ASPs: Breaking the Network Barrier

Solving the network connection problem
between the ASP and the SME customer

1 Statement of the Problem

Two of the biggest problems faced by Small and Medium Enterprise customers (SMEs) when considering a migration from in-house to ASP-provided services are:

- The cost of obtaining a data connection to the ASP of sufficient bandwidth and reliability to support business-critical applications

- Concerns over the security of Internet-based connections, both in terms of vulnerability of sensitive information, and the lack of service-level-agreements to guarantee lifeline services.

In a recent survey by The Research Group **http://www.researchgroup.co.uk**, the following question addressed the deterrents to deploying ASP services.

3. IF YOU WERE CONSIDERING ASP DEPLOYMENT WHAT DO YOU FEEL WOULD MOST DETER YOU?

The following answers topped the poll for concerns (the full report is available from The Research Group).

	All	Finance
Bandwidth	42%	48%
Security	41%	61%
Industry Maturity	41%	43%
Service Level Agreement (SLA)	33%	43%
Unable to change ASP applications	27%	30%
Knowledge	25%	26%
Budget	21%	48%

In their summary analysis of the results, they make the following observation:

The mantra of concerns echoes over the last twelve months is repeated once again. However, bandwidth worries appear for the first time, with a figure of 41%, barely managing to overtake the regular objections surrounding security.

Both of these problems need to be resolved in the network that links the SME customer premises and the ASP Datacentre. The network is generally outside the direct control of an ASP, though network problems such as these directly affect the marketability of the services the ASP offers.

This paper describes one solution to breaking through this "network barrier" between the ASP and their customer.

2 Background

2.1 ASP Marketplace

The ASP marketplace provides businesses with IT applications, generally on a subscription basis. They also supply the entire IT infrastructure and support services necessary to deliver these applications - ideal for companies looking to outsource their IT requirements. Applications are typically hosted at a remote data centre, and require a secure data connection back to the customer.

Since the late 1990s when the ASP industry began, the marketplace has been developing rapidly throughout the UK. However, although the ASP industry is developing and brings many benefits to the business environment, growth of this industry has been constrained due to network barriers.

2.2 ASP Opportunity

The ASP business model brings a number of advantages to businesses of all sizes. For companies utilising ASP services, they have continual access to the latest technology, without the risks, costs and administrative responsibilities associated with developing and maintaining IT infrastructure.

The financial benefits of the ASP opportunity are particularly beneficial for the SME marketplace. This includes reduced capital outlay on IT hardware and software, and reduced costs on support, as ASP customers are provided with dedicated, expert support. Companies using ASP services can therefore shift their focus from IT concerns, to their core business.

The advantages are not just financial. ASP services are also flexible and scalable and can cope with any amount of growth or development – key to IT requirements.

2.3 Key Problems for ASP Customers

Key issues that are impeding the development of UK based ASPs, are connectivity and security.

Whilst most ASPs do not own private networks, ASP customers require an 'always on' data connection back to their ASP. At present, choices are limited to a direct leased line connection or a connection over the Internet. There are issues with both of these choices.

2.3.1 Bandwidth

The ultimate solution for ASP connectivity is via a dedicated connection, direct to the ASP, as this eliminates network security issues (both snooping and denial of service) and contention ratios. However, the cost of bandwidth is an issue.

At the start of 2000, the leased line industry in the UK was estimated to be worth in excess of £1.4 billion a year. With few players in the leased line industry, BT has maintained a dominant market share and has set a premium rate for leased line connectivity in the UK.

This is a different story from the United States. Whilst BT controls the 'last mile' in most of the UK, the United States has been deregulated for a longer period and as a result, connectivity prices have adjusted through market forces, at least in the vast majority of business areas.

Whilst high-speed city-to-city costs have declined in the UK, BT still controls the last mile for 64kbps – 2 Mbps leased lines. At present, BT charges on a distance and speed matrix. Most Datacentres are located in London; therefore SMEs outside of London cannot afford connections.

As a result, ASP growth in the UK is being obstructed and only a small minority of companies in the UK have managed to embrace ASP services. ASPs are still looking for an affordable way to bridge the gap between customers and applications.

2.3.2 Security

An alternative way of connecting to an ASP is via an Internet connection. Whilst many organisations already have an 'always on' connection to their Internet Service Provider (ISP), ASPs are utilising this Internet connection to connect their customers in a cost effective way.

However, due to the public nature of the Internet, there are security issues that need to be addressed. Using a firewall and encryption method will increase security, but it can slow down the throughput of the data and ultimately, larger amounts of bandwidth are required for the encryption process.

3 The "Integrated Access Provider" (IAP) Concep

3.1 Business Model

IAPs provide affordable network connections for telephony and data services. Through using today's technology, voice and data services are converged and placed down the same connection to utilise bandwidth. The voice traffic travels over the IAP network connection and is billed by the IAP who provides the voice services, and the amount billed can be used to subsidise the cost of the data connectivity. This method ensures IAP customers experience dramatic savings on their IT expenditure – particularly customers who have previously required more than one leased line connection.

Regardless of the technology deployed, the key to a converged voice and data offering is the quality of service and support delivered by the IAP. As shown in the diagrams, the IAP model is an efficient way to deliver an array of communication services.

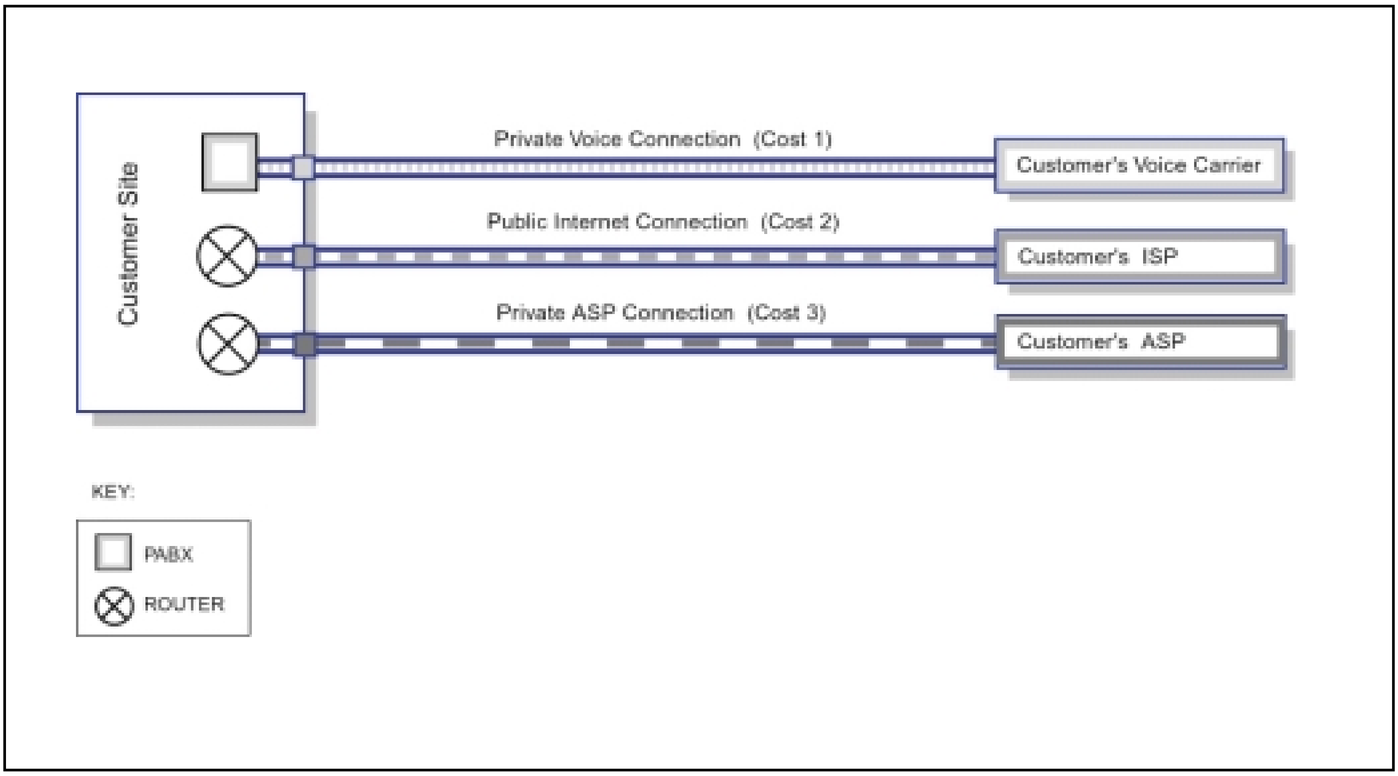

Figure 1 - Traditional Communications Model

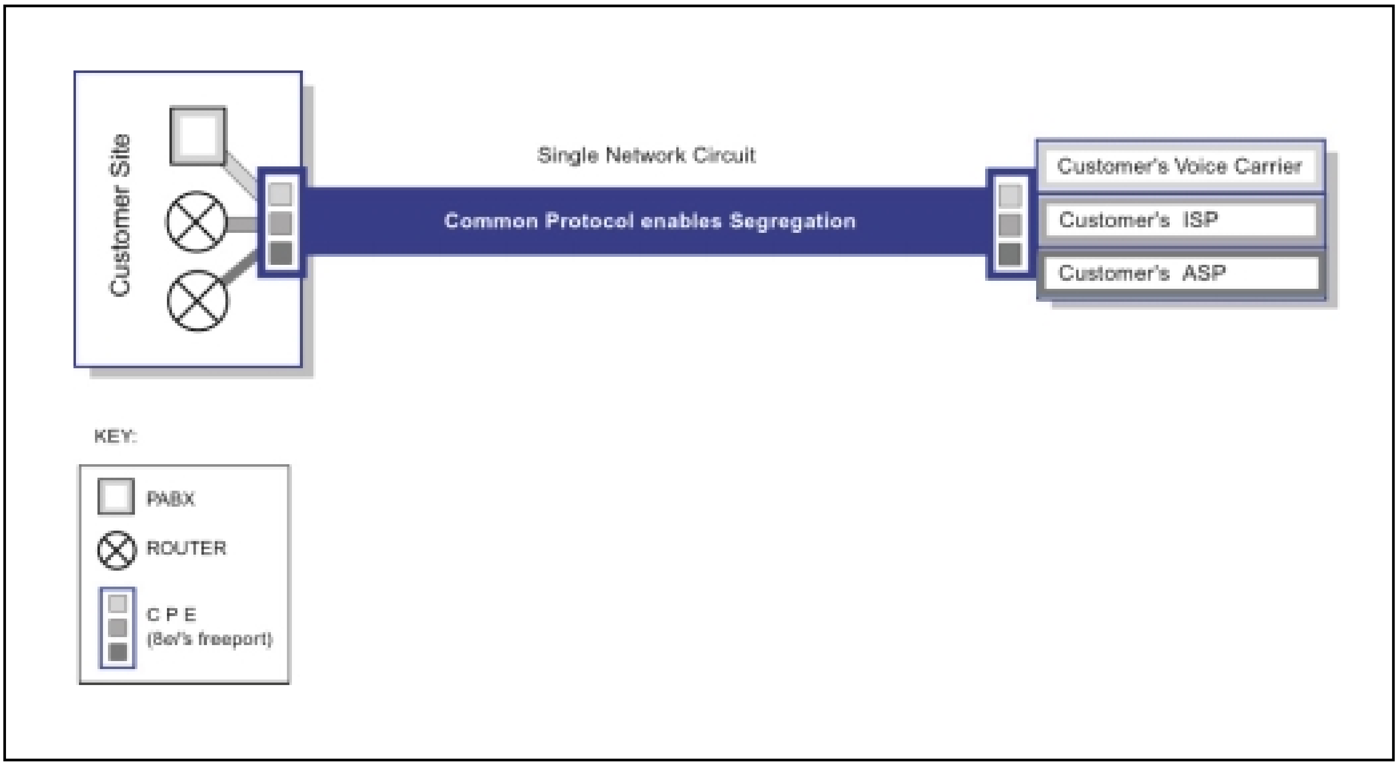

Figure 2 - IAP Communications Model

3.2 Aggregated Telecom Ltd (8el)

Aggregated Telecom Ltd, trading as 8el **http://www.8el.com**, built the foundations for the IAP model in the UK, and is currently using the model to deliver aggregated voice and data services.

3.2.1 History

Founded in 1999, 8el has access to over 100 Points of Presence (POPs) and currently provides services to a wide spectrum of companies.

8el has ensured all voice services delivered over the network are of a high standard, and has invested in excellent customer services and technical support.

The technical expertise in the company includes voice and data specialists, with a particular focus on security and encryption. The services provided by 8el utilises ATM technology, and provides all customers with a "customer premises equipment" (CPE), called freeport. The network management team therefore has visibility of each network connection from the customer site, back to the 8el core. Any network issues can be detected instantly, at both ends of the network connections.

3.2.2 Service Offering

8el provides voice and data connectivity for businesses in the UK. By utilising the IAP model, 8el converges voice and data over private network connections to provide a complete solution.

Connection speeds range from 64Kbps to 155Mbps, with flexible upgrade paths in place for growth and development. Each solution is tailored to suit the customer's requirements, and the customer can also choose the preferred supplier for data services.

8el are a Tier 1 voice carrier and use the subsidy approach for the data connections – i.e. the customer's monthly telephone call spend is used to subsidise some or all of the cost of bandwidth that is used for data. This approach has opened up bandwidth to the SME marketplace, and has particularly benefited ASPs.

While today most services are delivered over fibre SDH circuits, with the unbundling of the local loop (ULL) movements in the UK, 8el are well positioned to take advantage of a lower cost base and expand the service offerings to a broader section of the market.

8el provides connections from the customer site to service providers, and also connections for Virtual Private Network (VPN) facilities. The VPN data connections link customers' regional sites in a private and secure environment, enabling database and file sharing between sites.

All of the 8el services have been built on a structure that has concentrated on reliability, efficiency, security and scalability.

Some of the services currently being connected over the 8el network include

- Closed-Circuit Television
- LAN Monitoring
- Data storage / Warehousing
- Interactive Voice Response
- Video conferencing

3.2.3 Relationship with Microsoft

8el and Microsoft are both working to advance the concept of software as a service, by providing enabling technologies and services to Application Service Providers.

ASPs provide a wide range of applications, and many ASPs provide Microsoft applications. Microsoft reviewed the ASP marketplace and was very aware that the issues surrounding bandwidth in the UK are some of the key contributing factors holding back the ASP marketplace. With the aim of resolving these issues 8el and Microsoft are working together with joint business partners to help facilitate the growth of the ASP industry in the UK.

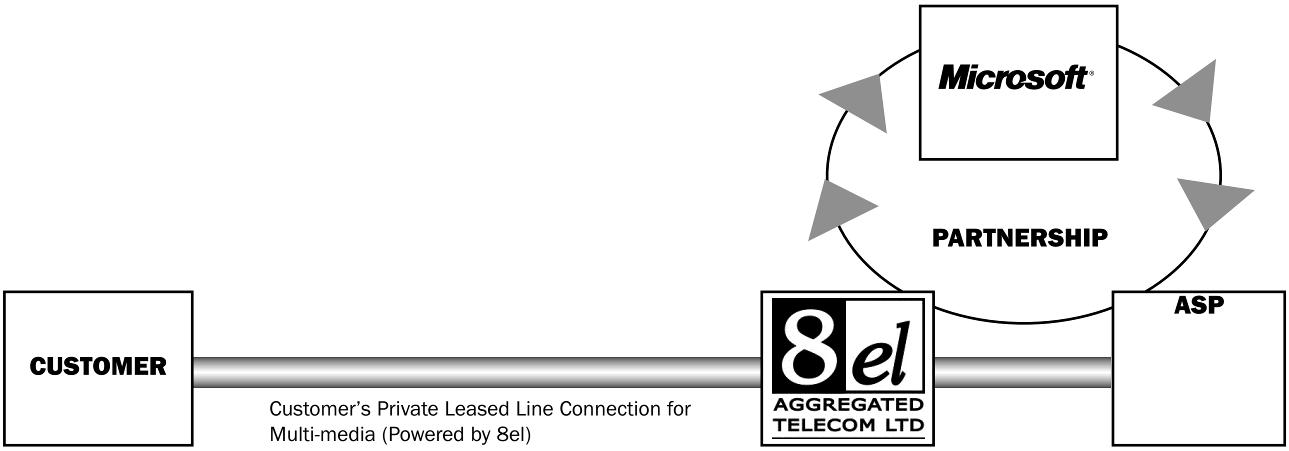

Figure 3 – Partnership model

8el's production and escalation process is integrated with the sales and marketing database in a Remedy solution developed in-house by the customer services-team. During the design of this system much time was spent on evaluating different database systems to run Remedy on. The choice of Microsoft® SQL Server™ was quite easy to make after assessing the maintainers time required for any of the other database vendors.

With a very user friendly administration and scenario-based configuration of all the actions that MS-SQL provides, the implementation of the whole main database system was done in only two days.

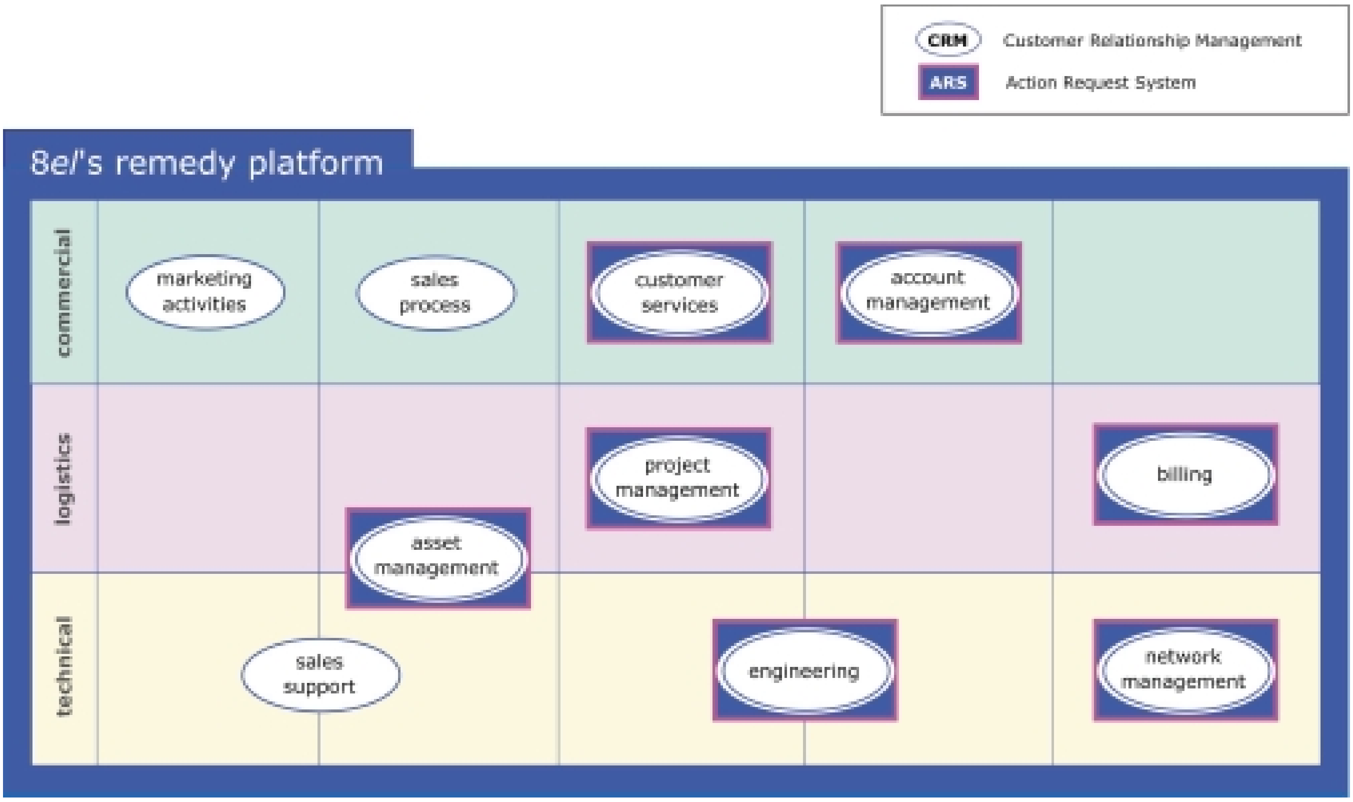

Figure 4 – 8el's Remedy infrastructure running Microsoft SQL Server

Microsoft is pleased to be working in partnership with 8el in the preparation of this whitepaper and associated business initiatives.

3.3 Service Architecture

The IAP provides an additional layer of service on top of existing leased line services. See below:

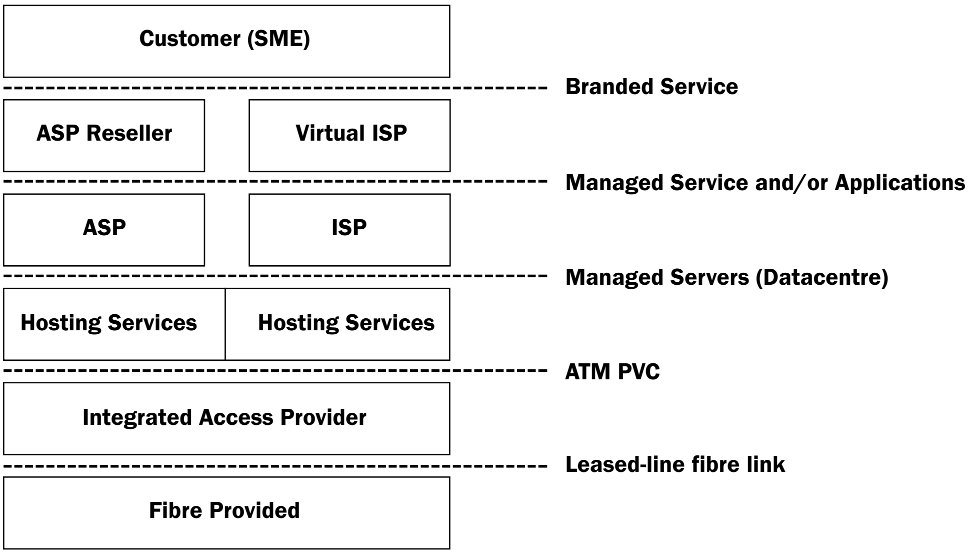

Figure 5 – IAP Service Offering

Each of these "layers" of the ASP business is dependent on the services provided by the layers beneath. Of course, in many cases there will be a single company offering a combination of "layers", such as providing a branded service based on their own applications running in their own Datacentre. However, there are few companies, if any, that have the resources and expertise to supply the entire top-to-bottom solution for a customer without the assistance.

The IAP does business by buying leased-line connectivity from a leased-line provider (PDH or SDH transmission), and providing access to this for both data (as shown above) and telephony services.

3.4 Challenges and Opportunities

The main challenge for the IAP is market maturity and awareness. As the Integrated Access Provider is a relatively new concept, it is new to the communications industry and many are not aware of the IAP solutions on offer. Another challenge is the reservations the industry has due to the new technology being deployed.

Market awareness can be easily increased through advertising campaigns and PR activities, but reservations are more difficult to eliminate. 8el has acknowledged the industry's reservations and find that educating prospect customers on the benefits of voice and data integration is part of the selling process. Business partners also require some form of technical training in converged services, as many companies have either focused in the voice arena or data arena.

To prove the technical feasibility, 8el have live customers (references) who benefit from the converged service.

The opportunities provided by the IAP model are unlimited for all industries. The subsidy approach offered by the IAP has opened up new opportunities for smaller companies. With IAP connections, SMEs can start benefiting from the opportunities that have been so widely talked about, but not affordable due to the cost of bandwidth – ASP applications, E-commerce opportunities, Data Storage centers and Virtual Private Network facilities – the list goes on.

By converging information down the same leased line, the infrastructure costs can be shared out, and multiple services can be provided. This also has huge potential for business partnerships.

3.5 Partnerships

The IAP model using ATM technology relies on good partnerships with voice and data services providers in order to achieve a complete solution. Whilst the IAP is a voice carrier and bandwidth supplier, business partners who specialize in their specific industry supply the data services.

Although an IAP could provide the entire solution, by remaining independent and working with a number of service providers, the IAP is offering customers a choice. The customer can decide on their preferred service providers and utilise one leased line connection for all solutions.

By maintaining independence and partnering with numerous services providers, the IAP is benefiting the entire industry by making data services more accessible to all sizes of companies. This is particularly beneficial for the ASP industry where the cost of bandwidth has hit hard.

4 Using ATM – The Aggregation Technology

The IAP concept described here is based heavily on Asynchronous Transfer Mode (ATM) technology. This section gives a very brief overview of how ATM works, and the way it is used for aggregation of network services by 8el.

4.1 Overview

(This section introduces a very large number of telecom-related abbreviations and acronyms. Please refer to the Glossary at section 11 for definitions).

ATM technology began in the research world of the big telephone companies, as part of their work to develop a "Broadband ISDN", or "Information Superhighway". As such, ATM was designed from the outset to support their existing telephony services, as well as the newer data services such as Internet Protocol.

ATM carries all traffic in small fixed-length packets called "cells", each of which includes a 5-byte header, and a 48-byte payload. The cells are switched very efficiently by hardware in ATM switches, and follow paths called "virtual circuits". Each virtual circuit can by either a Permanent Virtual Circuit (PVC) or Switched Virtual Circuit (SVC). PVCs are the most common, and are setup and cleared down by network management commands from an operator. By contrast, SVCs are setup and cleared automatically by signalling when they are needed, like telephone calls.

The ATM standards support different services by including an ATM Adaptation Layer (AAL) appropriate to the type of traffic. The two most common are AAL1 for circuit emulations (e.g. telephony), and AAL5 for packet data connections.

The ATM protocols and switches support several different Quality of Service classes. The most common are Committed Bit Rate (CBR), for constant predictable streams of data such as telephony circuits, and Variable Bit Rate (VBR) which is used for data where traffic varies a great deal over time, such as Internet data packets. Another class of traffic, "Unspecified Bit Rate" or UBR, while popular in early ATM deployments, has now been largely replaced with VBR.

These connections can be combined on a single physical link, such as the leased line between SME and IAP.

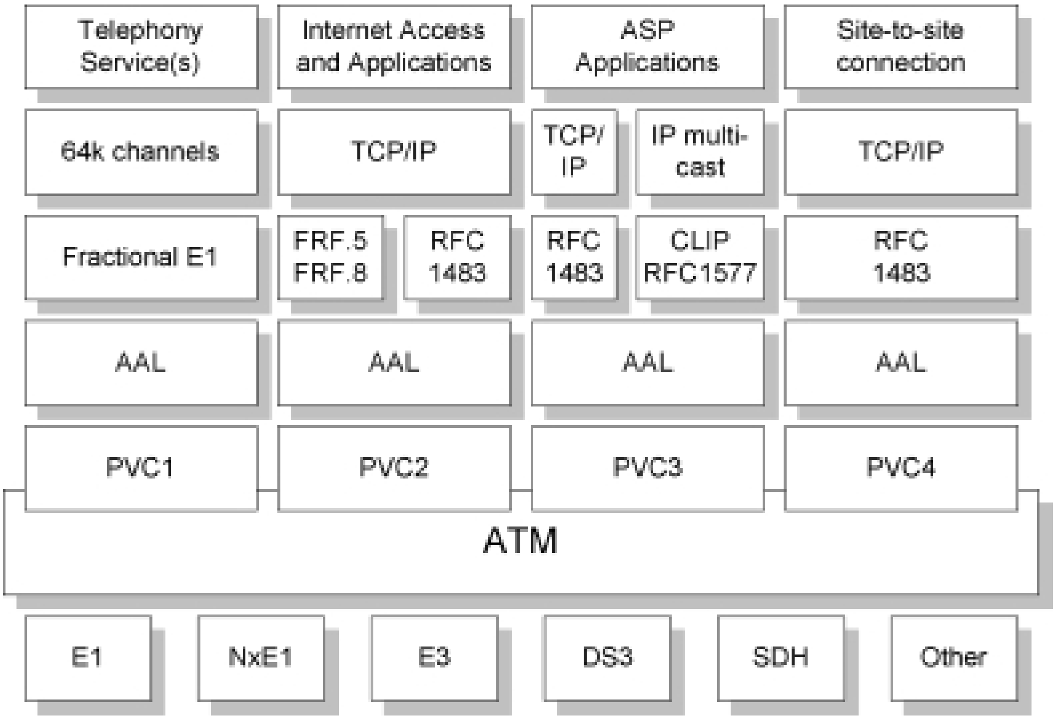

Figure 6 – Example of ATM service aggregation

As shown in Figure 6, an ATM network allows many diverse services to share the same underlying physical link. Each individual connection (either SVC or PVC) can carry a completely different protocol stack. In many cases, the specific protocols used will be decided by the ISP or ASP. The ATM network itself doesn't care what is carried in the VC; it just makes sure it goes to the right place.

Because ATM is transmission-independent, the underlying transport can be changed or upgraded without affecting the connections or protocols being carried. Also, there are many different physical options available for delivering PDH (e.g. E1, E3) and SDH connections across the last mile. The majority are currently on copper wire, but microwave, radio, laser, and fibre optic links can also be used. The IAP architecture described, being built on ATM, works equally well over any of these last-mile technologies.

(Don't worry about the specific protocols shown in the diagram. They are just some examples to show the flexibility of the ATM transport. ASPs and ISPs can use whatever they are most comfortable with.)

4.2 IAP Bandwidth Allocation

The following diagram shows how the IAP uses bandwidth in a typical deployment.

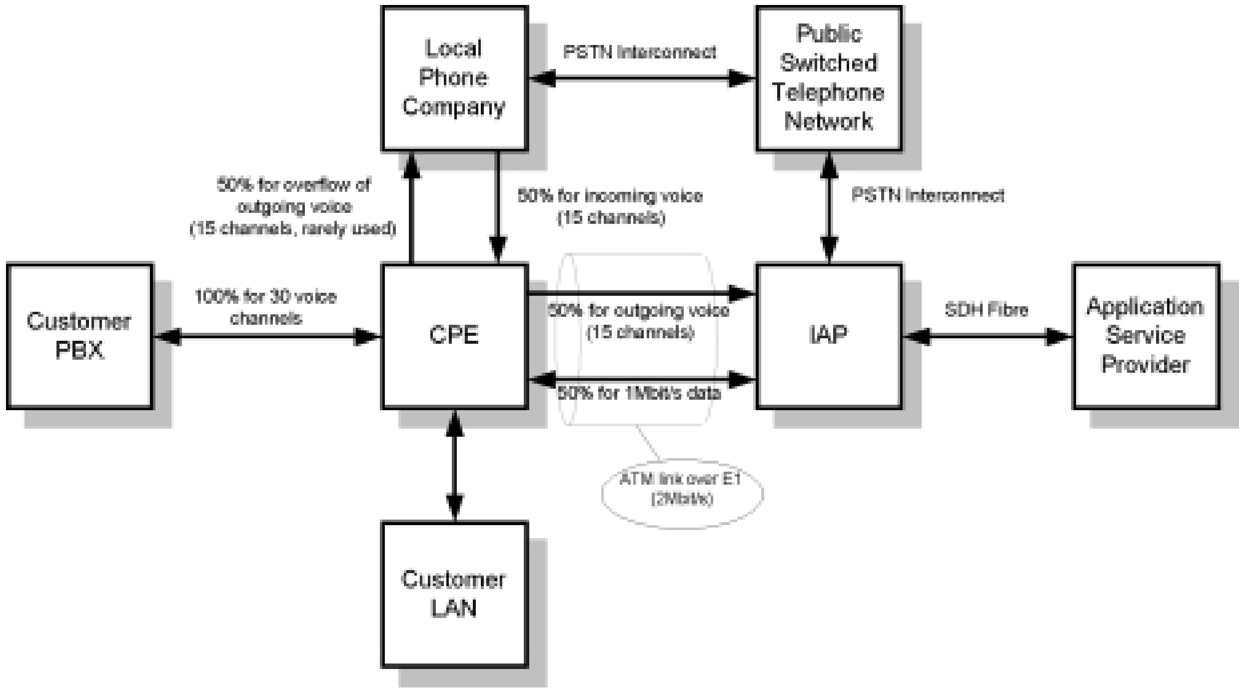

Figure 7 – Example of Bandwidth Allocation

This diagram shows a fully utilised system based on E1 transmission. Many customers will use less than the full number of voice channels, and some larger premises will use multiple E1s or even E3 or SDH to carry the ATM.

The existing ISDN connection to the local phone company (usually BT in the UK) is retained for incoming calls. It my also be used for outgoing calls due to overflow conditions on the ATM link, or during fault management situations.

The new link is used for outgoing calls from the PBX, and for data connectivity to the ASP. In this way the IAP gets to handle all outgoing (metered) calls from the customer site.

4.3 Supporting Telephony services

As noted above, ATM was designed from the outset to support telephony. Specifically, telephony in this case means the support of existing PBX equipment without modification, to a standard of quality and resilience equivalent to that of existing public phone systems. This is a very high bar to meet!

In the world of telephony, lifeline service means real 99.999% uptime, with absolutely NO downtime for planned maintenance, software upgrades, etc. And in the case of a life-threatening emergency, lifeline service can mean literally what it says. Any new voice carrier has to be prepared to answer hard questions on how their technology and service can meet these (often statutory) requirements.

Because the legacy ISDN connection (often a PRI) that previously connected the PBX to the local phone company is retained for incoming calls, there is no need to change the customer's published phone number or wait on the phone company to manage the change.

When calls are originated from the customer, they are routed over the ATM interface to the IAP for delivery to the national PSTN. However, the IAP is (optionally) able to insert a Calling Line Identity (CLI, sometimes referred to as Caller ID) in the call request that shows the number used for incoming calls to the customer on their public line. In this way, the IAP's telephony service appears identical to the local carrier service it replaces. There are no changes required at the PBX, the local phone company, or on the company literature and stationary.

4.3.1 Voice Circuit Emulation

In this case, the CPE is responsible for extracting the live voice channels from E1 from the PBX, and switching them onto a circuit-emulation over ATM, or if this is full, onto another E1 to the local phone company.

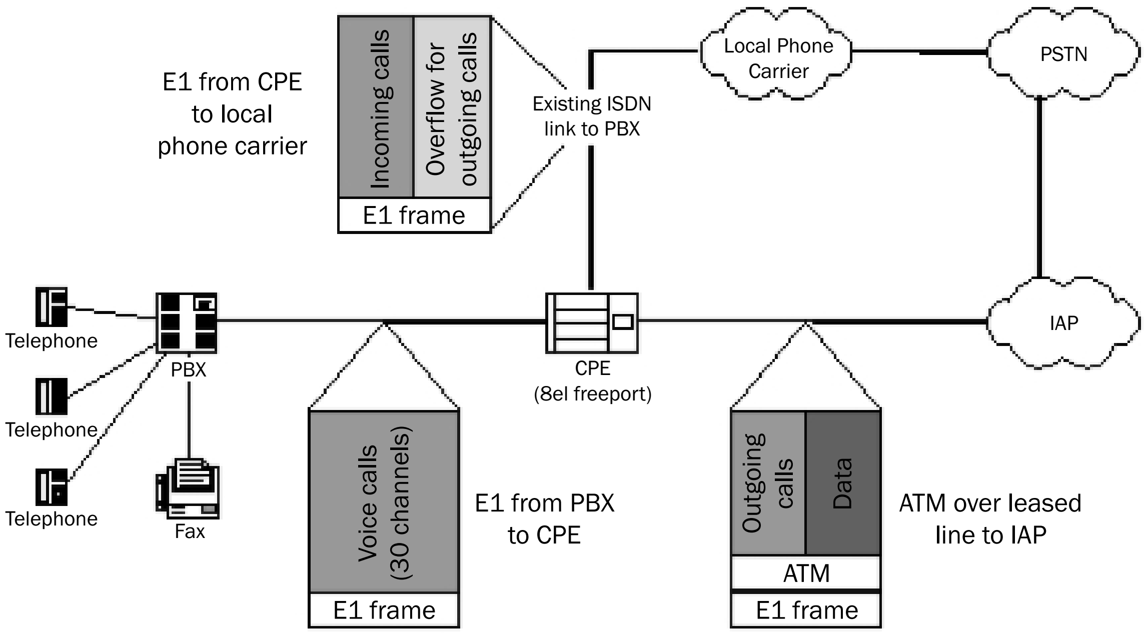

Figure 8 – Voice Connections

This arrangement allows surplus telephony traffic to be routed via the local phone carrier, and allows for failover to the public carrier in the event of a leased-line outage.

The voice channels over ATM are carried using AAL1 on a PVC. This provides for a stable timing environment, and very low latency. 8el's experience indicates the IAP can deliver the voice channels (uncompressed) from the PBX to the national PSTN with a latency of around 3 milliseconds, well below the threshold that would require any form of echo cancellation.

4.3.2 What about Voice-over-IP?

Voice-over-IP is a much discussed and hyped technology. In due course, it may be possible to move the IAP architecture over to VoIP in place of Voice over ATM. However there are several reasons why this is not recommended at this time.

- A good VoIP implementation requires the PBX to be VoIP enabled. This is probably beyond the budget of most SME companies planning to engage an ASP, and could impose quite a hurdle to adoption.
- The quality of voice communications on VoIP still does not meet the expectations of regular phone users. Delay can be quite noticeable (e.g. 300 milliseconds), and usually requires the presence of echo cancellation.
- There are many commonly-used services on the current voice telephone network that are not yet viable using VoIP.
- The price of VoIP equipment remains quite high in comparison to the solution described here. Of course this may change over time.
- Once compressed into H.323 or other IP-based protocol, quality has been lost and it is difficult to connect back to a conventional 64kbit/s telephony voice channel.

However, this architecture does not preclude the introduction of VoIP as a complementary or replacement service at a later date.

4.4 Supporting Internet services

Quite often, an ISP offering Internet access to a SME customer wishes to locate an ISP-managed router on the premises. If this is a requirement, a serial link can be provided from the CPE to the router, running protocols such as HDLC or Frame Relay.

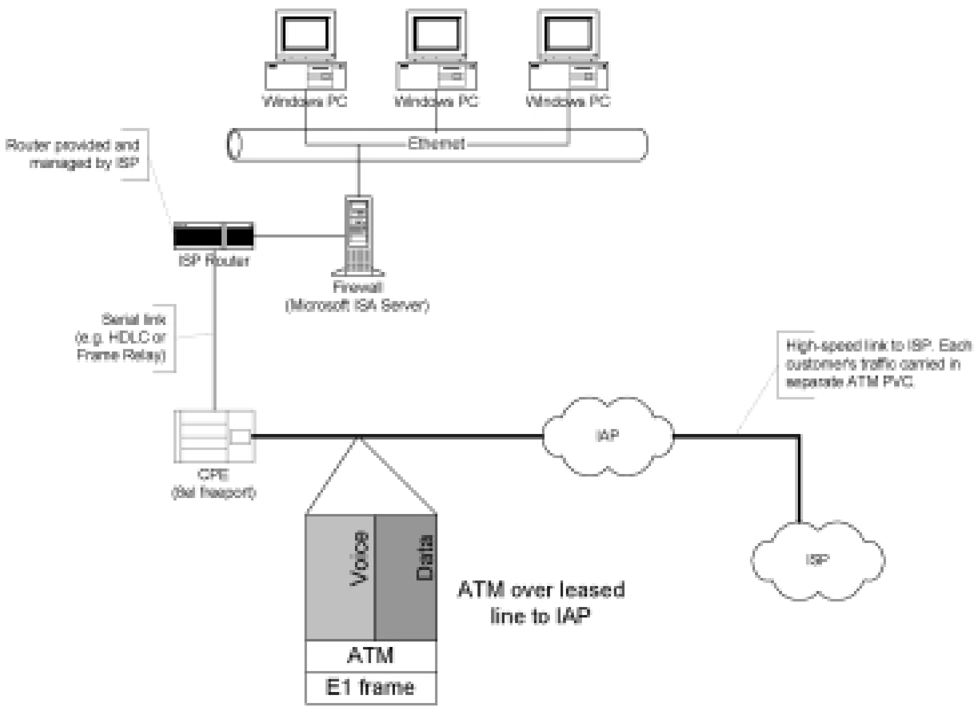

Figure 9 – ISP Connections

This arrangement allows the ISP to control their router using SNMP, just as if it were directly connected on a leased line.

4.5 Supporting ASP services

A similar configuration to 4.4 can be used to concurrently support a link to an ASP. To support this, 8el's Freeport CPE includes an Ethernet port (10/100BaseT) that connects over a PVC to the ASP's Datacentre. This is a dedicated PVC, separate from that used to connect to the ISP (if any).

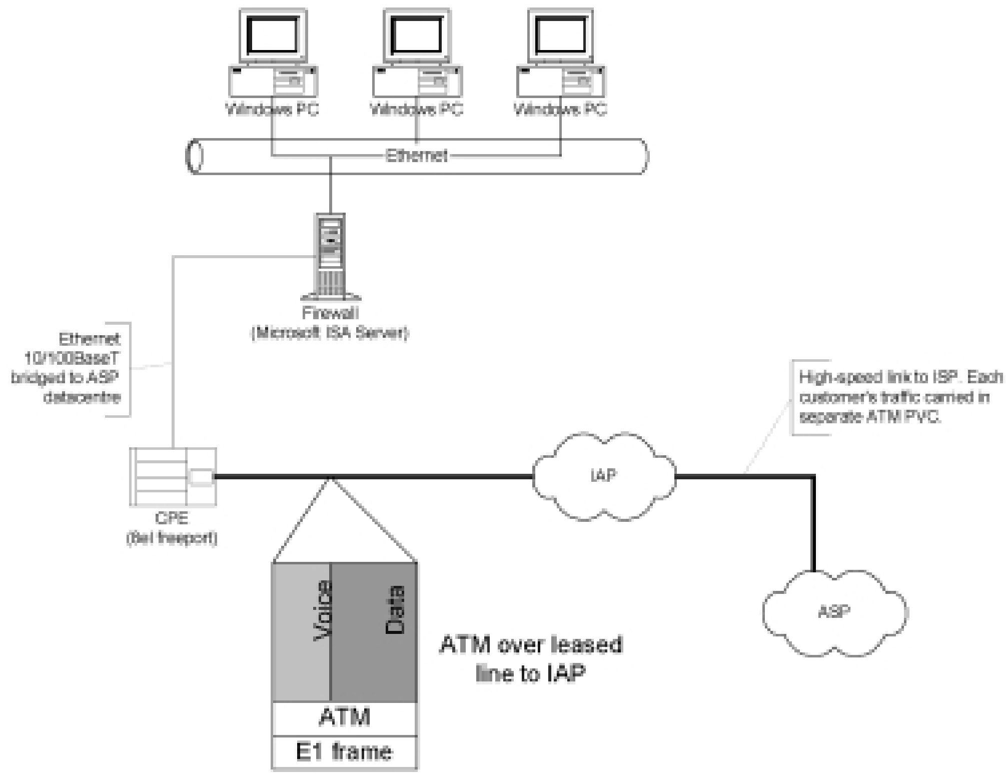

Figure 10 – ASP Connection

The firewall may or may not be required, but is recommended. Not least, those ASP services based on web pages can benefit from the caching available in the ISA server.

A single ISA firewall server can handle both ISP and ASP connections, as shown below.

Each interface from the firewall is assigned a separate IP address. The internal interface and the ASP interface are likely to be private IP addresses, while the interface to the ISP is likely to be a public interface, valid on the Internet.

A decision can be made as to whether the link to the ASP should be treated as external (with firewall protection and caching) or internal (simply routed). In the vast majority of cases, we recommend that both be treated as external, providing maximum security, caching and monitoring. The only reason not to do this would be if the ASP has applications that do not work through a firewall, though some creativity in the firewall configuration can probably resolve such an issue. See the ISA Server configuration guide **http://www.microsoft.com/isaserver** for further information.

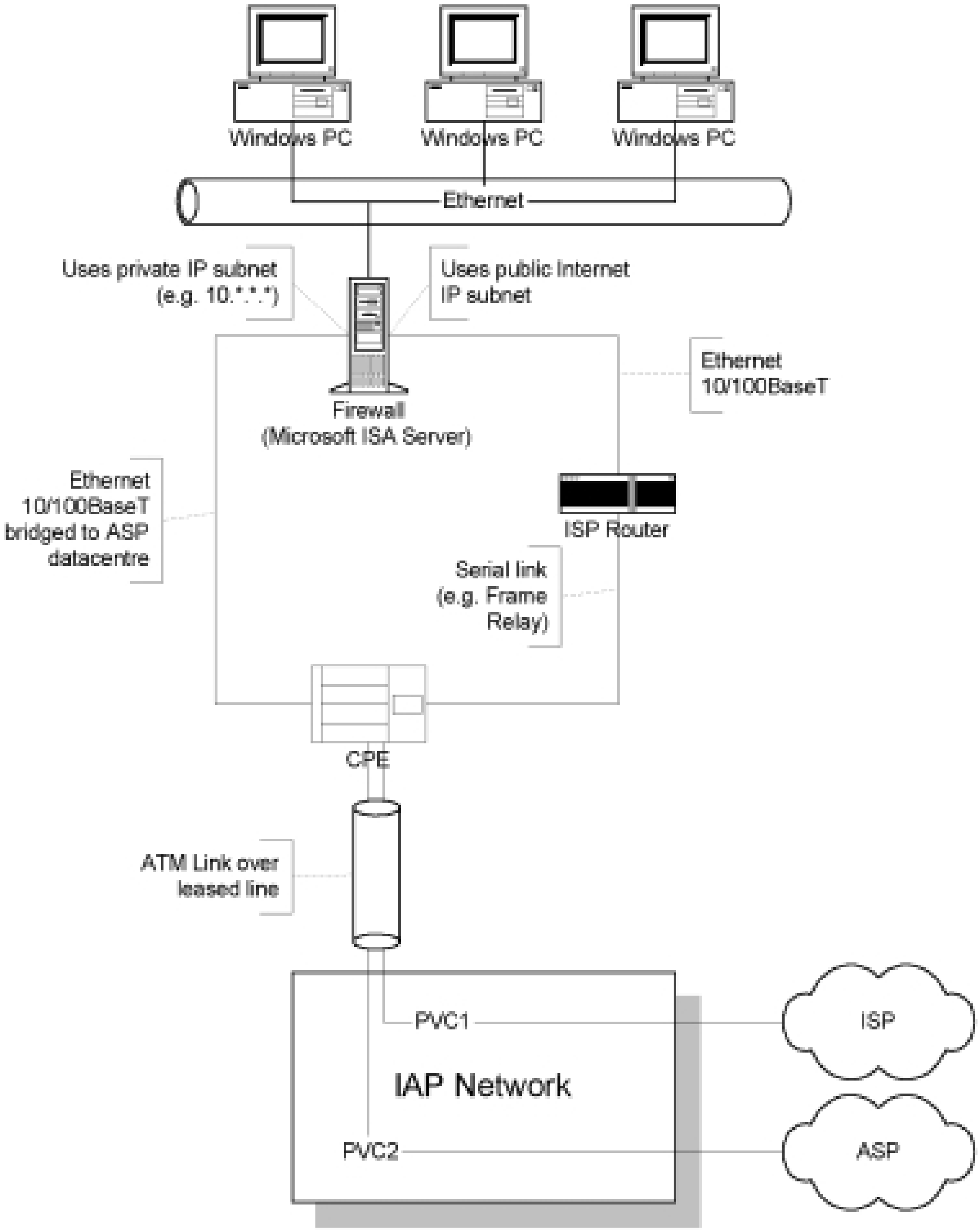

Figure 11 – Multiple Network Connections

It is also possible to provision additional connections to other network destinations, by mapping between an IP address on the router and a specific Frame Relay DLCI on the interface to the CPE. ISPs are almost always unwilling to provide this on their managed routers, so a router supplied and/or managed by the customer themselves, or by the IAP, is required.

4.6 National VPN services

For companies wishing to connect multiple sites, the IAP can provide end-to-end PVC connections across the ATM network. These must be terminated within each customer location, for example by allocation to a DLCI into a router, as shown in Figure 9 – ISP Connections.

In this case, though the DLCI/PVC mapping is managed by the IAP, the VPN traffic remains on the IAP's ATM network (i.e. it never touches the ISP or the Internet), and can use the private IP addressing scheme of the customer. Encryption is not usually required in this case, but could be provided by running an IPSec connection over the ATM connection and putting IP traffic on top of this.

4.7 International VPN access

Many multinational companies have global VPNs that terminate in major cities around the world. However they can run into problems connecting regional locations to these national access points. An IAP solution can therefore help by bridging the gap between a regional office (e.g. Bristol) and the nearest global VPN access point (e.g. London) at far lower prices than a dedicated leased-line.

5 Security

As noted above, the second major objection to using an ASP service after the cost of bandwidth, is worries over the security of the service. These concerns may involve worries about trust in the ASP, and/or concerns over possible intrusion/monitoring of the applications in use by sniffing data packets within the network.

Now, clearly the IAP cannot address concerns with the security of data in the ASP Datacentre, but the IAP network can certainly solve many concerns over security in the network.

5.1 Data partitioning

Firstly, the IAP architecture we describe provides a point-to-point layer-2 connection between the SME and the ASP. This is similar to a leased-line arrangement, in that there are no Internet routers involved, nor does the data share resources with any other customer or Internet user.
If the ASP and SME both use private IP addresses, the hosts in both locations will be physically unreachable from the Internet.

Using one PVC to carry ASP traffic, in parallel with a second to carry ISP Internet traffic (as shown in Figure 11 – Multiple Network Connections), removes the need to run a VPN connection (such as PPTP or L2TP) over a public Internet service. As such, the essential information that flows between ASP and SME is not exposed to any hacking, spoofing, denial-of-service, or any other risks associated with carriage over the open Internet.

5.2 Firewalling

By implementing a firewall at the SME (as shown in Figure 11 – Multiple Network Connections), the IAP, ASP or ISP is able to provide excellent security. Microsoft's Internet Security and Acceleration Server (ISA) is recommended for this role, as it combines excellent firewall protection and first-class caching to improve network performance.

The firewall needs to be configured to allow passage of traffic required for the ASP and ISP applications. In most cases, for ASPs this can be accomplished by allowing all IP traffic to a destination set including all the IP addresses used in the ASP Datacentre. In many cases these will be private addresses such as the 10.*.*.* network, there being little point in using scarce public IP addresses for applications accessed via leased line or IAP service. The ASP will themselves provide security across their own servers, so there is little benefit in adding further complexity at the SME premises. This approach works for applications where the client always initiates the connection, such as most client-server and web-based applications. However, it does not allow a server at the ASP to arbitrarily access a server or client within the customer premises. Some additional firewall configuration (e.g. “publishing” the protected computer by the ISA Server) may be used to get around this issue.

Whoever is managing the firewall servers can take advantage of the ISA Server Enterprise Edition by setting global policies for access to the ASP servers. This central management of ISA servers employs Microsoft® Windows® 2000 Active Directory™ as the distributed store for policy information, and all ISA servers can then be updated with a single management command.

If the SME can justify the expense, ISA will support two or more ISA servers running as an array with common policy, and full active/active load sharing. While a single ISA server can support very large throughput in SME terms, having a pair of machines provides better fault-tolerance in case of firewall hardware failure or interruption. Additional ISA servers can be added and will automatically configure themselves based on the current global/enterprise and array policies stored in the Active Directory.

For further information on the Microsoft Internet Security and Acceleration Server, please refer to **http://www.microsoft.com/isaserver.**

6 Fault-management

This section discusses the requirements for handling faults.

6.1 Requirements

An IAP needs to meet very high standards of fault-tolerance in the service they provide. Remembering that the direct competition is from physical leased-lines from public telephone companies sets the customer expectation.

As such, like the public phone company, the IAP network needs to be engineered from the ground up as a fault-tolerant system with no single point of failure. Customers may accept limited degradation of service under extreme conditions, but a complete outage is not acceptable, especially in terms of voice telephony services.

Of course, fault-tolerance incurs some cost, so there is also a need to make a variety of options available, so that customers and ASPs can make their own value decisions over the level of redundancy and depth they require.

6.2 General points

The leased-line itself can become a single point of failure. This is avoided by continuing to buy voice call capacity from the phone company on a separate PRI link. While this is still a cost, relatively few voice calls will be routed in this direction, so the phone bill should remain manageable.

6.3 Failure scenarios

The following scenarios describe the vast majority of possible outages that could be experience in the IAP network. Of course, there are additional failures that could occur within the SME premises or the ASP Datacentre, but these are well understood and in most cases are not the IAP’s responsibility to solve, although an IAP is ideally positioned to assist with fault analysis and location.

6.3.1 Leased-line outage

Firstly, the service depends on the leased-line from the SME to the IAP. These can fail due to problems in the underlying transmission network (often human error in the management of the transmission, or planned outage due to reconfiguration). Or, perhaps more common these days, failure due to inadvertent interference with the street cabling (the dreaded dig-up incident).

6.3.1.1 Voice services

For fault-tolerant voice services (as shown in Figure 8 – Voice Connections), the service relies on the existing PBX link to the local phone company. This link is normally used for incoming calls to the customer's existing telephone numbers, but can take up the slack with outgoing calls when the need arises.

When the CPE deployed by 8el detects loss of carrier on the leased line (ATM), it automatically starts routing calls through the local phone company using Indirect Access (based on a 1xxx prefix in the UK). The calls are routed to the IAP's voice switch, so that standard telephony services continue to work, and the IAP handles the billing for these rerouted calls.

6.3.1.2 Data services

With the loss of the main network bandwidth, there will be degradation of service in some form. However, some degree of lifeline service can be maintained by the Microsoft ISA server (firewall) dialing out as ISDN call via the PBX (or separate BRI lines) to the ASP datacenter. This ISDN connection can be multilinked, based on the capabilities of the PBX as an ISDN switch, which would allow greater than 64kbit/s connections to be established.

6.3.2 CPE failure

In this case, the IAP's CPE itself suffers an outage, either due to equipment failure or loss of power.

6.3.2.1 Voice services

The CPE deployed by 8el is designed to fail back to the existing local phone carrier PRI service. This means the PBX automatically becomes directly connected to the public phone network, and calls can be made as usual, though they will be billed by the local phone company. Lifeline service is maintained.

6.3.2.2 Data services

As in case 6.3.1.2, the firewall is responsible for rebuilding a data connection using an ISDN call through the PBX or BRI connection.

6.3.3 Firewall failure

6.3.3.1 Voice services

Because the voice service uses traditional and ATM transport, the firewall is not involved in voice communications. Thus the total loss of the firewall has no effect on voice service.

6.3.3.2 Data services

Assuming the firewall is used to protect and optimise all data connections, the loss of the firewall is a serious condition, and all data connections will be lost.

There are two mitigating factors that can be applied.

Firstly, when the network is designed it is quite easy to deploy a second firewall server alongside the first as an ISA Array. This array will provide fault tolerance as well as performance benefits.

Secondly, with the ISA server configured as a member of an Array (even if it is the only firewall machine in use), it can be quickly replaced with another machine should a fault develop. A replacement with the ISA server software installed only needs to join the ISA Array and it will at once be given all the configuration and policy information that its predecessor used, as this information is stored in and obtained from the Active Directory.

For the network demands of most SME businesses, an ISA server can be built on a fairly basic hardware platform, so the cost of providing a second server is not that high.

6.3.4 IAP Network outage

In the event that the entire IAP is taken out, due to massive network failure, major power outage, earthquake, fire, flood, or isolation of the IAP's installation, lifeline services must still be maintained.

6.3.4.1 Voice services

The CPE will become aware of the loss of ATM service on the leased-line link to the IAP. When this happens, it will try redirecting calls using the Indirect Access method described in 6.3.1.1. However, it will soon become apparent that calls are not being connected on the 1xxx number. In this case, the CPE will start dialling the calls direct to their destinations using the PSTN service of the local phone company.

6.3.4.2 Data services

The same solution as mentioned in 6.3.1.2 would apply; i.e. the ISA Server (firewall) would connect to the ASP using ISDN dial out through the PBX or separate BRI line. As with voice calls, this connection would be routed via the national public ISDN service.

6.4 Fault Management

As with any network operator, the IAP requires a clear process for the management of fault conditions. For example, 8el uses a customised version of the "Remedy" system for customer interactions including trouble ticketing. The Remedy system uses Microsoft SQL Server as the backend database, with Windows 2000 clients. A good IAP should be able (and willing) to notify customers that a problem exists even before the customer realises there is a problem.

Section 8.1.4 below discusses the expansion of network fault management into a web service that can be accessed directly by ASP and/or customer applications.

7 Scalability

The IAP needs to have means of expanding the service as customer and ASP needs mature over time.

7.1 Expanding voice capacity

The IAP architecture described allows for a number of voice channels from the PBX across the ATM link. Calls in excess of these can still be made, but will need to be routed by the IAP via their 1xxx Indirect Access service.

For example, a small business requires ASP data access and 10 voice channels (lines) from their PBX. Most of the time they manage with 6 voice channels, the other 4 being for extreme peak load only. The IAP configures the CPE to expose 10 live voice channels to the PBX, and to deliver 6 voice channels along with the ASP data across the ATM link. The remaining 4 voice channels, if used, will be routed to the IAP switch by Indirect Access through the public telephone company network.

As our example company develops, they decide they do not need as much data capacity, but would benefit from increasing the number of voice lines. A call to the IAP results in network management commands being sent to the CPE, which rebalances the load on the ATM link to carry the full 10 channels of voice traffic. They also agree to provide some additional overhead, and increase the lines to the PBX to 12, with the last two channels being routed by Indirect Access.

7.2 Expanding link capacity

ATM is a transmission-independent protocol. It can be used over many different types of physical transport and transmission. As such, should the customer fill up the whole of an E1 leased-line, the transmission link can be expanded to a higher capacity.

The next step up from E1 (2Mbit/s) is to use two, three, or four copper pairs of E1 with an inverse multiplexor (IMA) card in the CPE to make them appear as a single faster link. This increases the capacity to 8Mbit/s . Also, with diverse routing of the physical copper pairs, this approach provides additional protection against link failure, as the loss of a single copper gives reduced capacity rather than complete failure.

Beyond the 8Mbit/s IMA solution, the next level is probably E3 (34 Mbit/s), and then up to an STM-1. However these are quite bug jumps for a single company to make.

One way to ease this large jump is for a group of companies in a business park to share the capacity of an STM-1 link. The park would share an SDH ring, using Add/Drop multiplexors to drop multiple E1s, or an E2 or E3 link to individual premises.

A more sophisticated approach would be to concatenate at the ATM layer, using the CPE for Add/Drop of capacity. The specific transmission architecture will depend on local cost and equipment availability.

In any event, the logical architecture of the IAP service remains unchanged and fully scalable, with the same PVC architecture supporting both voice and data services.

7.3 Traffic Management and QoS

During times of peak usage, perhaps while waiting for the link capacity to be expanded, there is a need to carefully manage the use of bandwidth on the ATM link. Voice channels are dedicated, but the data path will also benefit from bandwidth management.

The recommended approach for this is to use the bandwidth management capabilities of Microsoft® Windows® 2000 Professional or Microsoft® Windows® XP on the customer computers, and the bandwidth policy control provided in the Microsoft ISA Server. This way, the most critical ASP applications can be assured of the resources they need to function correctly, and spare capacity can be distributed fairly, or according to specific policy, for other purposes.

ISA server also provides a good platform for monitoring the use of network resources, for capacity measurement and planning purposes.

[1] 8el’s Remedy system uses Microsoft SQL Server as the backend database, with Windows 2000 clients.
[2] Historical note: At one time there was a E2 transmission format running at approximately 8Mbit/s. This very rarely seen today, though it might be available in some cases.

8 Future of IAPs

In this section we examine some possible future enhancements of IAP services. None of the suggestions listed is currently deployed by 8el, and some are dependent on Microsoft technology that is not yet released. However this section should provide an insight into some exciting future possibilities.

The delivery of specific services will of course depend on customer demand. The intent of this section is to show some of the possibilities this new IAP architecture can support.

8.1 IAP Web Services

One of the central concepts of the Microsoft® .NET initiative is the deployment of web services that can easily be built together to make complex distributed applications. The IAP is in a unique position for offering services that link public network telephony services with ASP applications.

8.1.1 .NET Web Services concept

Historically, developers built applications by integrating local system services. This model gave developers access to a rich set of development resources and precise control over how the application would behave.

Developers have already largely moved beyond this model. Today, developers are building complex n-tier systems that integrate entire applications together from all over their networks and then add unique value on top. This enables developers to focus on their unique business value rather than building infrastructure. The result is less time-to-market, higher developer productivity, and, ultimately, higher-quality software.

We're entering the next phase of computing—a phase enabled by the Internet, specifically by the key Internet technology, Extensible Markup Language (XML). XML enables the creation of powerful applications that can be used by anyone, anywhere. It increases the reach of applications and enables the continual delivery of software. In this context, however, software is not so much something installed from a CD, but is a service—like caller ID or pay-per-view television—to subscribe to through a communications medium.

It does this by joining the tightly coupled, highly productive aspects of n-tier computing with the loosely coupled, message-oriented concepts of the Web. This style of computing is called Web services and represents the next evolution of application development. A Web service is an application that exposes its features programmatically over the Internet or intranet using standard Internet protocols like Hypertext Transfer Protocol (HTTP) and XML. It may be helpful to think of it as component-programming over the Web.

Conceptually, developers integrate Web services into their applications by calling Web application programming interfaces (APIs) just as they call local services. The difference is that this call can be routed across the Internet to a service residing on a remote system. For example, a service such as Microsoft® Passport could enable a developer to provide authentication for an application. By programming for the Passport service, the developer can take advantage of Passport's infrastructure and rely on Passport to maintain the database of users, make sure that it is up and running, backed up properly, and so on.

.NET is founded on this principle of Web services, and Microsoft® is providing the infrastructure to enable this evolution to Web services through each of the pieces of the .NET platform. The next generation of development tools and infrastructure, including Microsoft® Visual Studio.NET, the .NET Framework, Microsoft® Windows.NET, and the Microsoft®.NET Enterprise Servers, have been designed for developing applications on the Web services model. The Microsoft®.NET Building Block Services, the new .NET device support, and the forthcoming .NET user experience will provide the remaining pieces of the puzzle to enable the development of applications that take full advantage of the Web services model.

8.1.2 Billing

8el already provides customers with web-based access to their telephone bills, updated every hour. This is very good for casual examination or queries.

8.1.2.1 BizTalk

Although Call Data Records (CDRs) can be downloaded, a more sophisticated approach is to include the Microsoft® BizTalk™ Server as an alternative front-end to the IAP's billing application. This allows the ASP or customer to obtain the same billing information in a structured XML format, suitable for loading into an application or database (e.g. Microsoft SQL Server 2000). For more information on BizTalk, see http://www.microsoft.com/biztalk.

8.1.2.2 Web Services

However, by moving to a .NET Web Services model, even greater possibilities emerge. The telephony billing system could be made available for real-time examination by ASP applications. For example, a CRM application could actively interrogate to determine when the last call was made to a give client/contact, and the customer record could be updated with the date, time, and duration of all calls from that customer (assuming CLI is available).

8.1.2.3 Smart Tags

For example, a Smart Tag for Microsoft® Office XP could be developed to provide a simple mouse-click link between a telephone number typed in a Microsoft® Word or Microsoft® Excel document and the current call records for incoming and outgoing calls associated with the number typed.

8.1.3 Network Management

As noted in section 7.1, there are occasions where the end customer or the ASP may wish to have configuration changes made to the CPE, for example to change the number of telephony lines that are enabled. At present this is done by calling an operator at the IAP who then issues an instruction from a network management console to the CPE.

An obvious expansion of this would be to publish a web service that allows an ASP application to make this adjustment. With suitable authentication of the request, this would allow faster changes to meet the day-to-day needs of the customer. It would also allow an ASP to develop scheduling or calendaring services that can plan for changes in traffic load in advance, and make the changes when necessary.

It may also be useful to allow ASPs limited access to the network management information of the IAP, particularly for those elements that are likely to be helpful in diagnosing problems being experienced by a customer. The web service approach to distributed application design will make this much easier to implement.

8.1.4 Fault Management

When faults occur, trouble tickets need to be raised and tracked until the problem is resolved. Microsoft BizTalk server can provide a B2B interface for the submission and tracking of these forms in standard XML format, allowing the ASP and the IAP to tightly integrate their problem resolution systems.

Trouble tickets could also be raised and tracked using a Web Service interface, allowing real-time posting and status checks. In addition, critical problems can be raised to all interested parties using a notification service such as the Microsoft® Mobile Information 2001 Server. This would allow warnings of urgent fault conditions (and their successful resolution) to be sent to all interested parties by text SMS messages to mobile phones.

An expansion of this would be the inclusion of a web server using Microsoft .NET Mobile Controls to llow direct access to fault management information from WAP phones, PocketPCs and other wireless devices. This would be beneficial both for managers needing instant information, and engineers working to solve the problems. In fact, trouble tickets requiring field technician support could be directly sent to the engineer using SMS and WAP or PocketPC wireless messaging.

8.1.5 Voice Traffic analysis

The IAP is in the advantageous position of seeing all the telephone voice traffic to and from the customer (assuming all calls are routed via the CPE as shown in Figure 8 – Voice Connections), and can add considerable value by making some of this available to ASP application developers.

8.1.5.1 Successful calls

Firstly, the ASP could take information on completed calls and automatically include it in Customer Relationship Management records, giving the customer a complete picture of their (land-line) communications with their own customers and suppliers.
Such call tracking could also be used to study staff behaviour and activity levels, track down inappropriate use of telephone service, etc.

8.1.5.2 Unsuccessful calls

Because the IAP see all incoming calls, they can also make the ASP aware of calls that are not successfully connected to the customer, either due to congestion, busy lines, no answer, invalid extension numbers, etc.

This information can then be used to advise the customer on the need for additional capacity (see 7.1), better staff training, or the need for IVR services to automatically handle incoming calls.

For example, if an SME's customer complains that they have called many times and not been answered, the ASP could provide a list of the actual date and time of such unsuccessful calls with which to verify or refute this complaint.

8.1.6 Packet-based billing

The IAP is in a suitable position to provide packet-counting services to the ISP or ASP, should they wish to base an element of their billing on this information. This may be of particular interest once Quality of Service becomes an element of network services, to ensure fair access to scarce or high-cost services.

8.2 Conferencing Services

The IAP is well positioned to provide human and computer conferencing services that span both telephony and data environments.

8.2.1 Voice Conferencing

Telephone conferencing is well established for small numbers of callers, where one caller takes the responsibility to bring others in to the group. For larger meetings, there is a need to organise "virtual meeting rooms" with published numbers that callers can dial to access the conference. One of the most flexible solutions for this is to use the Microsoft® Exchange Conferencing Server.

This software allows a number of virtual meeting rooms to be specified, which appears as normal Resources in the Microsoft® Outlook address book. These can then be reserved by creating a Meeting event in the Outlook calendar, and inviting people as for a regular face-to-face meeting. The Exchange Conferencing server can then set up the necessary conference on the voice conference bridge, and notify the invitees of the number and access codes they need to use to attend.

In most cases the Exchange Conferencing server will be owned by and located at the ASP, and will communicate with the voice bridge at the IAP to arrange meeting resources.

8.2.2 Video Conferencing

Video conferences can be supported by dedicated VC equipment, or by the use of PC applications such as NetMeeting. The dominant standard for conferencing over IP is H.323, and a mixture of compliant equipment can be used within the same conference. However there is still a lot of ISDN-based VC equipment, using the H.261 standard and its successors.

The Exchange Conference server at the ASP can be used to manage many of these conferences. For dedicated ISDN conferencing systems, the Conferencing server makes the reservation of resources, and leaves the MCU to handle switching on ISDN circuits.

For PC-based conferencing, such as NetMeeting, see 8.2.3 below.

8.2.3 Data Conferencing

Where the terminal for conferencing is a PC, we have some specific issues to consider.

The client software will probably be Microsoft NetMeeting or an equivalent product, based on H.323 video compression over UDP/IP.

The best way to conference a large group of these clients is to use IP Multicast, so that each terminal is only required to transmit a single copy of the voice and video of the user. This requires the presence of an IP Multicast Server (such as Windows 2000 running the Windows Media Server component). However, smaller conferences can be supported by individual Unicast connections to the Exchange 2000 Conferencing server.

For suggestions on implementing a true multicast solution on an IAP / ASP network, see section 8.2.4.

8.2.4 IP Multicast Support

Accessing a multicast stream through a firewall can be problematic at present, so the service is likely to revert to a Unicast connection model, which is less efficient in the use of network resources. However, network equipment and software is gradually incorporating multicast capabilities, so we expect this situation to improve in due course.

If an ASP customer is very keen to have multicast support today, they need a direct bridged connection between their local LAN and the ASP's multicast router, with no intervening firewall. So a solution can be provided if required, at the expense of some security issues.

As a future evolution, the following approach may prove more effective.

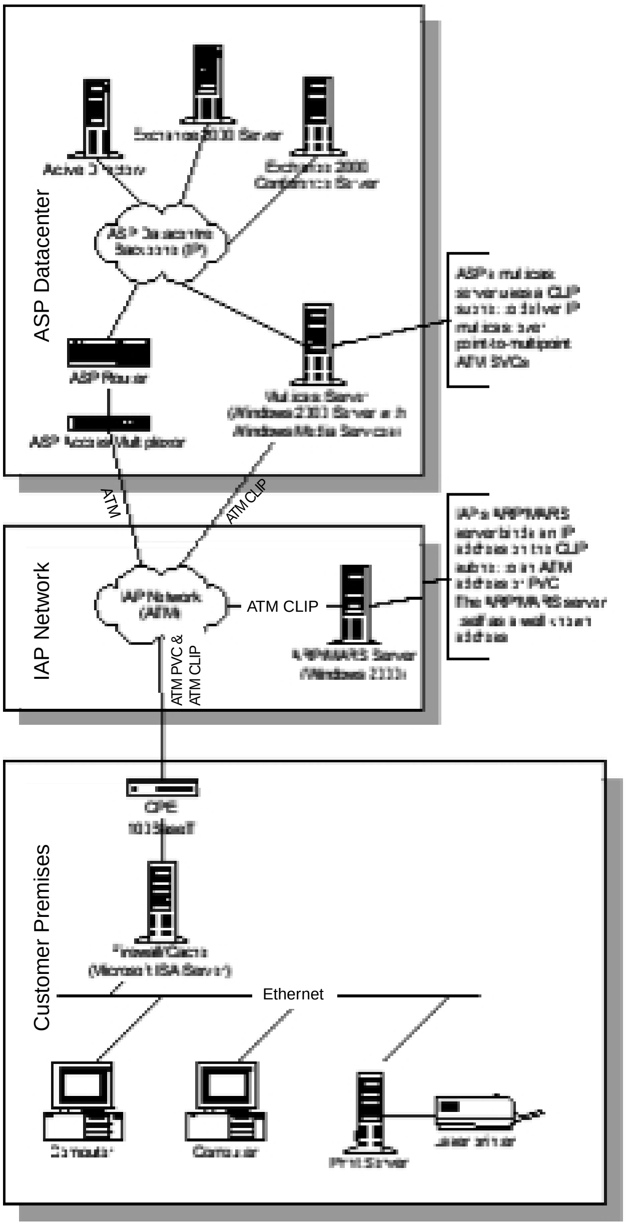

Figure 12 – Multicast over ATM

ATM switching is able to support point-to-multipoint connections, where data be sent just once to a large number of recipient users. The "CLIP" and "MARS" standards were developed by the IETF to map the IP multicast protocols onto this connection type.

The IAP can add a Classical IP over ATM (CLIP) subnet to support dynamic IP multicast over the ATM layer. CLIP is documented in the IETF RFC 1577 (see **http://www.ietf.org/rfc/rfc1577.txt?number=1577**).

Microsoft's implementation of CLIP is known as IPOA (IP over ATM), and is a shipping feature of Windows 2000.

A CLIP network will require RFC 1577 support in the CPE. It also requires an ARP/MARS server attached to the ATM network to resolve IP unicast and multicast addressing on ATM.

CLIP can be used on a PVC-based ATM network, but multicast is far more effective on an SVC based system, where IP multicast can be supported over point-to-multipoint ATM connections. This allows the ATM switch to provide replication of data to multiple recipients, removing the need for the ASP to transmit many copies of the same data to the IAP.

Because the IPOA subnet would only be accessible directly from the ASP or a customer, and would be dedicated to multicast communications, there is probably little need to provide firewalling on this connection, though there will still be as need to provide routing.

Though the multicast connection does not pass directly through the firewall, the ISA server is able to split the received audio and/or video stream to as many recipients as required, without additional load on the ATM link.

The willingness of IAPs to support efficient multicast is obviously tied to the level of demand for such a service, and the availability of network products to deliver it.

8.3 Streaming Media Services

Although streaming media is largely used for entertainment on the Internet, the number of business applications in steadily increasing, both as an item in its own right, and as a component of large product offerings.

As multimedia becomes a growing element of these new applications, ASPs will want to offer video streaming as a way to keep up with the services provided by locally installed software systems. Some of the issues are similar to those for data conferencing, while others are specific.

8.3.1 Presentation Broadcasting

Microsoft's current Microsoft® PowerPoint® presentation tool now includes a facility for broadcasting presentations, where the receiver gets a concurrent video of the speaker, audio, and the slides being presented. The tool can operate either live, or by recording for later on-demand viewing (or both).

Live broadcast replies on the use of a multicast server (see 8.2.4) above, though a presentation to a small audience of perhaps ten can be achieved directly without this. Similar to the data conferencing described above, the material to be shared is sent by unicast from the presenter's PC to the multicast server, then viewers (who visit a web page or receive an email) connect to the multicast feed to obtain their copy of the feed.

8.3.2 Live event streaming

Other events can be streamed directly from a Windows Media encoding machine, and could include material drawn from videotape or live television (with suitable licensing of course). A major event of interest to businesses, such as a government budget announcement or a major conference, can then be offered as part of a portal service to all customers of the ASP.

9 Conclusions

The IAP architecture described provides a good solution for many companies facing the two biggest hurdles to ASP deployment; purchasing network bandwidth, and assuring secure, fault-tolerant connections to services.

ATM is proving itself as the most effective network technology to support the IAP architecture, because it is able to deliver the telephony aspects without compromising on the strict requirements of existing PBX and carrier equipment.

The use of ATM allows the same physical link to be used for multiple voice and data connections, and allows for cross-subsidy of the data service from telephony call spending.

Moving a customer to an IAP service requires minimal change to data and voice networks at the customer site.

8el developed the IAP concept, and remains the UK leader in serving ASPs and customers with aggregated telecommunications solutions.

In addition to solving today's problems, we anticipate the IAP platform described will evolve to offer .NET Web Services that allow ASPs to integrate applications with their customer's telephony system.

10 Further Reading

Microsoft Corporation	**http://www.microsoft.com/uk**
8el	**http://www.8el.com**
Microsoft .NET information	**http://www.microsoft.com/net**
Microsoft Internet Services Network	**http://www.microsoft.com/isn**
Microsoft ISA Server	**http://www.microsoft.com/isaserver**
The ATM Forum	**http://www.atmforum.com**
ASP World Net	**http://www.aspworldnet.com**
ASPIC	**http://www.aspindustry.org/**
Oftel (UK regulator)	**http://www.oftel.gov.uk**
ASP Community	**http://www.aspcommunity.org/**

11 Glossary

8el	Trading name for Aggregated Telecom Ltd, of Reading, England, http://www.8el.com
AAL	ATM Adaptation Layer; several different mappings of services onto the underlying ATM fast-packet transport
ABR	Available Bite Rate; a QoS class in ATM designed for packet data, sophisticated but not much used in real ATM networks, UBR dominates for data transport Application Service Provider
ATM	Asynchronous Transfer Mode; popular Layer-2 fast-packet switching protocol, designed to scale to huge public networks and support widely different services
BRI	ISDN Basic Rate Interface; offers two 64 kbit/s data channels that can be combined into a single 128 kbit/s channel; see also ISDN2 and PRI
CBR	Committed Bite Rate; a QoS class for transport of synchronous circuits over ATM
CLI	Calling Line Identity; sometimes known as Caller id, tells the recipient of a telephone call the number of the person calling
CLIP	Classical IP over ATM, as described in RFC 1577 and successors
CRM	Customer Relationship Management
DS1	American PDH link running at approximately 1.5Mbit/s; sold to customers as T1 service (ANSI standard)
DS3	American PDH link running at approximately 45Mbit/s (ANSI standard)
E1	European PDH link running at approximately 2Mbit/s (ETSI standard)
E3	European PDH link running at approximately 34Mbit/s (ETSI standard)
ETSI	European Telecommunications Standards Institute; statutory standards body of th European Union, though also authoritative across most of the world outside North America, Japan and Korea.
IAP	(in this context) Integrated Access Provider; a company providing a combination of voice and data services
IETF	Internet Engineering Task Force; the voluntary group who design and set standards for the Internet (see RFCs)
IP	Internet protocol; used to route packets on the Internet and other TCP/IP networks
IPSec	Internet Protocol Security; an add-on protocol to TCP/IP providing encrypted secure communications between two computers (VPN)
ISA	Microsoft Internet Security and Acceleration Server; combined firewall and cache solution
ISDN	Integrated Services Digital Network; the global digital telephony and data network
ISDN2	Brand name of BT's BRI service in the UK

ISDN30	Brand name of BT's PRI service in the UK
IVR	Interactive Voice Response; the use of automated systems to filter incoming calls, provide voice mail, etc
L2TP	Layer-2 Tunnelling Protocol, secure VPN protocol based on IPSec
LEC	Local Exchange Carrier; the local phone company
MARS	Multicast Address Resolution Service; supports IP multicast in a CLIP environment
MCU	Multipoint Conferencing Unit; a bridge for video conferencing calls
.NET	Microsoft's technology to enable software-as-a-service; see **http://www.microsoft.com/net**
OC-3	SONET fibre link running at approximately 155Mbit/s; near equivalent to STM-1.
PDH	Plesiochronous Digital Hierarchy: the traditional transmission system used in public telecom networks. Now largely replaced by SDH/SONET on the backbone, but still present in leased lines and at the network edges; PDH circuits of all kinds can also be transported and delivered within SDH/SONET Virtual Containers
PPTP	Point-to-Point Tunnelling Protocol, an early but widely adopted VPN protocol
PRI	ISDN Primary Rate Interface; based on E1 transmission in Europe, and DS1 in the USA, the European version offers 30 channels at 64 kbit/s (e.g. 30 voice lines); see also ISDN30 and BRI
PSTN	Public Switched Telephone Network
PVC	Permanent Virtual Circuit; an ATM connection established and controlled by network management commands (see SVC)
QoS	Quality of Service
RFC	"Request for Comments"; identifier for Internet standards, issued by the IETF
SDH	Synchronous Digital Hierarchy: global standards for replacement for PDH; much more flexible and optimised for fibre-optic cables. SONET is the USA's near-equivalent.
SME	Small to Medium sized Enterprise.
SMS	Short Message System; two-way text messaging service on GSM mobile phone networks, messages are limited to 160 characters, but with assured delivery
SONET	Synchronous Optical Network: the American form of SDH, almost identical, but not quite
SS7	Switching System No 7; the signalling protocols used by telephone systems to create, clear, bill for, and manage telephone calls and data and connections
SVC	Switched Virtual Circuit; a "dial-up" ATM connection established on-demand by signalling requests from a customer (see PVC)
STM-1	SDH fibre link running at approximately 155Mbit/s

T1	American 1.5Mbit/s PDH service to an end user; transmission format is DS1
TCP	Transport Control Protocol; connection-oriented packet transport over IP
TCP/IP	Communications protocol suite used to build the Internet
UBR	Unspecified Bit Rate; a QoS class in ATM for best-effort packet transport, most commonly used for carrying Internet and other IP traffic over ATM
UDP	User Datagram Protocol; connectionless packet transport over IP, see TCP
ULL	Unbundled Local Loop
VC(1)	Virtual Circuit on ATM or Frame Relay; can be either switched (SVC) or permanent (PVC)
VC(2)	Virtual Container on SDH and SONET; used to transport a leased line circuit across the optical backbone; for example a "VC12" container is used to carry an E1 circuit
VPN	Virtual Private Network; any of several methods to provide a secure connection over an underlying public network; often used in an Internet context (e.g. PPTP, L2TP) but not exclusive to it

Notes

Notes