QuickC: Practical sub-millisecond transport for small cells

Rakesh Misra
Stanford University
rakeshmisra@stanford.edu

Aditya Gudipati
Stanford University
adityag1@stanford.edu

Sachin Katti
Stanford University
skatti@cs.stanford.edu

ABSTRACT

Cellular operators want to be able to deploy small cells with the same ease as WiFi access points, to quickly address traffic hotspots in dense urban areas. However, deploying small cells has not been easy. The reason is that due to scarcity of licensed spectrum, small cells need to use the same spectrum as the existing macro cells, and they need to explicitly coordinate their spectrum usage with each other to manage interference. The challenge is that this coordination needs to happen with latencies less than a millisecond, otherwise adding small cells does not help scale the overall network capacity. Implementing such tight coordination in dense urban deployments has not been easy in practice.

We present QuickC, a wireless transport technology that can simplify small cell deployment. QuickC enables small cells to coordinate with their neighboring macro and small cells with sub-1ms latencies over the operator's licensed spectrum but in a way that the users of this spectrum are negligibly affected. QuickC is designed to be an "add on" to existing cellular networks and does not require any invasive changes to the existing infrastructure or standards. We implement QuickC on a commodity system on chip from Texas Instruments used for building commercial cellular baseband and show that it can consistently deliver latencies less than 0.6 ms between neighboring cells. We also use system-level simulations to evaluate the wide-area impact of deploying QuickC in current networks, and show that it can deliver 5−10 Mbps of bandwidth between neighboring cells while causing less than 1−4% throughput degradation to the existing users of the spectrum.

CCS Concepts

•**Networks** → **Wireless access points, base stations and infrastructure;** *Physical links; Mobile networks;*

Keywords

Backhaul; Small cell; Ultra low latency; Air interface; Sub-1ms coordination; Interference management; Concurrent transmission; Successive interference cancellation; Densification; Coordinated multipoint; CoMP; LTE

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

MobiCom'16, October 03-07, 2016, New York City, NY, USA

© 2016 ACM. ISBN 978-1-4503-4226-1/16/10...$15.00

DOI: http://dx.doi.org/10.1145/2973750.2973773

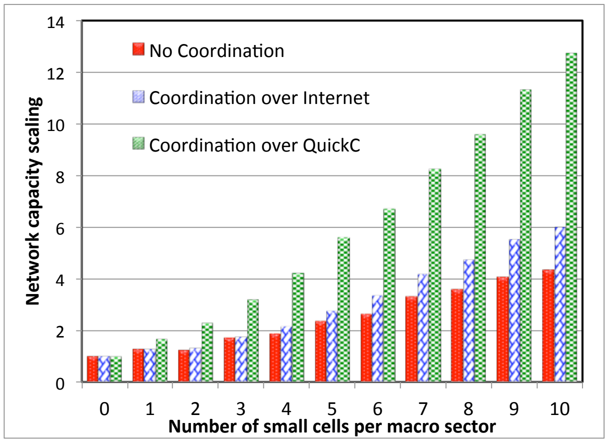


Figure 1: Network capacity with small cells (capacity of a macro-only network is normalized to 1). If small cells do not coordinate with their neighboring cells, adding more small cells per macro causes more interference in the network and therefore does not increase network capacity linearly. If small cells coordinate over a sub-1ms link like QuickC, the network capacity can be around 100% more than if they do so over public Internet! The simulation parameters are summarized in Table 2 (more details in Section 2).

1. INTRODUCTION

Ideally, deploying an LTE small cell should be as easy as deploying a WiFi AP. Deploying a WiFi AP is relatively simple: find power and an Internet connection, connect the AP and in almost all cases that is enough. Mobile network operators would like to have the same ease and simplicity for deploying small cells, it would make it much easier for them to quickly address traffic hotspots that arise in downtown areas, conferences, bus stops and so on. In fact that was the original vision and motivation for small cells, but deploying them has turned out to be significantly more complex. The primary reason is that operators have to ensure small cells can coordinate their spectrum usage with their neighboring macro and small cells[1] to manage interference, moreover the latency of this coordination needs to be less than 1ms round trip, see Figure 1[2]. This stringent latency requirement on the backhaul transport makes small cell deployment hard. Mobile operators have to ensure that the latency of any small cell to any other small cell or macro cell in the vicinity is less than 1 ms round trip. The average Internet connection that we might find in downtown or residential areas ex-

---

[1] especially with the macro cell which, due to its higher transmit power, subsumes the coverage area of the small cell.

[2] Several industry studies and field trials [5, 13, 26] have also reported similar results.

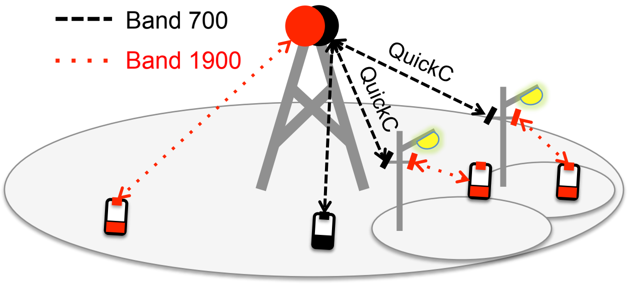


Figure 2: An example illustrating how spectrum is typically used in cellular networks and how QuickC fits in. Mobile operators typically own one sub-1GHz carrier and one or more 1-3GHz carriers (in this example, Bands 700 and 1900). The sub-1GHz carrier (Band 700) is used exclusively at the macro cell site as it is excellent for providing wide-area coverage. The 1-3GHz carrier (Band 1900) is used both at macro and small cell sites for additional capacity. The small cells and the Band-1900 macro cell interfere with each other and therefore need to coordinate their use of Band 1900. QuickC is a transport technology that can carry this coordination traffic to and from small cells over the sub-1GHz carrier (Band 700) with sub-1ms latency and with negligible interference to the users of Band 700.

hibits much higher latencies, typically in excess of 10-20 ms one way [21, 22], so mobile operators cannot just pick and lease an Internet connection from a wired ISP and install a small cell like they would for a WiFi AP.

Our goal in this paper is to ensure that small cells can be deployed with any commonly-available Internet connection and yet be tightly coordinated with their neighboring cells. We present the design and implementation of QuickC, a novel sub-1ms transport technology to *augment* whatever Internet connection is available for a small cell. This split-transport design is motivated by the observation that the small cell backhaul transport needs to carry two kinds of traffic each with a different set of requirements: (i) coordination traffic to and from neighboring cells which requires sub-1ms latency but only moderate bandwidth (∼5 Mbps per 20MHz 2×2 small cell [14]), and (ii) data traffic to and from the operator's core network which requires more bandwidth (∼50 Mbps) but is not sensitive to latencies on the order of $10-20$ ms. Our approach is to route the latency-insensitive traffic on the Internet backhaul whereas we design a novel QuickC link with sub-1ms latency and modest bandwidth to transport the latency-sensitive traffic.

QuickC's key innovation is to *reuse the existing licensed spectrum* of the operator to build a transport link between a small cell and its neighboring macro cell, see Figure 2 (neighboring small cells communicate in two hops). This might seem counter intuitive since licensed spectrum is an extremely expensive resource. In this paper, we make the following contributions:

1. a link-layer design that allows the coordination traffic to be communicated *concurrently* with the existing wireless traffic of the macro cell over the same spectrum, but in a way that the existing users of the spectrum are negligibly affected,
2. an air-interface design that allows QuickC to achieve an end-to-end round-trip latency of less than 1 ms while concurrently transmitting with LTE whose corresponding latency is nearly 20 ms,
3. a modular plug-and-play implementation using commodity equipment that requires no invasive changes to the existing cellular infrastructure or standards, thereby allowing QuickC to be an inexpensive solution.

We design and fully implement the entire baseband stack for a QuickC link on a commodity multicore system on chip used for designing LTE base stations and show that QuickC can consistently deliver an end-to-end round-trip latency of less than 600 μs between neighboring cells. We extensively evaluate the bandwidth of QuickC links and their impact on the macrocell access network, and show that QuickC can reliably deliver $5-10$ Mbps of bandwidth between neighboring cells in typical deployment scenarios while causing less than $1-4$% throughput degradation[3] to the existing users of the macro cell.

2. BACKGROUND AND MOTIVATION

To enable the vision of WiFi-like ease in LTE small cell deployment, a key question is whether a typical last-mile Internet connection is sufficient for LTE's needs. So it is useful to first understand the nature of such last-mile connectivity that can be used to backhaul LTE small cells in urban areas.

Last-mile Internet connectivity

To do so, we turn to measurements that the FCC conducts in all major US metro areas to characterize last-mile throughput and latencies [21, 22]. These measurements are at the IP layer between average broadband Internet connection points (such as cable, fiber, DSL) and a gateway node located in the same metro area. There are two key takeaways. *First*, average speeds are on the order of 50 Mbps and are likely to increase over the next few years as cable and FTTH (fiber to the home) deployments become more widely available. *Second*, average one-way last-mile latencies are at best 14 ms (cable), 15 ms (fiber) and 31 ms (DSL) [22, p. 43, chart 20], [21, pp. 48-49]. These last-mile characteristics reported by the FCC are representative of the kind of connections one might want to deploy small cells with, so this gives us a sense of what to expect in terms of the throughput/latency of links available to backhaul small cells.

Are such links sufficient for LTE small cells?

The backhaul connection connecting the small cell to the operator's core network serves two functions.

Data transport: The backhaul link carries mobile traffic from the small cell to the operator's core network, and back. In LTE jargon, this transport is referred to as S1 [6]. Typical LTE small cells generate 25 Mbps average and 80 Mbps peak S1 traffic [30, 14].

Coordination transport: The backhaul link is also expected to carry coordination traffic between neighboring base stations in a geographic area. This helps them coordinate interference and load states with each other to handle decisions such as interference management, handover and load balancing etc. In LTE jargon, this transport is referred to as X2 [7]. Peak traffic on X2 has been estimated to be less 5% of data traffic i.e., 4 Mbps per small cell [14].

From the above, it seems that commonly-available Internet connections can satisfy the bandwidth requirements of LTE small cells, both for data and coordination transports. However, their latencies turn out to be too high for LTE small cells. Specifically, latencies of several tens of milliseconds are fine for data (S1) transport, but the coordination (X2) transport needs to have a latency of *less than a millisecond*. We explain why.

Why is sub-1ms latency between neighboring base stations crucial?

The prime reason is that in dense deployments, *interference coherence times* and *traffic burst sizes* are both typically on the order of a

---

[3] In return, what networks can get with QuickC is nearly a 100% improvement in their overall network capacity, see Figure 1.

millisecond, so neighboring cells need to be able to act together at such fine time scales to best serve the clients at the intersection of their coverage areas. To give an intuitive sense for why interference varies so rapidly, think of a typical scenario where an operator has deployed 4-8 small cells per macro sector. In such scenarios, it is common for many clients to see strong channels from 2-5 cells, and the channel to any one cell decorrelates significantly (by more than 10%) typically in 4-5 ms, so the SINR (signal to interference and noise ratio, which is the combined effect of the channels to the different cells) of a typical client varies on the order of 1-2 ms. In addition, mobile data traffic is becoming very bursty; to give a sense, 80% of mobile web packets account for only 10% of the traffic volume [29] and last less than 1-2 ms. In such dynamic conditions, in order to serve each burst optimally, neighboring cells need to exchange network state information such as channel, interference and load measurements with each other faster than 1 ms, and adjust their scheduling and transmission parameters accordingly.

These observations are not novel, these are widely recognized in the LTE community and many coordination mechanisms exist to tackle such highly-dynamic interference and load. For example, a popular coordination mechanism that operators are actively evaluating for use in their future networks is Coordinated MultiPoint Joint Transmission (CoMP-JT) [4]. CoMP-JT is a mechanism by which multiple neighboring cells can transmit to multiple clients on the same time-frequency slot; they need to precode their transmissions using MIMO-like techniques (e.g. interference nulling) such that the SINRs at the clients is maximized. The abstraction is that distributed cells would mimic a virtual MIMO antenna array and turn interference into an advantage rather than a liability. To be able to implement CoMP-JT, each cell needs to exchange the channel measurements from the target clients with its neighboring cells so they can jointly compute the MIMO precoding vector needed for joint transmission. This exchange happens over the X2 interface between neighboring cells, and the latency of the interface dictates how effective CoMP-JT is, as we describe below.

Figure 1 shows how the capacity of a network employing CoMP-JT scales with density. When the X2 interfaces between neighboring cells are implemented over typical Internet links which have latencies in excess of 10 ms, the network capacity is more than 50% worse than if they are implemented over sub-1ms links like QuickC. The intuitive reason is that the CoMP-JT precoding vector calculation at each cell depends on the precise knowledge of channel states between the clients and neighboring cells. But as we described, the channel states in such dense networks become stale in 1-2 ms, hence by the time the channel measurements reach the neighboring cells via X2 over Internet, they are stale and the computed precoding vectors are incorrect which in turn leads to poor performance. 3GPP evaluations [5] and several industry studies [13, 26, 14] have reported similar results and concluded that *sub-1ms coordination latency*[4] is critical for small cell deployments to be successful.

The above analyses explain why one cannot deploy LTE small cells simply by hooking them to any available Internet connection in an urban area. Small cells need to implement the above tight coordination at least with the macro cell in whose coverage area they are deployed to manage interference, and often with several neighboring small cells. But an average Internet connection does not provide the sub-1ms latencies needed for such coordination.

3. QUICKC: OVERVIEW

QuickC is a novel transport link for small cells that enables them to coordinate with their neighboring macro and small cells with latencies less than 1 ms round trip. A small cell equipped with QuickC can be deployed simply by connecting to a power supply and any available Internet connection to serve as the data (S1) transport to the core network; QuickC provides the coordination (X2) transport to neighboring cells.

3.1 QuickC reuses the operator's licensed spectrum for coordination between small cells and its neighboring cells

The QuickC link is designed using the operator's licensed spectrum as the physical transport medium between any small cell and its serving macrocell. To understand this better, it is useful to understand how operators use their spectrum today to deploy macrocells and small cells.

Cellular frequencies: Cellular networks around the world are deployed in frequencies ranging between 700 MHz and 3000 MHz, in bands of 5, 10, 15 or 20 MHz (each such band is referred to as a *carrier*). For example, 4G LTE networks in the US are currently deployed using carriers in Bands 700, 850, 1700, 1900 and 2500 [40]. The sub-1GHz carriers, for example in Bands 700 and 850, owing to their better propagation, are well suited for providing wide-area coverage. However, sub-1GHz spectrum is extremely limited and expensive as most of it has already been licensed to other non-cellular services. There is relatively more spectrum available in 1-3 GHz, for example in Bands 1700, 1900 and 2500, although their coverage range is comparatively smaller.

Macrocell spectrum: The typical strategy for top cellular operators around the world is to own one sub-1GHz carrier and one or more 1-3GHz carrier(s), for e.g. in the US, AT&T, Verizon and T-Mobile own LTE carriers in Bands 700, 1700 and 1900, Sprint owns LTE carriers in Bands 850, 1900 and 2500 [40]. They use the sub-1GHz carrier to deploy macrocells nation wide for coverage, and augment it with one or more macrocells in the 1-3GHz carrier(s) for capacity where necessary.

Small cell spectrum: In dense metropolitan areas and other traffic hubs where these macrocells are not sufficient to meet the traffic demand, operators also deploy small cells at new ad hoc sites (like street poles, rooftops etc.) to add to their network capacity. Small cells typically use one of the 1-3GHz carriers which, owing to shorter propagation than sub-1GHz carriers, are better suited for their smaller coverage requirements.

Figure 2 illustrates a typical deployment scenario. In this example, the macrocells provide access to their clients over Bands 700 and 1900 while the small cells provide access over Band 1900. The Band-1900 macro cell and the two small cells interfere with each other and therefore need to coordinate their spectrum usage.

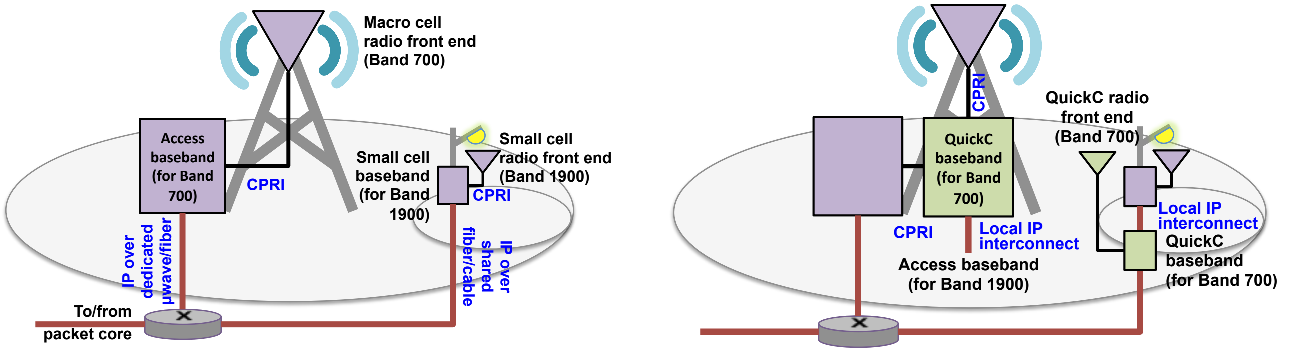
QuickC spectrum: QuickC consists of a new radio at each small cell that communicates with the macro cell using the sub-1GHz carrier[5]. For example, in Figure 2, the QuickC radios at the small cells use Band 700 as the medium for transporting coordination messages between the small cells and the Band-1900 macrocell. To do so, the small cell sends the coordination messages to the QuickC radio (over a local IP interconnect), the QuickC radio sends them to the Band-700 macrocell over Band 700, and the Band-700 macro cell in turn forwards them to the Band-1900 macrocell (over a local IP interconnect), and similarly in the other direction, see Figure 3b.

Note that the end points of a QuickC link in the example in Fig-

---

[4] Note that in general lower the coordination latency, better the performance. However in LTE the minimum unit of scheduling and transmission is 1 ms, so coordinated processing faster than 1 ms would not help since control cannot be applied faster than 1 ms.

[5] In general, QuickC can operate on any carrier that is exclusively used at the macro cell site and not by the small cell.



(a) Macro and small cells each consist of a baseband unit and a radio unit which are connected via an open interface like CPRI (over fiber). The baseband units interface with the backhaul transport network over IP. In relation to the example in Figure 2, the Band-1900 macro cell is not shown here for simplicity.

(b) Deploying QuickC involves adding a QuickC baseband unit with a QuickC radio unit at the small cell in series with its existing IP backhaul, and a QuickC baseband unit at the macro cell in series with CPRI. QuickC reuses the existing radio unit at the macro cell and requires no changes to the existing hardware/software units at the macro or small cell.

Figure 3: Illustration of how current cells are typically architected (left) and how QuickC can be inserted in current networks (right). Figure 4 shows more closely how the QuickC box at the macro cell site interacts with the existing boxes.

ure 2 are the QuickC radio and the Band-700 macro cell, so in the rest of the paper whenever we mention macro cell in the context of QuickC, it would imply the Band-700 macro cell. However, it is important to keep in mind that the traffic carried over QuickC is eventually going to/from the Band-1900 macro cell.

3.2 Potential benefits of reusing the operator's licensed spectrum for coordination

Reusing the operator's licensed spectrum to implement the coordination transport for small cells is attractive for several reasons.

Potentially very-low latency: Being a one-hop wireless link, it has the potential to be very low latency.

Ubiquity: It can be deployed universally because macrocell networks already exist and are architected to provide good coverage in urban non-line-of-sight environments.

Potentially inexpensive: It requires no additional spectrum (unlike microwave-based solutions) and no dedicated wired infrastructure to every ad hoc small cell location (unlike P2P fiber-based solutions), so it has the potential to be inexpensive.

3.3 Challenges with reusing the operator's licensed spectrum for coordination

One way of implementing QuickC over the operator's LTE spectrum might be to treat the new QuickC radio at the small cell as just another LTE user connected to the macrocell, probably even reserve a few dedicated subcarriers. However such a design has several drawbacks which makes it impractical.

Cannot achieve sub-1ms latency: The round-trip latency in LTE radio access network is in excess of 10 ms even excluding scheduling and retransmission delays [27, 31], so we cannot achieve sub-1ms latencies even if we reserve dedicated subcarriers for QuickC.

Huge impact on macrocell access: Macro cells have to now divide their spectrum resources between normal LTE users and QuickC. To provide a quick estimate of this overhead, current trends [33] suggest that there will be 4−10 small cells deployed per macro sector (a macro cell site typically has 3 sectors), and as discussed earlier each small cell requires up to 4 Mbps of bandwidth over the coordination transport, so each macro sector will need to reserve spectrum worth 16−40 Mbps on both uplink and downlink for small cells. A typical LTE macro sector serves around 160 Mbps on downlink and 80 Mbps on uplink, so this overhead (10−25% on downlink, 20−50% on uplink) will significantly affect the ability of macro cells to provide capacity and coverage.

Requires changes to existing infrastructure: Such a design is not modular and requires invasive changes to macrocell software (which is expensive and also means operators have to replace all their macro cells) as well as to existing LTE standards. That makes it an impractical choice for operators who are looking to densify their networks in the next few years [33].

3.4 QuickC's contributions and design goals

QuickC's approach is to divide the spectrum non orthogonally between LTE users and QuickC, in a way that is entirely transparent to the LTE users. To that end, QuickC makes the following contributions.

Underlay transmission/reception: QuickC contributes an underlay Tx/Rx technique that allows itself to coexist with the macrocell access network on the same licensed spectrum and yet have minimal impact (more in Section 4).

Sub-1ms air interface: QuickC contributes an ultra low-latency air interface that achieves sub-1ms latency round trip not just between a small cell and a macro cell but also between neighboring small cells (more in Section 5).

Prototype on commodity hardware: We show that QuickC can be implemented entirely using commodity hardware (more in Sections 6 and 7), this is important for QuickC to be inexpensive.

Accordingly, we set the following design goals for QuickC.

1. **Latency and bandwidth**: QuickC must deliver sub-1ms latency[6] from a small cell to any neighboring macro or small cell consistently. It must also provide up to 4 Mbps of bandwidth per small cell with LTE-like reliability so as to support all advanced coordination features of LTE.
2. **Impact on access**: QuickC's impact on throughput of LTE users in the macro access network must be minimal; as a reference, it must be an order of magnitude lesser than reserving dedicated channels for coordination i.e., less than 1−2.5% on the downlink and 2−5% on the uplink.
3. **Modularity**: QuickC must be a modular "add on" for existing macro and small cells, and must not require changes to the existing stacks or standards.

---

[6] The latency of each QuickC hop is measured as the delay between IP packets coming in to and out of the QuickC basebands at the two ends of the link.

3.5 QuickC's architecture and deployment

Before we explain the architecture of QuickC and how it can be deployed in current cellular networks, it is useful to understand how current cells are architected.

Existing cellular architecture: Figure 3a illustrates how current macro and small cells are architected. Every cell consists of a baseband unit and a radio unit. The baseband unit connects to the radio unit over an open interface like CPRI [18][7]. The baseband unit also interfaces with the backhaul transport network over IP. Small cells transmit at a much lower transmit power (30 dBm or 1 W) as they need to serve a much smaller area than macro cells (who transmit as high as 50-60 dBm or 100-1000 W).

QuickC architecture/deployment: Figure 3b shows how QuickC is architected and inserted in existing cellular networks. Deploying QuickC involves inserting the following components.

At the small cell: The QuickC unit at the small cell is inserted in series with its existing backhaul, and interfaces with the small cell unit over a local IP interconnect. This QuickC unit forwards the latency-insensitive traffic over its existing IP backhaul while switching the latency-sensitive traffic over the QuickC link. This unit consists of a *QuickC baseband unit* (Section 6) which runs on commodity SoC, and a radio front end. The *QuickC antenna* (Sections 4 and 7) on the radio front end is an off-the-shelf directional[8] antenna with gains in excess of 15-18 dBi and a form factor that is appropriate for installation with small cells, for e.g. [9] (the specifications for this antenna are studied in Section 7.2). The QuickC radio, as we explain in Section 4, transmits at a much lower power than LTE devices i.e., lower than 23 dBm.

At the macro cell: The QuickC unit at the macrocell is only a separate baseband unit that reuses the radio front end/antennas of the macrocell. The QuickC baseband runs on commodity SoC and sits between the macro cell's baseband unit and the radio unit, and interfaces with CPRI on both ends. The QuickC baseband is also accompanied by a *QuickC scheduler* which reads the measurements into and scheduling decisions out of the macro cell's scheduler every LTE subframe (1 ms) via a read-only interface, and accordingly decides how QuickC transmissions must be concurrently scheduled on the same licensed spectrum (Section 4). Figure 4 shows how the QuickC baseband unit interfaces with the existing units[9].

QuickC is architected such that it requires no invasive changes to any existing cellular infrastructure or standards.

4. QUICKC'S UNDERLAY TX/RX

The key ideas that allow QuickC to communicate its traffic concurrently with the LTE access traffic of the macro cell, and yet have a negligible impact on the performance of the macro cell's LTE access network, are the following:

(i) make the QuickC links stronger than the access links *by design*,

(ii) exploit the stronger QuickC links to transmit the QuickC traffic concurrently with the LTE access traffic but *at a much lower power*. We explain how below.

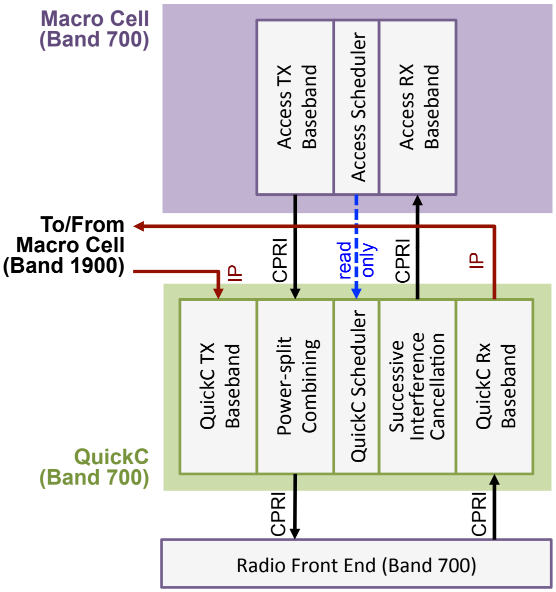


Figure 4: External interfaces of the QuickC box with the existing boxes at the macro cell site, ref. Figure 3b. QuickC can be inserted in a plug-and-play fashion, and does not require any invasive changes to any of the existing boxes.

How do we design a QuickC link to be much stronger than a typical access link? We do so by employing directional antennas for the QuickC radio at the small cell with gains in excess of 15-18 dBi, as described in Section 3.5. Note that LTE devices use antennas with much-lower gains, typically 0 dBi or lower, because of their smaller form factor and the constraint to be omnidirectional (to make performance independent of device orientation). But the QuickC radio on a small cell enjoys a larger form factor and has a static orientation so it is capable of using a high-gain directional antenna. In this way, a QuickC link is designed to be inherently 15-18 dB stronger than an access link[10].

Note that the QuickC radio does not need a line of sight (LoS) to the macro cell; QuickC uses the same LTE spectrum and works in the same non-LoS environments that are typical to LTE devices.

How do we exploit the stronger QuickC links? We exploit the stronger links to transmit the QuickC traffic at a lower baseband power than the access traffic so as to minimally affect the performance of the macrocell access network. In other words, the QuickC traffic is communicated as an *underlay* of the LTE traffic.

We explain the above in more detail for uplink and downlink separately, using a simple 3-node example in Figure 5 consisting of 1 macro cell, 1 LTE client and 1 small cell with QuickC.

How does underlay communication work on the uplink? With reference to Figure 5a:

(i) The LTE client transmits as before, totally oblivious to the presence of QuickC.

(ii) The QuickC client at the small cell transmits concurrently with the LTE client but at a lower baseband power (to limit the interference it introduces in the access network, as we describe later).

(iii) The macrocell radio receives a sum of the QuickC and the access signals, with the QuickC signal at a higher power than access due to its stronger link. The QuickC unit at the macro cell per-

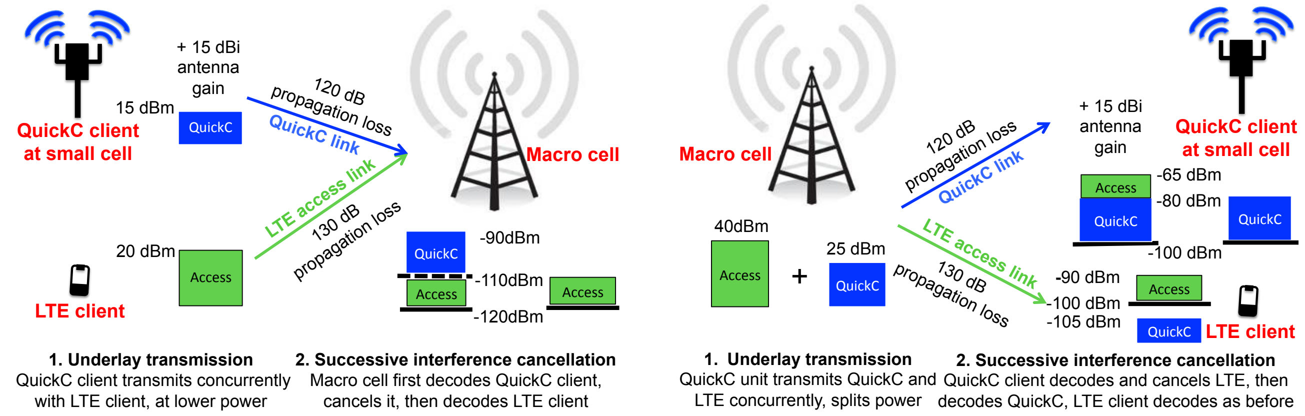
---

[7] CPRI stands for Common Public Radio Interface, and is an industry standard for the interface between a baseband unit and a radio unit to support vendor interoperability

[8] Note that QuickC antennas are required to be directional towards the macro cell not necessarily in the physical space but in the signal space so there is maximum energy transferred over the QuickC link.

[9] In this paper, we focus on FDD (frequency division duplexed) LTE which is the more popular version of LTE and where Tx and Rx happens simultaneously over different frequency bands. Although the principles of QuickC carry over to TDD (time division duplexed) LTE, we leave the implementation and evaluation of QuickC for TDD LTE to future work.

[10] In practice, the QuickC radio which is static and mounted 5-10 m above the street sees a better channel to the macrocell than an LTE device which is mobile and closer to the street. So the difference in link qualities is often more than 15-18 dB.



(a) Underlay communication on the uplink. The QuickC unit at the macro cell performs traditional SIC. The LTE client is oblivious to the presence of QuickC.

(b) Underlay communication on the downlink. The QuickC client at the small cell performs symbol-level SIC. The LTE client is oblivious to the presence of QuickC.

Figure 5: Conceptual illustration of QuickC's underlay communication.

forms successive interference cancellation (SIC) [11]; it decodes the QuickC signal by treating the access signal as noise, cancels it from the received signal and forwards it to the macro cell who then decodes the access signal as before.

How does underlay communication work on the downlink? With reference to Figure 5b:

(i) The QuickC unit at the macrocell combines the QuickC signal with the access signal from the macro cell such that the QuickC signal is at a very low power compared to access, and transmits the combined signal. For example, if the maximum transmit power of the macrocell radio is 100 W (50 dBm), the QuickC scheduler might allocate 1 W (30 dBm) to QuickC and the remaining 99 W (49.96 dBm) to access.

(ii) The LTE client decodes the access signal as usual, totally oblivious to the fact that there is a QuickC signal mixed in as interference. The QuickC scheduler controls this interference by adjusting the fraction of power allocated to QuickC.

(iii) The QuickC client at the small cell performs SIC; it first decodes the access signal (even though it is not intended for it), cancels it and then decodes the QuickC signal.

At first glance, it seems this is enough. The concept of SIC, which has been widely explored [11, 28] as a technique to decode concurrent transmissions when one transmission is received at a much-higher power level than the other, seems to solve the problem of making QuickC co-exist with LTE without having a significant impact. However, a careful analysis tells us that the following challenges still remain. In the following subsections, we explain each challenge and describe how QuickC solves them.

1. SIC is infeasible on the downlink, see Section 4.1.
2. Underlay transmission on the downlink causes interference at the LTE clients, see Section 4.2.
3. Underlay transmission and SIC on the uplink interfere with the macro cell, see Section 4.3.

4.1 Symbol-level SIC on the downlink to cancel LTE on a per-OFDM symbol basis

The challenge: SIC is not a feasible solution on the downlink for achieving sub-1ms latency. The reason is that LTE codewords span 1 ms (the minimum unit of transmission in LTE), which means that the QuickC client at a small cell has to wait for 1 ms to receive a complete LTE codeword before it can cancel it. As a result, the QuickC client cannot cancel LTE access signals faster than 1 ms, which in turn means that QuickC cannot achieve a sub-1ms round-trip latency if it uses SIC on the downlink[11].

QuickC's solution: QuickC solves this challenge by using a variant of SIC on the downlink that we term *symbol-level SIC*. The QuickC receiver at the small cell, after estimating and equalizing its channel, makes a hard decision about the LTE constellation symbols and cancels them at the symbol level. The QuickC scheduler signals the modulation format used for the LTE symbols in every resource block (which it reads off an interface from the LTE scheduler, see Figure 4) in the control channel for QuickC. There is however a chicken-and-egg problem; the QuickC client cannot decode the QuickC control channel before it has canceled the LTE signal, and it cannot cancel the LTE signal without knowing the LTE modulation format. To get over this, the QuickC client performs symbol-level SIC blindly for all possible modulation formats in LTE uplink (namely, QPSK and 16QAM), decodes the QuickC control channel in every resource block for each of the possibilities and then selects the format that matches the information in the QuickC control channel.

The reason hard slicing works for QuickC is because the LTE access signal is received at the QuickC client with a much-better link quality than it had been encoded for (for e.g., in Figure 5b, the QuickC client at the small cell receives the LTE signal at 15 dB where it was intended to be decoded by the LTE client at 10 dB). QuickC exploits this and simply cancels estimates of the LTE constellation symbols every OFDM symbol; this allows the QuickC client to decode QuickC on a per-OFDM symbol basis (as opposed to a per-codeword basis with traditional SIC) thereby making it possible for QuickC to achieve sub-1ms latency, see Section 5.

4.2 Intelligent power splitting on downlink to minimize impact on LTE

The challenge: The concurrent QuickC transmissions on the downlink cause interference at the LTE clients. For negligible impact, this interference must be kept well below their noise floors; it is impractical for QuickC to know the noise floor of every LTE client in the network.

QuickC's solution: QuickC solves this challenge by exploiting the knowledge of the channel measurements that the LTE clients report to the macro cell every 1 ms, which the QuickC scheduler reads off

---

[11] Note that this is not a problem on the uplink since the QuickC signal is decoded first.

an interface from the macro cell's LTE access scheduler, see Figure 4. With this knowledge, the QuickC scheduler can intelligently split the total available power at the macro cell between the LTE and the QuickC signals such that *the QuickC signal is received below the noise floor by design at each LTE client*, without requiring to know or estimate their noise floors. For example, in Figure 5b, to keep the QuickC signal 5 dB below the noise floor at the LTE client when its true SNR is 10 dB, the QuickC scheduler splits the total power such that the QuickC signal was transmitted 10 + 5 = 15 dB lower than the LTE access signal. In our implementation, we ensure that the QuickC signal is always received at least 6 dB below the noise floor at the LTE clients, this limits the access LTE SNR loss to 1 dB. As we show in Section 7, this in turn limits the median throughput degradation of LTE clients to 1%.

Note from Figure 4 that QuickC's power-split combining (PSC) block is part of its baseband unit that sits between the LTE access baseband and the RF front end of the macro cell, and performs this PSC in a manner that is transparent to the existing macro cell. This PSC has to be done differently for every pair of LTE and QuickC clients that are scheduled concurrently. A challenge then is to be able to perform PSC per subcarrier when the inputs to the PSC block are time-domain samples. To do so, the PSC block converts the time-domain samples back to the frequency domain, combines the symbols according to the power split and converts them back to time-domain samples.

4.3 Intelligent scheduling and power control on uplink to minimize impact on LTE

The challenge: *First*, SIC is not perfect in practice, even the best implementations of SIC [23, 24] that we are aware of have reported a residual interference of 1.5-2 dB after cancellation. *Second*, when QuickC is deployed in a multi-cellular network, its uplink transmissions act as a source of additional interference to neighboring macro cells which, according to our evaluation, can degrade the IoT (interference over thermal noise) by 4 dB or higher[12]. In addition, these degradations can be variable depending on the relative link qualities of QuickC and LTE. Such high and unpredictable degradation is catastrophic for macrocell LTE access networks.

QuickC's solution: QuickC solves these challenges by being conservative in how it chooses its schedules and its uplink transmit powers. In time-frequency slots where the macro cell's LTE scheduler has scheduled low-SINR LTE clients, the QuickC scheduler avoids scheduling any QuickC client whenever possible so as to not hurt the LTE links at all. In other time-frequency slots, the QuickC scheduler restricts the QuickC uplink transmit power such that *the QuickC signal is received at the macro cell radio at a constant level relative to the LTE signal* irrespective of the QuickC link strength, in other words it tries to keep QuickC's uplink SINR constant[13]. It does so for two reasons. First, the residual interference after SIC is a function of the QuickC SINR, so by keeping it low and constant, it causes a low and predictable degradation to LTE across all time slots and subcarriers (the amount of degradation can be controlled by tuning the QuickC SINR). Second, it also keeps the IoT rise due to QuickC across the entire network low and predictable. In our implementation, we keep the QuickC SINR constant at 15 dB, as we show in Section 7 this keeps the total degradation on the uplink less than 1.5 dB and the corresponding median throughput degradation less than 4%.

5. QUICKC'S ULTRA LOW LATENCY AIR INTERFACE

At the heart of QuickC is an ultra-low-latency air interface that makes *guaranteed* sub-1ms round-trip latency possible.

5.1 Interface between macro and small cell

The challenge: The challenge in designing such an ultra low latency air interface is that the scheduling, transmit processing, transmission, receive processing and retransmission if any on both uplink and downlink should all complete *predictably* within a total budget of 1 ms. To put things in perspective, the round-trip latency in LTE radio access networks is nearly 20 ms [27, 31] which includes processing and transmission delays of over 10 ms, scheduling delay of over 6 ms and an average retransmission delay in excess of 1 ms. WiFi, on the other hand, can potentially achieve sub-1ms round-trip latency if there is no contention, but since it is not a centrally-scheduled protocol, the latency can become unpredictably high in the presence of contention, often several milliseconds to tens of milliseconds [37].

QuickC's solution: QuickC solves this challenge by exploiting its unique design requirements to make the following key modifications to the LTE air interface. Note that this modified air-interface design is only for the QuickC link, this does not require any changes to or support from the existing LTE air interface or devices, or the existing software at macro/small cells.

Shorter transmission times: In LTE, the minimum unit of transmission is 14 OFDM symbols spanning a total of 1 ms which results in a processing plus transmission delay in excess of 10 ms round trip. QuickC is designed to be able to transmit and receive at the finest granularity, that of an OFDM symbol; in other words, QuickC's minimum unit of transmission can be as low as 1 OFDM symbol spanning 1/14th of a millisecond (71.4 μs). This might be a bad design choice for any bandwidth-intensive interface as it means paying a large bandwidth overhead (%age of non-data bits per transmission), but QuickC can make this trade off for latency as its bandwidth requirements are modest. Such shorter transmission times also mean that the average retransmission delay is now less than 100 μs round trip.

Pre-allocated resources for uplink session initiation: In LTE networks, at the start of any uplink session, the LTE clients do not know which time-frequency resources to transmit on, therefore LTE has defined scheduling request-and-grant procedures that take in excess of 6 ms just to initiate an uplink session. The reason LTE has these sophisticated request-grant protocols is to support an arbitrary number of clients with arbitrary traffic requirements. QuickC gets rid of this delay altogether by pre allocating resources to the QuickC clients at the small cells for initiating their uplink sessions. The reason QuickC can afford to do this is that the QuickC scheduler at the macro cell knows the QuickC clients in its coverage area that it needs to support. Specifically, it can be configured to know the number of small cells with QuickC, the average coordination traffic requirement per small cell[14] and their locations relative to the macro cell. As a result, the QuickC scheduler can pre allocate resources to the QuickC clients at the small cells instead of requiring them to request for resources every time they want to start an uplink session. Once an uplink session has been initiated, the QuickC

---

[12] This is not a problem on the downlink because the macro cell's total power is split to accommodate QuickC, so the total transmitted power in the environment is the same as before.

[13] This is in contrast to LTE's uplink power control which allows clients with better links to be served at higher SINRs, and is similar to the power control in CDMA where clients closer to the base station transmit at a lower power than clients farther away.

[14] The coordination traffic requirement of a small cell can be estimated based on its bandwidth and the list of LTE coordination features it supports.

scheduler takes over and can reschedule the QuickC clients potentially every 1 ms so that QuickC's underlay Tx/Rx has minimal impact on LTE clients.

QuickC's air interface inherits the orthogonal frequency division multiple access (OFDMA) time-frequency resource grid of LTE that has been designed to work well in urban non line of sight environments. QuickC also inherits LTE's modulation and coding scheme (MCS) selection, so it exhibits the same reliability as LTE in terms of bit error rate performance.

5.2 Interface between neighboring small cells

The challenge: With QuickC, neighboring small cells communicate in two hops - source small cell to macro cell and macro cell to destination small cell, or four hops per round trip. The resulting end-to-end round-trip latency between neighboring small cells with the design we have described so far is in excess of 1 ms.

QuickC's solution: The QuickC unit at the macro cell uses a special forwarding technique for frames destined to one small cell from another that helps cut the round-trip latency to less than 1 ms. Specifically, it performs what we term *noisy-symbol forwarding*; it (i) performs OFDM demodulation (FFT) to recover noisy versions of the QuickC constellation symbols from the QuickC client at the source small cell, (ii) maps them to new time-frequency slots if necessary to match rates, (iii) performs OFDM modulation (IFFT) to forward the noisy QuickC symbols to the QuickC client at the destination small cell. This saves heavily on latency since the QuickC unit at the macro cell does not need to run the channel decoder or encoder while forwarding which are computationally heavy.

The reason noisy-symbol forwarding works for QuickC is its strong directional links. Mathematically, if the SNRs of the two QuickC hops are α each, the effective SNR of the two-hop link with noisy-symbol forwarding is $\frac{\alpha^2}{\alpha+1}$, so when $\alpha \gg 1$ this effective SNR is $\approx \alpha$ which is the best that could have been achieved, with a traditional decode-and-forward strategy. Thus, noisy-symbol forwarding performs nearly optimally while saving heavily on latency.

6. QUICKC'S BASEBAND PROTOTYPE

We prototyped QuickC's baseband units for both small cells and macro cells on TCI6638K2K [39], a multi-core DSP system on chip (SoC) from Texas Instruments. This SoC is used for designing commercial LTE base stations and hence is representative of the commodity platform on which QuickC baseband will be deployed in practice. We used an off-the-shelf evaluation module, TMDSEVM6670LE [38], to develop and test our prototype with this SoC.

To implement QuickC's ultra-low-latency baseband and underlay Tx/Rx mechanism, we scripted 27 signal processing blocks (listed in Table 1) in around 6000 lines of C code. Next, we stitched together these blocks and computed how to schedule them on the available DSP cores using the recently-developed Atomix framework [12][15]; Atomix provides a high-level scripting interface to lay out which blocks will execute on which of the DSP cores and in what order. It also enabled us to realize the inter-block communication using explicit and predictable data transfer mechanisms, which is critical to ensuring that the strict timing guarantees of QuickC are met. In our current implementation, we only needed to use 4 of the 8 available DSP cores to fully realize each of the four QuickC baseband stacks (uplink Tx, uplink Rx, downlink Tx and downlink Rx).

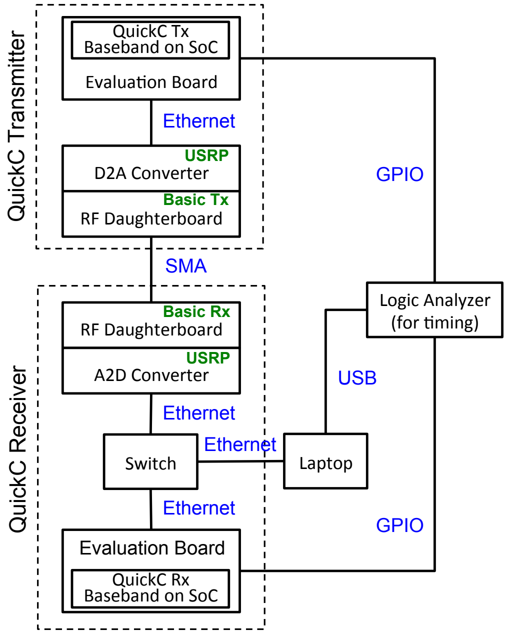


Figure 6: Block diagram of the experiment setup to evaluate the end-to-end latency of a QuickC link.

To realize the stitching together of the signal processing atoms, the data transfers and the scheduling of the DSP cores, we needed to write another 2500 lines of code in the Atomix scripting interface.

QuickC's baseband stacks execute with consistent ultra-low latencies on commodity DSP platform: Table 1 profiles the latencies of QuickC's baseband stacks on the TI DSP platform. To compute the execution times of the individual processing blocks in each stack, we added a timer script[16] to log the cycle count of each block, and then translated the cycle counts to microseconds (our DSP cores clock at 1.2 GHz). When a QuickC client at a small cell communicates with the QuickC unit at a macro cell, the baseband latencies of QuickC's Tx and Rx stacks are 53 and 80 μs respectively on the uplink, and 59 and 99 μs on the downlink, for a total of 291 μs round trip. When the QuickC clients of neighboring small cells communicate with noisy-symbol forwarding at the macro cell, the round-trip baseband latency is $(53+29+13+99)\times 2$ or 388 μs. Note that without noisy-symbol forwarding, this latency would have been 291×2 or 582 μs, so noisy-symbol forwarding improves the total baseband latency by a precious 194 μs. We repeated these measurements 100 times; the maximum variation in the latency of any stack was 1 μs. This proves that the latency of the QuickC baseband is not just low but also extremely consistent.

7. EVALUATION OF QUICKC AND ITS IMPACT ON LTE ACCESS NETWORKS

7.1 End-to-end latency of a QuickC link

Our objective in this section is to evaluate the end-to-end latency of a QuickC link. Figure 6 shows a block diagram of the setup for this experiment which represents one hop of a QuickC link end to end. The QuickC Tx/Rx baseband stacks, implemented on the TI DSPs, interface via Ethernet to USRP2 front ends which perform

---

[15] Atomix [12] is a framework to build modular signal-processing software with hardware-like performance guarantees on commodity multicore DSP platforms.

[16] Note that adding timers to atoms slightly increases their execution time, however we verify later using a logic analyzer (Section 7) that these timings are accurate to $\pm 2\%$.

Table 1: Execution times of QuickC baseband processing blocks while sending/receiving one QuickC frame of size 10 MHz × 1 OFDM symbol using QPSK mapping and 3/4 coding. The numbers (0−3) beside each block is the DSP core on which it executes. The estimates of QuickC round-trip latencies, obtained by adding the individual latency components, are independently verified end to end in Section 7.1.

Source small cell		**Macro cell**				**Destination small cell**	
Uplink Tx	**Timing (cycles)**	**Uplink Rx**	**Timing (cycles)**	**Downlink Tx**	**Timing (cycles)**	**Downlink Rx**	**Timing (cycles)**
0:CRC Insert	2700	*0:CFO Est-Ref*	*2030*	0:CRC Insert	2700	0:CFO Est-Ref	2030
0:Scrambler	20100	*0:CFO Correct-Ref*	*11000*	0:Scrambler	20100	0:CFO Correct-Ref	11000
1:Turbo Encoding	18436	*0:Tr Samples-Ref*	*6500*	1:Turbo Encoding	18436	0:Trans. Samples-Ref	6500
2:Puncturer	7500	*0:Symbol Read*	$*$	2:Puncturer	7500	0:Symbol Read	$*$
2:Interleaver	2584	*0:CFO Correct-Data*	*6200*	2:Interleaver	2584	0:CFO Correct-Data	6200
2:Mapper	3440	*0:OFDM Demod*	*8821*	2:Mapper	3440	0:OFDM Demod	8821
3:OFDM Mod	8821	0:Tr Data Tones	2020	*3:OFDM Mod*	*8821*	0:Tr Data Tones	2020
		1:Channel Est-Ref	(28915)	*3:Power Splitting*	*6200*	1:Channel Est-Ref	(28915)
		1:OFDM Equalizer	5803			1:OFDM Equalizer	5803
		1:Soft Slicer	4400			1:Hard Slicer	10327
		1:Deinterleaver	2584			1:Symbol Canceller	13000
		1:Puncture 1	2347			1:Soft Slicer	4400
		1:Puncture 2	(5200)			1:Deinterleaver	2584
		2:Viterbi Decode 1	20650			1:Puncture 1	2347
		2:Viterbi Decode 2	(13779)			1:Puncture 2	(5200)
		3:Descrambler Setup	(4315)			2:Viterbi Decode 1	20650
		3:Descrambler	20100			2:Viterbi Decode 2	(13779)
		3:CRC Check	2700			3:Descrambler Setup	(4315)
						3:Descrambler	20100
						3:CRC Check	2700
63581 (53 μs)		*34551 (29 μs)* + 60604 (51 μs)		54760 (46 μs) + *15021 (13 μs)*		118482 (99 μs)	

Additional latencies per hop = 2*20 μs (Front End) + 36 μs (Frame Alignment) + 71 μs (Transmission Time) = 147 μs

Macro to small cell latency = 280 μs (uplink, 53 + 147 + 29 + 51), 305 μs (downlink, 46 + 13 + 147 + 99), so **585 μs** round trip

Small cell to small cell latency = 488 μs (one way, 53 + 147 + 29 + 13 + 147 + 99), or **976 μs** round trip

LEGEND: $*$ = Counted separately as the Transmission Time, (.) = Not counted in total run time due to pipeline parallelism

Italic = The only blocks executed by macro during noisy-symbol forwarding (otherwise, it executes both italicized and non-italicized).

DAC/ADC. Basic Tx/Rx [20, 19] daughterboards provide SMA interfaces to the USRPs. We connect the Tx and the Rx daughterboards via SMA to emulate an AWGN channel without any multipath reflections (we test the impact of realistic multipath channels separately in Section 7.3 using an LTE channel emulator). Our experiments were conducted in baseband so our setup does not include any latency due to RF processing. We toggle a GPIO line on the Tx and the Rx SoCs at the start and end of packet processing respectively, and use a logic analyzer (Saleae Logic 8 [34]) across the two SoCs to measure the delays between these events.

QuickC's end-to-end round-trip latency is 574 μs between a small cell and a macro cell, and 964 μs between neighboring small cells: We tested the QuickC link with different modulation and coding schemes up to QPSK and 3/4 coding rate. The mean end-to-end latency was 275 μs on the uplink and 299 μs on the downlink, for a total of 574 μs round trip. The mean end-to-end latency in the small cell to small cell setup was 482 μs one way and 964 μs round trip[17]. Moreover, these latencies exhibited extremely-low variability, they were within ±2 μs of their means in all cases.

7.2 Performance of QuickC in LTE networks

Our objective in this section is to evaluate how QuickC performs in wide-area LTE networks comprising multiple macro and small cells. To do so, we set up a MATLAB framework to simulate a multi-cellular LTE radio access network (RAN) in the presence of QuickC. We rely on simulations because the wide-area impact of QuickC intrinsically depends on several factors such as the network layout, the spatial distribution of LTE devices, the antenna orientations etc. which cannot be recreated realistically in an academic testbed. We simulated the 3GPP Case 1 [3] scenario which is an industry reference for evaluating new technical proposals for LTE radio access networks and which is representative of the scenarios where QuickC is expected to be deployed. Table 2 summarizes the corresponding simulation settings.

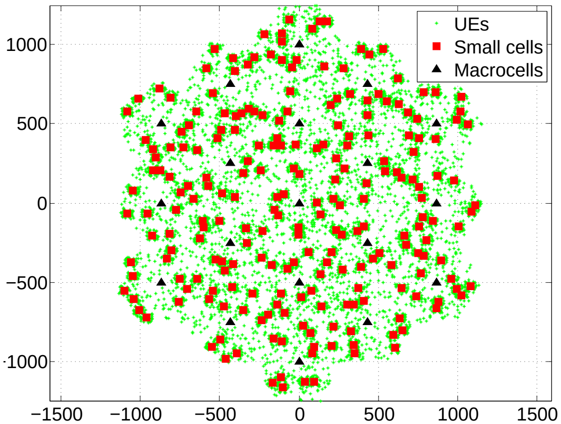


Figure 7: A snapshot of the simulated network comprising 19 3-sector macrocells and 4 small cells per sector, for a total 57 macro sectors and 228 small cells.

Deployment and traffic model: Figure 7 shows a snapshot of the simulated network. Each of the 19 macro cell sites consists of 3 120° sectors and 4 small (micro) cells per sector. In order to evaluate the *capacity* of QuickC links, in other words the maximum throughput they can provide if they had traffic to carry all the time, we used a full-buffer traffic model for QuickC. In addition, we wanted to stress test the performance of the QuickC links by maximally loading the access links, so we also used full-buffer traffic for the access links.

Channel models: We used a MATLAB implementation [35] of

---

[17] These results also validate the numbers we had estimated using our timer scripts in Table 1 which were 585 μs between small cell and macro cell, and 976 μs between small cells.

Table 2: Summary of simulation parameters

Parameter	Macrocell	Small cell	Client (UE)
Deployment	Uniform, 19 cell, 3 sector	Uniform at random, 4 per sector	P = 2/3 around a small cell
Height	32m	5m	1.5m
Tx Power	49 dBm	30 dBm	23 dBm
Antenna	0dBi/70°/70°	Access: 0dBi/omni/40° QuickC: 18dBi/45°/7°	0dBi/omni
Minimum Distance	Inter-macro: 500m	Macro-small: > 75m Small-small: > 40m	Macro-UE: > 35m Small-UE: > 10m
Channel Model (3GPP Spatial Channel Model)			
Macrocell	urban_macro with no line of sight		
Small cell	urban_micro no line of sight		
Traffic Models			
Access	Full buffer continuous		
QuickC	Full buffer continuous		
System Parameters			
Tx Mode	2 × 2 MIMO, Tx Mode 3		
Bandwidth	20 MHz @ 2 GHz		

3GPP Spatial Channel Models [8] to simulate the channel conditions. These are empirical channel models that 3GPP has defined for different LTE deployment scenarios to serve as references for benchmarking performance results in the LTE RAN. We used the `urban_macro` channel model with *no line of sight* to simulate the links to the macro cells. This 3GPP channel model captures the effects of path loss, shadowing and multipath reflections that are common in urban macro-cellular environments and realistically models the conditions on LTE as well as QuickC links. We refer our readers to [8] for more details of the model itself and to Table 2 for the heights, Tx powers and antenna settings used in our simulations for the different nodes. Note that our evaluation of QuickC explicitly ensures that there is no line of sight between small cells and their serving macro cells.

Table 3: Increase in uplink IoT in LTE networks due to QuickC

QuickC antenna		Average increase in uplink IoT (dB)		
Gain	**H/V HPBW**	**Q = 12dB**	**Q = 15dB**	**Q = 18dB**
18 dBi	20°/ 20°	1.09	1.58	2.34
18 dBi	45°/ 7°	0.29	**0.61**	1.34
25 dBi	7°/ 7°	0.02	0.04	0.04

Table 4: Sensitivity of uplink IoT to the orientation of the QuickC antennas at small cells

Maximum error (in degrees) (for an 18dBi/45°/7° antenna)	**Avg. increase in UL IoT** (for $Q = 15$ dB)
±0	0.61 dB
±1	0.70 dB
±5	1.00 dB

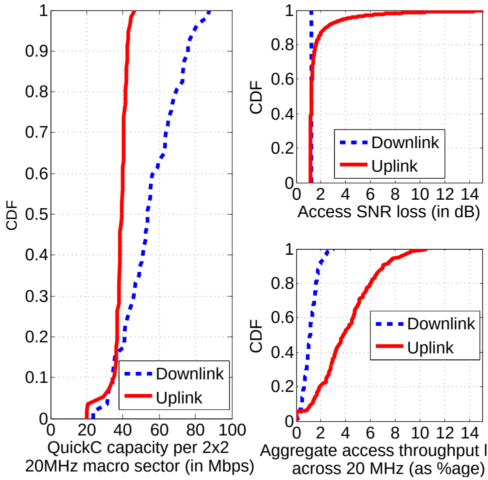


Figure 8: QuickC capacity per 20MHz 2×2 macro sector and its impact on LTE link quality and throughput.

QuickC's impact on uplink IoT and choice of antenna: Recall from Section 4.3 that QuickC transmissions on the uplink increase the IoT at neighboring macro cells which QuickC tries to keep low and predictable by restricting its SINR (say Q) to be constant regardless of its link quality. Table 3 presents the increase in uplink IoT for different values of Q as well as for different choices of QuickC antennas (all of them are specs of commodity LTE antennas with similar form factors). Lower Q-value means lower uplink transmit power for QuickC and therefore lower IoT, but it also means lower uplink capacity for QuickC. Similarly, higher directionality means less energy is leaked to neighboring cells, but the cost is a higher sensitivity to antenna orientation. Our evaluation suggests that a good practical tradeoff to choose for QuickC is $Q = 15$ dB and an 18dBi/45°/7° antenna. This increases the uplink IoT by only 0.61 dB while still allowing sufficiently-high QuickC uplink capacity and low antenna sensitivity as we show next.

Sensitivity of QuickC IoT to antenna orientation: What happens if the orientations of QuickC antennas at small cells happen to deviate from their best directions? To answer this question, we added a random error in both H- and V-planes of each QuickC antenna in our simulations. Table 4 presents our results. QuickC's power control algorithm now readjusts the uplink transmit power so as to maintain Q at 15 dB, so its capacity performance remains similar, however QuickC antennas may now be leaking more energy to their neighboring cells than before. Interestingly, even with a ±5° error (which, we found from RF field engineers of cellular operators, is a generous margin to assume for directional antennas in practical networks) in an antenna that has a V-plane HPBW of 7°, the IoT increases only from 0.61 dB to 1.00 dB. We discovered from our evaluation that a higher orientation error does not necessarily translate to a significant increase in IoT; IoT takes a hit only when the error is such that energy leaks directly into the directions of the neighboring macro sectors or cells.

QuickC's capacity and its impact on access: How much capacity can QuickC provide while still meeting its design goal of causing less than $1-2.5\%$ throughput degradation to LTE clients on the downlink and $2-5\%$ on the uplink? In order to stress test the capacity of QuickC, we maximally loaded the access network, as that is the most adverse setting for QuickC's capacity. Figure 8 plots the results.

Our simulations show that even in the most adverse setting, the QuickC unit at a macro sector can provide a median capacity of 56 Mbps on the downlink and 40 Mbps on the uplink, which corresponds to an average capacity of 14 Mbps and 10 Mbps respectively for the QuickC clients at each of the 4 small cells deployed per sector. Even with up to 10 small cells per sector, this translates to 5.6 Mbps and 4 Mbps of coordination transport bandwidth per small cell on downlink and uplink respectively, which comfortably meets the bandwidth goals we had set in Section 3.4. The corresponding median degradation in LTE link quality is 1.1 dB on both uplink and downlink, and this distribution is very tight around the median proving that QuickC is able to keep this degradation low and predictable. Recall that the QuickC scheduler avoids co-scheduling QuickC with low-SNR LTE client for whom a 1.1dB hit can be fatal, so this degradation is seen only by LTE clients with relatively good link qualities. The corresponding median degradation in their throughputs is 1% on the downlink and 4% on the uplink.

Recall that QuickC by design has similar bit error performance as LTE, namely 10^{-3}-10^{-6} with 1-2 retransmissions.

7.3 Impact on LTE links, on hardware

In the previous section, we showed using MATLAB simulations that QuickC causes a median degradation of 1.1 dB to the LTE links. In this section, our objective is to profile the same degradation but using our QuickC prototype, under a variety of LTE multi-path channel conditions.

Figure 9 shows our setup for this experiment which mimics the 3-node example in Figure 5. We use an LTE baseband generator and channel emulator [10] to generate baseband signals and apply LTE multi-path channels. We reuse our QuickC receiver setup from Section 7.1 to receive and decode the QuickC baseband signal. The USRPs in our set up have 12-bit ADCs, and our DSP implementation uses 16-bit representations for I/Q samples. For both uplink and downlink, we use the emulator to generate LTE-compliant I/Q samples to transmit over the LTE access link, and we load QuickC-compliant I/Q samples with appropriate power scaling to transmit over the QuickC link. On the uplink, the emulator applies a 2×1 channel and feeds in an analog baseband signal in real time to the QuickC receiver at the macro cell (USRP + DSP + laptop 1); the DSP decodes QuickC, performs SIC and ship the resulting I/Q samples after cancellation to laptop 1 which decodes the LTE signal using MATLAB. On the downlink, the emulator applies a 1×2 channel to the two I/Q sample streams and feeds two analog baseband signals, one to the QuickC receiver at the small cell (USRP + DSP) which performs symbol-level SIC and decodes QuickC, and the other to the LTE receiver (USRP + laptop 2) which decodes the LTE signal while treating QuickC as noise.

Figure 10 plots the degradation in LTE link quality (in dB) on both uplink and downlink as we increase the strength of the QuickC link, for a variety of LTE channel models. As we see, the mean degradation on the downlink is 1.1 dB, same as what we had seen in our simulations in Section 7.2. The mean degradation on the uplink however goes up to 1.8 dB from 1.1 dB. The reason is that the degradation on the downlink at the LTE receiver is independent of the SIC that happens at the small cell, on the uplink however the degradation is affected by SIC which in turn has more imperfections due to the quantization noise introduced by the QuickC receiver's ADC and the fixed-point operations on the DSP. Note that the degradation is fairly tight around its mean (limited to ± 0.2 dB) which is an artifact of QuickC's design.

8. DISCUSSION

Does QuickC require modifications to the existing scheduler at the macrocell? The access scheduler at the macrocell is a complex and proprietary piece of software that is difficult to modify, so we have designed QuickC such that the access scheduler requires no change whatsoever. As shown in Figure 4, all that QuickC needs from the access scheduler is a read-only interface, all vendors have such an interface internally for testing/debugging their schedulers. This interface however is not standardized yet, so while QuickC in its current version is plug and play from the perspective of any single vendor, it needs vendors to agree on a common definition for this interface if it has be vendor interoperable (when access and QuickC baseband units are manufactured by different vendors).

Does QuickC transmission at the small cell interfere with the small cell's own reception, and vice versa? No, as we had described in Section 3.1, QuickC operates on a different carrier (typically a sub-1GHz carrier) than the carrier used by the small cell (typically a 1-3GHz carrier).

9. RELATED WORK

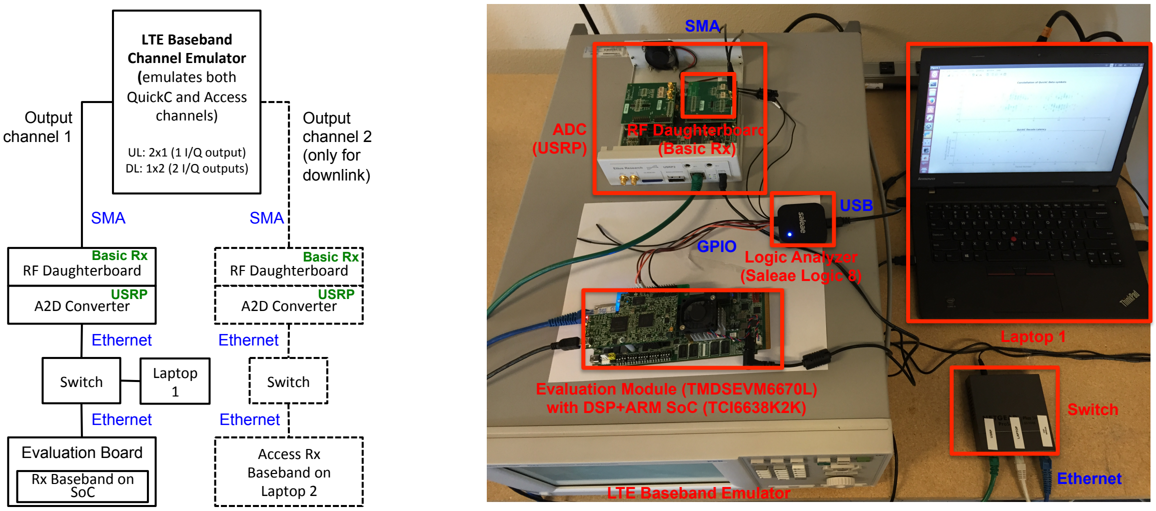
The principles and techniques used in the design of QuickC are related to prior work along the following axes.

Alternative technologies for sub-1ms transport: Point-to-point (P2P) fiber- and microwave-backhaul based sub-1ms transport already exist for macrocells [1, 2], however they are relatively expensive to be used widely for small cells [15]. The reason is that small cells do not bring operators as much revenue as a macrocell, so operators are looking for a cheap transport solution that will work universally for their small cells. P2P fiber does not exist universally and making it accessible in every urban location where operators want to deploy small cells is prohibitively expensive (note that public fiber is not sufficient as it has much higher latencies, see Section 2). Microwave-based backhaul do not work well in urban non line of sight (nLoS) geographies as those frequencies have very short propagation characteristics. QuickC on the other hand is inexpensive, requires no additional wired infrastructure or wireless spectrum, works well in similar nLoS environments as LTE since it uses the same spectrum, and offers a universal solution for operators as it works wherever macrocell networks exist.

Orthogonal dedicated control channel: The idea of having an orthogonal dedicated channel for control is common and is employed even in LTE [3]. However, as we argued in this paper, reserving dedicated orthogonal LTE subcarriers for carrying coordination traffic to/from small cells is impractical as it involves invasive changes to the existing macrocell software and cellular standards, apart from having a significant spectrum overhead. The highlights of QuickC are that it uses non-orthogonal scheduling to coexist with LTE so its spectrum overhead is negligible, and it is a modular add-on feature that does not require any invasive changes either to the existing cellular infrastructure or the cellular standards.

Concurrent transmissions and SIC: The concept of SIC has been studied extensively in the wireless community [11, 28, 32] and the idea of designing a side channel for control using concurrent transmissions has been explored earlier [16, 42]. These work involve *overlaying* control over regular transmissions by sending narrowband high-power control pulses. Such an approach however does not work for us as it involves changing all the existing LTE clients and macro cells to be able to perform SIC on downlink and uplink respectively to cancel out the stronger control signals before decoding their own signals. The closest work to QuickC's *underlay* mechanism is [41] which is able to achieve a bandwidth on the order of only 100 kbps over the ISM bands. QuickC's highlight is that it can stay entirely transparent to LTE while still offering bandwidths on the order of tens of Mbps over 20 MHz.

Sub-millisecond cellular air interface: There is considerable interest and several efforts in the industry to design and define a sub-1ms cellular air interface for 5G networks [17, 25, 36]. To the best of our knowledge, QuickC is the first-ever implementation of a sub-1ms cellular air interface that provides LTE-like throughput and reliability, and which also demonstrates that it is possible to realize such an air interface on commodity equipment.



(a) Block diagram of the setup (uplink and downlink). (b) Snapshot of our experiment setup for the uplink.

Figure 9: Experiment setup to evaluate the impact of QuickC on the quality of LTE links.

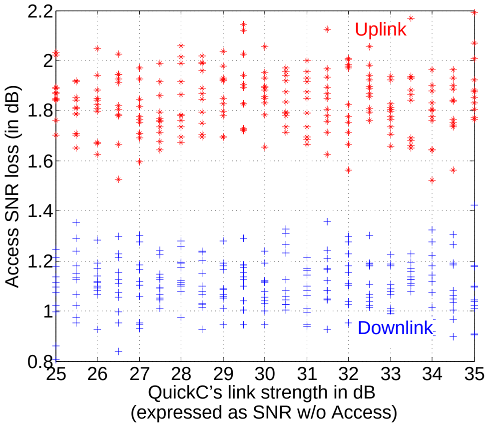


Figure 10: Degradation in quality of the LTE link (in dB) as a function of the strength of the QuickC link (in dB).

10. CONCLUSION

Historically, mobile access networks have scaled capacity along two broad dimensions, (i) by adding more spectrum, and (ii) by improving spectral efficiency (including MIMO). But advances along both these dimensions are now yielding diminishing returns. The next wave of capacity scaling is widely expected to come from *densification* - by deploying more cells per unit area and aggressively reusing spectrum to obtain spatial reuse gains in capacity. The ability to deploy small cells ad hoc in traffic hotspots is key to realizing this vision of densification. However, small cell deployment has not taken off as expected due to two main practical challenges: (i) acquiring sites in urban hotspot areas where small cells can be deployed for maximum impact, and (ii) finding an inexpensive backhaul solution in these areas that meets requirements.

QuickC directly addresses the backhaul challenge and offers a technology that promises to be an inexpensive solution which meets the bandwidth/latency requirements of small cell backhaul. Operators can use QuickC to deploy a variety of radio resource management applications inexpensively which in turn can lead to a tremendous increase in capacity, see Figure 1. While QuickC does not entirely solve the site acquisition challenge, by virtue of being a wireless backhaul that can augment any commonly-available Internet backhaul, it removes the requirement that a P2P fiber/microwave access point for sub-1ms latency be located near the small cell site, and thereby makes it considerably simpler for operators to find deployment sites for small cells. We therefore believe that QuickC will greatly simplify small cell deployment, and consequently enable operators to aggressively extend the capacity, coverage and capabilities of their radio access networks.

11. ACKNOWLEDGMENT

We would like to thank Manu Bansal for his help in prototyping QuickC using Atomix. We would also like to thank Aaron Schulman and members of Stanford Networked Systems Group (SNSG) for all the valuable discussions during the writing of this paper. We are thankful to our anonymous reviewers and shepherd for their useful comments and suggestions that helped bring the paper to its current form. The first author was supported by Stanford Graduate Fellowship (SGF).

12. REFERENCES

[1] Aruba Networks. http://www.arubanetworks.com.
[2] Cambridge Broadband Networks. http://cbnl.com/.
[3] 3GPP. Evolved Universal Terrestrial Radio Access (E-UTRA); Further advancements for E-UTRA physical layer aspects. TR 36.814, Mar. 2010.
[4] 3GPP. Coordinated multi-point operation for LTE physical layer aspects. TR 36.819, Sept. 2013.
[5] 3GPP. Coordinated multi-point operation for LTE with non-ideal backhaul. TR 36.874, Dec. 2013.
[6] 3GPP. Evolved Universal Terrestrial Radio Access Network (E-UTRAN); S1 general aspects and principles. TS 36.410, Dec. 2014.
[7] 3GPP. Evolved Universal Terrestrial Radio Access Network (E-UTRAN); X2 general aspects and principles. TS 36.420, Dec. 2014.
[8] 3GPP. Spatial channel model for Multiple Input Multiple Output (MIMO) simulations. TR 25.996, Sept. 2014.
[9] Advanced RF Technologies. AD-SBSP-1900-18-45HFB: PCS outdoor server antenna. http://adrftech.com/download/126/, 2011. [Online].

[10] Agilent. N5106A PXB Baseband Generator and Channel Emulator. http://www.keysight.com/en/pd-1455099-pn-N5106A/pxb-baseband-generator-and-channel-emulator. [Online].

[11] J. G. Andrews. Interference cancellation for cellular systems: A contemporary overview. *IEEE Wireless Communications*, 12(2):19–29, April 2005.

[12] M. Bansal, A. Schulman, and S. Katti. Atomix: A framework for deploying signal processing applications on wireless infrastructure. In *12th USENIX Symposium on Networked Systems Design and Implementation (NSDI 15)*, Oakland, CA, May 2015. USENIX Association.

[13] S. Brueck, L. Zhao, J. Giese, and M. A. Amin. Centralized scheduling for joint transmission coordinated multi-point in LTE-Advanced. In *ITG WSA*, 2010.

[14] Cambridge. Backhauling X2. http://cbnl.com/sites/all/files/userfiles/files/Backhauling-X2_0.pdf, 2012. [Online].

[15] Ceragon Networks. Mobile Backhaul: Fiber vs. Microwave: Case Study Analyzing Various Backhaul Technology Strategies. http://www.digitalairwireless.com/files/Fiber-vs-Microwave-White-Paper_1333235596.pdf, Oct 2009. [Online].

[16] A. Cidon, K. Nagaraj, S. Katti, and P. Viswanath. Flashback: Decoupled Lightweight Wireless Control. In *Proceedings of the ACM SIGCOMM 2012*, SIGCOMM '12, pages 223–234, New York, NY, USA, 2012. ACM.

[17] Ericsson. 5G Radio Access. http://www.ericsson.com/res/docs/whitepapers/wp-5g.pdf, Feb 2015. [Online].

[18] Ericsson AB, Huawei Technologies Co. Ltd, NEC Corporation, Alcatel Lucent, and Nokia Networks. Common Public Radio Interface (CPRI); Interface Specification v7.0. http://www.cpri.info/downloads/CPRI_v_7_0_2015-10-09.pdf, Oct 2015. [Online].

[19] Ettus Research. BasicRX Daughterboard 1-250 MHz Rx. http://www.ettus.com/product/details/BasicRX. [Online].

[20] Ettus Research. BasicTX Daughterboard 1-250 MHz Tx. http://www.ettus.com/product/details/BasicTX. [Online].

[21] FCC. 2016 Broadband Progress Report, Jan 2016.

[22] FCC's Office of Engineering & Technology and Consumer & Governmental Affairs Bureau. 2015 Measuring Broadband America Fixed Broadband Report: A Report on Consumer Fixed Broadband Performance in the US, 2015.

[23] A. Gudipati, S. Pereira, and S. Katti. AutoMAC: Rateless Wireless Concurrent Medium Access. In *Proceedings of the 18th Annual International Conference on Mobile Computing and Networking*, Mobicom '12, pages 5–16, New York, NY, USA, 2012. ACM.

[24] D. Halperin, T. Anderson, and D. Wetherall. Taking the Sting out of Carrier Sense: Interference Cancellation for Wireless LANs. In *Proceedings of the 14th ACM International Conference on Mobile Computing and Networking*, MobiCom '08, pages 339–350, New York, NY, USA, 2008. ACM.

[25] Huawei. 5G: A Technology Vision. http://www.huawei.com/5gwhitepaper/, 2013. [Online].

[26] R. Irmer, H. Droste, P. Marsch, M. Grieger, G. Fettweis, S. Brueck, H.-P. Mayer, L. Thiele, and V. Jungnickel. Coordinated multipoint: Concepts, performance, and field trial results. *Communications Magazine, IEEE*, 49(2):102–111, 2011.

[27] M. Laner, P. Svoboda, P. Romirer-Maierhofer, N. Nikaein, F. Ricciato, and M. Rupp. A comparison between one-way delays in operating HSPA and LTE networks. In *Modeling and Optimization in Mobile, Ad Hoc and Wireless Networks (WiOpt), 2012 10th International Symposium on*, pages 286–292, May 2012.

[28] N. I. Miridakis and D. D. Vergados. A Survey on the Successive Interference Cancellation Performance for Single-Antenna and Multiple-Antenna OFDM Systems. *IEEE Communications Surveys Tutorials*, 15(1):312–335, First 2013.

[29] C. M. Mueller. On the importance of realistic traffic models for wireless network evaluations. In *European Cooperation in the Field of Scientific and Technical Research*, Bologna, Italy, Nov 2010.

[30] NGMN Alliance. Small cell backhaul requirements, 2012.

[31] Nokia Siemens Network. The impact of latency on application performance. http://networks.nokia.com/sites/default/files/document/LatencyWhitepaper.pdf, Jul 2009. [Online].

[32] Y. Park. 5G Vision and requirements. In *5G Forum*, 2014.

[33] Qualcomm. 1000x: More small cells. Hyper-dense small cell deployments. https://www.qualcomm.com/media/documents/files/1000x-more-small-cells-.pdf, Jun 2014. [Online].

[34] Saleae. Logic8. https://www.saleae.com/. [Online].

[35] J. Salo, G. D. Galdo, J. Salmi, P. Kyosti, M. Milojevic, D. Laselva, and C. Schneider. MATLAB implementation of the 3GPP spatial channel model (3GPP TR 25.996). http://www.ist-winner.org/3gpp_scm.html. [Online].

[36] Samsung DMCR&D Communications Research Team. 5G Vision. http://cambridgewireless.co.uk/Presentation/RadioTech_30.01.14_HowardBenn.Samsung.pdf, 2013. [Online].

[37] K. Sui, M. Zhou, D. Liu, M. Ma, D. Pei, Y. Zhao, Z. Li, and T. Moscibroda. Characterizing and Improving WiFi Latency in Large-Scale Operational Networks. In *Proceedings of the 14th Annual International Conference on Mobile Systems, Applications, and Services*, MobiSys '16, pages 347–360, New York, NY, USA, 2016. ACM.

[38] Texas Instruments. TMS320C6670 Evaluation Modules. http://www.ti.com/tool/tmdsevm6670#tmdsevm6670le. [Online].

[39] TexasInstruments. TCI 6638K2K Multicore DSP+ARM KeyStone II System-on-chip. http://www.ti.com/product/tci6638k2k. [Online].

[40] Wikipedia. List of LTE Networks. https://en.wikipedia.org/wiki/List_of_LTE_networks. [Online].

[41] K. Wu, H. Tan, Y. Liu, J. Zhang, Q. Zhang, and L. Ni. Side channel: Bits over interference. In *Proceedings of the Sixteenth Annual International Conference on Mobile Computing and Networking*, MobiCom '10, pages 13–24, New York, NY, USA, 2010. ACM.

[42] X. Wu, S. Tavildar, S. Shakkottai, T. Richardson, J. Li, R. Laroia, and A. Jovicic. Flashlinq: A synchronous distributed scheduler for peer-to-peer ad hoc networks. In *Communication, Control, and Computing (Allerton), 2010 48th Annual Allerton Conference on*, pages 514–521, Sept 2010.