

Operator and Radio Resource Sharing in Multi-Carrier Environments

Pongsakorn Teeraparpwong, Per Johansson, Harsha V. Madhyastha and Amin Vahdat
University of California, San Diego
{pteerapa@cs, pjohansson@soe, harsha@cs, vahdat@cs}.ucsd.edu

Abstract—Today’s mobile networks prevent users from freely accessing all available networks. Instead, seamless network composition could present a win-win situation for both users and operators. Users can gain better quality of service with more resources to choose from, while each individual operator can provision lesser bandwidth since resources can be shared during times of peak demand. In this paper, we analyze the benefits of operator cooperation using real trace data of cellular data access. We leverage the difference in burstiness at small timescales across network providers to shed the peak usage of one operator on to another. Our results show that even when an operator provisions network capacity below the peak load, cooperation with other network providers can help maintain quality of service for most sessions. In addition, we investigate the performance delivered by various kinds of cellular data cards. Our results confirm that WiFi 802.11b/g consistently delivers superior performance compared to 3G. It will take the next generation 4G technologies such as LTE to deliver end-user performance comparable to widely-deployed 802.11 networks.

I. INTRODUCTION

In recent years, the mobile and wireless networking world has undergone rapid evolution. Driven by increasing demand and a highly competitive market, the choice of network providers and access technologies has significantly increased. Many novel access technologies have been developed including 3G/UMTS, WiFi 802.11b/g, WiMAX, and LTE.

Given current trends, a key obstacle that the mobile network world faces is the lack of seamless network access. Particularly, although the coverage of mobile networks is ubiquitous, agreements with service providers typically limit users to only their home operator’s network. Instead, network composition can enable users to gain better service coverage and quality of service. Additionally, network composition could enable operators to provision only for average bandwidth usage rather than the peak. Network load at the peak can be offloaded to other networks that have spare resources at that time. Such sharing of networks improves resource utilization while reducing provisioning costs. The key trade-off here is that the quality of service can be impacted if the provisioned bandwidth is too low.

Several efforts have attempted to implement network composition such as the European Ambient Networks (AN) project [1] and the resource reservation framework proposed by Al-Fares et al. [2]. However, the concept is yet to be evaluated in a realistic environment.

In this work, we extend the framework proposed by Al-Fares et al. to analyze the benefits of operator cooperation

using real trace data. We leverage the difference in burstiness at small timescales to shed the peak bandwidth usage of one operator onto another even when the average load of all operators are comparable. We use the Swing [3] traffic generator to extract traffic characteristics and generate traffic based on traces captured from an operational 3G data network.

Our experiments show that even when the capacities of all operators are limited to less than the peak usage of their respective networks, operator cooperation can help maintain quality of service comparable to when bandwidth is unconstrained. As a result, operators can benefit from cooperation by having to provision for lesser bandwidth, thus reducing capital expenditures.

An additional challenge is to decide which radio access technology (RAT) users should use to obtain the best performance or to minimize the cost/performance ratio. From an operator’s point of view, the challenges include which combination of access technologies is more appropriate to deploy to support user demand at minimum cost.

To provide insight on these challenges, we investigated the performance delivered from an end-user’s perspective by 3G, WiFi, and LTE for various application classes by utilizing Swing’s abilities to tune the parameters of the network topology. Our results show that WiFi can provide significantly better performance compared to 3G. Therefore, WiFi hotspots can be a good choice to augment 3G services in a small highly-populated area such as an airport. On the other hand, we expect the performance of LTE to be comparable to or even better than WiFi, especially when the number of users increases. Also, a key benefit of LTE is the ubiquity of its coverage.

In summary, to the best of our knowledge, we are the first to show the potential benefits of network composition using real 3G data traces. Also, we confirm that 802.11b/g still delivers qualitatively better performance than 3G. Finally, we quantify the level of real-world performance required from 4G technologies to become competitive with 802.11b/g.

II. RELATED WORK

Composition of mobile networks was previously proposed in the European *Ambient Networks* (AN) project [1]. The vision of AN is to allow instant composition of networks on demand without any prior configuration between operators. This can be done by establishing so-called *Composition Agreements* (CA) dynamically. This agreement facilitates the sharing of services and resources across networks and the

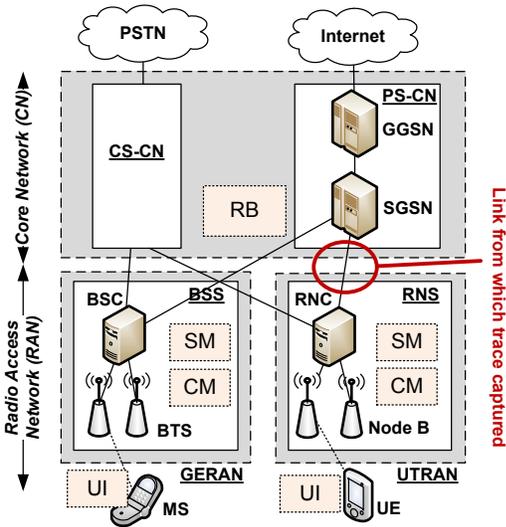


Fig. 1. 3G/UMTS network architecture and location of resource reservation components

compensation of usage can be regulated. The AN project delivered a system framework design [4] as well as validation and evaluation of the framework [5], [6].

As part of the AN project evaluation, Pöyhönen et al. [7] investigated the benefits of operator cooperation by using performance metrics for both operators and end-users. Their analysis showed that cooperation can provide better service availability and quality, leading to improved user experience and more benefits to operators.

Johansson et al. [8] analyzed the cost and performance of deploying heterogeneous access networks. Their results show that deployment cost can be significantly reduced and radio resources can be utilized more effectively. The results from the AN project [6] also show that all operators can reduce costs by using multiple radio access technologies and supporting cooperation with other operators.

There have already been numerous studies on the performance and usage of WiFi [9], [10] and 3G [11]–[13] deployments. Nevertheless, the performance evaluations of LTE are still limited. For example, Dahlman et al. [14] evaluated the performance of LTE compared to WCDMA and HSPA in a simulation environment.

III. BACKGROUND

In this section, we provide some background on the different wireless access technologies we compare in our study, and systems from prior work that we use in our experimental infrastructure.

A. Wireless Access Technologies

1) *3G/UMTS Network Architecture*: Figure 1 shows the architecture of a 3G/UMTS network [15], [16] which consists of a *Radio Access Network (RAN)* and a *Core Network (CN)*. An operator can maintain both *UMTS/HSDPA RAN (UTRAN)* and *GPRS/EDGE RAN (GERAN)* for backward compatibility

of GSM devices. In UTRAN, the *User Equipment (UE)* connects to one of the base stations (*Node B*'s) via an WCDMA (Wideband Code Division Multiple Access) wireless interface which supports downlink and uplink capacities upto 384 and 128 Kbps, respectively. Some modern systems also support *High Speed Packet Access (HSPA)* enhancement to WCDMA, allowing UE to achieve up to 14.4 Mbps in downlink speed.

In our work, we focus on the *packet-switched domain (PS-CN)*, which routes user IP packets to external PS networks, e.g., the Internet. PS-CN contains the *serving GPRS Support Nodes (SGSNs)* and *Gateway GPRS Support Nodes (GGSN)*. The SGSN performs mobility management and access control. The GGSN then serves as a gateway to the external PS networks.

2) *WiFi*: 802.11-wireless LAN (WLAN), known as WiFi, is widely used in homes, enterprises, and public areas. The standards supported by most routers are 802.11b and 802.11g which have a maximum range of around 45 and 90 meters for indoor and outdoor respectively. The peak physical data rates is up to 54 Mbps for 802.11g. Due to the standardization, WiFi is easy to deploy with low cost. However, the main limitations include its range and mobility support.

3) *Long-Term Evolution (LTE)*: LTE [17] is a new radio access standard developed by 3GPP aiming towards mobile broadband 4G. The system should support peak data rates of 100 Mbps and 50 Mbps on the downlink and uplink, respectively. The system allows rates upto more than 300 Mbps for the configuration with more antennas. Also, the round-trip time to RAN should be less than 10 ms. Many carriers globally have already started or plan to deploy LTE soon.

B. Resource Reservation Framework

Al-Fares et al. [2] proposed a resource reservation framework to enable network composition across multiple wireless access technologies and operators. The proposed framework is based on the GENI [18] resource reservation system documented in [19]. In order to access the resource, the user must first have credits which are represented by signed *Tokens*. Tokens can be exchanged for a signed *Ticket* which is a promise to access the resource. The value of the token is open and abstracted. Any resource is also abstracted and described by an XML-style document.

Al-Fares et al. implemented a prototype of all components in Java. The Apache XML-RPC library is used as a communication medium between components.

This framework has four main components as shown in Figure 2.

- *Site Manager (SM)*: Every network has a SM that continuously monitors the local resources and donates available resources to the resource broker.
- *Resource Broker (RB)*: RB receives resources donated by SMs. It also matches requests from users with appropriate resources and issues the ticket based on users' tokens.
- *Component Manager (CM)*: CMs grant users access to the resource specified in the user's ticket.

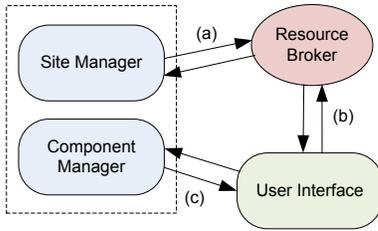


Fig. 2. GENI resource reservation components, which consist of following interactions: (a) resource donation, (b) ticket (resource) request, and (c) ticket redemption (resource allocation).

- *User Interface (UI)*: UI has resource discovery mechanism. When the user wants to acquire the resource, it first exchanges its tokens for a ticket through the resource broker (Figure 1(b)). Then it can redeem that ticket at the component manager to access the resource (Figure 1(c)).

Each operator can generate and give tokens to users when they pay their bill. Users can spend these tokens on any networks. Then those network providers will bill user’s home operator for network usage.

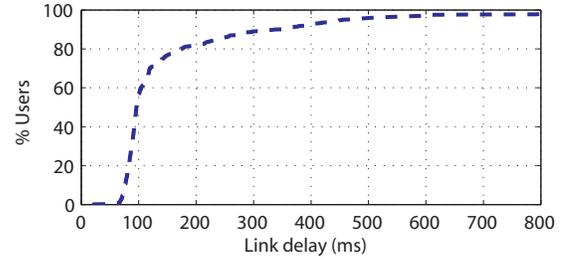
For our experiments, we consider the resource reservation framework to be mappable to the 3G and WiFi network scenarios as follows.

- *3G/UMTS Network*: As shown in Figure 1, the UI component can be embedded in the mobile station’s USIM/SIM. The SM and CM of a network would be co-located at each RNC and BSC. The SM can monitor resources at each RNS/BSS and donate available resources to the RB periodically. Finally, the RB can be located in the core network.
- *WiFi Hotspot*: Here too, the UI can be implemented on the user’s devices such as laptops. The SM and CM can be embedded in the access point router. The access control of users can be dynamically configured by manipulating firewall rules in the access point. Finally, the RB would reside at the core network of the service provider.

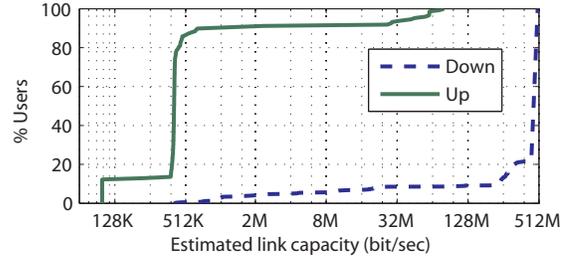
C. Traffic Emulation

We emulate these network scenarios using the *Swing* traffic generator tool [3]. We use Swing to first infer a network’s characteristics from a trace from the network and then reproduce traffic that accurately captures burstiness across a wide range of timescales. Traffic generated by Swing comprises TCP connections and is responsive to custom network conditions.

Swing generates a simple dumb-bell topology for the emulation. The topology consists of a *target link* which corresponds to the link from which the original trace was captured. Nodes are attached to both sides of the target link, representing end-hosts in the original trace. Due to the limited environment capacity, Swing collapses all emulated end-hosts into around 1,000 traffic generators. Swing then assigns a traffic generator to each node of the topology, with each generator assigned to play out TCP connections based on the traffic distributions that Swing inferred from the trace.



(a)



(b)

Fig. 3. Estimated (a) link delays and (b) link capacities from 3G trace.

Swing utilizes ModelNet [20], a network emulator, to emulate the dumb-bell topology. ModelNet consists of physical edge machines running multiple end-node applications, which are the Swing traffic generators and listeners in our case. All packets are routed through a ModelNet core machine which emulates the link bandwidth, latency, and loss rate of the topology being emulated. Previous work [20] shows that ModelNet can emulate traffic accurately up to 1 Gbps using a single core machine.

IV. TRACE DATA

The traces we use in our experiments were captured live from an interface between an RNC and an SGSN in a commercial 3G/UMTS network during Winter 2008. The traces comprise only packet headers of user-plane data. In addition, the traces include traffic only from 3G PC card devices.

In our experiments, we use 30 minutes of the peak period as the input trace data for Swing to infer traffic characteristics. This period comprises approximately 20 GB of traffic in total, with 78% of the traffic on the downlink. On average, 696 users sent or received data in each minute, producing around 90 Kbps of aggregate bandwidth usage. Also, on average around 16,000 TCP flows are active every minute. Approximately 5% of traffic is non-TCP and is excluded from Swing traffic generation.

Figure 3(a) shows the distribution of link delays from client devices to the target link estimated by Swing. The median delay is 96 ms. The cause for the high latency is the 3G wireless link between the client and base station. Figure 3(b) shows the distribution of estimated uplink and downlink capacities by Swing. Bursty downlink traffic due to packet queueing results in most of the estimated downlink capacities reaching Swing’s default maximum of 500 Mbps.

Nevertheless, our evaluation of the traffic generated by Swing using wavelet-based analysis [21] showed that it is able to preserve burstiness in traffic at various time granularities. On the other hand, most estimated uplink capacities are close to their median of 400 Kbps which is typical for a 3G network.

On the other hand, regardless of the amount of traffic, only around 2% of users experience non-zero packet loss. We attribute very low loss rate to the effective loss recovery protocol, namely RLC [22], in the 3G wireless link layer and the typically stationary nature of laptop clients.

V. MULTI-OPERATOR EXPERIMENTAL SETUP

We next evaluate various aspects of resource sharing between multiple operators by leveraging a resource peering framework. We explore the benefits of a scenario when every operator in the same area experiences roughly similar network load which is typically similar to real-world case. Due to the traffic burstiness, average network usage may be much lower than the peak usage. With operator cooperation and resource sharing, an operator might be able to offload some of the peak traffic onto other operators that have spare resources at that time. Consequently, we expect that operators can provision lesser bandwidth than their peak load while maintaining the same quality of service (QoS). Such sharing would also lead to higher resource utilization in each operator's network.

Since our traces comprise the aggregate traffic from many base stations at the RNC, we explore the situation when the capacity is limited at the RAN or in the core and higher up network. These capacities can be considered as a scarce resource in addition to radio resources since the provisioning comes with significant transmission cost.

Note that the network composition is not fully adopted in practice yet. Any real deployment would require more investigation into many aspects such as the right traffic sharing mechanism or pricing structure. Instead of getting into these issues, in our work, we are exploring the best case benefits of network composition.

A. Experimental Scenario

In our experiments, we use Swing to reproduce 30 minutes of traffic during the peak period from the trace described in Section IV. We emulate an environment comprising two operators providing 3G/UMTS service. Each client is randomly assigned to one of the operators, i.e., both operators have a similar number of subscribers. Also, in this scenario, our aim is to reproduce aggregate characteristics of the trace, and hence, we do not separate the traffic into application classes.

We emulate three different network scenarios.

- 1) First, we let clients generate traffic normally through their home operator. We impose no bandwidth constraints on either operator's network, and operator cooperation is not employed. The QoS measured in this setting serves as a baseline.
- 2) Next, we limit the downlink capacity of both operators. In this case too, we do not allow for any resource sharing. We measure the QoS to gauge the impact.

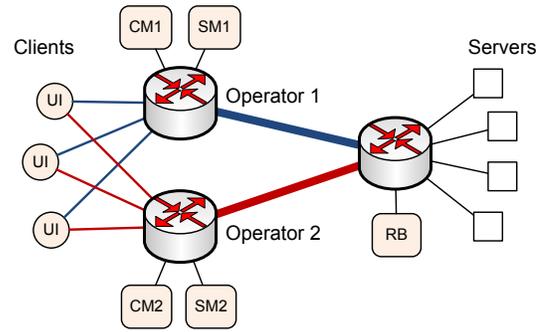


Fig. 4. Two-operator topology with resource reservation components

- 3) Finally, we enable resource sharing between operators while retaining the capacity limits of the previous case. In this case, a client's session could be assigned to either operator's network based on the cooperation policy. The measurements in this case yield the benefit or harm of network composition.

We employ the following policy for resource sharing. The client will access its home operator's network if the available bandwidth is above certain threshold. If not, the client will access the network that has more available bandwidth. We set the threshold to 20% of the network's capacity in our experiments. This threshold is low enough that clients will still use their home operator when there is enough bandwidth, avoiding unnecessary access to other service providers. The threshold is also high enough to be responsive to congestion.

Regarding resource reservation procedure, at the beginning of every session, the client first contacts the resource broker, which is shared between the two operators. The resource broker decides the operator for the client to access based on the above sharing policy and issues a ticket to user in exchange for user's tokens. We assume that all parties involved have enough token to obtain the best case performance. The client then redeems the ticket with the component manager of the assigned operator and begins sending traffic. Thus, operator selection is performed for every new session and the operator is fixed for the entire session. Additionally, each operator has a site manager which continually monitors the available bandwidth and donates it to the resource broker every half a second. All control traffic is sent over separate control channels.

B. Implementation

Our experimental setup consists of running Swing and ModelNet on a cluster of eight machines. One machine is setup as a ModelNet core, while the rest are edge machines running applications on end nodes of the emulated topology. All machines are powered by Intel Xeon 2.8 Ghz with 2 GB of memory and a Gigabit NIC. Each edge machine is responsible for emulating approximately 150 to 250 end hosts.

1) *Multi-operator Topology*: As mentioned earlier, Swing uses a dumb-bell topology. In this topology, the links connecting end nodes on one side of the target link represent the

paths from mobile clients to the RNC. Links on the other side represent the path from the SGSN to server hosts across the Internet. The target link itself represents a link between the RNC and the SGSN.

In our scenario, we assume congestion happens at the RNC or within the core network. Therefore, we consider the target link as the bottleneck link. Consequently, we can add another target link to represent a second operator as shown in Figure 4. One end of the new target link represents the RAN of the new operator. The other end can be shared with the first operator, assuming both operators have the same latency from the core network to the Internet. Every client node will have an additional link connected to the new operator, creating an alternate path to all destination servers. By assuming the latencies from clients to base stations and the load on base stations is the same for both operators, the characteristics of links connecting clients to the new operator will be the same as those to the first operator. Additionally, since Swing tends to overestimate link capacity, we cap the downlink and uplink capacities of clients to 3.6 Mbps and 512 Kbps based on characteristics of the 3G technology.

2) *Resource Reservation Framework Integration*: In order to support cooperation between operators, the resource reservation framework components described in Section III-B are integrated into the topology as shown in Figure 4. UIs are located at each client’s node. We do this by integrating the UI into Swing’s traffic generators. A generator starts sending traffic for a new session only once its UI has contacted RB and learned which operator it should access. For each operator, a SM and a CM are connected to the RAN side of the target link. The RB is connected to the other side which represents the core network. Additional links are added to carry control traffic. Further, we implemented a new component that runs on the ModelNet core machine to monitor the available bandwidth of links in the emulated topology. It logs its measurements to a file for each operator’s SM to read periodically.

3) *ModelNet Modification*: ModelNet uses a set of fixed routing paths between every pair of nodes. These paths are computed and loaded at startup. In our scenario, however, each client has more than one path to each server, each of which corresponds to an operator. We therefore modified ModelNet to load multiple sets of paths for every pair of nodes. A new ModelNet interface is implemented for selecting the path (operator) for each pair of nodes to use. At startup, each client will be set to use the path of its home operator. Then during runtime, the CMs can use the interface to change the path or operator of each client-server pair dynamically.

VI. MULTI-OPERATOR EXPERIMENTAL RESULTS AND ANALYSIS

In this section, we present results from the experimental setup described in the previous section. First, we explore the traffic and bandwidth usage for either operator. We then compare the quality of service when operators cooperate compared to when there is no cooperation. In making this comparison, we also investigate the overheads of the resource reservation

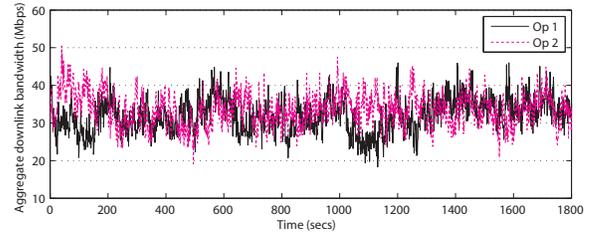


Fig. 5. Aggregate downlink bandwidth for each operator every 1 second when there is no capacity constraint

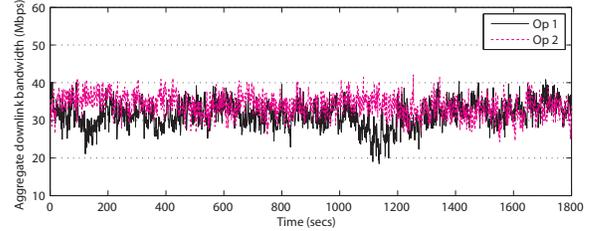


Fig. 6. Aggregate downlink bandwidth for each operator every 1 second when the capacities are limited to approximately 41 Mbps.

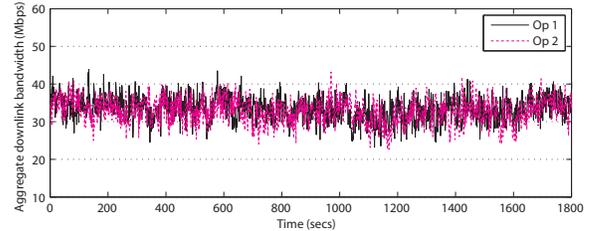


Fig. 7. Aggregate downlink bandwidth for each operator every 1 second when operator cooperation is supported and the capacities are limited to approximately 41 Mbps

system. Finally, we discuss potential alternate scenarios for resource sharing.

A. Traffic

Figure 5 shows the aggregate downlink bandwidth at the target link for both operators at the granularity of 1-second bins. Though the average usage is roughly the same, at 33.8 and 33.5 Mbps, the two networks differ in the burstiness at smaller timescales.

Rather than have each operator provision for the peak bandwidth usage, we next evaluate three different scenarios where the downlink capacities of both operators are capped to their 80th, 90th, and 95th percentiles of bandwidth usage, which are approximately 37, 39, and 41 Mbps. Figure 6 shows the aggregate downlink bandwidth when the capacities are capped to 41 Mbps. The time series (not shown) looks similar when a cap of 37 Mbps or 39 Mbps are applied. In all three cases, there are periods when the usage of one operator’s network reaches the limit while the other does not.

When resource sharing is enabled, network usage is distributed fairly equally across both operators, as shown in Figure 7 for the case of 41 Mbps cap. This is a consequence of the sharing policy, which specifies that at high network load,

TABLE I
PERCENTAGE OF AMOUNT OF TRAFFIC SWITCHED TO ANOTHER OPERATOR FROM EACH ORIGINAL OPERATOR

Cap	37 Mbps		39 Mbps		41 Mbps	
Original Op	1	2	1	2	1	2
% switched traffic	52.0%	52.1%	47.9%	46.7%	37.7%	37.7%

TABLE II
AVERAGE TRANSFER RATES (KBPS) AT DIFFERENT PERCENTILES AND THEIR PERCENTAGES COMPARED TO THE NON-CONSTRAINED CASE FOR THE FIRST 200 SECONDS WITH BANDWIDTH CAPPED AT 37 MBPS.

	w/o constraint	37-Mbps constraint	
		w/o coop	w/ coop
25 th	4.69 (100%)	4.62 (98%)	4.84 (103%)
50 th	19.82 (100%)	18.91 (95%)	20.27 (102%)
75 th	52.63 (100%)	50.13 (95%)	53.57 (102%)
90 th	113.52 (100%)	107.05 (94%)	111.19 (98%)
95 th	170.37 (100%)	157.36 (92%)	161.14 (95%)
99 th	292.91 (100%)	276.49 (94%)	287.82 (98%)
99.5 th	363.07 (100%)	315.65 (87%)	345.7 (95%)
99.9 th	902.08 (100%)	618.64 (69%)	753.03 (83%)

traffic is assigned to the lesser loaded network. Further, when the cap is set to 37 Mbps (not shown), usage of both networks remains close to the cap for the entire duration, which indicates that the cap is relatively too strict.

Table I summarizes the amount of traffic switched to the operator other than the user’s home operator. Since the amount of traffic generated by users of either operator is roughly the same, as is the fraction of switched traffic, both operators get assigned a similar amount of total traffic. The amount of switched traffic is highest when the cap is 37 Mbps, when the network is congested the most.

In a real deployment, when a user uses the service of another provider, the home operator needs to pay that provider for the network usage. Our results show high percentage of switch of traffic in general because our policy can cause unnecessary switch. Nevertheless, both operators still receive the same amount of traffic in the end. Hence, neither operator is at a disadvantage.

B. Quality of Service

Next, we measure quality of service by each session’s average downlink transfer rate. We compute the average transfer rate of a session by dividing the total bytes downloaded by the total active time, i.e., the duration of the session minus the periods when the session is idle. Note that we do not include latency of the resource reservation procedure to the transfer rate calculation.

For this evaluation, we focus on the first 200 seconds of the experiment when bandwidth is capped at 37 Mbps. Table II compares the distribution of average transfer rates of sessions during this period, in the three cases evaluated: 1) without any bandwidth cap, 2) with bandwidth capped but without any cooperation between operators, and 3) with bandwidth capped and cooperation enabled. Though the average transfer rates are similar at lower percentiles in the three cases, at

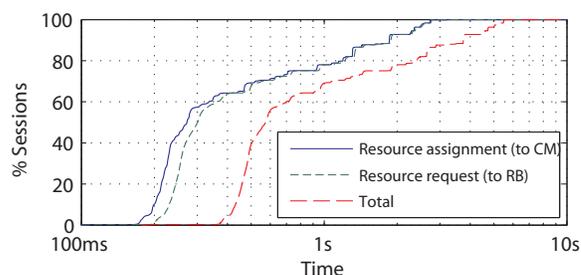


Fig. 8. CDF of session’s resource reservation overheads in log scale

the 99.9th percentile, performance measured as a fraction of that obtained when no cap is applied considerably improves from 69% without cooperation to 83% with cooperation. Thus, when provisioned capacity is limited, operator cooperation can help improve user performance close to that when there is no capacity constraint. Note that we do not measure the packet loss separately because it is already accounted for by TCP throughput.

C. Resource Reservation Overheads

Next, we evaluate the overheads imposed by the resource reservation framework. Figure 8 shows the CDF of per-session latency overhead experienced at the beginning of each session for the 37 Mbps constraint case. These distributions are independent of the bandwidth cap because the control traffic has its own dedicated channel. This overhead includes two XML-RPC calls from the client—first to the RB, and second to the CM at the assigned operator. The overhead distribution is heavy-tailed with a relatively high median of 573 ms for the total latency.

The average processing time at the RB and CM is only 43 ms in total. The dominant component of the remaining overhead latency is the 3G wireless link latency. Since 3G clients already experience a significant connection setup delay, it may be possible to merge the delays imposed by the resource reservation system into the existing setup routine.

Note that the high setup delay can affect the performance of resource sharing since it introduces a gap between when an operator is chosen for the user and when the user can actually access the network. However, in our experiments, we do not see significant gain when the client’s link latency is ignored for control traffic.

With regard to resource donation, the average latency for SMs to make a donation to the RB is 32 ms with relatively low deviation since these components are connected by wired links.

The control traffic accounts for only about 0.8% of data traffic on average. In a real deployment, the RB, which both processes resources donations and matches requests from users with resources, could be a bottleneck. However, it is possible to deploy multiple RBs, each of which is responsible for a subset of sites, and let them exchange information periodically.

D. Alternate Scenarios

We also explored an alternate resource sharing policy which assigns a session to the other operator only when the home operator has reached maximum capacity. However, switching occurs only when congestion has already set in. Congestion can be hard to detect when the traffic gets bursty at a timescale smaller than the monitoring interval. As a result, this policy is not as effective as we hoped.

Also, another potential scenario is to leverage burstiness in traffic to reduce the contention at the base station (radio resources) by shedding the load onto the base station of other service providers within a cell or sector area. We expect to see more potential benefits of operator cooperation from this scenario, since users typically tend to experience drop in network quality due to congestion at the base station. We leave this for future work when we obtain more detailed trace data and has better emulation tool.

VII. ALTERNATE AND FUTURE ACCESS TECHNOLOGIES

In this section, we compare the performance delivered by 3G, WiFi and LTE access networks for various application classes. Although each access technologies have distinct parameters in terms of bandwidth, we try to simulate and measure the real throughput perceived by users by exploiting Swing capability to capture and generate traffic and network model based on realistic traffic.

A. Experimental Setup

First, we classify the traffic in our trace data (Section IV) into one of the following application classes. The percentage within braces indicates the fraction of total number of bytes accounted for by the respective application class.

- **HTTP (34%)**: HTTP and HTTPS protocols. This can also include Flash video streaming under HTTP protocol.
- **Streaming (5%)**: RTSP (Real Time Streaming Protocol), RealPlayer, Quicktime, and Shoutcast streaming. Some protocols send data separately via UDP which is not included in this class.
- **P2P (40%)**: BitTorrent, eDonkey, and Gnutella.
- **Other TCP (21%)**: Any other TCP applications, such as email, file transfer, and instant messaging. However, these known protocols constitute less than one percent of total traffic. Majority of this class comprises traffic on non-standard ports which we were not able to classify.

We use Swing's dumb-bell topology and modify the characteristics of links on the client's side based on the access technology being studied.

a) *3G Link Characteristics*: For 3G, we use the same characteristics as described in Section IV, capping the downlink and uplink bandwidth to 3.6 Mbps and 512 Kbps.

b) *WiFi Link Characteristics*: We used Swing to estimate WiFi link characteristics from the WiFi tcpdump traces from the Jigsaw project [23], [24]. The traces, captured at the gateway that interfaces the UCSD campus network with the CSE wireless VLAN, consist of aggregate traffic from 40 802.11b/g access points. There are about 46 users using the

network on average in every minute, producing an aggregate bandwidth of 2.8 Mbps on average. We estimated only uplink capacities and used it for both uplink and downlink because Swing highly overestimates the downlink capacity in this case. The median link capacity is 18.6 Mbps and 95% of users have capacity less than 30 Mbps. For link delay, the median is 1.39 ms and 95th percentile is 4ms. These characteristics are significantly better than those of 3G access. However, approximately 16% of users experience packet loss, which is higher than with 3G. We assigned the link characteristic values to clients based on these distributions and used Swing to generate traffic as before.

c) *LTE Link Characteristics*: Since LTE technology is still not in commercial service, real trace data is not available. Thus, evaluations of LTE published so far are mostly based on simulations and vary significantly depending on the simulation parameters. Consequently, we chose to estimate LTE link characteristics by scaling the estimated 3G link characteristic distributions based on the 3G and LTE specifications. For link bandwidth, the maximum downlink bandwidth that clients in our 3G trace can achieve is no more than 2-3 Mbps due to the early version of HSPA. On the other hand, LTE can support up to 100-300 Mbps in an ideal condition. Consequently, we scale the bandwidth of all clients by a factor of 20. For link latency, 3G WCDMA/HSPA has RAN RTT of approximately 100-150 ms, while the RTT of LTE can be as low as 10 ms. Therefore, we scale down all 3G link latencies by a factor of 10.

B. Experimental Results and Analysis

Table III compares various percentiles (across sessions) of average transfer rates for different application classes with 3G, WiFi, and LTE. We see that the median transfer rates with WiFi are roughly double those with 3G. The differences are even greater for higher percentiles. This is due to a big difference in both latencies and capacities of clients between the two access technologies. Median link latency of WiFi is almost two orders of magnitude lower than 3G, and median bandwidth of WiFi is greater by about a factor of 5.

On the other hand, LTE can deliver performance comparable to WiFi in most cases. The streaming application class can even perform better using LTE in many percentiles. Although the link capacities of LTE clients are better than those of WiFi clients, the link latencies are not. Consequently, streaming applications which involve large volume transfers can gain more benefits from LTE's higher bandwidth, while WiFi's better latency has greater impact on performance for applications which comprise small flows.

Our estimates of WiFi link characteristics are extracted from traces captured from a lightly-loaded network. When the number of users is increased, we expect to see a significant drop in performance for WiFi. On the other hand, LTE has been developed to scale much better than WiFi. Therefore, LTE is likely to outperform WiFi under heavy load.

Overall, we conclude that WiFi is a good candidate to augment 3G services in a small highly-populated area such

TABLE III
DISTRIBUTIONS OF SESSION'S AVERAGE TRANSFER RATES (KBPS) BY APPLICATIONS WITH 3G/UMTS, WiFi 802.11b/g AND LTE ACCESS TECHNOLOGIES.

	HTTP			P2P			Streaming			Other TCP		
	3G	WiFi	LTE	3G	WiFi	LTE	3G	WiFi	LTE	3G	WiFi	LTE
25 th	11.2	20.0	18.9	7.7	15.9	14.4	14.7	34.8	32.9	1.5	3.4	3.0
50 th	32.8	57.0	54.5	20.7	51.3	44.7	473.1	714.9	774.1	6.8	14.0	12.6
75 th	78.9	160.4	148.5	45.1	143.4	112.1	1604.9	3502.8	3391.9	21.0	46.8	41.2
90 th	166.3	434.6	383.7	87.1	358.8	240.5	2219.0	6783.0	6876.7	47.1	118.6	99.2
95 th	272.1	785.5	671.7	123.8	678.5	392.3	2381.6	7510.9	7774.6	71.4	221.9	177.3
99 th	851.7	2661.5	2369.9	225.3	1422.0	906.3	2756.1	10267.1	9896.3	178.6	684.0	526.0

as a coffee shop or an airport, so operators can save the cost of investing in more expensive 3G infrastructure. On the other hand, we expect the performance of LTE to be comparable to or even better than WiFi, especially when the number users increases. Another key benefit of LTE is the ubiquity of coverage. Therefore, LTE can be a good candidate to replace 3G as well as the current WiFi in the future. We are yet to evaluate the 802.11n standard which is increasing in prevalence. We believe both LTE and 802.11n are potential access technologies that can dominate the mobile communication world in the future.

VIII. CONCLUSION AND FUTURE WORK

In this paper, we analyzed the benefits of resource sharing across mobile networks via operator cooperation. We leverage the difference across networks in the traffic burstiness at small timescales to shed the peak data bandwidth usage of one operator onto another. Our results, based on traffic generated from a 3G network trace, show that even when the capacities of individual networks are limited to less than the peak bandwidth usage, operator cooperation can help maintain quality of service for most of the user's sessions. As a result, all operators can benefit from cooperation by relaxing requirements on capacity provisioning. We also investigated the performance delivered by 3G, WiFi, and LTE for various application classes. Our results show that WiFi can provide significantly better performance compared to 3G. On the other hand, we expect the performance of LTE to be comparable to or even better than WiFi, especially when the number of users increases.

There remain several interesting aspects to explore further such as the business model for network composition. In this paper, we have focused on reducing operator costs but this could in turn lead to lowering of user costs as well. With operator cooperation, the notion of home operator or service provider could become less significant. A user's service provider could be a virtual operator which bundles different access networks and exports a seamless interface to users. On the other hand, users themselves could be given the power to choose amongst any available access network. In either case, new models of business competition and new pricing strategies will come into a play.

REFERENCES

[1] Ambient Network Partners, "Ambient networks," <http://www.ambient-networks.org>.

[2] M. Al-Fares, M. Johnsson, P. Johansson, and A. Vahdat, "Flexible resource allocation and composition across GSM/3G networks and WLANs," in *ACM MobiArch*, 2008.

[3] K. Vishwanath and A. Vahdat, "Swing: Realistic and Responsive Network Traffic Generation," 2009.

[4] M. Johnsson *et al.*, "An system description (d18-a.4)," February 2008, <http://www.ambient-networks.org/deliverables.html>.

[5] D. Zhou *et al.*, "Validated composition and compensation architecture (d26-g.2)," December 2007, <http://www.ambient-networks.org/deliverables.html>.

[6] R. Jennen *et al.*, "System evaluation results (d27-h.5)," December 2007, <http://www.ambient-networks.org/deliverables.html>.

[7] P. Pöyhönen, J. Markendahl, and O. Strandberg, "Analysis of benefits of operator cooperation using end-user and operator performance metrics," in *ITS*, 2008.

[8] K. Johansson, A. Furuskar, and C. Bergljung, "A Methodology for Estimating cost and Performance of Heterogeneous Wireless Access Networks," in *IEEE PIMRC*, 2007.

[9] M. Afanasyev, T. Chen, G. Voelker, and A. Snoeren, "Analysis of a mixed-use urban wifi network: When metropolitan becomes neapolitan," in *IMC*, 2008.

[10] V. Brik, S. Rayanchu, S. Saha, S. Sen, V. Shrivastava, and S. Banerjee, "A measurement study of a commercial-grade urban WiFi mesh," in *IMC*, 2008.

[11] W. Tan, F. Lam, and W. Lau, "An Empirical Study on 3G Network Capacity and Performance," *INFOCOM*, 2007.

[12] M. Kohlwes, J. Riihijarvi, P. Mahonen, F. GmbH, and G. Dusseldorf, "Measurements of TCP performance over UMTS networks in near-ideal conditions," in *IEEE Vehicular Technology Conference*, 2005.

[13] P. Svoboda, F. Ricciato, W. Keim, and M. Rupp, "Measured WEB Performance in GPRS, EDGE, UMTS and HSDPA with and without Caching," in *IEEE WoWMoM*, 2007.

[14] E. Dahlman, H. Ekstrom, A. Furuskar, Y. Jading, J. Karlsson, M. Lundevall, and S. Parkvall, "The 3G long-term evolution-radio interface concepts and performance evaluation," in *IEEE Vehicular Technology Conference*, 2006.

[15] H. Kaaranen, S. Naghian, L. Laitinen, A. Ahtiainen, and V. Niemi, "UMTS Networks: Architecture, Mobility and Services," 2005.

[16] J. Bannister, P. Mather, and S. Coope, "Convergence Technologies for 3G Networks: IP, UMTS, EGPRS and ATM," 2004.

[17] E. Dahlman, S. Parkvall, J. Skold, and P. Beming, *3G evolution: HSPA and LTE for mobile broadband*. Academic Press, 2007.

[18] N. S. Foundation, "Geni: The global environment for network innovations," 2009, <http://www.geni.net/>.

[19] T. Anderson and A. Vahdat, "Geni distributed services," November 2006, <http://groups.geni.net/geni/attachment/wiki/OldGPGDesignDocuments/GDD-0%6-24.pdf>.

[20] A. Vahdat, K. Yocum, K. Walsh, P. Mahadevan, D. Kostić, J. Chase, and D. Becker, "Scalability and accuracy in a large-scale network emulator," *ACM SIGOPS*, 2002.

[21] P. Abry and D. Veitch, "Wavelet analysis of long-range-dependent traffic," *IEEE Transactions on Information Theory*, 1998.

[22] 3GPP, "Radio link control (RLC) protocol specification (Release 5) TS 25.322," 2002.

[23] Y. Cheng, P. Verkaik, P. Benkö, J. Chiang, A. Snoeren, S. Savage, and G. Voelker, "Automating cross-layer diagnosis of enterprise wireless networks," *ACM SIGCOMM*, 2007.

[24] —, "UCSD CSE wireless traces," 2007, <http://sysnet.ucsd.edu/wireless/traces/sigcomm2007/>.