

Performance and Capacity Considerations for Mobile VoIP over WiMAX

Subir Varma

Technology Development

1. Executive Summary

1.1. Objective and Scope

The initial deployment of the Sprint WiMAX network (XoHM) features fixed VoIP service, in addition to mobile high speed data. However Sprint plans to introduce mobile VoIP over WiMAX at a future date, and one of the open questions is: How does the capacity of Mobile VoIP over WiMAX compare with other mobile VoIP alternatives, such as VoIP over legacy 1xRTT, VoIP over EVDO Rev A or VoIP over LTE. The objectives of this report are the following:

- To enumerate the various factors that influence the performance and capacity of mobile VoIP of WiMAX and their current status, either from the implementation or standardization point of view.
- Provide an overview of existing results on mobile VoIP over WiMAX system capacity, and how it compares with capacity results from other technologies.

2. Introduction

The deployment of all IP networks such as EvDO and WiMAX makes possible the migration of voice services over an all IP based infrastructure. Previous studies undertaken within Sprint have shown that this can result in large savings in network cost required to put in additional capacity for network expansion. Sprint is already along the path of testing mobile VoIP over EvDO Rev A networks, with the objective of deploying this service sometime in 2009.

Sprint plans to deploy fixed VoIP during the initial rollout of the WiMAX network, and is putting the Core Network infrastructure support required to do this. The Technology Development organization has initiated the process of investigating the issues involved in deploying a mobile VoIP service over WiMAX, and this report is a step in that direction. We focus on the issue of WiMAX VoIP system performance and capacity and enumerate the various factors that influence these. We also trace the evolution of additional features influencing VoIP capacity that are planned for WiMAX Release 1.X and IEEE 802.16m. We also provide a survey of some existing simulation results that provide VoIP capacity estimated for WiMAX, and compare these with known capacity estimated from other technologies such as EvDO Rev A and LTE.

3. Factors that Influence VoIP Capacity: WiMAX Wave2 Systems

In this section we enumerate the factors that influence VoIP capacity in WiMAX Wave 2 systems. These factors include:

1. The Uplink Downlink Split in TDD Systems: An asymmetric uplink downlink split reduces VoIP capacity, since VoIP is a symmetric application.
2. Header Compression: The data content of a VoIP packet is small, hence capacity can be increased by reducing the overhead to packet headers.

3. Choice of Codec: Using codecs that are designed specifically for wireless operation improves voice quality and also capacity.
4. Spatial Multiplexing and Collaborative MIMO: These multiple antenna technologies can increase VoIP capacity by increasing link capacities, but not in all cases, as explained in Section 3.4

The initial deployment of WiMAX VoIP will be based on the technologies present in Wave 2 systems. Simulation results presented below, show that the capacity of these systems is comparable to that achieved by 1xRTT or VoIP over EvDO Rev A systems.

Technology	Codec	Header Compression	Up:Dn Split	Capacity/Source
1xRTT	EVRC		1:1	95/Lucent
VoIP over EvDO Rev A	EVRC-B		1:1	123/Lucent
VoIP over WiMAX	EVRC	No	29:18	71/Nokia
VoIP over WiMAX	EVRC	Yes	29:18	100/Nokia
VoIP over WiMAX	AMR	Yes	1:1	140/NGMN
VoIP over WiMAX	G.729	No	29:18	45/Samsung
VoIP over WiMAX	G.711	Yes	29:18	32/Nokia

The capacity numbers in this table clearly show the influence of the first 3 factors in VoIP capacity. These numbers have been obtained without the use of uplink collaborative MIMO, which is expected to provide a further boost in capacity.

In the rest of this section, we provide further details on each of these determining factors.

3.1. Uplink Downlink Split in TDD Systems

The initial deployment of the XoHM system features an uplink-downlink split of 29 downlink symbols and 18 uplink symbols per frame. Since VoIP is a symmetric application, as a result the uplink constitutes a bottleneck for the voice capacity of the system. Hence the most straightforward way of increasing the voice capacity is by increasing the number of symbols in the uplink portion of the frame. Note that FDD systems feature a 1:1 split between the uplink and downlink, hence even without taking any other factor into account, just the 29:18 split puts the XoHM voice capacity at a disadvantage when compared to FDD based systems with the same voice carrying capacity.

Some more details on the uplink-downlink link capacity numbers for WiMAX are provided next: Of the 29 downlink signals, 7 to 11 symbols are taken up by the Preamble + Control packets, while on the uplink 3 to 6 symbols are taken up by control information. This leaves only 18 downlink symbols and 12 uplink symbols available for data transmission, in typical cases. The downlink and uplink data space is divided into 2-dimensional units called slots. There are 270 downlink slots and 140 uplink slots, in the example above. Each downlink slot can accommodate 30 bytes of MAC data, in the best case of 64QAM modulation at a 5/6 rate, which leads to a best case downlink rate of 12.96 mbps. The corresponding number for the uplink is 4.032 mbps (since the best case uplink modulation is 16QAM at a $\frac{3}{4}$ rate).

3.1.1. Requirements

Sprint should look into using a more symmetrical downlink uplink split, such as 23:24 symbols, to generate additional uplink capacity, in case symmetrical applications such as VoIP constitute a major portion of the traffic.

3.1.2. Recommended Vendor Strategy

All three major infrastructure vendors allow a configurable downlink uplink split.

3.2. Header Compression (Based on RoHC)

The data portion of a VoIP packet usually does not exceed 20 bytes in length, and when sent using the UDP/IP/RTP protocol stack, is tagged on with an additional 40 bytes of packet headers. VoIP system capacity can be increased by suitably compressing these headers, and this is what the RoHC algorithm is designed to do. RoHC is able to reduce the header size from 40 bytes to 2-3 bytes, and does so in a manner which enables it to recover reliably from lost packets which are common on a wireless link.

The following considerations arise when applying RoHC to WiMAX systems:

- **Location of the RoHC End Points:** This could be located in the ASN gateway or the BTS. By locating it in the ASN gateway the system can take advantage of RoHC compression over the backhaul link in addition to the airlink. The NWG1.5 has also made this choice.
- **RoHC Signaling:** A technique has to be defined for carrying RoHC signaling from the ASN gateway to the MS, for the initial RoHC negotiations. NWG1.5 has made the following choices: The DSA-REQ/RSP and DSC-REQ/RSP messages are used for ROHC negotiation across R1. In ASN architecture, the Data Path Registration Procedure delivers those parameters over the R6 and R4 reference points by means of the *ROHC parameter payload* TLV.
- **RoHC Feedback Channel:** The RoHC protocol specifies a feedback channel from the de-compressor to the compressor in order to improve the performance and reliability of the compression algorithm. NWG 1.5 has specified that the channel for DL ROHC payload packets shall also carry Feedback-Acks for UL channel that

is identified by 802.16 CID (on R1) and GRE key (on R6). Analogously, the channel for ROHC UL payload shall transport feedback for DL ROHC packets.

- Issues related to classifying RoHC encoded packets at the BTS: Packets that arrived at the BTS after RoHC encoding at the ASN gateway no longer have upper layer protocol fields needed for classification. Discussion on how to deal with this problem are still underway in the IEEE and WiMAX Forum.

Some further details on the RoHC algorithm are provided next: The ROHC algorithm is similar to video compression, in that a base frame and then several difference frames are sent to represent an IP packet flow. This has the advantage of allowing ROHC to survive many packet losses in its highest compression state, as long as the base frames are not lost. A ROHC compressor is in one of 3 main states. In Initialization and Refresh (IR) state, the compressor has just been created or reset, and full packet headers are sent. In First-Order (FO) state, the compressor has detected and stored the static fields (such as IP addresses and port numbers) on both sides of the connection. The compressor is also sending dynamic packet field differences in FO state. Thus, FO state is essentially static and pseudo-dynamic compression. In Second-Order (SO) state, the compressor is suppressing all dynamic fields such as RTP sequence numbers, and sending only a logical sequence number and partial checksum to cause the other side to predictively generate and verify the headers of the next expected packet. In general, FO state compresses all static fields and most dynamic fields. SO state is compressing all dynamic fields predictively using a sequence number and checksum. The size of the sequence number field governs the number of packets that ROHC can lose before the compressor must be reset to continue. The size of the sequence number in 1 and 2 byte ROHC packets is either 4 bits (-1/+14 frame offset), or 6 bits (-1/+62 frame offset), respectively, so ROHC can tolerate at most 62 lost frames with a 1-2 byte header.

3.2.1. Requirements

Sprint has required all its WiMAX infrastructure vendors to support RoHC, as defined by the WiMAX NWG 1.5 specification.

3.2.2. Recommended Vendor Strategy

All three infrastructure vendors have committed to supporting RoHC. Samsung and Motorola will be including RoHC in releases that are due later in 2008, while Nokia will provide RoHC in a release to be delivered in mid 2009.

3.3. Choice of Codec

Voice codecs are chosen to provide the best performance given network constraints, and usually wireline service providers choose G.711 or G.722 codecs, while wireless service providers choose ECRC or AMR codecs. Most of the older codecs are narrowband (NB) which means that signals in the range 100-3500 Hz are coded, while some of the newer codecs are wideband (WB) in which voice signals in the range 50-7000 Hz are coded. WB codecs offer higher fidelity than NB codecs while using bitrates that are comparable to them.

The AMR codec has been developed by Nokia and preferred by 3GPP, while the EVRC codec family (EVRC-A, EVRC-B and EVRC-WB) has been developed by Qualcomm and preferred by 3GPP2. EVRC-A is currently in use in Sprint's 1xRTT network, while EVRC-B has been recommended for use in the EvDO Rev A network.

Sprint recently concluded a performance study of the following 4 codecs: AMR-WB, VMR-WB, EVRC-B and EVRC-WB. The following conclusions were drawn from the study:

- EVRC-B performs better than other NB codecs under frame error and background noise conditions and at a lower average data rate than the other narrowband codecs evaluated.
- All WB codecs have similar performance characteristics, however EVRC-WB seems to operate at a slightly lower average data rate.

Based on these results, EVRC-B is preferred NB codec, while EVRC-WB is the preferred WB codec, for use in Sprint's networks.

Some of the desirable features in a wireless oriented codec are the following:

- The codec should have the ability to dynamically adapt its bit rate to current network conditions.
- Avoid the use of transcoding whenever possible, since it leads to voice quality de-gradation and increases latency.
- Advanced techniques to combat network jitter, such as time warping. This technique temporarily slows down the play of speech samples if packets are arriving slower than they should and speeds up the play if they are arriving faster. This enables the link to combat network jitter, without requiring a large de-jittering buffer, which increases latency.
- The codec must be robust enough to withstand a certain degree of data degradation during transmission, due to lost frames.
- The codec should support capabilities such as discontinuous transmission mode (DTX), which reduce bandwidth consumed by sending low rate "noise information" frames when silence is detected.

3.3.1. Detailed Investigation Findings

- Donnie Townsend, "VoIP Codec Recommendation," Technology Evolution Plan (TEP) document

3.4. Downlink Spatial Multiplexing and Uplink Collaborative MIMO

Downlink Spatial Multiplexing (SM) has the effect of increasing the channel bit rate to specific mobiles which are experiencing good channel conditions. However it is of no use in increasing the voice capacity of the system, since capacity is measured using at most one voice call per mobile. SM does increase the data capacity of the system, hence it is of benefit in a mixed voice + data scenario in increasing total capacity.

Uplink Collaborative MIMO (UC-MIMO) on the other hand, is more effective in increasing VoIP system capacity for the following reasons:

- UC-MIMO enables simultaneous transmissions from up to 2 mobiles to the BTS, which leads to a direct increase in uplink VoIP capacity, even under the constraint of a single VoIP call per mobile
- With the 29:18 downlink/uplink split, the VoIP system capacity is constrained by the uplink, as pointed out in Section 3.1. Hence any technique that can increase uplink capacity, such as UC-MIMO, leads directly to an increase in total VoIP system capacity. With a more equitable downlink/uplink split, the effect of UC-MIMO on system capacity will not be as pronounced, but will be useful in overcoming other uplink rate handicaps, such as the 16QAM uplink modulation restriction.

3.4.1. Requirements

Sprint has requested that its WiMAX infrastructure vendors support UC-MIMO.

3.4.2. Recommended Vendor Strategy

Samsung and Motorola have committed to supporting UC-MIMO, and will make it available in releases coming later in 2008. Nokia has not committed to supporting UC-MIMO at this point.

4. Factors that Influence VoIP Capacity: IEEE 802.16 R2/WiMAX R1.X Based Systems

IEEE 802.16 and the WiMAX Forum are currently working on a follow on to IEEE 802.16e specification, which goes by the name Release 1.X in the Forum and Release 2 in IEEE 802.16. Products compliant to this specification are expected to be available in 2009. Release 1.X is going to include a number of features that will help improve VoIP performance, including the following:

- Persistent Assignments: Improves capacity by reducing the MAP overhead, which is one of the biggest inhibitors to VoIP capacity in a WiMAX Wave 2 network.
- AMC Carrier Permutations: Improves capacity by reducing the PHY overhead for pilot carriers, and also enabling more efficient scheduling strategies.
- Reduction in Handover Latency: Reduces handover latency by eliminating some of the control message exchanges involved during handover.

In Section 6 we discuss simulation results that show that the addition of persistent assignments to WiMAX causes the VoIP capacity estimates to increase to a level that is comparable to the capacity estimates being projected for VoIP over LTE technology.

4.1. Reduction in Downlink Control Packet Overhead using Persistent Assignments

In WiMAX R1.0, every voice packet that is scheduled in the TDD frame requires an Information Element (IE) descriptor. IEs for both downlink and uplink transmissions are present in the downlink portion of the frame, in the DL_MAP and UL_MAP control

messages respectively. These messages are typically transmitted using QPSK $\frac{1}{2}$ modulation, with repetition 4 or 6.

Consider the following: Sub-burst IE size of 35 bits, which corresponds to a IE for HARQ MAP with Chase combining. With QPSK $\frac{1}{2}$ repetition 6, this IE by itself will take up 5 slots. On the other hand the VoIP packet itself will probably not take up more than 2 to 3 slots (assuming 20 bytes for VoIP data, 3 bytes for RoHC compressed IP headers and 6 bytes for MAC headers and 2 bytes for CRC, sent using 16QAM $\frac{3}{4}$ rate modulation). Hence we can see that due to the high MAP overhead, the IE descriptor ends up occupying more space in the frame than the actual VoIP packet itself.

In order to reduce the MAP IE overhead, R1.X is working on a technique called Persistent Assignments, which works as follows: VoIP packets are generated on a periodic basis, usually one packet every 20 ms, when the voice source is in the active state. Persistent Assignment takes advantage of this regularity by assigning a fixed region in the frame for the VoIP flow using a special IE, with the understanding that further IEs for this flow are not needed in the following frames, unless something in the flow changes. A change could be caused by the source becoming idle, changes in the modulation/coding level or a link error leading to HARQ re-transmissions. When this happens, the BTS reverts back to sending regular IEs for the flow. Analysis by Intel has shown that this technique can reduce the DL MAP size from 91 bytes to 77 bytes, and the UL MAP size from 113 bytes to 70 bytes, leading to a 15% increase in the number of users per sector.

Note that the Persistent Assignment technique will be effective in increasing VoIP system capacity only if the VoIP bandwidth bottleneck is due to the downlink. In the initial Sprint deployment, the VoIP bottleneck is due to lack of uplink capacity, hence Persistent Assignments will not help. Once we transition to a more symmetric bandwidth distribution, this technique can be put into use. Vendors who have provided simulations showing the benefits of Persistent Assignments usually assume a downlink/uplink split of 23:24.

4.2. More Efficient Carrier Permutations

The initial Sprint deployment is based on WiMAX Wave 2, which features PUSC carrier permutation in both downlink and uplink directions. The PUSC permutation provides frequency diversity which helps to make the channel robust in situations involving multipath. Release 1.X will enable AMC2X3 carrier permutation in both directions, which comes with the following benefits:

- The pilot carrier overhead in AMC2X3 is much lower as compared to PUSC. For example, UL PUSC has 33% pilot overhead, while uplink AMC only has 11% pilot overhead. This translates into a larger number of sub-channels for UL AMC (48 vs 35 for UL PUSC). The net increase in throughput for UL AMC compared to UL PUSC is about 15%.
- Even though UL AMC cannot take advantage of frequency diversity, it can realize an equivalent benefit by taking advantage of Frequency Selective Scheduling (FSS). The 5 best AMC logical bands, among the 12 available AMC logical bands, are reported by the mobile back to the BTS, every 30 ms. Based on this information, the BTS scheduler assigns the available slots to mobiles whose link is

performing better at that point in time, using schemes such Proportional Fair (PF).

Note that Release 1.X also improves the downlink MIMO algorithm, by using Closed Loop MIMO (CL MIMO). According to simulation results from Nortel, CL MIMO in combination with the AMC permutation, can improve the downlink throughput by 23%, as compared to open loop MIMO on PUSC. However as noted earlier, improvements in downlink throughput due to MIMO schemes will not improve VoIP capacity, since the assumption is that there is at most VoIP flow per mobile.

4.3. Other Factors: Reduction in Handover Latency

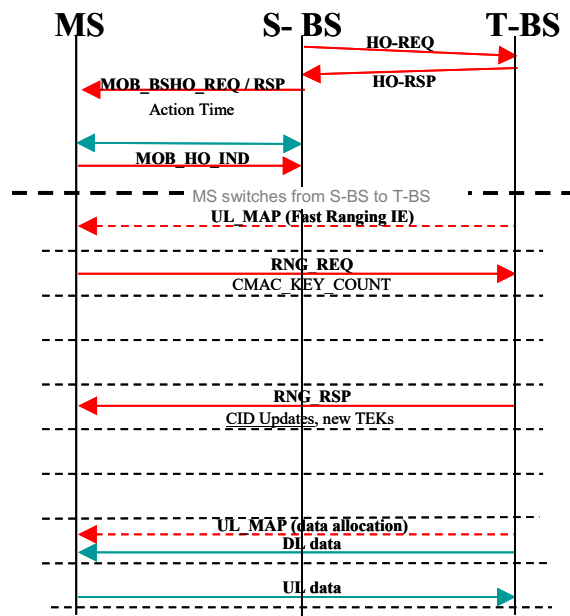


Fig 1a: Fully Optimized HO in R1.0

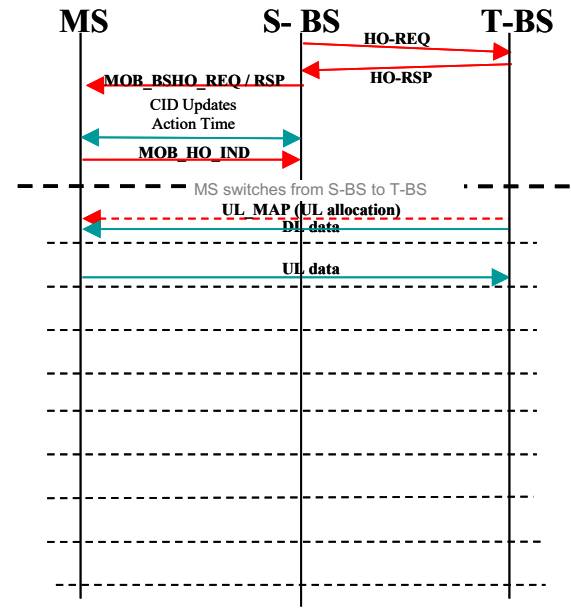


Fig 1b: Fully Optimized HO in R1.X

The controlled HO process is divided into two parts. During the first phase, known as the HO preparation phase, the mobile remains connected and operational on the source BTS. During the second phase, known as the HO execution phase, the mobile disconnects from the source BTS after sending the MOB_HO_IND message, and is able to transmit data again (to the target BTS), after exchanging a RNG_REQ/RSP message with the target, as shown in Fig 1a. In the best case, the mobile remains disconnected for a time period of about 8 frames, or 40 ms. This analysis does not take the PHY layer re-connect delay into account, which will add some additional delay to this process.

Figure 1b shows the HO execution phase in Release 1.X. The main difference between the two is that the RNG_REQ/RSP has been eliminated in Release 1.X. This is done by information about the new CIDs and TEKs (Traffic Encryption Keys) in the target BTS to

the mobile while it is still connected to the source BTS, as shown in Figure 1b. As a result of this optimization, the HO delay is reduced to 5 ms (+ PHY re-connect delay).

5. Factors that Influence VoIP Capacity: IEEE 802.16m Systems

IEEE 802.16 has started work on the next generation version of their specification, which goes by the name 802.16m. Their objective is finish work on the specification by the end of 2009, so that products based on it will start appearing in the market towards the second half of 2010. One of the objectives of the 802.16m specification is to satisfy the performance requirements set forth in the IMT Advanced spec., one of which pertains to the voice capacity of the system. The final form of the 802.16m spec. is not known today, since the technical proposal stage of the process has barely begun, however it is possible to make a few broad observations of the probable nature of the spec. We outline a few features that are likely to make it to the 802.16m spec, and which will have a significant impact on the VoIP capacity and performance of the system:

- **More Efficient Framing Structures:** In particular, IEEE 802.16m adopts the concept of a sub-frame of about 640 μ s in duration (as opposed to the 5 ms duration frame in 802.16). A shorter frame significantly reduces the bearer latency over the airlink, and makes recovery from link errors much faster. Multiple sub-frames are then concatenated together to form a frame, in both downlink and uplink directions. The number of sub-frames in the downlink and uplink portions can be adjusted dynamically, in response to changing traffic patterns.
- **Reduced MAP Overhead:** IEEE 802.16m plans to adopt more efficient MAP IE designs that will significantly reduce the downlink control channel overhead. In particular tree based MAP descriptors are being considered, which imposes a binary tree structure on the description of the available slots in the frame, thus allowing multiple slots to be described by a single field. This also facilitates the inclusion of persistent scheduling schemes of the type that have already been included into Release 1.X.
- **Higher Level Uplink Modulations:** IEEE 802.16m is likely to allow for 64QAM modulation in the uplink, which will remove one of the significant causes of lower uplink data rates.

6. Other Factors

6.1. QoS Support

VoIP over the WiMAX airlink can be supported using either the Unsolicited Grant Services (UGS) service class or the Extended Real Time Polling Services (eRTPS) service class.

- The UGS service class is allocated higher priority than the service classes. In both the downlink and uplink directions, UGS packets are subject to

Peak Rate regulation, based on a configured value. This in combination with Admission Control ensures that they do not hog all the available bandwidth on the link. In the uplink, UGS flows are allocated automatic grants of a fixed size, at fixed intervals. This is done to serve typical voice sources, without requiring the source to send each packet using contention access.

The main issue with the UGS service class is that once the UGS flow is established, the uplink grants are continuously granted, without regard to whether there are any packets to send or not. Hence voice sources that use silence suppression are not able to reduce their bandwidth requirements when using UGS. This characteristic of UGS flows also means that they can only be used with Dynamic Flow support in place. With Static Flow support, the bandwidth allocated to the UGS flow cannot be used by other flows for the duration that the device is active.

- The eRTPS service class was also designed with voice sources in mind, but includes features that make it more suitable for use with a codec that uses the silence suppression feature. In order to do this the eRTPS service class does the following: When the voice source is active, data grants are given at regular intervals just as in the UGS service class. When the source stops sending packets, the BTS detects this event and stops the automatic grants. In order to re-start the grants, the mobile can either send a contention based bandwidth request to the BTS, or utilize one of the CQICH based uplink slots (which is more efficient).

The first phase of the Sprint WiMAX network will use the eRTPS service class for VoIP flows. These flows will be statically provisioned, at the time the mobile does initial registration with the network, and will stay active for the duration that the mobile is in the active or sleep mode. Note that due to the nature of the eRTPS scheduling, an idle or inactive VoIP flow will not consume any resources on the airlink. However since call admission control is done only at initial registration, it leads to the following issue: The bandwidth that is set aside for VoIP calls will be over-provisioned in order to increase call capacity. This can lead to the situation in which if a larger than expected number of users start talking on their device, then bandwidth from other service classes will be infringed upon. In order to prevent the situation described above, the system should support Dynamic Service Flows. These flows are set up at the time the user call is established, and torn down when the call is over. The call set up and tear down are triggered by SIP control messages which in turn trigger QoS control messages using IMS protocols such as PCRF.

In addition to the airlink, QoS for VoIP flows should also be supported over the backhaul link. The Sprint WiMAX network plans to do this using the DiffServ protocol, which provides the Expedited Forwarding (EF) service class for VoIP type flows. Bandwidth reservation over the backhaul link is co-ordinated with BW reservation over the airlink, in order to provide end to end QoS support over the RAN.

During a handover, an active voice call needs to reserve bandwidth over the new airlink and backhaul links, otherwise it results in a call drop. In order to reduce the frequency of this event the network should aside some bandwidth that is reserved for active flows in the handover state. This bandwidth cannot be used by newly created VoIP flows in that

cell. Similar policies have been used in older wireless networks, and have been shown to significantly reduce the incidence of dropped calls.

6.2. E911 Services

Mobile VoIP enabled devices must have an embedded geo-positioning capability and use RF based network location determination to be compliant with Phase II FCC E9-1-1 mandates. The latter feature is required for cases when either GPS is not present in the mobile, or if the mobile is located indoors where it cannot receive the satellite signal. The WiMAX forum is working on satisfying this requirement as part of its NWG1.5 work effort. One of the protocols being considered is OMA SUPL, which provides a general framework by means of which location services can be supported across various access networks.

6.3. ASN GW/HA Issues: Local Breakout Function

The WiMAX networking architecture is based on Mobile IP, which requires that all traffic to a mobile user be routed through the Home Agent (HA). The Core Network architecture allows for 1 or 2 HAs per regional distribution network. It is also a well known fact that the majority of calls are between users that are in close geographical vicinity, hence forcing this traffic to be routed to the nearest HA which may located remotely increases the load on the backbone network. In order to solve this problem, Sprint is investigating techniques that bypass the HA based routing, and directly connect the VoIP traffic between mobiles through their ASN gateway, also known as Local Breakout. This project is already underway as part of the optimizations required for VoIP over EvDO Rev A, and need to be extended for the WiMAX case as well. The following needs to be done:

- ASN Gateways located in a regional data network need to have direct links connecting them, over which local VoIP calls can be routed.
- ASN Gateways need to examine all VoIP traffic originating from the mobiles, and if they determine that the destination mobile is connected directly to it, or to one of the other ASN Gateways to which it has a direct connection, then it re-directs the packet to the appropriate path, instead of routing it to the HA.

6.4. Voice Call Continuity – With EvDo Rev A

6.5. Operational Support Systems (OSS)

6.6. Security

6.7. CALEA

7. VoIP Capacity Comparisons

This section discusses a few VoIP capacity comparisons from the WiMAX vendors including Intel, Nokia, Nortel, and it also addresses results from the NGMN consortium.

There are a number of factors that influence VoIP capacity, and unfortunately the simulation results from the vendors cannot be directly compared, since they make their own set of assumptions about these parameters. The Table below provides a high level summary:

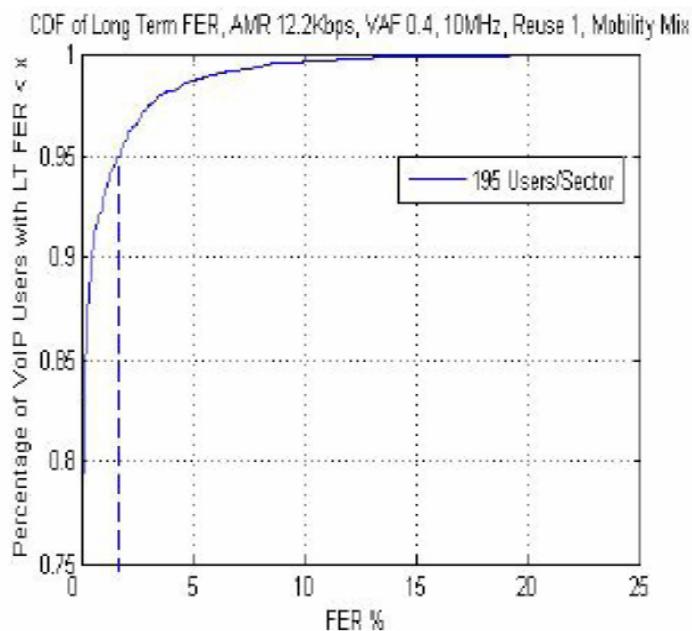
Table 1

Factors that Influence VoIP Capacity	Intel	Nortel	NSN	NGMN WiMAX TDD	NGMN LTE TDD
TDD Ratio	23:24 DL Control: 10	25:22	29:18	1:1	1:1
Header Compression	RoHC		RoHC	RoHC	RoHC
Codec	AMR		EVRC		
Persistent Scheduling	Both results available (see below)	Both results available (see below)	Yes	No	Yes
Adaptive Modulation	Adaptive	Adaptive		Adaptive	Adaptive
MIMO	2X2 DL SM and UL Collaborative MIMO	2X 2 STC	2X2 STC	2X2 DL SM and UL Collaborative MIMO	
Frequency Reuse	1		3	1	1
Channel Model	Mixture of Ped B and Veh A	Ped B		Ped B	Ped B

Capacity per 10 MHz Channel	200 w/o Persistent Scheduling 230 with Persistent Scheduling	140 w/o Persistent Scheduling 200 with Persistent Scheduling	100	140	237
--	---	---	-----	-----	-----

Nokia's *conservative* simulations for WiMAX 2008 systems as submitted for MSA predict 100 users/sector. Nokia's simulations take into account the deployment characteristics of the Xohm network (frequency re-use of 3, 29:18 TDD split, MIMO, EVRC and RoHC. Without ROHC, the capacity is expected to be 70 users/10 MHz carrier. Note that the 29:18 split ratio is the constraining factor, if we apply this ratio to the NGMN numbers, then we get a capacity of 107 users/sector, which is close to the Nokia number.

The Intel study assumes availability of ROHC, 12.2 kbps AMR codec, eRTPS QoS, Collaborative MIMO, silence suppression, 1:1 UL/DL ratio, and considers a system outage < 3%.



As can be seen from Figure 1, the VoIP capacity of 195 users/10 MHz channel is available with FER target of 95th percentile below 3%.

The NGMN Technology Evaluation group did a performance comparison between WiMAX and LTE. The following vendors provided simulation results: Nortel, ZTE, Qualcomm, Huawei, Motorola,

Nokia/NSN, Samsung, China Mobile, CATT, Intel and Alcatel Lucent, and the results are shown below.

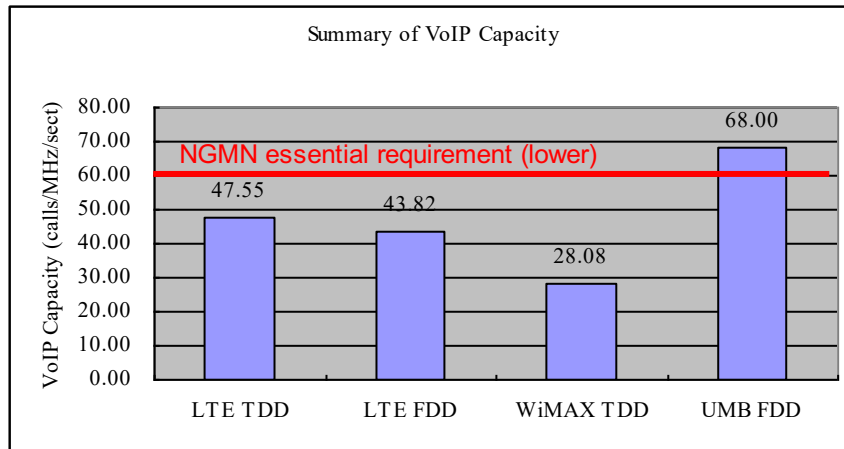


Figure 3: Data averaged over all vendor submissions to NGMN by technology.

As Figure 3 shows, the 2008 WiMAX system VoIP capacity for a 10 MHz channel with a 1:1 TDD split is expected to be around 140 users. (This assumes UL Collaborative MIMO), whereas LTE numbers are on average around 220 for FDD and 236 for TDD. The LTE assumptions included use of AMC whereas WiMAX assumed PUSC.

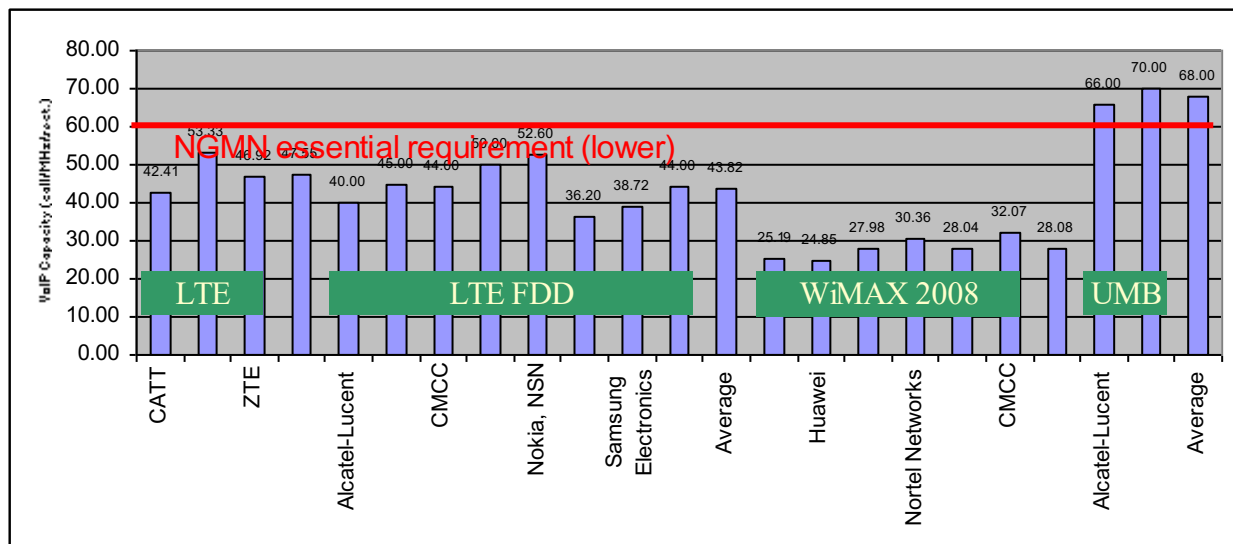


Figure 4: Individual vendor results submitted to NGMN

As 802.16e evolves to 1.x and 802.16m, the system capabilities are enhanced in several areas that impact VoIP performance such as control channel enhancement, lower pilot overhead, smaller MAC headers and frame sizes, and persistent MAP allocation. Furthermore, antenna techniques combining beam-forming and MIMO as well as interference cancellation capabilities are expected to increase VoIP capacity in WiMAX. Nortel expects the 1.x release to support 200 users and the 802.16m release to support 300 users. The persistent scheduling feature expected in 1.x has been shown to boost the VoIP capacity by 43% in Nortel simulations. Lucent provides estimates around 257 users/10 MHz carrier with similar enhancements. In comparison, forecasted LTE numbers are in the 180-260 range per NGMN results.

7.1. Comparison Between WiMAX and LTE

The difference in VoIP capacity between the Wave 2 version of WiMAX (140 users) and the TDD version of LTE (237 users), can be explained as due to the following factors:

- LTE has lower MAC control channel overhead due to the use of Persistent Scheduling for VoIP. This feature is planned as part of WiMAX Release 1.X.
- LTE uses an AMC type permutation, which has lower pilot channel overhead as compared to the PUSC permutation used by WiMAX Wave 2. The AMC permutation also enables Frequency Selective Scheduling, which boosts throughput in cases of low user mobility. AMC is supported in WiMAX Wave 2, but using its use precludes the use of MIMO. This issue is being addressed as part of Release 1.X, which will allow MIMO in the AMC zone.
- LTE also has other differences, such as the use of Incremental Redundancy rather than Chase Combining and finer granularity in modulation and coding control, but it is not clear whether these features have a significant effect on VoIP capacity.

It is expected that with WiMAX Release 1.X the capacity difference between WiMAX and LTE will narrow considerably, as evidenced by the Intel and Nortel simulation results shown in Table 1.