Is IP going to take over the world (of communications)?

Paper # 1730 Total number of pages: 9

***Abstract*—While it is technically pleasing to believe that IP will dominate all forms of communication, our delight in its elegance is making us overlook its shortcomings. IP is an excellent means to exchange data, which explains its success. It remains ill-suited as a means to provide many other types of service; and is too crude to form the transport infrastructure in its own right. To allow the continued success of IP, we must be open-minded to it living alongside, and merging with, other techniques (such as circuit switching) and protocols that are optimized to different needs.**

I. INTRODUCTION

Whatever the initial goals of the Internet, there are two main characteristics that seem to account for its success: Reachability and Heterogeneity. IP provides a simple, single, global address to reach every host, enables unfettered access between all hosts, and adapts the topology to restore reachability when links and routers fail. IP hides heterogeneity in the sense that it provides a single, simple service abstraction that is largely independent of the physical links over which it runs. As a result, IP provides service to a huge variety of applications and operates over extremely diverse link technologies.

The growth and success of IP has given rise to some widely held assumptions amongst researchers, the networking industry and the public at large. One common assumption is that it is only a matter of time before IP becomes the sole global communication infrastructure, dwarfing and eventually displacing existing communication infrastructures such as telephone, cable and TV networks. IP is already universally used for data networking in wired networks, and is being rapidly adopted for data communications in wireless and mobile networks. IP is increasingly used for both local and long-distance voice communications, and it is technically feasible for packet-switched IP to replace SONET/SDH.

A related assumption is that IP Routers (based on packet-switching and datagram routing) will become the most important, or perhaps only, type of switching device inside the network. This is based on our collective belief that packet-switching is inherently superior to circuit switching because of the efficiencies of statistical multiplexing, and the ability of IP to route around failures. It is widely assumed that IP is simpler than circuit switching, and should be more economical to deploy and manage. And with continued advances in the underlying technology, we will no doubt see faster and faster links and routers throughout the Internet infrastructure. It is also widely assumed that IP will become the common convergence layer for all communication infrastructures. All communication services will be built on top of IP technology. In addition to information retrieval, we will stream video and audio, place phone calls, hold video-conferences, teach classes, and perform surgery.

On the face of it, these assumptions are quite reasonable. Technically, IP is flexible enough to support all communication needs, from best-effort to real-time. With robust enough routers and routing protocols, and with extensions such as weighted fair queueing, it is possible to build a packet switched, datagram network that can support any type of application, regardless of their requirements.

But for all its strengths, we (the authors) do not believe that IP will displace existing networks; in fact, we believe that many of the assumptions discussed above are not supported by reality, and do not stand up to close scrutiny.

It is the goal of this paper to question the assumption that IP will be *the* network of the future. We will conclude that if we started over - with a clean slate - it is not clear that we would argue for a universal, packet-switched IP network. We believe that in the future, more and more users and applications will demand predictability from the Internet; both in terms of the availability of service, and the timely delivery of data. IP was not optimized to provide either, and so it seems unlikely to displace networks that already provide both. We take the position that while IP will be the network layer of choice for best-effort, non-mission critical and non-real-time data communications (such as information exchange and retrieval), it will live alongside other networks such as circuit-switched networks, that are optimized for high revenue time-sensitive applications that demand timely delivery of data and guaranteed availability of service.

We realize that our position is a controversial one. But regardless of whether or not we are correct, as researchers we need to be prepared to take a step back, to take a hard look at the pros and cons of IP, and its likely future. As a research and education community, we need to start thinking how IP will co-exist with (and possibly control) other networking technologies.

II. IP FOLKLORE

In what follows, we try to identify some folkloric assumptions about IP and the Internet, and examine each in turn. We will start with the most basic, and easiest assumption to dispel: that the Internet *already* dominates global communications. This is not true by any reasonable metric: market size, number of users or the amount of traffic. Of course, this is not to say that the Internet will not grow over time to dominate the global communications infrastructure; after all, the Internet is still in its infancy. It is possible - and widely believed - that packet-switched IP datagrams will become the *de-facto* mechanism for all communications in the future. And so we will move on to consider the assumptions behind this belief and ask if packet-switched IP offers inherent and compelling advantages that will lead to its inevitable and unavoidable dominance. This requires us to examine some "sacred cows" of networking; for example, that packet-switching is more efficient than circuit-switching, that packet-switching is simpler, it lowers the cost of ownership, and it is more robust when there are failures in the network.

A. IP Already Dominates Global Communications

Although the Internet has been a phenomenal success, it is currently only a small fraction of the global communication infrastructure consisting of separate networks for telephones, broadcast TV, cable TV, satellite, radio, public and private data networks, and the Internet. In terms of revenue, the Internet is a relatively small business. The US business and consumer-oriented ISP markets have revenues of $13B each (2000)[1] [20][21], by contrast, the TV broadcast industry has revenues of $36.9B (1997), the cable distribution industry of $35.0B (1997), and the radio broadcast industry $13.5B (1997) [4], [23], and the phone industry $268.5B (1999), of which $111.3B correspond to long distance and $48.5B to wireless [19]. The Internet reaches 60% of US households [24], compared to 94% for telephones and 98% for television. [26], [25]. It is interesting to note that if the revenue per household remains the same, the total revenue for the ISP industry can at most double. If we restrict our focus to the data and telephony infrastructure, the core IP router market still represents a small fraction. As shown in table I the expenditure in core routers worldwide was $2.4B, compared to $44.8B for transport circuit switches. So in terms of market size, revenue and number of users, it is safe to say that the Internet does not currently dominate the global communications infrastructure.

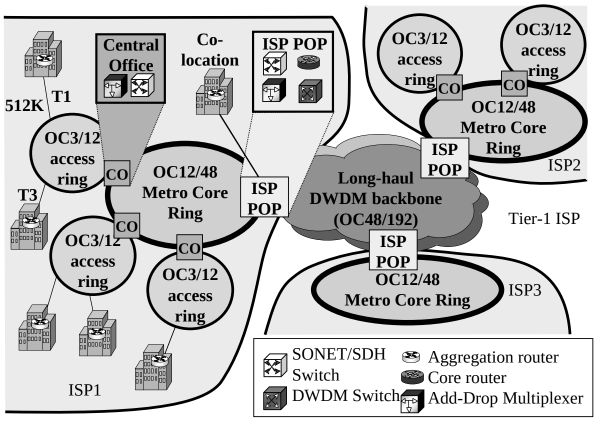

Figure 1. Architecture of the public Internet. There are also many large private voice and data networks that consist of IP routers, LAN switches and voice switches at customer premises.

Segment	Market size
Core routers	$2.4B
ATM	$5.6B
SONET/SDH	$18.8B
WDM	$10B
Telecom OSS	$16B
Class 4/5 Voice switches	$34B
VoIP gateway/softswitches	$0.3B

TABLE I
MARKET BREAKUP FOR THE PUBLIC TELECOMMUNICATIONS INFRASTRUCTURE IN 2001. SOURCE: [22].

Figure 1 illustrates the devices currently used in the public network. The current communication infrastructure consists of a transport network - made of circuit-switched SONET and DWDM devices - on top of which run multiple service networks. The service networks include the voice network (circuit-switched), the IP network (datagram, packet-switched), and the ATM/Frame Relay networks (virtual-circuit-switched). Notice the distinction between the circuit-switched transport network, which is made of SONET and optical switches that switch coarse granularity ($n \times STS-1$, where an STS-1 channel is 51Mb/s), and the voice service circuit switches, which include Class 4 and Class 5 systems that switch 64Kbps voice circuits and handle various telephony-related functions. When considering whether IP has or will take over

---

[1] These numbers probably decreased in 2001, although numbers are not yet available.

the world of communications, we need to consider both the transport and service layers. In other words, for universal packet transport we are considering using a packet network to replace the transport infrastructure; and for voice-over-IP (VoIP) we are considering an application built on top of an IP network that replaces the traditional class 4/5 TDM voice switches.

In what follows, we will be examining the merits of a packet-switched IP network; and to do so, we need to compare it with an alternative. The most obvious alternative is circuit-switching. In one respect this is not an apples-with-apples comparison; the packet-switched IP data network today already operates over a circuit-switched transport infrastructure. If we consider only the core of the network, there is essentially a central core of circuit-switching surrounded by IP routers. It helps to think of the comparison as a question as to which of two outcomes is more likely: Will the packet-switched IP network grow to dominate and displace the circuit switched transport network; or will the (possibly optical) circuit-switched transport network grow to displace the IP routers?

B. IP is more efficient

> "Analysts say [packet switched networks] can carry 6 to 10 times the traffic of traditional circuit-switched networks" – **Business Week**.

From the early days of computer networking, it has been well known that packet switching makes efficient use of scarce link bandwidth [28]. With packet switching, statistical multiplexing allows link bandwidth to be shared by all users, and work-conserving link sharing policies (such as FCFS and WFQ) ensure that a link is always busy when packets are queued-up waiting to use it. If instead the Internet had been based on circuit switching, with each application flow assigned to its own channel, a channel could go idle even if other flows are blocked from accessing the network. Packet switching (and thus IP) makes more efficient use of the bandwidth than circuit switching would, which was particularly important in the early days of the Internet when long haul links were slow, congested and expensive.

It is worth asking: What is the current utilization of the Internet, and how much does efficiency matter today? In [5], [6] Odlyzko and Coffmann report that the average link utilization in links in the core of the Internet is between 3% and 20%, (compared to 33% average link utilization in long-distance lines in the phone network). The reasons that they give for low utilization are threefold; First, Internet traffic is extremely asymmetric and bursty, but links are symmetric and of fixed capacity, second it is difficult to predict traffic growth in a link, so operators tend to add bandwidth aggressively, and finally with falling prices for coarser bandwidth granularity as faster technology appears it is more economical to add capacity in large increments.

There are other reasons to keep network utilization low. When congested, a packet-switched network performs badly, becomes unstable, and can experience oscillations and synchronization. Many factors contribute to this. Complex and dynamic interaction of traffic means that congestion in one part of the network will spread to other parts. And because control packets (such as routing packets) are transmitted *in-band* in the Internet they are more likely to get lost when the data-path is congested. When routing protocol packets are lost due to network congestion or control processor overload, it causes inconsistent routing state, and may result in traffic loops, black-holes and disconnected regions of the network, which further exacerbates congestion in the data path. Today, the most effective way for network providers to address these problems is by preventing congestion and keeping network utilization low.

But perhaps the biggest reason that network providers overprovision their network is to give low packet delay. Users want predictable behavior, which means low queueing delay. We already demand, and are willing to pay for, huge over-provisioning of Ethernet networks (the average utilization of an Ethernet network today is about 1% [6]) just so that bandwidth is available to us whenever we want it, so that our packets can pass through without queueing delay, and so we do not have to share the network with others. We will demand the same behavior from the Internet as a whole.

Therefore, even though conceptually a statistical multiplexed *link* can potentially yield a higher network utilization and throughput, in practice, to maintain a consistent performance and reasonably stable *network*, the network operators significantly over-provision their network, and keep the network utilization low.

But simply reducing the *average* link utilization will not be enough to make users happy. For a typical user to experience low utilization, the *variance* of the network utilization needs to be low, too. There are two flavors of variance that affects the perceived utilization: variance in time (short-term increases in congestion during busy times of the day), and variance by location (while most links are idle, a small number are heavily congested). If we pick a user at random and consider the network utilization their traffic experiences, our sample is biased in favor of users who find the network to be heavily congested. This explains why, as users, we know the average utilization to be low, but find that we often experience long queueing delays.

Reducing variations in link utilization is hard. Without sound traffic management and traffic engineering, the performance, predictability and stability of large IP networks deteriorate rapidly as load increases. Today, we lack effective techniques to reduce the unpredictability of performance introduced by variations in link utilization. It might be argued that the problem will be solved by research efforts on traffic management and congestion control (to control and reduce variations in time), as well as work on traffic engineering and multipath routing (to load-balance traffic over a number of paths). But to-date, despite these problems being understood for many years, effective measures are yet to be introduced.

We can expect that over time users will demand lower and lower queueing delay in the Internet. This means that as users, we collectively want network providers to stop using statistical multiplexing, and to instead overprovision their networks, *as if it were circuit switched*. [34], [35]. To date, network providers have responded to our demands by over-provisioning, by publishing delay measurements for their network, and by competing on the basis of these numbers. In the long term, the demand for lower delay will drive providers to decrease link utilization even lower than it is today, and network utilization will continue to decrease as the world economy becomes more dependent on the Internet.

We can take the demand for low delay one step further, and ask whether users experience the lowest response times in a packet switched network. Here, we define the user response time to be the time from when a user opens an HTTP-over-TCP connection, until a requested file finishes downloading. HTTP downloads represent 65% of Internet usage today (as a fraction of the total number of bytes transferred), and so is representative of typical user behavior. Now consider two types of network: One is the current packet-switched network in which packets share links and each flow makes constant, albeit slow, forward progress over congested links. The other network is a hypothetical comparison. Each new application flow triggers the creation of a low bandwidth circuit in the core of the network, similar to what happens in the phone network. If there are no circuits available, the flow is blocked until a channel is free. If all flows have the same duration, the authors show in [30] that the average user response time in the circuit switched network is half that of the packet switched network. At the core of the network, where the rate of a single flow is limited by the data-rate of its access link, simulations and analysis suggest that the average user response time of both techniques is the same.

In summary, we have observed that while packet switching can lead to more efficient link utilization, unpredictable queueing delays force network operators to operate their networks very inefficiently. We conclude that while efficiency was once a critical factor, it is so outweighed by our need for predictability, stability, immediate access and low delay that network operators will be forced to run their networks very inefficiently. Network operators have already concluded this; they know that their customers care more about predictability than efficiency, and we know from the dynamics of queueing networks, that in order to achieve predictable behavior, network operators must continue to utilize their links very lightly. As a result, they are paying for the extra complexity of processing every packet in routers, without the benefits of increased efficiency. We can conclude that the original goal of “efficient usage of expensive and congested links” is no longer valid, and would provide no benefit to users.

C. IP is robust

> “The internet was born during the cold war 30 years ago. The US Department of Defence [decided] to explore the possibility of a communication network that could survive a nuclear attack.” – **BBC**

The Internet was designed to withstand a catastrophic event where a large number of links and routers were destroyed. This goal is in line with users and businesses who rely more and more on the network connectivity for their activities and operations, and who want the network to be available at all times. Much has been claimed about the reliability of the current Internet, and it is widely believed to be inherently more robust and capable of withstanding failures of different network elements. Its robustness comes from using soft-state routing information; upon a link or router failure it can quickly update the tables and route packets around the failed element. By contrast, a circuit switched network needs to reroute all affected active circuits, which can be a large task for a high-speed link carrying hundreds or thousands of circuits.

The reliability of the current Internet has been studied by Labovitz et al. [2]. They have studied different ISPs over several months, and report a median network availability equivalent to a downtime of 471 min/year. By contrast Kuhn [1] found that the average downtime in phone networks is less than 5 min/year. As users we have all experienced network down-time when our link is unavailable, or some part of the network is unreachable. On occasions, connectivity is lost for long periods while routers reconfigure their tables and converge to a new topology. Labovitz et al. [3] observed that the Internet recovers slowly, with an average BGP convergence time of 3 minutes, and frequently takes over 15 minutes. On the other hand, SONET/SDH rings, through the use of pre-computed backup paths, are required to recover in less than 50ms; a glitch that is barely noticeable in a network connection or phone conversation.

While it may be argued that the instability and unreliability of the Internet can be attributed to its rapid growth and the the ad-hoc and distributed way that it has grown, a more likely explanation is that it is fundamentally more difficult to achieve robustness and stability in packet networks than circuit networks. In particular, since routers/switches need to maintain distributed routing state, there is always the possibility that the state may become disconnected. In packet networks, inconsistent routing state can generate traffic loops and blackholes, and disrupt the operation of the network. In contract, inconsistent routing state will result in blocked calls in a circuit network. While it is not desirable to have new requests rejected, the positive aspect is that none of the established circuits are affected — i.e., inconsistent routing state is not *service-impacting* in circuit networks. In addition, as discussed in Section B, the likelihood of a network getting into a inconsistent routing state is much higher in IP networks because (a) the routing packets are transmitted in-band, and therefore are more likely to incur congestion due to high load of user traffic; (b) the routing computation in IP networks is very complex, therefore, it is more likely for the control processor to be overloaded; (c) the probability of mis-configuring a router is high. And misconfiguration of even a single router may cause instability of a large portion of the network. It is surprising is that we have continued to use routing protocols that allow one badly behaved router to make the whole network inoperable. In contrast, high availability has always been a government-mandated requirement for the telephone network, and so steps have been taken to ensure that it is an extremely robust infrastructure. In circuit networks control messages are usually transmitted over a separate channel or network. This has the added advantage of security for network control and management. In addition, the routing in circuit networks is much simpler.

On the face of it, it seems that packet-switched IP networks experience more failures, and take longer to reestablish connectivity. However, it is not clear that reliability and fault tolerance are a direct consequence of our choice of packet-switching or circuit-switching. We can attribute a lot of the growth of the Internet to the ad-hoc and distributed way that it has grown; so it should not be surprising that there are frequent misconfigurations of routers, and poorly maintained equipment.

The key point here is that there is nothing inherently unreliable about circuit-switching, and we have an existence proof that it is both possible and economically viable to build a robust circuit-switched infrastructure, that is able to quickly reconfigure around failures. There is no evidence yet that we can define and implement the dynamic routing protocols to make the packet-switched Internet as robust. Perhaps the problems with BGP will be fixed over time and the Internet will become more reliable. But it is a mistake to believe that packet-switching is inherently more robust. In fact, the opposite may be true.

D. IP is simpler

> "IP-only networks are much easier and simpler to manage, leading to improved economics." – **Business Communications Review**

It is an oft-stated principle of the Internet that the complexity belongs at the end-points, so as to keep the routers simple and streamlined. While the general abstraction and protocol specification are simple, implementing a high performance router and operating an IP network are extremely challenging tasks. In terms of router complexity, while the general belief in the academic community is that it takes 10's of instructions of process an IP packet, the reality is that the complexities of a high performance router has to as much to do with the forwarding engine as the routing protocols (BGP, ISIS, OSPF etc), where all the intelligence of the IP layer resides, and the interactions between the routing protocols and forwarding engine. A high performance router is extremely complex, particularly as the line rates increase. One subjective measure of the complexity is the failure rate of the start-ups in this space. Because of the perceived high growth of the market, a large number of well-financed start-ups with very capable talents and strong backing from carriers have attempted to build high performance routers. Almost all have failed or are in the process of failing— putting aside the business/market related issues, none have succeeded technically and delivered a product-quality core router. The core router market is still dominated by two vendors, and many of the architects from one came from the other. The bottom line is that building a core router is far from simple, mastered by only a very small group of people.

If we are looking for simplicity, we can do well to look at how circuit-switched transport switches are built. First, the software is simpler. The software running in a typical transport switch [11] is based on about three million lines of source code, whereas Cisco's Internet Operating System (IOS) is based on eight million [8], almost three times as much. Routers have a reputation for being unreliable, crashing frequently and taking a long time to restart.

So much so that router vendors frequently compete on the reliability of their software, pointing out the unreliability of their competitor's software as a marketing tactic. Even a 5ESS telephone switch from Lucent, with its myriad of features for call establishment and billing, has only about twice the number of lines of code as a core router [9], [10].

The hardware in the forwarding path of a circuit switch is also simpler than that of a router. At the very least, the line card of a router must process the packet header, find the longest-matching prefix that matches the destination address, generate ICMP error messages for expired TTLs, process optional headers and then buffer the packet (a buffer typically holds 250ms of packet data). If multiple service levels are added (for example, differentiated services [17]), then multiple queues must be maintained, as well as an output link scheduling mechanism. In a router that performs access control, packets must be classified to determine whether or not they should be forwarded. And in a router that supports virtual private networks, there are often different forwarding tables for each customer.

On the other hand, the linecard of an electronic transport switch typically contains a SONET framer to interface to the external line, a chip to map ingress time slots to egress time slots, and an interface to a switch fabric. Essentially, one can build a transport linecard by starting with a router linecard and then removing most of the functionality [32], [33].

One measure of this complexity is the number of logic gates implemented in the linecard of a router. An OC192c POS linecard today contains about 30 million gates in ASICs, plus at least one CPU, 300Mbytes of packet buffers, 2Mbytes of forwarding table, and 10Mbytes of other state memory. The trend in routers has been to put more and more functionality on the forwarding path: first, support for multicast (which is rarely used), and now support for quality of service, access control, security and VPNs. (And we thought that all the complexity was in the end system!). On the other hand, the linecard of a typical transport switch contains a quarter of the number of gates, no CPU, no packet buffer, no forwarding table and an on-chip state memory (included in the gate count). Because they use simpler hardware, electronic circuit switches consume less power, allowing more capacity to be placed in a single rack. It should come as no surprise that the highest capacity commercial transport switches have about four times the capacity of an IP router [13], [12], and sell for about one third as much per gigabit per second.

One might argue that the reason the circuit switches cost less is that they solve a simpler problem. Instead of being aware of individual application flows, they deal with large trunk lines in multiples of 51Mb/s. So for the sake of comparison, it is worth considering the cost and complexity of building a core transport switch that could establish a new circuit for each (TCP) application flow. Let's assume that each user connects to the network via a 56Kb/s modem, and so this will define the granularity of the switch. While such a small circuit might not be the best way to incorporate circuit switching into the Internet, using such a small flow granularity provides an upper bound on the complexity of doing so. A 10Gb/s linecard needs to manage at most 200,000 56Kb/s circuits. The state required to maintain the circuits, and the algorithms needed to quickly establish and remove circuits, would occupy only a fraction of one ASIC. This suggests that the hardware complexity of a circuit switch will always be lower than the complexity of the corresponding router.

It is interesting to explore how optical technology will affect the performance of packet switches and circuit switches. In recent years, there has been a lot of discussion about all-optical Internet routers. There are two reasons why this does not make sense. First, a router is a packet switch and so inherently requires large buffers to hold packets during times of congestion, and there are currently no economically feasible ways to buffer large numbers of packets optically. The buffers need to be large because TCP's congestion control algorithms require at least one bandwidth-delay product of buffering to perform well. For a 10Gb/s link and a round-trip time of 250ms, this corresponds to 2.5Gbits of storage, which is a large amount of electronic buffering, and (currently) an unthinkable amount of optical buffering. The second reason that all-optical routers do not make sense is that an Internet router must perform an address lookup for each arriving packet. Neither the size of the routing table, nor the nature of the lookup, lends itself to implementation using optics. For example, a router at the core of the Internet today must hold over 100,000 entries, and must search the table to find the longest matching prefix - a non-trivial operation. There are currently no known ways to do this optically.

Optical switching technology is much better suited to circuit switches. Devices such as tunable lasers, MEMs switches, fiber amplifiers and DWDM multiplexors provide the technology to build extremely high capacity, low power circuit switches that are well beyond the capacities possible in electronic routers.

The perceived simplicity of IP comes from the fact that the it has a relatively low initial (fixed) cost of communication. If one were to build a box that should only forward one byte of information during its lifetime, packet switching would be simpler than circuit switching, since we do not need to have to deal with any signaling, state management, resource reservation or crossbar programming. On the other hand packet switching has a high variable cost, i.e. it costs more to send the n-th piece of information using packets than circuits, because circuits take advantage of prescheduled arrivals of information to eliminate operations to be performed on the incoming information, there is no need to buffer information, detect loops in the routes, decide where to go next, etc. When one has to move a lot of data around, it is the variable cost what dominates the overall complexity. So even if packet switching might be simpler for low data rates, it becomes more complex for high data rates, and that is why it cannot achieve the same data rates as circuit switching. In other words IP's "simplicity" does not scale.

E. Support of Telephony and Other Real-time Applications Over IP Networks

> "All critical elements now exist for implementing a QoS-enabled IP network." – **starting sentence on an IEEE Communications Magazine article**

There is a widely-held assumption that IP network can support telephony and other real-time applications. If we look more closely, we find that the reasons for such an optimistic assumption are quite diverse. One school holds the view that IP network is ready today. There are two reasons for such a belief. First, IP networks are and will continue to be heavily over-provisioned, and the average packet delay in the network will be low enough to satisfy the real-time requirements of these applications. Second, most interesting real-time applications including telephony are *soft* real-time in the sense that they can tolerate occasional packet delay/loss and *adapt* to these network variabilities. While today's IP networks are heavily over-provisioned, it is doubtful whether a new solution (far from complete yet) that provides a worse performance can displace the reliable and high quality service provided by today's TDM-based infrastructure (which is already paid-for).

Another school believes that for IP to succeed, it is critical for IP to provide QoS with the same guarantees as TDM but with more flexibility. In addition, the belief is that there is no fundamental technical barrier to build a connection-oriented service (Tenet and IntServ) and to provide guaranteed services in the Internet. The technical ingredients for a complete solution include efficient packet classification and scheduling algorithms. Unfortunately, after more than 10 years of extensive research and efforts in the standards bodies, the prospect of end-to-end per-flow QoS in the Internet is nowhere in sight. The difficulty seems to be the fact that there is huge culture gap between the connection and datagram design communities. By blaming the failure on "connections", a third school holds the view that a simpler QoS mechanism such as DiffServ is the right way to go. Again, we are several years into the process, and it is not at all clear that the "fussy" QoS provided by DiffServ (with no route pinning support and no per flow QoS scheduling) will be good enough for customers who are used to the excellent QoS provided by the existing circuit-switched transport networks

III. DISCUSSION

Up until this point, we have considered some of the folklore surrounding the packet-switched Internet. Our overall goal is to provoke discussion and research on fundamental issues that need to be addressed so that IP can continue to revolutionise the world of communications. We hope to provide a vantage point for the IP community to reflect upon the problems that still need to be solved. As a research community we need to think beyond the daily challenges of maintaining and optimizing the expanding Internet, and move on to consider the enormous challenges that lie ahead.

It seems that there are two main limitations to the widespread adoption of IP: Dependability, and the right way for it to co-exist with circuits. In what follows, we will discuss each in turn.

A. Dependability of IP Networks

High dependability, in the broadest sense, is a must if IP is to become a successful transport technology (to compete or displace circuit-based transport networks), and if the Internet is to become the universal infrastructure for high value applications. For example, voice services are a high-revenue, and very profitable business. Trusting them to today's unreliable, and unpredictable IP networks would be an un-necessary risk, which is why — despite predictions to the contrary — telephone carriers have not done so.

High dependability means several things: robustness and stability, traffic isolation, traffic engineering, fault isolation, managability, and last but not least, the ability to provide predictable performance in terms of bounded delay and guaranteed bandwidth (QoS). In its current form, the Internet excels in none of these areas. Although it is clearly a challenge to achieve each of these goals, they must all be solved for IP to become dependable enough for use as a transport mechanism.

B. How should IP interact with circuits

The curernt Internet is based on packet switched routers, interconnected by a circuit switched transport network. Given the benefits of circuit switching, it would seem perverse for the packet switched network to grow to subsume the transport network. It is inconceivable to us that the network providers would remove the existing, robust, reliable, predictable and larglely paid-for transport network, and replace it with a technology that seems more complex, less reliable, more expensive and not yet installed.

What seems more likely is that packet switching will continue to exist at the edge of the network, aggregating and multiplexing traffic from heterogeneous sources for applications that have no delay or quality requirements. In other words, packet switched IP will continue to provide a simple service abstraction for a variety of applications.

At the core of the network, we expect the circuit switched transport network to remain as a means to interconnect the packet switched routers, and as a means to provide high reliability, and performance guarantees. Over time, more and more optical technology will be introduced into the transport network, leading to capacities that (necessarily) electronic routers cannot achieve.

One remaining question is whether or not the circuit switched network will be controlled by IP. In other words, will the IP network decide dynamically when to create new circuits between routers? For example, a router could monitor the occupancy of its queues and periodically add or remove circuits to other routers based on current demand [29], [30], [31]. Such a system has the benefit of enabling IP to gain the benefits of fast optical circuit switches in the core, yet maintain the simple service model for heterogenous sources at the edge.

However, while this approach seems appealing to IP, we need to remember that the majority of the revenue for the circuit switches will still be from other applications, such as voice. Because the packet switched network is unlikely to provide the predictability needed for voice traffic, it will continue to operate over its own, separate circuit switched edge network and be carried over the shared transport network at the core. In this environment, it is unlikely that the routers will be allowed to control the entire capacity of the transport switches, unless the revenue for the Internet significantly exceeds that of telephony. In 2001 only 6% of all international voice traffic used VoIP [18], and the fraction does not seem likely to increase much in the near future. This leads us to believe that it is more likely that the routers will be allocated a fraction of the circuit switched transport infrastructure, which they can control and adapt to best serve their needs.

With the dynamic control of circuit-networks (possibly by an IP-based control plane), it is also conceivable that the IP routers at the edge can signal to the transport network to dynamically create new circuits or change bandwidth of existing circuits

C. What if we started with a clean slate

In the preceding discussion, we predicted an outcome based on historical reasons, and in the context of a pre-existing circuit switched transport network. So if we started again, with the benefit of hindsight, would we build a network with circuit switching at the core, and packet switching at the edge? We believe that we would, and that it would look something like this:

- **Addressing scheme.** A simple, unique and universal addressing scheme (like IP's) would allow us to communicate with any sort of device or application anywhere in the world. This addressing scheme defines the routing algorithms in the intermediate network nodes, but it is completely independent of the forwarding or switching mechanisms that they use.
- **Switching in the edges of the network.** Packet switching would be used in the edges of the network as well as those links where bandwidth is scarce (like some satellite links and underwater cables). The reasons for this are threefold. First, packet switching makes a very efficient use of the bandwidth in these cases. Second, it can greatly improve the end-user response time by borrowing all available link bandwidth when other users are not active. Finally, packet switches can be simpler for lower link rates. The packet-switched network should ideally gather traffic from disparate sources, and multiplex it together in preparation for carriage over a very high capacity, central, circuit switched core. In this environment, local switching at the edge of the network is an optimization that may or may not be necessary. Without it, the packet switched network is simply a hierarchy of statistical multiplexers, with little or no forwarding decisions necessary. All traffic can be multiplexed towards the core, then demultiplexed again towards the edge. While less efficient, it provides a simplified environment in which to deploy the delay guarantees needed by telephony. And so it might be feasible to carry the traffic from access voice switches to the core over the statistically multiplexed edge network.
- **Switching in the core of the network.** At the core of the network, there seem a number of compelling reasons to use circuit switching. First, circuit switching has already demonstrated its robustness, and its ability to quickly recover from failures. Circuit switching is inherently simpler than packet switching, requiring less work to forward data, and so will cost less as a result, will consume less power, and will take up less space. Last, though probably first, circuit switch-

ing provides an easy way to adopt the huge potential of high capacity optical switches. Without electronics on the forwarding path, we can expect optical switches to provide abundant capacity at low cost.

- **Integration of both switching mechanisms.** Rather than working independently, both of these mechanisms would be tightly integrated, in such a way that a change in behavior in one provokes a change in behavior in the other. For example, packet switching would have to export the QoS and connection oriented nature of the circuit switched core to the applications that require it. On the other hand circuit switching has to respond to the increases in activity of packet switching, by adapting its capacity among core/edge gateways accordingly.

REFERENCES

[1] R. Kuhn, Sources of Failure in the Public Switched Telephone Network. IEEE Computer, Vol. 30, No. 4, April 1997.

[2] C. Labovitz, A. Ahuja, F. Jahanian, Experimental Study of Internet Stability and Wide-Area Network Failures. University of Michigan Technical Report CSE-TR-382-98.

[3] C. Labovitz, A. Ahuja, A. Bose, F. Jahanian, “Delayed Internet Routing Convergence”. IEEE/ACM Transactions On Networking, Vol. 9, No. 3, June 2001.

[4] A. M. Odlyzko. Content is not king. First Monday 6(2), http://firstmonday.org/, Feb 2001.

[5] A. M. Odlyzko. Data networks are mostly empty and for good reason. IT Professional 1 (no. 2), pp. 67-69, Mar/Apr. 1999.

[6] K. G. Coffman and A. M. Odlyzko. Internet growth: Is there a “Moore’s Law” for data traffic?. Handbook of Massive Data Sets, J. Abello, P. M. Pardalos, and M. G. C. Resende, eds., Kluwer, 2001.

[7] B. Wellenius, C. A. Primo Braga and C. Z.-W. Qiang Investment and growth of the information infrastructure: summary results of a global survey. Telecommunications Policy, Pages 639-643, Sep 2000.

[8] Chris Edwards, Panel weighs hardware, software design options, http://www.eetimes.com/story/OEG20000607S0043, June 8, 2000.

[9] The Code Decay Project, University of Maryland, http://www.cs.umd.edu/%7Eaporter/html/evolution.html

[10] S. G. Eick, P. Schuster, A. Mockus, T. L. Graves and A. F. Karr. Visualizing Software Changes. National Institute of Statistical Sciences, Technical Report 113, Dec 2000

[11] Private Communication

[12] Cisco Systems. Cisco 12416 Internet Router: Data Sheet. http://www.cisco.com/warp/public/cc/pd/rt/12000/12416/prodlit/itro_ds.htm

[13] Ciena. CIENA MultiWave CoreDirector. http://www.ciena.com/downloads/products/coredirector.pdf

[14] V. Fineberg. A Practical Architecture for Implementing End-to-End QoS in an IP Network. IEEE Communications Magazine, pp 122-130, Jan 2002.

[15] J. Postel. Internet Protocol. RFC 791, Sep 1981.

[16] F. Baker. Requirements for IP Version 4 Routers. RFC 1812, June 1995.

[17] S. Blake, D. Black, M. Carlson, E. Davies, Z. Wang, and W. Weiss, “An Architecture for Differentiated Services”, Internet Engineering Task Force, RFC 2475, http://www.ietf.org/rfc/rfc2475.txt

[18] G. A. Chidi,”VoIP taking 6 percent of international calls”, ITworld.com, http://www.itworld.com/Net/3303/IDG011106VoIPvolume/, Nov 2001.

[19] Industry Analysis Division, Common Carriers Bureau, Federal Communications Commission, Trends in Telephone Service, Aug 2001, http://www.fcc.gov/Bureaus/Common_Carrier/Reports/FCC-State_Link/IAD/trend801.pdf

[20] Cahners, In-Stat Group. 2001 Business ISPs Service, Size, and Share. Advanced Carrier Business Report, http://www.instat.com/abstracts/tx/2001/tx0109sp_abs.htm, Oct 2001.

[21] Cahners, In-Stat Group. Information Alert Newsletter. Volume #23, http://www.instat.com/infoalert/2001/alrt2001-23.htm, July 30, 2001

[22] T. Luke, S. Levy, A. Green, “2001 Guide To Communication Networks & Equipments”, Lehman Brothers, 2000.

[23] US Census, Industry Quick Report, http://www.census.gov/servlet/IRQBrowseServlet

[24] Nua Internet Surveys, “How Many On-line?”, http://www.nua.ie/surveys/how_many_online/n_america.html, Aug 2001.

[25] A.L. Penenberg, ”The war for the poor”, Forbes Magazine, http://www.forbes.com/1997/09/26/feat.html, Sept 1997.

[26] National Telecommunications and Information Administration, “Falling through the Net: Defining the Digital Divide”, http://www.ntia.doc.gov/ntiahome/fttn99/, 1999.

[27] AT&T Press Release. Ianna Outlines Plan to Evolve the AT&T Network. http://www.att.com/technology/ip/iannaplan.html, Mar 2, 1999.

[28] Paul Baran, “On Distributed Communications: I. Introduction to Distributed Communications Network”, Rand Corporation, Memorandum RM-3420-PR, http://www.rand.org/publications/RM/RM3420/, Aug 1964.

[29] A. Banerjee, J. Drake, J. Lang, B. Turner, D. Awduche, L. Berger, K. Kompella, Y. Rekhter, “Generalized Multiprotocol Label Switching: An Overview of Signaling Enhancements and Recovery Techniques”, IEEE Communications Magazine, vol. 39(1), pp. 144-150, Jan 2001.

[30] P. Molinero-Fernández, N. McKeown, “TCP Switching: Exposing Circuits to IP”, IEEE Micro magazine, vol. 22(1), pp. 82-89, Jan/Feb 2002.

[31] M. Veeraraghavan, M. Karol, R. Karri, R. Grobler, and T. Moors. "Architectures and Protocols that Enable New Applications on Optical Networks". IEEE Communications Magazine, vol. 39(3), pp. 118-127, Mar 2001.

[32] PMC-Sierra, “Diagram of a 10 Gigabit Core Router Architecture”, http://www.pmc-sierra.com/products/diagrams/CoreRouter_lg.html

[33] PMC-Sierra, “Diagram of a Sub-wavelength Optical Cross Connect”, http://www.pmc-sierra.com/products/diagrams/ SubWavelengthCrossConnect_lg.html

[34] Matrix.net, “Internet Ratings”, http://ratings.miq.net/

[35] The Internet Traffic Report, http://www.internettrafficreport.com/