RTT TECHNOLOGY TOPIC
July 2005

Convergence Costs and Margins

In last month's Hot Topic, (Broadcast Over Cellular) we reviewed some of the practical design challenges involved in developing cellular handsets with integrated digital broadcast receivers. We pointed out that spectral value was intrinsically dependent on delivering robust radio systems with adequate link budgets that could ensure a consistently good quality user experience.

Broadcast over cellular is apart of a bigger shift towards delivering converged mobile and fixed wireline and wireless service platforms that can handle real time and streamed multi media content.

Developing business plans that scale across wireless/wireline/copper/optical access however depends on being able to develop cost models that deliver accurate 'cost of delivery' metrics that can be used to validate pricing policy. This in turn requires an understanding of some of the engineering issues involved - the subject of this month's Hot Topic.

Telecom Cost Complexity
Telecom costing has always been a complex process involving amortisation of fixed facilities with a life span of 50 years or more (think telegraph poles and traditional telephone exchanges). Universal service obligations and access guarantees have always implied a substantial over provisioning of access and network bandwidth.

Popular myths - packet is cheaper
The transition from circuit switched to packet routed networks has been promoted as a mechanism for improving network and transmission bandwidth multiplexing efficiency which in turn is supposed to deliver a decrease in the cost of delivery (measured in euroherz or dollarherz or eurobytes or dollarbytes of network transmission bandwidth).

Multiplexing efficiency is dependent on buffering. Buffering has two cost components - the cost of the memory and processor bandwidth needed to store, manage and prioritise buffered packet streams and the (largely hidden) cost of supporting differentiated classes of service.

The end to end to end differential quality of service (QoS) needed to support multimedia content over a multi service network produces a cost and efficiency metric which is intuitively dependent on the offered traffic mix - conversational traffic costs more to deliver than best effort because it cannot be buffered.

This is of course only true if the multiplexing efficiency of buffering exceeds the cost of over provisioning transmission bandwidth beyond the point at which buffering would no longer be required.

If a significant percentage of offered traffic is real time/ conversational, the assumption of a multiplexing gain from buffering is invalid.

Additional network efficiency is supposedly achievable by deploying flexible end to end routing. Again, any efficiency gain achievable is directly dependent on the offered traffic mix. If a significant percentage of offered traffic is real time/conversational, the assumption of a multiplexing gain from routing flexibility is invalid. The amount of bandwidth used increases as the number of hops increase and/or the routing trajectory becomes progressively more indirect.

IP protocol and signalling overheads

At protocol level, the overheads of SS7 have been or are being replaced by the 'lighter signalling load' of the TCP/IP or UDP protocol stack working with traffic shaping and IP prioritisation protocols (MPLS, Diffserv, RSVP etc). Any efficiency gain achievable is directly dependent on the offered traffic mix. If a significant percentage of offered traffic is real time/conversational, the assumption of an efficiency gain from IP signalling and traffic shaping is invalid.

If average packet lengths are 1500 bytes or so, IPV4 or IPV6 address overheads are relatively trivial. However, if a significant percentage of traffic is multi media and time sensitive or time interdependent, IP packet traffic (IP voice, IP audio, IP video) will be mapped on to fixed and short (typically 92 bytes or less) fixed length packets and IP address overhead will be substantial. In radio systems, this can be equivalent to taking between 4 and 5 dB of link budget out of the radio system. Given that one dB of link budget (rather approximately) equates to a 10% increase in network density then it is clear that IP addressing and signalling overhead adds directly to the dollardensity costs of deploying a wireless network. Similarly, a multimedia multiplex will require fixed length packets in order to preserve the time domain properties of the multiplex. The flexibility of variable length packets (one of the claimed advantages of IP when compared to ATM) therefore no longer applies.

Connectionless is better than connection oriented?

The always on connectivity implicit in a packet network delivers efficiency benefits when handling small bursts of data. In a GSM network, the 1 to 2 seconds of set up delay in a circuit switched call (or 3 to 4 seconds in a Release 99 network) represents a major overhead when handling short bursts of data. However, these packet efficiency benefits are only valid with a short session duration. Most present evidence suggests that in a multi media network, sessions are substantially longer than traditional telecom voice calls and are increasing over time. This suggests that any supposed efficiency benefits achievable from having a connectionless end to end channel are becoming increasingly invalid.

This also affects the overall efficiency of call set up protocols such as SIP. A SIP message generates about 8 kilobytes of signalling load. A change in session property (a new user or additional content stream) requires a new SIP message exchange. As sessions become more complex over time, SIP becomes progressively less efficient and requires substantially higher signalling bandwidth than existing SS7 signalling.

The hidden costs of IP MMS
An IP MMS (IP Multi Media Subsystem) enabled multi service network is rather like a postal or courier service but instead of two classes of service (first class or second class for postal services, priority or standard for couriers) there are 4 levels of service. Note that the postal service in the UK loses money on second class mail and makes money on first class mail. This is partly because first class mail has a higher value but also because second class mail costs more to deliver. The additional costs are incurred because the storage costs (and related storage administration) now exceed the benefits of holding back mail to fully fill increasingly inexpensive trucks (rather bizarrely the real costs of running a large delivery van up and down the motorway are going down rather than up over time). With international carriers, the same tipping points apply and as aviation costs have reduced (larger, cheaper more efficient aeroplanes) it becomes cheaper not to store and forward but just to forward. The same principle applies to indirect routing. Although the cost of indirect routing may decrease over time, the process still consumes additional bandwidth. This is also true of wireless mesh networks, which can be remarkably spectrally inefficient due to routing inconsistencies and associated user signalling overheads.

So is the internet an efficient transport system?
Not really. It's robust because that's what it's designed for. It was however never designed to handle multi media traffic and if you expect it to deliver the same end to end channel performance in terms of latency and jitter as an ATM or circuit switched network, then any supposed efficiency benefits will rapidly disappear. All Internet protocols are inherently inefficient and either waste bandwidth through transmission retries (TCP/IP) or lost packets (UDP) or a mixture of both.

So why is the Internet so cheap and why does it work so well?
Because it is presently grossly over provisioned and has been financed by the pension funds that lost investors money in the dot com dot bust cycle.

So if IP networks are neither more efficient nor lower cost than circuit switched or ATM networks, why are telecoms operators deploying them?
Well that's a good question and one of the answers is that to an extent hardware costs are reducing as hardware (and to a lesser extent software) in the network becomes a commodity but this is in reality a modest incremental process. Complex ATM hardware is being replaced with complex soft switch platforms with high clock speed processors, extravagantly provisioned high performance buffer memory cards and unstable traffic shaping protocols. Not really a big leap forward.

As the Internet becomes more aggressively loaded over time, the real costs of delivery will start to reappear and will need to be factored in to end to end delivery cost calculations.

The real drivers- transmission system and storage system efficiency
The second (more convincing) answer is that there are technology drivers that have nothing to do with the circuit switched to packet switched transition, or with the Internet but everything to do with the efficiency of the underlying transmission systems and persistent storage capabilities which will jointly drive next generation telecom business models.

Lets look at **transmission systems** first.

	Copper	Electrons
Silicon	Optical	Photons
	Radio	Electrons

Device Bandwidth Effects
In last month's Hot Topic we talked about device scaling and the reduction in geometry from 95 to70 to 65 to 45 to 32 nm devices. Decreasing device geometry directly delivers a bandwidth gain both in terms of volume and value.

The image processing chain in a camera phone (see March and April Hot Topics on camera phone and camcorder functionality) is a present example. Note it's not just electron flow that's increasing. As CMOS sensors get smaller we can capture more and more photons. Imaging volume and value increases every time that we move to a new device generation. As resolution increases, new applications emerge. Bar code scanning and face recognition and/or shape recognition are present examples.

The same bandwidth principle applies to voice and audio bandwidth. Each device generation provides new opportunities to improve voice quality or voice value. As audio and voice resolution increases, new applications emerge. Voice and speech recognition are present examples.

We can then use the additional clock cycles available to us to compress source coded content without loss of perceptual quality. This generates bandwidth gain and value gain - an 'apparent bandwidth'/'perceptual bandwidth'/'illusional bandwidth' effect.

We can then use the additional clock cycles available to us to channel code the content to make it more resilient to the propagation characteristics of the (copper, optical or radio) transmission medium.

We can then use the additional clock cycles available to us to improve the efficiency of the transmission medium either by using higher level modulation and/or OFDM techniques to improve throughput.

These techniques are common to radio, optical and copper.

Optical transport efficiency
In the optical domain, the ability to create 'narrow band' 25 GHz channels at L Band (1560 to 1620 nm) and S band (1280 to 1350 nm) has created a 'new bandwidth' effect which directly reduces per byte long distance transmission costs. Note that it is important to differentiate real costs and actual costs. The actual cost of optical bandwidth is presently artificially low due to over supply/over provisioning - the excess of dark fibre financed by pension funds during the dot com dot bust cycle. At some stage, demand balances out with supply. Real bandwidth costs reassert themselves and need to be factored in to end to end delivery cost budgets. At that stage, real costs rather than actual costs become important as do technologies that

substantially reduce those costs.

Copper transport efficiency
With copper, similar principles apply. Copper and optical are both more consistent in terms of their behaviour than radio. An ADSL modem is specified to deliver error rates of 1 in ten to the ten. In radio, error rates of one in ten to the three are considered to be good and error bursts (the result of a fading channel) are inherently problematic and wasteful in terms of power. All three media are of course power limited but all three have untapped 'bearer potential' in terms of their ability to support additional bandwidth and an ever-decreasing cost curve.

It is this effect that will drive future telecom margin and value. In this respect, telecoms is not significantly different to the oil industry in that automotive technology (more efficient combustion) increases the 'bandwidth' (miles per gallon) available irrespective of fluctuations in supply and demand delivering a net real gain in terms of industry value.

Storage system efficiency
And last but not least, the same device bandwidth effect can be seen to be changing the persistent storage value proposition. It is important here to differentiate buffer storage in a queued network (which we have argued increases cost without a clear-cut cost or revenue benefit) and persistent network based or network accessed storage which is where long term storage of customer/subscriber content can be used to increase archive value (and related customer dependency metrics).

As device geometry decreases (and device functionality increases) our ability to compress and store and manage and retrieve and redeliver information increases and our cost per megabyte decreases. Similarly, handset resident persistent memory facilitates a complete new sub set of 'apparent bandwidth'/ 'perceptual bandwidth'/'illusional bandwidth' applications in which the user experience is determined not by the amount of network bandwidth available but by the memory and processor clock cycles available in the handset itself. Multi player games are a present example.

The (new) telecoms technology value model
The future of telecoms will be based on a technology driven device driven decrease in delivery and storage costs and a technology driven device driven increase in delivery and storage revenues. These cost reductions and revenue gains are real rather than actual and as such deliver a real (long term) rather than illusory (short-term) industry gain.

The loading overheads of the Internet (the inefficiencies of the Internet as a transport system) will bring forward the point at which additional investment will be needed or justifiable on economic grounds to improve present optical, copper and radio bandwidth efficiency. The same principle applies to persistent storage.

This presents a wide range of new opportunities for vendors with strong technology portfolios that can directly address this need.

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