RTT TECHNOLOGY TOPIC
August 2005

WiFi Voice and Video

A number of vendors and operators (including recently Skype, Kineto and Boingo) are actively promoting the provision of IP voice over public access (Hot Spot) wireless local area networks. In parallel, IP voice is being propositioned as an add on to existing WiFi data networks for corporate and consumer applications.

In this month's Hot Topic, we discuss some of the capacity, coverage and quality issues implicit in WiFi voice and (longer-term) **IP voice and video** service provision.

The WiFi PHY and MAC

Our August 2004 Hot Topic (WiFi PHY) reviewed WiFi PHY(physical layer) and WiFi MAC(medium access control)functionality in some detail but just as a reminder, an 802.11 b and g access point will typically be configured in the 2.4 GHz ISM band to support 3 non overlapping 20 MHz radio channels with centre frequencies spaced either 30 MHz apart (Europe, Channels 1,7 and 13) or 25 MHz apart (channels 1, 6 and 11 in the US).

Access points using the 5 GHz ISM band (802.11 a) will typically be configured to support up to 8 channels spaced 20 MHz apart (Europe Band 1 and US Band 1 and 2), with additional channels available in Band 2 in Europe and Band 3 in the US .Raw data rates of 54 M/bits/second are achievable in strong C/I conditions using high level modulation and an OFDM multiplex to improve channel resilience.

Multiple channels capable of supporting multiple users co sharing a common 54 Mbps looks like a lot of bandwidth but in practice there are significant MAC overheads and link budget constraints that result in substantially lower net throughput rates.

These **MAC overheads increase when periodic two-way time bounded services** (ie real time voice and/or real time voice and video) **need to be supported along side best effort services.**

Voice and video frame lengths and arrival rates

Voice frames are typically 92 bytes long and arrive every 20 milliseconds (the frame rate is determined by the syllabic rate). Video frames are 1464 bytes long and arrive every 40 milliseconds (assuming a 25 frame per second video rate). A 92 byte voice packet arriving every 20 milliseconds implies a voice data rate of 36.8 k/bits/second (92x8x50). A 1464 byte video packet arriving every 40 milliseconds implies a video data rate of 292.5 k/bits/second.

Table 1 Voice frame lengths, video frame lengths and arrival rates

Voice frames	92 bytes	Every 20 milliseconds
Video frames	1464 bytes	Every 40 milliseconds (25 fps)
Data frames	1500 bytes	
Fast data frames	3000 bytes	

A combined voice and video call would have a combined data rate of 329 k/bits per second. This is however the rate to support unidirectional voice and video. Although people do not (generally) speak at the same time, the MAC layer has to provision bi-directional periodic bandwidth (known as voice and video transmission opportunities) so the bandwidth occupancy is effectively doubled to 73.6 k/bits/second to support a bi-directional voice call, 585 k/bits per second to support two way video and 658 k/bits per second to support two way voice and video. This suggests a capacity of 13 voice channels per m/bit, 1.7 video channels per m/bit or 1.5 voice and video channels per m/bit. (2 way video calls).

Data throughput - distance and MAC overheads

Data throughput is dependent on the modulation used and channel coding. In 802.11 a and g, between 48 and 54 Mbps of gross data rate is available if 64 QAM is used in a lightly coded (3/4) channel but this is dependant on having a strong C/I (carrier to interference ratio). As the C/I worsens, the gross data rate reduces to 23/36 Mbps (16 QAM), then 12-18 Mbps (QPSK), then 6-9 Mbps (BPSK) and the channel coding overhead increases from 3 /4 to 1 /2 (one error protection bit for each data bit).

Table 2 Data rates, modulation and coding in 802.11 a and g

Data rate	6/9 mbps	12/18 mbps	24/36 mbps	48/54 mbps
Modulation	BPSK	QPSK	16 QAM	64 QAM
Coding	1/2 or 2/3 or 3/4			

Although there are more channels available in 802.11 a, the propagation loss is higher and the net throughput therefore falls faster as a function of distance (though this also means that access points can be positioned closer together so channel reuse can be a bit more aggressive). In 802.11g, a Request to Send (RTS) and Clear To Send (CTS) message is needed if bandwidth is co shared with an 802.11 b transmitter (known as working in mixed mode). This produces a significant decrease in real throughput. The effect of distance in 802.11 a and the impact of RTS/CTS overhead in 802.11g when used with 802.11 b is clearly shown in the table below. Note how **real throughput rates in 802.11g and 802.11a** quickly fall to levels that are not much higher and **sometimes lower than standard 802.11b.**

Table 3 Effect of distance on 802.11 a, b and g throughput, effect of mixed mode b and g signalling overhead on 802.11g throughput.

Distance(ft)	802.11b mbps	802.11a mbps	802.11 g only	802.11 g mixed mode
10	5.8	24.7	24.7	11.8

50	5.8	19.8	24.7	11.8
100	5.8	12.4	19.8	10.6
150	5.8	4.9	12.4	8.0
200	3.7	0	4.9	4.1
250	1.6	0	1.6	1.6
300	0.9	0	0.9	0.9

RTS/CTS is a poll and response algorithm and therefore implies a scheduling delay in addition to introducing a significant protocol and time/bandwidth overhead.

Note that '**Mixed mode**, actually implies two sets of MAC overheads.

Firstly the way the contention MAC is managed in 802.11b is different from 802.11a and g. Both use time slot back off but 802.11 b uses a 20 microsecond slot width and a and g use 9 microseconds. If 11b devices are inter operating with 11g devices then the 20 microsecond slot length must be used. This means that **contention overheads will be higher.**

Similarly with 11b devices, there is a choice of a long 192 microsecond and/or short 96 microsecond preamble. The OFDM preamble is 20 microseconds. In mixed mode, either the long or short 11b preamble will need to be used to support 11b devices. This means that **preamble overheads will be higher.**

Secondly. mixed mode now also implies that the MAC will be **simultaneaously loaded with time bounded (periodic)and best effort traffic**. This will have **a significant impact on throughput and capacity.**

Taking these MAC overheads into account, table 4 (below) shows typical throughputs for TCP/IP best effort data and/or UDP throughput. The table is from an Atheros White Paper and includes their proprietary bonded(40 MHz) channel solution giving a max throughput of 108 Mbps.

Note that time bounded services would normally used UDP (Unacknowledged Datagram Protocol) rather than TCP (with transmission retries).

Table 4 MAC overheads when supporting TCP or UDP.

		Max TCP	Max UDP
802.11 b	11 mbps	5.9	7.1
g plus b	54 mbps	14.4	19.5
g only	54 mbps	24.4	30.5
a	54 mbps	24.4	30.5
a turbo	108 mbps	42.9	54.8

SIP (Session Initiation Protocol) places an additional bandwidth overhead on the UDP throughput of approximately 8 kilobytes every time a voice, video or voice and video session starts or is modified.

Mobility, handover and power control overheads.
The above also excludes any measurement and signalling overheads introduced by the need to support mobility. As long as a user stays in one place then these overheads can be avoided. This may be /probably is the case for Skype laptop users but may not/probably will not be the case for people expecting to use WiFi voice from their mobile phone either at a public access Hot Spot or on a corporate, SOHO or home wireless LAN.

Mobility overheads include the need to do measurement reporting (802.11k), the use of measurement reporting to manage handover algorithms (802.11f) and/or the use of measurement reporting to manage per packet power control (802.11h).

Most transceivers can now collect RSSI (Received Signal Strength Indication) at a per packet level. Given that the channel is reciprocal (same RF channel on the uplink and downlink) it is easier to do measurement reporting with WiFi than it is with cellular (which uses a different radio channel on the uplink and downlink each with different propagation properties).

However, its all very well collecting this information but then you have to decide what to do with it. There has not been much point up to now in doing power control with best effort data. If the RX level is good then you just send the data faster using higher order modulation and/or with reduced channel coding.

Voice, video, and/or voice and video combined are however different in that they occupy periodic bandwidth with typically a longer (more persistent) session length. If the user is close to the base station then it is worth reducing power to a level at which packet error rates can be kept above the coding threshold of the channel encoder and the error threshold of the source decoder. Reducing the TX power helps reduce battery drain on the uplink but also tends to improve receive sensitivity so the downlink usually benefits as well but care needs to be taken to make sure the power control doesn't take up more bandwidth and/or power than it saves.

Table 5 Power control dynamic range in 802.11

Power in dBm	**Power in milliwatts/microwatts**
20	100 milliwatts
17	50 milliwatts
15	30 milliwatts
13	20 milliwatts
7	5 milliwatts
0	1 milliwatt
-10	100 microwatts

The range of RSSI measurement in a 802.11 transceiver is typically 60 dB. The range of power control is typically either 20 or 30 dBm. This is less than you find in a wide area radio interface (35 dB for EDGE, 80 dB for 1XEV or Rel 99 UMTS) but still potentially useful. The usefulness depends (as with handover algorithms) on how mobile the user is likely to be.

Table 6 Mobility thresholds

Transmit power in intervals of 2 dBm	Distance (m)	Difference (m)
0-2	7	
2-4	9	2
4-6	10	1
6-8	21	11
8-10	26	5
10-12	36	10
12-14	46	10
14-16	70	24
16-18	90	20

http://www.cs.colorado.edu/department/publications/reports/docs/CU-CS-934-02.pdf

The above is taken from a University research project on mobility thresholds. It assumes two users a certain distance apart with one of the users walking away from the other user at 1.5 m (metres) per second. The closer the distance between the two transceivers the faster the rate of change in terms of required output power. For example, at a distance of 9 m, walking 2 m away (just over a second of elapsed time) results in a 2dB power change. At a distance of 70 m, a distance of 20 m has to be covered before a 2dB step change occurs (13 seconds of elapsed time). It is therefore plausible that a per packet power control algorithm could be deployed that could be reasonably stable when used in this type of 'gentle mobility' application and that could yield some worthwhile power savings and related link budget benefits. From a PA design perspective however, it does mean the operating point of the amplifier will be constantly changing and this in itself has an impact on error vector magnitude and harmonics.

The ability to power control handsets and access points also provides an opportunity to optimise the radio system in terms of channel reuse and coverage. This will become increasingly important if voice users become a significant percentage of the offered traffic mix.

Impact of receive sensitivity on the link budget

The link budget is a composite of the TX power, TX accuracy (typical error vector magnitude for an 802.11g transmitter should be <1.5% and <2% for 802.11a but is often much worse), path loss and sensitivity. Sensitivity is a function of data rate. With 802.11 a and g higher data rates are achieved by using higher order modulation, every time the modulation state is doubled (eg from BPSK to QPSK), another 3 dB is needed on the link budget, moving from 16 QAM to 64 QAM implies a 6 dB increase. In practice, the fact that on the TX side, EVM tends to get worse with higher order modulation means that the real implementation losses are higher. The table below shows some typical receive sensitivity figures (and some claimed sensitivity figures) at different data rates.

Table 7 Typical 802.11 a and g and b receive sensitivity

Data rate	54 Mbps (a and g)	11 Mbps (b)	5.5 Mbps (b)	2 Mbps (b)	1 Mbps (b)
Typical sensitivity (dBm)	-75 dBm	-85 dBm	-88 dBm	-89 dBm	-92 dBm
Claimed max sensitivity					-101 dBm

Linearity requirements
Table 8 compares three generations of cellular transceiver with a WiFi (802.11 a and g) transceiver in terms of peak to average ratio (PAR), peak to mean ratio, whether the radio channels are full or half duplex and the power control dynamic range.

Table 8 Linearity comparisons between cellular and WiFi (OFDM)

Generation	System	PAR(dB)	PMR(dB)	Duplex	Power control(dB)
1G	AMPS	0	0	Full	25 dB
	ETACS	0	0	Full	25 dB
	J TACS	0	0	Full	25 dB
2G	GSM	0	0	Half	30 dB
	PDC	3-5	>10	Half	30 dB
	US TDMA	3-5	>10	Half	30dB
3G	EDGE	>3	>10	Half/Full	35 dB
	1XEV	>5	>10	Full	80 dB
	HSDPA	>5	>10	Full	80 dB (Rel99)
WiFi	OFDM	17 dB	>20dB	Half	30 dB

The use of OFDM in 802.11 a and g delivers some significant benefits in terms of channel resilience and ISI performance (a constant and relatively low symbol rate) but the cost is a substantial envelope variation on the composite modulated waveform (although the example of 20 dB peak to mean is a worst case condition with all 52 sub carriers lining up over a symbol period). This requires additional linearity from the PA which is difficult to realise in a power efficient manner.

In contrast, Bluetooth 2.0 EDR (which uses GFSK, 4 phase DQPSK or optionally 8 phase DPSK) is arguably more power efficient.

GSM, PDC, TDMA and EDGE are described as half duplex in that they don't transmit and receive at the same time (except for EDGE Class 13 through 18) but they still have an RF duplex separation between transmit and receive which translates directly into an improved sensitivity figure.

WiFi is half duplex in that it uses the same RF channel which is time division duplexed to separate the uplink and downlink. This means the sensitivity will always be less than a full duplexed cellular system using separate RF channels for the uplink and downlink.

This matters because sensitivity is part of the link budget and the link budget determines coverage (range) and capacity. On this basis it could be argued that WiFi is not particularly spectrally efficient. The additional linearity needed also means it is not particularly power efficient when compared to cellular or Bluetooth.

WiFi's spectral efficiency and power efficiency limitations
So the WiFi PHY is arguably less spectrally efficient and less power efficient than cellular, and probably less spectrally efficient and certainly less power efficient than Bluetooth.

The WiFi contention optimised MAC when used for connection oriented time bounded traffic is arguably less efficient than existing connection optimised MACS used in cellular and Bluetooth voice applications.

So why use WiFi for IP voice and video?
IP voice, IP video and IP voice and video are all potentially supportable on WiFi radio systems but require careful implementation in terms of PHY management (channel reuse) and MAC management (the co sharing of common bandwidth between time bounded voice and best effort data).

Whether WiFi is efficient or not when compared to other options is to an extent irrelevant if costs are sufficiently low to drive adoption though an inefficient PHY and MAC will always have a cost in terms of additional battery drain.

The addition of **OFDM increases processing overhead in the receive chain** (the cost of the receiver FFT) and processing overhead in the TX chain (the inverse FFT).

The **additional linearity** implied by the envelope of the composite waveform also **reduces TX power efficiency** when compared to other radio systems.

However, OFDM is really the only way to realise data rates in the region of tens of Mbps (direct sequence spread spectrum starts to run into jitter problems at these higher speeds and you need the OFDM multiplex to slow the symbol rate down in order to control ISI and increase multi path resilience).

In the longer term, WiFi with MIMO (multiple input multiple output) is one way of getting speeds in the order of 100 Mbps or more (the other way is using UWB). In other words, if you want bit rates of tens of Mbps then WiFi is probably quite an acceptable option and at least provides a plausible route to delivering 100 Mbps over the next 3 to 5 years, albeit over a very short distance.

Summary- the engineering story
So this is one of those circular arguments. The WiFi PHY and MAC were never designed to support a significant mix of time bounded services. It is reasonable to assume that IP voice and in the longer term IP video and IP voice and video will become a progressively more important part of the offered traffic mix (and by implication a more important part of offered traffic value). This implies that handsets and access points will need to support higher data rates.

Higher data rates are achieved by implementing mixed mode 802.11 b and g, which implies additional contention overhead. The connection based nature of voice and voice and video calls also adds contention overhead and signalling bandwidth load.

The fact that users might expect to walk around when using WiFi IP voice implies the need to manage mobility which implies the need to introduce network assisted handover which implies the need to implement RSSI or packet error measurements. If you are doing RSSI or packet error measurement you may as well implement per packet power control which will help to improve capacity and coverage (by lowering the overall noise floor and by improving sensitivity in the receiver). This in turn will help reduce some of the power budget issues. The higher data rates are needed partly because time bounded services absorb bandwidth but also because time bounded services are more expensive to deliver in terms of PHY and MAC utilisation.

Summary - the business story

Finally this all has to make sense to the network operator. Probably the most significant shift here will be the inclusion of network operator specific browsers in the next generation of SuperSIM smart cards. Rather like existing application layer WiFi browser products like Cirond, these will identify preferred networks which of course means preferred networks from a network operator perspective. At this point, WiFi becomes a profit opportunity not a threat.

Whether we like it or not, WiFi is here to stay and will become an integral part of most future cellular phones supporting simultaneous wide area (cellular) local area (WiFi) and personal area (Bluetooth) PHY and MAC functionality.

WiFi voice and in the longer term WiFi voice and video will become an important part of the application mix and will be complementary to wide area (cellular) and personal area (Bluetooth) voice and video platforms.

Whether WiFi will scale successfully (through MIMO implementation) to support shared channel data rates of 100 Mbps or above is arguably more open to question. From an engineering perspective, **mono pulse or multi carrier OFDM both offer substantially better power efficiency at these higher data rates (the bandwidth gain effect of the ultra wideband channel).**

What is for sure is that the co existence and combination of these multiple radio systems both in the handset and access network will remain as one of the dominant design challenges over the next three to five years.

About RTT Technology Topics

RTT Technology Topics reflect areas of research that we are presently working on.

We aim to introduce new terminology and new ideas to clarify present and future technology and business issues.

Do pass these Technology Topics on to your colleagues, encourage them to join our

Push List and respond with comments.

Contact RTT

RTT, the Shosteck Group and The Mobile World are presently working on a number of research and forecasting projects in the cellular, two way radio, satellite and broadcasting industry.

If you would like more information on this work then please contact

geoff@rttonline.com

00 44 208 744 3163