Internet on the Move: Challenges and Solutions

Arjuna Sathiaseelan
Computer Laboratory
University of Cambridge
Cambridge, UK
arjuna.sathiaseelan@cl.cam.ac.uk

Jon Crowcroft
Computer Laboratory
University of Cambridge
Cambridge, UK
jon.crowcroft@cl.cam.ac.uk

This article is an editorial note submitted to CCR. It has NOT been peer reviewed. The authors take full responsibility for this article's technical content. Comments can be posted through CCR Online.

ABSTRACT
The Computer Laboratory, University of Cambridge hosted a workshop on "Internet on the Move" on September 22, 2012. The objective of the workshop was to bring academia, industry and regulators to discuss the challenges in realizing the notion of ubiquitous mobile Internet. The editorial summarises a general overview of the issues discussed on enabling universal mobile coverage and some of the solutions that have been proposed to alleviate the problem of having ubiquitous mobile connectivity.

Categories and Subject Descriptors
C.2.1 [Network Architecture and Design]: Wireless communication; C.2.2 [Network Protocols]: Protocol Architecture; C.2.6 [Internetworking]: Standards.

Keywords
Mobile, Internet, Resource pooling, Congestion

1. INTRODUCTION
There is a complete paradigm shift in how users access the Internet. Significantly more people access the Internet via mobile rather than computers. The International Telecommunication Union (ITU-T) [1] estimates that there are almost 1.2 billion mobile Web users in the world and mobile broadband subscriptions have grown 45 percent annually over the last four years. Mobile-broadband subscriptions have now outnumbered fixed broadband subscriptions 2:1. In many developing countries mobile broadband is often the only access method available to people.

Both wireless and mobile technologies have rapidly evolved over the years. Mobile technologies have evolved from 1200 bps Nordic Mobile Telephone (NMT - 1G) to 100 Mbps Long Term Evolution (LTE - 4G) to the future 200 Mbps 5G services in 2030 [2] while wireless standards have evolved from 2 Mbps 802.11b to 600 Mbps 802.11n. Although the speeds have increased, the notion of ubiquitous access to the Internet anytime and anywhere continues to be a mirage. The high level purpose of the workshop was to end up sharing information to see how to achieve better performance for the mobile wireless user by looking at a variety of technical improvements, including bonding and merging of services through a range of techniques at all layers of the protocol stack. Stakeholders from industry (equipment and service providers, as well as the measurement community), academia, regulators and users were invited.

In this editorial, we summarize the outcome of the workshop by categorizing the problems inundating mobile/wireless broadband into three, and discuss some of the solutions that have been proposed to alleviate the problem of having ubiquitous mobile connectivity.

2. UBIQUITOUS COVERAGE
Achieving ubiquitous mobile broadband coverage is currently seen as not feasible by major mobile operators as direct investment in local infrastructure may be uneconomic. For e.g. in the UK, the major telecom operators claim there is 90% or more 3G coverage [3], however the recent BBC conducted crowd-sourcing survey from 44,600 volunteers showed that 3G coverage is far more patchy than mobile operator coverage maps indicating that there are still many 'not-spots' – this ironically includes major towns and cities [3]. Current wireless and mobile access networks have been developed in a fragmented way. In the future a great deal of flexibility is needed in terms of how networks are constructed and operated, how spectrum is used most efficiently between several operators and technologies for managing such flexibility. Such advances can help users with cellular coverage to attain higher speeds, but do not extend coverage. Achieving ubiquitous coverage requires policy changes within telecom regulators to mandate the need for providing 100% mobile broadband coverage in return for selling 4G spectrum licences to mobile operators – Ofcom in the UK has a mandatory coverage obligation policy for 4G operators to provide coverage to least 98% of the UK population by the end of 2017 [4]. The workshop discussed resource pooling as a potential remedy until the notion of 100% coverage becomes a reality.

2.1 Cross-layer Resource Pooling
Resource pooling can be achieved by bonding or aggregation at different layers of the protocol stack.

Physical layer bonding Physical layer bonding can be achieved using joint multi-user beamforming (JMB) [5], a system that enables independent access points (APs) to beamform their signals, and communicate with their clients on the same channel as if they were one large Multiple-Input and Multiple-Output (MIMO) transmitter. JMB builds on the well-known Multi-user beamforming technique that enables a MIMO transmitter to deliver multiple independent packets to receivers that have fewer antennas. JMB introduces a new low-overhead technique for synchronizing the phase of multiple transmitters in a distributed manner. Using JMB, a wireless LAN can efficiently scale its throughput by continually adding more APs on the same channel.

Link-layer bonding Channel bonding (also called as Carrier Aggregation) combines two or more network interfaces in a device to offer redundancy or increased throughput. This could be achieved with either 3G or WiFi.

- **3G bonding** Bonded 3G technology uses the plethora of mobile broadband channels provided by various mobile operators and aggregates these physical channels to provide one expanded pipe that will appear as a single broadband connection to the end user, offering increased throughput and redundancy. This requires a number of SIM cards from different providers and utilising the available network connectivity and bandwidth from each of these connections. A three-channel bonded solution can deliver speeds up to 21 Mbps, while up to 42 Mbps is possible with the six-channel version; both dependent on cell site congestion and signal strength.

  The bonding technology takes the incoming packets and sends them across multiple links. This could be done either through round-robin scheduling to achieve load balancing or measuring the available bandwidth across each of the individual links and scheduling packets proportionate to the available bandwidth (adaptive balancing). The bonded connection will usually appear as a single IP address to the overlying application. In order to bond multiple channels, a secure tunnel is setup between the user's aggregation router and a proxy server provider by the provider who is providing the aggregation service. All packets are proxied and sent parallel over several aggregated links through this secure tunnel [6].

  3G Channel bonding provides several benefits such as increased coverage, throughput and offers reliability. However, the solution is still not considered as an affordable solution – mainly used by enterprises and has not yet found mass market. It is expected that 4G technologies will enable much easier carrier aggregation, as they are the key elements in the specifications for LTE Advanced, Worldwide Interoperability for Microwave Access (WiMAX 2) and future iterations of High Speed Packet Access (HSPA+).

- **WiFi bonding** WiFi channel bonding can be achieved by bonding two adjacent WiFi channels to double the bandwidth. The IEEE 802.11n standard allows wireless devices to operate on 40MHz-width channels by doubling their channel width from standard 20MHz channels. Although channel bonding provides increased bandwidth, it also substantially increases the risk of interfering with nearby Wi-Fi networks due to the increased spectrum and power consumption. The adoption of MIMO technology in 802.11n, devices can now exploit the increased transmission rates from wider channels at a reduced sacrifice to signal quality and range. However WiFi bonding has still not received wider adoption due to telecom regulations (especially in the UK).

Network-layer bonding Network layer resource pooling can be done by bonding multiple routers to appear as a single virtual router, hence maintaining connectivity even during a failure. This requires multiple routers in the network path to participate to create a single virtual router. Network layer bonding can be achieved using the VRRP (virtual router redundancy protocol) (Internet Engineering Task Force (IETF) RFC 5798). VRRP is based on Cisco's proprietary Hot Standby Router Protocol (HSRP) a proprietary redundancy protocol by Cisco for establishing a fault-tolerant default gateway (IETF RFC 2281). The static default routed environment can cause single point of failures. VRRP eliminates this by specifying an election protocol that dynamically assigns responsibility between multiple routers for a virtual route. VRRP creates the notion of virtual routers, an abstract representation of multiple routers (a combination of master and backup routers). VRRP assigns the virtual router to be the default gateway and if the physical router that is routing packets on behalf of the virtual router fails, another physical router is selected to automatically replace it. The physical router that is forwarding packets at any given time is called the master router. Network layer bonding can offer the necessary redundancy.

App-level bonding New mobile applications have emerged that allow users to share their mobile broadband connections (e.g. OpenGarden [7]) to create mobile overlay mesh networks with other similar subscribed users. App-layer bonding applications such as OpenGarden also offer multiple connection logic (for e.g. inter-technology mobility i.e. when WiFi is available, offload traffic to WiFi from 3G/4G) and link level bonding support.

2.2 Crowd-shared Resource Pooling

Another potential solution would be the notion of crowd-shared resource pooling - where users share their resources with others (either as a cooperative wireless mesh network (for e.g. see Guifi [8]) or as a standalone Internet connection (for e.g. BT FON [9]).

Solutions such as BT FON have already proved the efficacy of providing ubiquitous mobile network connectivity where home users share their home broadband connection with the public and earn credits which will enable them to access other similar users' access points. Although these methods are gaining worldwide acceptance, they are usually viewed as an extension of a user's paid service. The Engineering and Physical Sciences Research Council (EPSRC) has recently funded a project called PAWS (Public Access WiFi Service) [10], which takes the approach of community-wide participation, where broadband customers can volunteer to share their high-speed broadband Internet connection for free with fellow citizens. As it is essential to ensure that the free user traffic does not impact perceived performance of the bandwidth donor, we will explore the free services available through Less-than-Best Effort (LBE) access (also known as the Scavenger Class) to the network [11]. These methods allow a person to use a shared link without competing for the resources of those who have shared their Internet connection.

App-level bonding can also be used for crowd-sharing for e.g. as we saw in the case of OpenGarden which enables creation of mobile overlay mesh networks with other similar subscribed users. The research community have explored solutions using Pocket Switched Networking (PSN) [12] to explore the use of localized connectivity to provide a multi-hop opportunistic network made available by human mobility to transport mobile data (for e.g. Shair – a solution to allow sharing of a user's unused contract minutes with other users [13]).

3. MOBILITY SUPPORT

Mobility can happen due to several reasons: User mobility (user moving from one physical location to another), Host mobility (where a device associates itself to a new network attachment point (the interface between the network and link layer) also referred to as connection migration) or Session mobility (sessions moving to different server for load-balancing as in the case of server farms).

Endpoints use network attachment points as identifiers to communicate with its peer end points. A mobility event such as a host moving from one location to another could cause an endpoint to associate itself with another network attachment point. If this change is not notified to the communicating peer endpoint, then this would result in a disconnection since the peer end point will continue to address packets to the previous network attachment point. Such disconnections become particularly problematic for session-based applications such as streaming media.

Once a connection is terminated, creating a new connection means that an endpoint must first discover its new attachment point and communicate this information to its peer endpoints. In the network layer level, several solutions have been proposed in the past to solve the problems imposed by mobility. Solutions such as Mobile IP (that allows location independent IP packet routing allowing mobile users to move from one network to another while maintaining permanent IP address) have been standardized at the IETF (RFC 5944, RFC 4721, RFC 6275). Other solutions include Host Identity Protocol (HIP) (IETF RFC 4423), Nimrod (IETF RFC 2103) etc.

Once both the endpoints have synchronised on the new attachment points, the endpoints must terminate the old transport connections and establish new transport connections. Transport protocols that provide reliability (for e.g. TCP) may result in the loss of packets that were not yet successfully transmitted on the initial connections during the process of changing the network attachment points. A number of researchers have proposed mechanisms to allow connections to adapt to changes in attachment points. The proposals were to either extend TCP to support connection migration (e.g. TCP-R, Multi-homed TCP etc), standardize new transport protocols that support multi-homing such as Stream Control Transport Protocol (SCTP) and Datagram Congestion Control Protocol (DCCP), introduce higher level mechanisms to combine multiple TCP connections into one virtualized connection or the recently proposed Multipath TCP which establishes more than one path between two endpoints hence providing the necessary redundancy (make before break) and reliability by treating the multiple paths into a single resource pool.

The sporadic mobility of users places a considerable burden on applications that rely on seamless Internet especially multimedia applications. User mobility can introduce temporary periods of disconnection (this may be due to no network coverage) or the user may have moved into a higher density coverage area where the available bandwidth per user maybe low for an acceptable application performance. Current wireless networks are often point solutions, tightly bound to the overall architecture chosen by the operator (e.g. 3GPP), relying on careful provisioning and control of the network to ensure stable operation. There is no framework to support seamless adaptation of multimedia content across a range of mobile networks. Previous work [14][15] suggests significant performance penalties arise not only from the mismatch between codec expectations and the congestion control algorithms, but also because the traditional Berkeley Sockets API does not expose sufficient information to allow effective adaptation. Research is required to establish a new transport framework that is both adaptive to the needs of applications and also to the diversity of network conditions over which it must operate. Such a framework needs to embrace developments in media codecs and then move beyond the TCP/IP stack of today's Internet to offer high-availability to multimedia services, and provide mechanisms that can support user expectations as they migrate around the network requiring traffic to pass over links with vastly different properties.

4. BACKHAUL CONGESTION

The recent years have seen a proliferation of new Internet enabled devices ranging from smartphones to connected cars. According to the Olswang report published in early 2011, as of March 2011 22% of UK consumers had a smartphone, with this percentage rising to 31% amongst 24 to 35-year-olds and the rate of smartphone adoption is expected to accelerate. Connected cars with features like Internet-enabled navigation and streaming media will soon be the norm with a predicted 50 million being sold every year by 2017 [16]. It is estimated that 66% of the world's mobile data traffic will be video by 2014 and that mobile video will grow at a compound annual growth rate of 131% over the next five years [17]. This trend coupled with demand for connectivity, which is outstripping supply will create an overload on backhaul networks as well as an overcrowded radio frequency (RF) spectrum, which could result in poor user experiences. Mobile operators have to rethink their infrastructure with an eye on managing RF spectrum resources and must quickly find new ways to economically increase capacity and extend network coverage. Expanding the network by simply doubling 3G/4G capacity is not the right solution as additional traffic will not bring additional revenue.

4.1 Offloading and Onloading

Cellular and Wireless Offloading The WiFi has come as a boon to mobile operators at a time when these operators were searching for a cost effective solution to solve the problem of backhaul congestion. The ubiquitous nature of WiFi hotspots now create several transmission opportunities and is seen as an attractive option (compared to other alternative solutions) to mobile operators to offload data seamlessly and reliably from the congested cellular networks (3G/4G). Hence mobile operators have now started considering integration of WiFi with their mobile network infrastructure. Seamless integration with the services provided by existing cellular network core is extremely important and this would require changes to all equipment within the cellular network from the edge devices (mobile devices, customer premise equipment (CPE)) to the core network services - which recently have seen advanced capabilities to enable transparent interactions such as Home Location Register/Home Subscriber Service (HLR/HSS), Policy Changing and Rules Function (PCRF) and Authentication, Authorization, Accounting (AAA) between cellular and WiFi networks. The 3GPP release 10 supports WiFi offloading as part of its standard [18].

3G Onloading Recently there have been proposals on doing 3G Onloading (3GOL) [19], as a way to improve performance for residential users, for those applications that are bottlenecked by the wired network. 3GOL is used to augment existing connections by moving part of the traffic onto the 3G infrastructure. [19] shows that throughput augmentation could scale linearly with the number of 3G devices in the downlink although limited by the 3G/4G uplink technology. The results also showed significant reaching 3x downlink and 12x uplink even with a small number of 3G devices.

The research community have explored more adventurous solutions such as MADNet (Metropolitan Advanced Delivery Network) [20], a heterogeneous wireless network architecture that enables offloading mobile data traffic using opportunistic communications and WiFi networks. MADNET is a deployable architecture that utilizes available technologies to aggregate the power of mobile social networks, opportunistic communications, and collaborations among cellular operators, WiFi providers, and mobile users.

4.2 Offloading for energy

Mobile computing models rely exclusively on cloud and local resources for different purposes: from assisting sensors such as A-GPS to code offloading. This model is not optimal in many situations as accessing resources in the cloud is subject to network availability and latency while also imposing an energy overhead on the handset. As a result, both the user experience and the battery life of the handset can be severely diminished.

ErdOS [21] is an energy-aware social operating system in which mobile handsets can collaborate with neighbouring machines in order to assist each other and share their network, sensors and computation resources using low energy wireless connectivity such as low power Bluetooth and Qualcomm Flashlinq. ErdOS aims both to improve the handset usability and extend the battery life by an efficient management of all the mobile resources present in the environment by incorporating social networks as a fundamental part of the operating system.

5. CONCLUSION

In this editorial, we summarize the outcome of the recently held workshop at Cambridge to discuss the problems inundating mobile/wireless broadband and explore some of the solutions that have been proposed to alleviate the problem of having ubiquitous mobile connectivity. We envisage next generation mobile devices and customer premise equipment to be equipped with intelligent connection managers that would have inherent support for:

- **Inter-Radio Access Technology (RAT) Mobility** to enable seamless mobility between 4G and 3GPP Network. When LTE is rolled out, it is not expected to replace 3G/2G immediately, but will augment it. By enabling inter-RAT mobility, the devices will maintain their connections when they move out of 4G coverage by falling back to 3G/2G.
- **Inter-Technology Mobility (WiFi Offloading)** – When an accessible WiFi connection is available, the sessions will be transferred to WiFi saving cost and energy (during low-throughput data transfer) to the user. This can be achieved even without operator tie-ups although interactive multimedia such voice may require service guarantees.
- **Manage Multiple Connections** – 3G devices could only support a single voice or data connection. With 4G technology, multiple connections with varying QoS may have to be actively supported. 4G technologies will also enable much easier carrier aggregation, as they are the key elements in the specifications for LTE Advanced, WiMAX 2 and future iterations of HSPA+.
- **Mobility support** such as Mobile IP and Hierarchical Mobile IPv6 (HMIPV6).
- **Multipath TCP** support with fall back to Standard TCP over Mobile IP support.
- **Energy efficient flow switching** support (different interfaces are energy optimal in different throughput regions).
- **QoS-aware** packet scheduling support.
- **Cross-layer application programming interfaces (API's)** for adaptive application support (applications can choose media rate based on available bit rate on runtime for e.g scalable video coding (SVC)).

6. ACKNOWLEDGMENTS

This work was funded by RCUK through the Horizon Digital Economy Research grant (EP/G065802/1) and EPSRC grant EP/K012703/1.

7. REFERENCES

[1] ITU facts and figures, http://www.itu.int/ITU-D/ict/facts/2011/material/ICTFactsFigures2011.pdf, 2011.

[2] UK To Get 200Mbps 5G Network Trial In 2013, http://www.techweekeurope.co.uk/news/5g-trial-uk-2013-4g-95408, October 2012.

[3] 3G mobile data network crowd-sourcing survey by BBC news, http://www.bbc.co.uk/news/business-14574816, August 2011.

[4] Assessment of future mobile competition and award of 800 MHz and 2.6 GHz, Ofcom, July 2012.

[5] H. Rahul, S. Kumar, D. Katabi, JMB: scaling wireless capacity with user demands, ACM SIGCOMM, Helsinki, August 2012.

[6] Viprinet principle, http://www.viprinet.com/en/technology/principle.

[7] Opengarden, http://opengarden.net/.

[8] Guifi.net, http://guifi.net/en.

[9] BT FON, www.btfon.com.

[10] A. Sathiaseelan, J. Crowcroft, M. Goulden, C. Greiffenhagen, R. Mortier, G. Fairhurst, D. McAuley, Public Access WiFi Service (PAWS), Digital Economy All Hands Meeting, Aberdeen, October 2012.

[11] A. Sathiaseelan, J. Crowcroft, The free Internet - a distant mirage or near reality? UCAM Tech Report, February 2012.

[12] J. Scott, P. Hui, J. Crowcroft, C. Diot, Haggle: A Networking Architecture Designed Around Mobile Users, IFIP WONS, France, January 2006.

[13] P. Hui, R. Mortier, K. Xu, J. Crowcroft, V. Li, Sharing airtime with Shair avoids wasting time and money, Tenth Workshop on Mobile Computing Systems and Applications (HotMobile), California, February 2009.

[14] L. Gharai, C. S. Perkins, Implementing Congestion Control in the Real World, IEEE International Conference on Multimedia and Expo, Switzerland, August 2002.

[15] A. Saurin, Congestion Control for Video-conferencing Applications, MSc Thesis, University of Glasgow, December 2006.

[16] Internet connected cars soon to become the norm, http://www.techradar.com/news/car-tech/internet-connected-cars-soon-to-become-the-norm-1093107, August 2012.

[17] Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update, 2010–2015, http://newsroom.cisco.com/ekits/Cisco_VNI_Global_Mobile_Data_Traffic_Forecast_2010_2015.pdf, February 2011.

[18] 3GPP TS 23.261 v10.1.0 (2010-09), "IP flow mobility and seamless Wireless Local Area Network (WLAN) offload; Stage 2".

[19] N. Rodriguez, V. Erramilli, Y. Grunenberger, L. Gyarmati, N. Laoutaris, R. Stanojevic, K. Papagiannaki, When David can help Goliath: the case for cellular augmentation of wired networks, ACM HotNets, Seattle, October 2012.

[20] Y. Li, G. Su, P. Hui, D. Jin, L. Su, L. Zeng, Multiple Mobile Data Offloading Through Delay Tolerant Networks, 6th ACM International Workshop on Challenged Networks (CHANTS 2011), Las Vegas, September 2011.

[21] N. Rodriguez, J. Crowcroft, ErdOS: Achieving energy savings in mobile OS, ACM MobiArch, Maryland, June 2011.