

1 Introduction

The Internet's design principles enabled a remarkable growth. However, these design principles also resulted in inconsistent service quality and sub-optimal security. The challenge is to design lightweight modifications of the architecture that preserve its benefits while correcting its shortcomings. The key ingredient to answer this challenge is to introduce architecture components that enable markets for service and security. With those markets, the providers, vendors, and users will have the incentives to improve the network. In Section 1.1, we describe our goals. In Section 1.2, we provide examples of the current missing markets.

1.1 Project Goals

In our view, the inconsistent service quality and lack of security of the Internet exist not because users aren't willing to pay more for improved service and security [14]. Rather, they exist because the current Internet architecture does not give users any mechanism to signal their desire and willingness to pay. If these mechanisms existed, providers could enrich the services they offer to users, thereby increasing their revenues. Moreover, such mechanisms would improve the network by rewarding parties for contributing to its overall security.

Service and security limitations are not disappearing. History has shown that applications are regularly invented that make use of improved network bandwidth and delays. Service limitations will always restrict the possible applications. Similarly, new security attacks regularly appear that circumvent existing remedies. The focus of our study is the design of an architecture that enables markets for service and security.

Designing such an architecture requires a balance between simplicity and capabilities. In the past, several alternative architectures have been proposed to offer service choice. These architectures, such as ATM and Intserv require users to reveal detailed service quality specifications [6]. However, these service models have not become widely adopted. We argue that the primary reasons for this failure are (i) that they complicate the Internet service model and thus greatly complicate implementation and (ii) that there is no good way for providers to share revenues and to monitor compliance. All the experience to date tells us that a complex architecture that requires structured forms of differentiated service is likely not to be successful.

Our approach is to develop a lightweight architecture that does not overly complicate or constrain the service model but offers just enough structure to create a market for user-specific choice. Rather than imposing rigidly defined classes of service from above, we provide a market mechanisms for users to find out the relevant information from the network to make individual choices which then implicitly reveal the appropriate information about them to the network. Through these markets, providers would have an incentive to improve the network.

We believe that the lack of an appropriate market is again the key missing ingredient to providing a secure network. Negligent users who do not protect their computer by regularly updating their antivirus software and operating system are clearly putting their own computers at risk [15]. But such users, by connecting to the network a computer that may become a host from which viruses can multiply and spread, put all computers on the network at risk. If a market existed that would force negligent users to bear the cost of the risk they impose on others, they would be

more diligent in protecting their computers.

Designing a market for security is tricky because one cannot always trace back security attacks to the computers from which they are launched and spread. Another problem is that making users liable for security problems might create an unacceptable worst case risk for users. Even careful users might get unlucky, and have a virus spread from their computer that eventually leads to a denial of service attack on a major e-commerce data center that causes millions of dollars of damage – just like a careful driver might have a costly car accident.

We explore what would happen if the network were designed with a form of incentivized insurance (InI). Besides risk redistribution, in the presence of negative externalities, as in networking, InI could serve to internalize these externalities. For example, auto insurance curtails reckless driving and speeding. Similarly, incentivizing users to comply with a due care standard for networked devices would make hacking less attractive by making it more difficult to penetrate the network.

In our architecture, users are incentivized to register with a certifying agent (CA). CAs are analogous to car insurance companies; they stamp the user packets with a certificate in a way we describe later. The CAs have limited liability for attacks that are traced back to their subscribers, and thus CAs have an incentive to encourage their subscribers to follow a due care standard.

Since privacy is a key issue in networking, our concept of InI does not mandate that all users be insured. Rather, the incentive is provided by having certified users receive better treatment from the network during times of attack. Users who need perfect privacy should be willing to accept the risk of poorer treatment by the network at such times. We propose a market for security based on such an InI model, which we argue will cure the current demand failure and lead to a more secure network.

1.2 Examples of Missing Markets

To further elaborate on the failure of the network to capture markets for services and security, we discuss a few illustrative examples. Most examples of missing markets that we present are driven by network externalities. These externalities lead to prohibitively high transaction and information costs. Our research combines technological and economic tools to alleviate these externalities.

Provider Revenues: You are a network provider and you see your voice revenues declining because of VoIP. To compensate for this decline, you need to enrich the offering of transport services for service providers. You could offer differentiated transport through your DiffServ network. However, the service provider does not have the capability of charging for premium service.

Streaming Service: A content provider wants to make his video programs available for a fee to customers. The delivery requires significant bandwidth from the network. To increase the volume of this service, the network provider must invest in higher capacity. Without being able to charge for such services, the network provider has limited incentive to improve its capacity. Both the network and content provider miss revenue opportunities.

Urgent Wireless Browsing: You access the network with WiMax and you need some urgent information. You are willing to pay some additional fee to speed up your access. However, the premium access requires a monthly subscription that you cannot justify for a one-time use.

Business-Grade Video Conference: We regularly use NetMeeting and similar applications for video conferences with colleagues and friends. The quality of these applications is reasonable

but not sufficient for important business presentations. Although one might be willing to pay a significant sum for a high-quality video conference, this option is not available.

Emergency Services: New sensors and low-cost radios enable a wide range of emergency services. Such services could be transported by the cellular infrastructure, WiFi hot spots, or future WiMax base stations. Protocols are missing for rapid authentication and automatic location. Technological solutions exist but the market mechanisms are lacking for devices that are not attached to a specific network and service provider.

Networked Sensors and Actuators: Networked sensors will reduce pollution through precise monitoring of the environment, the weather, chemical factories, power plants, and internal combustion engines. Networked sensors and actuators can reduce the water and fertilizer usage of the agriculture industry. Networked actuators can reduce the peak electric power usage by shedding loads judiciously. Supporting widespread reliable sensor and actuator traffic requires nontrivial investments.

New Applications: Clever inventors regularly design new applications. You can view your digital video recorder anywhere over the Internet. You can distribute your TV programs via peer-to-peer. Automobiles are getting networked for status monitoring, theft protection, and traffic information. The wide use of such applications requires network improvements.

Security Breaches: Major security breaches occur frequently where thousands of confidential records get stolen. There are many causes for such security lapses. Preventing them requires implementing some measures that have a substantial upfront cost. Organizations have no direct appreciation of the return on security investments.

Denial of Service: Unlike electric power failures, the cost of denial of service attacks is difficult to measure. Moreover, the source of the attack may be hard to identify and so are the responsibilities of software vendors, network providers, and users. Consequently, liabilities for such attacks are not easy to assess.

2 Markets for Service

It is intuitively clear that providing choice should improve customer satisfaction. Moreover, if the revenues of a provider increase if customers choose its service more frequently, then choice induces competition among providers and creates an incentive to match the desires of customers, thereby further increasing customer satisfaction. These are the familiar arguments for the invisible hand of the free market. More choice and suitable profit allocation should improve the network.

Today, some service providers offer a choice in the form of a premium service that users can subscribe to. However, such choice requires a long term commitment with a monthly fee. Many users may not wish to commit to a higher fee if they need the premium service only sporadically. Also, note that such an implementation is generally inefficient since subscriber to the premium service end up getting it even when their application does not require it. For instance, their file transfers may deteriorate the real-time applications of other premium subscribers. To generate more revenues, providers offer services that are inefficient and not aligned with the interests of users. This situation results from the lack of capabilities of the current architecture.

There are multiple ways of defining service choice. At one extreme one has specified services that guarantee end-to-end throughput, loss rate, delay, and delay jitter (the so-called quality-of-service, or QoS, description). At the other extreme one has a set of different services without

specification that the users learn to choose from based on their immediate experience. In-between, one can think of non-specified services that offer some consistency of characteristics over time and that the users value for their predictability.

We do not exclude any type of service qualification. These different service types correspond to different costs and different user demands. Providers will optimize the services they offer based on customer response. Our concern is with the creation of mechanisms that stimulate the introduction of services, not with the specific types of services that should be introduced.

To illustrate one possibility of service choice, we discuss a model of choice that we call *red-and-blue*. We choose that name to make it clear that the services are not specified by their characteristics. In one implementation, a user engaged in a connection, say through a browser or some conference application, has the option of clicking on a blue icon to switch to a premium service. The blue service costs an additional 10 cents per minute (or is paid through some advertising displayed on the customer's screen). The actual implementation of the blue service is hidden from the customer.

Here are a few possible implementations of service choice: 1) Priority scheduling in WiMax access network; 2) Higher weight in WFQ scheduler at DSLAM of DSL access provider; 3) Guaranteed bandwidth through DOCSIS in cable access; 4) Routing through separate metropolitan access network implemented via an overlay network or through some class-based routing; 5) Priority DiffServ class through a set of providers; 6) Priority in server of content provider.

2.1 Preliminary Results

We illustrate the types of analysis we have in mind with some examples. These examples show the following facts: 1) Offering a premium service allows a network provider to increase its revenue from web browsers; 2) A network provider can increase revenues by differentiating voice service; 3) In a multi-provider setting, individual providers have an incentive to differentiate services. Also, in these examples, we propose an efficient auction mechanism for interaction between network and service providers and we introduce a preliminary formulation of the interactions between content and network providers.

2.1.1 Urgent vs. Casual Web Browsing

Imagine that you are in the middle of an important project where you need to check a number of web pages. You might be willing to pay an extra fee to speed up the browsing. A system providing such an option should generate more revenues for the network providers because it adds value for the users.

Our model explores the effect of such a choice when many users can make it. The interaction of users is an essential aspect of networking. Whereas the perceived value of a bottle of premium wine does not depend on how many other people buy it, the quality of a network service typically depends on the number of users.

Consider a service provider that serves a population of users. The example shows that the service provider may have an incentive to introduce a more expensive premium service. Our model is similar to fluid models investigated by other researchers [23, 22].

A population of $M + N$ web users share a link that can deliver C pages per second, where $N < C < N + M$. Each user spends one second, on average, to read a page he has downloaded

and then asks for the next page. The users have a utility proportional to the rate at which they get pages. Specifically, M users of type A have a utility equal to ar and the other N users of type B have a utility br when they get pages at rate r , where $0 < a < b$. Thus, type- B users have a higher valuation for the speed of download than type- A users. The system charges the users at a rate per minute.

In this model, each user selects the service that maximizes his utility. We compare two versions of the system: a single-class system and a two-class system with strict priority for the premium class. In the single-class system, the charge is p per unit time and each user decides whether to stay with the service or not. In the two-class service, the premium class costs p_1 per unit time and the other class costs p_2 ; the user chooses which class to stay with, if any. The choice may be based on previous experience with the services, on instantaneous experience, or on the reputation of the service as reported by other users. Rules in the user's computer may select the service automatically. For instance, the computer may measure the download rates r_1 and r_2 for both services and compare $ur_1 - p_1$ and $ur_2 - p_2$ where u is the utility value that the user specified for the application ($u = a$ or b in our example).

Fact 1 *The provider revenues are larger for a two-class system and these revenues increase with the system capacity. More precisely, designate by J_1 the maximum revenues of a single class with price p and by J_2 the maximum revenues of a system with two priority classes with prices p_1 and p_2 . The maximization is over the prices. Then*

$$J_1 = \max\{aC, bN\} \text{ and } J_2 = Nb + a(C - N) - (N/M)(b - a)(C - N).$$

The ratio J_2/J_1 is always in $(1, 2)$ and is close to 2 when $aC \approx Nb$ and $N \ll C$. Thus, under some cases of customer demand, service differentiation can almost double the network provider revenues.

Not surprisingly, the benefits of service differentiation depends on the valuations of the users. Interestingly, if the valuations are uniformly distributed, then differentiation does not increase service revenues. Consequently, it is important to study the benefits of service differentiation. This study is non-trivial because of the network externalities caused by interactions among the users and because of the differences in the valuations of various users and applications.

2.1.2 Heterogeneous Applications

We imagine a service provider that serves a mix of web browsers and users of real-time streaming applications such as VoIP and video conferencing. For short, we will call all such real time applications "voice." We want to understand how network architecture affects potential revenue, and to do that we must model the willingness to pay of the customers. There are many effects that influence the willingness to pay of customers of both types, but among the most important effects, and the effect we focus on in this model, is the relative sensitivity to delay of different customer types. Voice users usually require packet delays of less than 100ms while web browsers can tolerate delays that are many times larger as long as the throughput is adequate. Our formulation for the behavior of users and providers is that of Game Theory [16].

2.1) Single Network

We consider a network of capacity C that serves a rate V of voice traffic and a rate W of data traffic. Voice users join as long as the delay is less than T_v and the price is less than u_v . Data users join as long as the delay is less than T_d and the price is less than u_d . There is a total potential demand R_v for voice traffic and R_d for data traffic.

First, we consider single-class network where all the traffic is served in a single queue. For a large number of users, one can approximate the delay through the system by

$$T_d \approx \frac{W}{W+V} \left[\frac{V/W}{C} + \frac{1}{(C-V-W)} \right] \text{ and } T_v \approx T_d - \frac{1}{C}.$$

Second, we consider a two-class network that gives high priority to the voice traffic. Under the same assumptions one can show that

$$T_d \approx 0 \text{ and } T_v \approx \frac{1}{C-V-W}.$$

We analyze the behavior of users and the provider on a numerical example where $R_v = 1, R_d = 2, C = 3, T_v = 1, u_v = 2, T_d = 10, u_d = 1$.

Fact 2 *The maximum revenue in a single-class network is 2.3 and it is 3.9 in a two-class network.*

In a single-class system, the provider maximizes his revenues by charging $p = 1$. In that case, all the data users connect and a fraction 1/3 of the voice users connect.

In a two-class system, the provider charges a price p_v for the high priority queue and p_w for the low priority queue. The maximum revenue corresponds to $p_v = 2$ and $p_d = 1$. In that case, all voice users connect and 95% of the data users connect.

2.2) Tandem Network

The previous discussion assumed that the service quality seen by users is dominated by the architecture of a single provider's network. However, most often, different pieces of the network are owned by different network service providers, and the quality a user sees depends on the configuration of both networks. In this subsection, we explore whether or not interconnected providers have an incentive to unilaterally move from a shared queue to a priority architecture. We discuss the simplest case of two networks connected in tandem.

We consider two networks in tandem, each with capacity C . The networks are owned by two different providers. The choice for both networks is whether to implement a single class or a two-class system with priority for voice. Moreover, the providers optimize their pricing to maximize their individual revenues. Once again we examine a numerical example where $C = 3.5, u_v = 1.5, T_v = 1$. The other parameters are as in the previous example.

Fact 3 *If both networks are single-class, a Nash equilibrium corresponds to each provider charging a price equal to 0.5 and getting revenue of 1.18. If one provider then upgrades to a two-class system, his revenue at the optimal prices increases to 1.93. The second provider then increases his revenue by upgrading to a two-class network.*

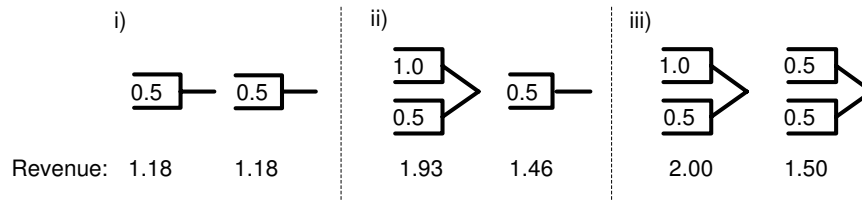


Figure 1: The effect of network configuration choice on revenues. i) Both providers use a shared-queue configuration, charge a price of 0.5 and each earns revenue of 1.18. ii) The first provider upgrades to a priority configuration, and raises his revenue to 1.93. iii) The second provider also switches to the priority configuration and raises his revenue to 1.50.

2.3) Conclusions

This example shows that there can be a “virtuous circle” effect that incentivizes owners of different pieces of the network to improve their network architecture. Figure 1 illustrates the results. However, one can also construct examples where this is not the case. For instance, if we had a tandem network with less capacity, it might be that the delay in each half of the tandem network is more than acceptable for voice. If one provider unilaterally changed configurations to reduce the delay to voice traffic, those improvements would not overcome the poor performance of the other provider’s half of the network. The provider who made the changes would see little or no increase in voice traffic, and thus not see any improvement in revenue.

2.1.3 Interactions between Service and Providers

An architecture that enables efficient utilization of the Internet backbone network is necessary for efficient Internet operation as a whole. Currently, the backbone is owned by various network providers such as MCI, Qwest, AT&T, etc. The service providers such as AOL, SBC, Vonage, etc. buy data and voice circuits from the network providers to provide service to their customers. While there are only a small number of players who are network providers in the United States, the number of service providers is huge. In this context, recall the very fragmented calling card market. Currently, the exchange between the network providers and the service providers happens primarily on a bilateral basis, though bandwidth exchanges are also used. Players on both sides act strategically and, thus, it is difficult to ensure efficient exchange purely through bilateral bargaining processes. What is missing are market mechanisms that ensure provably efficient exchange of network resources.

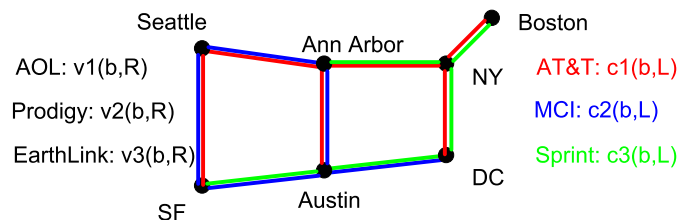


Figure 2: A communication network with buyers and sellers

Let us illustrate the underlying technical problem through a simple example. Consider the network shown in Figure 2. Network providers such as AT&T, MCI and Sprint own bandwidth on various links of the network. Network provider j has a cost $c_j(y, l)$ for making the bandwidth y available on link l . The service providers such as AOL, EarthLink, and Prodigy would like to buy bandwidth on various links to construct routes and build a network, to provide Internet access to their subscribers. Service provider i derives the utility $v_i(x, R)$ when it gets bandwidth x on route R . The route R here is a path through the network, i.e., a set of links.

We would like to determine allocations that maximize the social welfare function $\sum_i v_i(x_i; R_i) - \sum_j c_j(y_j; l_j)$. This function represents the sum of the utilities derived from the services minus the cost of providing them. Thus, the social welfare is the net benefit derived from the services. We call allocations that maximize this function *efficient*.

In [18] we propose an architecture for a network market for efficient resource allocation. The scheme is essentially as follows. A buyer i makes a bid specifying its willingness to pay \$ b_i per unit for up to δ_i units on a route R_i . A seller j makes a bid specifying its willingness to accept \$ a_j per unit for up to s_j on a link l_j . The auctioneer then chooses the allocation that maximizes the social welfare based on the revealed utility functions. It calculates prices according to VCG-type payment functions [32].

Fact 4 [18] *Every Nash equilibrium of the market mechanism proposed above is efficient. Furthermore, the mechanism is ex post budget-balanced, i.e., the sum of payments by buyers equals the sum of payments to the sellers at equilibrium.*

Thus, the mechanism provides the means for distributed optimal allocation of network resources and services. And having designed such a network market mechanism, the next problem is to design a network market architecture and protocols which will implement it. We plan to study how the various technologies for service differentiation such as GMPLS, DiffServ, etc. can be incorporated in this architecture and the additional signaling information required [30] in this regard.

Another market mechanism and a software infrastructure that implements the algorithm is described in [17, 20]. This mechanism, however, suffers from computational difficulties which the current proposal resolves.

2.1.4 Interactions between Content and Service Providers

Recently AT&T and Verizon threatened to block Google's video service unless it shares the revenue generated by these services [7]. This points to the continued vertical integration of the network providers, service providers, and content providers. Lack of an amicable agreement between the content providers and service providers can lead to significant impact on development of Internet services, and investment in developing new network technology and research. Thus, it becomes imperative to study the interactions between these providers and suggest mechanisms that result in efficient network operations.

A reasonable framework to study such interactions is the principal-agent model [24, 25]. The principal (e.g., Google) might offer a contract to one or multiple agents (e.g., AT&T and Verizon). Each agent has a capacity that is unknown to the principal. A provider can make an investment to change its capacity. The principal while offering a contract must take this into account. The contract specifies the payments to the agents as a function of the bandwidth they carry. When

they get these contracts, the agents decide whether to invest and announce their capacity to the principal. The principal then determines a price that it charges to consumers; that price affects the demand and the principal splits the traffic optimally between the providers based on the contracts.

In the economics literature this is called a problem of moral hazard with hidden information but with network effects [5, 25]. The capacity of the agents is unknown to the principal and the other agents. Moreover, there is competition between the agents. The decision taken by each agent affects the other as well. Such problems seem intrinsic to communication networks and apparently have not been studied in the economics literature.

It is perhaps worth mentioning here that such issues are important not only in the interaction between content providers and service providers but also in the context of peer-to-peer networks. AT&T may require that each of the DSL users pay a certain fee if they want to download or share media files. The users might be offered several contracts to choose from. Thus, very similar issues arise in this context as well. If such contractual transactions become common, it might even be necessary to automate them and develop protocols that implement them.

3 Markets for Security

In networked communications, confidentiality, integrity, and availability are generally viewed as the primary ingredients of security [4, 21]. Of these, confidentiality and integrity are generally provided by cryptography. In principle these can be provided over any underlying network architecture, as they involve modifications in the entities transported by the network, e.g. packets. Some aspects of availability are local and may similarly be viewed as being largely independent of the network architecture. Examples of local attacks against availability include cutting the power supply, locking a door, or stealing the hardware. However, for the most part, availability is very much tied to the network architecture.

The fundamental tradeoff that needs to be addressed in providing network availability is that between privacy and accountability. The way the Internet has been designed makes it too easy to be unaccountable. This has led to the situation today where there are a large number of ad-hoc fixes for resolving individual problems of availability, such as SYN flood attacks, stream DoS attacks, HTTP proxy attacks, ICMP flood attacks, reflector attacks, etc., while new problems keep cropping up [1, 2, 9, 10, 11, 19, 26, 27, 34]. Given the enormous economic cost of denial of service attacks, addressing how to provide availability must be an essential part of any new architecture.

The Coase theorem [12] asserts that an optimal resource allocation is achievable through market forces, irrespective of legal liability assignment, if information is perfect and transactions are costless. When the assumption of perfect information does not apply, or when transaction costs are present, intervention by an external force such as a standards body may be desirable [28, 29]. We suggest that a standards body such as the IETF introduce an architectural standard to facilitate a more efficient market for security. From the Coasian viewpoint, the architecture enhances efficiency because it reveals more information about the behavior of users. An ideal standard would be structured enough to reveal enough behavior information to enhance efficiency, flexible enough to allow future innovations, and of course it must be implementable. We outline what one such architectural standard could look like.

3.1 Security Mechanism

We propose an architecture where users are incentivized to register with a certifying agent (CA), which is analogous to the concept of auto insurance. The CAs create an incentive for their clients to use due-care standards by offering them better service. Moreover, the scheme provides an incentive for software vendors to make their products more secure.

In broad terms, in order to get a certificate, the user has to pay a CA. The CA classifies its clients into groups partly based on the software installed in their computer. These headers are created to guarantee that the packets are coming from the correct source or have the correct permissions. The CA then gives the user a key with which to stamp its packets as being good; the key is also shared with filters in the network. This idea may be thought of as a weak form of digital signature, where what is being authenticated is that the user is certified. However, we do not need strong keys for our approach, so the computational overhead should be manageable, see for instance [8]. Further, as we see below, we allow for scenarios where the authentication is compromised.

The goals guiding our vision are the following:

- We need privacy-aware technology. Data about traffic should be available for surveillance (and backtracking) but should nevertheless protect the privacy of users when there is no attack.
- Users have the option of having perfect privacy anyway by not subscribing to a certifying agency. They pay the implicit cost that when a lack of availability situation arises they get lower priority. More generally one can imagine quantifying various levels of privacy short of absolute privacy. Designing these would be a research issue.
- In reacting to lack of availability issues, speed is of the essence. The notion of a stamp has an advantage in that it is possible to quickly react to such an event, first by immediately dropping unstamped packets, which one would expect are the main cause of such events if the certification is done intelligently. Secondly, since there would be relatively few CAs, compared to the number of users, the network can more quickly interact with the CAs.

Our ideas are closest in spirit to those of Crocker [13] and Bellare [3], see also [31] and [33].

3.2 Security Mechanism Implementation

An approach for certification that builds off a key distribution architecture similar to the one already in use today suggests a basic implementation for our ideas. This is a preliminary proposal and will evolve as we better understand its possible weaknesses.

Consider a certifying agent with name A . Agent A periodically picks a key K and a hash function on user names, h . It sends (K, A) to all routers.

Consider user U who is certified by agent A . We think of U as a location specific name (or a good hash of such a name) such as a MAC address, so that it is very difficult to spoof as if the packet originated from U when it does not. Agent A sends user U the pair (K, h) and a number S that identifies characteristics of the software installed in U 's computer. For instance, S could indicate whether U has installed the latest security patches for his operating system. Suppose the user U wants to send data D with header H . The user will first create $N = h(U)$, a hash

of its name, then $C = f(D, H, N, A, K, S)$. The string (C, N, A, S) is added to the packet, so the overall packet is $[D|H|A|N|S|C]$. When needed, a router can check the certificate is okay by computing $f(D, H, N, A, K, S)$ and verifying if this equals C . We think of the pair (N, S) as a group associated to the agent A .

Some filters judiciously placed in the network monitor statistics for each group for each CA and for uncertified packets. They detect an availability attack using decision rules based on these statistics. When there is an availability attack the filters first immediately drop all uncertified packets (this means packets that do not meet the certificate check). If the attack still persists, the filters then give lower priority to misbehaving groups and inform the corresponding CAs. The incentive of the CAs is then to isolate the potentially misbehaving users in separate groups to avoid exposing well-behaving users to poor treatment. To improve this isolation, the CAs may install sniffers that maintain traffic history from their clients; they may also install software monitors in clients' computers; they could also have a process for certifying specific applications. The CAs may separate in different groups users that meet different degrees of certification. This process would create incentives for software certification.

All that is needed for this protocol is for K, h and S to be securely delivered to filters and to certified users. This sort of thing is already being done nowadays via key exchange protocols.

Implementing a scheme such as this will require installing some filters at strategic locations. The network providers have the incentive to install such filters because they improve the service they offer to their users. By making the standards open, we leave the filtering rules flexible and open to innovation. The hardware requirements for the filters are typical of line-rate encryption and traffic monitoring similar to what is installed in current systems.

Other schemes for implementing the certification idea, based for instance on public key techniques, could also be envisioned. Comparing different schemes according to their merits will be a research topic. It will be an important part of our research to investigate potential flaws in our schemes and make them as watertight as possible. It is also important to make the protocol as light as possible in terms of additional processing that needs to be done at the routers, and in terms of the overhead imposed on the packets.

3.3 Benefits of Security Choice

In this section we illustrate via a simple model how one could analyze the potential benefits of security choice in the framework of our idea.

Consider a network being used by K users (K is large). Each user may correspond to multiple source-destination pairs. For the purposes of this illustrative example, assume that user k on route $p \in \mathcal{P}_k$ offers deterministic flow f_{kp} . Let C_e denote the capacity on link e that the network has allocated to the collection of users. Thus:

$$R_e(F) = \sum_{k \in \mathcal{K}} \sum_{p \in \mathcal{P}} f_{kp} 1(e \in p) \leq C_e ,$$

where $F = [f_{kp}]$ is the $\mathcal{K} \times \mathcal{P}$ matrix of flows, $\mathcal{P} = \cup_{k \in \mathcal{K}} \mathcal{P}_k$, and $1(e \in p)$ denotes the indicator that link e lies on path p .

Let $d_e(\cdot)$ be a measure of performance along link e ; this is a function of C_e and R_e . Thinking of C_e as fixed for the purposes of this illustrative example, $d_e(\cdot)$ depends only on R_e . Let $D_p(\cdot)$

denote the performance along path p . This is a function of $(d_e, e \in p)$. The performance seen by user k depends on $((D_p, f_{kp}), p \in \mathcal{P}_k)$.

Now, assume that a *malicious* user enters the network through multiple (source, destination) pairs. A malicious user is modeled as introducing excessive flow into the network according to some pattern. This changes the realized flows of the legitimate users, thereby causing all of them to suffer a performance loss till the malicious user is detected. We think of the malicious user as someone who has managed to get certified by the network while nevertheless intending to later attack the network: attacks by uncertified flows are thought of as being dropped essentially instantaneously when a severe performance degradation is detected.

Let π denote the probability of a malicious user attacking the network. Let τ denote the time it takes to detect the attack by the malicious user.

Requiring the users of a network to be certified can be modeled as reducing the probability that a malicious user can enter the network. The charge for this certification process reduces the utility of legitimate users, but it can be applied towards the cost of better detection methodologies, thereby reducing the time it takes to detect an attack by a malicious user. The tradeoffs involved can be used to give guidelines for the design of the certification process.

Specific problems that such a toy model can address include: the statistics that the filters should maintain (this determines the speed with which it can detect an attack by a malicious user that it has certified, but saving more data would require more cost); how should one distribute among the CAs the cost incurred through performance loss during an attack so as to appropriately incentivize the CAs to better manage their users?

4 Justification for Collaborative Effort

We expect the proposed collaboration with Bell Labs, Lucent Technologies to provide a “reality check” in our research and the resultant recommendations. This collaboration would be very useful in two important ways.

(1) Lucent has extensive experience in a number of technologies that could facilitate the architecture we are proposing. For example, recently Lucent has been aggressively pursuing the IP Multimedia Subsystem (IMS) based architecture for both wired and wireless networks. This architecture is very well suited for implementing dynamic pricing, policy decisions and enforcement, resource management, and novel security functions. Another example of a desirable environment at Lucent that would benefit our research is its VitalSuite network and service management software product. This product has a number of features relevant to our proposal including the support for on-demand network performance monitoring. (2) Lucent has significant experience with a vast range of network design activities for many network/service providers. Lucent would bring into our deliberations information concerning the nature of the Service Level Agreements (SLAs) (between networks as well as between an enterprise and a network/service provider), the willingness to pay by a user for various QoS and security features, the real-life constraints on various network based QoS and security solutions, and the implementation feasibility of the economic mechanisms we are proposing. Some of the existing network/service provider customers of Lucent would also be good sources of feedback on the practicality and desirability of the new architectural ideas emerging from our research.

John Musacchio brings an expertise in game theory and pricing of network services in addition to his skills in modeling large systems. Moreover, his industry contacts through the Silicon Valley Center of UCSC will provide valuable feedback on our work. John and his student will develop models and contribute to their analysis and simulations.

5 Research Plan

5.1 Project Plan

Since we are proposing new mechanisms, data is not available about how users would respond to them. Moreover, the state of understanding of the issues involved is insufficient to design meaningful large-scale experiments. Ultimately, once the fundamental ideas are validated by the academic community, the providers will conduct the appropriate experiments. Accordingly, we will use modeling to validate the fundamental principles of our architecture proposals. Our research will consist of the following tasks.

User Service Models

- T1.1: Comparison of service models: efficiency and complexity for different specifications; service information made available to users.
- T1.2: Analysis of benefits of service monitoring: how much market efficiency is gained by monitoring and at what cost.
- T1.3: Quantification of the inefficiency introduced by the end users being unable to discriminate between the contribution of different providers?

Service-Level Agreements (SLAs) between Providers (Network, Access, and Application)

- T2.1: Study of the effect of SLA definition granularity on the markets for inter-provider services.
- T2.2: Examination of the roles of measurements and enforcement in inter-provider contracts.
- T2.3: Comparison mechanisms for profit allocation among providers.
- T2.4: Study of the architectural changes required for implementing these SLAs.

5.1.1 Market-Enabling Architectures for Security

- T3.1: Invent and contrast certification schemes from the point of view of attacks on such schemes, and user privacy.
- T3.2: Study the statistical issues needed for the information aggregation from the point of view of what needs to be stored by CAs and how to efficiently carry out traceback.
- T3.3: Study how to define CA liability: how much and how to share the liability. Study pricing strategies for the InI scheme.

5.2 Timetable

Figure 3 shows a breakdown of the research agenda into subtasks with a tentative timetable.

	Year 1				Year 2				Year 3				Year 3					
	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4		
T1.1 Service Models	█				█													
T1.2 Service Monitoring					█	█				█								
T1.3 Visibility									█	█								
T2.1 Provider SLAs					█	█				█								
T2.2 Measurement									█	█				█				
T2.3 Revenue Sharing	█	█			█													
T2.4 Architecture for SLAs									█				█					
T3.1 Certification Schemes	█	█			█													
T3.2 Aggregation Issues					█	█				█								
T3.3 Liability Issues									█	█								

Figure 3: Schedule of Tasks

6 Results from Prior NSF Support

6.1 Venkat Anantharam

Vertically-integrated primitives for a bufferless all-optical packet switched network

ECS-0123512

Constance Chang-Hasnain, Alan Willner, and Venkat Anantharam

Nov 1, 2001 - Oct 31, 2005

The project investigates design of a bufferless all-optical packet switched network by combining erasure recovery coding at the packet level with deflection routing of optical packets and spatial statistical multiplexing. The project studies the problem across all layers, including devices, system integration, and architecture. Anantharam's work focused on systems level issues. Chang-Hasnain focused on building a novel all-optical switch and Willner focused on all optical addressing and routing issues.

Work done by Anantharam includes: (i) invented a technique to exactly emulate priority queues using a switch with clocked delay lines; (ii) developed an environment to study deflection routing based bufferless networks; (iii) studied codes immune to deletions (erasure can be viewed as a deletion at the packet level).

[1] Anand Sarwate and Venkat Anantharam, "Exact emulation of a priority queue with a switch and delay lines". *Queueing Systems : Theory and Applications*, to appear.

[2] Sagnik Ghosh and Venkat Anantharam, "Bufferless All-Optical Networking with Erasure Codes". *Berkeley Scientific Journal*, Vol. 9, No. 2, pp. 115 -118, Fall 2005.

[3] Lara Dolecek and Venkat Anantharam, "Run-length properties of a Red-Muller RM(1,m) code with applications in channels with at most one synchronization error", *Proceedings of the 42nd Allerton Conference on Communications, Control, and Computing*, Champaign, Illinois, 29 Sep - 01 Oct. 2004, pp. 270 -279.

[4] Lara Dolecek and Venkat Anantharam, "On Array-based LDPC Codes in Channels with Varying Sampling Rate". Report No. EECS-2006-7, EECS Department, University of California,

Berkeley, 2006.

[5] “Bufferless all-optical networking with erasure codes”, Venkat Anantharam, Proceedings of the 2002 IEEE Workshop on Information Theory, Bangalore, India, October 20 -25 2002, pg. 19.

6.2 Jean Walrand

Flexible MAC Protocols for Configurable Radios (2004-2007)

This project studies the design of a new family of MAC protocols that has two innovative features. First, it allows any devices with vastly different data rate, communication range, and transmission power requirements to communicate with each other as long as their MAC protocol belongs to this family. This family of MAC protocols use a common physical layer framework based on OFDM to provide connectivity among a wide variety of devices. Second, this new family of MAC protocols will allow applications to make explicit tradeoffs among data rate, communication range, and transmission power when communicating to their neighbors.

[1] “Comparison of Multi-Channel MAC Protocols,” Wilson So, Jeonghoon Mo, and J. Walrand. MSWiM 05.

[2] “Design of a Multiple Channel Medium Access Protocol for Ad-Hoc Wireless Networks,” Ph.D. Thesis, Wilson So, March 2006.

Economics Mechanisms for Networks (2004-2005)

This proposal studies the games that network users and providers face when making strategic choices. These games are dynamic and the players have asymmetric information. A central issue concerns the incentives for network evolution: Do providers have investment incentives to improve the network or are they impeded by the limitations of other providers. The proposal explores mechanisms that provide suitable incentives and examines their scalability, security, potential for incremental deployment and extensibility.

[1] “Pricing and Revenue Sharing for Internet Service Providers,” Linhai He and Jean Walrand. To appear in JSAC 2006.

[2] “Pricing Differentiated Internet Services,” Linhai He and Jean Walrand, INFOCOM 2005

[3] “Pricing and Revenue Sharing Strategies for Internet Service Providers,” Linhai He and Jean Walrand, INFOCOM 2005.

[4] “An Efficient Auction Mechanism for Multiple Divisible Goods,” R. Jain and Jean Walrand, submitted to CDC 2006.

REFERENCES CITED

- [1] T. Alpcan and T. Basar, "A Game-Theoretic Approach to Decision and Analysis in Network Intrusion Detection," *Proceedings of the 42nd IEEE Conference on Decision and Control*, Maui, Hawaii, Dec. 2003, pp. 2595 -2600.
- [2] J. S. Baras and T. Jiang, "Cooperative Games, Phase Transitions on Graphs, and Distributed Trust in MANET," *Proceedings of the 43rd IEEE Conference on Decision and Control*, Atlantis, Paradise Island, Bahamas, Dec. 14 -17, 2004, pp. 93 -98.
- [3] S. Bellovin. "RFC 3514: The Security Flag in the IPv4 Header". Available at : <http://www.ietf.org/rfc/rfc3514.txt>
- [4] M. Bishop. *Introduction to Computer Security*. Addison-Wesley, 2004.
- [5] P. Bolton and M. Dewatripont. *Contract Theory*. MIT Press, 2005.
- [6] R. Braden, D. Clark, and S. Shenker, "Integrated Services in the Internet Architecture: an Overview," RFC 1633, Internet Engineering Task Force, 1994.
- [7] H. Bray, "Telecoms want their Products to Travel on a Faster Internet," *The Boston Globe*, December 13, 2005.
- [8] Broadcom product brief BCM 5841. "Multi-Gigabit Security Processor". <http://www.broadcom.com/collateral/pb/5841-PB03-R.pdf>.
- [9] G. Carl, G. Kesidis, R. R. Brooks, and S. Rai, "Denial-of-Service Attack-Detection Techniques," *IEEE Internet Computing*, Vol. 10, No. 1, Jan. -Feb. 2006, pp. 82 -89.
- [10] R. Chen and J.-M. Park, "Attack Diagnosis: Throttling Distributed Denial-of-Service Attacks Close to the Attack Source," *Proceedings of the 14th International Conference on Computer Communications and Networks*, ICCCN 2005, 17 -19 Oct. 2005, pp. 275 -280.
- [11] R. K. C. Chung, "Defending against Flooding-Based Distributed Denial-of-Service Attacks: A Tutorial," *IEEE Communications Magazine*, Vol. 40, No. 10, Oct. 2002, pp. 42 -51.
- [12] R. H. Coase, "The Problem of Social Cost," *Journal of Law and Economics*, Vol. 3, 1960, pp. 1-44.
- [13] S. D. Crocker, "Protecting the Internet From Distributed Denial-of-Service Attacks: A Proposal," *Proceedings of the IEEE*, Vol. 92, No. 9, Sep. 2004, pp. 1375 -1381.
- [14] R. Edell, and P. Varaiya, "Providing Internet Access: What we learn from INDEX," *IEEE Network*, Vol. 13, No. 5, Sept.-Oct. 1999, pp 18-25.
- [15] M. Fisk, "Causes and Remedies for Social Acceptance of Network Insecurity," *Workshop on Economics and Information Security University of California*, Berkeley May 16-17, 2002.
- [16] D. Fudenberg and J. Tirole. *Game Theory*. MIT Press, 1991.
- [17] R. Jain and P. Varaiya, "Efficient Market Mechanisms for Network Resource Allocation," *Proc. IEEE Conf. on Decision and Control*, 2005, pp. 1056 -1061.

- [18] R. Jain and J. Walrand, "An Efficient Auction Mechanism for Multiple Divisible Goods," 2006, preprint. Available at <http://robotics.eecs.berkeley.edu/~wlr/Papers/JainWalrand06.pdf>
- [19] G. Jiang and G. Cybenko, "Temporal and Spatial Distributed Event Correlation for Network Security," *Proceedings of the 2004 American Control Conference*, Boston, Mass., Jun. 30 -Jul. 2, 2004, pp. 996 -1001.
- [20] C. Kaskiris, R. Jain, R. Rajagopal and P. Varaiya, "Combinatorial Auction Design for Bandwidth Trading: An Experimental Study," *Proc. Intern. Conf. on Experiments in Econ. Sciences*, 2004.
- [21] C. Kaufman, R. Perlman, and M. Speciner. *Network Security: Private Communication in a PUBLIC World*. Prentice Hall, 2002.
- [22] P. Key, L. Massoulié, A. Bain, and F. Kelly, "Fair Internet Traffic Integration: Network Flow Models and Analysis," *Annales des Télécommunications*, 59, 2004, pp 1338-52.
- [23] S. Kumar, and L. Massoulié, "Integrating Streaming and File-Transfer Internet Traffic: Fluid and Diffusion Approximations," submitted to *Queueing Systems*, 2005. <http://www.stanford.edu/~skumar/preprints.htm>
- [24] J-J. Laffont and J. J. Tirole. *A Theory of Incentives in Procurement of Regulation*. 2nd ed., MIT Press, Cambridge, Massachusetts, 1994.
- [25] J-J. Laffont and D. Martimort. *Theory of Incentives: The Principal-Agent Model*. MIT Press, 2001.
- [26] R. Mahajan, S. M. Bellovin, S. Floyd, J. Ioannidis, V. Paxson, and S. Shenker, "Controlling High Bandwidth Aggregates in the Network," *Computer Communications Review*, Vol. 32, No. 3, Jul. 2002, pp. 62 -73.
- [27] J. Mirkovic and P. Reiher, "D-WARD: A Source-End Defense against Flooding Denial-of-Service Attacks," *IEEE Transactions on Dependable and Secure Computing*, Vol. 2, No. 3, Jul. -Sep. 2005, pp. 216 -232.
- [28] D. North. *Institutions, Institutional Change and Economic Performance*. Cambridge University Press. 1990.
- [29] M. Olson. *The Logic of Collective Action*. Harvard University Press, 1971.
- [30] J. Shu and P. Varaiya, "Pricing Network Services," *Proc. INFOCOM*, Vol. 2, 30 March -3 April, 2003, pp. 1221 -1230.
- [31] K. J. Soo Hoo, "How Much Is Enough? A Risk-Management Approach to Computer Security," *Ph.D. Thesis*, Stanford University, June 2000.
- [32] W. Vickrey, "Counterspeculations, Auctions, and Sealed Tenders," *J. Finance*, 16:8-37, 1961.
- [33] B. Weis, "RFC 4359: The use of RSA/SHA-1 Signatures withing Encapsulating Security Payload and Authentication Header," January 2006.
- [34] J. Yen and R. Popp, "Homeland Security," *IEEE Intelligent Systems and Their Applications*, Vol. 20, No. 5, Sep.-Oct. 2005, pp. 76 -86.