Active Routing

N. F. Maxemchuk, *Fellow, IEEE* and S. H. Low

***Abstract*—Active routing permits individual customers, network managers, or network owners to control the paths that their data takes through the network. The objective is to allow routing mechanisms that provide quality of service (QoS), mobility, etc., to be quickly deployed, without waiting for standards, and to allow different routing mechanisms, that provide similar services, to compete. The current work on label switching (MPLS) can also be used to give high level customers, such as virtual private networks (VPNs), more control over their paths. We show how active routing can extend the capabilities of MPLS. We address several implementation issues, including pricing and distributed sandboxes. Pricing or policing must be used to limit the resources that customers acquire, in order to encourage them to use network resources economically. Sandboxes must be used to limit the resources that the participants acquire, in order to limit the harm that they can inflict on other participants. Active routing creates a free market system where network providers compete to sell their resources and implementers compete to sell their active routing programs. We establish a framework to quantitatively compare networks and service providers. As an example, we route Internet protocol (IP) telephony over combinations of circuit and packet networks.**

***Index Terms*—Active networks, IP voice, pricing, routing.**

I. Introduction

PROTOCOLS are used by communications devices to transfer data. Traditionally, protocols reside in the hardware or software in each device and are standardized so that devices that are manufactured by different vendors can communicate. Active networks reduce the need for standard protocols by transmitting the instructions that execute a protocol along with the data, in a capsule. Every capsule may carry and use a different protocol. The protocol must still be correct in order to deliver the data but the protocol does not have to be a standard.

We propose using the instructions in a capsule to select a path through the network, active routing. The rules for selecting a path can either be carried in every packet so that every packet is routed independently or they can be carried in the first packet of a sequence of packets, so that a virtual circuit is set up for the duration of a connection. With active routing, every packet or sequence of packets can use its own rules. Datagrams and virtual circuits can coexist on the same network, along with new paradigms.

Active routing may be used by several layers of network participants. An individual user may use active networking to select between competing networks. Internet service providers (ISPs) may encapsulate the capsules from its users that must cross regions of the network that are outside its control. By retaining control of its user's path, the ISP is more likely to be able to provide end-to-end guarantees. A carrier may use active routing to aggregate packets that require the same quality of service (QoS) or to route around network failures.

Many of the functions that can be performed with active routing are currently being addressed by standards organization, such as the Internet Engineering Task Force (IETF), using techniques such as label switching. In Section II, we will expand on the advantages that active routing has over the current standards based approaches. In Section III, we will deal with a number of implementation issues, including evolving from the current networks and controlling the network resources in "distributed sandboxes." In Section IV, we develop a model for comparing networks that support active routing. In Section V, we describe applications of active routing. We show how it can extend the capabilities of label switching, and how it can be used for restoration, QoS routing, policy routing, mobility, and receiver controlled routing.

Finally, in Section VI, we tie the sections together with a specific example in which customers make voice calls over combinations of circuit and packet networks. The networks may provide different guarantees on the QoS at different times, and each of the customers may use a different tradeoff between quality and price. We show that a flexible solution, that can use different networks and resources, improves a customer's utility.

II. Advantages of Active Routing

Active routing allows network customers to use different criteria and different rules to select a path to their destinations. The use of a network can change fluidly between best effort and reserved resources as the number of customers who use a particular routing mechanism changes. In addition, new routing mechanisms can be tried without waiting for new standards.

The evolution of routing standards is becoming slower because of the success of the Internet and the number of carriers that are competing to make a profit from the network. In the 1970s, the ARPAnet was a nonprofit network that was operated by the government for a small technical community. Many adaptive routing mechanisms were tested to redistribute the load and improve the performance of the network [1], [2]. New mechanisms only required the approval of a small technical committee. Most packets in the current Internet follow a shortest path [3]. There is interest in using more flexible routing for traffic engineering [4]–[6]. However, changing routing rules requires the approval of most of the competing carriers and manufacturers that profit from the Internet. New routing standards are likely to encounter even more resistance in the future, as the Internet,

Manuscript received April 15, 2000; revised November 1, 2000. This work was partly done during a visit to the University of Melbourne, Melbourne, Australia.

N. F. Maxemchuk is with AT&T Research Labs, Florham Park, NJ 07932 USA.

S. H. Low was with the University of Melbourne, Parkville, Vic., Australia and is now with the California Institute of Technology, Pasadena, CA 91125 USA.

Publisher Item Identifier S 0733-8716(01)01788-7.

telephone networks, cable TV networks, and mobile networks converge into a single networking infrastructure.

Active routing is an alternative to the new standards that are needed to interconnect competing networks. In the current Internet, traffic that must cross competing networks, called autonomous domains, use the border gateway protocol (BGP). BGP can determine the domains that packets traverse but not the path within the domain or the delay on that path [7]. A new gateway protocol is needed because BGP cannot be used to provide QoS guarantees or enforce security policies across networks. In active routing, the capsules carry the rules or criterion that select between the available paths to a destination. The source network can retain control of a path while the capsule is in a different autonomous domain, without having a standard method for specifying packet requirements and without having interdomain agreements.

Active networks allow many routing techniques to coexist and creates a free market environment. Network users can choose between software vendors who use different routing mechanisms to obtain a service. Competing networks, that use different technologies, can provide different sets of resources that the routing programs chose between. New routing techniques and new networks are introduced by the vendors. The routing techniques and networks that can "best" satisfy customers thrive, while those that are less useful or are flawed, go out of business.

III. Implementation

In this section, we discuss several implementation issues. The first issue is the evolution of active routing. We consider both the evolution from current routing techniques to active routing and the evolution from a version of active routing that is economically competitive with the current routing techniques to a version of active routing that is presently precompetitive.

The next two issues we discuss are network sandboxes and whether to use pricing or policing to enforce the sandboxes. We describe the need for sandboxes and how the distributed sandboxes in a network differ from the localized programming sandboxes that are used by computer languages, such as JAVA. We show how pricing can be used to enforce distributed sandboxes and explain how pricing can be implemented without changing how customers are charged for communications.

In the following sections, we describe metarouters that operate between autonomous networks, with similar or different technologies, and information servers that monitor the different networks. Both of these issues must be addressed in order for active routing to make the best use of the current infrastructures and to create the competition between networks that is described in Section IV. In addition, addressing these two issues can create two new businesses. The first business is a super network that does not own any facilities of its own but provides service to its customers by making use of physical networks that compete with one another, either on price or technology. The second business is information gathering facilities that sell information to active routers so that they can make the best use of the networks.

The final issue is an extended notion of network state. We expand the concept of stateful and stateless connections to include the memory that may be used to store active routing programs and use pricing to eliminate the differences between hard and soft states.

A. Evolution

Active networking implies that there is intelligence in the network that can operate on the instructions in the capsules. The more general that we make the language that is used for the instructions, the more flexible we can make the network but we require more processing in the network and have to transfer more information in the program segment of a capsule. There are routers that can switch terapackets per second. We can build programming environments, like JAVA, that can safely perform the instructions in a capsule. However, we cannot construct a terapacket router that can execute a JAVA script for every capsule that is also economically competitive with current routers.

Economic trade-offs involving processing have changed, and will continue to change. Both the ARPAnet and the current digital telephone network were constructed in the 1970s. The cost of memory and processing made the ARPAnet a precompetitive technology. In order to compete with analog technologies, the digital telephone network used 64-kb/s voice coders that were shared by several telephones, rather than more complex 32-kb/s coders. Clearly, the tradeoff between transmission and processing has changed. The current Internet is competitive, even though it uses considerably more processing and memory than the 1970s ARPAnet. Active networking, with a general purpose programming environment, is currently precompetitive. However, as the trend from the 1970s continues, this will change.

Unlike the 1970s packet networks, which could not compete with the older technologies, a version of active networking can be made competitive now. In addition, there is an evolutionary strategy from the simple, competitive versions of active networks to more complex versions, that are currently precompetitive. The types of "programs" that are proposed for active networks range from a few bits in a capsule's header that select a protocol to scripts in a general purpose programming language that directly implement a protocol [8], [9]. Most of the current standard protocols transmit option bits that select between the functions that the protocol may perform. At the low end of the range, the implementation of capsules is almost identical to the implementation of the current standard protocol. Therefore, low end active networks can be implemented competitively.

The migration strategy between versions of active networks is similar to the strategy that is used to introduce new generations of the IP protocol in the current Internet [10, p. 347]. A new generation of IP is introduced on "islands" of network devices that understand and use the new protocol. Gateways at the boundaries of an island receive packets that use one of the protocols and places them inside packets that use the other protocol. The new packet is the "best" approximation of the protocol functions that can be performed in the region that the packet is about to enter.

A new generation of active networks, that uses a different programming language, is introduced on islands in the current network, in the same way as a new version of IP is introduced into the current Internet. In order for a capsule to cross a region of a network that uses a different language, the capsule is encapsulated in a new capsule that has instructions that can be interpreted in that region. An advantage of active routing over the current Internet procedures is that the encapsulation does not have to be the same for every capsule that crosses the border. The capsule can contain its own translation or the rules to perform the translation.

Carrying a few bits in the header of a capsule to select a protocol, rather than carrying the code, can be viewed as a short hand notation to specify a protocol. The bits select code which either resides at the processor, or can be obtained by the processor. An advantage of the shorthand notation is that the number of bits needed to transport the selection is less than the number of bits needed to transport the program. Therefore, the shorthand notation makes more efficient use of the network bandwidth. A disadvantage with the shorthand notation is that there is less freedom to select protocols, and changing or adding protocols can be as difficult as the current standards process.

There are several ways to make active routing practical sooner, in addition to the language evolution.

1) We can reduce the amount of processing per capsule by establishing a path for a sequence of capsules, similar to the call set-up procedures that are used in circuit switching. Circuit set-up is particularly appropriate for streaming, large data transfers on the WEB, and dedicated resources for a virtual private network.
2) Before intelligence is introduced at every router in a network, it can be introduced at the gateways between autonomous regions and at strategic routers within a region. Tunnels through the current network can connect the routers and gateways that execute active programs.
3) As intelligence becomes more economical but before routers can operate on every capsule, capsules can use techniques that tunnel ahead two or three routers at a time to reduce the average processing at each router. This is comparable to using digital or optical cross connects to reduce the processing in our current routers.

B. Network Sandboxes

In active networks, users introduce code into the network to acquire and use network resources. The code from one user may intentionally or accidentally have an adverse affect on the network or the other users. To prevent or limit the damage that a user may cause, the set of resources that the user can acquire is constrained to a "sandbox." The objective when defining the sandbox is to safely give the user access to as many resources as possible. Sandboxes are used in the JAVA programming environment [11] to allow foreign code to run on a user's computer. Network sandboxes are similar to programming sandboxes but have several new dimensions.

A JAVA sandbox limits the computer memory that a program can access and the processor cycles that the program can acquire. The purpose of the JAVA sandbox is to protect the computer from an untrusted program, while taking advantage of the service that the program provides. There is a balance between the services that the program can provide and the protection provided by the sandbox.

Protecting a network from the mobile code in a capsule, while allowing the code to perform useful routing functions, is more difficult than protecting a single computer from a JAVA script. To start with, the program in a capsule sequentially runs on a set of processors that are distributed around the network rather than in a single computer. In order to constrain the resources that a user acquires, the network must keep track of the distributed resources. Failure to track these resources can result in a user locking up resources far in excess of that allowed in his sandbox, and adversely affecting other users. For instance, if a user's routine produces a routing loop, the capsule may cycle indefinitely and use infinite bandwidth.

The network sandbox controls more types of resources than the programming sandbox [12]. The network sandbox must control the use of bandwidth and time in order to provide QoS guarantees. In addition to the processor memory that is used during the execution of a program, the sandbox must control fast memories in the routers and switches that operate at the network transmission rates and persistent memories that retain state information for sequences of capsules.

Active routing uses persistent memory to limit the average amount of processing per capsule and the overhead per capsule. In order to reduce the average processing time per capsule, the results from routing a capsule can be stored in persistent memories to route the next few capsules in a stream or an entire connection. In order to reduce the average overhead per capsule, the code can only be carried in the first capsule in a stream and stored in the persistent memory. Persistent network memory is also used to reserve resources, such as bandwidth or timeslots. The network must control the duration of the use of persistent memory or the resource may "leak" out of the network. In Section III-F, we describe techniques to control the extended network state, which includes persistent memory.

In the current programming sandboxes, resources are allocated then returned. The paradigm in networks is more complicated. The resources may also be reserved, scheduled for later use, or shared among a set of users. Users in the circuit switched voice network reserve bandwidth or schedule time slots on specific links in the network. Internet users share the bandwidth with other users. Combinations of shared and reserved resources are currently being proposed to provide QoS guarantees for an aggregated set of users [13]. As networks that use different technologies, such as circuit and packet switching, are interconnected, many more combinations of reservation and sharing mechanisms will be possible. These combinations will make new types of services available but complicate the management of an individual user's sandbox.

In addition to the stationary components of a user's sandbox, active routing can use a mobile sandbox to manage and pay for the distributed resources. A capsule is partitioned into areas that the user can modify and areas that are protected. The unprotected areas in the capsule are part of the user's sandbox. The protected areas can be used for accounting, such as counting the number of nodes that the capsule has traversed to end routing loops, or carrying digital cash to pay for resources, as described in Section III-C. As long as the capsule passes through trusted network components, the memory in the capsule can be protected by the same rules that protect the memory in a processor.

However, capsules may cross less trusted networks, and different techniques may be needed to protect the memory.

Finally, there are many types of participants in a network that may use active routing in the same capsule. The participants include individual users, ISPs that rent and resell facilities, competing network that are traversing another domain, and the network owner. The different participants require different types and sizes of sandboxes. For example, a user's sandbox must prevent that user from deallocating resources that have been dedicated to other users. Meanwhile, an ISP may sell interruptible and uninterruptible grades of service to different customers. In order to implement this type of service, the ISP's sandbox must have access to the resources that are allocated to all of its customers, even on links that are outside its own autonomous region. As a second example, the commands that are inserted in a capsule by the carrier that owns a network may be used to change the connectivity of its facilities switches—to change the topology of its network to match the load or to route around failures. However, an individual user or another network provider whose capsules are in transit should not be able to reconfigure the switches. In [12], there is a detailed discussion of the different size sandboxes that can be assigned to different participants in a network. Since there are several participants who may associate code with a capsule, and since the capabilities of the participants are different, the code areas of the capsule must be in different mobile sandboxes.

C. Pricing or Policing

At present, there is a trend to charge flat rates for Internet access, regardless of use. However, with active routing, an end user may improve his own performance by using network resources wastefully. For example, if the routing objective is to pick a path with the smallest delay, a wasteful user may duplicate the capsule, send a copy over each of the paths and discard all but the first capsule that arrives at the receiver. The result of this procedure is that wasteful users receive better performance, while users who are not wasteful receive poorer performance. This encourages all of the users to be wasteful. However, if all of the users are wasteful, their performance is worse than if no one was wasteful. In order for the network to operate at peak performance, users must be discouraged from wasting resources.

The transmission control protocol (TCP) protocol [14] is a well known example of a user controlled protocol that can encourage waste and is potentially harmful to the network. TCP users are responsible for controlling their own data rate. When the network becomes congested and packets are lost, the users who lose packets send less often. When all of the users follow the same rules for reducing and increasing the number of packets they transmit, everyone gets the same share of the network bandwidth. However, if some users send more aggressively, they receive a larger share of the bandwidth [15] while the other users get less. As a result, if some users operate more aggressively, the others are forced to follow suit or lose their share of the bandwidth. If the network is congested and no one backs off, fewer messages get through and everyone is worse off.

There are two techniques that we can use to encourage users not to waste network resources, pricing [16], [17] and policing [18]. With pricing, users are charged for the resources they use or the resources that they reserve as part of their sandbox. With policing, users are observed and if their network use is outside of acceptable bounds, they are punished.

Policing imbeds standards, in the form of "acceptable use." Acceptable use may prevent new routing mechanisms from being implemented until a new policy is agreed upon and new policing mechanisms are put in place. Policing can slow changes in routing as much as other standards.

Per use pricing is contrary to what Internet customers expect and maintaining detailed bills is currently one of the most expensive components of telephone service. However, pricing does not necessarily imply that the network maintain billing information or institute per unit charges. For instance, as an alternative to billing for resources, capsules can carry electronic cash [19]–[25] and pay for the resources as they are used. Electronic cash may be placed in a capsule by the user, or by the user's ISP. The cash that is placed in the capsule by the ISP may reflect the grade of service that the user has contracted, and not a per use charge.

The original electronic cash mechanisms require too much processing and communications to be used to pay for each resource used by each capsule. These mechanisms may be appropriate to pay for resources on a connection basis, like paying for a three-minute packet voice connection with quality guarantees. There are electronic cash mechanisms that are designed for micropayments on the Internet [26]–[28]. These mechanisms come closer to our objective of paying by the capsule but may still be too complex. It is worth considering other mechanisms that take advantage of the unique characteristics of our problem.

One difference between network payments and general digital cash mechanisms is that we may be able to stop the user from handling the cash by keeping it out of his sandbox. When digital cash is handled by a user, we must make certain that the user is not duplicating cash or printing his own currency. Counterfeit prevention causes most of the processing and communications in digital cash. On the network, we can require a trusted ISP to place the cash in the capsule. As long as capsule is only processed by trusted network components, we can guarantee that the user's script will not modify the cash.

There is still the problem described in the Section III-B, when the capsule traverses a different network. The network may not be trusted to restrict access to the sensitive parts of packet. The conventional cryptographic approach to this problem is to place a sequence number in the data to detect duplication and to use a digital signature to detect tampering. Once again, the processing and overhead bits associated with digital signatures may not be justified and we should look at the unique characteristics of our problem.

The value of payments in the capsule is small, the lifetime of the cash is short, and we can require the networks to "cash in" the electronic cash at the bank that gave it to the ISP. It may be sufficient for the bank to attach a short random sequence, one-time pad, to each payment unit that it issues and only "cash it in" if it has the correct sequence number. For instance, the nth unit of cash is issued in time interval T_i and has the attached random number X. It can only be recovered once during that time interval, if it has the correct random number. The probability that a unit of cash can be duplicated decreases as the length of the random number increases. As long as the value of the unit of cash is small, we can use just a few bits and stop cashing in units from networks that have frequent occurrences of invalid units.

Because the payment values are small, we can consider more radical alternatives. For instance, many European public transportation systems trust the users to pay for the service. Occasionally, the system checks the riders and fines those who have not paid. One network may trust another network to protect capsules and to handle the cash properly but may occasionally send a test capsule through to "check up on" the network.

While we have given several examples of how electronic cash may be implemented, we should emphasize that in an active network there should not be a standard. Any mechanism that is agreed upon by the user and facilities provider should be allowed, and any mixture of mechanisms should be permitted. If the two parties are willing to pay for the processing and communications associated with conventional digital cash, and are willing to be limited by the constraints that this places on their throughput, then they should be allowed to do this. If an entrepreneur buys a guaranteed bandwidth tunnel, and figures out how to guarantee quality of service through the tunnel by issuing cash (tokens) with a short life, the network should not prevent it.

In addition to paying for facilities, electronic cash can be used to control a distributed sandbox. Electronic cash can prevent a user from intentionally or inadvertently exceeding any limit that is placed on the allowed resources. For instance, consider a routing mechanism that duplicates capsules for multicast [29] or dispersity routing [30]. If there is a flaw in the protocol, the capsule may be duplicated at every node that receives a copy and flood the network. However, if each router removes some cash and divides the remaining cash among the copies of the capsule, the total number of duplicates that can be generated is limited by the amount of cash in the original capsule. As a second example, consider a routing mechanism with a loop, the capsule eventually runs out of cash and is discarded. In both of these examples a distributed sandbox is implemented by carrying state in the capsule that summarizes the resources that the capsule has used. We have discussed limiting the total resources that a capsule can use but more precise controls can be placed on a sandbox by using different types of cash to pay for different facilities.

An advantage of electronic cash is that the network does not dictate which resources a capsule may access. A user who has a real need to flood the network, and is willing to pay for it, can. The amount of cash in a capsule determines the amount of network resources that a user can access and provides different grades of service without networks standardizing a set of grades of service. It is the responsibility of the facilities providers to limit the amount of cash that they print, so that they do not sell more facilities than they own.

D. Metarouters

There are networks that use similar or vastly different technologies that are competing for users in the core of the network. There are telephone networks using circuit switched technologies and Internet providers using packet-switched technologies. Active routing can be used to select the most advantageous combination of networks with different technologies, or the best among a set of similar network, by installing metarouters at their junctions. For instance, active routing may select a best-effort network when measurements indicate that the network can provide an acceptable QoS, or it may bypass a congested portion of a best-effort network with a dedicated circuit. If two telephone networks are available for a segment of a path, the router may select the network with the lowest tariff. In a metarouter, different capsules can select the different networks based upon network measurements, price, prearranged contracts, or any other criterion that is important to a particular user.

There are also networks competing for local access to the core network, conventional twisted pair, digital subscriber loops, community antenna television system (CATV), fixed and mobile wireless, etc. Metarouters at the edge of a network can select the appropriate path depending upon bandwidth, cost or availability. The metarouter may also convert the data to fit the constraints of the particular technology. For instance, segments of the data in a high rate multicast may be discarded if the access link cannot support the rate.

The metarouters construct a super-network that is composed of all of the existing networks. They convert capsules to the transmission format of the selected network so that the underlying network does not have to use or even be aware of active routing. Customers may elect to be connected to a super-network, rather than one of the underlying, single technology networks. Several super-networks may compete based upon the location of their metarouters and their connectivity to and agreements with the underlying networks.

In the example in Section VI, metarouters are used to establish a voice connection across a combination of circuit and packet-switched networks. In this example, the different users have different cost/quality tradeoffs that can result in different combinations of circuit and packet connections between the same two points. The different users may also establish the paths for different durations: some of the users may set up a path for an entire connection, while others monitor the performance and modify the path as the characteristics of the connection changes. In addition to selecting the path, the metarouters perform the translation functions of "hop on/hop off" nodes in Internet protocol (IP) telephony networks. The metarouter performs the translation so that neither the voice nor packet network need to be changed.

E. Information Servers

Routing implies knowledge of the network. At the very least, a router must know or be able to discover the connectivity of the network in order to select a path. Selecting a route that provides QoS requires additional information. In a circuit-switched network, establishing a route that has a specified quality includes determining if there is spare bandwidth on a path. In a packet-switched network, we must determine the delay bounds on the path, which requires information about the utilization of the links.

In active routing we expect a user to obtain and pay for information about the network in the same way that he obtains and pays for any of the other resources that he needs. There are three sources of information:

1) network owners;
2) third party information centers; or
3) direct measurements.

The network owners are in the best position to collect information on their own resources. However, traffic from a source to a destination is likely to cross several autonomous domains that use BGP to exchange information. Collecting and passing information between networks requires bilateral agreements and a new standard protocol. It is difficult for the information exchanged by a standard protocol to keep up with the information that is needed in an environment where the services change rapidly.

Third party information centers collect or buy information from network owners, make their own measurements, and process the information. Tools have been deployed by ISPs that make these measurements [31]. The information centers can sell information to the network owners so that they can provide end-to-end service guarantees for their users or sell the information directly to the users. In a competitive environment with several information centers, any information that can be sold at a profit will quickly become available.

Information centers that collect the same information for a group of users should operate more economically than individuals who collect their own information. However, in the infancy of a new service or for users with unique requirements, a user has the option to probe the network and make his own measurements. A network probe may simply be occasional packets that take alternate paths to the destination and report to the source which performs better. Alternatively, a capsule may spawn a "ping" from the intermediate node along the alternate path to the destination, and set a state variable at the node if future packets should follow the alternate path.

F. Network State

The telephone network is considered a "stateful" network because the switches assign resources to each user with a connection through the switch. The Internet was considered a "stateless" network because the routers avoided assigning resources to specific users. Per user state information may be added to the Internet in order to provide QoS guarantees.

In order to deal with some of the problems associated with stateful networks, the Internet introduced the concept of hard and soft states [32]. Hard states are set up by a protocol and remain in existence until the protocol changes them. The disadvantage with hard states is that a protocol error or a failure may lock up resources. Soft states are also set by a protocol but are automatically removed after a specified amount of time. The disadvantage with soft states is that a protocol must periodically reassert its need for the resource.

Returning resources is a more important problem in active networks than in the current networks. To start, individual users in an active network can acquire more resources of more types than in conventional networks. More importantly, the user supplied protocols in an active network are more likely to have errors that fail to deallocate resources than standard protocols.

In the spirit of active routing, the network should not dictate that an individual user use hard or soft states. An alternative is to allow a user to rent resources for a specified time, as part of his sandbox. The amount of time is determined by the user and, in effect, becomes the soft state refresh interval for that user. The interval may be very different for different users, and may approach hard states for some users. For instance, a corporation's private network may rent bandwidth by the month and remain in existence for years, while an individual voice call may rent bandwidth by the second and only remain in existence for a few minutes.

IV. NETWORK PRICING MODEL

In order to determine the effect of active routing on pricing, we model a system in which networks compete for individual customers or classes of customers. We first describe a customer model, then a network model, and finally their interaction. The interaction resembles a fair auction. When a customer must buy a single set of resources, he buys the resources from the network that is willing to sell them at the lowest price but pays the price at which the second lowest priced network would have sold the resources.

With active routing, a customer can substitute equivalent resources by selecting different networks. Our objective is to demonstrate that customers can strike a better deal when they are willing to use different resources from different types of networks. The efficient use of resources has been extensively studied while the substitution of resources has been barely considered. In the future, the gains that customers obtain by choosing between competing resources may be greater than the incremental gains that they can obtain by further improvements in how efficiently they use a particular resource.

In this section, we do not discuss the implementation details of the auction. Different active routing programs may use very different strategies that either implement or approximate a fair auction.

- A program that is used to route the links in a large corporate virtual private network (VPN) that will be operational for a long time may traverse networks several times with offers from competing networks to directly implement an auction.
- A program that routes a voice connection that will only last a few minutes, may access a tariff database before selecting a network. The customer depends on competition in the market place to approximate an auction for frequently purchased resources.
- A program that moves large amounts of data between proxy servers may visit sites where network providers report discounts, due to a short term over supply, and buy bits on the spot market. The discounts offered by competing networks that can see the discounts offered by the other networks, also approximates an auction.

In our model:

$\vec{r} = (r_1, r_2, \ldots)$ is an n-tuple of resources. The resources include:

- the memory and processing cycles that are needed to run the program in the intermediate switches and routers;

- persistent memory to store state or programs along a path between a source and destination;
- information about the state of the network; and
- transmission resources, such as total bits, guaranteed bandwidth, guaranteed delay, and queuing priority.

p is a payment that the customer will make or that the network provider will receive.

For a specific customer, c, $U_c(\vec{r}, p)$ is the utility of purchasing $\vec{r}$ at p.

If $U_c(\vec{r}_1, p_1) > U_c(\vec{r}_2, p_2)$, then c prefers purchasing $\vec{r}_1$ at p_1 to purchasing $\vec{r}_2$ at p_2.

If $p_a < p_b$, then $U_c(\vec{r}, p_a) > U_c(\vec{r}, p_b)$. ($U$ is a monotonically decreasing function of p.)

$p_{c,\max}(\vec{r})$ such that $U_c(\vec{r}, p_{c,\max}(\vec{r})) = 0$ is the maximum price that c will pay for $\vec{r}$.

$\mathbf{R}_c = \{\vec{r}: U_c(\vec{r}, 0) \geq 0\}$ is the set of resources that c will consider buying.

For a specific network n, $V_n(\vec{r}, p)$ is the value of selling $\vec{r}$ at p.

If $V_n(\vec{r}_1, p_1) > V_n(\vec{r}_2, p_2)$, then n prefers selling $\vec{r}_1$ at p_1 to selling $\vec{r}_2$ at p_2.

If $p_a < p_b$, then $V_n(\vec{r}, p_a) < V_n(\vec{r}, p_b)$. ($V$ is a monotonically increasing function of p.)

$p_{n,\min}(\vec{r})$ such that $V_n(\vec{r}, p_{n,\min}(\vec{r})) = 0$ is the minimum price that n will sell $\vec{r}$.

The group of networks that can sell $\vec{r}$ to c is $N(\vec{r}, c) = \{n: p_{n,\min}(\vec{r}) \leq p_{c,\max}(\vec{r})\}$. If $N(\vec{r}, c)$ is empty, then there are no networks that can provide the resource at a price that the customer will pay.

If there is only one member in $N(\vec{r}, c)$, n and c will negotiate p such that $p_{n,\min}(\vec{r}) \leq p \leq p_{c,\max}(\vec{r})$. The network will try to keep the price as high as possible and the customer will try to keep the price as low as possible. In most instances, the network is in a stronger bargaining position than the customer because a customer suffers more hardship if he cannot communicate than a network suffers if it fails to sell its resource to one of many customers. Of course, there are exceptions. A large individual customer may have the position of strength. A large number of small customers may negotiate through a consumer's union. Or, a regulatory agency may force a network to publish a single tariff for all customers, in which case the network may maximize its value over all customers by reducing p below the $p_{c,\max}(\vec{r})$ that the highest paying customers are willing to pay.

In the remainder of this work, we will assume that the network bargains with individual customers from strength. When *one* network bargains with a customer who must buy *one* resource, signified by the superscript (1, 1), the price of the resource is

$$p^{(1,1)} = p_{c,\max}(\vec{r})$$

c receives utility

$$U^{(1,1)} = U_c(\vec{r}, p_{c,\max}(\vec{r})) = 0$$

and n receives value

$$V^{(1,1)} = V_n(\vec{r}, p_{c,\max}(\vec{r})) \geq 0.$$

In this instance, there is no competition between networks. The excess profit that n makes, above the profit that it was willing to accept, is

$$P_n^{(1,1)} = p^{(1,1)} - p_{n,\min}(\vec{r}) = p_{c,\max}(\vec{r}) - p_{n,\min}(\vec{r})$$

and, the excess savings that c obtains, below the price that he was willing to pay, is

$$P_c^{(1,1)} = p_{c,\max} - p^{(1,1)} = 0.$$

When there are two or more networks in $N(\vec{r}, c)$, the networks compete with one another for the customer. We can order the networks as:

$$n_1, n_2, n_3, \ldots$$

so that

$$p_{n_1,\min}(\vec{r}) \leq p_{n_2,\min}(\vec{r}) \leq p_{n_3,\min}(\vec{r}) \leq \cdots.$$

After a fair (second price [33]) auction, between *multiple* networks that bid to sell *one* resource, signified by the superscript $(m, 1)$, n_1 offers the resource at

$$p^{(m,1)} = p_{n_2,\min}(\vec{r}) - \varepsilon$$

which we will consider to be $p_{n_2,\min}(\vec{r})$. c receives utility

$$U^{(m,1)} = U_c(\vec{r}, p_{n_2,\min}(\vec{r}))$$

and, n_1 receives value

$$V^{(m,1)} = V_{n_1}(\vec{r}, p_{n_2,\min}(\vec{r})).$$

In effect, the most competitive network is forced to offer the resource at the price that the next most competitive network is willing to accept, rather than the maximum price that the customer will pay.

Competition decreases n_1's excess profit to

$$\begin{aligned} P_n^{(m,1)} &= p_{n_2,\min}(\vec{r}) - p_{n_1,\min}(\vec{r}) \\ &= P_n^{(1,1)} - [p_{c,\max}(\vec{r}) - p_{n_2,\min}(\vec{r})]) \end{aligned}$$

and increases c's excess savings to

$$\begin{aligned} P_c^{(m,1)} &= p_{c,\max}(\vec{r}) - p_{n_2,\min}(\vec{r}) \\ &= P_c^{(1,1)} + [p_{c,\min}(\vec{r}) - p_{n_2,\min}(\vec{r})]. \end{aligned}$$

Let us now assume that c can make use of any one of a set of resources $\mathbf{R}_c$. The customer selects the network that offers the highest utility. However, the network would like to win the auction with as low a utility as possible. The lower utility increases the price that the network can charge for any resource, and therefore, increases the network's value.

The highest utility that n can provide to c, $U_n = \max_{\vec{r} \in \mathbf{R}_c} U_c(\vec{r}, p_{n,\min}(\vec{r}))$.

As before, order the networks as: $n_1, n_2, n_3, \ldots$, so that $U_{n_1} \geq U_{n_2} \geq U_{n_3} \geq \cdots$.

Network n_1 wins the auction and must provide utility $U_{n_2} + \varepsilon$ (which we consider to be U_{n_2}).

For each $\vec{r} \in \mathbf{R}_c$, n_1 can charge up to $p_{n_1,\max}(\vec{r})$ such that $U_c(\vec{r}, p_{n_1,\max}(\vec{r})) = U_{n_2}$, and obtain value $V_{n_1}(\vec{r}, p_{n_1,\max}(\vec{r}))$.

Network n_1 selects r_1 such that $V_{n_1}(\vec{r}_1, p_{n_1,\max}(\vec{r}_1)) \geq V_{n_1}(\vec{r}_2, p_{n_1,\max}(\vec{r}_2))$ for all $\vec{r}_2 \in \mathbf{R}_c$.

The result of the *multiple* network auction for *multiple* resources, signified by the superscript (m, m), is that c receives utility

$$U^{(m,m)} = U_{n_2}, \text{ using resource } \vec{r}_1 \text{ at price } p_{n_1,\max}(\vec{r}_1)$$
$$\text{and } n_1 \text{ receives value}$$
$$V^{(m,m)} = V_{n_1}(\vec{r}_1, p_{n_1,\max}(\vec{r})).$$

Network n_1's excess process is

$$\begin{aligned} P_n^{(m,m)} &= p_{n_1,\max}(\vec{r}_1) - p_{n_1,\min}(\vec{r}_1) \\ &= P_n^{(m,1)} - \left[p_{n_2,\min}^{(m,1)}(\vec{r}_1) - p_{n_1,\max}(\vec{r}_1)\right]\Big) \end{aligned}$$

where $p_{n_2,\min}^{(m,1)}(\vec{r}_1)$ is the price from an $(m, 1)$ auction for $\vec{r}_1$, and c's excess savings is

$$\begin{aligned} P_c^{(m,m)} &= p_{c,\max}(\vec{r}) - p_{n_1,\max}(\vec{r}_1) \\ &= P_c^{(m,1)} + \left[p_{n_2,\min}^{(m,1)}(\vec{r}_1) - p_{n_1,\max}(\vec{r}_1)\right]\Big). \end{aligned}$$

$P_n^{(m,1)}$ is the profit from a single service auction for $\vec{r}_1$. If there is only one network that can provide resource r_1, then $P_n^{(m,1)} = P_n^{(1,1)}$ and $p_{n_2,\min}^{(m,1)}(\vec{r}_1) = p_{c,\max}(\vec{r}_1)$.

Whether or not the customer is better off in the (m, m) than in $(m, 1)$ auction depends on the resource that n_2 must use to provide $U_{n_2}^{(m,m)}$. If n_2 can use $\vec{r}_1$ to provide $U_{n_2}^{(m,m)}$, then

$$U_{n_2}^{(m,m)} = U_{n_2}^{(m,1)}$$

and the customer is no better off for considering multiple resources. If n_2 must use $\vec{r}_2 \neq \vec{r}_1$, then

$$U_{n_2}^{(m,m)} > U_{n_2}^{(m,1)}$$

which infers that

$$p_{n_1,\max}(\vec{r}_1) < p_{n_2,\min}(\vec{r}_1)$$

and therefore,

$$P_c^{(m,m)} > P_c^{(m,1)}.$$

We can conclude that a customer is always better off when he adds a new resource that wins the auction by providing a higher utilization. If the new resource cannot win the auction but can become the second best choice, then it can improve the customer's bargaining position so that the customer receives a higher utility with the old resource. If a resource does not fill one of these two slots, it does not improve the customers position.

In Section VI, we apply the auction to routing voice over a combination of circuit and packet networks. We show examples of how the $(1, 1)$, $(m, 1)$, and (m, m) auctions result in different utilities for the customers and different values for the networks.

V. Applications

In this section, we describe several applications of active routing. The first two applications, extended label switching and policy-based routing, are traditionally intended for network operators. However, with active routing, an individual user may use his own restoration techniques if the reliability of the network is not sufficient or may create his own policies to protect sensitive information. The first application is an extension of the multiple protocol label switching (MPLS) techniques that are currently being introduced into the Internet, and shows that active routing is a logical extension of the current routing trends.

The last three applications, QoS routing, mobility, and receiver-controlled routing, are primarily intended for individuals. However, we do note that QoS routing may also be useful to virtual network providers or to source networks with traffic crossing different autonomous regions. We describe new uses of the last two applications, including network support for proxy servers and controlling the final distribution of multimedia information.

A. Extended Label Switching

MPLS uses label switched routers (LSRs). In an LSR, select IP addresses are recognized and these packets are encapsulated in packets that have much shorter addresses called labels. The labels can be much shorter than IP addresses because they only specify a connection in a single router, rather than a unique network destination. When a labeled packet is switched, the label is either exchanged for a label that has meaning in the next router, or removed to expose the original IP packet. The exposed IP packet then uses conventional IP routing. The label switched path (LSP), from the router that encapsulates the IP packet to the router that removes the encapsulation, is a virtual circuit. On the path, the operation of the LSR on labels is identical to the operation of an asynchronous transfer mode (ATM) switch [34] on the addresses in an ATM cell.

Active routing can extend label switching to:

1) provide more flexible access to LSPs;
2) perform a wider range of operations on stacks of labels;
3) implement new rules for constructing stacks of labels;
4) enable a user to establish his own label-switched paths;
5) use more flexible restoration techniques;
6) remove the need to increase the address space in IP v4.

At present, the IP addresses of the packets that are encapsulated by a label are loaded into the edge router when a virtual circuit is set up. With active routing, we can allow capsules to gain access to the path by other means. For instance, a network may allow access to a high priority LSP to any capsule that pays a toll. The active capsule can use one of the digital payment options described earlier.

More flexible access options may also be useful in VPNs. LSPs provide security in VPNs by limiting access based on IP addresses. However, it may be useful to allow employees who are traveling or working from home to access a corporate VPN. An active capsule can implement a protocol to identify itself as having permission to enter. Different companies can use different degrees of security, depending upon how concerned they are with illegitimate access and how much the network charges to implement the protocol.

Labeled packets can encapsulate other labeled packets, as well as encapsulating the original IP packet. Stacks of labels are being considered to route around a failed link, or to merge or split the virtual circuits in a VPN, in order to get better control over the topology. for restoration, traffic engineering, etc. At present, MPLS only allows simple push and pop operations on label stacks that implement a single path, and the rules for forming stacks of labels are programmed into the router. It is difficult, for independent applications providers to create new uses for labels.

With active routing, capsules can perform more complex operations on the stacks of labels and carry their own rules for creating new stacks. A useful extension would be conditional paths. Instead of having a single stack of labels, a capsule may carry a stack that splits at a midpoint to provide two paths to the final destination. At the split, the capsule may take one path or the other dependent upon which links have failed, the expected delay on the paths, the delay that the capsule has incurred to this point, a sign that was left by the owner of the VPN, or many other conditions. Path selection can be used in a packet voice network where delay is a concern. At the split, the capsule may select a low or high priority tunnel, dependent upon the delay that it has incurred. Path selection can also be used for restoration, traffic engineering, etc. by large corporate VPNs or by new carriers that are being created as VPNs on existing carrier networks.

At present, label paths are established by the owner of the network using one of the label distribution protocols [5]. In an active network, a capsule can be used to set up a new label switched path. A corporation may use this ability to add new paths to its VPN for short periods of heavy load or when a network link has failed. The path can be set up as a label distribution capsule moves through the network, using conditions that are important to the user.

Bypassing failed links is an important application of label switching [35]. Each node has a label-switched tunnel that does not use the direct link to each neighboring node. The tunnel is used to route around failed links. With active routing, each capsule may respond differently at the failure. Some capsules may take the detour recommended by the network, while others may spawn a warning message to the source or pick their own alternate paths.

The analogy between MPLS and ATM can be used to reduce the address space in IP packets, which is a prime motivator to move from IPv4 to IPv6. ATM uses unique identifiers to set up a circuit but does not carry the identifier in an ATM cell. IP switching carries the unique identifier in each packet but can operates without setting up circuits. With MPLS, we can use shorter IP addresses without setting up per user virtual circuits.

Label switching operates on virtual circuits, the same as ATM. Permanent MPLS tunnels can be established between subnetworks, such as Australia, the ATT network in the USA, .edu, the GE corporate VPN, etc. Instead of using IP addresses to uniquely identifying all users, IP addresses can be used to identify users in a subnetwork, so that users in different subgroups can have the same IP address. An active capsule can be used to select the MPLS tunnels to the destination subnetwork, then conventional IP routing can be used in the subnetwork. This approach sets up a relatively small number of permanent virtual circuits between subnetworks, rather than setting up a large number of virtual circuits between each source and destination.

If we were certain that we knew the best way to construct and interconnect subnetworks, we could create standards and use conventional tunneling and encapsulation techniques to reduce the IP address space. With active routing, we can try different constructs and correct them as we learn.

B. Policy Routing

Policy-based [36] routing was initially proposed to constrain the paths taken by different users when the civilian and military ARPAnets were merged. At present, policy routing is used in firewalls [37] to control the packets that can pass between private and public networks.

VPNs [38], [39] that use shared resources on the public Internet require new types of firewalls that are more flexible and distributed. Access to the MPLS or IP tunnels in a VPN is constrained. As previously noted, the advantage of active routing is that the VPN's owner can specify his own criterion to access the network.

An interesting twist on virtual *private* networks that use active routing is virtual *public* networks that use active routing. The public network may set aside part of its resources for users with trusted credentials. The rest of the resources being made available to everyone. The objective is to create a VPN that has better characteristics than the general public network. For instance, the VPN may protect known, good customers from the general population. If an untrusted user mounts a denial of service attack, the network firewall stops him from acquiring the resources that are set aside for trusted customers. (Establishing trust can include proving your identity so that known customers can be traced if they mount a denial of service attack.) The type of credentials that a network accepts determines how much processing is required and how well the reserved resources are protected. In a competitive environment, different networks can use criterion that attract different customers.

Returning to the original Milnet application, policy routing can be used to make e-commerce and electronic stock trading more secure. In active routing, the capsule controls the parts of the network that it is willing to traverse. The policy can make certain that credit card numbers and buy or sell orders remain in parts of the network that are trusted.

C. QoS Routing [40], [41]

As the circuit and packet-switched networks converge and compete, QoS has become one of the most important topics in networking. We would like to put services like telephony and video broadcasts, that expect low delay, low loss, and guaranteed bandwidth, on networks that, up to now, have provided best effort service.

QoS is an excellent candidate for active routing because of the following.

1) QoS can be supplied by several different participants in a network, the network owner, a source network that is traversing a different autonomous region, a virtual private network, or the source — the actions taken by any participant depend upon the actions that are taken by the others

and may change as a capsule progresses through interconnected networks.

2) QoS is based upon individual perception — each user may have different tradeoffs between quality and cost and a small number of classifications may not be sufficient to satisfy all of the users.
3) A large number of alternative strategies for QoS are developing and rapidly changing — on the Internet, protocols such as RSVP [42] and Diffserv [43] are competing with one another as well as with network-based techniques, such as forwarding equivalence classes [44].
4) Active routing provides a means for protocols and networks to compete, rather than a winner being dictated by standards.

QoS routing provides a good example of how active routing can misuse network resources, and also why pricing is preferable to policing. If the objective is to pick a route with the smallest delay and there are no charges associated with transmission, the best solution, from an individual's point of view, is to transmit copies on all of the routes and accept the first copy that arrives. If a significant fraction of the users follow this strategy, the network becomes congested and everyone's performance suffers [30]. However, there may be critical applications, like fire alarms or stock market orders, that should replicate the data. By using pricing, instead of policing, these applications can purchase the additional resources.

D. Mobility [45], [46]

There are a number of ways to send data to a receiver that moves around the network, either on a short term basis, such as traversing cells in a mobile network, or on a longer time scale, such as an individual from the east coast who is temporarily working in an office on the west coast. The mobility techniques include:

1) sending data to an intermediary that the receiver keeps apprised of its location, and that forward the data;
2) trying several locations where you think that the receiver may be; or
3) sending a broadcast message over part or all of the network.

With active routing, customers can implement different mechanisms, and the best will prevail.

The newest example of mobility in the Internet is proxy servers. Copies of data that is expected to be requested, or that has recently been requested, is cached at proxy servers nearer to where it is likely to be needed. When the data is requested, it is transmitted from a nearby cache. In effect, the data is mobile. At present, proxy servers are implemented at the applications layer. In a network with active routing, proxy servers can also be supported at the network level.

E. Receiver-Controlled Routing

So far, the source in our the applications has controlled routing by placing a program in a capsule. We can also allow a receiver to download programs into the network to control how data is routed to it.

One application of receiver control is the type of mobile network that we have just described. The receiver can store state in the switches at various points in the network. The state can be interpreted by capsules to decide on the best path to a mobile receiver, or the most convenient location of a proxy server.

The best known example of receiver control is the multicast backbone (Mbone [29]) on the Internet. In this application, receivers communicate with the network devices that replicate the signal, so that the signal is not transmitted on parts of the network where it is not needed. One of the problems with multicast is that the source transmits at a rate that can be received by the least capable receiver. An early application of active networks has been to use processing in the network to filter the signal, so that it can be received by receivers with different capabilities [47].

As the Internet, telephone network, mobile radio network, and cable TV network converge, each user is likely to have several network connections. The connections will operate at different bit rates and support devices that can receive different types of signals. The user must be able to route signals to the appropriate device or convert between formats. At present, on wireless Internet connections, images are removed from WEB pages in order to get the information in a reasonable time. A second example where the network must tailor the data to the user interface, is universal messaging [48]. If an email message is being forwarded to a standard telephone, it is routed through a processor that does text to speech conversion. In the present systems, conversion is supported by the network and is limited by the network operators. Active routing allows the receiver to route the data through any intermediate processors and can open format conversion to competition.

VI. Example—IP Telephony

In this section, we consider routing a voice connection between two users over a combination of circuit and packet switched networks. The networks are connected by metarouters that can move the connection from one network to the other at several points along the path. In the current IP telephony networks, these meta-routers are called HoHo's, for hop-on, hop-off nodes. In this application, there is a tradeoff between price and quality and different customers prefer different solutions.

A. Service Model

We do not expect customers in this type of an application to write the code that routes the data, know which network resources are used to establish their connection, or specify their utility function. The customers buy or lease a program. The program assumes a utility function, uses the utility function to negotiate with the networks for resources, sets up the connection, and the bills the customer.

Each software vendor may use a different utility function. The utility function may be different for different classes of customers and a vendor may or may not adapt the utility function to an individual customer. For instance, one vendor may assume that there are just two classes of customers. In one class, there are customers that prefer a low delay, low loss connection and

TABLE I
LINE COSTS FOR CIRCUIT AND PACKET NETWORKS ON TWO PATHS FROM LA TO LONDON

Path	Link	$p_{n,\min}$ (¢/min)	
		Cir	Pkt
1	LA-Chi	5	2
	Chi-NY	5	2
	NY-Lon	10	2
2	LA-Dal	6	3
	Dal-Atl	6	3
	Atl-Lon	12	3

will not consider degraded service. In the second class, there are customers that prefer the lowest cost connection that can exceeds a minimum quality. The vendor will design utility functions that reflect the difference between the two groups and let the customer select his initial group. The vendor may then adjust the customer's utility function as he uses the service. For example, if the customer indicates that he would like to consider higher quality, more costly connections, the vender may adjust the utility function, show the customer the price per minute that he is paying and let the customer decide if he prefers this service to the service that he has been receiving.

The software vendors may also use different sets of resources to implement the application. In a circuit/packet voice network, there are a range of implementations. Some of the implementations are possible immediately, while others require additional infrastructure or technical advances.

In order to demonstrate the range of solutions, consider a specific example where the service programs establish a voice connection between Los Angeles and London. Assume that:

- there are two possible paths, one path from L.A. to Chicago to N.Y. to London, and the other from L.A. to Dallas to Washington to London;
- there are metarouters that can switch traffic between the circuit and packet networks in each of the cities;
- the connections start and end in the circuit domain, and the metarouters charge 2¢/min. to convert between packets and circuits; and
- the minimum prices that the networks charge for each of the links are as shown in Table I.

Different programs can:

1) consider single or multiple paths based on the long term data that has been gathered on the paths;
2) send out scout packets during the circuit set-up procedure and select the best path at the time of the connection;
3) make measurements during the silent intervals in voice so that the best connection can be selected for each active interval; or
4) make decisions on a packet basis.

In order to implement any of these services, the software vendor must decide upon an acceptable grade of service, and make measurements or buy information about the links. The cost of obtaining information must be added to the price that he charges the customer.

As an example, the vendor may only consider links if the variable component of the delay on the path is less than 100 m/s at least 99% of the time. If the vendor builds out the delay by 100 m/s before converting from packets back to a circuit, less than 1% of the packets are lost because of late arrivals. A vendor that selects a QoS that is too high will lose the ability to lower a customer's cost by considering alternative networks, and a vendor that selects too low a criterion will not provide circuits that satisfy the customer.

Assume that long term measurements on utilization indicate that any one but not more than one of the three links on the second path can provide adequate quality on almost all connections. A vendor in the first class, who only considers the Chicago route, decides that the packet alternative will not meet the customers demand, and can only select the circuit network. When this vendor negotiates with the network, he will pay $p_{c,\max}$, the maximum amount that the customer will tolerate. A vendor that considers both paths can select the circuit network from LA–Dal–Atl, and the packet network from Atl–Lon. The minimum price, including two conversions between circuits and packets, is $p_{n,\min} = 19$¢/min. on the second path. The customer's satisfaction with the packet connection may be lower than the circuit connection but this vendor should be able to negotiate a higher utility from the networks than the first vendor. In Section VI-B, we present an example of the negotiation.

A vendor in the second class sends scout packets along both links prior to setting up the connection. Let's assume that half of the time he finds the utilization on the packet network on the LA–Chi–NY link to be low enough that it will provide adequate QoS for the next few minutes, then, $p_{n,\min} = 18$ on the first path, and the vendor can use that link to improve his bargaining position. Half of the time, this vendor makes the same choice as the first vendor but half of the time, he can increase the utility for the customer. The added expense is the cost of scouting the network before setting up a connection.

Finally, a vendor in the third class monitors the packet links during each silent interval in speech. If the utilization is low enough to provide adequate QoS for the next few seconds, the vendor uses this link to improve his bargaining position. The vendor may use a different utility function for a circuit that changes paths for each active interval, since the customer's perceived quality may be lower. Let us assume that the vendor has a packet connection from LA–Chi–NY, as a result of the second vendor's negotiation, and finds that for half of the active intervals he can send the packets all of the way to London. The cost of the first path is reduced to $p_{n,\min} = 10$, which improves the bargaining position relative to the second vendor. The added cost is the measurements that the vendor makes during silent intervals.

In addition to selecting between networks, in the near future there would be priority service on the packet network. We assume that the premium service would cost more, otherwise everyone will select the service, and there will be no high priority. The same three techniques that were used to select the circuit or packet network can be used to select the high priority or standard packet network.

As user accessible processing in the nodes increases, active routing may also be used to make decisions on a capsule basis. For instance, if a capsule is charged for each router that it traverses and extra is charged for priority service, the capsule may change its service depending upon the time that has elapsed since it was transmitted. If the delay exceeds a specified amount, at a given point on the path, the capsule priority can be raised

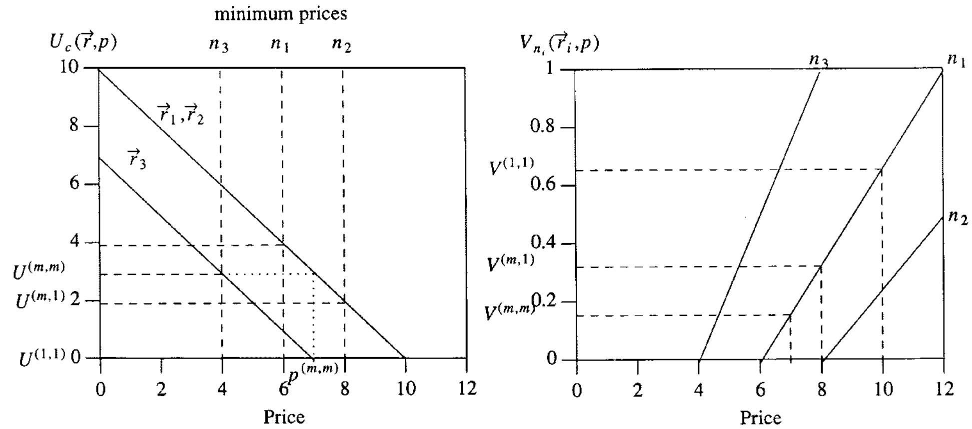

Fig. 1. Example: Two circuit networks and a packet network compete to provide IP telephony. The utility as a function of price for the customer and the value as a function of price for the networks.

for the remainder of the trip to the destination. However, if the delay exceeds the reconstruction delay at the receiver, the capsule can be discarded since it will not arrive in time to be used.

B. Network Competition

In the IP telephony example, a software vendor decreases the price that is paid to the network by increasing the number of networks that compete and the sets of network resources that can be used. However, the sets of resources may provide different QoSs, that customer's value differently. In Section IV, we describe how the vendor and network negotiate. IP telephony is a common service and the same result would occur each time there is a negotiation. Therefore, the negotiation does not have to occur each time a connection is set up or changed. Instead, the prices can be established with a particular vendor who uses a particular utility function until a new set of resources or a new network is introduced into the mix.

The service program negotiates for resources on each link on a path, determines the utility of the path, then selects the path with the highest utility. In this section, we give an example of a negotiation on a link that has two competing circuit networks and a packet network.

In order to negotiate, the service program assumes a utility function for the customer c. Let a particular service program model all utility functions with the form:

$$U_c(\vec{r},\, p) = U_0\,(1 - (p/p_{c,\,\max}(\vec{r})))$$

where U_0 is the utility if $\vec{r}$ is free. There are many possible utility functions that satisfy the set of constraints that we gave in Section IV. The program that picks the function that most closely models a user's true utility is the program that makes the user most satisfied.

Number the three competing networks so that n_1 and n_2 are the two circuit networks and n_3 is the packet network, and let $\vec{r}_i$ be the resources that c would buy from n_i. The utility of $\vec{r}_i$ is defined by the pair, $(U_{0_i},\, p_{c,\max}(\vec{r}_i))$. Let the utility of $\vec{r}_1$ and $\vec{r}_2$ be defined by the same pair, $(10,\, 10)$, since these are both circuits with the same quality. Let the utility of $\vec{r}_3$ be defined by the pair $(7,\, 7)$. The utility as a function of price, is plotted in Fig. 1. This utility function assumes that the customer prefers the packet connection to the circuit connection if he can save more than three price units.

Assume that the networks have value functions

$$V_{n_i}\,(\vec{r}_i) = \frac{p - p_{n_i,\,\min}\,(\vec{r}_i)}{p_{n_i,\,\min}\,(\vec{r}_i)}$$

where $p_{n_1,\,\min} = 6$, $p_{n_2,\,\min} = 8$, and $p_{n_3,\,\min} = 4$. The network value functions are also shown in Fig. 1. These value functions show one circuit switched network being less competitive than the other and the packet network as being able to provide the lowest price.

Consider three service programs, the first can only use the circuit switched network, n_1, the second can select either circuit switched network, and the third can also select the packet network. In the first situation, n_1 can charge as much as the customer will pay for his service, $p_{c,\,\max}(\vec{r}_1) = 10$, and the added value that the network receives, above what it is willing to accept is $V^{(1,\,1)} = 2/3$. The customer's satisfaction, when he must pay the full amount that willing to pay for a service is $U^{(1,\,1)} = 0$.

In the second situation, n_1 competes with n_2 to deliver the same service. Since the services are identical, the competition is on price. Network n_1 still provides the service but cannot charge more than the least amount that n_2 is willing to accept $p_{n_2,\,\min} = 8$. The customer's satisfaction increases to $U^{(m,\,1)} = 2$, and the added value that the network receives is cut in half to $V^{(m,\,1)} = 1/3$.

In the third situation, the services provided by the packet and circuit networks are not the same. The customer prefers the circuit to the packet service if the prices are the same. If the packet network bids four, the lowest price that it can, and the circuit network bids just under seven, the customer prefers the circuit service, and n_1 wins the auction with $p^{(m,\,m)} = 7-\varepsilon$. In effect, the auction takes place on utility. Network n_1 can provide utility as high as four, while n_3 can only provide utility 3, therefore, n_1 wins the auction at $U^{(m,\,m)} = 3+\varepsilon$. The network's added value decreases to $V^{(m,\,m)} = 1/6$.

In this example, the same network provides the same service in all three situations. However, the ability to route to other networks, that provide the same, or comparable services, reduces the amount that the customer pays that network from ten units to seven units.

VII. Conclusion

Active routing makes it possible for the customer to participate in routing decisions. The customer obtains two important advantages by controlling his path. First, new services, such as those described in Section V, can be deployed by an individual customer instead of waiting for the network to make the service available. When there are multiple approaches to a service, all of the approaches can compete for users. As a result, services can evolve more rapidly and the "best" services, as decided by any metric that the customers consider important, should succeed.

The second advantage is that the customer can improve his satisfaction by forcing networks to compete with one another, as described in Section IV, and demonstrated in the IP telephony example in Section VI. Competition on an end-to-end level occurs, when customers choose between competing carriers. Competition for segments of a connection can now occur in IP telephony networks, especially when customers must move from the circuit to the packet network to complete connections. Intelligence within a network can increase the number and type of choices that are available to the customer.

There are technical challenges that must be addressed. Routing within a network has always been under the control of the network, and routing optimization has been a global optimization to improve the use of the network resources. The networks must put mechanisms in place to control the local optimizations that individuals will perform. In this work, we have described two mechanisms, sandboxes and pricing. Sandboxes protect the resources that have been assigned to a customer and also prevent that customer from acquiring unassigned resources. Pricing can be used to make a customer's best interest and the network's best interest the same. For instance, the network may charge a customer less if he selects the resources that the network would have selected. These indirect mechanisms do not provide the same degree of control as totalitarian decisions.

A number of unexpected problems are bound to occur. For instance, one of the reviewers asked if active routing can cause instabilities. Our instinct is that routing instability, such as occurred with some of the adaptive routing mechanisms that were used in the 1970s ARPAnet, are less likely in an active network. First of all, instability usually occurs when everyone does the same thing. For instance, if everyone listens to the same morning traffic report, and a reporter announces that the delay at the Lincoln tunnel is much longer than at the Holland tunnel, by the time you can get to the Holland tunnel, it may have the longer delay. If everyone drives back and forth between the tunnels, without going through, the routing is unstable. In active routing, there are a wider variety of things that a customer may do and, therefore, it is less likely that the combined effect can be as detrimental. Second, service programs are expected to compete for customers. If some programs operate badly, because of instabilities, they will stop being used. While we do not expect instabilities to be a serious problem, there will probably be instabilities while faulty network implementations are weeded out.

We have not recommended a particular solution in any section of this work. Instead, we have described several alternatives. This is intentional. Active routing opens the network to competition and makes change easier. It does not point toward the "right" way to operate. If there is a right answer today, it is likely that it will change for a different customer set or application, or that the metric will change tomorrow.

References

[1] J. McQuillan, I. Richer, and E. Rosen, "The new routing algorithm for the Arpanet," *IEEE Trans. Commun.*, vol. COM-28, pp. 711–719, May 1980.

[2] N. F. Maxemchuk and M. El Zarki, "Routing and flow control in high speed networks," *Proc. IEEE*, vol. 78, pp. 204–221, Jan. 1990.

[3] C. Metz, "At the core of IP networks: Link state routing protocols," *IEEE Internet Comput.*, pp. 72–77, Sept./Oct. 1999.

[4] D. O. Awduche, "MPLS and traffic engineering in IP networks," *IEEE Commun. Mag.*, vol. 37, pp. 2–47, Dec. 1999.

[5] A. Ghanwani, B. Jamoussi, D. Fedyk, P. Ashwood-Smith, and N. Feldman, "Traffic engineering standards in IP networks using MPLS," *IEEE Commun. Mag.*, vol. 37, pp. 49–53, Dec. 1999.

[6] G. Swallow, "MPLS advantages for traffic engineering," *IEEE Commun. Mag.*, vol. 37, pp. 54–57, Dec. 1999.

[7] T. G. Griffin and G. Wilfong, "An analysis of BGP convergence properties," in *Proc. ACM SIGCOMM*, Cambridge, United States, Aug. 30–Sept. 3, 1999, pp. 277–288.

[8] D. L. Tennenhouse and D. J. Wetherall, "Toward an active network architecture," *Comput. Commun. Rev.*, vol. 28, no. 2, pp. 5–18, Apr. 1996.

[9] C. A. Gunter, S. M. Nettles, and J. M. Smith, "The switch ware active network architecture," *IEEE Network*, pp. 29–36, May 1998.

[10] J. Kurose and K. Ross, *Computer Networking*. Reading, MA: Addison-Wesley, 2001.

[11] D. S. Wallach, D. Balfanz, D. Dean, and E. W. Felten, "Extensible security architectures for Java," in *Proc. 16th ACM Symp. Operating Systems Principles*, Saint Malo, France, Oct. 5–8, 1997, pp. 116–128.

[12] N. F. Maxemchuk, "Active networks in telephony," in Open Arch, New York, Mar. 26–April 7, 1999.

[13] A. Terzis, L. Zhang, and E. L. Hahne, "Reservations for aggregate traffic: Experiences from an RSVP tunnels implementation," in *IWQoS*, pp. 23–25.

[14] D. D. Clark, "The design philosophy of the DARPA internet protocols," in *ACM SIGCOMM*, Stanford, CA, Aug. 1988, pp. 106–114.

[15] I. Stoica, S. Shenker, and H. Zhang, "Core-stateless fair queuing: Achieving approximately fair bandwidth allocations in high speed networks," in *Sigcomm*, Vancouver, BC, Canada, Sept. 2–4, 1998.

[16] S. H. Low and P. P. Varaiya, "A new approach to service provisioning in ATM networks," *IEEE/ACM Trans. Networking*, vol. 1, pp. 547–553, Oct. 1993.

[17] S. Shenker, "Making greed work in networks: A game-theoretic analysis of switch service disciplines," *IEEE/ACM Trans. Networking*, vol. 3, pp. 819–831, Dec. 1995.

[18] E. P. Rathgeb, "Modeling and performance comparison of policing mechanisms for ATM networks," *IEEE J. Select. Areas Commun.*, vol. 9, pp. 325–334, Apr. 1991.

[19] D. Chaum, "Privacy protected payments: Unconditional payer and/or payee untraceability," in *Proc. Smart Card*, D. Chaum and I. Schaumuller-Bichl, Eds., Holland, 1989, pp. 69–93.

[20] S. Dukach, "SNPP: A simple network payment protocol," in *Proc. Computer Security Applications Conf.*, Nov. 1992, pp. 173–179.

[21] G. Medvinsky and B. C. Neuman, "Netcash: A design for practical electronic currency on the internet," in *Proc. 1st ACM Conf. Computer, Communications Security*, Nov. 1993.

[22] T. Okamoto and K. Ohta, "Universal electronic cash," in *Crypto*, Santa Barbara, CA, Aug. 11–15, 1991, Abstracts, 8.7-8.13, pp. 324–337.

[23] M. A. Sirbu, "Internet billing server prototype scope document," Carnegie Mellon Univ., Information Network Inst., INI Techn. Rep. 1993-1, Oct. 14, 1993.

[24] S. D'Amiano and G. Di Crescenzo, "Methodology for digital money based on general cryptographic tools," in *Proc. Eurocrypt*, 1994, pp. 151–162.

[25] T. Eng and T. Okamoto, "Single-term divisible electronic coins," in *Proc. Eurocrypt*, 1994, pp. 311–323.

[26] L. Tang and S. Low, "Chrg-http: A tool for micropayments on the world wide WEB," in *6th USENIX UNIX Security Symp.*, San Jose, CA, July 1996.

[27] A. Herzberg and H. Yochai, "MiniPay: Charging per click on the web," in *Proc. 6th Int. World Wide Web Conf.*, Santa Clara, CA, 1997.

[28] J. Brustoloni and J. Garay, “MicroISPs: Providing convenient and low-cost high-bandwidth Internet access,” *Comput. Networks*, vol. 33, no. 1, pp. 789–802, 2000.
[29] S. Deering, “Multicast routing in internetworks and extended LAN’s,” in *ACM SIGCOMM*, Stanford, CA, Aug. 1988, pp. 55–64.
[30] N. F. Maxemchuk, “Dispersity routing,” in *Proc. ICC*, San Francisco, CA, June 1975, p. 41-10, 41-13.
[31] A. Feldmann, A. Greenberg, C. Lund, N. Reingold, and J. Rexford, “NetScope: Traffic engineering for IP networks,” *IEEE Network*, pp. 11–19, Mar./Apr. 2000.
[32] L. Zhang, S. Deering, D. Estrin, S. Shenker, and D. Zappala, “RSVP: A new resource reservation protocol,” *IEEE Network*, vol. 7, pp. 8–18, Sept. 1993.
[33] D. Fudenberg and J. Tirole, *Game Theory*. Cambridge, MA: MIT Press, 1996, ch. 1.
[34] J. S. Turner, “Design of an integrated services packet network,” *IEEE J. Select. Areas Commun.*, vol. SAC-4, pp. 1373–1380, Nov. 1986.
[35] T. Chen and T. O. Oh, “Reliable service in MPLS,” *IEEE Commun.*, vol. 37, pp. 58–62, Dec. 1999.
[36] D. D. Clark, *Policy Routing in Internet Protocols*. Cambridge, MA: MIT. Press, May 1989.
[37] B. Cheswick, “The design of a secure Internet gateway,” in *Proc. Usenix Summer Conf.*, Anaheim, CA, June 1990, pp. 233–237.
[38] P. Ferguson and G. Huston, “What is a VPN?,” *Cisco Systems; Telstra Internet*, Mar. 1998.
[39] H. T. Kung and S. Y. Wang, *TCP Trunking*. Cambridge, MA: Harvard Univ. Press, Nov. 1998.
[40] R. Sharma, S. Keshav, M. Wu, and L. Wu, “Environments for active networks,” in *Proc. IEEE Int. Workshop Network, Operating System Support Digital Audio, Video*, Piscataway, NJ, 1997, pp. 77–84.
[41] R. A. Guerin and A. Orda, “QoS routing in networks with inaccurate information: Theory and algorithms,” *IEEE/ACM Trans. Networking*, vol. 7, pp. 350–364, June 1999.
[42] D. J. Mitzel, D. Estrin, S. Shenker, and L. Zhang, “An architectural comparison of ST-II and RSVP,” in *Proc. IEEE Infocom*, Toronto, ON, Canada, June 1994.
[43] N. Seddigh, B. Nandy, and P. Pieda, “Bandwidth assurance issues for TCP flows in a differentiated services network,” in *Proc. IEEE GLOBECOM*, Rio de Janeiro, Brazil, Dec. 1999, p. 6.
[44] A. Stephenson, “Diffserv and MPLS—A quality choice—The diffserv and MPLS specs both address IP QOS, only they go about it in different ways,” *Data Commun.*, pp. 73–, Nov. 1998.
[45] S. Bhattacharjee, K. Calvert, and E. Zegura, “Reasoning about active network protocols,” in *IEEE Int. Conf. Network Protectopm*, Austin, TX, Oct. 13–16, 1998, pp. 31–40.
[46] X. Zhao, C. Castelluccia, and M. Baker, “Zhao-X; Castelluccia-C; Baker-M,” in *Proc. IEEE/ACM MobiCom*, Dallas, TX, Oct. 25–30, 1998, pp. 145–156.
[47] E. Amir, S. McCanne, and R. Katz, “An active service framework and its application to real-time multimedia transcoding,” in *Proc. ACM SIGCOMM*, Vancouver, BC, Canada, Aug.31–Sept. 4, 1998, pp. 178–189.
[48] J. F. Meech and S. Abu-Hakima, “Intelligent seamless messaging,” in *Int. Conf. Systems, Man, Cybernetics*, vol. 2, Oct. 11–14, 1998, pp. 1241–1244.

N. F. Maxemchuk (M’72–SM’85–F’89) received the B.S.E.E. degree from the City College of New York, NY, and the M.S.EE. and Ph.D. degrees from the University of Pennsylvania, Philadelphia, PA.

In 1968, he joined the RCA David Sarnoff Research Center, Princeton, NJ. In 1975, he joined Bell Labs Research, Murray Hill, NJ. In the summer of 2000, he joined the Electical Engineering faculty, Columbia University, NY, as a Professor. During his career of more than 30 years, he has authored many papers, been issued many patents, and received numerous awards.

S. H. Low received the B.S. degree from Cornell University, Ithaca, NY, and the Ph.D. degree from the University of California, Berkeley, CA, both in electrical engineering.

He was a Consultant for NEC, CA, in 1991. He was with AT&T Bell Laboratories, Murray Hill, NJ from 1992 to 1996 and the University of Melbourne, Australia from 1996 to 2000. He is now with the California Institute of Technology, Pasadena, CA, as an Associate Professor. His research interests are in the areas of control and optimization of communications networks and protocols, and network security and privacy.

Dr. Low was a corecipient of the IEEE William R. Bennett Prize Paper Award in 1997 and the 1996 US R&D 100 Award. He is on the Editorial Board of IEEE/ACM TRANSACTIONS ON NETWORKING. He has been a Guest Editor of the IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS.