

Presence and collaboration in IP communications and the impact of new standards

This white paper is part of a series on Voice over IP and Voice over Wireless produced by Samsung. This document seeks to explore the development of new IP based communications applications and the evolution of the standards that support them. Presence and collaboration are key enablers of the Virtual Enterprise and have a direct impact on the take up of VoIP and Voice over wireless.

This paper has three roles:

1. To discuss new applications and the changes in behaviour they create
2. To discuss some of the technical background to how these technologies are developing
3. To provide into developing standards

This paper does not seek to be the final word from a technical point of view. Much useful information and details of the standardisation process are available free of charge (in most cases) from the web sites of the key standards bodies: in this case, ISO, the IETF and the ITU. To simplify the text a glossary of key terms is provided at the end.

Presence and Identity – the foundations for collaboration

Presence

Presence is not merely a technical capability – it actually changes the way things are done. Understanding the impact of changes in behaviour explains why some of these technological developments take time to mature. It's not the technology that is difficult; it's the cultural and behavioural change they require or enable that takes time.

In a presence-enabled context communication (whether it is voice or data-centric) is only initiated when the various parties are available. This represents a fundamental change in the way calls are made and is similar to how Instant Messaging (IM) applications are used. Voice systems have traditionally shown the current call state of an extension on the system. What was not provided was information about whether the person was actually at their desk even when the phone was idle. IM applications provide presence information based on keyboard strokes or mouse movement. This information can be linked to the user's calendar and geographic location to deliver more meaningful information: "In a meeting", or "Don't call he's in Japan where it is currently 3am". Combining this functionality with the telephone is a natural next step and the SIMPLE protocol (see later section and glossary for more detail) is currently being worked through the IETF to support both instant messaging and presence. This in turn changes the model for how telephony services are charged. The prevailing model is likely to involve personal subscriptions (i.e. a per user charge) rather than a line or extension charge, plus a call charge for calls that go off net. This is similar to how mobile phones are charged which in turn reinforces a "personal handset" model where the users most commonly used numbers are stored as part of the subscription or

on the handset itself. With Voice over Wireless handsets the form factor and styling is likely to echo mobile and users will carry them around the office or when they go out and about. Dual function phones with both Voice over Wireless and true 2G/3G capability will be available in identical form factors and styling options to the pure mobile handsets they replace.

Presence further develops the idea, borrowed from IM, that communication is enabled with people only when they are available. This comes with its own set of pro's and con's. Firstly users have to be online a good percentage of the time for it to be valuable (or in the case of our children, at predictable times) but the key advantage is time isn't wasted communicating speculatively. Every call is productive.

Presence has two main benefits. It makes communications richer by providing a sensible basis for choosing the appropriate medium with which to communicate, and it makes communications more productive by reducing wasted calls and ensuring the two parties are in the best possible state to have a worthwhile exchange. It has the secondary benefit of making complex technology such as conferencing more usable and makes the experience more pleasurable because it is more appropriate.

A good example of this can be seen looking at Skype. There are several aspects to the Skype service that change the way we use telephones. Firstly, like with many mobile calls, you don't dial a number. You look at whether your contact is online and free and then open a call in almost exactly the same way as you initiate an Instant Messenger session. In fact the Skype interface looks most like an IM client and even supports chat capabilities. This type of multi function capability pinpoints a trend towards a single "window" for controlling all of your communications systems whether they are email, voice, chat, video or fax. Some years ago there was a trend towards personal numbering and Unified Messaging. This may well now have arrived, if via a torturous route.

Identity

As organisational boundaries change, become transparent or disappear, the ability to assess whether another party is someone one wishes to initiate or even permit communication with becomes even more important. There will always be organisations and individuals who want to restrict their visibility and these tools will enable this as much as they will improve visibility. It is key that the individual has control over these parameters for these tools to become accepted. With identity comes security to protect organisations and individuals from illicit or damaging behaviour.

Although this may seem to reduce the opportunities for making new contacts, and clearly has a benefit in screening nuisance calls, it does also allow the opportunity for people to broadcast when they are open to contact. As these applications get more sophisticated and the networking interface improves people will put more effort into developing their online identity and further improving communication.

The ability to establish identity is not without its challenges, and this topic is discussed in greater detail in the section on SIMPLE later in this document.

Once an identity is established, a range of applications that enable us build relationships become feasible and this is examined in the section on networking.

If these two aspects are put together they create the drivers for Voice over Wireless. Flexibility and changing working practices drives a need for the sort of flexibility and dynamics that wireless networks deliver. People's changing use of telephones drives the need not just for VoIP but also for networking and collaboration applications that drive greater value.

New Applications

Collaboration

Once presence and identity management exist in the network it becomes possible to foster collaboration. Collaboration applications come in various flavours:

- Communications-based collaboration
- Knowledge-based collaboration
- Workflow collaboration (shared documents and processes) at a task or project level.

One simple way to understand the impact of collaboration is to compare the capabilities and the activities we have in a face to face meeting and think of their online analogues. In a meeting we can:

- Schedule the meeting, invite attendees and set an agenda
- Identify and meet people
- Build relationships
- Hold a conversation
- Discuss things
- Make a presentation
- Share data and information
- Work on things together using white boards and notes
- Make agreements
- Distribute actions
- Schedule follow up meetings
- Have someone make notes and prepare a record of the meeting so that agreements and actions are followed up

Online collaboration (and networking) seeks to create a digital equivalent for any and all of these activities so that people can work together even when they are not face-to-face.

In a face-to-face meeting we don't normally even think of what we are doing from a communications perspective. Communications based collaboration includes the ability to make and receive calls whether audio or video, the ability to set up conferences of various types, to manage email, texting, and IM. These lead to web logs, chat rooms and various other types of discussion forum, covering the full range of electronic communications both asynchronous (non real time) and synchronous.

Knowledge-based collaboration concerns itself with the ability to identify and search, and covers directory and storage capabilities. One of the key capabilities that result from this is the ability to network. Looking at the meeting analogue for collaboration, there are a number of activities in a meeting, which have been absent from the online equivalent, which have tended to focus on communication and workflow aspects. Meeting people and building relationships have intangible but no less worthwhile benefits of face-to-face meetings. The section below on networking explores this in more detail.

Workflow collaboration covers a number of disciplines:

- Meeting management and scheduling
- Document sharing
- Resource allocation
- Task and Project management
- Process management
- Reporting and measurement

The ability to collaborate at these levels creates a business accelerator effect especially for small businesses, as they have traditionally been the domain expertise of the larger organisation.

Networking

Online networking has become one of the “killer” applications for Broadband over the last couple of years. As insurance against the isolation that can be a feature of remote and homeworking, networking has been the final piece in the jigsaw that has enabled many people to strike out on their own and set up their own online micro-business. The independent consultant/contractor has naturally migrated to this model and other start-ups are starting to see the benefits in both social and business contexts. Apart from friendship, and distraction these networks can be a powerful marketing tool for the self-employed, helping them access networks and contacts they would not normally have access to. Apart from some initial scepticism as to the value of some of these relationships and a certain resistance to the “dating” aspects of such meetings, when blended with face-to-face meetings, online networking has a key role to play. It will never entirely replace meeting people in the flesh (and the more sophisticated services don’t pretend that they can) but as part of the collaboration tool kit it has an important role to play.

Knowledge-based collaboration creates the lead into networking, particularly if users can search by certain parameters that match their requirements. Presence and identity are key enablers. In fact this kind of online networking is becoming an application in its own right with the explosion of usage on networks such as Ecademy, Ryze and LinkedIn. These are business-oriented sites, even though the process can seem a bit like a lonely-hearts club. In fact business has always been done via networking – originally via church, the Masons or other business clubs such as Round Table – it’s just that now it is done online. As a way for finding customers, suppliers or partners its use is growing especially amongst small businesses looking to expand the scope of their coverage. So identity becomes a critical way of marketing both at an individual and at an organisational level. Business has often been more about who you know than what you know, and the growth of online networking proves this.

As people invest more time in online networking issues of trust and value inevitably arise. Identity becomes more than a trivial search parameter. Online identities will gradually become more sophisticated. Initially they look very like a directory in the traditional sense with minimal information. Over time they become more like a personal web page or micro site with a range of preferences and interests. Hosting these directories becomes a key service for service providers.

Conferencing

Conferencing, whether audio or video (and especially internationally) has been another application classically limited to larger organisations. As organisational boundaries break down and collaboration tools really deliver on the promise of the virtual organisation it is very likely that these tools will be in common use in the smaller business. The ability to manage down the cost of long distance calls and set up dynamic conferencing on the basis of availability will deliver benefits to the first time and the experienced user as traditional conferencing can be awkward and time consuming to coordinate. This in turn migrates into a “chat” methodology where calls are enabled rather than made in the formal sense. Video conferencing has suffered in the past from the limitations of the underlying technology, especially the bandwidth limitations and cost of ISDN based services. Moving to IP networking opens up the market but identity and presence are also important. The final hurdle has been quality.

Follow me (Unified Messaging)

As the number of phone numbers users “own” has increased the ability to find the right person without making multiple calls has become a challenge. With the development of Intelligent Networking (IN), driven in large part by the growth of call centres, calls can be routed across a telephone network based on a number of commands or policies. Most people’s experience of this starts with the Intelligent Voice Response (IVR) systems commonly in use in Call Centres. Most people have become used to having calls queued and being given choices to make as to how they are handled. The inverse of this is having a system that finds the user. Early versions of this required individual schedules and locations to be set up. IP technology and presence means that calls can be routed dynamically to wherever the user shows up on the network, but can also be overridden if they do not wish to receive the call. This level of control creates a virtuous loop with presence management. The user can now look to see, not only if the other party is available but also whether it is a good time to call at all.

The evolution of standards in IP communications

The importance of SIP

Session Initiation Protocol is a standard being developed specifically for multi media communications by the IETF. Previous attempts have tended to focus on one media type (such as H.323 for voice/telephony). SIP is a peer-to-peer protocol. This means it is ideal for supporting communications which are initiated by a user, at the edge of the network compared with other protocols which often rely on a centralised control methodology for “call” set up. It is also a relatively simple protocol and yet has sufficient sophistication to enable it to cover most application types including multicasting and basic Automatic Call Distribution (ACD). Importantly for

collaboration it supports different message types within the invitation structure enabling users to negotiate the appropriate application environment, and so is particularly useful where a user may have variable access to different capabilities depending on location and terminal equipment. It supports mobility by using a proxy to forward requests to a user current location. Because it is an application layer protocol it is designed to operate independently of the underlying infrastructure and is ideal for masking complexity.

SIP addresses (URL) can be embedded in Web applications making it easier to integrate, and names can be mapped or redirected. It can also be gatewayed to other protocols for interoperability. Some basic principles were addressed with SIP, built on the experience of more complex protocols such as H.323:

1. Where possible services would operate end-to-end
2. Extensions would be allowed where they were generic rather than specific to a particular session type
3. Simplicity, compatibility with other IP protocols and interoperability were key

Although it is an important protocol for VoIP implementation it is much broader than other voice only protocols making it increasingly the de facto standard for collaboration application development.

A technical overview of SIP

SIP, as its name suggests, is a protocol for setting up and managing sessions across an IP infrastructure. Each session can be between multiple end points (called user agents) and may contain different media. Rather than just expecting a static set of users it specifically includes the control necessary to locate users who are on the move or may have multiple identities. SIP works in conjunction with other protocols which are targeted at the transport of media and characterises the session so that media and transport can be defined. It also enables the setting up of proxy servers to manage these exchanges

SIP defines five aspects of establishing and terminating multimedia communication sessions:

- User location: identification of the end systems
- User availability: determination of the willingness of the called party to engage in communications;
- User capabilities: determination of the media and media parameters to be used;
- Session setup: "ringing", establishment of session parameters at both called and calling party;
- Session management: including transfer and termination of sessions, modifying session parameters, and invoking services.

Because SIP is a protocol focused on call management it supports a number of protocols which actually transport the data. The key ones are: Real-time Transport Protocol (RTP) for transporting real-time data and providing QoS feedback, Real-Time Streaming Protocol (RTSP), Media Gateway Control Protocol (MEGACO) for

controlling gateways to the Public Switched Telephone Network (PSTN), and Session Description Protocol (SDP) for describing multimedia sessions.

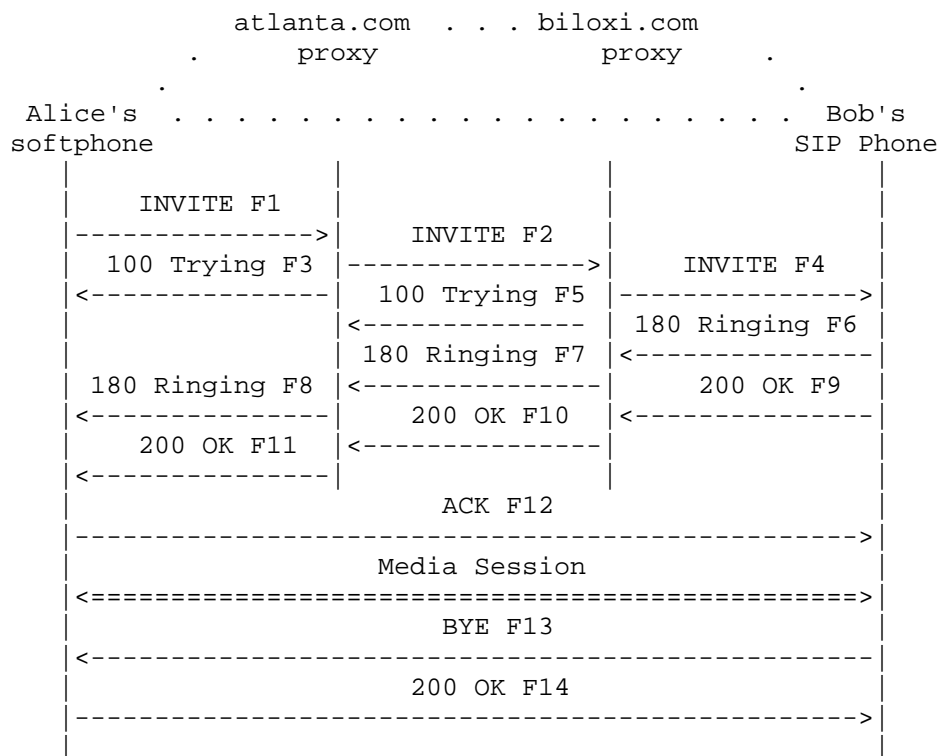
It should be understood that SIP does not depend on any of these protocols nor does it, itself, provide services. However, the types of services which SIP supports tend to be sensitive to the user and so SIP provides a suite of security services, which include denial-of-service prevention, authentication (both user to user and proxy to user), integrity protection, and encryption and privacy services.

SIP works with both IPv4 and IPv6.

An example of SIP call set up

The diagram below is a simple example of a message exchange between two users, one using a softphone on a PC and the other a SIP handset. There are two proxy servers facilitating call set up

Party A (Alice) calls party B (Bob) using his SIP Uniform Resource Identifier (URI), which is analogous to an email address. The first part of this exchange is an INVITE request followed by a number of header fields with additional information about the request:



Messaging exchange

```

INVITE sip:bob@biloxi.com SIP/2.0
Via: SIP/2.0/UDP pc33.atlanta.com;branch=z9hG4bK776asdhd
Max-Forwards: 70
To: Bob <sip:bob@biloxi.com>
From: Alice <sip:alice@atlanta.com>;tag=1928301774
Call-ID: a84b4c76e66710@pc33.atlanta.com
  
```

CSeq: 314159 INVITE
Contact: <sip:alice@pc33.atlanta.com>
Content-Type: application/sdp
Content-Length: 142

SIP uses the UTF-8 character set but the request and response message and header field syntax is identical to HTTP/1.1. Key to these sequences is the Content-Type header field which defines the format of the content within the message.

How a user agent works

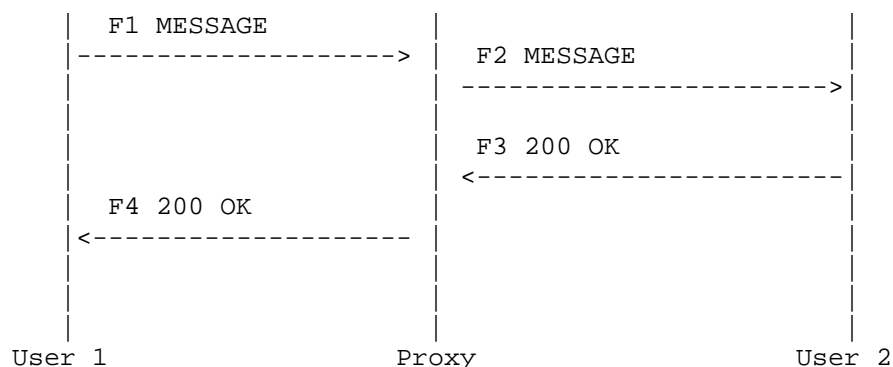
A user agent contains a user agent client (UAC), which generates requests, and a user agent server (UAS), which responds to them. A UAC is capable of generating a request based on some external stimulus (the user clicking a button, or a signal on a PSTN line) and processing a response. A UAS is capable of receiving a request and generating a response based on user input, external stimulus, the result of a program execution, or some other mechanism. When a UAC sends a request, the request passes through some number of proxy servers, which forward the request towards the UAS. When the UAS generates a response, the response is forwarded towards the UAC.

Before a session can start to exchange data a dialog is entered into between User Agents to establish the relationship between the agent and controls the sequencing and routing. A session is initiated with an INVITE message as seen above, although a session can be modified on the fly.

SIP extensions for Instant Messaging and Presence

As has already been discussed, SIP is a session oriented protocol whereas Instant Messaging is a near real-time exchange of messages which are independent of each other. The original standards for Instant Messaging were all proprietary so it was necessary to incorporate them within SIP by an extension method called MESSAGE. With this method, the body of the message is included with the request and can be plain text or message/CPIM (Common Presence and Instant Messaging) which is the format created to ensure compatibility with different IM systems.

A simple IM MESSAGE flow might look like this:



Message F1 looks like this:

```
MESSAGE sip:user2@domain.com SIP/2.0  
Via: SIP/2.0/TCP user1pc.domain.com;branch=z9hG4bK776sgdkse
```

Max-Forwards: 70
From: sip:user1@domain.com;tag=49583
To: sip:user2@domain.com
Call-ID: asd88asd77a@1.2.3.4
CSeq: 1 MESSAGE
Content-Type: text/plain
Content-Length: 18

User1 forwards this message to the server for domain.com. The proxy receives this request, and recognizes that it is the server for domain.com. It looks up user2 in its database (built up through registrations), and finds a binding from sip:user2@domain.com to sip:user2@user2pc.domain.com. It forwards the request to user2. The resulting message, F2, looks like this:

MESSAGE sip:user2@domain.com SIP/2.0
Via: SIP/2.0/TCP proxy.domain.com;branch=z9hG4bK123dsghds
Via: SIP/2.0/TCP user1pc.domain.com;branch=z9hG4bK776sgdkse;
received=1.2.3.4
Max-Forwards: 69
From: sip:user1@domain.com;tag=49394
To: sip:user2@domain.com
Call-ID: asd88asd77a@1.2.3.4
CSeq: 1 MESSAGE
Content-Type: text/plain
Content-Length: 18

Further work on these standards is now incorporated in RFC 2779 and is managed by the SIMPLE working group as detailed below. This enables the content of Instant Messaging to support other media types and allows the more complex states implied by Presence capability to be reflected in Instant Messaging applications.

The role of SIMPLE in Presence and Collaboration

SIMPLE stands for “SIP for Instant Messaging and Presence Leveraging Extensions”. It seeks to apply SIP to new services such as Instant Messaging and Presence. What makes these important, in turn, is that they support the interactive and real-time aspects of multi media (unlike email which is asynchronous) and act as key foundations for the range of applications, which are becoming referred to as collaboration.

Perhaps the most important aspect of presence is that it is not just a technology but signals a completely different paradigm for communications and whole new set of user behaviour. As an example, traditional telephony requires one party to make a call with no knowledge of the called party’s ability to answer. That party is represented by a number, which doesn’t really refer to them but to a line or handset. Mobile goes some way to towards personalising this experience and most calls are made by identifying an individual in a directory, rather than calling a number. One example of new behaviour is using text as a cheap way of checking the preparedness of the other party for receiving a call before the call is made.

Presence conveys not just the ability but also the willingness or desire of a user to communicate across a range of devices. Core to this definition is a new term “presentity” which combines devices, services and personal information to create a full picture of a user’s presence status on the network. Within this definition a user may have multiple aliases. Part of presentity is the URI, which is not just a name but can also indicate status. The personal information includes not just static information but also dynamic information about the current location and status of the individual. “Devices” defines the devices available at the time and services are those to which the individual has access at that time.

One of the most interesting new terms used in the SIMPLE documentation is that of “watcher” rather than the more familiar “caller”. Sat in front of a presence enabled application the user “sees” the status of his target not just in terms of his availability to communicate but also which his preferred communications method might be, which itself is affected by the context in which he finds himself. So the target may be “offline” from email, but online via a PDA for Instant Messaging, a mobile call or even a text. Video might work but it might be variable in quality. The “watcher” can then decide which is the most productive way of communicating which might well be “try later”. The ability to reduce the wasting of time and resources and reduce frustration levels promised by this new paradigm is truly revolutionary.

One of the drawbacks of traditional communications which presence and collaboration seeks to overcome is a direct result of this mismatch of availability. As we get busier and the incidence of failed calls increases, voice mail has become ubiquitous creating a bizarre ritual which we can call the “voicemail dance” where two parties repeatedly call each other and, despite a desire to talk, keep failing and leaving each other messages. A similar problem has beset video conferencing. Conferences require fairly firm scheduling, which requires all parties to be in fixed locations at certain times. When people are late or fail to make the meeting problems occur in re-synching communications. Presence enables you to establish communications dynamically, based on availability and gives you a number of methods to signal your intentions so that the message gets through based on the recipients preferred method of receiving it. Thus if a potential conference participant is “in a meeting” but that meeting is casual or of low priority he may elect to join a conference which he perceives to have a higher priority or interest level.

MIME

As part of the evolution of multi media communications new standards are being set in all areas, even email. The MIME standards (RFC 2045 inter alia) seek to improve interoperability of sophisticated email services by defining standards for message headers and content. The development of the MIME standards within the IETF is closely tied to the development of SIP and SIMPLE creating a strong relationship between text and other communications media.

In MIME message headers enable different clients to establish not only which version of the standards is being implemented, but also what is in the content. The content falls into five main categories:

- Text. Old versions of mail standards only supported US-ASCII excluding the use of complex character sets such as those employed by Asian languages and sophisticated formatting
- Image. With a default standard of JPEG. This standard enables mail clients to recognise and open documents with images rather than calling other applications
- Audio. A basic standard definition enables native interpretation of simple audio with more complex formats support by extensions in the header
- Video. Video often includes synchronised audio and the default standard is MPEG
- Applications. This covers everything from binary data and Postscript to sophisticated application executables

In the case of an inability to identify a particular type of content the default format is “application” and the user then specifies which application to run against the content.

As part of this development, new standards of security for messaging are being developed (S/MIME) which not only cover encryption standards but also how messaging fits into the development of PKI standards (Public Key Infrastructure).

Standards for VoIP

Voice over IP has been “in development” for nearly a decade and three main standards have evolved, driven by different sets of requirements and perceptions of what is important.

H.323

H.323 is a mature standard administered and developed by the ITU, which is a body which seeks to organise the telecommunications industry. H.323 has become widely used for “trunking” traffic between different telecommunications operators although it will support audio, video data across any IP network. H.323 is an extremely sophisticated suite of protocols and its precisely this complexity which has kept the cost of equipment high and opened the door for other simpler protocols

MGCP.

Media Gateway Control Protocol is a proprietary protocol designed to simplify the use of H.323 and to reduce the cost of handsets and other CPE. It is frequently used where the main control of the network is centralised such as in a Centrex system. The relationship between gateways and Call Agents is deliberately quite rigid and formal.

SIP.

Session Initiation Protocol. A simplified protocol for peer-to-peer multimedia communications easily integrated with other applications. Supported by industry giants such as Microsoft and increasingly embedded in a wide range of standard office systems. SIP is discussed in more detail in the previous section

As costs of handsets fall and the range of supported applications increases it is likely that SIP will become the dominant protocol in the public eye and that H.323 and MGCP will be relegated to specialist roles

H.263 and standards for Video

The H.263 standard was published by the ITU (International Telecommunications Union) in January 2005 and covers “Video coding for low bit rate communication”. Designed to address some of the former failings of video transmission over IP networks, what makes the new work stand out is the level of sophistication in terms of quality and services supported. This shows up in terms of the ability to support multiple channels and sub channels and the ability to improve picture quality through enhanced picture prediction. As a result we are likely to see video conferencing take its place as an important and valuable tool in the IP communications kitbag.

Conclusion

For new technology to take off sometimes requires not just a range of supporting applications and services to be available but also for a change in behaviour or culture to become widespread. This is the “sun, moon and stars” theory. What it means is that the mere existence of a new technology or product may not be enough. A good example is VoIP. It has been around for about ten years, and there is an argument that there is too much technology rather than not enough. Large corporations in private networks have used it for years, as a way of avoiding sending calls over public tariffed networks. It has also taken off in new buildings where the cost of installing a single cabling system makes the “risks” worthwhile. Mobile telephone companies have used it extensively for trunking traffic because it compresses predictably. There is an argument that pretty much any telephone conversation now passes over VoIP for at least part of its journey. But it has not taken off in the public mind as a natural alternative to proprietary, PBX based communications in ordinary businesses. This may be about to change for the following reasons:

1. People have got a handle on the issues of mixing voice and data traffic on a single network
2. The cost of handsets makes them more attractive at a single user or consumer level
3. A range of applications that add real value make the change worth considering
4. Communications are becoming so pervasive and complex we need a better way of handling them

When we look back on the point at which VoIP and Voice over Wireless took off it will be some combination of the above which will be the trigger. Whether a change in behaviour and culture will be the most important trigger or merely the most important consequence of this change is yet to be seen, but the sun, moon and stars are finally in alignment.

Glossary of terms/abbreviations

Centrex

Centrex stands for “central office exchange service”. It is a system for centralising switching in a telephone network, which reduces the cost of customer equipment, and removes the need for a switch (or PBX) on the customer’s premises. More popular in

the US than in Europe, it is being replaced by IP Centrex, which overcomes many of the inflexibilities which have given Centrex a bad name with some customers

ITU

The International Telecommunications Union is part of the United Nations and based in Geneva. It undertakes standards setting and coordination of global telecommunications networks

IETF

The Internet Engineering Task Force is an open community of interested parties, which monitors and develops standards for the evolution of the Internet

ISO

The International Organisation for Standards takes its name not from an abbreviation but from a Greek word. It is an association of the national standards bodies of a number of countries

JPEG

The Joint Photographic Experts Group set a standard for compression of digital images. Although this enables compression and therefore smaller file sizes some quality is lost in the process. An ISO standard

MIME

Multipurpose Internet Mail Extensions are a set of standards developed by the IETF to support multimedia content within mail. Original mail standards only defined plain text in US-ASCII. S/MIME looks at standards for mail security

MPEG

The Moving Picture Experts Group is another standards group within ISO, focuses on video compression and file formats, and seeks to improve on proprietary standards. The current version is MPEG 4

PKI

Public Key Infrastructure is a system for managing digital certificates, which help establish identity and trust

SIMPLE

SIP for Instant Messaging and Presence Leveraging Extensions is an IETF initiative to apply SIP to Instant Messaging and Presence services.

SIP

Session Initiation Protocol is a standard of the IETF. SIP is a text based protocol for initiating interactive communications sessions between users, including voice, video, chat, games and virtual reality.

URI

Uniform Resource Identifier is a generic term to cover all types of names and addresses used on the Internet