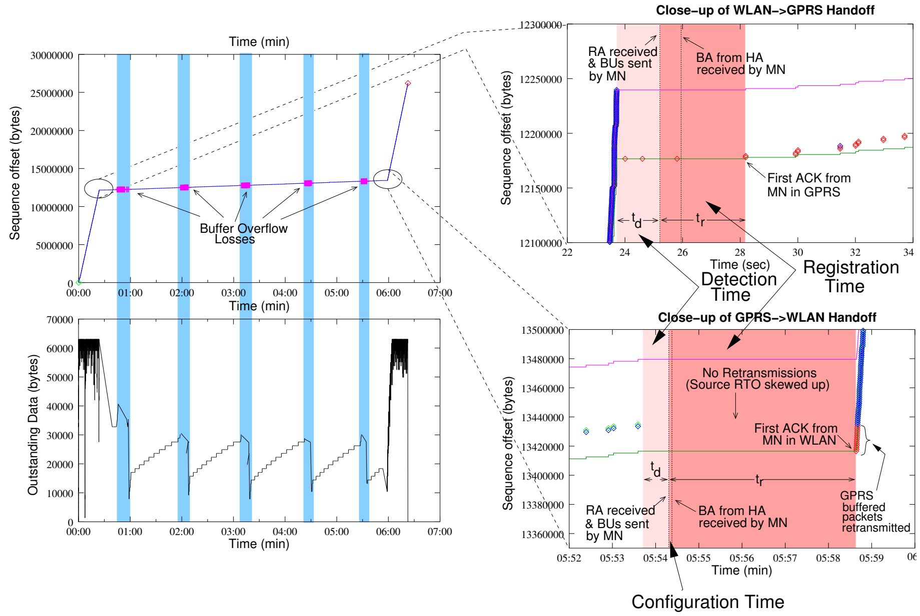

Fig. 2. Figure 2 show the time-sequence plot for vertical handoffs using Mobile IPv6 that involved transfer of a 25MB file over WLAN, with a vertical handoff to GPRS and then a reverse handoff back to WLAN. Top-right of figure 2 shows the close-up of WLAN→GPRS handoff, while bottom-right show GPRS→WLAN handoff. Managing mobility when there is significant disparity in the link-layer characteristics between different wireless links (WLAN, GPRS/3G etc.) will pose significant challenge. (Detailed analysis and measurements available in [7].)

instances of ‘holes’ in the coverage offered, since calls can be dropped often while travelling in a train or a car, and especially when passing through tunnels (in many tunnels in Europe, one can find reasonably good coverage). However, even if coverage from one cellular operator is lost, the same may not hold true for the coverage from other operators. In any case, most cellular operators usually attempt best (taking the quality of service and that service price into account) to be more omnipresent in terms of coverage. Apart from the GSM/GPRS base stations in figure 2, we can also locate base stations of UMTS 3G networks from Hutchison, which provides somewhat faster wireless data access than GPRS.

Based on extent of such wireless diversity available, we do find that we can make use of our MARS router to connect to number of wireless cellular networks simultaneously, and able to exploit the network diversity in the wide-area to provide a fairly reliable communication link. The advantage from such network diversity is apparent; network interfaces in a MARS router connect to the base-stations of different operators that are located differently (unless the site is shared, and which is rather rare) to ensure that it provides a single more reliable wireless link in the wide-area, even if some links had to suffer from link-outages and wireless-related loss effects. In this way, MARS can benefit from the inherent network diversity.

III. Experiences with Mobility

As we have discussed, network diversity can be exploited in a wireless overlay and across cellular networks for systems that will use multiple networks simultaneously, such as MARS. Even if we consider that coverage offered from cellular network infrastructure have substantial overlap, MARS’ ability to exploit distributed diversity from networks will be for most cases limited to perform good bandwidth aggregation. However, there might also be cases when MARS router will have to handover across different networks (e.g., GPRS, 3G, and WLANs). This is particularly true for the case when MARS has to exploit network diversity using WLANs. Since a MARS router is a highly mobile device, each network interface in a MARS will be usually associated with an IP address; however, there will be instances where network interface have coverage less usual (for e.g. moving *in-and-out* of ‘hotspots’ WLANs). For all such cases, we will need to analyse and understand the impact mobility can have on overall (vertical) handover performance.

We have studied the effects of mobility on vertical handovers, and have highlighted many such challenges with Layer 3 (IP) mobility [7]. Using a real GPRS–WLAN testbed, we have evaluated the impact hard and soft handoffs have on transport protocols such as TCP in a wireless overlay networks. Here, we account for some of our practical experiences using Mobile IP (in this case using Mobile IPv6) for layer 3 mobility. A thorough description of our practical experiences with GPRS–WLAN vertical handovers is available in the form of a separate technical report [7].

In these tests, we conducted vertical handovers between two networks – GPRS and WLAN – using a multimode mobile device located in a WLAN hot-spot and also under GPRS coverage. We then experimentally evaluated the impact of hard and soft handoffs on TCP performance using file downloads from a web-server. We elicit two most important observations:

Hard Handoffs Difficult: Optimizations doesn’t help much. Hard handovers that occur at layer 3 are usually more complicated; the mobile device stops listening on one interface (or a network) and simultaneously up (start listening from) the other.

As a result, packets that are already in-flight, or are destined (and those that already made it) to the previous network interface, unfortunately, are not read. These packets have to be retransmitted by the source, which leads to under-performance.

Figure 2 shows a TCP connection behaviour during vertical handoff between GPRS↔WLAN. We find that the time to handoff from GPRS→WLAN and WLAN→GPRS to be considerably high. A number of factors contribute to such high handoff latency, but mainly because of the disparity in the link-layer characteristics between these two networks [7]. We have also applied a number of optimizations to improve such handoff performance. Although such schemes help improve performance, we have found that the overall gains here might not be well-worth the efforts required for such optimisations.

Soft Handovers are always better. An important benefit from a wireless overlay and from cellular networks, besides using multiple access for a reliable wireless communication link, is the macro-diversity available for use during handover decisions. Such handovers, for example, can be between GPRS↔WLAN and/or GPRS↔3G. While only limited link-layer disparity between GPRS and 3G wide-area wireless links can somewhat mitigate the effects on hard handoffs between GPRS↔3G, handoffs between GPRS↔WLAN (and 3G↔WLAN) will continue to pose a significant challenge [7].

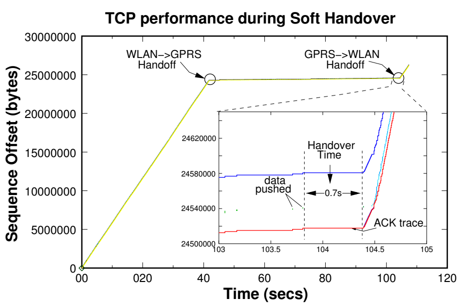

Fig. 3. Exploiting Macro-diversity during (Soft) Handovers.

However, unlike hard handovers, a better technique to handoff across coverage offered by networks in a wireless overlay and/or cellular networks can exploit the inherent macro-diversity available in order to make handovers *soft* to help improve performance. Figure 3 shows TCP performance using soft handovers. Here we have found that the performance of TCP protocol during soft handoffs is significantly better than the one shown earlier (with hard handoff in figure 2). Traditionally, soft handovers have already been successfully exploited for link-layer handovers in cellular networks [6].

Nevertheless, our MARS router would be mostly involved in bandwidth aggregation, striping packets across multiple wireless links. Only in relatively few cases (when exploting diversity from WLANs) it may decide to handoff, and in which case, it should use the soft handoff approach to mitigate the handoff effects.

IV. MARS – Architecture and Components

Based on the motivation and the measurements presented in this paper, a case has been made for exploiting the network diversity in wireless access for providing a reliable communication channel. We are exploiting the diversity in wireless overlay and cellular network access in our MARS router.

MARS is *dual-proxy* system consisting of a MARS client – a multimode mobile device used as the mobile accesss router and connected to different wireless networks (e.g. GPRS, 3G, WLAN etc.) and operators (e.g., Vodafone, Orange etc.) simultaneously – to communicate with a MARS server proxy located in the wired infrastructure (see figure 4). The MARS client acting as mobile access router can be placed in a car, bus, train etc., to perform bandwidth striping (aggregation) across multiple network interfaces to exploit the network diversity available from different wireless networks to provide a reliable wireless channel.

In the MARS system, we are focussed on the downlink as we perceive this to be most important to the MARS users. Downlink caters to most of the important applications such as web, e-mail, ftp etc. Besides web proxying in MARS, requirements of various other applications has led us to make use of two other proxies that sits between MARS client proxy and the server:

- **Transparent TCP Proxy.** The transparent TCP proxy is useful for applications (other than Web) that make use of TCP such as FTP, IMAP etc. The transparent TCP proxy ensures that such TCP connections are taken care of.

- **Generic Protocol Proxy.** A generic proxy in MARS can take care of the application protocol packets other than TCP, such as UDP, ICMP or even IPSec tunnel packets. The proxy encapsulates these protocol packets, provides sequence numbers before striping them across multiple wireless links, which are then reconstructed.

Between the client and server proxy, MARS implements the following features to improve web performance:

MARS Transport. MARS exploits the knowledge of the link-layer information to perform intelligent striping of data across network interfaces. Experiments conducted have shown that Internet transport protocols like TCP may not be adequate for better web experience over wireless wide area such as GPRS and 3G [12], [8]. Instead, a protocol that is optimised for wireless link characteristics, and one that minimises link traversals (especially DNS look-ups and TCP's 3-way handshake), and responds efficiently in the event of the patterns of packet loss we commonly observe over such wireless links is better (see [9]). MARS uses a transport based on flow-control that runs over UDP, applies no congestion control like TCP, to implement an ordered, reliable message transfer that is highly optimized for striping packets over multiple wireless links. It achieves substantially better link utilization than TCP.

MARS Session Protocol. One of the key functionalities of MARS is that it can aggregate the available bandwidth in all wireless interfaces. This aggregated bandwidth is then offered as a larger, more stable pipe to the end users. A simple approach to aggregate bandwidth from multiple links is to use link-layer

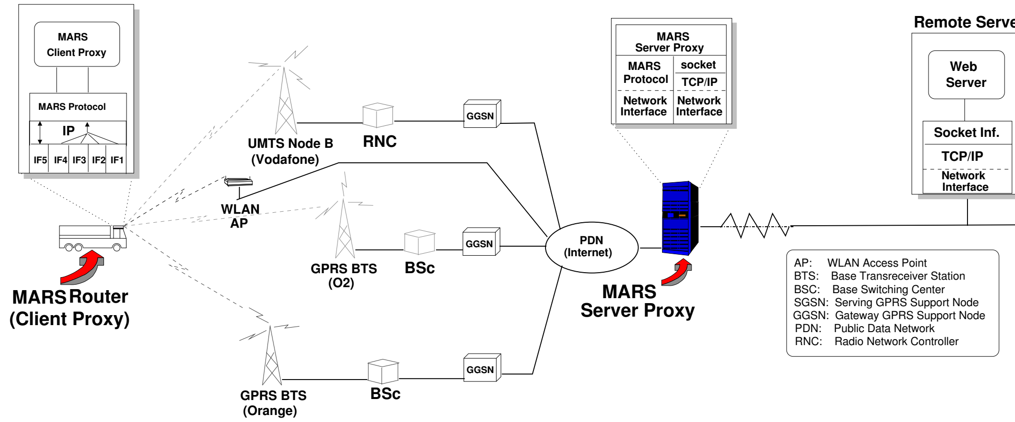

Fig. 4. MARS System Architecture and Components.

striping techniques [13], [14]. However, such schemes work quite poorly in wireless links with high fluctuations [11]. Other approaches such as PTCP have also been proposed at the transport layer for bandwidth aggregation [11]. However, these schemes rely on the congestion window to be a tight approximation of the available bandwidth-delay product to be able to efficiently stripe different packets in the different interfaces. In many real deployments (e.g. GPRS, CDMA 1xRTT), wireless systems include deep buffers in the Base Station Controllers to mitigate burstiness and therefore artificially inflate the congestion window to much larger values than the true bandwidth-delay product. This clearly reduces the applicability of such schemes in real environments.

Instead, the MARS platform uses a new MARS session-level protocol between the MARS client-proxy and the server-proxy. The MARS session protocol is wrapped over the transport to combine dynamic parallel-download techniques [10] and FEC [15] to aggregate the available bandwidth in multiple wireless interfaces without requiring a precise estimation of the bandwidth or the delay in the wireless links. This new aggregation protocol does not require changes to the end systems (clients and servers) and automatically adapts to changing wireless conditions by instantaneously shifting the load from under-performing wireless interfaces to better performing ones.

MARS Extended Caching. MARS client-side proxy caching improves performance by eliminating network round-trips and reducing the amount of data exchanged over the wireless link. As discussed earlier, MARS transport protocol performs intelligent striping of data across multiple wireless links, and closely integrates with the MARS client-side proxy cache to decide on the data to stripe. Furthermore, a custom caching protocol is used between the MARS client and server proxies that enables better hit rates by using fingerprints of objects to determine whether they have actually changed or not. The protocol can thus eliminate unnecessary object transfers to stripe across wireless links, and makes better use of the limited size cache available in the mobile access router. MARS server-side proxy also implements a traditional HTTP cache to reduce bandwidth requirements on the wired network, and thus, can take the place of existing proxy caches that are already commonly deployed by ISPs.

Application-Level Adaptation. Since wireless bandwidth is scarce, and also expensive, MARS compresses data before sending/striping it over the wireless links, reducing transfer size and thereby improving response time. Data is compressed unless it is already in a compressed format (e.g. JPEG images, Zip Archives). A separate string table is used for HTTP headers, resulting in better compression. When the server-side MARS proxy detects that a previously cached object has been updated, it tries using a difference algorithm to encode the differences between the old and new objects. The compressed *deltas* are sent in place of the new version if they would result in a smaller transfer.

DNS Migration/Push-when-idle. Most web pages contain a number of images and other objects that make-up the page structure, e.g. button graphics, spacers, style sheets, frames etc. These objects are requested by a browser after parsing the HTML documents. A round trip delay is normally incurred before transfer of these objects can commence. The *push-when-idle* mechanism in the MARS server proxy parses HTML objects, and pro-actively starts *pushing* objects towards the MARS client proxy cache if the link would otherwise be idle. The *Push-when-idle* feature is similar to *parse-and-push* used in GPRSWeb [9]. DNS look-ups migration to the Server proxy ensures that name resolutions do not ential high link (e.g. GPRS, CDMA 1xRTT) RTTs.

V. Related Work

Berkeley's BARWAN project made several important observations for wireless overlay network [1]. The wireless overlay network concept is a way to combine the advantages of wireless coverage while still achieving the best possible bandwidth and latency for mobile devices at any point of time. The objective in MARS is only somewhat different, the idea here is to exploit the network diversity not just from wireless overlay networks, but also from the network diversty of the virtually pervasive cellular infrastructure. While BARWAN project is mainly focussed on low-latency inter-network handovers between overlays, MARS aims to exploit the network diversity inherent in wireless access to aggregate bandwidth that can be offered as a larger and stabler pipe to the end users.

In some ways, the idea of exploiting network diversity for sustainable data rates in communications channels from wireless overlay and cellular networks has quite similar objective to that of resilient overlay networks [4]. In RONs, applications exploit the advantage of network paths in the Internet routing, whereas in MARS the client (proxy) uses the diversity in the wireless access to benefit from many link simultaneously, thereby improving end-to-end reliability and performance.

MIT's Personal Router (PR) project [3] has broader objective than MARS. The main idea behind Personal Router is to provide technological infrastructure that support mobile access to wireless services, along various dimensions such as network support with fast handoff, pricing, QoS, network traffic monitoring and user modelling. While Personal Router Project evaluates many key issues related to wireless access; the main objective in MARS (of exploiting network diversity for reliability and performance) is different from that in PR. However, some innovations in PR might still be applicable for MARS.

Related research projects include the Mobile People Architecture (MPA) [17], the ICEBERG project [16], and the TOPS architecture [18]. All the three projects attempt to provide user level mobility within one or more network types. The MPA uses a person-level router, the Personal Proxy, that tracks a mobile user's location, and accepts communication on the user's behalf, performs any conversions, and then forwards communications to the user. The ICEBERG and TOPS approach depends upon tracking proxy (e tracking router) nodes within the network.

The research closest to MARS is the GPRSWeb proxy system [9]. Although GPRSWeb does not exploit network diversity by using multiple network interfaces, MARS borrows a number of proxy-specific features from GPRSWeb that significantly improves web performance over wireless (here GPRS).

VI. SUMMARY

In this paper, we discussed the limitations of the current wireless access systems of the wide-area. To that end, we made a case for exploiting the network diversity in wireless access. We have argued that distributed diversity in wireless access from different wireless overlay and cellular networks can be used as an advantage to provide a sufficiently sustainable and reliable wireless communication channel. Based on a real-world example, we have shown that there is a substantial overlap in terms of coverage being offered by most of these operators (e.g. Vodafone, Orange etc.) and also across networks (e.g. GPRS, 3G, and/or WLANs).

We have used the network diversity model as potential advantage in MARS - a Mobile Access Router *on-the-move* for bandwidth aggregation during mobile Internet access.

While initial measurements and some experiences with the implementation of MARS indicate encouraging results, there are some interesting questions to address:

- *Does network diversity imply communication reliability?* Since exploiting diversity from wireless and cellular access networks is the biggest motivation for building MARS, we are performing a thorough quantitative study of the benefits available from network diversity. However, initial experiences here are very promising.
- *Can we provide MARS users with end-to-end security (e.g. https)?* A drawback using custom protocol between MARS client and server proxies over wireless links is that it can not provide end-to-end security (`https`). Designing a proxy-based solution that can also provide end-to-end security is still an open topic for research. Here, we are exploring an alternative to https with an object-based encryption rather than session-based encryption.
- *Can MARS Server proxies collaborate?* Since wireless and cellular networks can deploy their own MAR server proxy, it is quite possible that these proxies can be made to co-ordinate and avoid circuitous traffic routing. We are also investigating how MARS server proxies could benefit from such collaboration.

REFERENCES

[1] R. H. Katz and E. C. Brewer, "The Case for Wireless Overlay Networks", *In Proceedings of SPIE Multimedia and Networking Conference (MMNC'96)*, San Jose, CA, Jan 1996.

[2] B. Walke and V. Kumar, "Spectrum issues and new air interfaces", *Computer Communications*, 26(2003), pages 53-65, Elsevier.

[3] D. D. Clark and J. T. Wroclawski, "The Personal Router Whitepaper", MIT Laboratory for Computer Science Technical Report. Version 2.0 – March 2001.

[4] D. G. Andersen, H. Balakrishnan, M. F. Kaashoek, and R. Morris, "The Case for Resilient Overlay Networks", In *Proceedings of HotOS–VIII*, Schloss Elmau, Germany, May 2001.

[5] D. B. Johnson, C. E. Perkins and J. Arkko, "Mobility Support in IPv6" (draft-ieft-mobileip-ipv6-23.txt), May 2003. `http://www.ieft.org`.

[6] N. D. Tripathi, J. H. Reed and H. F. Vanlandingham, "Handoff in Cellular Systems", *IEEE Personal Communications*, December 1998.

[7] R. Chakravorty, P. Vidales, L. Patanapongpibul, K. Subramanian, I. Pratt and J. Crowcroft, "On Inter-network Handover Performance using Mobile IPv6", University of Cambridge Computer Laboratory– Technical Report, June 2003.
`http://www.cl.cam.ac.uk/users/rc277/overlay.html`

[8] R. Chakravorty, S. Katti, I. Pratt and J. Crowcroft, "Flow Aggregation for Enhanced TCP over Wide-Area Wireless", in *Proceedings of INFOCOM 2003*. `http://www.cl.cam.ac.uk/users/rc277/gprs.html`

[9] R. Chakravorty, A. Clark and I. Pratt, "GPRSWeb: Optimizing the Web for GPRS Links", In *Proceedings of the ACM/USENIX MobiSys 2003*, May 2003, San Francisco, USA.

[10] P. Rodriguez and E. W. Biersack, "Dynamic Parallel-Access to Replicated Content in the Internet", In *ACM/IEEE Transactions on Networking*, August 2002.

[11] H.-Y. Hsieh and R. Sivakumar, "A Transport Layer Approach for Achieving Aggregate Bandwidths on Multi-homed Mobile Hosts." In Proceedings of the ACM MOBICOM, Atlanta, GA USA, September 2002.

[12] M. C. Chan and R. Ramjee, "TCP/IP Performance over 3G Wireless Links with Rate and Delay Variation", in *Proceedings of ACM Mobicom 2002*, September 2002.

[13] J. Duncanson, "Inverse multiplexing", *IEEE Communications Magazine*, 32(4):34-41, Apr. 1994.

[14] C. B. Traw and J. Smith, "Striping within the network subsystem", *IEEE Network Magazine*, 9(4):22-32, July 1995.

[15] J. W. Byers, M. Luby, and M. Mitzenmacher, "Accessing Multiple Mirror Sites in Parallel: Using Tornado Codes to Speed Up Downloads", in *Proceedings of INFOCOM 1999*, New York, NY, USA.

[16] A. D. Joseph, B. R. Badrinath and R. Katz, "The Case for Services over Cascaded Networks", In *Proceedings of WOMMOM'98*, 1998.

[17] M. Roussopoulos, P. Maniatis, E. Swierk, K. Lai, G. Appenzeller, M. Baker, "Person-Level Routing in the Mobile People Architecture", In *Proceedings of the USENIX Symposium on Internet Technologies and Systems*, October 1999.

[18] N. Anerousis, et. al., "TOPS: An Architecture for Telephony Over Packet Networks", In *IEEE Jounal of Selected Areas of Communications*, 17(1), January 1999.