RZ 3316 (# 93362) 01/08/2001
Engineering & Technology 16 pages

Research Report

Everything over IP, IP over Everything

Robert Haas* and Raffael Marty‡

*IBM Research
Zurich Research Laboratory
8803 Rüschlikon
Switzerland
rha@zurich.ibm.com

‡Swiss Federal Institute of Technology
ETHZ-INF
8092 Zurich
Switzerland
raffy@raffy.ch

Everything over IP, IP over Everything

Robert Haas

IBM Research, Zurich Research Laboratory, 8803 Rüschlikon, Switzerland

Raffael Marty

Swiss Federal Institute of Technology, ETHZ-INF, 8092 Zurich, Switzerland

Abstract –

The title of this report, "Everything over IP, IP over Everything", refers to the convergence role of the Internet Protocol. In this report, we illustrate how IP has become the "de facto" networking protocol, with particular emphasis on economical aspects. After a brief review of the wide range of services running over IP today, we examine the most used core networking technologies such as DWDM, SONET, ATM, 10-Gigabit Ethernet, and MPLS, and how future networks dedicated to IP traffic will lead to a contraction of layers. Finally, we review the challenges faced by IP, in terms of scalability, security, and quality of service, and propose an outlook for the technology's use in future networks based on economical grounds.

1. Introduction

The reason the so-called Internet Economy came about is that the Internet Protocol (IP) was designed in such a way that it could scale indefinitely with respect to applications and users it can support in an environment of unpredictable growth, allowing "networks of networks" to be constructed.

IP has come to dominate the networking market for several reasons:

- It is open and available to everyone, encouraging rapid innovation.
- It is application-independent, requiring no proprietary application-layer gateways.
- Services are placed at the edges of the network rather than integrated into the network itself; this allows services to evolve without impacting the network and keeps complexity out of the network core.
- The ability of packets to carry globally meaningful addresses enables network nodes to make autonomous decisions in processing each packet. This allows the distribution of work throughout the nodes, providing redundancy as well as improving scalability.

Our first aim is to show why IP is increasingly becoming the focus in the networking world. This is done in Chapter 2, which discusses some common technologies that IP does and will have to support. Chapter 3 looks at the layers that lie below the IP layer. Here we focus on “access technologies” as well as “core technologies”, with the emphasis clearly being on the latter part. To round off, we look at challenges facing IP in Chapter 4. We discuss services that IP will have to support in the near future. The final chapter summarizes our findings and briefly recalls the most important points.

2. Services over IP

Currently, there are a variety of services that run over the Internet Protocol (IP). A number of technologies such as email, messaging services (e.g. ICQ, Instant Messenger, …), file transfer, and web are already widely used. But as bandwidth is becoming cheaper and broader, new services such as video on demand or audio streams are gaining popularity. IP will also have to cope with all these demands.

2.1 Application Service Providers

Application Service Providers (ASPs) [4] deliver and manage applications and computer services from remote data centers to multiple users via the Internet or a private network. With data processing performed off-site by a vendor, organizations can focus on their areas of core expertise.

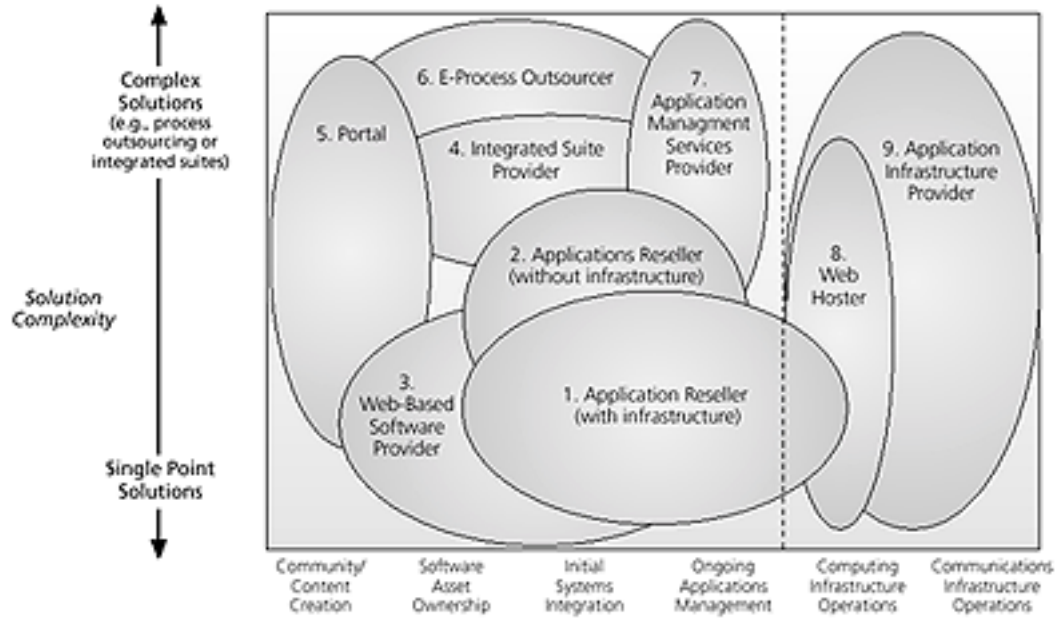

Figure 1: Scope of vendor activities, from [19].

Figure 1 shows that ASPs provide services in the entire range of Internet applications. If IP were not the underlying technique used to provide all services, it would be very difficult for ASPs to offer something that is both transparent to the customer (the user of the Internet) and easily deployable for the ASP itself.

The services that we shall briefly introduce in the following three sections and many more are good examples of things an ASP could provide. For example, a customer can buy a messaging service from an ASP, which in turn will handle all communication needs. The customer itself will have nothing to do with the communication network, hardware and support, except using these things.

2.2 Messaging over IP

A variety of message sources, stores and retrieval devices exist on the market. The major telco-equipment manufacturers have realized this and now try to access that market. Ericsson, for example, is working on a "network for messaging". Their idea is to use the IP layer as communication facility. As Lundquist and Svensson state in [13], "Messaging-over IP is deployed in a distributed messaging network that makes use of the existing transport infrastructure. This approach provides the necessary scalability to meet current and future needs. The choice of IP as the foundation for the networking of components gives messaging-over-IP the same scalability and plug-and-play characteristics as the Internet."

With this approach, the existing messaging means, such as email, SMS, phone or fax, can be combined and interconnected, making it possible to read email over the cell phone or send a fax from a handheld organizer. There will be a single application provider that supports the entire communication range, thereby rendering everything interoperable per se. In the example of Ericsson, where the decision was made to use IP as a communication layer, it will be possible to support almost the entire range of existing communication networks and protocols. This will have the advantage that only well-known protocols and infrastructure are being used. No new mechanisms have to be invented, deployed and tested. Moreover no additional investments into new technology have to be made. The entire existing IP infrastructure can be used. The fact that a telco-equipment manufacturer wants to establish such a messaging architecture may result in decisions that will make it difficult for competitors to use the same service. Imagine a scenario in which only Ericsson cell phones can be used to access this network. Nokia, Siemens, Motorola and all other cell phone manufacturers will not be able to participate. Their only solution would be to set up another separate messaging infrastructure.

2.3 Virtual Private Networks

In order for communication endpoints to communicate securely, companies have dedicated lines that they lease from the telcos. To achieve absolute confidentiality, additional special encryption devices that render data unreadable for third parties have to be installed. A concept that can achieve the same level of confidentiality is that of Virtual Private Networks (VPN) over the Internet. From an economical point of view, the advantage clearly is that no special communication medium is needed apart from the already existing Internet connection. In further comparing the two approaches, we see that both are absolutely transparent to the application layer as encryption is doneon a lower level. Flexibility-wise, a VPN usually is a layer-3 solution, where in the leased line scenario we can have layer-2 (Link Layer) or layer-3 (Network Layer) encryption. Layer-3 encryption provides more freedom in terms of encryption endpoints. In layer 2 we do not have quite the granularity that might be desirable. If we look at interoperability, current VPN solutions are built on IPsec, making them highly interoperable, whereas in the other cases where proprietary technologies are applied, interoperability suffers.

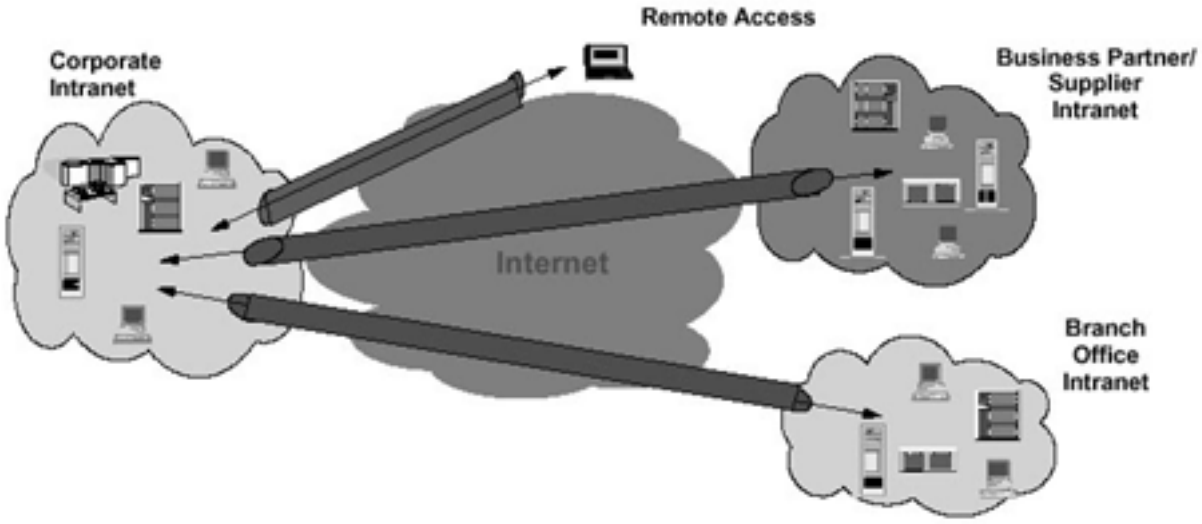

Figure 2: Possible VPN deployment scenario, from [16].

Figure 2 shows a scenario in which a corporation interconnects its business network (left-hand side) with business partners (upper right), a branch office (lower right) and employees working remotely. The only infrastructure the company needs is an internet connection on every side of the “internet cloud” and some VPN equipment at the access points. There is no need for the company to install expensive links between the individual communication endpoints.

2.4 Voice, video and audio over IP

Even though it means cannibalizing their own phone business, carriers are now aggressively investing into voice over IP (VoIP). The main aspect in pushing VoIP is cost. Comparing the equipment required in the telco and the internet world, the price per kb/s for an ethernet hub is three orders of magnitude lower than for a PBX (Private Branch Exchange) with an equal number of ports [17]. The other aspect is the capability to introduce sophisticated services in a much more cost-effective fashion. Whereas enhanced services in proprietary PSTN (Public Switched Telephone Network) platforms are usually mass-marketed and controlled by the provider itself, VoIP services can be customized for specific business needs and provided in an open market. A rough estimate [22] places the development of new switch functions at 5 millions $ and two years, Advanced Intelligent Network (AIN) functions at 1 million $ and one year, and VoIP services at 50 k$ and one month development time. In terms of bandwidth, a VoIP call needs roughly 16 kbps whereas a PSTN call reserves 64 kbps, the gain coming from compression and silence detection.

Carriers have to provide VoIP, otherwise they risk losing customers to new competitors: it is expected that revenues will instead come from value-added services, such as unified messaging. Voice, real-

time video, streaming video, and audio require careful handling by the network. IP being datagram-based, it is apparently more difficult to provide similar quality as a virtual circuit or a leased line service. Therefore, efforts have been undertaken to overcome this problem in the IP layer, as we discuss in Chapter 4.

3. IP over Everything

Every year, there is a tremendous increase in the number of hosts, web sites, and entire corporations that interconnect their sites. Simultaneously, the underlying mass technology for public Internet access allows always faster transmission rates. This all leads to bandwidth being the most demanded commodity at the core of the Internet.

It appears that Internet traffic is multiplying by a factor of approximately four to ten annually. Clearly, the ability to scale up networks to meet the soaring demand is the most critical issue for every carrier. New fiber networks are being installed in every major city, across countries, the oceans and in addition existing equipment is upgraded to allow faster and faster transmission rates. Improvements in current technologies allow faster provisioning, provide better bandwidth efficiency, and scale in a more flexible fashion to higher demands. Cost is a complex function that only partly depends on acquisition costs. Past investments in the currently installed base were huge, therefore interoperability with any new technology is the key. Moreover, cost of operation and maintenance dictates the cost of ownership of any new technology during its entire lifespan. The necessity to „protect“ existing investments while increase the speed of a given network often leads to heterogeneous networks, composed of various technologies that can overlap in terms of functionality and create inefficiencies as we will show below. New network providers starting with a clean sheet do not have such a „heritage“ and can therefore invest aggressively into the latest technologies.

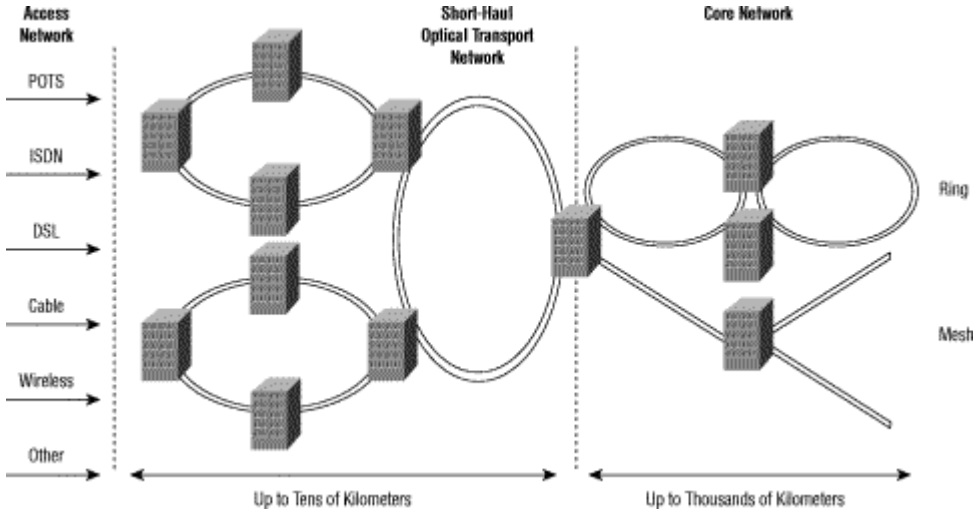

Figure 3: General network architecture.

In this section, we concentrate on some of the access network technologies, which have a direct impact on the traffic that has to be handled by the core network (see Figure 3). In our descriptions, the short-haul as well as the core network are treated together because they use similar technologies.

Wireless access technologies are gaining support, both in the form of fixed wireless, such as WLL (Wireless Local Loop) that allows copper loops owned by incumbent carriers to be bypassed, and in the form of third-generation wireless phones (UMTS, Universal Mobile Telephone System). We shall not discuss wireless in this report, altough it clearly will have a profound impact on the Internet economy as well.

3.1 Access Technologies

In the home market, a few technologies are used to access the Internet. For a technology to become widespread, two factors need to be considered. Most important for the home user is that he or she can use the existing infrastructure to transport data. This means that a fiber link from the user's home to the ISP is not an option. Existing infrastructure needs to be reused. This fact applies to all of the following access technologies:

- POTS (Plain Old Telephone System)
- Integrated Services Digital Network (ISDN)
- Digital Subscriber Line (DSL)
- Cable
- Powerline

The powerline technology is currently being developed. There are test networks that are in that area, but the end user cannot yet buy the technology. In contrast to this, DSL and cable access are becoming increasingly popular, as access providers are updating their infrastructure enormously.

Table 1: Comparison of wireline access technologies from [1].

Technology	Downstream	Upstream	Permanent connection	Modem cost (in CHF)	Cost per kbps (in CHF)	Media
POTS	56 kbps	56 kbps	No	100	~ 2.00	Phone connection
ISDN	128 kbps	128 kbps	No	250	~ 2.00	ISDN connection
SDSL	768 kbps	768 kbps	Yes	650	~ 0.90	2-wire copper
HDSL	2'048 kbps	2'048 kbps	Yes	1'500	~ 0.75	4-wire copper
ADSL	7 Mbps	640 kbps	Yes	2'500		2-wire copper
Cable modem	~ 8 Mbps peak	~ 8 Mbps peak	Yes	500	~ 0.06	Updgraded CATV

If we look at the price comparison for the Swiss market in Table 1, we see that the cable modem approach yields the best return on investment. The problem with this technology is the existing infrastructure. The cable network operators have a network, but only the downstream is fully usable. This is because the TV network only needs data to be sent from the provider to the customer and not vice versa. To enable the network to send data in the opposite direction, considerable investments are necessary.

In Europe we face another problem with DSL technologies [24]. Some years ago the telcos invested heavily into the ISDN network. Now that DSL promises more bandwidth than ISDN, the telcos need a strategy to sell their ISDN services. What they are currently doing is to slow down the DSL development. This does not apply to the US, where the telco companies did not invest a lot in ISDN technology. But as the customer demands will grow for DSL, the telcos wil not have a choice, not even in Europe. In the USA, DSL providers have access to the central offices (CO), where the last-mile copper loops terminate. The COs are owned by the incumbent local exchange carriers (ILECs), who have to let the competitive local exchange carriers (CLECs) install the DSL equipment in their COs. The business of these CLECs (such as Covad, North Point Communications, and Rythm) focusses on DSL access, therefore much effort is put into installing such equipment [9]. In Europe and Switzerland, incumbent carriers such as Swisscom, British Telecom or Deutsche Telekom are slowing down the unbundling of their networks, thereby discouraging new entrants to install DSL equipment in the COs. The European Union reacted to this behavior and forced the incumbents to unbundle their networks, whereas in Switzerland, the network legally still belongs to Swisscom. But there are efforts to challenge this situation.

If we look at the technology aspects, we find good reasons why DSL could be very interesting for service providers. By using ATM with DSL, up to 16 phone lines can be multiplexed over the same copper wire. Here ATM cells are preferred owing to the low delay jitter that can be guaranteed for

voice traffic while simultenaously sharing the link with data traffic. Other technologies cannot provide this advantage at all.

3.2 Core networks

The design of a core network is the key to providing competitive services. It is a well-known fact that carriers are reluctant to disclose the exact internal architecture of their core networks. But by examining information from suppliers of networking equipment, the general architecture can be deduced. Here, we briefly review the advantages of existing core-network architectures, show the problems encountered, and how newer technologies address such issues. Note that these core networks not only serve as support for IP traffic exclusively. However, the share of IP traffic, whether originating from the public Internet or corporate extranets, is clearly outgrowing IPX, SNA and other legacy protocols. The amount of data traffic has in fact already passed that of voice traffic. For public carriers, IP is critical for future revenue growth. The market researcher CIMI Corp. expects that from 2000 on, 80 percent of service provider's profits will be derive from IP-based services.

The technologies deployed in core networks to meet the requirements above encompass the following: IP over SONET, DWDM (Dense Wavelength Division Multiplexing), ATM (Asynchronous Transfer Mode), and MPLS (Multi-Protocol Label Switching).

3.2.1 SONET

SONET [15] (Synchronous Optical Network)[1] is a standard method to interconnect fiber optic systems. Its bandwidth ranges from 51.84 Mbps at the OC-1 level to 9953.28 Mbps at the OC-192 level[2]. SONET uses TDM (Time-Division Multiplexing) to multiplex multiple channels. To have two distinct paths between any two systems, and therefore withstand accidental fiber cuts or electronic equipment failures, SONET systems are built around rings, with fast protection-switching schemes. Each system in the ring has an ADM device (Add-Drop Multiplexer) to pull/inject a point-to-point channel from/into the entire signal. Rings can be interconnected with cross-connects using optical-to-electronic conversion (O-E-O) to perform switching. According to certain vendors [7] of pure optical cross-connects (O-O-O), such high-speed O-E-O cross-connects are not yet widely deployed, and therefore automatic end-to-end provisioning of services (such as an point-to-point 155 Mbps connection) is not possible.

Carriers offer SONET to interconnect corporate sites at very high speeds, either within one SONET ring, i.e. the MAN (Metropolitean Area Network) or across the WAN (Wide-Area Network) with linear SONET connections, such as C1-D1-A1 and B2-A4, see Figure 4. The carrier itself likely provisions the linear SONET connections as hops across its own DWDM/SONET rings.

[1] The European equivalent to SONET is SDH (Synchronous Digital Hierarchy), with only minor differences.

[2] The electrical signal notation corresponding to OC-3n (Optical Carrier) is STS-3n (Synchronous Transport Signal), and the SDH equivalent to both optical and electrical signals is STM-n (Synchronous Transfer Mode). The SONET hierarchy starts with OC-1, whereas the SDH hierarchy starts with STM-1 (corresponding to OC-3).

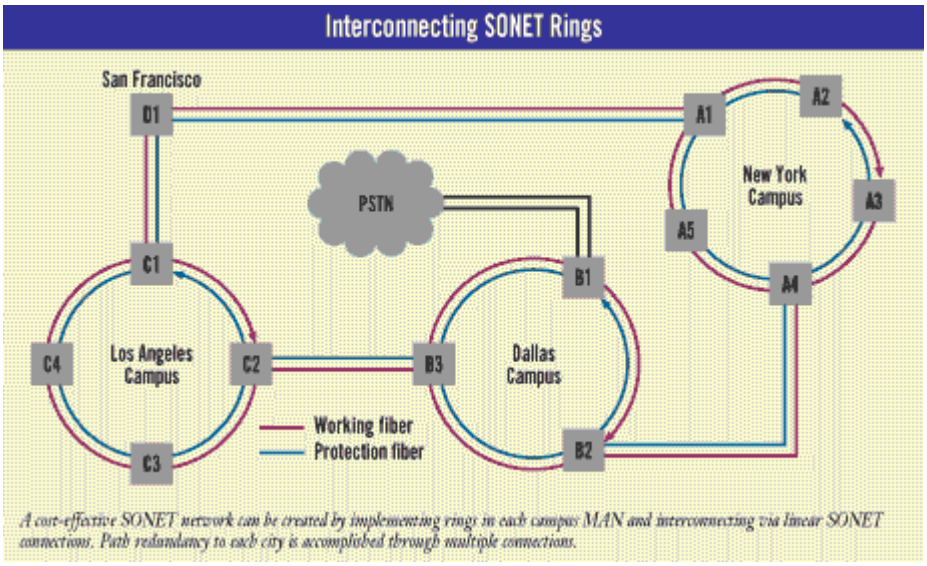

Figure 4: An example of a SONET network interconnecting four remote sites from [8].

Drawbacks are slow provisioning times (a route through the network of interconnected rings has to be found manually, and human intervention might then be required), and coarse bandwidth granularity. Network management systems such as MISA (Management of Integrated SDH and ATM networks) [10] exist that allow automated provisioning of SONET services, but the deployment of such systems is very limited because it requires flexible ADMs and SONET crossconnects. Figure 5 shows how multiple types of traffic are multiplexed together into a single OC-48 stream. Note that in order to benefit from statistical multiplexing, IP and ATM traffic are aggregated into a sub-channel, for instance OC-12.

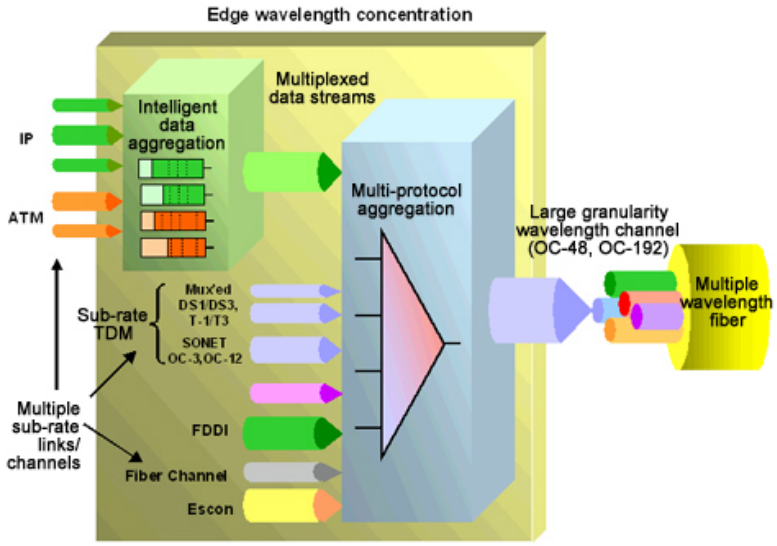

Figure 5: SONET (in blue) and DWDM (in yellow) multiplexing.

To evaluate this technology from an economical point-of-view, the pricing of SONET services is shown in Figure 6, but only for the US market. It gives the monthly charge at various rates for a SONET ring interconnecting three nodes in a MAN. Note the enormous price difference between the linear SONET OC-12 connections and the equivalent switched ATM service offered by Sprint, which consists of 3 PVCs.

SONET Pricing

Local Markets (three-node ring, six miles of fiber)	One-Time Charge	Monthly Charge OC-3	Monthly Charge OC-12	Monthly Charge OC-48
AT&T	$0 to $3,155	$12,340 to $20,771	$23,580 to $34,935	$39,800 to $87,986
BellSouth	$1,020	$5,675	$10,075	$20,875
MCI WorldCom	$1,250	$13,500	$22,800	$53,100
SBC	$650	$6,345	$12,300	$25,600
U S West	$2,910	$9,300	$21,486	$67,944
Coast to Coast		Chicago to Washington	Los Angeles to Chicago	Washington to Chicago
AT&T	$0	$1.32 million	$2.67 million	$3.41 million
Sprint (switched ATM service*)	$4,200	$29,500	$29,500	$29,500

*Sprint does not offer private SONET services

Figure 6: SONET pricing for a three-node MAN and an OC-12 coast-to-coast link from [8].

IP packets can be carried directly in SONET using the PPP protocol encapsulation [14], as shown in Figure 7. The efficiency of such an encapsulation can be calculated as follows: PPP header: 4 bytes, FCS (Frame Check Sequence) either 2 or 4 bytes (4 bytes give the same performance for error correction and detection as ATM AAL-5 encapsulation does), and 1 byte of framing, i.e. a total of 9 bytes per packet of overhead. For both PoS (Packet Over SONET) and IP/ATM/SONET, the SONET overhead is the same. PoS is clearly more efficient than ATM for transporting IP packets: ATM AAL5 and SNAP header/trailer overhead is 16 bytes per packet, plus the cell tax of 5 bytes per 48 bytes of payload. Based on usual packet-size distributions, the IP/ATM overhead is around 25%, whereas the PoS overhead is 2% [2].

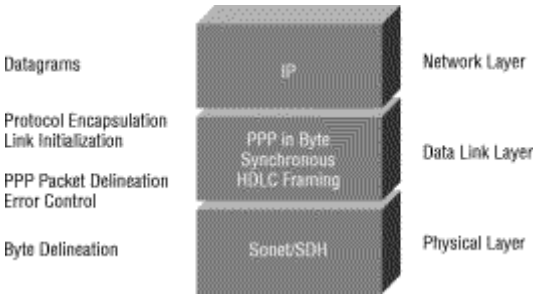

Figure 7: PoS (IP Packet over SONET).

3.2.2 DWDM

[11] is a rather new technology that allows multiplexing over different wavelengths, typically between 32 and 64 [18], the current maximum being 192, thereby virtually multiplying the available capacity per individual fiber. Cost savings on equipment are huge compared to an equivalent number of individual fibers, especially in the long haul where amplifiers are required on each fiber. For a carrier that needs to upgrade its SONET network, adding DWDM makes it possible to keep the existing SONET investment, and scale up the remainder of network based on the newly available wavelengths.

Similarly to SONET, DWDM systems are built around rings, with Optical ADMs in each system. The major difference is that traffic is handled purely optically, and only converted electronically where necessary. Optical crossconnects are also available that switch entire wavelengths optically. Optical

burst switches are in preparation, in which approximately 10 ms bursts of traffic (corresponding to 100 Mb of data at OC-192 rate) can be switched, irrespective of the wavelength.

All-optical DWDM networks in combination with condominium or municipal dark-fiber networks, 10-Gigabit Ethernet and CWDM (Coarse WDM) will allow LAN architectures, concepts and, most importantly, LAN economics (low price per port, simplicity of management) to extend its reach into the WAN, starting with the MAN. The main reason is that in the MAN dark fiber is not expensive, as customers can share the costs of installing a fiber strand. The price can then drop to 100$/month [3] (compare this to an OC-3 MAN charge between $5,000 and $20,000, as shown in Figure 6), because fiber is an investment that lasts 20 years and longer, in contrary to SONET equipment that become quickly obsolete. In addition, traditional SONET networks are built to the highest degree of reliability required by the most demanding customer, whereas the costs must be shared by all customers. This no longer applies when each customer decides which technology to employ over his own dark fiber. Moreover, an all-optical network, where no intermediate optical-to-electronic conversion occurs, renders the SONET capabilities to detect bit errors and carry out automatic network restoration superfluous. Originating from the LAN world, Gigabit Ethernet now becomes a strong competitor for MANs, replacing costly SONET equipment. This requires of a fully transparent DWDM optical network.

3.2.3 ATM

Similarly to SONET, ATM is circuit-based, with the main difference being that ATM circuits are virtual. Instead of performing TDM, each fixed-size cell carries the ID of the virtual connection to which it belongs in its header.This allows one to benefit from statistical multiplexing gain on the link, and therefore make better use of existing ressources. ATM is still the only transport technology capable of guaranteeing Quality of Service, and therefore offers “integrated services”. ATM circuits are also called “software” circuits, and they can be set up and removed quickly. Major drawbacks of ATM are the high overhead it incurs for IP packets and the difficulty to interface IP datagram-based technology on top of circuit-based ATM. ATM circuits can either be set up statically at each ATM switch (PVC, Permanent Virtual Circuits) or dynamically with out-of-band signaling (SVC, Switched Virtual Circuit). VCs (Virtual Circuits) belong to VPs (Virtual Paths), which can be switched as a whole. ATM can support data rates between 34 and 622 Mbps. From the OC-3 rate on (155 Mbps), ATM uses SONET framing. Therefore, ATM switches are commonly used to „groom“ traffic from various sources before it is sent onto SONET rings, so that multiplexing gain can be achieved (see the green box in Figure 5).

3.2.4 Future Protocol Architectures of Core Networks

Today’s long-haul core networks mostly rely on a 4-layer architecture, as shown in Figure 8. At the lowest layer, point-to-point DWDM allows the number of installed fibers to be virtually multiplied, thereby saving the cost of laying new fibers into the ground. At each end of these fibers, SONET equipment provides point-to-point physical transport, although with very slow provisioning capabilities. To perform traffic engineering (QoS support and provisioning), the ATM layer is used with its much faster provisioning times than the SONET layer. Finally, the IP layer on top provides the transport layer.

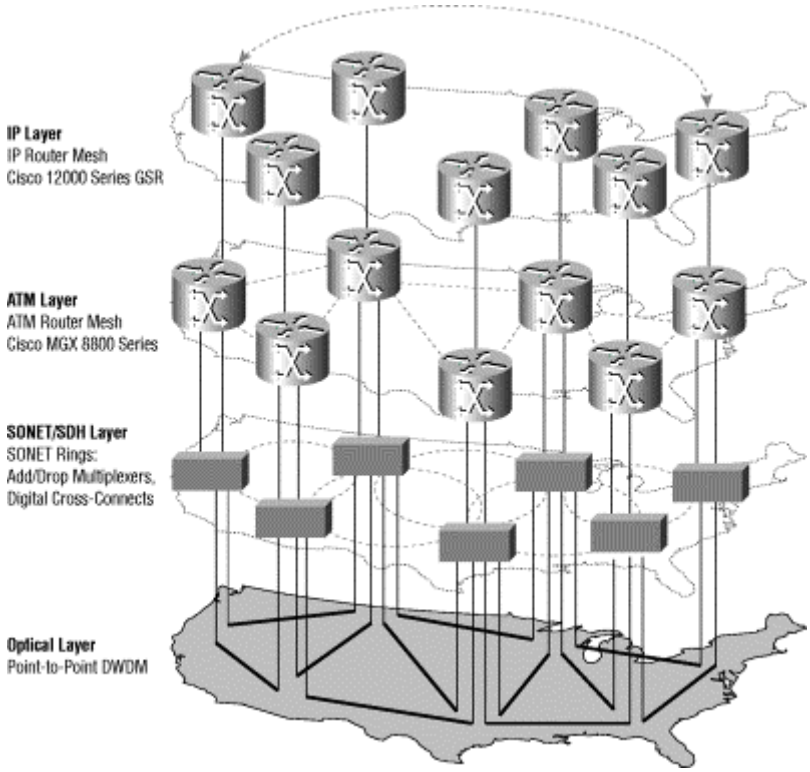

Figure 8: Four-layer model network.

Note that the dynamic QoS-routing feature of the ATM layer often is not present (PNNI) and that instead PVCs (Permanent Virtual Circuits) are set up statically throughout the network. The SONET network consists of rings, interconnected with ADMs (Add-Drop Multiplexors). Setting up circuits through multiple rings still is essentially a manual task, as cross-connects (switches) are not deployed widely. Rings have better fault-tolerant characteristics than star networks do (two alternate distinct paths are available between the same pair of nodes), but lead to a less bandwidth-efficient design as nodes in-between the pair of nodes cannot use the same circuit.

Mostly, such four-layer networks suffer from slow provisioning, dictated by the underlying SONET layer, and a functional overlap. Fault-tolerant features are found at all layers: the SONET layer performs protection switching, ATM reroutes the VCs (only for soft-PVCs with an ATM-PNNI network) and IP finds alternate routes. The combined effect can lead to instabilities at all these layers, and cost inefficiency (most of the time , SONET back-up fibers remain unused). The goal is to include provisioning capabilities directly into either the optical or the IP layer, and drop the intermediate layers, as shown in Figure 9.

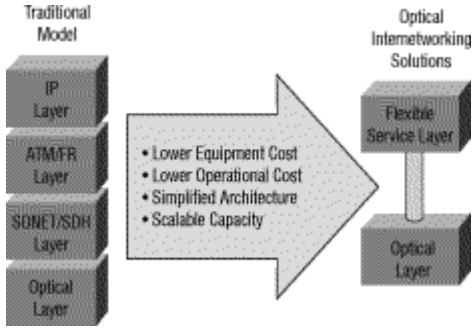

Figure 9: Towards leaner networks.

The trend towards leaner networks with fewer layers relies on the following observations:

- IP router interfaces work at much higher speeds now (OC-48 and OC-192) that are equivalent to the SONET speeds on a wavelength. Wavelength switching in the optical layer can therefore provide similar features as ATM VP switching, albeit with coarser granularity.
- A new protocol called MPLS (Multi-Protocol Label Switching) provides traffic engineering features similar to ATM. Used at the optical layer, MPLS provides the traffic-engineering capability at wavelength granularity (i.e., it replaces VP switching), while used at the IP layer, it provides packet-granularity traffic-engineering.
- Fault tolerance is no longer required of the SONET layer, and it is instead provided by the MPLS control, freeing up much of the back-up fibers.

The trend is therefore to create networks composed of only two layers, namely, the IP and the optical layer, and that have the necessary traffic engineering at each layer, namely MPLS. Proprietary solutions based on SONET for framing of data are being developed (SONET-lite), in which the management overhead of SONET has been streamlined. To provide finer granularity switching while staying at the optical layer, optical packet switches are being developed, thereby imitating the ATM switching concept at the optical layer.

The IP and the optical layer can either function in overlay or in peer mode. The overlay mode, in which IP routers ignore the underlying topology of optical switches, can create scalability problems for routing: establishing a full-mesh between all routers is costly. In the integrated mode, IP routers and optical switches act as peers in the control plane, all running the IP routing protocols. Thereby, the full mesh can be avoided, and routers peer only with their one-hop-away neighbors. Of course, in the data plane, optical switches do not perform IP forwarding, but simply switch based on the MPLS label. The other advantage of the MPLS approach is that traffic is not routed over the shortest route only (leading to imbalances and poor usage of certain links): MPLS paths can be set up over various routes connecting the same end-points, and traffic can be balanced over these paths. Clearly, this requires a certain knowledge of the amount and distribution traffic to be expected, so that these paths can be set up appropriately. This is known as “traffic engineering”.

4. Challenges facing IP

For IP to truly become the convergence layer, it needs to offer seamlessly equivalent capabilities as its underlying technologies do, e.g. being able to exploit QoS from any underlying technology (if any) while providing its own QoS mechanisms, or providing security equivalent to a private leased-line network. Simultaneously, because of the success of IP, several issues arise owing to scaling problems. We review these issues as well, and then discuss the complexity and associated costs incurred by all these challenges.

4.1 Quality of Service

Quality of Service (QoS) [20] [21] in the IP layer is an approach that will have to be dealt with in the near future. Everyone tends to have the IP layer as close to the physical layer (fiber) as possible. This implies that IP will need some QoS functionality. In fact, IP provides an end-to-end communication traversing networks with varying bandwidths and link layers, where bottlenecks can occur. One could argue that network bandwidth will grow substantially, and that servers will then be the main source of congestion. But regardless whether QoS be network-centric or server-centric, it has to be handled by the IP layer. Moreover, networks in which resources are naturally constrained (such as wireless networks) will require means to arbitrate QoS. In addition, an accounting mechanism for charging has to rely on the treatment given to packets to differentiate between tariffs, and this encompasses issues such as QoS as well as security and fault tolerancy. IPv4 currently does not enforce any QoS demands. It provides what is called a “best-effort” service. Unlike VC technologies such as ATM and Frame Relay, IP does not make hard allocations of resources. This provides much more efficient use of the available bandwidth, as well as more flexibility. Typical network traffic is bursty rather than continuous. Being datagram-based, IP uses the available bandwidth most efficiently by sharing what

is available as needed. This also allows IP to adapt more flexibly to applications having varying needs.

There are currently two approaches to obtain a certain QoS in an IP network: a quantitative approach (int-serv) [6][23], in which necessary resources are reserved in each hop of the network by signaling the required QoS of a flow, and a qualitative approach (diff-serv) [5], in which packets indicate the treatment (similar to a priority) they want to receive, but no resources are dedicated to any flow. It is clear that either approach has its drawbacks. Int-serv requires that routers keep state for a large amount of flows, in contradiction to the connectionless Internet design, and hence is not scalable. Diff-serv cannot provide hard guarantees unless proper admission control and traffic provisioning has been performed to ensure that the network does not receive more traffic than it can handle.

Underlying technologies such as SONET/SDH do not provide QoS, apart from a guaranteed-rate circuit (OC-n). ATM virtual circuits on the other hand support various types of services, such as CBR (Constant Bit Rate), VBR (Variable Bit Rate), ABR (Available Bit Rate) or UBR (Unspecified Bit Rate). The int-serv Guaranteed Rate and Controlled Load services easily map onto the corresponding ATM services, and a mapping for diff-serv is under way. Ethernet has certain QoS features, albeit much more limited ones than ATM.

4.2 Security

IP has the property that all data it carries is in plain language, i.e. unencrypted, thus enabling a malicious user to monitor a conversation without any trouble. This approach leaves it up to higher-level protocols to add confidentiality (encryption). Furthermore, IP does not support any kind of authentication. The recipient of an IP packet could look at the source IP address of a packet to find out where the packet originated, but it is easy to forge a source IP address (IP Spoofing). If a packet carried a digital signature, which signs the message and its header, the recipient could check the signature and notice any modification to the packet. IPSec, the security add-on of IP encrypts and/or signs IP packets. The sending party will generate the signature. The receiving party checks the signature and decides whether the packet is authentic. This check can only be made if the recipient knows the public key of the sender. In this scenario arises the problem how the recipient can authentically obtain the public key of the sender. This problem still needs to be resolved. There are approaches using Public Key Infrastructures (PKI), but these alone do not solve the problem. A global solution needs to be found in trust management. On a theoretical level, the problems have already been solved, but technicaly, economicly and politically, several open issues remain. On the technical side, the IP protocol needs a mechanism to obtain the CA's (Certificate Authority) public key in a secure manner (authentic). One could imagine that this would already happen in the manufacturing process of network devices. On the political and economic sides, there is the need for a global public key infrastructure. The exact topology can be arbitrarily complex. Currently available on the market is a construct, in which some companies (such as Verisign and Baltimore) have created an infrastructure themselves. They are all completely independent of each other. This implies that cross validation is not possible. There certainly are economic reasons for this. Political boundaries are the third aspect where solutions need to be found. There are questions such as "How easy is it to deny access to certain countries?" Reasons for wanting this are manifold. Another question would be "Is it desirable to have the functionality to exclude anybody? What about revocation of certificates?" As we see, there are plenty of solutions to be found.

Security support in IP is a key element for the growth of IP-based corporate networks (extranets). A majority of such private networks still rely directly on private leased lines, Frame Relay connections, ATM VCs, or even Ethernet VLANs (Virtual LANs) in which security is "physically" guaranteed. However, moving towards an IP-based VPN requires that security features are added into the IP protocol itself.

4.3 IP Multicast and Broadcast

Built into the heart of IP is the ability for traffic to be broadcast and even multicast. Nowadays if a user requests an audio or video stream from the web, the stream is sent as any other traffic in a point-

to-point manner. Only very few organizations and products make use of the multicasting possibilities. As the web grows and more and more audio and video is made available, ISPs (Internet Service Providers) will have to think about the use of multicast. The end users on the other hand do not care whether the traffic reaches them via multicast or a plain point-to-point connection because their access lines will be used in the same way in both cases. But the ISP will probably care that its backbone is being used (and charged) as many times as there are users subscribing to the same stream. On the other hand, service providers nowadays need a huge connection to the net in order to meet customer demands in terms of bandwidth. If they could use multicast, bandwidth would shrink enormously. Note that, unlike ATM or Ethernet, SONET/SDH does not provide support for multicast.

4.4 Addressing and routing

A threat to IP's scalability is the growing size of routing tables in major exchange points as well as the shrinking of the available address space. Routing tables with more than 100 k entries are commonplace. This can be explained by the number of exceptions in the aggregation of prefixes: organizations that request a portion of the IP address space get portions of the address space that do not necessarily aggregate with the prefix of their ISP. The reason is the exhaustion of the IPv4 address space, where only smaller segments can still be given away.

Although fast IP routers can cope with current table sizes and MPLS allows of short-circuit address lookup in the core networks, IP routing tables will continue to grow in size, endangering scalability of IP routing protocols and forwarding schemes. The current answer to these problems is IPv6. First, it will have a sufficiently large address space to assign IP addresses to every mobile phone and even more. Second, IPv6 avoids a significant number of exceptions in the aggregation by adequate renumbering of addresses to comply with the upper-level prefix.

4.5 Impact of these challenges

The challenges mentioned above will be solved only at a price, which we try to evaluate and justify in the following:

To provide int-serv QoS, all routers on the desired path have to support sophisticated queuing and scheduling mechanisms. For scalability reasons, when the number of flows expecting a certain QoS is high, per-flow int-serv QoS becomes too expensive (processing of signaling messages and hardware requirements are important). Diff-serv does not incur such costs, but requires instead intelligent provisioning over the entire the network and policing at the edges, where the number of flows is significantly lower. MPLS traffic-engineering capabilities are then used to create switched paths through the core network. Compared to ATM, diff-serv with MPLS signaling is not significantly different. These features can be integrated into existing IP equipment, thereby saving the cost of modifying the underlying ATM equipment, is a strong promoter for QoS support in IP. This does not take into account the difficulties to manage an IP and an ATM network, compared to a single network.

Security support in IP requires encryption capabilities, either in hardware or software, which are available today as add-ons to router interfaces. Although it is difficult to evaluate the management cost of an IP-based VPN (but in leased-line private networks it cannot be ignored neither: the effort to create the point-to-point connections between various sites can be lengthy), IP clearly is a much more flexible infrastructure than a static private network. For instance, employees can connect to their corporate VPN simply by connecting to the Internet. This is cheaper than maintaining modem banks at the corporate site (dial-up and hardware maintenance costs).

Partial relief of network as well as server congestion, can be obtained by using IP multicast together with caching. Multicast is supported in ATM and Ethernet (as broadcast), also when they run directly over DWDM rings (i.e. not via SONET as intermediary; even Packet over SONET uses PPP, which does not provide multicast support) and is also supported natively over cable-TV networks. If multicast is to be deployed on a larger scale over heterogeneous technologies, IP has to provide the necessary overall control. Some experiments have already acknowledged the substantial time gain of a multicast database synchronization between headquarters and branch offices, thereby saving server-

CPU time and expensive network bandwidth. Caching, on the other hand, can only be executed at the higher layers (HTTP). The flexibility of the open IP infrastructure allows the seamless integration of these caching mechanisms, another proof for the validity of the end-to-end concept.

Finally, also the need for transition to IPv6 will become more pressing, when large numbers of mobile handsets receive their own IP addresses. This challenge is probably the most expensive one in terms of time and money. It impacts not only network nodes but also end nodes, as networking stacks have to be updated everywhere, and some software has to be rewritten to use the latest IPv6-compliant socket APIs. Therefore, only a gradual transition will take place.

5. Conclusion

"Everything over IP over Everything" can be understood in two different ways, and both ways are important. First, in terms of transport protocols, IP is the clear winner over SNA and IPX (according to a survey made by the Gartner Group). The acceptance of a uniform network layer, together with the corresponding transport protocols, fuelled the creation of a vast number of services to be run on top of it. With this this in mind, the WWW is the natural extension of IP, and acted, as we know, as a "booster" for the public Internet. The growth of IP-based VPNs will contribute to the phasing out of the costly circuit-based private networks: IP-VPNs are ubiquitous because they extend beyond Frame Relay or ATM networks, and the deployment costs of IP-VPNs are significantly lower than for similar technologies (costs are not distance-dependent).

The second and more direct interpretation of the title shows how all known transport technologies are nowadays used to carry IP packets, from switched circuits to shared wireless LANs, from optical wavelenghts to cable TV networks. But paradoxically, most of these technologies were not developed from scratch to specifically support IP packets. Instead, they were either developed to respond to consequences of the success of IP, namely the growth of data traffic, or retro-fitted to accommodate the ubiquitousness of IP. A network designed for IP has yet to come.

Clearly, given the wide range of services running on top of IP, enhancements are required, such as header compression for better voice support, QoS for predictable delays and bandwidth, security to allow business transactions to take place safely, and billing mechanisms, to name just a few. This will eventually make IP a more complex protocol, and may lead to dedicated alternatives being preferred in specific cases.

As technology has evolved, the common denominator has been rising through the OSI layers from IP to HTTP. Web browsing over mobile phones, without any IP layer underneath, is just one example. In business applications, it is the middleware layer, such as CORBA, that assumes the convergence role. This can be viewed as natural evolution: more complex applications are designed, and therefore rely on always higher-level building blocks.

Communications can be divided into three types: human-to-human (sensitive to delay and jitter), computer-to-human (such as web, video streaming, etc), and computer-to-computer (email, ftp, database synchronization, storage network). In the latter two cases computers can provide buffering to cope with delay variations, and best-effort service is sufficient if networks are correctly dimensioned. IP supports all three types of communications, although historically it is best-effort-oriented. Applications running over an IP network are inherently flexible, and this requires that also underlying technologies be flexible (dynamic provisioning, i.e. placing bandwidth where needed when needed). The Intelligent Optical Network, in which wavelengths are set up dynamically like circuits, is an answer to this need in the core networks.

Given the rate at which IP traffic increases, the infrastructure has to scale accordingly. Therefore, it is necessary to streamline the underlying technologies to avoid rapid obsolescence and costly upgrades, as proposed in the IP over DWDM concept, which bypasses SONET. IP over DWDM requires traffic-engineering capabilities provided before by the ATM and SONET layers: this is replaced nowadays

by adding MPLS to IP. Given the speeds at which router interfaces operate, the SONET multiplexing capability is not necessarily required in high-speed backbones: IP can then directly run over DWDM if certain restoration capabilities, originally provided by SONET, can be integrated either into the DWDM or the IP layer. By streamlining these layers, IP becomes aware of the underlying resources and can make better use of them, such as using protection fibers for best-effort traffic. Cost efficiencies are also derived from this streamlining process: fewer layers means less equipment and a more unified management, just to name a few.

6. References

[1] Access Networks, "Last Mile: Die Ankoppelung an den Information-Highway", 1999, http://www.accessnetworks.ch/home.thtml/access/dsl

[2] Jon Anderson et al., "Protocols and Architectures for IP Optical Networking", Bell Labs Technical Journal, January-March 1999, http://www.lucent.com/minds/techjournal/common/arc_issues.html

[3] Bill St. Arnaud, "Wide Area and Long Haul Gigabit Ethernet: The LAN is invading the WAN", CANARIE, http://www.canet3.net/library/papers/wideandlonggigabit.html

[4] ASP Industry Consortium, November 2000, http://www.aspindustry.org

[5] S. Blake, D. Black, M. Carlson, E. Davies, Z. Wang, W. Weiss, "An Architecture for Differentiated Services", IETF RFC 2475, December 1998.

[6] R. Braden, L. Zhang, S. Berson, S. Herzog, S. Jamin, "Resource ReSerVation Protocol (RSVP) - Version 1 Functional Specification", IETF RFC 2205, September1997.

[7] Cisco, "Whitepaper: Scaling Optical Data Networks with Cisco Wavelength Routing", 1999, http://www.ieng.com/warp/public/cc/so/neso/olso/wartso/opdn_wp.htm

[8] Joel Conover, "No Competition Among Local Providers", Network Computing , May 15, 2000, http://www.networkcomputing.com/1109/1109f2full.html

[9] Robert X. Cringely, "Through an ILEC Darkly: How DSL Works and Might Even Make Us Rich", September 14, 2000, http://www.pbs.org/cringely/pulpit/pulpit20000914.html

[10] Alex Galis, "Multi-Domain Communication Management System", CRC Press, 2000.

[11] N. Ghani, S. Dixit, T.-S. Wang, "On IP over-WDM Integration", IEEE Communications Magazine. Vol. 38, No. 3, March 2000, pp. 72-84.

[12] L. Kahn, "A New Medium for IP; Telecommunications", International Edition journal, Vol. 32, No. 7, July 1999, pp. 58-62.

[13] J. Lundquist, B. Svensson, "Messaging-over-IP – A Network for Messaging and information services", Ericsson Review, No. 3, 1999, pp. 142-147.

[14] A. Malis, W. Simpson, "PPP over SONET/SDH", IETF RFC 2615, June 1999.

[15] J. Manchester, J. Anderson, B. Doshi, S. Dravida, "IP over SONET", IEEE Communications Magazine, Vol. 36, No. 5, May 1998, pp. 136-142.

[16] Martin W. Murhammer, "Virtual Private Networks (VPN)", IBM International Technical Support Organization, ITSO Workshop, Raleigh, NC, 1998.

[17] Henning Schulzrinne. “Scaling the Internet”, IFIP WG 7.3 Performance Evaluation workshop, Keynote Address, Eastsound (Orcas Islands), WA, June 20, 1997, http://www.cs.columbia.edu/~hgs/research/talks.html

[18] Siemens, “IP Over DWDM: New Business for City Carriers”, TransXpress Waveline product information, 2000.

[19] David Sovie and John Hanson (Mercer Management Consulting), “Application Service Providers”, November 2000, http://www.mercermc.com

[20] Stardust.com, “IP QoS FAQ”, http://www.qosforum.com/docs/faq/, September 5, 1999.

[21] Stardust.com, “QoS Protocols & Architectures”, http://www.stardust.com/qos/whitepapers/protocols.htm, August 1999.

[22] Michael Tesler (Broadsoft), “Next Generation Services Requirements and Opportunities”, NGN 99 Conference Panel on “Transforming Telephony: Value-Added Applications”.

[23] J. Wroclawski, “The Use of RSVP with IETF Integrated Services”, IETF RFC 2210, September 1997.

[24] A. Zerdick et al, “Die Internet-Ökonomie – Strategien für die digitale Wirtschaft”, European Communication Council Report, Springer, Heidelberg, 1999, Section 3.2: Herrscher der Netze: Wer besitzt die Infrastruktur? Analyse des Telekommunikations-Sektors, pp. 61-99.