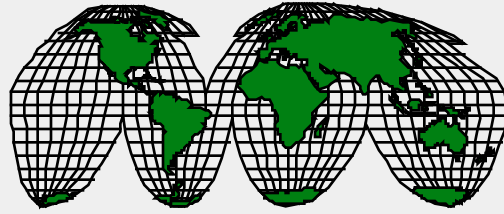


The COOK
Report on



Internet

ATM: Grand Unifying Technology or Brain-Damaged Transport Product?

In *COOK Report* Interviews InternetMCI Reveals Total ATM Commitment for Launch Before April 30

Introduction: Design Philosophy May Determine Choice of Network Technology

ATM was designed by the telephone companies as *the* grand unifying transport technology. In order to satisfy the demands of real-time delivery of voice and video, it relies on a steady stream of 53 byte long cells. As a result of its small cell size, its architecture is not compatible with any standard of efficiency in the TCP/IP world. In short ATM is a telephony technology disparaged by many of the designers of the Internet as being an utterly alien environment to the TCP/IP protocol that has become the language of choice for data networks world wide.

Yet, because it is a transport technology, ATM can be used in switches that act like high speed express channels between points on a net and can aggregate sessions at a cheaper cost per port than routers. Some providers are over-laying an ATM switching fabric between backbone hubs from which packets are routed to their final destination. Those who do this claim they are cost effectively punching more traffic through their backbones with ATM than they could without. However, their critics claim that they loose the gains in throughput, either at routing bottle necks, or through what they derisively refer to as the ATM cell tax.

There is another element of ATM that, in view of its other architectural short comings, causes its proponents to push ahead. Namely, unlike TCP/IP, ATM is a connection-oriented transport ser-

vice. That it is to say it must run over either permanent or Switched Virtual Circuits (SVCs) where a slice of network bandwidth is dedicated and assigned to a single route to get from one user to another. Datagrams flowing through a connectionless network are almost impossible to account and bill for. Consequently TCP/IP service is billed by the size of the leased line or pipe through which traffic flows and not by the amount of traffic actually pushed through. SVCs may change this.

ATM permanent virtual circuits are nailed up much like leased lines and without much thought of trying to measure explicit traffic. However, switched virtual circuits are now on the horizon. Twelve to 18 months from now it is anticipated that these circuits can be set up and torn down automatically by software in end user applications. Such applications can range from real-time audio and video conferencing, to high bandwidth interaction with one or more databases or distributed computing sessions done in real time. SVCs will allow reserved bandwidth (also called quality of service) applications to network end users. Such applications, needless to say, can be billed for individually at significantly higher cost than sessions that have to compete with all other network traffic and are not individually identifiable and billable. SVCs are very attractive to the phone companies because they hold out the first practical hope of being able to create different priorities of service and bill by usage for some Internet applications.

NSF Program Promotes Priority of Service

On March 14th the National Science

Vol. V, No. 1, April 1996

© Cook Network Consultants

ISSN: 1071-6327

Foundation released its new meritorious high bandwidth connections solicitation - one specifically designed to promote the development of bandwidth reservation services. According to the NSF program announcement: "The National Science Foundation has introduced a new twist to its connections program: emphasizing innovative solutions that may have broad implications for all Internet users. The program will look for meritorious applications that require high performance networking, and will then fund development by university and college campus network service providers. Technology developed for this program will likely affect future Internet operation.

The technology will introduce the idea of prioritization to Internet traffic. For example, if planning to use the U.S. Postal Service to send a package, you have options: overnight mail, first-class service, or third-class service. The rate

On the Inside:

ATM: CIF & Cell Tax	pp. 1 - 3
Bandwidth too Cheap	pp. 3 - 6
MCI's ATM Plans	pp. 7 - 13
Gartner, Access I, etc	pp. 13
AGIS Problems pt. 2	pp. 14 - 20
Web Standards	pp. 21 - 22
AT&T Internet Debut	pp. 22, 24

of the package delivery is contingent on how it is designated. Freeways around major cities often have either express toll roads or high occupancy-vehicle lanes to bypass congested areas. Similarly, NSF's connections program is expected to spur the development of switches and routers to help alleviate bottlenecks of information.

"There is no single solution. We hope this grant program will stimulate the development of a technological option for the Internet, to introduce prioritization and provide a new style of connection that gives a guaranteed level of service at a national level," said Mark Luker, manager of NSF's connections program. [(703) 306-1950 or mluker@nsf.gov. See also <http://www.cise.nsf.gov/cise/ncric/connect96.html>.]

Currently on the Internet, all packets of information are treated alike. While this worked fine before the popularization of the Internet, it now interferes with some uses that require high performance service. One example is to use high performance connections of multiple small computers to create a large workstation cluster distributed across the nation. The Internet is currently too congested for such a system. Teleconferencing or videoconferencing also places too great a need on the current capacity."

Is Ethernet As the Last Mile Technology a Barrier to ATM?

Wisdom says that SVCs, to be used to their maximum advantage, must be user-controlled and will therefore have to rely on end-to-end ATM transport

Wisdom says that SVCs, to be used to their maximum advantage, must be user-controlled and will therefore have to rely on end-to-end ATM transport right to the user's desk top. Traditional wisdom says getting all this in place looks quite expensive.

Furthermore, because ATM is not compatible with ethernet, it seems to demand the very expensive replacement of the technology of choice for these local sites. Thus it would appear that there is a significant cost barrier to maximum market penetration.

right to the user's desk top. Traditional wisdom says getting all this in place looks quite expensive.

Furthermore, because ATM is not compatible with ethernet, it seems to demand the very expensive replacement of the technology of choice for these local sites. Thus it would appear that there is a significant cost barrier to maximum market penetration.

This was brought home to us by a conversation that we had with Scott Bradner, IETF Co-Area Director for IPng. To Scott we said: Steve Tabaska MCI VP of Network Engineering had told us that he wanted to use PVCs for short and bursty IP traffic while saving SVCs for user initiated calls for reserved bandwidth, also referred to as Quality of Service Demands (QoS). Do you think this is a realistic and efficient way to use bandwidth?

Scott replied: PVC for most & SVC for longer things sounds fine until you start to think about QoS. I'm not sure that such an architecture is realistic unless there is a way to get the QoS controls back to the user & that will be hard to do over the Ethernet plus token ring "last mile" to the desktop To which we said: Steve Tabaska was saying that we had to wait for the QoS stuff to get embedded in the actual applications. Would that be a way around the ethernet difficulty you mentioned?

Scott responded: you can't do Quality of Service on Ethernet, it does not support the function. I see no reason for ATM-dependent QoS to get put in applications unless one assumes that in *all* cases the end-to-end path, server to client is ATM. But this simply will not happen - there is just too much Ethernet and token ring. Fast Ethernet (100 megabits per second) is just too cheap relative to ATM. Furthermore, gigabit Ethernet will be here "soon", for far less than ATM. ATM will be an important technology but it is not *the* network technology.

But What About Cells in Frames?

An NSF-funded project at Cornell called Cells in Frames (CIF) is designed to run ATM over Ethernet. We queried Richard Cogger, the project Principal Investigator, who with Scott Brim is developing a means of using 10BASE-T Ethernet infrastructure to transmit ATM.

Richard replied: I must emphasize: CIF is *real* atm over ethernet hardware. It is *not* emulated ATM, ethernet-to-ATM

One Cells in Frames Attachment Device is placed at each user port on the 10 BASE T ethernet hub. The cost of the device at roughly \$100 per work station must be compared to an estimated cost of five to six hundred dollars necessary to rewire and bring ATM directly to each workstation. Cogger estimates that changing the Cornell University campus to an end-to-end Cells in Frames ATM network will enable the University to effectively merge its data and telephone networks and replace 75 to 90% of its PBX lines with phones that plug into workstations. The goal is to unify the institution's voice, video and data networks into a single ATM network with potentially huge telephony savings.

conversion, etc. You will run the same ATM stack and signaling in a CIF-attached End System that you will run in 25 Megabit atm, 100 M atm, 155 M atm, etc. You will run LANE, Classic IP, MPOA in the end system for legacy applications You will have the same application interfaces (APIs) for applications to access QoS parameters. (Note that the API's are not standardized yet).

Cells in Frames is just: (1) another speed for ATM (with potential upgrade to 100 Megabit ethernet); (2) a technique for re-using existing hardware, cables, etc. (3) a technique for offloading some processing from the host to the CIF-Attachment Device, in a manner analogous to the way that a typical ATM adapter offloads some of the processing from the host. But in order to use the existing ethernet adapter, we have to move this processing to the other end of the wire; (4) a way to be able to afford much sooner/wider deployment of ATM and access to QoS, Traffic-management. This point is the essence of the whole crusade, and the only thing that makes it worth it. As to what the techniques are and how they will be implemented, some of that is available on our website (<http://cif.cornell.edu>) and some is being worked out in ongoing email discussions of the CIF Alliance participants.

Later, in a telephone interview with Richard Cogger we learned that Cells in Frames is not intended for use in a shared-media ethernet LAN environ-

ment. Rather it is designed to re-use part of the equipment typically used in a 10BASE-T installation where each user is separately connected in a spoke and wheel topology to a central hub. The CIF "Attachment Device" replaces the repeater hub, providing a high-speed uplink to an ATM backbone, while the Ethernet card in the user's workstation and wiring is undisturbed. Note that each user now has a separate 10 Mbps path to an ATM switching fabric rather than a share of a 10 Mbps broadcast environment. The cost of the attachment device, targeted at roughly \$100 per work station (discounted), must be compared to an estimated cost of five to six hundred dollars necessary to change out the Ethernet card for, say, a 25Mbps ATM card and provide a 25Mbps ATM switch port at the other end of the wire.

Cogger estimates that changing the Cornell University campus to an end-to-end Cells in Frames ATM network will enable the University to effectively merge its data and telephone networks and replace 75 to 90% of its PBX lines with phones that plug into workstations. The goal is to unify the institution's voice, video and data networks into a single ATM network with potentially huge telephony savings. Sources unrelated to Cogger and Cornell told us that were it not for the Cells in Frames project, they did not think that universities would be able to afford the cost of rewiring necessitated by the NSF's new connections program. Many ATM sceptics have dismissed the viability of ubiquitous ATM SVCs because of the last mile Ethernet hurdle. If Cells in Frames can put enough of a dent into that last mile, ubiquitous ATM to the desktop becomes much more feasible although by no means assured.

The ATM Cell Tax

In the meantime carriers using ATM for switched backbone fabrics will have to deal with architectural issues which, according to critics, rob OC-3 ATM of between 25 and 32% of its total TCP/IP carrying bandwidth.

On March 19 Jim Forster from Cisco explained the "tax" in the following way: It is the overhead resulting from carrying variable length IP packets in fixed length ATM cells. Consequently the last cell of an AAL5 frame will contain 0 - 39(?) bytes of padding, which is wasted bandwidth. Assuming random length distributions (which they're not), the average waste is about 20 bytes per packet.

My rough estimate, based on an average

Internet packet size of 200 bytes (used to be correct, not sure anymore), is that the waste due to cell padding is about 10%. Due to the highly skewed packet size distribution the actual overhead might vary substantially. Note that this 10% is on top of the nearly 10% overhead due to the 5 byte ATM cell headers, and various other overheads (some of which are also present in frame over SONET schemes). There's beginning to be some expectation that there will be a transmission capacity crunch in the carrier's SONET nets, and this approximately 25% ATM cell tax may be looked at carefully as packet over SONET solutions emerge.

Cisco's Paul Ferguson then added: I think it will be critically important, especially in long-haul interconnect scenarios. Historically speaking, as providers have replaced congested links with faster ones, they have also become quickly congested. And since these high-speed long-haul links are not cheap, you can bet that these organizations want to get the most bang for-the-buck. I can't imagine that roughly 25% bandwidth overhead would be acceptable, especially when more efficient alternatives exist. [Editor: PPP over SONET is apparently one such alternative.]

On March 20, Tim Salo of the Minnesota Supercomputer Center wrote "the following numbers are from John Cavanaugh's [Minnesota Supercomputer Center paper]: Line Rate is 155.520 Mbps. But Available to ATM after the SONET payload is removed is 149.760. ATM payload 135.632 available to AAL. John then computes the overhead for three MTUs, and yields rates available to IP and TCP:

MTU	576	9180	65527
Available to IP	125.198	135.102	135.547
Available to TCP	116.504	134.513	135.464

These are the maximum available rates, namely they assume MTU-sized packets. Reader's can apply their favorite packet size distributions to these numbers.

Having said all that, I am not sure where that leaves us. One could theoretically remove the SONET overhead, but then one loses the ability to manage the SONET link. One could remove the ATM overhead, but then one has a point-to-point link, rather than a link over which data from many sources can be multiplexed."

On March 22 Tim Salo, responding to a

question from us, wrote: MTU is the Maximum Transfer Unit, the largest sized packet which can be transferred across a network. For example, the MTU on Ethernets is 1500 bytes. The default across the Internet is 576 bytes. The default for ATM networks is about 9K bytes. [Editor's note: on page 602 of the Rose & Lynch, *The Internet Handbook*, we found that these figures are derived by taking one thousandth of the one second bandwidth of the slowest link on the network.]

Salo concluded: "With respect to the "available to TCP" data, when you are comparing the performance of various solutions, you need some common measurement. We like end-to-end TCP performance because it is what most closely reflects what users see."

We asked a respected member of the internet community who describes himself as neutral to ATM but hostile to its hype: Why would a company like MCI make such a huge investment in ATM on its backbone given the overhead ATM cell tax that cuts 155 mbits down to about 125 to 116 range unless it wanted to be prepared for end to end svcs with meterable salable bandwidth on demand??? Is this where the ATM folk think the pot of gold lies?

The source replied: ATM is currently the best technology to provide for subdividing a fast pipe (T3 or faster), it does this (using PVCs quite well) an overhead that cuts 15% off the usable bandwidth is a whole lot better than having to get a pipe twice as big as you need because you have no way to break it down. I expect that there are even rather good reasons to be able to provide specific bandwidth between specific customer sites - this is rather different than assuming that the various Internet Service Providers will set things up so that bandwidth reservations will pass between them. The ATM people are constantly having to shift the location of the pot of gold cuz they assume there is one even if they have no specific evidence.

We provide this introduction of concerns about the end-to-end viability of SVCs and wasted ATM overhead as a bit of an antidote to the very pro ATM views of Stephen von Rump and Steve Tabaska of MCI that appear later in this issue. Finding out that we had interviewed them after Dave Sincoskie, one source chided us for talking so far only to pro ATM advocates. Asking those folk about ATM is like asking a dyed-in-the-wool Republican whether Dole should be in the White House, he said.

Bandwidth & Resource Reservation as Factors in One's Network Provisioning Philosophy -- Can Bandwidth Ever Be Too Cheap to Meter?

An Interview with Noel Chiappa

Editor's Note: We intend to do several more ATM interviews as we probe the other major carriers for their strategies. In May we shall publish an interview with Sprint which has adopted a non ATM strategy. This article is an interview with Noel Chiappa. It discusses differing philosophies of network design that Noel believes will essentially determine the technologies used. The most basic question that must be asked according to Noel is are bandwidth demands such that networks can be built by over-engineering capacity, or must they allow for resource reservation?

J. Noel Chiappa is currently an independent researcher in the area of computer networks and system software. He has been a member of the TCP/IP technical community since 1977. While a member of the Research Staff at MIT, he worked on packet switching and local area networks, and was responsible for the invention of the multi-protocol router; he later worked with a number of companies to bring networking products based on work done at MIT to the marketplace.

We asked Noel, about the interface between the economics of a viable internet business model, the technology of backbone routers and the presumed need for and interest in TCP/IP bandwidth on demand for audio and video applications.

Resource Reservation Versus over Engineering

Chiappa: The single biggest open question in networking these days is whether provision of new multi-media services, such as voice and video conferencing, is going to require explicit resource reservation in the switches, as some people claim (i.e. either flow support, or virtual circuits, depending on which exact camp people fall into), or whether it's simply possible to build in enough extra capacity so that a pure datagram network can provide these services (sometimes called the "overengi-

neering" camp). It's really hard to say much about what the network of the future will look like without answering this question.

This is an issue where I listen to both sides. Both sides have very smart people who have thought about this for a very long time and they have come to diametrically opposite conclusions and they can't both be right but I'll be damned if I can figure out for sure which side is right. I tend to lean towards the resource reservation crowd but probably not for reasons that are very good. Some of the reasons have something to do with old parts of the infrastructure.

COOK Report: Like?

Chiappa: Like if you are on an ISDN line and you have either 64 or 128 kbs, then you need resource reservation in there if you are going to run both internet packet voice and data over that line. In the copper plant part of the network where major bandwidth increases are not an option and bandwidth is not infinitely expandable, you likely will need resource reservation for activities which require certain amounts of resources.

Now given a blank piece of paper and a whole new infrastructure, the whole question of whether you can actually engineer a system with enough bandwidth in it that you don't need resource reservation seems a bit dubious to me and here is why.

If bandwidth is totally free and unmetered without controls, what stops people from doing things like setting up a video conference between me and my friend Fred's desk and leaving it up all day? I think there has to be some sort of feedback loop between consumption or resources and paying for them.

COOK Report: Are you saying that the dynamics making it necessary for this feedback loop to be defined exists above some certain threshold "x"? We send email with abandon for example. But the bits taken by email are relatively trivial.

Chiappa: Sure. You have to start by asking a whole bunch of questions. For example: looking at the network ten years from now what are the majority of the bits going to be? Video and voice? Or data transactions. You have to characterize your traffic so you know how long average transactions are. If 92% of your traffic is purely digital traffic composed of a couple of packets at a time, resource reservation is a waste. But if 92% of traffic is video and audio transmissions that go on for a long time, then maybe you need resource reservation and the reason why would be, I think, this economic feedback issue.

Here is why. Assume I have a pipe of "n" megabits per second into the infrastructure. Presume it is enough for me to have a couple of video conversations going on where I get to see five or six other people. So you will need a lot of bandwidth. Bandwidth going all the way out to the leaves of your network is going to be extremely high. Now unless there is some sort of feedback mechanism to induce a cost for the use of that channel, then there is no reason for everyone not to use the full bandwidth available all the time. And if all 200 million people in America generated several megabits-per-second, the central core backbones would fall over.

COOK Report: Regardless of whether there is enough fiber to carry it, is the question of being able to route it by any available means problematic?

Chiappa: To me that is not the right question to ask. You have to decide whether you are going to need resource reservation or not. Consider the entire topology of the network from end user to end user and ask yourself whether if everyone had to have essentially unlimited bandwidth you would have to undertake the huge expense of building backbone type infrastructure into all campuses and businesses?

It is possible that there is a way to reduce the bandwidth requirements of the backbone without going to full resource reservation. But I can guarantee you that if we go into a situation where you can send audio and video, and, there is no charge differential depending on

whether you are or you aren't, you are just asking to swamp the backbones.

COOK Report: So Gilder's idea that bandwidth should be free is a utopia that will fail?

Chiappa: Yes. My feeling is over the next 20 to 30 years we are going to have to have some feedback loop between the consumption of bandwidth and charging for it. This is especially so when you can consider the amounts that people using things like video can consume in a very trivial way.

COOK Report: So one major reason for resource reservation is to be able to put in a feedback loop, there by measuring and charging, for the reasons you have indicated?

Chiappa: It might be that there is some other way to control usage. Perhaps there would be the equivalent of an electric meter on the pipe between you and the phone company and they come along every month and read it. That could cause people to moderate their usage without having all the formal infrastructure necessary for resource reservation.

People are beginning to talk about telephony over the internet. It turns out that phone companies have to work very hard to ensure that they meet delay constraints with the voice network. Phone conversations where the end-to-end delay is larger than some threshold level tend to degrade. It is not clear to me that you can meet these constraints a high enough percentage of the time in a network that purely relies on over engineering unless it is grossly over engineered.

Self-similar traffic means very bursty traffic. One characteristic of it is that when you add a bunch self-similar channels together the aggregate of this addition will also be self-similar. Imagine looking at the total network work load and seeing that it is basically a spiky self-similar sort of thing. You can see there are going to be spikes that get clipped off unless you way way over-engineer the network. So the question is: whether way way over engineering the thing to the point where people get an acceptable degree of service most of the time will be cheaper over the long run than adding resource reservation.

There are times when you have to shed load. This is a lot easier to do in a system with resource reservation than without it. Imagine that you have a link that will usefully hold ten voice conversations and twelve people come along and

try to use it. They attempt to split the bandwidth twelve ways. The quality of voice falls below what is acceptable and probably some random number of them simply drop out.

Now you asked the question of what the future is going to look like if we can build switches or routers that operate at this speed? What ever the future is will be greatly influenced by the answer to that question of yours as to "whether you do or do not need resource reservation?" If you decide that you need it, the kind of solutions that you look at start out by looking very different from the solutions that offer themselves in the absence of a resource reservation effort. The answer to these questions will also drastically impact the kind of switching architecture you adopt.

Assuming that you do need reservation, you will find out that there is one class of answer called ATM which is clearly a level 2 fabric that provides resource reservation. But people have some problems with ATM. There is another group of people who think resource reservation is the answer but that ATM is not the means to strive for achieving it. These are the people working on things like RSVP and Int-serve. But there is another whole group of people who reject the idea of resource reservation. They tend to just let the packets fly while over building and over engineering.

The Over Building Approach

It turns out that, based on some of the conversations that I had that what you start to build when you start to get really really high speeds is not stuff that looks like a connectionless IP router any more.

COOK Report: What about the Ipsilon switch?

Chiappa: As far as I understand it the goal for Ipsilon is to build something that looks like an IP router with an order of magnitude increase in the possible speeds. My understanding is that the Ipsilon router is basically built along the no resource reservation model.

Lets assume that you are a backbone provider. What I seem to see people gravitating to is not routers and underlying transmission technology that developed independently, but rather hybrid systems in which the routers are

What ever the future is will be greatly influenced by the answer to that question of yours as to "whether you do or do not need resource reservation?" If you decide that you need it, the kind of solutions that you look at start out by looking very different from the solutions that offer themselves in the absence of a resource reservation effort. The answer to these questions will also drastically impact the kind of switching architecture you adopt.

slightly weird and produce stuff that makes sense only with a certain kind of substrate.

COOK Report: Is it almost the very idea of routing IP directly over SONET? Does it go in that direct?

Chiappa: I honestly don't know what to tell you but the impression I get is that people are no longer building stock off the shelf routers. But that they are rather building almost what amounts to a series of half routers where there is one half the router at the network edge where the customer can see it and the other half interfaces to some weird stuff which interfaces to the first half router on the other end.

COOK Report: And to get it across the network they don't switch it but engage in some sort of black art? Of course the reason that we are talking black art is because what is actually happening is covered by non disclosure agreements?

Chiappa: Correct. It is a technology which is neither pure datagram nor pure switching. As the network speeds up, they have been driven in this direction by the necessity to get what appears to be datagram service but be able to run it at very high speeds. The question is whether over the next ten years of data networking this sort of paradigm will spread out towards the end nodes. Is this the direction networking as a whole is going?

When you look at those people who say you can do it with pure datagrams, you also have to remember that they generally are not applications experts.

COOK Report: So you are saying

that there is a certain kind of art in fitting different kinds of applications into datagrams?

Chiappa: Absolutely. With voice one of the pieces of the art is that you have to have a fairly low and stable round trip end-to-end delay. If you don't have that the voice channel doesn't work anymore. Think in terms of elastic or inelastic. FTP is elastic. You will be satisfied by service over a wide range of bandwidths. Whether the file travels at one kilobit per second or 100 kilobits, it eventually will get there. Voice is inelastic. Above 16 kilobits per second it is pretty much useless in the sense of "the extra bandwidth does you no good" and below a few kilobits per second it is useless too.

COOK Report: We gather there are sophisticated routing technologies under development and that there are people trying to bring them to market as of yesterday.

Chiappa: There are some very very clever people out there and there is a lot of money involved. When you put these two factors together amazing things can happen. I don't know whether deployment schedules for these are six months from now or a year from now. I do know it is not a five year time frame.

In the meantime, looking at this from the point of view of general economics, one of the problems that we see developing, is that the way the network is configured with a hub and spoke layout, the problem of trying to handle at these hubs (also known as NAPs & MAEs) is that you wind up building a very small number of very expensive routers. With small builds of big routers, you might have trouble getting the price down to a reasonable level. These relatively small numbers do not play to the strengths of the computer world, which likes to get good cost/benefit ratio by building large runs of lower performance units. Look at supercomputers - people now by a couple of DEC Alphas rather than a Cray.

By hub and spoke architecture I mean a relatively small number of high speed lines converging at a small number of high speed inter-connect points or NAPs. What this is doing is cramming you up into one corner of the cost/performance curve. It also produces a network which is more vulnerable to trouble; with fewer components, a single failure takes out more capacity. I understand that it offers less overhead in physical plant, etc, but I still wonder

whether it is a wise idea.

COOK Report: Are you saying that someone like MCI could more cost effectively engineer its backbone with 30 or 40 less powerful switches in it as opposed to the 10 very large ones they are starting with?

Chiappa: That's a possibility. The economics are very complicated. Building a powerful router in very small numbers is a difficult market niche to be in.

COOK Report: But if you are talking about the top five or six ISPs which run the default free core backbones of the Internet, you have a niche defined by them that requires such routers. Right?

Chiappa: Correct. What the people at a major router vendor have told me is that in going along their current product line, if you take the speed of each router in packets-per second, and divide that by the price, you do find that the cost of routing packets, rated in packets-per-second routed per dollar of router cost, does decline along their product line.. To date they can get efficiencies of scale. I wonder however if this will continue?

To me it is bizarre that they get a higher packets-per-second-per-dollar ratio out of their top end products than they do out of their low end ones, because you have to figure they are selling tens of thousands of their lowend cookie sheet routers. How they get a better cost/benefit ratio at the top of the curve than at the bottom is something I am still a bit confused about. This is especially true because all of these companies have better margins at the top of their product line than at the low end.

It is certainly true that at the higher speeds you have a router shortfall being presently unable to directly route OC-12. It makes sense to me to use parallelism to focus on this engineering problem. In effect you run a bunch of small routers in parallel with each other to get the total aggregate throughput up. Using cheaper components and using a whole slew of them in parallel seems to me to be the intelligent way to do things rather than try to build a very fast router, out of some exotic technology.

In figuring all this out, you really have to look at a whole series of inter related questions focusing on such things like what is the overall architecture of the network, what size pipes do you have and how many packets per second will they have to push in order to do what is asked of them?

Running Current Routers at OC-12

COOK Report: We want to make sure we understand what is it is necessary to do in order to use current routers at OC-12?

Chiappa: The only thing I can see doing is to take the OC-12 pipe and split it into four OC-3 channels with a router at the end of each of the four channels. Realize that there is no way to mix the channels together. They are really going to have to enter and leave the switching fabric as OC-3 channels. If you do this you get the benefit of the lines directly connecting each switch being OC-12, but if you look inside that OC-12 line, you will see that you basically have four OC-3s being multiplexed.

The phone companies however aren't fazed by this because they have a whole history of multiplexing together bunches of voice grade bandwidth into very high speed channels that actually contain multitudes of lower speed channels. This is how the voice network works. They take a bunch of slow speed pipes and lump them together into high speed channels that actually contain bunches of slow speed channels. I know that the phone company has had T-3 links and that T-1s are multiplexed to put that system together. The switches at the center of the network need only deal with T-3 streams. They actually handle T-3 channels without the switches understanding that there was more stuff inside each channel.

Translate this into the Internet world and you would expect to find something similar going on. Here is what is not clear to me. Can you have an ATM switch with an OC-12 coming in one side with maybe x hundred ATM connections, and split them up so that they are sent out of four different OC-3 ports. Directing certain circuits to certain ports would be a routing issue that would be taken care of at the time ATM virtual circuits are enabled.

But the most tricky thing is indeed the interface between level 3 and 2. Somehow Internet switches have to understand that they need to set up ATM connections (SVCs) between themselves. To the extent that they want to provide resource reservation services at the internetwork layer, they are going to have to figure out some way to have those resource reservations translated down to the ATM layer. The IP over ATM working group is examining these details, I gather.

InternetMCI Bets its Future on ATM

Data Services Marketing & Data Services Engineering Vice Presidents Explain MCI Strategy to COOK Report

Stephen von Rump had been a part of the MCI Data Services Division. It was headed by Lance Boxer, Vint Cerf and Stephen. Data Services have now been mainstreamed into the various functional organizations. As a result Stephen currently heads a large marketing organization responsible for all of MCI Data Services including all Internet related offerings. We had two interviews with Stephen. The first on February 20, the second on March 11.

COOK Report: So where does ATM lie in the scheme of things in MCI Data Services?

Von Rump: Our Hyperstream ATM commercial service has been in operation for about a year. We have a number of customers on it running at speeds from OC-3 on down to DS-1. The most note worthy of our customers is the VBNS running at OC-3 at the five supercomputer centers.

COOK Report: Is the VBNS the only customer running TCP/IP?

Von Rump: No most are running TCP/IP. The VBNS is just the only one running at OC-3. But from a Hyperstream ATM perspective we try very hard to keep these things separated organizationally, logically, and financially so there is no inappropriate leverage of the VBNS agreement. From the Hyperstream ATM perspective we try very hard to look at the VBNS as just another customer riding on that network. However as part of our cooperative agreement with the National Science Foundation which is funding the VBNS, we are doing collaborative research on high speed networking.

Today we have a DS-3 backbone on the Internet. There is currently no ATM connection. We are just using IP routers over ordinary DS-3 links. We are now in the process of upgrading this backbone to OC-3 for obvious growth reasons.

COOK Report: By the way there was a comment on the inet-access list this morning (Feb 20, 1996) that an ISP was told by his MCI salesman that you were putting a 45 day hold on backbone installs because of capacity problems. Is

that correct?

Von Rump: I can't speak to the individual sales person's comment because I don't know the context it was made in. I will tell you that we are experiencing a lot of growth in the Internet and that we are having to accelerate our build plans. It is possible that in certain sections of the country there may be some extended install intervals for a while until we get back farther out ahead of the curve.

COOK Report: People on the list are saying better to delay installs than to crowd more people on an already burdened network.

Von Rump: We feel that the integrity of our network is one of our strongest assets as a company. If we were to mess that up, it would damage our credibility all across the board.

Meanwhile we are in the process of major upgrades. Adding ports obviously. And as we change from DS-3 to OC-3 the OC-3 will actually be an ATM backbone. This will mean that we will insert an ATM switched layer two into our Internet. The DS-3 connections between the IP routers will be placed on ATM circuits at OC-3. We will run this OC-3 ATM internet backbone completely independently of our commercial ATM backbone. We will complete this by the end of April. Completion will mean all our InternetMCI traffic would be on dedicated circuits running at OC-3 over ATM.

COOK Report: So from the technical point of view how do ATM switches fit into what you are doing?

Von Rump: First of all we are using the General Datacom Apex Switches which we have had running on our ATM backbone for some time now. I don't specifically know what their buffer sizes are. I do know the engineering rule of thumb is about 200 milliseconds for coast to coast transmission. So, whatever speed you are operating at, the buffer size had been able to handle the amount of data that you can transmit during that 200 millisecond time period.

COOK Report: So the General Datacom switch has adequate buffer size to

support OC-3?

Von Rump: Yes. They have adequate size today, and just as important, they have satisfied us they will continue to increase the buffer size as we move upward in the SONET chain.

COOK Report: Because a year from now we could be talking OC-12?

Von Rump: Yes and I am not even sure it is a year away because growth is incredibly rapid.

COOK Report: How about the framing issue - where a cell making up a part of an IP packet is mangled there is no way to realize that the mangling has occurred until it gets to the other side of the network and cannot be properly be reassembled by the router? The problem is that the ATM switch has no knowledge of the IP layer and bandwidth is wasted because the packet is transported across the entire network only to be dropped on the other side.

Von Rump: Right. It is my understanding that this is exactly how IP over ATM would work today. If it mangled a cell, it would not know enough to identify the cells in front and the cells in back as part of the same piece of information. I am not aware of any ATM switch that is capable of this kind of recognition, at speeds of OC-3 and above, before transmitting the packet. It can recognize a bad cell and drop it before the transit. But it cannot recognize the relationship of a bad cell to other cells carrying other parts of the same IP packet. Therefore a bad cell has no impact on whether or not to carry neighboring cells.

The Time Table for SVCs

COOK Report: From your point of view what has to happen before ATM can come in and solve the problems of the Internet?

Von Rump: I think that vendors need to finish the implementation of the standards work that has been completed so far - whether it is switched virtu-

al circuits or standards for running fractional T-1s. The promise of ATM is that it gives you a very efficient backbone for a variety of service requirements for example bandwidth reservation and the ability to mix and match a bunch of different traffic types over a very efficient switched backbone. This is the promise of ATM and what we are hopeful that we will be able to offer, not only as a service to customers but also as a more efficient way of running our own network. ATM is clearly moving in this direction and we have a commercial service in place to show that it is.

But for it to become a ubiquitous switching layer for many kinds of services there are some things that are either not yet in place in the standards bodies or certainly have not yet been implemented in any commercially sturdy way.

COOK Report: How about switched virtual circuits (SVCs)? You would have get those out of the standards bodies first and then develop the software - right?

Von Rump: That is correct. But there is a third piece to this which a lot of people overlook. When the software is done and is available in the switches, you have to have a way for the carrier that offers the service to tie it into its ordering and billing system.

COOK Report: In a flat rate, charge-by-the-size-of-the-pipe IP world, this would be irrelevant. However, one of the reasons you want SVCs is to establish bandwidth reservation quality of service options where users can buy slices of bandwidth for specific durations and be billed for these individually. Correct?

Von Rump: Correct. Remember also that ATM audio and or video services that can also be delivered by means other than IP and that we would need to be able to measure and bill for these as well.

COOK Report: What is the time line for the completion of SVC standards work?

Von Rump: I think it will be some time next year before we really see the maturity of these standards and their implementation in switches to the point where we could offer a viable commercial service.

COOK Report: Once the standards are

done what is the time needed for their implementation in software? Six months?

Von Rump: That's a good rule of thumb. The big variable here is how well are the standards negotiations going now? This is another way of asking how close we are to the final standards and to what extent are vendors willing to take a chance on implementing in chip sets what has been accomplished so far? All in the hope that what they produce will be fully interoperable with product produced when the standards are finally finished? But we are still talking about six months anyway, even if the vendors take some moderate risks on the final standards. The early adapters will either be lucky or they will end up with proprietary solutions.

COOK Report: But the third part - getting the software linked with billing systems - is that part sequential to or parallel with the other software development process?

Von Rump: Depending on the risks that the systems developers are willing to take, it can be partly parallel. Once you have the specifications for the software nailed down, you can, in theory, go ahead and begin developing the billing the systems that tie into a given switch. If you do this you are taking risks that you will have to make changes because the other parts aren't set in cement yet. So speaking for MCI once we see things get firmed up, we will begin systems development work before the switched software is actually commercially available. We are also going to insist on a very firm standards environment. We want to make sure that we as a company do not get involved in proprietary or potentially proprietary solutions in the ATM world.

COOK Report: So for the next 12 to 24 months we're stuck with permanent virtual circuits that have to be manually configured. What type of limitations does a world of PVCs place on what we can do in the meantime with ATM over internetMCI?

Von Rump: With regard to the Internet the absence of SVCs is not going to put a major damper in our plans. You can still set up PVC paths under ATM and realize significant switch speed advantages over a pure router environment today. So the absence of SVCs for the next 12 to 24 months is not going to hinder or slow down our migration to an ATM switching layer for MCI's Internet.

If you are talking about a relatively small number of nodes for the ATM backbone, on the order of tens but not hundreds, for

example, you can make a reasonable inference that the number of PVCs we'd have to deal with would be quite manageable.

COOK Report: So for each path between nodes you would have to have a PVC?

Von Rump: Yes. If you really wanted to make the IP and ATM layers independent such that they could be engineered and operated completely on their own, then yes you need a fully meshed or almost fully meshed ATM backbone. Once you have done this as your ATM network grows, you begin to see a requirement for SVCs emerge.

But in the interim, we are managing it as an integrated entity where we are using the ATM to increase our switching speed and go in one hop from one side of the backbone to the other instead of having to go through a string of routers to get from one side of the country to the other. This offers us far more efficiency. And it is this that we want from ATM over the next 12 months. To gain this increase in speed and efficiency and work the ATM and IP layers in harmony with one another we don't need to have a fully meshed ATM backbone.

COOK Report: Would you quantify what you will get out of this?

Von Rump: Our engineering people tells me that it is between 5 and 6 times faster than a backbone driven solely by routers.

COOK Report: But we are looking for a contextual advantage. Is it that you can obviously set up and then tear down requests for video conferences a lot more quickly. Is there a more immediate and desirable consequence of this increased speed?

Von Rump: Yes - the most immediate and desirable consequence is simply better performance of the IP backbone. We expect to be able to go to much higher capacities and still be able to maintain good throughput and low latency. Infrastructure efficiency is what we are looking for in ATM right now, not for any enhanced ATM over IP service capabilities like voice or video conferencing.

Interview Part 2

On March 11 Stephen advised us that Stephen Tabaska, MCI's Vice Presi-

dent of Engineering would be talking to us about more technical issues. Having said this, he added that the new OC-3 backbone would be independent of the commercial Hyperstream backbone and "function as a separate network overlay using Fore rather than General Datacom switches."

Von Rump: However it is still our intention and direction to merge this with our commercial backbone. We do not intend for this to be a separate overlay forever.

COOK Report: So you are going to a separate overlay and the most powerful ATM switch available to you to squeeze the very most bandwidth possible from this upgrade?

Von Rump: Exactly and remember that we can do switching at the ATM layer faster than the routers are capable of routing now.

COOK Report: When you say overlay network, you mean taking a slice of bandwidth from a single glass fiber and dedicating that bandwidth exclusively to the new internet MCI backbone, while the Hyperstream backbone might be running on its own slice of bandwidth perhaps even in the same fiber?

Von Rump: Yes. The ability to do this is dependent on SONET multiplexing equipment that we have installed in order to do this and many other separated activities over our network.

COOK Report: To conclude let me read you a quote from an article by Christine Hudgins-Bonafield in *Network Computing* some months back: "Vint Cerf, the father of the Internet and MCI's Internet guru, says he's probably the token pessimist when it comes to ATM. Cerf hopes the technology will work. MCI, after all, is using it for Internet backbone transmission, but he isn't convinced ATM will scale. He doesn't want to refight the battle, but Cerf is concerned that ATM's cell size (which was set to reduce jitters in voice and video) isn't the best size for data."

Can you give us some idea from a technical point of view on where Vint is coming from in being the token pessimist?

Von Rump: I think the quote says it rather clearly. He is concerned about the fact that ATM was fundamentally a

compromise between the voice and the data guys when the 53 byte cell size was set in the late 1980s. Had the data folk had their way the cell size would have been a lot larger. People like Vint are worried that when we get up to many hundred megabit and then giga bit bandwidth that we may not have the hardware to do things fast enough.

COOK Report: So while ATM at OC-3 and even OC-12 is available now getting it to work at OC-48 speeds or higher is an unresolved issue?

Von Rump: Certainly it is not resolved to Vint's satisfaction.

COOK Report: Alright. Thank you Stephen. We're looking forward to talking with Steve Tabaska.

Steve Tabaska is the MCI Vice President of Data Services Engineering and is based in Richardson, Texas. In this role, he is responsible for development and support of all MCI virtual data services, including frame relay, SMDS, ATM, Internet access and X.25 packet services. Prior to joining MCI in April, 1994, he was director of Advanced Technology at WilTel, now LDDS/Worldcom..

COOK Report: Please tell us more about the Fore switches that you are using. How long have they been out and where have they been used before.

Tabaska: Fore systems has a whole family of switches. They have made their mark mainly in the government arena where they have a very significant share of the government market for LAN and campus type backbones. What we are buying is their largest, the ASX1000 that costs depending on configuration about \$100,000 a copy.

COOK Report: How many will you be deploying by the end of April?

Tabaska: We are looking at roughly ten of them to start. We have actually used this Fore switch on a couple of government bids for some four to five months where we have put the equipment onto the customer premise. In those contracts we have installed, configured, managed, and maintained the switch.

Because Fore has made their mark in high speed campus LAN environments with TCP/IP networks, they very well understand what is needed to handle high speed IP over ATM.

They have a larger buffer than the Gen-

eral Datacom, some 13,000 cells in size which is over 700,000 bytes. They have also solved the framing issue to the extent that they can catch about 90% of problem cells before the packets containing them are sent over the network.

NAP Bottlenecks?

COOK Report: What bottlenecks exist at NAPs? What would you have to do to get the vBNS OC-3 circuit connected all the way through a NAP and routed to a campus?

Tabaska: The big issue here is routing. How much information can you bring into a NAP and how many places can you route it to and from? At a NAP you are going to have to be bumped up to level 3 and go through a router before you can get back down to speedy level two again. The routing is where the black magic is and where most people agree we have a bottleneck point. Most people are using Cisco 7,000 routers with silicon engines inside. If we bring in an OC-3 pipe of commercial data, and another OC-3 of VBNS data, and Sprint and UUNET each bring in their OC 3s to a NAP, we have just way exceeded the capacity of Cisco's biggest box. In such a case you have to figure out how to cluster routers and do route aggregation.

Higher speed inter connect and distributed routing is clearly what is going to be needed. We are also convinced that hierarchical routing will be needed. We only see a couple of vendors doing this. One of these is the Giga-router by NetStar that we are using now on the VBNS. They are doing routing at very high speeds. We are not running them up to these speeds but they are claiming that they will go all the way up to OC-12 [622 megabits]. They do have routing capability at these high speeds but unfortunately they lack the full suite of routing protocols.

COOK Report: So its a question of getting Cisco's software into a box of hardware that can do the job?

Tabaska: Exactly. I see the NAPs as tending always to be congested, so, even if we alleviate things there by clustering routers we still need some high speed interconnection points and have to figure out some traffic engineering between them. This gives six to nine months more breathing room before the congestion crunch hits

again. The NAPs are the most congested points in the entire Internet and so while I'd like to say we have an answer, until we move to much higher speed routers and hierarchical routing I don't think we do have an answer.

COOK Report: Can you explain what you mean by hierarchical routing?

Tabaska: It is really one of how do you interface to a high speed stream of 622 megabits per second and make the decision to send it to ten Cisco 7,000s? You need a single device that can handle route aggregation at OC 12 speeds and then split that stream down to many different lower speed routers like Cisco 7000s where the Ciscos would distribute the traffic further. So you do route summaries and make quick and simple routing decisions at the higher speeds.

COOK Report: So in effect you are adding a new top level routing hierarchy?

Tabaska: I think that when we are talking about these higher speeds we will simply have to do this.

How Do These Issues Affect the NAPs?

COOK Report: Then it seems there is a related issue. MCI and Sprint are having some private interconnects to exchange traffic from one backbone to the other rather than wait to do it at a crowded NAP. Are these hand offs much less resource intensive? In other words if I have to be prepared to hand off to any of five or six other major players, how much am I increasing the demands on my routers at those major exchanges? Is it five or six times more resource intensive?

Tabaska: The more exchange points we have the more we can distribute the load. Really the reason for the private interconnect points that you mentioned is the desirability off loading traffic where possible from the big NAPs. We are connected to the NAPs, and to CIX and MAE East and West and we are connected to a number of private interchange points. The goal is really one of distributing the load by aiming more exchange points. Right now I don't recall the exact number of private exchange points we have but they are certainly increasing in number.

COOK Report: How do you define a private exchange point? They are putting together a New York MAE. Does

this fall into such a category or are you talking only about and exchange between you and UUNET, or you and Sprint?

Tabaska: Private exchange points are really ones created just between two companies with no third party involvement. It makes sense to us to exchange traffic with the larger providers directly. We eliminate the surcharges associated with NAPs and CIX and we can also distribute our load.

Really what we want to do is get to a point where we don't have to invent new technology like I think will have to be invented to make the NAPs work as defined today because the loads there are just getting so high. Everyone would like to peer at OC-3. Unfortunately I don't see the technology available to make all this happen without some new invention.

COOK Report: So is the problem of higher speed interconnects at NAPs, with five to ten large players interconnecting, that you have these scaling issues while, on the other hand, if you had dozens and ultimately perhaps hundreds of private exchanges, you could still get the traffic routed to where it finally needs to go? Speaking hypothetically you might have 7 or 8 with Sprint, 6 with UUNET, 4 each with PSI and ANS? All of them scattered around the country. You can do one-to-one exchanges reasonably efficiently in a lot of different places - instead of coming into a "NAP" like MAE East which is rather like O'Hare Airport at rush hour?

Tabaska: Yes that is true. The technology just isn't here to scale the NAPs to the size they need to be. The sizes are getting to the point where I think they are almost out of control. New routing technology has to be invented to make the design of the current NAPs viable. In the meantime, while waiting for the new routing technology, we are having to get on with our business by using growing numbers of private exchanges. In order to get by we expect to be doing this for the near term - at least for the next nine months.

COOK Report: And if in the next nine months the NAPs are scaling better, great! And if this is not happening, presumably you just go on with a fairly sizeable number of one-on-one exchanges?

Tabaska: Yes. And it wouldn't be inconceivable for us to end up with a three way exchange. We all have major interconnect points. If you overlay our networks, you will find that we all hit the top ten to fifteen cities. So it would not be inconceivable for a group of us to exchange in

one of more of these large cities making such an exchange or exchanges something like pseudo-NAPs but I would still call them private exchanges because a third party is not working the "NAP" part of the exchange.

COOK Report: Because you and UUNET might find yourselves in virtually the same wiring closet in the Westin building in Seattle?

Tabaska: Yes. If we had such a location, why not do a three way instead of three separate two way exchanges?

COOK Report: If we are talking about scaling issues at NAPs, and about private interconnects as a result between large providers, to what extent do you think we might see a situation where, among the five to ten largest providers, private interconnects become the main means for transfer of traffic? In the meantime, for the next 20 to 50 middle sized ISPs, the NAPs become the place where they do their traffic exchange with each other and then also exchange traffic with the big players?

Tabaska: That is a possible scenario. I think that the way all this will play out is undetermined. Until we get technology scaling to the size needed at the NAPs it is difficult to see how we will really rely on them and, as a result private interchanges will continue to increase. Whether the technology turns around and gives us the routing capability we need at the NAP points, will determine our future direction. If it happens quickly, I can see the NAPs becoming the predominant points for exchange of traffic, if it doesn't occur

COOK Report: If the proliferation of private exchanges continues and the larger players find that they are relying less and less on the NAPs, the possibility for the middle size players to go to the NAPs to be able to connect right at the top of the Internet pyramid disappears and they are pushed back down to just being customers of the big players and relying on them to do their peering and transport for them. Is that a fair conclusion?

Tabaska: Sure. Very fair.

cur, I think the proliferation of private exchanges will continue. It really has become a case of the technology tail wagging the entire dog.

COOK Report: If the proliferation of private exchanges continues and the larger players find that they are relying less and less on the NAPs, the possibility for the middle size players to go to the NAPs to be able to connect right at the top of the Internet pyramid disappears and they are pushed back down to just being customers of the big players and relying on them to do their peering and transport for them. Is that a fair conclusion?

Tabaska: Sure. Very fair.

Ease of the Next Speed Increase

COOK Report: If you are looking at moving from OC-3 to OC-12 possibly by year's end, what is it going to be necessary to do in order to ratchet up to the next level of speed?

Tabaska: The ASX1000 already has OC-12 available. We have lab tested it but are not quite ready to use it because the demand isn't quite there yet. We would just have to take the trunk interface cards and replace them with OC-12 connected to OC-12 private lines. Then on the fan out side of the switch where you would go to individual routers, the highest speed interface is still only OC-3, so you'd have four OC-3 pipes going to four different routers.

COOK Report: So you are saying that right now the only thing to do with OC-12 is to split the output 4 ways and shove it into 4 Cisco 75XX series routers?

Tabaska: Right.

COOK Report: And at some point some number of months from now Cisco can be expected to have an OC-12 interface for its routers?

Tabaska: That's right.

COOK Report: Might NetStar have it sooner?

Tabaska: Yes, except NetStar has some other problems. We don't think it has a full suite of routing capabilities.

COOK Report: So their software is

deficient while Cisco's is not?

Tabaska: Yes.

COOK Report: What about Ipsilon?

Tabaska: They could help with the routing software, but I'm not aware if there is any relationship in place" ..

An ATM Cloud?

COOK Report: If you look at your ATM backbone, are you in effect talking about having a high speed nationwide ATM cloud? And, if so, what, if anything, would keep you from selling NAP access via the cloud?

Tabaska: Sure. For everyone who is buying an Internet service with a private line, buying ATM technology means buying into speeds higher than what frame relay can presently deliver. Also most interconnects at the NAPs will be moving to ATM. So the idea of plugging into a national ATM cloud to get to the NAPs would make sense. We have had a high speed ATM service now for over a year. If we have Internet access via frame relay why not have it via ATM?

COOK Report: So it could be fully priced service where the cost of x bandwidth to Y NAP would be known to all well ahead of time?

Tabaska: Yes. But there will also be some issues about whether MCI connects to the NAP and pays the NAP fees or whether the ISP buying our transport service is responsible for that?

COOK Report: So are you actually working on such a service? And should we expect to see it priced by the end of the summer.

Tabaska: We haven't seen enough demand for this to put it on an actual development schedule yet.

COOK Report: So it is more of a hypothetical capability than a planned product?

Tabaska: Yes. There would be very little technology challenge to making this happen. Most think in terms of a DS-3 when interconnecting with NAPs. If they are truly going to consume a DS-3 to a NAP, the best technology to use is a private line. But if they only use some fraction of that or their needs are for bursty traffic, then

an ATM cloud would make sense. When you start getting down into the second tier of Internet service providers such a service might well make sense.

ATM and Bandwidth Reservation

COOK Report: The possibility of video and audio bandwidth on demand seems to be a major reason why people are enamored of ATM technology within the Internet. Some say that switched virtual circuits are a key component of such technology that we do not as yet have. I understand that the standards work is not yet done. We hear 12 to 18 months before SVCs really arrive. Would you explain what is holding us up? Does RSVP hold any promise as an alternative here?

Tabaska: I think you will find that there will be experimentation with SVCs before the end of this year. We feel that you need not only the signaling protocol, but also something that triggers it. Let me give you an example. If a switch or a router on a customer premises required an operator to go over to a separate console and issue a command in order to enable a network connection, this would be a lot less attractive that if an IP datagram would do the necessary configurations for you.

COOK Report: What you have defined is essentially the difference between and PVC and an SVC? Correct?

Tabaska: From a network perspective - as long as you send the signaling message from you to me and the network sets it up automatically - both are really SVCs. So, if you go to the operator console, and you send the command from the customer premise switch into the network and it the command is a signaling message that the network receives and process on the fly, it is still really an SVC. I am talking about a person at the customer premise having to take action as opposed to something that is embedded in the protocol and application to be able to go right from the end user out and across the network.

COOK Report: Are you describing something that is in the midst of transition to a full SVC?

Tabaska: Let me give you an analogy with the phone system today.

When I pick up my phone and dial a number out, my PBX understands and it signals the network to complete the call. Now if I didn't have the flexibility of being able to dial the digits myself from my desk top phone and I had to call an administrator in my building who made the connection across the network for me, it wouldn't be that usable. As long as the network itself makes and breaks the connection on command it's an SVC. The analogy is that if I had my computer sitting here and I wanted to something across the network for which no connection was established, my router or switch translates my command into a signaling message which makes the connection across the network for me and when I am done it tears it down.

It has to be this easy, and right now no SVC mechanisms are actually built into the applications. This area is definitely immature while the signaling interface that comes to the network from the customer premises over the UNI is fairly well defined. But there is still some unfinished activity in this area.

COOK Report: So you have been defining the conditions that must exist in order to have SVCs really user friendly and desirable?

Tabaska: Correct. I think you will see experimentation with SVCs this year. Commercial viability does extend into the 1997 time frame. I could probably deploy a switch that meets SVC signaling specifications by the summer time frame. Should I be doing this? Would it be profitable to deploy my resources in that direction? It doesn't seem like the market is going to mature in this time frame.

COOK Report: When you have this capability, do you then have the capability for reserved bandwidth on demand?

Tabaska: Yes. Exactly.

COOK Report: What about RSVP? A couple of people have told me that RSVP may be able to do quite a bit of what we have just discussed. Do you have an opinion on this?

Tabaska: RSVP is also in its quite early stages and I think you are also going to see experimentation with it. People who are very strong IP advocates will probably put their complete trust behind RSVP. I think a combination of RSVP and ATM qualities of service makes a lot of sense. ATM as

the switching technology and IP as the routing technology and RSVP combined with ATM qualities of service is really the way that things will work out.

A Dual PVC - SVC Architecture

COOK Report: We have heard some criticisms of ATM as being a very poor match for certain types of traffic. For example you wouldn't want to have the network set up and tear down SVCs every time any sent a command to a web server. Or even if you are not using SVCs, people say you can't use ATM efficiently for things like domain name service look ups.

Tabaska: I think that what people are getting at is that with small packets and small hold times it doesn't make sense to set up a connection oriented circuit and then have to tear it down. But there are ways to actually handle the setting up of PVCs to servers to use for these short but frequent transmissions. With PVCs you can actually maintain a connectionless environment. For example on the VBNS we provide today essentially a connectionless IP service over a PVC.

COOK Report: So in your world of the future you could have some portion of your bandwidth dedicated to PVCs for short and bursty traffic and SVCs for the other type of traffic if we are hearing you correctly?

Tabaska: Yes. You have a switching layer which is usually going to be layer two with ATM or at the low end, frame relay. You create a PVC to connectionless data servers either ones with an IP flavor or one of an SMDS type. Behind these you have a bunch of services, domain name servers being but one example. All your quick short hits come by your connectionless PVC route.

COOK Report: A couple of months ago in an interview John Curran of BBN gave us the impression that if you look at big IXCs like Sprint and MCI they had enough bandwidth and other infrastructure so that they didn't have to worry much about putting ATM switch into the core backbones at this time. Yet you at MCI are now doing what Curran very recently thought you had no need to do. Why? Is demand for bandwidth so rampant that you believe you must do this to squeeze all the efficiencies possible out of your transmission? But are you also doing this because you will

then have all the infrastructure in place to do SVCs and bandwidth on demand? This puts you in a position to get out of the starting gate before your competitors. Correct?

Tabaska: Some would make you believe that RSVP by itself without any kind of switched environment could give you the bandwidth reservation that you want. Other people would say we are just in the early stages of where this growth will take us. While some find a 155 megabit per second backbone amazing I don't see it stopping there. I see it going to 622 and beyond.

In packet level processing like frame relay there is much more processing required for variable sized packets than for fixed size cells. And that the only way to really scale the network is to bet on fixed size cells. Now when we go to 622 megabits people may well prove me wrong and build silicon routers like NetStar that will do the job. But at the same time that we have routers of that speed, ATM switches will be operating at 2.4 gigabits. Or even higher to as much as ten gigabits per second worth of individual interface capabilities because the entire ATM switch already functions at a total of ten gigabits today.

COOK Report: Well lets close by looking at the quote of Vint Cerf's in Network computing some months back where as we said already to Steve Von Rump, Vint calls himself the token pessimist on ATM at MCI because he thinks the cell size is just plain too small to be useful. Can you help us understand what Vint was talking about? It almost sounded like you thought it would scale because of the cell size.

Tabaska: Most protocols do want to see something close to 1500 bytes or bigger. The bigger they get the higher through put you can ultimately have. But if I want to move data across a device - call it a switch or a router - where something comes in, an address look up is made that determines where it goes on the other side, if I wanted to move an aggregate bandwidth of ten gigabits across such a device, and I ask can I do it with a router? The answer is no. Can I do it with a frame relay switch? The answer is no. Can I do it with an ATM switch? The answer is yes.

When we talk about moving information at high speed, it doesn't have to

be ten gigabits from one single source to another. In fact the biggest problem with regard to high carrier networks like this is how do you aggregate lots of customers be they T-1 or 45 megabits? The end station will be the bottleneck and the over all through put doesn't to travel at 10 gigabits between two work stations

What we need is adequate switching capacity. ATM switches start in the five to ten gigabit range and may move up into hundreds of gigabits. Routers starting with megabit speeds may be able to be pushed to a gigabit speed.

COOK Report: So, to make a simple example, is the hope to have a major take off point in New York and a landing point in San Francisco and have multiple gigabits being thrown from New York to San Francisco where, if worse comes to worse, instead of having one or two routers that are able to route that firehose you might have to have quite a few routers dealing with what - demultiplexed streams from all this? We guess we are back to the question of what do you do with the interface from that huge fire hose of switched data back into the routers themselves?

Tabaska: You essentially use the switch on each end of it to provide the multiplexing and demultiplexing from OC-12 to OC-3 and route virtual circuits into those routers which will ultimately terminate the ATM stream and then move it up to level 3 for routing. And, in your end-to-end performance, you would be throttled back to the speed of the single slowest link which would be OC-3. So you come in with OC-12 into an ATM switch and then you go to at least four if not ten routers - all of them attached to OC-3. You have virtual circuits aggregated into the switched ATM core for the virtual circuit multiplexing demultiplexing function. But you essentially run across the network at 622 megabits per second. And in a few years 2.4 gigabits and in a few years after that 10 gigabits.

COOK Report: And if you are talking about a single strand of fiber what is the total maximum that you can pour down it today? Two point four gigabits by means of your SONET transmission facilities?

Tabaska: Right now we are pumping 2.4 gigabits over our fiber plant. We might have one or a few very thin routes of ten gigabits. So a fiber now

can actually carry ten. There is also going to be wave length division multiplexing going on. So you may have multiple 2.4 or 10 gigabit bit streams flowing over the same strand of fiber. So the aggregate bandwidth that you can transmit over a single fiber can today exceed 40, 80 or even 160 gigabits per second. Today 2.4 is common and 10 gigabits runs on an exception basis.

Now what can go into an ATM switch? No manufacturer has so far made a available a 2.4 gigabit interface. Many have designed the architecture into their switches at 2.4 modularities for switching purposes so that they can accept a 2.4 gigabit or OC-48 interface but today no one has commercialized this. OC-12 or 622 megabits its currently the fastest path into and out of an atm switch. The vendors are saying 2.4 gigabits as soon as you tell us what the demands are and what we are saying is probably 1998.

Gartner Group's Lack of Professionalism Irks COOK Report

In the Business Section of the March 24 Sunday *New York Times* the Gartner Group ran a multi-page advertisement for its Internet and Electronic Commerce (TM) Exposition. On page 27 there is a Glossary of Internet, World Wide Web and Networking Terms.

On Tuesday we received a phone call from a friend asking if we had seen our web page glossary in the *New York Times*. We were amazed and flattered and ran outside to snatch it from the hands of the recycling garbage pick-up. There we read at the bottom of a 48 term Glossary: "Some of these terms and definitions were suggested [an ambiguous term if ever there were one] by The Cook [sic] Report Guide for the Perplexed: Understanding Internet Jargon. [http://pobox.com/cook/glossary/\[sic\].html](http://pobox.com/cook/glossary/[sic].html). Others were culled from Gartner group research reports. Gives the impression of a 50/50 mix doesn't it?

When we started reading the glossary to figure out what terms were ours, our pleasure evaporated. There were three of the 48 (modem, NSP, and SMPT) that had been lifted with some rephrasing from our work. What was worse however was that the quality of most of the remaining material was horrible. Consider: ISOC - The Internet Society, a member organization consisting of

individuals and International organizations. (Definition 26). Impressed by how informative that was? We weren't. Now ISOC was judged so important that it was defined twice. Definition 20. The Internet Society - not ours either - was disjointed but somewhat better. Then there was the incredibly informative (5) CATV: cable television. And (41) URL: Pronounced "earl," it stands for . . . You get the picture.

So Gartner has published a grade D glossary which leaves readers with the impression it was largely our work. And Gartner, hearing that everything on the Internet is there for free, grabbed some of our work, tossed it in with their own mangled mess, in such a way as to leave readers with no knowledge of who was responsible for what. Moreover, they did all this without asking any permission.

On the Access Indiana Front

During February and March use of the COOK Report supplied mail list ai_inexile@pobox.com took off to the extent where discussion was as heavy and lively as discussion on the pre censored internal state list last summer. Noting that, Mike Huffman of the Indiana Department of Education announced a new state list, unmoderated, and uncensored called Access Indiana-in-Indiana. Unfortunately, Huffman has been unable to make the state machine function properly and for the moment discussion on both lists has been stanchued. The absence of Mark Whitman, Ed Tully, Ray Ewick and John Sullivan from the policy discussions of the program for which they are responsible is remarkable.

DNS Service as Another Factor in the Stress of the Network

We have seen some interesting private discussion which we have been unable to get permission to publish. We have also seen some public mention of the problem but not enough to do a detailed discussion. For the time being let us say only that there is growing belief in the network technical community that the sheer number of DNS calls is putting tremendous stress on backbones and their routers. Remember every hit on a web site is also the same as a Domain Name Service call. The situation is exacerbated by the improper configuration of some DNS servers by new ISPs operating with staffs having less than adequate technical know how.

AGIS Closes Down NET99 DNS

Inept Handling of Customers Continues

Editor's Note:

Many have been wondering why we have been harsh on AGIS. We have been harsh on them for for several reasons. Primary among them is the fact that they bought NET99 and pledged to us on the night of the purchase to continue NET99's policy of being a backbone provider committed to offering high quality of service to the small ISP. It is our contention that they have since abandoned the pledge. They have effectively disassembled NET99. As readers of this article will see, in the words of their very own customers, they have given a sizable number of customers amazingly poor service.

They have displayed an ignorance of and contempt for Internet culture cheerily as evidenced by Peter Kline's remarks last fall that they were not interested in cooperation but would instead use lawyers to impose order on an unruly Internet. Peter's position was that the use of lawyers was the best means of establishing quality of service.

Any reader of these installments will get a glimpse of business practices that in our opinion do not meet the highest ethical standards. We will continue to be harsh on them. AGIS is trying to act like a major player. With the Internet hype and speculative fever continuing they could convince the unwary that they are a major player. They are welcome to try. In the meantime we shall hold up the mirror of the reality of their own making in front of them.

Part 2 Continued from Our March Issue

Jan 30 **Rich Drewes:** Then the week before last I started seeing *major* degradation in my T1 link: 80% packet loss to major sites(!). Also mixed in were *total outages*. There was a mysterious email on AGIS' own customer list about MFS having a major problem at MAE East . . . but then I call AGIS NOC and find out that "all the engineers are in San Jose" for a major crisis there. Hmmm. East or West, where is the problem? . . . After a few days things seemed to be getting better. Packets were getting through reliably . . . then wham! total loss of connectivity to portions of the net, other portions were OK. I reported the prob-

lem to the NOC via email, don't hear anything back. I call the NOC voice, which took the information efficiently (they have that down). However, they were unaware of my email. Gulp. . . . My feeling is: with MCI possibly charging by byte soon, and Sprint filtering routes, and UUNet costing what, four times AGIS' price for a T1?: we should pull for AGIS. But we have to hold them to a high level of service so that don't go away by pissing all their customers off.

AGIS: your customers will be happier knowing about problems as they are developing, even if you think they are so serious they might spook us, rather than hearing about them *ex post facto*. I understand it is hard to communicate to everyone via mail

Let's hope AGIS is getting a control on things internally. I sort of was expecting some kind of trauma period of outages around the acquisition. It almost seemed inevitable to me that it would happen when one organization takes control of another, including their network (no matter how good or overstressed that network may be). Let's hope this was the trauma.

COOK Report: In late August in a phone call to Phil Lawlor we explained how Stroup and other Net99 folk had used inet-access very effectively to answer customer concerns and questions. I told Phil that I thought it imperative that he do the same thing. He seemed very surprised by the suggestion.

Dave Siegel: Not only that, but Stroup used it as a problem detector. When he saw shit on inet-access, he would come runnin' into my office (even if it was only one small complaint). "Goddamit, we've got customers bitchin' about the news service", or "the service in LA sucks," or whatever. "Dave. What do we do to fix it? I need it fixed NOW."

He didn't always like my answers, but he always made a point to get back to the customer, if only to say a three line response "Our engineers are working on it" even though in many cases, it might have been a small notation on one of three 8 foot x 4 foot whiteboards in my office.

Amanda Robinson [Amanda is DNS admin for AGIS]: Mark Cole asked - Has anyone ever been to their facilities? Are they run out of a house, a business, a garage?

Since most of our major equipment is in POPs, the office does not need to be impressive. There is a total of 14 people in the Dearborn office, including secretaries, sales staff, and Peter (who doesn't spend much time here). There are 5 people in the Phoenix office (formerly Net99 world headquarters), all of whom are sales, install coordinators, or secretaries. I have no idea of the exact number [of AGIS customers] , but I can assure you that there are more than 5 or 6.

Rich Drewes: I have found Amanda Robinson good to work with on a few DNS issues, and a few news issues . . . but the recent *total news meltdown* must have her resource crunched on more important things (I am guessing here. Amanda?).

Amanda: That would be an understatement. Although I find it heartening that when my performance is less-than-stellar, people attribute it to "resource crunch" rather than "incompetent" or "uncaring" or less tactful combinations of the two. That's a good part of what keeps me going, and I thank you (all) for it.

Rich Drewes: AGIS/Net99's news feed was never great when I had it, but it at least gave me something to bargain with when I set up a feed swap that is now saving my cookies. If I were starting up now I'd be much more concerned since I wouldn't have any feed to swap with to build up redundancy.

Amanda: If you were starting up now, and had no other feeds, you would have gotten a feed. Feeds were down for two weeks because I believed it when I was told "You'll have working hardware next week." After two weeks of no machines, I started improvising, and to my knowledge everyone who had no or inadequate backup feeds got a feed from SOMEWHERE. And as long as the topic has come up, I would like to publicly thank all of the customers who allowed me to "borrow" their news machines and who set up feeds to sites previously fed by the dead AGIS machine. On the other hand, I will admit that after a while I

had to start refusing new customers who had no newsfeed, as I ran out of places to put them.

Rich Drewes Then I got a pro-active call from AGIS (Tom I think) the next day asking if the problem I reported was resolved. I said yes, thanks. This call is a good sign, I think, and I appreciate their calling me to check. I hope it had nothing to do with setting up this list.

Amanda: I doubt it, as I think I'm the only AGIS staff member on this list, and very probably the only one aware of its existence. We do have a web page for the sole purpose of reporting outages. Unfortunately, when outages occur, the first reaction of the NOC staff members is to try and get it fixed, and we frequently do not remember to update the web page until well after the outage is fixed.

Rich Drewes Let's hope AGIS is getting a control on things internally.

Amanda: I'm not sure how well I can address this without stepping into confidential areas, but since I'm the sole AGIS representative at the moment, I'll do my best:

We're working on it. We have one new news machine up. It is still in a beta-test phase at the moment, but it appears to be quite stable. Its primary function will be to feed customers that I previously had to refuse a newsfeed to, followed by customers who are currently being fed by other AGIS customers (see my public "Thank you!" above), followed by feeds currently fed by news.agis.net. This machine will hopefully be the first of several new news machines.

Two former nameservers are being shipped to our Dearborn office, reloaded as nameservers, and re-deployed in other areas. We had some trouble getting these machines shipped previously. These machines will help to reduce the strain on our current nameservers and improve DNS all around.

My knowledge in other areas is severely limited, so I can't address most of those issues. However, I do know that there has been considerable talk of hiring more NOC staff, and that should alleviate some of the resource crunch and problems with getting responses.

February 1: **Mark Cole** - Well - this "Agis" list sure seemed to die down in a hurry... Maybe I can pick it up again.. - but just in case, I'm copying this to the inet-access list. First time I've gone public on the inet-access list, but it's

time - Agis is not doing the right thing...

Just called into Agis to report another routing problem (actually the same one I've been reporting for the last three days). Spoke with a person named Eric (first time I ever heard of him). He said there were serious problems at Mae-East and *everything* was screwed up. When I tried to nail him on what was screwed up but he danced around that one...

Any way - I realized that *this* was a person who I had not spoken with before so I thought I'd try to find out who he was... His name was Eric, and he said he worked in the NOC 6 months ago but has since moved into national sales. Aha! I said - a salesman! Perhaps he would be sympathetic to a customer. So I proceeded to dump... Boy - was I wrong!!!!

After he stopped laughing at my woes (I mean it - he thought I was joking!) I asked him if he was aware of the Agis mailing list. He laughed again, and said that was a *list of people with nothing better to do with their time*. I BLEW MY TOP!!!! I work 8am to 11pm, 6 days a week, and a lot of the day on Sunday. Much of my time is spent figuring out what's wrong with their network, calling them, writing emails that are ignored, explaining to customers why their website is down, why the routes don't route, etc.

I told him that I have ordered a T from UUnet - and I don't even believe that bothered him. I guess they have more customers than they need. And By the way - I still have not heard from Peter or Phil, by phone or email.... I give up - does anyone else out there share my opinion, or you are you all just hanging around this list 'cause you've got nothing better to do???

Rich Drewes: As most of you are aware, Peter Kline sent several letters to Agis' own customer list Wednesday and Thursday warning of a hardware problem in Santa Clara. Hooray! Their communication was much better on this one. Status updates and everything! Also, I didn't directly even notice the outage, though some of my customers reported problems that may have been related. However, my customers were more tolerant since I put them on notice of potential problems thanks to Peter's kind note.

I think the real value of this list will come about if and when there are more widespread problems, since we can compare notes. However, if AGIS is as forthcoming with information if and when the next outage occurs as they were on this last one, then this list will have accomplished its objective and will not be necessary. But that doesn't mean I'm going to delete it! I intend to report to this list every AGIS outage I experience. I encourage all of you to do the same.

To the extent that AGIS values ISPs as a future customer base, your messages to inet-access *will* make a difference. It may not make a difference this instant, but it will long term. When I was selecting my upstream provider, I read the inet-access list and I bet more than half of you did the same. Your message may have dissuaded a few potential customers, who knows.

Carlos M. Gutierrez: I am not a current customer of AGIS, so I don't receive the other notices that you normally would. I would encourage folks to keep posting to this list... so that potential fellow customers can get the full scoop on the current AGIS situation...

As a prospective customer, I would like to commend Amanda's interest in reading both inet-access and agislist, plus responding to (difficult) questions. One thing I liked about Net99, was having Joe Stroup answering questions - no matter how nasty- on inet-access and com-priv. That's just not available with Sprint, or MCI... I can't figure out why Agis would buy Net99... to immediately dismantle it! If they had kept the Net99 personnel, probably Mr. Kline would have enough free time now to do some "public relations" on this list and others!

I think everybody here has a lot of interest in seeing them shape up and then some. Some because they are already connected, and everybody because Agis is one of the few national options to Sprint, MCI, UUNET.

COOK Report: One of our subscribers reported that AGIS made the following claims: "AGIS carries 15% of domestic traffic, 12% of global traffic. Through special arrangements with other carriers worldwide, they have over 200 POPs (for leased line-type connectivity not dial-up). They have over 300 dedicated customers with T-1 or greater." Our personal opinion is that this represents a greatly exaggerated view of reality.

Feb 2, **Dave Nye:** I'm beginning to get a bit spooked here with all this. I still have 2 years on my contract from hell with these folks. I've had some nasty outages in the last few months, some over 24 hours. Most were blamed on MFS and fiber cuts, etc. This has caused my customers and ourselves pain and suffering due to being left out in the cold after several calls and no notifications as to WHEN and IF the link would return.

I've gone so far as to order a second line from another provider even though I'm not using anywhere near my full T-1 bandwidth. I guess that's just part of

the business these days, if you're not multi-homed you're not gonna make it.

The turnover of folks there is a bit worrisome too. Why is it happening, why haven't they hired any top notch folks? I had to tell one of their NOC folks this morning how to use 'hash' on ftp because he never heard of it before.

February 5, **Dan Graupman**: I'm running a startup ISP located about 45 miles west of Chicago. We have been up since the end of November '95. My only complaints concerning AGIS/NET99 is the News feed. I've just recently installed a Pagsat news feed, so I'm not totally dependent on Net99 for news anymore. In the near future I plan on looking to exchange feeds, or purchase feeds from other sites. I guess there is another complaint, that would be the lack of response to the various email addresses, routing, newsadmin etc. The times that I've call the NOC, though, have always been productive.

Eric Kozowski: Well, I reported this [routing] problem this morning @ 0923 PST by sending email to routing@agis.net, noc@agis.net, peter@agis.net, phil@agis.net and have yet to receive a response. I've also tried calling their noc and either get a busy or no answer. Another fine example of AGIS' poor network and customer support.

The Downfall of Amanda

February 6, **Jeff Weisberg**: Uh Oh: this just found its way into my mailbox: Amanda Robinson is no longer with AGIS. Please direct all correspondence to news-admin@agis.net, dns-admin@agis.net, or noc@agis.net, as appropriate. Why do I suddenly not feel so well?

Dan Graupman: Fast turn over of personnel is never reassuring, whatever the reason. What bothers me is that she seemed to be the main news and dns person. When I first signed up for a connection, I was assigned 3 sales people sequentially, before we came online. I received my first routing information from Carlos, who then disappeared. My first News info came from someone, I don't remember the name. Who was then replaced by Amanda it seems, and now she has left also.

February 7: **Amanda Robinson**: After reading a (rather large) number of messages that found their way into my personal mailbox, I get the impression that everyone expects me to take this opportunity to "tell all" about how awful and evil and completely terrible Phil and Pe-

ter were to work with.

They weren't. Sorry to disappoint you. For at least 60% of my employment (the first three of five months), I was completely happy at AGIS. I -liked- coming to work every day. That's not to say that everything was sunshine and roses, either. But every company and every employer and every human being on the face of this planet has a good side and a bad side. Phil and Peter are no different.

Having said that, though, I'll probably be focusing more on the bad side, as I will freely admit that I am biased on this issue. But please keep in mind that everything I say about AGIS -is- biased, and that I am not making any concerted effort to point out any of Peter or Phil's many good qualities. AGIS's biggest problem is that they are understaffed. And the staff they do have is burned out-- especially Peter. The average stress level at AGIS is extremely high. And that tends to make tempers run hotter than normal. I won't go into specifics, but about a month and a half ago, several tempers exploded at once. The morale since then has been quite low, including (and probably especially) my own.

I made several attempts to discuss problems and possible solutions that went nowhere. Customers aren't the only ones who can't get responses out of email sent to Phil and/or Peter. Talking to either of them face-to face is nearly impossible, as they (particularly Peter) are usually out of the office.

I probably got the worst of that, since right around the same time, news3 died. And the nameservers started crashing (or, more specifically, started crashing more often). So my stress level went up even higher, and my frustration at not getting responses out of Phil or Peter soared. I seriously considered quitting at least a half-dozen times, and decided to stick with it "a little longer" for a myriad of reasons (one time it was because a salesman who depended on me to answer technical questions figured out what I was thinking about and begged me not to leave). I eventually set a mental deadline-- if things didn't get better by Feb 1st, I would give notice (then I looked at a calendar and discovered that Feb 2nd was a Friday, and revised my mental deadline). However, I did not mention that deadline to anyone else.

Feb 2, Peter finally got around to my 90-day review (two months late), and offered me the same pay, the same headaches, and a different title (in his defense, many of the headaches would have gone away on a long term scale). I told him I'd been thinking of quitting, but would consider his offer and give him an answer on Monday (after all, it wasn't a

bad deal long-term). However, it didn't take me that long to make my decision, and I knew by the time I got home Friday that I'd be putting in two weeks notice on Monday.

As it turns out, AGIS took the decision out of my hands. I tried to log in and read my email Saturday, and discovered that my password didn't work. I called Peter to find out why not, and was asked if I'd ever heard of agis-list@interstice.com. Naturally, since I've been subscribed to it for several weeks, I said I had, and was informed that by subscribing to this list without telling management about it, I had "turned in my resignation" and that AGIS had accepted it. (I seriously disagree with that-- if they want to fire me, that's fine, but let's call it what it is, shall we?)

I'd like to take a moment to point out that this list has ALWAYS gone to my personal account, and not my AGIS account. I'll let you figure out exactly how well my "resignation" meshes with the First Amendment and the idea of "privacy".

COOK Report: Sounds like Peter can't delegate any authority. Don't think it was that way at Net99 before the buy out.

Siegel We had a good deal of empowerment to make decisions on small amounts of money. The goal was to make the customer happy to the best of our abilities. If it was hardware we needed, we asked him for it, and he got it for us as soon as he could.

Cook Report: Even then Joe Stroup had driven himself to the point of total physical exhaustion.

Siegel: True. At one point, he had fallen asleep at the keyboard at home, pinched a nerve in his neck, cut off the blood supply to his arm, and damned near lost usability of his arm forever... He was definitely pushing the envelope for a while there.

COOK Report: Kline and Lawlor are trying to run possibly a larger organization very likely with FEWER than Joe's 17 employees.

Siegel: Yep. I don't envy their position.

Jim Butt: Have you seen any Net99 ads lately?

Siegel: Have you *ever* seen one? The only advertising done was Joe Stroup on com-priv, inet-access, and other associated "spamming" in news groups.

[Editor's Note: Dave's memory is faulty. For a period of roughly three months they did advertise in Internet world and other Internet publications.]

Butt: How many of the Net99 employees were sales?

Siegel: 3 plus Joe Stroup.

Butt and how many were techs?

Siegel: Me, Pete Rusinek, Brian Pettin-gell, Joe Charbeneau. For provisioning we had Mary Lynn Jamison, and Pat Pederson. (They got router and CSU's installed by MFS, scheduled remote-hands, made sure the T1 was where it was supposed to be, etc.) We had Mark Holmes on UseNet News part time... We had Jot Powers on security of in-house Unix machines part time...

Butt: How many Circuits did net99 run up per day or week?

Siegel: We were turning up between 3 and 7 a week.

Butt: How many is AGIS doing?

Siegel: No idea. Sorry. . . . I don't think *any* of us are arguing that. We all want to see prices stay at "reasonable" levels. Personally, I hope that AGIS get's things fixed, and get's people happy again, and that more like AGIS get started. This will not be a bad thing for any of us.

February 9: **Howard D. Leadmon** As most people here are, I am a Net99/AGIS customer. When I had a problem with the network several months ago, I contacted the NOC, I sent repeated Email about the problems, and also sent many repeated messages to Peter. Every time I spoke to someone about our problems with the AGIS/Net99 network, I was flat out told that I would have to talk to Peter. I have in many conversations asked for Peter to contact me (even left my work, home, car, and pager numbers). I also sent Peter repeated messages asking to be contacted about some of the problems/issues we have with our connection.

Bottom line, it has been several months now, and never once has anyone at AGIS/Net99 done a darn thing to help with any problems we are having with our link. Also never once has Peter contacted me to discuss our problems. So I can believe he is busy, I know how many hours I usually work every day, but to not contact a customer in over two months, this is total crap!

OK, since this is directly related to ME, let's address this here in the open forum,

since everything else has failed (never received a response). First you say the Net99 network is overloaded where I am connected, and that I need to transition to the AGIS network. So I ask the following questions. What am I supposed to tell my customers who have been calling for over two months complaining of lousy throughput and heavy packet loss?? Remember this has been going on for 2-3 months now! When will this be corrected, as nobody at AGIS has any idea as to when this so-called conversion will take place.

Second, when I have been talked to about the Net99 to AGIS backbone I was informed that I would be converted from my point-to-point link over to the wonderful AGIS backbone that is running over Frame Relay, and my port would be assigned a 256K CIR. Now I ask you, since when is moving from a dedicated p-to-p T1 connection to a 256K CIR Frame Relay connection an UPGRADE! Only AGIS could try and sell this to someone. . . . Now on to my third and last point. Our throughput over our T1 connection has been so bad most of the time over the past several weeks that my throughput has dropped as low as 100 kbps MAX over my T1. . . .

Bottom line, things are so bad that I have now had to have a second T1 from another source connected at our location. So now that I have two T1's, maybe things will improve by routing around AGIS. While working on getting this other T1 installed, I told AGIS I now need to use BGP4 so I can perform our best case route announcements to the network. What am I told, "Baltimore can't support BGP4, you will have to wait for the upgrade to the AGIS network."

This is Not Backbone Service

February 10, Dan Maus: This is *not* backbone service. This is "a few frames now and then". Where the f*** is my \$1650 per month T1. I have a dialup link 12 hops back on MCI that can outrun this joke that I pay over \$2400 per month for. 400 ms turnaround is a Zoom 14.4 modem, not a T1. I just threatened to go full occupancy on my pipe with icmp. Johnl@agis.net said "Peter will probably turn off your icmp if you do that".

I *PAY* for service, service which you are once again denying me, while still charging me full price. I *PAY FOR ALL PROTOCOLS AT ALL TIMES*. 70% frame loss, how about a 70% reduction in link charges, plus you pay 70% of my leased line charge. Except even that is a

sucky deal for me because the 30% I'm left with is unusable, and effectively worthless.

[Editor's note: Feb. 10 also brought news of another AGIS staff departure.]

Thomas Earl: I left AGIS to seek employment elsewhere. But Why? AGIS is a company the is growing at an alarming rate and unfortunately they are SERIOUSLY understaffed. Phil and Peter gave me the opportunity to become management, but I turned it down for many reasons. I am not totally out of the AGIS loop though, I work for one of their customers now.

Feb. 11, **Thomas Earl:** Sorry, I didn't realize how bad that sounded. I left because I was burning out, losing my relationship with my girlfriend, and loosing sleep, all because they are understaffed. I might still be there if they had gotten some people in there back when they originally said, BUT I want to state that AGIS is NOT a bad company to be a customer of (or an employee of), well as long as you are on the AGIS backbone and not the NET99 bone.

Michael Lucas (mikel@agisgate.agis.net) Yes, we are understaffed. AGIS is currently accepting applications and the staff is actively encouraged to send in the names of good recruits. I was hired through such recruitment -- Amanda Robinson gave Peter my name. Peter sought me out. I had an interview two days later. I could have started the next day, but I gave my old employer the customary two weeks notice. They're trying very hard to get qualified people. I like the people I work with. I was friends with Amanda before I started here, so I'm somewhat biased in her case. Tom was cool, and we remaining NOC technicians get along well.

In emergencies, I have woken Peter scant hours after he's finished an eighteen-hour day. He has always been perfectly polite, and promptly taken care of problems. He *always* says "thank you" when I wake him at some ungodly hour. Most people couldn't be that polite. I know that if someone called me after such a day and said "Here's a major headache, and it's all yours," I'd say a whole string of four-letter words and seriously consider hanging up the phone.

Since Amanda's departure, I have become DNS admin. (I assisted Amanda with DNS before she left.) DNS is a very complex subject. I used to work in a library, however, and DNS is far simpler than the Library of Congress' AACR2 cataloging rules. At the moment I'm trying to straighten out some glitches with reverse DNS, such as what Dan Maus complained

about. In the next few weeks, I will be solving the problem by building new nameservers from scratch and moving customers onto them, but a solution simply can't wait until then. We're all working as hard as humanly possible.

It's easy to burn out under such conditions. We're pretty much allowed to work as many hours as we like. Every individual must decide for himself how much is enough, however. I set a pretty firm limit of 50 hours a week (although I must admit, a few hours extra overtime is a nice option when my fiancée's birthday is coming up).

February 12 **Philip J. Lawlor:** Folks, I have been watching this list since the first AGIS email address was added. The current demands on my time are such that I cannot respond to each of the threads on this list.

What I will try to do is give appropriate information to this list as I can. I doubt that I will take the time to respond to any particular point or counterpoint, but rest assured that I will take the time to read them.

For the record: The financial and ownership information given to this list is neither confirmed or denied. I am the Chairman of the Board and President of AGIS. AGIS is a privately held corporation, and as such, does not disclose financial and ownership information. [Editor's Note: We stand by the information we have provided.] I have not seen any information published here, or anywhere that is entirely accurate. AGIS' customers can rest assured that AGIS is solvent and meeting its financial obligations.

AGIS is a great place to work. Several employees have, after working here just a few short months, increased their market value in multiples. As an employer, I am sorry to see anyone leave the company, however, as it says on our web page, we don't compete with our customers. In some cases, though, our customers do compete with us. One thing that our customers can't do is offer their employees AGIS stock. Several of our employees already have earned this valuable asset, and more will, I am sure.

AGIS has over 250+ dedicated circuit customers. Our satisfaction rate is high, however we continue to work on raising our quality. I again reiterate that I probably won't get involved in any individual thread. AGIS has never had a policy of "Don't post on public lists." It does have a policy of "No flaming." As far as this particular list goes, I am going to make sure that all AGIS employees are subscribed. I need to go for now.

February 14: **Siegel:** [Peter Kline had cited among the possible reasons for Amanda's departure: "the divergence of the in-

terests of the employee and the employer."]

This one case leads to the occurrence of the other nasties that you mentioned, and is by far the most avoidable of every problem you have listed. Is it, or is it not, the responsibility of the employer to go the extra mile to see that the interests of the employer become the interests of the employee? Such is the essence of inspiration. Without inspiration, little can be accomplished. If you've done your best as an employer to inspire and meet the needs of your employees, and their interests still diverge, then by all means, they belong somewhere else.

Brian Pettingell: I agree with Dave, but I would also like to add that there was something missing from the AGIS "work environment." I can remember working for Dave and I can remember working for Peter Kline. Here, IMHO, was the major difference. I guess first I should note that I worked for Dave Siegel as a Network Engineer at Net99 and that I left Agis/Net99 2 weeks after Dave. Now, here was the major difference. Dave took his inspiration, his "love" for his work and through TEAMWORK and leadership, instilled it in everyone that worked for him. He also did something that I have not had many employers do before; he never put himself above anyone else. He never tried to be the BMOC. Dave and I would often stay up late into the morn working on various projects and he would usually come down the hall from his office to sit in the NOC with us and interact on a (for lack of better phrase) junior engineer level. Dave could route circles around us and troubleshoot problems faster than any of us, but he never made us feel like we had to respect him for that, rather we naturally respected him for the way he treated us, and the way he made us all feel like we were wanted and needed as part of a well oiled machine.

Now without saying anything negative towards AGIS or Peter Kline, lets just say that IMHO, AGIS was lacking the "work environment" that I just described above. IMHO, because I can not speak for all the other Net99 Techs, I believe this to be the major contributor to the leaving of all the Net99 Tech staff. I say this only because in retrospect to the whole AGIS/Net99 buy out, the Net99 Tech staff, although now split and working for different companies, still communicates on a regular basis and if you asked anyone of us, we would still to this day tell you that we are a TEAM. I think that if you asked Amanda, she would probably not say the same about her experiences.

A Hypothesis to Explain Events that Don't Otherwise Make Sense

On March 5th we published the following to the agislist: What follows is admittedly hearsay evidence. I am not naive enough for a moment to think that it is substantial

enough to be proven in a court of law without first hand testimony and supporting documents. All of what follows however, I either have direct knowledge of - having heard it directly from the principals involved - or it has been told to me by people who themselves participated directly in the events described.

1. Joe Stroup's efforts to sell Net99 began to blossom very early in 1995. Within about four to five months after he founded the company.

2. As early as May/June 1995 time frame Joe began complaining to me about MFS's "poor service." His litany of complaints about service and their alleged business dealings grew and grew until it became a case of: why should they be paid for doing NET99 harm?

3. People who know Joe's record over a period of many years claim that in the past he has left major telco's with unpaid bills.

4. The letter of intent for the sale was signed on Sunday. The sale was consummated the following Weds. Joe had been up all night long Tuesday to wednesday and early morning hours was pessimistic that the sale would go through. Roughly 11 hours later it was a done deal.

5. It is very likely that, with the existence of an unpaid bill to MFS, the discovery proceedings were such that by Wednesday the AGIS buyers were well aware of the "problem". They were very definitely looking at NET99 payables and receivables.

6. If there were an unpaid MFS backbone bill, why didn't AGIS pay the bill in full? Well, it is known that the NET99 backbone had been built on the cheap and technical advice to AGIS was that it had only a few months to go before it would have to be re-engineered at great expense.

7. Therefore there may have been no monetary incentive to AGIS to pay the NET99 backbone bill. The incentive presumably was to buy NET99, acquire its customers, disassemble its staff, and, possibly, to ultimately put what was left into bankruptcy.

8. AGIS reportedly had, at the moment of the buyout, 17 customers. Reportedly their backers were expressing grave concern as to: why so few?

9. At one stroke with the expenditure of only \$800,000 with an additional \$200,000 to be paid by November 1 (the second installment reportedly was never paid), AGIS acquired 120 NET99 leased line customers with an additional 40 or so installs on order.

10. They did not merge the companies (although some early bills were paid by Net99 customers either to AGIS or AGIS/NET99). Recently bills from the net99 side of the business have by paid solely by

NET99.

11. The intent - based on circumstantial evidence - seems to have been to transfer Net99 customers asap to the agis backbone and close the net99 operation down.

12. My conversations with Phil Lawlor at the end of August were marked by a great deal of anxiety that I knew as much about his newly acquired company as I did. Phil seemed to me to be excessively paranoid.

13. Congruent with this strategy AGIS allowed an environment to exist such that the key Net99 technical staff left within 6 weeks of the buyout and that almost all staff were gone by year's end. The number reported still in the Phoenix office ranges from 5 to 3 with 3 being, I believe, the more reliable figure.

14. From the point of view of customer satisfaction and a reasonably well operating backbone NET99 was a viable concern. Financially viable is another question and one of which I have no direct knowledge.

15. *Is there or is there not an unpaid bill to MFS?* I queried MFS myself more than a month ago. The answer I got was something along the lines of we can't comment on that. It was certainly not: their (NET99's) accounts are all in good standing.

16. I have been told by someone who worked there (AGIS - NET99) and claimed to have seen it with his own eyes that there was a large unpaid bill.

17. My point is that what has happened SINCE the buyout just doesn't make sense, in the *absence* of the hypothesis about the unpaid backbone bill. Given the hypothesis -it makes a great deal of sense.

18. If my chain of reasoning is accurate - how long before NET99 is placed into Chapter 11?

AGIS employees flamed us without mercy for this posting but challenged none of our specific statements. AGIS customers while publicly silent told us privately that they thought our comments reasonable.

Announcing the Demise of NET 99 DNS

On March 10 Peter Kline announced Effective April 2, 1996, ns1.net99.net will be retired. 204.157.3.2 (and 204.157.3.3) are going away. Effective immediately we have shortened the expire on all zones ns1.net99.net handles to one week, and beginning 10 days before April 2, expire will be shortened to one day. April 1st, expire will be shortened to one hour (no joke). April 2, ns1.net99.net gets strapped to a hay bale and thrown into that water aqueduct thing that goes through Phoenix that you can see from the airplane (pretty amazing when you think about it: the river

beds are all dry, but the water doesn't evaporate from the canal. No wonder everyone drinks bottled water there).

Also on March 10, **Robert W. Foster** commented: Peter, You act as if this is a good thing. You've taken a perfectly good system and bastardized into a giant frame relay mess!!!! I hope I speak for all of us that have had good service and excellent throughput on Net99, I truly wish you would strap The AGIS network to the Net99 server when you throw it, and I hope the water is deep enough, (oops, of course AGIS will probably float), so please hold on to the rope! If I could find a decent packet switch in your entire network, I'd beat you with it!

March 14 Gas Main Explosion and Outage

At 2:29 pm Pacific Time on March 14 Peter Kline posted the following to the agislist. At approximately 12 Noon Pacific Time this afternoon, a gas main exploded near the WilTel POP in Santa Clara, CA. The fire department has evacuated the building, cut the power, and has forbidden WorldComm from starting their generator. Our equipment ran for about another hour before the UPS's ran down. There also is no air conditioning. It appears that our site is completely out of service at this time.

All customers directly connected to Santa Clara are down. All customers attached to Seattle are down. The rest of the net is affected to the extent that a) some routes visible only on the west coast are now unreachable and b) the rerouting of the rest of our traffic to Chicago, Trenton, and Washington will cause congestion there. There is no estimated time to repair. You may, however, check status on our web page, or call and ask for ticket 609.

On the evening of the 14th Robert Gourley of Worldcom wrote: "Allow me to tell the LDDS Worldcom side... A gas main did explode near the POP. The fire dept. turned off the power in the surrounding area, and would not let us start our generators. We do however have battery backup, which carried our systems for the entire duration of the incident.

AGIS Had No Battery Back Up

The only outage was to co-located customers who do not provide their own battery backup. Colocated customers, like AGIS, have several options for powering their equipment. They can pay to use our battery backup DC (if their equipment provides for DC power) or they can run off AC and provide their own UPS systems. As a courtesy, I believe we will provide them AC from our generator when it is running. We don't provide them AC via our battery backup. I be-

lieve this is because AC inversion is inefficient.

During the incident today, only customers which didn't provide their own battery backup were affected. This amounted to about 6 circuits total. The duration of the outage for these customers was 2 hrs 17 minutes. Local POP techs worked closely with the fire dept. to restore power as soon as it was safe. Customers which choose to operate co-located equipment with any sort of battery backup are running at some minimal risk.

The same evening Jim Dixon wrote: Our circuit to Agis in Santa Clara was down for about five hours. When I spoke to people at the Agis NOC, one of their lines was "it's not just YOU, there are a lot of people down!" And MFS said that they were flooded with trouble calls. We were still getting the same story from Agis at 7 pm California time, about 7 hours after the explosion.

AGIS Knew

But what I didn't really understand until after a long and sometimes heated conversation with Peter Kline and his boss was that Agis KNEW that all of the circuits out of Santa Clara were going to go down for AT LEAST AN HOUR before they did go down. Their equipment was running on [UPS] battery backup which failed at roughly 2 pm. *NO ONE AT AGIS HAD THE COMMON SENSE TO TELL THEIR CUSTOMERS.* What does it take? Five man-minutes to email the agis customer list and say 'we are running on battery backup in Santa Clara, we will be down at 2 pm or so'? Nothing. Not a word. Our notice from Agis was a red light on the CSU/DSU.

We learned about the explosion from MFS and the very helpful people at Net99 in Phoenix; they spent quite some time trying to find out what was going on after I told them that we had been calling Agis unsuccessfully for half an hour or so. They said that they had many telephone calls from other former Net99 customers whose circuits were out (and couldn't get any joy out of Agis).

Peter Kline eventually sent out a notice, but at 5 pm his time, when none of the affected customers could possibly receive it. And oddly enough that notice was addressed to agislist, so that people who DON'T subscribe to what he calls the "Agis sucks" list have been told nothing. Perhaps this means that Peter believes that the entire Agis customer base has subscribed. I could understand that belief.

We called Agis repeatedly only to get a busy signal, an answering machine, or occasionally NOC staff who seemed to be completely incapable of explaining what

was going on or returning a call. We did spend about two hours total on hold later in the day. "Just a sec!" they would say. Thirty minutes of dead silence. Hang up and redial, to get the same treatment.

We are primarily a European network, but we have started investing heavily in new US facilities. We need reliability. Things will go wrong. But problems can be handled a hell of a lot better than this.

Peter Kline asked me this evening if I thought that he blew up the gas line in Santa Clara. No, I don't think he did. But he knew about it a long time before I did, and he didn't tell us, so we couldn't warn our customers, and they couldn't warn theirs. And no one at Agis seems to have done anything to mitigate the effects of the disaster. They knew it was coming, but they did nothing, and they said nothing. They lynch people for things like this where I grew up.

An on March 18 the following exchange occurred between Richard Dawes and Jim Dixon:

On Sun, 17 Mar 1996, **Rich Drewes** wrote: ITEM #1: MAJOR MFS OUTAGE Tuesday 3/12/96: Not much to say about this one, except "ouch". The MFS outage brought me (as well as several other non dual-homed local ISPs) down cold for 8 hours.

Dixon: Have you read Alice in Wonderland? I hope so. We saw hours of downtime on Tuesday, Wednesday, Thursday, and Saturday. This was blamed on MFS, Wiltel, the gas explosion, attacks by flood pinging, you name it.

Drewes: ITEM #2: WILTEL EXPLOSION-RELATED PROBLEMS SINCE THURSDAY 3/14/96 Something exploded up in Santa Clara and it's screwing me up bigtime here in San Jose. What do the rest of you see?

Dixon: I don't want to be rude to anyone, but to the best of my knowledge, the current troubles have very close to nothing to do with the explosion in Santa Clara. There WAS a spectacular gas explosion. When I heard this story I was exasperated and frankly didn't believe it, so I drove over, looked around and talked to people there. There was NO physical damage to equipment at the Wiltel PoP. The explosion was by the entrance to the parking lot; Wiltel (and Agis) are at the far end of an untouched nearby building. A tree was badly scorched, but I saw no signs of any critical Agis connections to it.

Drewes: According to Chris, Peter Kline and the CEO were meeting tonight to discuss the situation. I could not get an ETA from Chris for when the Santa Clara situation, and presumably therefore my own situation here in San Jose, would be fixed. The AGIS NOC WWW status page has not been updated in three days at the time of this writing and still shows

that the Santa Clara POP is totally down. Are Santa Clara customers still down?

Dixon: No. We connect to the Santa Clara router and were back up after 'only' five hours or so. As I understand it, other Santa Clara customers were up faster.

Drewes: Chris mentioned that the remaining problem at the Santa Clara POP--the problem somehow triggered by the explosion yet not directly related to it--has to do with a Cisco instability.

Dixon: We couldn't possibly comment on these interpretations ...

Drewes: Anyone have more information on these events, please send it here.

Dixon: This is an interesting and relevant traceroute: [Editor: traceroute omitted to save space.] That is from San Jose to San Jose. Via LA, Washington DC, and that old unreliable trashy etc Net99 backbone, you know, the one that works.

We have been running a lot of traceroutes over the last day or so, and getting similar meandering results. To get to www.tlg.com (TLG attaches to MAE West about 100 yards from here) we go via Seattle; just 2000 miles or so out of the way. I think that we reach most MAE West-attached sites by the LA-DC-Net99 route, but what's another 6000 miles? Packets to sites that we reach through the PacBell San Francisco NAP, which used to be about 9 ms away, now take a little detour down to LA first and come back up to the NAP courtesy of ans.net.

It looks to me like the Agis Santa Clara router has lost its connections to all of the Northern California peering points: the CIX, MAE West, and the SF NAP. Most of the connectivity appears to have been lost on Saturday, when the Santa Clara router was upgraded. We don't pay any attention to suggestions that someone damaged a board when they were putting the Cisco 7000 back together.

During mid March an AGIS customer in Florida that was connected through Net Runner, cancelled his service with AGIS. He alleged that AGIS then 'maliciously' transmitted false routes that threatened the viability of his second connection with MCI. For several days he and Peter Kline called each other liars over the network. AGIS also began to announce to former customers that it was reclaiming their address space. While AGIS was had the legal right to do this it went about it in an underhanded way stating falsely that it was the interNIC that was forcing AGIS' action. It also asked for the return of address blocks holding up to 16,000 hosts in 21 days. Clearly an insufficient amount of time to allow those customers to renumber and clearly in violation of the Tony Li Yakov Rekhter draft RFC recommendation of six months.

AGIS' "Capacity"

At the end of February, the *Detroit News* wrote: "Out of a suite of tiny offices riddled with cables, AGIS now provides about 15 percent of the Internet's domestic capacity -- an amount equivalent to the entire Internet just three years ago. AGIS's president, Phil Lawler, thinks he can bring his market share to 20 percent."

We don't think this is a credible figure. Consider routes at MAE East. In November Sprint and MCI were advertising about 7,500 each, UUNET, PSI & ANS in the range of 2,000 to 5,000, Net99 almost 1000 and AGIS almost 100. Or combined at around 1100. We know number of routes advertised is not a perfect correlation as to traffic and size. It is a rough index.

We have not seen anyone with solid figure on how much internet traffic a given carrier has. In the absence of centrally gathered NSFnet backbone statistics, it is impossible to know.

Here are our gut level estimates derived from talking to all the players and keeping our ears to the ground. MCI 24%; Sprint 21%; UUNET 16%; BBN 12% and rising; PSI 10%; ANS 8%; AGIS 3%; NETCOM 3%; Others 3%. We place the on-line services under their respective backbone providers.

The AGIS figure of 15% domestic traffic is simply not credible. Of course the article uses the word capacity. Not traffic. Is AGIS talking in terms of the speed of the backbone times the number of miles? Internet capacity is a strange term and one not generally in use. Why? Because you could buy an OC-3 connection on your own a lot of capacity. On the other hand if you had few customers pumping data through such a connection, we assume you would not be in business long. So lets suggest that what is meant is actually the data generated by customers attached to the national service provider's backbone.

If that is the case then in looking at traffic generated, we'd have to ask about CERFnet which has a T-3 from California to MAE-east? And how about Cable and Wireless which is entering the US market - not as just as a carrier but as a provider with CAIS and is contemplating going into the NAP business? How about the RBOCs and so on? PSI and UUNET as publicly held companies have stated they each have leased line customers numbering in the thousands. They have spent tens of millions on infrastructure. AGIS has said more than 250 leased line customers. We have seen no evidence that it is spending tens of millions on its infrastructure. Readers should compare what is known about the size of AGIS's competitors and draw their own conclusions.

Wither the Standards Process?

Should New Internet Applications Standards Be Developed within the IETF?

A Dialogue Between Peter Flynn and Tony Rutkowski

Editor's Note: as the commercial Internet expands, new companies like Netscape can bring hot new products to market very quickly. Public offerings generate new companies with hitherto unheard of market valuations - sometimes in not much more than a year after the company first comes onto the scene. In this gold rush atmosphere, the open and seemingly slow process of IETF standardization may seem inadequate to some. With the World Wide Web the engine that is seeming to power the Internet's continued expansion, there has been the appearance to some that web consortia may be ready to take further development of HTML into less public and hence more proprietary paths.

While we do not follow web developments very closely and certainly don't consider ourselves experts in the standards process, we are aware of the importance of the open IETF standards process as being the single most important driver in getting the internet to the stage where an application like the web could be launched. One of the most important issues that will determine the future evolution of the Internet is how standards development for core Internet technologies and for applications like the the Web will take place in the future. As a result, the following conversation on a private internet writers list caught our attention. We reprint a summary of the dialogue with the permission of the participants.

Who's Driving Future Web Development Standards?

On March 3, Peter Flynn, author of *The World Wide Web Handbook*, wrote: As some of you may be aware, the World-Wide Web Consortium (W3C) has not been having a particularly happy relationship with the IETF (from what I can see), largely centering around the way in which standards like HTML are devised and derived. The HTML Working Group, who have actually been doing the donkey-work, appear to be due for the axe, and the rumor machine indicates that the W3C may well intend to pursue some form of standard on its

own, without the existing [IETF] expertise.

As the W3C is composed mainly of the large or resource-heavy organisations who founded it, there is considerable unease about the direction this work will take. Obviously the pressure from Microsoft and Netscape, among others, will be to use (perhaps competing) proprietary models for the Web, because it will satisfy their need to exert control. [Peter indicated that he would like to hear whether others agreed or not.]

Tony Rutkowski, Director of Internet Development for General Magic replied: Your portrayal of W3C neglects to consider that it is headed by Tim Berners-Lee who conceived the Web, and developed and guided html and most of the basic Web specifications. Tim's vision and energy and determination to develop the best open specifications are a fundamental component of the Consortium, and the core basis for support and funding by most of the Consortium's members. (General Magic is a member, and I'm its representative.

The W3C also has the resources and the institutional home (MIT's legendary Computer Lab and INRIA's equally famous facilities in both Sophia Antipolis and Rocquencourt) to employ teams of people who focus on the development of common open Web protocols and specifications.

Most of the work, by the way, is open for others to see. See www.w3.org

I think what is being misread here is the entirely natural emergence of multiple new specialized standards bodies as the Internet environment has matured and enlarged. Even a few years ago, specialized MIB groups began to spin off. Then the W3C emerged. And of course a number of standards were effected in the ISO. There is now an Internet Mail Consortium, a VRML Consortium, a CUSeeMe Consortium, a Financial Services Consortium, among others. Some of this work also finds its way into the IETF, much of it

not, as the IETF concentrates on core Internet network technologies. And then, of course, there is standards work represented by companies moving ahead on their own with standards and specifications, drawing multiple companies into de facto open standards forums, and migrating the standards into the various standards bodies. Java (Sun), Telescript and MagicCap (General Magic), VRML (Silicon Graphics), HTML extensions (Netscape), are several examples among many.

Standards processes are an inextricable component of a larger, very complex environment of open markets, product development, individual initiative, and public policies. Simple characterizations fail to account for this complexity and the tradeoffs.

Flynn: The problem doesn't seem so much to be the politics of location or who "houses" whom, as the reluctance of the corporate members of the W3C to accept that standards are not made in a day, and that a public standard in particular needs public debate. In an interview in Personal Computer World a bit more than a year ago Tim Berners-Lee put it succinctly: "The ideal thing is that [standards controllers] allow any changes that have to be made to the standard to be done in the light of public discussion, in an open process." [Unfortunately,] the companies who are driving this are doing the same as most newcomer companies do on the Net: they expect ready-made solutions for free.

There are of course some companies who _have_ put significant resources (staff time, travel time and money, and equipment) into the effort, and they need applauding for so doing, but these are the minority, and do not include the large players, who (perfectly understandably) want the job done for free.

Now the W3C criticises the IETF development methodology as well, from what I hear, so [I fear that] we are faced (as predicted) with the withering of HTML to the level of hobby toy, and the rise of a new, shiny marketing tool for remote-control PowerPoint, rather than the information system which was originally

intended.

Rutkowski: [I have a different opinion.] I just returned from Paris and the meeting of the Int'l WWW Conference organization - which you will be pleased to know is remaining independent and incorporating as an international organization in Switzerland. Several INRIA staff who were newly hired for W3C were present, and I'm impressed generally by the competence and change of attitudes in the French networking community. Even in the past, however, some at INRIA have long been Internet leaders like Christian Huitema (former IAB chair), and Jean-Francois Abramatic (director of Internet development).

Flynn: That is excellent news. Christian I know well from my RARE days, one of the best people at Sophia Antipolis then. In fact it was his work for the IAB which was one of the major spurs to the ditching of X.25 and the rapid *volte face* which both RARE and DGxiii undertook.

Rutkowski: Some of the best Consortia staff papers, by the way are being introduced at the 5th Int'l WWW Conference at Paris in May, among the more than 300 papers received. The shifting focus in Europe toward the Internet is quite dramatic, and will be exemplified in the workshops, exhibitions, and developer sessions at the Conference. And, would you believe, the European

Union is sponsoring a one day event on the use of the Internet for small and medium business enterprises. Lastly, even France Telecom is expected to be previewing some highly innovative new Internet offerings.

Flynn [what concerns me is] the ability of these bodies to disengage their normal corporate (read: marketing) *modus operandi* and enter into partnership with the rest of the network, rather than try to dominate it by attempts to impose ill-considered "solutions" of their own private concoction.

Rutkowski: I share your concern and perspective here. The strength of the IETF paradigm is its function as an open development and technology transfer tool - fostering alternative perspectives, innovations, blunt criticism, operating code, and rough consensus among developers. The "standards" function is arguably just a secondary byproduct. Smart, strategic corporations realize this, and the rest are learning. The better consortia are also emulating.

Cook Report: Tony, please elaborate on what you mean here? Are you excluding IETF standards as a desirable IETF activity? Leaving them to other consortia?? Do you advocate that consortia take over what IETF working groups develop that may fall into contentious debates, arguments etc and have difficulty reaching IETF

standardization? what do you mean by *secondary* by product?

Rutkowski: My remarks referred to function rather than form - and go to the heart of why Internet standards activities, and those of the IETF in particular, have generally been far more successful than those of other standards bodies in the IT field.

ATIS, by the way, is not an obscure standards body, but the primary telecom standards organization in the U.S. - and the feeder body to the ITU-T - having been established by the FCC at the breakup of AT&T to undertake these functions. It is, however, a decidedly different realm and lacks most of the IETF's functional characteristics to which I was referring.

Standards are still a desirable activity.

COOK Report: Your use of form vs function is still not clear. Do you mean you consider the process leading up to the proposed standard more important than the standard itself??

Rutkowski: Yes, in part. However it also goes to the larger question of "what's happening" at these fora. My suggestion was that although the resulting standards were important, it's the collective brainstorming by diverse parties who freely provide ideas, criticism, code, and consensus that's the most important.

AT&T Misses the Internet Market?

BBN Will Teach AT&T But It Will Take a Long Time for Them to Learn

We greeted the AT&T WorldNet announcement with a yawn. It neither means the end of the small ISP nor it did not mark a major date in Internet history. It did dramatically demonstrate that AT&T is fundamentally clueless about the nature of the Internet.

AT&T did do one thing right. In partnering with BBN it has an alliance with a National Service Provider with a cadre of talent that only MCI can begin to match. BBN is the player who better than anyone else can understand and possibly deliver on the automotive industry's needs to put as many as 100,000 parts suppliers on the internet with guaranteed quality of service. Only problem is that the build out re-

quired might take a couple of years and many hundreds of millions of dollars. BBN has the know how but not the capital nor the fiber. AT&T lacks the know how but certainly has both the capital and the fiber.

What is at stake is the building of a vertical industry market that accounts for one seventh of all US employment. Once the way is shown for the automotive folk, sources tell us that petroleum industry and then the banking industry will follow. We are talking hundreds of thousands of commercial, dedicated accounts demanding high quality of service and paying monthly connect fees of a minimum of \$1,000. The average would likely be \$2,000 a month. Three hundred thousand such accounts

would generate \$600 million dollars a year. \$900 million if the average turned out to be \$3,000 a month.

But what market did AT&T pick? Why the low margin over crowded dial up market of course. The one served by AOL, Prodigy, CompuServe, MSN and over five thousand independent internet service providers. There are perhaps 8 million dial up accounts averaging \$20 a month. 160 million dollars a year. Lets say AT&T gets a million customers in the first year. Twenty million dollars in revenue. Yawn. Not very impressive compared to a \$600,000,000 market it neglected to jump start and there by to own.

The dial up ISP market is served far

cont'd p.24

Executive Summary

ATM: Grand Unifying Technology or Brain Damaged Transport Product? pp. 1-3

ATM is a level 2 connection oriented transport technology composed of 53 byte long cells. As such its architecture does not mesh at all with connectionless TCP/IP and its variable length packets. Nevertheless we find a heavy interest in ATM among the major phone companies which have invested billions in the technology. One reason appears to be the desire to multiplex switched traffic between backbone nodes of a network before sending it else where. Another is to bring Switched Virtual Circuits (SVCs) to market. These are connection-oriented circuits between users that are set up and torn down by software upon command.

Reserved, on demand bandwidth will depend on SVCs - or possibly on an IETF "developing" protocol: RSVP. The National Science Foundation has just announced a new high speed Connections Program to develop the capability to define and reserve differing priorities of data service.

Ironically Ethernet is as the last mile to the desktop is seen as a major barrier to SVCs. ATM and Ethernets are not compatible. However, we report on Cells in Frames, a project at Cornell that is developing an inexpensive ATM "attachment device" that will drive a 10 megabit ATM connection to a workstation from a 10-BASE T Ethernet hub. Converting LANs to this technology will permit merging of PBX phone lines with the ATM network.

Meanwhile, in wide area backbones, ATM suffers from various overheads, referred to by its detractors as a Cell Tax, that according to a Minnesota Supercomputer Center report would drive OC-3 155 mega bit bandwidth available to IP down to 116 mbits per second.

Resource Reservation Impacts, pp. 4 - 6

Noel Chiappa, in an interview with

the *COOK Report*, discusses two competing philosophies of network design: resource reservation and over engineering. Chiappa: "What ever the future is will be greatly influenced by the answer to that question of yours as to "whether you do or do not need resource reservation?" If you decide that you need it, the kind of solutions that you look at start out by looking very different from the solutions that offer themselves in the absence of a resource reservation effort. The answer to these questions will also drastically impact the kind of switching architecture you adopt."

InternetMCI Stakes Future on ATM pp. 7 - 13

In interviews with Steve von Rump, MCI Data Services MarketingVP, and Steve Tabaska, MCI Data Services Engineering VP we discuss in detail MCI's backbone expansion and its transition to ATM.

MCI expects to be able to do SVCs desktop-to-desktop before the end of next year. In the meantime, it will open a switched ATM OC-3 backbone fabric by the end of April 96. It has changed from General Datacom switches to Fore ASX 1000 switches and is starting with 10 of these \$100,000 a copy top-of-the-line switches.

Tabaska expressed doubts about technology arriving at the NAPs that would allow them to scale adequately to permit multiple peering sessions at bandwidths of 155 megabits per second and higher. MCI, he said, is beginning to rely in part on private exchanges with one or two other large NSPs in places other than NAPs.

By this time next year MCI anticipates running OC-12 on their backbone. To do this OC-3 cards in the Fore switches will need to be replaced with OC-12 cards and Cisco will have to come out with an OC-12 interface for its routers. They have looked at the NetStar Gigarouter by are not quite ready to go with it. They plan some months from now to run a dual architecture of PVCs to handle things like telnet, and DNS lookups and SVCs to allow reservation of bandwidth on demand.

Gartner Miscue, p.13

We repudiate the Gartner Group's misleading appropriation of parts of our

web page glossary for an advertisement in the Business section of the March 24 *New York Times*. Gartner did not ask permission to use our material which appeared as an indistinguishable part of a poorly done glossary of their own making.

Rough Road for AGIS Continues pp.14 -20

AGIS has been consolidating NET99 customers into its infrastructure and will we believe soon close Net99 down. In the meantime the Net99 customers are extremely unhappy and continue to make their unhappiness well known on the agislist@interstice.com. We present a summary of the discussion from early February through mid March. We are harsh with AGIS because they bought the one provider which came into existence to help the small ISP and have effectively dismantled it. They appear to be unresponsive to their customers who also have complained about their business practices. Readers may remember that AGIS was the company which, last October, said it would use lawyers to impose order on the internet. We would not like to see the Internet reshaped in its image, and will continue to confront it editorially with the reality of its own making.

Tony Rutkowski & Peter Flynn Talk about Standards Issues for the Web pp. 21 - 22

Many Web standards are now being developed outside the IETF. Flynn expresses concern. Rutkowski shares a bit of the concern, but explains why he is generally optimistic.

Impact of AT&T's Dial Up Internet, pp. 22, 24

As Kent England said "BBN will teach them (AT&T) how to run a commercial Internet service. But it will take a long, long time in Internet Years for them to learn." We wonder why AT&T entered a low profit and already crowded mass-market rather than focus with BBN on the industrial grade quality of service market that the auto industry would like and BBN could help deliver. Dial up service nationally doesn't scale when compared to service that good small ISPs can provide locally.

continued from p. 22

more cost effectively by local independent companies. AT&T with its many levels of management, high overhead, and hefty salaries can't compete. And if it thinks it can out source customer support into the hands of telemarketing types paid \$7 to 8 dollars an hour, it is very much mistaken. Internet customer support is not anything in which you can train someone adequately in a week's time.

On March 13 Kent W. England ex-director of CERFnet posted to comp-priv the most shrewd appraisal of the AT&T move that we have seen. "I have," he wrote, "some empirical evidence regarding AT&T's success likelihood. As one of the first BBNers in the door at AT&T in April 1991 to pitch the idea of a commercial Internet service, I saw first-hand how AT&T's internal organization approached the Internet. At that time it wasn't pretty."

"My most distinct memory was of shell-shocked Bell Labs researchers wandering the halls trying to figure out the Internet. When I asked them about their backgrounds they would say things like "Up until last month I

designed VLSI circuits and now I am on the Internet development team." AT&T turned BBN down in 1991 but came back four years later to ask BBN if they wouldn't please try again. With a new deal-maker CEO fresh from IBM, BBN made the WorldNet deal."

"Today, I note that Steve Bellovin of Bell Labs is working full-time on Internet issues for AT&T, so they finally figured out who in Bell Labs understands anything about the Net. And BBN will teach them (from the outside or the inside) how to run a commercial Internet service. But it will take a long, long time in Internet Years for them to learn."

Perhaps there is a ray of hope? On March 26 we saw the first detailed evaluation of AT&T's software from Chris Locke <clocke@panix.com>. Chris found it very user friendly - evaluating it as far superior to anything he had seen before in the general internet consumer Market.

1996 COOK Report Subscription Rates

1. Individual - now in ascii form only - \$85
2. Individual - hardcopy via first class US Postal Service - \$150
3. Academic Libraries and Higher Education - \$175
4. Small business, small government agency (hardcopy) - \$200
5. Large government agency or business over \$10 million a year gross revenues (hardcopy) - \$350 (price unchanged)
6. Corporate or government agency site license (same as 5) except subscribing business receives desktop published hardcopy and electronic copy and permission to redistribute either or both to as many of their employees as they will. We will email directly to an addressee of your choosing or an internal mail reflector. We also encourage you to reproduce the hardcopy and route internally to as many people as you wish. - \$650
7. Deluxe Corp Site License - same as 6 with added right to place electronic text on employee only corporate wide web server - \$1,000. see <http://pobox.com/cook/>

Gordon Cook, President
COOK Network Consultants
431 Greenway Ave
Ewing, NJ 08618, USA
Telephone & fax (609) 882-2572
Internet: cook@cookreport.com

The COOK Report on Internet
COOK Network Consultants
431 Greenway Ave.
Ewing, NJ 08618