CONSIDER THE DOOR

M. Graff, Sun Microsystems, Inc.
23 June 1997

INTRODUCTION

Thank you. I will begin my prepared remarks this afternoon by introducing my company and myself.

Sun, as many of you know, manufactures workstations and produces UNIX system software. We sell more workstations than any other company in the world. We like to think of ourselves as one of the keys to the success of the World Wide Web and the Internet. Recently, of course, we have become known as the company that invented Java™.

I think that is about all you will hear from me today about our product line. You see, my job is not to work on product, but rather to design or specify the software, procedures, and strategies that protect our own internal network. We call our network the "SWAN", for Sun Wide Area Network. It encompasses over 51,000 workstations, and spans about 85 countries. We do all the company business on it.

This afternoon I will share with you some ideas we are working on to help improve the security of our network. These ideas are not now part of any product, and may never be. They are part of our plan to better serve our internal customers. Still, the best of Sun's innovations have had away of becoming part of the way the Internet does business, so it may be that what you hear from me today will be directly relevant to the way you run your own networks in the future.

THREE SECURITY DILEMMAS

The redesign of our network security scheme seeks to resolve three dilemmas. These dilemmas, experience has led me to believe, are inherent in the distributed computing model in use today around the world. I expect you will find them familiar, if you, too, are responsible for networks, or audit them. These are very practical, mundane issues.

The first issue is the perceived conflict between security and ease of use. Are these two design considerations mutually exclusive? They are commonly believed to be. And of course, no security measure can be effective if the person who must put it into practice will not cooperate. As architects, we accomplish nothing by providing encryption tools to protect electronic mail if our users continue to send messages in plain text. As makers of policy, we waste everyone's time if we insist that no one shall share passwords--especially if, as is so often the case, we offer no reasonable alternative--then take no steps to ensure (or even measure) compliance. It is indeed an unusual community of users who will comply with security measures that are obstructive, restrictive, or simply not well explained. But aren't these commonplace complaints today?

A second, related challenge is that in today's enterprises we tend to enforce uniform policies and procedures throughout the network. This is especially true

in a company like Sun, which deploys workstations at every desk. The practice arises in part because, if one computer inside the corporate firewall is compromised, the security of all others is endangered as well. Once an intruder has established a foothold inside, it's very difficult to withstand his attacks, so every computer should be hard to break into. We end up applying high standards all over. Have you ever tried to explain this reasoning to a frustrated engineer? I have, and I came away resolved to find a new way to do my job and not get in the way of his.

Now for my third dilemma. Here is a sample scenario. I'm sure it will sound familiar. The department responsible for network security standards and policies issue an edict. Let's say it concerns all of the company's external Web sites. And just for fun, let's assume that the edict mandates that, for security purposes, all advertising copy describing the firm's products must be encrypted, and kept secret! That way, the security memo explains, we will keep our competitors in the dark about what we are doing. Now, if you were the vice president of product development, would you advise your troops assembling the Web site to obey this ridiculous security policy at all costs? No. Wouldn't you tell them to proceed as planned? Then, you would take the heat from the security department, explaining (in small slow words, perhaps) just how much revenue the company would lose if you carried out the policy.

Here we see the dilemma: how does a company resolve the conflict between security measures and business needs? Today, I believe, they are usually resolved in favor of the more powerful executive, or more popular department. (This means that the security department usually loses.) There must be a better way.

At this point, like any good engineer, I will lay out for you in recapitulation the problems I am trying to solve. They are:

1. The seemingly inherent conflict between security and ease of use.
2. The need to apply to every user the most stringent protections required of any user.
3. The lack of an objective method to resolve conflicts between security and business needs, especially those, which involve differing business units inside the same company.

Now, at the risk of removing any suspense from these proceedings, I will list for you the solutions I propose to each of the three conundrums.

1. For ease of use, I propose to make security invisible. In re-designing the network, I will try to act like a magician and make the obstacles disappear. Mind you, I don't know how to do this yet, but I have an approach in mind. I will share it with you in a few minutes.
2. To be able to offer customized security services and standards within the SWAN, we plan to break our monolithic network into a set of virtual Private Networks. I don't know how to do this, either, but we are close to a solution here.
3. Finally, in order to facilitate the balancing of security versus business needs, I'd like to see us institute what I call "unsecurity credits". It's a variation on the usual IT charge back scheme. Since it's the simplest to put into practice, I'll explain this idea in detail first.

UNSECURITY CREDITS

"Unsecurity Credits" proceed by analogy with environmental pollution credits. If you're not familiar with those, it's a system fairly popular in the U. S. that allocates to industrial manufacturers the legal right to pollute the environment, as a byproduct of their operations, up to a certain limit. If the company pollutes past that limit, fines ensue. If they exceed expectations, and operate more cleanly than required, they can sell the leftover credits to other companies, at whatever price they can get for them. The genius of this arrangement is that it gives all concerned parties a rational means to decide on an optimal course. For such a credit program to work, naturally, the issuing authority must be able to measure pollution quite accurately--and assign a meaningful cost for it, too.

I argue that the network is our computing environment. Unsecure practices pollute it. By instituting credits we would make it easier for the company as a whole to make balanced security decisions. Suppose, for example, that a business unit wants to operate an unsecure Web site for 90 days. In most companies today, the decision whether or not to go ahead would come down to a power struggle. In a few lucky companies, a risk analysis would be undertaken, producing an estimate, say, that there is a 10% likelihood of a $10,000,000 embarrassment. The lucky companies could then balance the expectation of a $1,000,000 risk {that's the estimated loss multiplied by its probability) against the prospective new business, and make its case.

There's an advantage to augmenting these risk assessment calculations with an "unsecurity credit" scheme. Unsecurity credits would force the company to make policy decisions one time, up front, at the time credits are assigned. Evaluate your assets, determine the total amount of "unsecurity" you are willing to sustain, and allocate the credits. The arm-wrestling will take place then, in the full light of day with all of the decision makers engaged. Later, when specific opportunities arise, the individual business units can decide for themselves how to invest the degree of unsecurity they have been allocated.

Listening to myself describe this arrangement, I realize that I am preaching the advantages of making up budgets, to a group of auditors!

It is a lot like an unsecurity budget, isn't it?

By the way, I haven't sold this scheme at Sun yet. I would be interested to know what members of this audience think of the idea, especially if it has been tried before, and failed!

INTERARTICULATED VIRTUAL PRIVATE NETWORKS

I'll return now to the idea I mentioned earlier about breaking up the monolithic WAN at Sun. This project we definitely plan to undertake. In fact, a small prototype is scheduled to be rolled out next week.

The idea is to use encryption to implement a set of Virtual Private Networks, or VPN's. To simplify, let's say that we will put all of the Accounting Department on one VPN, the company executives on a second, and all of Engineering on a third. The idea is to class together users who have comparable security needs, and tend to use the same sets of software applications. Accountants, for example, we might constrain to use a complete set of accounting tools, but deny access to compilers. The data they work on would have the strongest possible protection. Executives, too, would find their systems {and all of their VPN)

strongly protected as well. The set of applications they would have access to would be even smaller, perhaps being limited to electronic mail and calendar programs. But they would have a special capability: since they travel a great deal, they would be able to operate all of their applications while on the road--something we might never let the accountants do. Engineers would not have access to the other two VPN's, but would in compensation be accorded a virtual playpen, full of toys.

The idea of partitioning a network in this way is not new. What is new is that:

1. We can implement the VPN's on top of a network that is still physically connected all around. The network group can still treat it as a monolith, when that is desirable (for example, during backups, or security sweeps).
2. We can regulate the passing of information among the VPN's, even though they are still physically connected. After all, how many Engineers need access to the Human Resources database? A few, perhaps, but not thousands. Today there is no bar to any of our more than 50,000 systems initiating a connection to any other system. That is 2,500,000,000 different possible machine-to-machine pathways. Only a handful of those possibilities are legitimate. Why allow the rest?
3. Because we will be using point-to-point encryption, the VPN's can span the firewall with good security. This means that traveling executives, or engineers working from home, will see the same environment while on the road that they do at home. It also makes it possible for us to set up a VPN that includes our strategic partners, and have that VPN be part of some networks outside our firewall, and some inside.

The main objection to this scheme, as I have quickly explained it, is that many employees will not fit comfortably into any category. We have executives, for example, who are fine engineers. We would not want to (and, frankly, would not be allowed to) cut off their access to the compilers and other engineering tools they want to use occasionally. What can we do about this?

For now, we will treat the exceptional users as a separate class. Their VPN can span the entire physical network. I'll explain this a little more as I discuss the remaining of my three proposals. Remember? It's my magic trick.

INVISIBLE SECURITY

I want to make security invisible. I don't know if I can do it, but I believe I can get close. How? Consider the humble door.

I have been happy, this past month, to lie on the floor of our family room with my infant son. He plays with the door. I watch. I see him exert pressure with one hand on the door's vertical surface, then stretch out the other hand, intending to catch the door as it closes. He has yet to catch it, though: when he pushes, it moves, like a satellite orbiting the hinge, falling always short of his outstretched hand and curving into the door frame. We have repeated the experiment countless times, to unending wonder--his and mine. He never catches the door. He expects it to move in a straight line, you see: but the door always seems to swerve at the last moment and engages the latch.

I spend these moments pondering the nature of doors (and watching out for little fingers). Now that I have seen the door through his eyes, it looks different to me.

Do we adults consider the door in its frame, the knob and lock, as barriers we must overcome to gain access? We do not. Rather, we call the door the prototypical access device!

Yet it is built into the wall. It is part of the wall, part of the barrier. But because it can swing out of our way--and because we know that it will--we think of it as an access point. Never--except for those times when the two-year-old accidentally locks the bathroom door from the inside--never do we think of the door itself as a barrier.

The wall, the door, the knob, lock, and key represent not only successive refinements in the historical technology of security, but also a series of channels, which, successively, narrow our focus. As we approach, they guide us pacifically into the frame and through the wall. The door, you see, swings away.

So we shall build a doorway into our network. Guests, members of the Sun family, and friends will pass through it into the entry hall of a vast mansion. Inside, they will find an array of inviting corridors and many rooms, laid out differently for each visitor: a personal VPN. They will navigate around by using whatever hardware device they came to the door with, whether it be a workstation, network computer, or laptop.

As visitors perambulate inside the mansion, they will be accompanied by a virtual assistant. For some, he may take the form of a butler. For others, he may resemble a St. Bernard (that would be Network, the Sun mascot). The assistant will answer questions, and report suspicious activity to one of the groundskeepers. Hey, that could be me!

CONCLUSION

As you can see, there are some important details yet to be worked out. I am sure part of these ideas sound like a fantasy. I believe it can be done. (The magician's secrets are great design and a little misdirection.) Furthermore, in my view, I am not leading the way, but rather following the demands of our users and customers as they insist on an easy, practical way to do what they need to do.

I remember a plaque I kept on my office wall during my days in management. It reads, "Which way did they go? How many of them were there? When did they pass? I must find them: I am their leader!

From my point of view, you see, I am just trying to catch up. I suppose, in some ways, we all are. So, because we're all in this together, I feel entitled to ask for your help, your guidance, and your indulgence as we go about inventing our little piece of the future. And remember: if we knew what we were doing all of the time, we wouldn't be making history.

Thank you.