RTT TECHNOLOGY TOPIC
November 2004

3 G Cost and Quality

In this month's Hot Topic, we review some of the cost and quality issues implicit in present wide area wireless IP network deployments, looking specifically at the 3GPP IP RAN and IP core network and the cost/performance trade offs inherent in the Release 6 IP multimedia sub system (IP MMS) specification.

The Promise - cost reduction, more services, better quality

The IP Multi Media Sub system (IP MMS) is part of the 3GPP Release 6 standards process. IP MMS promises (sooner or later) the seamless delivery of IP voice, IP video, IP audio, IP images, IP text and IP data - a composite mix of multimedia including real time conversational services delivered over an IP core network via an IP RAN (radio access network). The implication is that these services will be delivered at equivalent or better quality than existing services and at lower cost.

Cost savings are to be realised through a combination of lower hardware costs including radically lower prices for IP RAN hardware and the use of commoditised processor hardware to route traffic across the core network. Better quality voice together with simultaneous video, audio, images, text and data will create new 'quality based' added value opportunities. More added value and lower costs will result in higher profits.

Well possibly but let's qualify some of the thinking behind the 'profit promise'.

Cost saving complications

Cost items first. It is simplistic but still defensible to argue that processor and memory bandwidth will continue to halve in price every 18 months or so. Given that the UMTS physical layer moves most of the signal processing to baseband then these cost savings should be realisable over the life span of a 3G IP RAN. Similarly high capital cost hardware switches are being replaced with low capital cost routers.

Lower hardware costs in the IP RAN and core network do not however directly relate to overall costs in the network. Consider for example, software component count, software complexity and software cost. A first generation analogue cellular phone had 10,000 lines of code, a GSM phone 100,000 lines of code and a UMTS phone between one million and 1.5 million lines of code. In a router, a simple kernel implementation of a TCP/IP protocol stack might take 'only' 15,000 lines of code but on top of this you have to put policy engines that can manage multiple per user packet streams each with their own particular quality of service metrics with the ability to manage security metrics and access policy rights. Both in the handset and the network, hardware engineering costs are replaced with software engineering costs.

Defining voice quality- the PSTN benchmark

Then there is the issue of quality.

Defining quality of service in a PSTN circuit switched voice call is straightforward, a combination of blocked call rates and voice quality. Voice quality can be defined using mean opinion scoring. In an MOS scored system, (the ITU PESQ-LQ standard), a score of 5 is excellent, 4 is good, 3 is fair, 2 is poor and 1 is bad.

The UK PSTN has a Mean Opinion Score of 4.3 and almost no blocking.

Mobile networks have to add in dropped call rates as a key performance metric and voice quality is (as you would expect) far more variable- typically a GSM network will have an operating range of between 2.9 and 4.1.

PSTN networks have a delay of 35 milliseconds or so and no delay variability. Source coding, channel coding and interleaving in mobile networks adds 50 to 60 milliseconds to the delay budget but if the mobile call is circuit switched, there will be little or no delay variability.

There is therefore a well-established existing benchmark for voice quality (voice QoS) in mobile networks. The QoS may of course vary through a call and will certainly vary from handset to handset - small form factor handsets tend to have lower MOS scores partly due to their physical acoustic qualities, partly due to coupling effects between the hand and the handset and partly due to the performance limitations of small internal antennas.

Mobile voice QoS is therefore dependent on handset performance, the codec used, the quality of the radio link at any given moment and network effects. Network effects include the impact of multiple transcoding and, in IP networks, the impact of packet loss, packet delay (end to end jitter) and packet misinsertion. Additional header overheads in IP voice also have an impact on the radio link budget.

The problem of packet loss, jitter and misinsertion can be mostly overcome by setting up virtual paths and circuits through the network, by using traffic shaping protocols and avoiding buffering. This is what is meant by ' conversational class' service in the 3GPP specification. Conversational class implies dedicated network bandwidth that in practice means the bandwidth cost is equally expensive irrespective of the transport medium. Similarly, the quality impact of the radio layer can be moderated by using adaptive source coding and channel coding and more aggressive power control although all three of these mechanisms have a cost in terms of occupied bandwidth and power. The simple message is that voice quality comes at a cost - the better the quality, the more it costs to deliver.

Defining video quality

The same principle applies to real time video. It is not just that video needs more radio and network bandwidth than voice, it needs better quality bandwidth. Quality in this context implies lower error rates at the radio layer, evenly distributed errors at the radio layer, low packet loss rates through the network and evenly distributed loss rates. Video quality is also dependent on the quality of the original images (colour depth, resolution and frame rate), the compression used, the error concealment techniques deployed in the encoder/decoder and the resolution, colour depth, contrast ratio and refresh rates of the display used in the receiver. Quality therefore has a cost not only in terms of the quantity and quality of the radio and network bandwidth needed but also the processor bandwidth (and delay) needed in the handset encoder/ decoder. More processor bandwidth in the handset also implies the need for more battery bandwidth. Note also video quality is impaired both by error rate and the distribution of errors in time. A service level agreement that specified a certain minimum bit error rate or minimum packet loss rate through the network but would not relate accurately to actual video quality. Burst errors on the radio channel and bursty packet loss through the network will destroy video quality even if the overall error rate and packet loss rates are low.

Defining image quality

Non real time images are easier to send and therefore cost less to send than real time video but there is still a bandwidth cost/image quality trade off. Image quality can be defined in terms of the 'Q' of the image, a metric commonly used in digital cameras. An image taken in fine camera mode will have a Q of 90 and will produce a file size of 172,820 bytes which would take just over 40 seconds to send on a 33 k/bit uncoded channel. An image with a Q of 5 would produce a 12 kilobyte file size which would take less than 3 seconds to send on the same channel but the quality would be extremely poor.

Defining audio quality

The quality /bandwidth trade off also applies to audio. Note that audio may also need to be real time if being sent with real time video. Note also how user expectations of audio quality are changing. A low cost I Pod has a digital/solid state microphone capable of capturing an audio signal up to 20 KHz and headphones with a frequency response from 20 Hz to 20 KHz. Gone are the days when 7 KHz of audio bandwidth was considered adequate or acceptable for consumer applications.

Defining text quality

Even text is going to become more expensive to send - the 3G TXT work group within 3GPP is defining the protocol and network requirements needed to support simultaneous texting - the ability to add real time sub titles to a conversational real time video stream. This means the text data stream must be sent on an isochronous end to end channel with no buffering - real time radio and network bandwidth is expensive bandwidth. Text quality can also be defined in terms of actual visual quality- the ability to italicise, kern and embolden which in turn requires an ability to do text rendering in the handset - a processor bandwidth issue.

Defining data and application quality

And finally data. Data is the one 'product' in the media mix where you might expect to get away with buffering in the end-to-end channel - the 'best effort' delivery proposition. Possibly so, but most IT managers expect to have service level agreements that describe typical end to end delays, packet loss rates in UDP and/or retry rates in TCP/IP. The administration of a data service level agreement or service level guarantee is fraught with complication- who measures the statistics, what happens when the network operator fails to deliver against a service level guarantee. Service level agreements and service level guarantees are a direct consequence of moving from circuit switched to packet-routed networks. They are a source of friction between network operators and end users and introduce hidden customer relationship and customer support costs that have a direct impact on AMPC (average margin per customer) metrics. Additionally, the data proposition may well involve access to server or storage bandwidth. An application Service Level Agreement will typically specify maximum transaction response times. Server performance may or may not be under the control of the network operator but becomes an additional cost and risk in an application service level guarantee. If an application service level guarantee is breached there may be economic harm implications.

The cost of quality

So voice, video, image, audio, text and data all have an associated cost which is quality dependent - a 24 bit colour depth 20 frame a second real time 'conversational' video with simultaneous 24 k/bit AMR coded voice and text subtitling will cost more to send than a 12 bit colour depth 15 frame a second video with a 4.75 k/bit encoded voice stream, which will cost more to send than an interactive or streamed session, which will cost more to send than a best effort data file. It is not just the quantity of bandwidth needed but also the quality of the bandwidth that adds to the cost. Video is more sensitive to error rates and error distribution on the radio channel and to packet loss and packet misinsertion in the network. Multimedia bandwidth is expensive bandwidth.

The value of quality

Given that we know it will cost more to support, for example a 'rich media' multimedia conversational exchange, the question has to be will a user pay more for the experience and will the tariff premium cover the additional delivery cost. The answer is probably no given that we expect voice quality as a right rather than a privilege, we expect video and audio quality as a right rather than a privilege, we expect to have our images delivered without a major perceivable loss in quality, text to be displayed in a comfortable readable format and data files to be delivered without a loss of integrity.

However, what we can say with reasonable certainty that increasing the application bandwidth of the user experience will increase billable bandwidth. Application bandwidth is the composite bandwidth made up of each of the simultaneous traffic streams- voice, video, audio, image, text and data. The more components supported in a session, the longer the session. Provided network operators continue to bill by time then billable bandwidth will increase.

We can also say with reasonable certainty that increasing the quality bandwidth of the user experience will increase billable bandwidth. Quality in this context is quality as perceived by the user and is a composite of voice quality, video quality, audio fidelity, image quality, text quality and data integrity. As the quality and consistency of the user experience improves, session lengths will increase. Provided network operators continue to bill by time then billable bandwidth will increase.

The increase in billable bandwidth should (deep breath needed her) be sufficient to offset the additional cost of delivery. The higher the loading on the network, the more likely it will be that hardware and software costs can be amortised and a real return be achieved on radio bandwidth and network bandwidth investment.

Storage bandwidth value as a future source of profit generation

Which brings us on to storage bandwidth value as a profit opportunity. Given that imaging bandwidth in camera phones is typically now at 1.3 mega pixels and that device sampling is presently either 2 or 3 mega pixel, it is safe to assume that there will be a substantial increase in imaging bandwidth in future handsets. Improvements in processor efficiency and battery capacity will also increase offered traffic in the uplink.

This suggests that storage bandwidth may become a major future profit generator with differentiated QOS provided on the basis of access rights, security and server bandwidth available for redelivery and redistribution.

Storage bandwidth is enjoying the same cost reduction curve as processor bandwidth though additionally as processor bandwidth increases, our ability to compress multimedia files also improves. The capital value of archived media also increases over time.

The cost and margin pressures implicit in 3G network roll outs suggest a fundamental rethink may be needed in the way that network value is defined and delivered - storage bandwidth rather than delivery bandwidth added value.

In a world of cautious accounting, storage bandwidth should be treated as an appreciating asset, delivery bandwidth as a necessary, but costly, liability.

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