The Great Disconnection?

Michael F. Schwartz

CU-CS-521-91 February 1991

Department of Computer Science
University of Colorado
Boulder, Colorado 80309-0430
(303) 492-7514
electronic mail contact: schwartz@latour.colorado.edu

Abstract

In this paper we present measured data about the types of sites reachable by upper layer services (such as mail and "finger") on the global TCP/IP Internet. We analyze changes in this type of reachability by comparing data from two world wide measurements, conducted 6 months apart. Our impetus for this analysis is to examine the extent to which sites are reducing their accessibility from the Internet, in response to increasing security concerns. We consider upper layer service connectivity instead of basic IP connectivity because the former indicates the willingness of organizations to participate in inter-organizational computing, which will be an important component of future wide area distributed applications. Surprisingly, we find that while some sites are disconnecting or otherwise distancing themselves from the Internet, the vast majority of sites have retained full or nearly full Internet connectivity. Moreover, we estimate that the number of sites accessible via the Internet has grown by approximately 31% in the past 6 months, significantly outpacing the rate at which sites are distancing themselves from the Internet. Our measurements are broken down by distancing mechanism and institution type/location.

1. Introduction

At the Fall 1990 Interop conference public session on security, David Clark of MIT mentioned the possibility of "The Great Disconnection". By this he meant the possibility that many sites would reduce their connectivity with the global TCP/IP Internet, in response to rising security concerns. Such concerns have increased markedly in the past few years, after a number of well publicized events, such as a series of espionage attempts directed at U.S. government research laboratories [Stoll 1988], the Internet Worm of November 1988 [Spafford 1989], and other risks of being interconnected with interorganizational networks [National Research Council 1991].

There are a number of different ways that sites might reduce their closeness of association with the Internet. The most extreme measure is simply to disconnect from the Internet. Because of the tremendous advantages of Internet access, however, many sites prefer less extreme measures. Common measures include disabling inter-site services (such as file transfers); and insulating a site's internal networks with gateways that allow no traffic directly into the internal network, but which selectively forward certain types of traffic (such as mail).

Clearly, the Internet is a tremendous resource, and "The Great Disconnection" will be a terrible loss if it ever happens. We are particularly concerned about this possibility, because for the past several years our research has become increasingly focused on the Internet as an experimental environment in which to deploy and measure wide area distributed services [Schwartz 1991]. The bigger issue beyond our particular research aspirations is that diminished Internet connectivity will hinder or prevent the deployment of interesting new types of network services.

One of the prototypes we have deployed is an Internet "white pages" tool called netfind [Schwartz & Tsirigotis 1991a]. In order to estimate the scope of the directory provided by netfind, we ran a measurement experiment in August 1990, which, among other things, sought to connect to finger [Harrenstien 1977] and Simple Mail Transfer Protocol (SMTP) [Postel 1982] servers at a large number of sites around the Internet. By repeating this experiment in February 1991, we have obtained direct measurements concerning changes in Internet upper layer service reachability. These measurements are the subject of the current paper. We consider upper layer service reachability in our measurements instead of just IP connectivity, because the former indicates the willingness of sites to participate in inter-organizational computing. Such willingness will be an important component in determining the success of future wide area distributed services.

Throughout this paper when we refer to "sites" we mean organizational groupings inferred by the Domain Naming System [Mockapetris 1987]. For example, the machine "abingdon.eng.sun.com" falls within the domain "eng.sun.com", which is a different domain (and hence site) than "central.sun.com", even though both domains belong to a single corporation (Sun MicroSystems, Inc.). When discussing domain names explicitly, we use the term "domain". We use the term "institution" to refer to a collection of sites related to a single organization (Sun in the above example). The purpose of distinguishing between sites and institutions is to permit a more fine-grained analysis of the patterns of Internet disconnection and growth. In particular, a number of institutions allow direct Internet access to some of their sites, while restricting such access to other sites (e.g., allowing Internet access for a research branch of a company, while restricting such access for a product development branch). Our measurements incorporate this level of detail.

The remainder of this paper is organized as follows. In Section 2 we overview the measurement methodology. In Section 3 we present and interpret our measurements. In Section 4 we offer our conclusions, and discuss areas for future work.

2. Methodology

The data used for this paper were not originally collected for the purposes of measuring general Internet upper layer service reachability. For example, these measurements do not consider reachability of services like telnet [Postel & Reynolds 1983] and FTP [Postel & Reynolds 1985]. So that the reader understands the reason behind collecting the data in the fashion that we did, we begin by briefly overviewing the function and structure of netfind. This overview also provides an example of a class of wide area distributed applications whose usefulness would be reduced if Internet site reachability were to diminish significantly.

2.1. Overview of Netfind

Given the name of a user and a rough description of where the user works (e.g., the company name or city), netfind attempts to locate telephone and electronic mailbox information about that user. Rather than requiring users to register with an administratively centralized service (as with the SRI Network Information Center WHOIS service [Harrenstien, Stahl & Feinler 1985]), or that special directory servers be run at many sites around the Internet (as with the CCITT X.500 directory standard [CCITT 1988]), netfind focuses on the ability to use a number of existing protocols and highly decentralized sources of relatively unstructured information. The mechanism used by netfind operates as follows.

We begin with a database of "seed" data, which provides hints of potential machines to probe when a search is requested. This database is built by gathering information from the headers of USENET [Quarterman & Hoskins 1986] news messages over time. These headers typically list the user name, organization name, city, and electronic mailbox for users who post messages. When a search is requested, the seed database is consulted to locate the names of a number of machines associated with institution keywords specified in the search request. Requests use the format "*UserString InstString [InstString ...]*", where *UserString* identifies the user (typically by last name), and the conjunction of one or more *InstString*s identify the institution where the user works. For example, a search could be requested for "schwartz university colorado" or "schwartz boulder".

If the machines found in the seed database fall within more than three naming domains (an example of one domain being "colorado.edu"), the user is asked to select at most three domains to search. The Domain Naming System is then contacted, to locate authoritative name server hosts for each of these domains. The idea is that these hosts are often central administrative machines, with accounts and/or mail forwarding information for many users at a site. Each of these machines is then queried using SMTP, in an attempt to find mail forwarding information about the specified user. If such information is found, the located machines are then probed using the "finger" protocol [Zimmerman 1990]. The results from finger searches can sometimes yield other machines to search as well. Ten lightweight threads are used to allow sets of DNS/SMTP/finger lookup sequences to proceed in parallel, to increase resilience to host and network failures.

This architecture has a number of implications, including the ability to function in the presence of partial remote protocol support, which provides good reliability and far-reaching scope. For further discussion and measurements of netfind, the reader is referred to [Schwartz & Tsirigotis 1991a]. For further discussion of the general techniques introduced by netfind, the reader is referred to [Schwartz & Tsirigotis 1991b].

2.2. Reachability Measurement Methodology

To estimate the scope of the directory provided by netfind, in August 1990 we conducted a measurement experiment in which we sought to execute the search protocol (Domain lookup, SMTP probe, and finger probe) on a small number of machines in each domain found in the seed database. A success was counted as contacting either an SMTP or a finger server in a domain (since either server could potentially yield white pages information). Once one success or more than ten failures occurred for a domain, no more probe attempts were made on machines in that domain. The experiment was run over a number of sessions at varying times of the day. For each session, the measurement program read the log output of the previous session, and tried contacting machines that either had not been tried before (because they were newly added to the growing seed database), or that had failed due to timeouts on previous tries, in domains for which fewer than 10 failures had occurred. Measurement sessions were run until the results of a session differed by little from the results of a previous session, indicating that most of the reachable domains had been counted.

The results in the current paper were derived by repeating this measurement experiment in February 1991, and then comparing the two data sets a number of different ways.

A possible source of concern in using this measurement methodology for the current paper is the fact that the set of sites to probe is not generated from a global list of all Internet sites, but was instead discovered by monitoring USENET transmissions. Because USENET "administration" in not coordinated with Internet administration, this database could be missing many Internet sites. However, this problem is inherent in generating *any* list of Internet sites, because of the decentralized nature of the Internet. Since host information exists within autonomous Domain servers, the only way to derive a list of sites is through some process of *discovery*, one example of which is the one used to generate the netfind seed database.

Intuitively, it is likely that the seed database collection mechanism will capture many of the Internet domains, because it is much easier and cheaper to join USENET than the Internet, and hence the vast majority of Internet sites are probably also "on" USENET. Indeed, the scope of the seed database contents compares favorably with that of the SRI Network Information Center's (SRI NIC) "HOSTS.TXT" file [Feinler et al. 1982], which contains a list of Internet hosts, gateways, and networks that have registered with the SRI NIC. At the time of this writing, the SRI

NIC list contacted 1,819 domains. In contrast, the seed database contains 6,317 domains.[1] By performing name lookups on a randomly selected subset of 2% of the hosts in the seed database, we found that 61% of the seed database entries are currently on the Internet. The total number of domains represented in the seed database as of this paper writing is thus approximately 3,860. Another difference between the SRI NIC list and the contents of our seed database is that the former is quite U.S.-centric. For example, the SRI NIC list contains only 14 German sites, compared with 313 such sites in the seed database (of which 134 are on the Internet).

Even though the seed database does not contain information about all Internet domains (for example, Mark Lottor's "Zone" Domain Naming System traversal tool was able to discover approximately 4,800 domains [Lottor 1990]), the fact that we used the same methodology for both runs means that any bias introduced in the first run would probably be qualitatively similar to bias introduced into the second run. Hence, while comparing the two runs may not yield quantitatively exact measurements, percentage changes are probably fairly accurate.

3. Measurement Results

We begin in Section 3.1 with measurements of sites that could be reached via SMTP or finger (which we refer to as "service reachable sites"). We then break these sites down further in Section 3.2, to characterize the types of institutions that have joined or left Internet upper layer service reachability. Finally, in Section 3.3 we characterize the types of institutions that isolate subdomains from the Internet using a gateway that selectively forwards only certain types of traffic, and the number of sites isolated by this mechanism.

3.1. Service-Reachable Sites

The sites that could be reached via SMTP or finger in the August 1990 and February 1991 measurement experiments are summarized in Table 1, sorted by number of sites of each type. This table also indicates the percentage of sites of each institution type listed in the seed database that could be reached (e.g., 870 (76%) of the 1,139 sites in the "edu" top-level domain could be reached in the August measurement run). Also, the final column indicates the percentage change of reachability from August to February. For example, 36% more "edu" domains could be reached in February than in August. Table 2 explains the meaning of each top-level domain in Table 1, sorted alphabetically by top-level domain name.

Sites that could not be reached could either indicate that a site was not on the Internet (e.g., as is the case with small companies that only connect indirectly, and have mail forwarded to them from an Internet site), or that a site had disconnected from the Internet. We analyze changes in service reachability in more detail in Section 3.2.

The largest percentage growth in Internet upper layer service reachability occurred in European sites, while the largest numerical increases occurred in U.S. educational and commercial institutions. The only reachability reduction occurred in domains named with the old-style ".arpa" name, which is being phased out of existence because of the advent of hierarchical naming in the Domain Naming System.

3.2. Changes in Service Reachability

To uncover more specific information about the changes that led to the measurements in Table 1, Table 3 indicates the distribution of sites that used to be reachable that are no longer reachable, as well as the distribution of sites that were not reachable previously, that are currently reachable. The final column of this table indicates the net change in reachability (newly connected - disconnected).

There are two primary reasons why a site would lose service reachability. The first reason is that the site stops running upper layer services. The second reason is that the site has been removed from direct Internet connectivity.

To discern which of these reasons governed sites that are no longer reachable, we performed domain name lookups on each of the 214 domains summarized in the left hand side of Table 3. Of these, we found that 50 sites no longer had Internet addresses associated with them, indicating that the sites were no longer on the Internet. We consider measurements of how sites disconnected from the Internet further in Section 3.3. The remaining 164 sites were on the Internet, but had shut off SMTP and finger access. In contrast, 809 new sites connected to the Internet.

[1] In July 1990 we augmented the seed database with entries from the "HOSTS.TXT" file, to capture host names for sites that do not typically post messages on USENET, such as sites on Milnet. Doing this increased the seed database size by approximately 40% at the time. Since that time the seed database has grown another 23% from continued USENET monitoring.

August 1990 Results			February 1991 Results			
Top-Level Domain Name	Reachable Sub-Domains	% Seed Database Entries	Top-Level Domain Name	Reachable Sub-Domains	% Seed Database Entries	%Change From Reachable Domains in August 1990
edu	870	76	edu	1,184	84	+36
arpa	310	25	com	322	23	+35
com	238	23	arpa	245	20	-20
gov	80	73	au	114	59	+60
ca	80	56	ca	113	62	+41
au	71	57	gov	93	70	+16
mil	68	56	mil	89	69	+30
se	34	40	de	49	39	+206
nl	23	30	se	46	44	+35
org	22	24	nl	37	39	+60
net	19	70	org	33	25	+50
fi	19	56	jp	28	23	+75
jp	16	19	fr	25	50	+92
de	16	18	fi	25	57	+31
no	15	79	net	23	77	+21
fr	13	32	no	20	77	+33
dk	11	37	dk	14	42	+27
nz	8	40	nz	11	39	+37
it	5	36	it	11	52	+120
us	4	07	ch	10	38	+400
uk	2	01	us	5	06	+25
mx	2	100	il	5	71	(new)
ch	2	12	at	5	36	(new)
pr	1	50	uk	4	02	+100
			is	4	50	(new)
			kr	3	100	(new)
			mx	2	100	+0
			pr	1	50	+0
			in	1	100	(new)
			gr	1	33	(new)
			es	1	7	(new)
Total	1,929		**Total**	2,524		

Table 1: Breakdown of Top-Level Reachable Domains

Subtracting disconnected from connected sites, this amounts to an net increase of 595 service-reachable sites on the Internet, which is a 31% increase over the 1,929 service reachable sites measured in August. Interestingly, even U.S. military sites are a growing presence on the directly connected Internet.

There is one other, less interesting cause of sites appearing to lose service reachability, that occurs when a site changes names.[2] This commonly occurs when a site adds a new level of depth to its naming hierarchy (e.g., the arizona.edu domain became cs.arizona.edu between August and February). One other noteworthy example of name changes seen in Table 3 is the move away from the old-style ".arpa" naming convention, which is being phased out of existence.

2 Over time, the netfind seed database can accumulate dated machine names.

Top-Level Domain Name	Description
arpa	Old-style ARPANET node names
at	Austrian Institutions
au	Australian Institutions
ca	Canadian Institutions
ch	Swiss Institutions
com	Commercial Institutions
de	German Institutions
dk	Danish Institutions
edu	U.S. Educational Institutions
es	Spanish Institutions
fi	Finnish Institutions
fr	French Institutions
gov	U.S. Government Institutions
gr	Greek Institutions
il	Israeli Institutions
in	Indian Institutions
is	Icelandic Institutions
it	Italian Institutions
jp	Japanese Institutions
kr	Korean Institutions
mil	U.S. Military Institutions
mx	Mexican Institutions
net	Institutions named by network connections
nl	Dutch Institutions
no	Norwegian Institutions
nz	New Zealand Institutions
org	Non-profit Institutions
pr	Puerto Rican Institutions
se	Swedish Institutions
uk	British Institutions
us	U.S. Institutions

Table 2: Top Level Domain Naming Legend

Finally, there is the issue of how service-reachable sites break down in these measurements. Since the measurement experiment is based on hosts discovered through USENET transmissions, there are actually two factors that could contribute to how the reachable site count increased. The first is that USENET sites listed in the seed database became connected to the Internet. The second is that new USENET sites were discovered, which allowed the measurement experiment to try more sites for possible Internet upper layer service reachability. To gauge the relative weight of each of these factors, we note that the August 1990 seed database contained 5,138 domains, of which 1,929 were service reachable. The February 1991 seed database contained 6,317 domains, of which 2,524 were service reachable. Therefore, the USENET domain count increased by 23% from August to February, while the reachable Internet domain count increased by 31%.

3.3. Internal Domains Isolated by Gateways

A common means of isolating the security risk of a site is to have only a small number of external gateways on the Internet, that selectively forward only certain types of traffic into the site's internal network. Usually, this means that the external gateway will accept SMTP connections for incoming mail, and then forward the mail into the internal network. This mechanism prevents any traffic other than mail from entering the internal network. Moreover,

Reachable August 1990, No Longer Reachable February 1991		Not Reachable August 1990, Newly Reachable February 1991		
Top-Level Domain Name	Count	Top-Level Domain Name	Count	Net Change
arpa	145	edu	344	+314
edu	30	com	98	+84
com	14	arpa	80	-65
mil	10	au	45	+43
nl	3	ca	34	+33
gov	3	de	33	+33
au	2	mil	31	+21
uk	1	nl	17	+14
se	1	gov	16	+13
org	1	se	13	+12
nz	1	org	12	+11
no	1	jp	12	+12
net	1	fr	12	+12
ca	1	ch	8	+8
		no	6	+5
		it	6	+6
		fi	6	+6
		net	5	+4
		il	5	+5
		at	5	+5
		nz	4	+3
		is	4	+4
		uk	3	+2
		kr	3	+3
		dk	3	+3
		us	1	+1
		in	1	+1
		gr	1	+1
		es	1	+1
Total	214	**Total**	809	+595

Table 3: Breakdown of Reachability Changes, by Top-Level Domain

SMTP (or any other traffic) cannot directly reach machines on the internal network. Instead, machines on internal networks are registered with the Domain Naming System using only "mail exchange" records that point to the external mail gateway. In effect, no interactive traffic (such as netfind's use of SMTP to locate mail forwarding information) is permitted from external to internal network nodes. A number of implementations of this mechanism are possible. See, for example, [Cheswick 1990].

To measure the extent to which this mechanism is being used in the Internet, we constructed a program that read through the August 1990 and February 1991 logs, and checked each unreachable domain to see if there was a higher-level domain that was reachable. For example, the domain "sun.com" is reachable, but the domain "eng.sun.com" is not reachable. Each such subdomain was then flagged as "isolated", in the sense that the domain existed but could not be reached, because a higher domain was not passing traffic through. We also flagged each such higher level domain as "isolating". Between these two flags, we could determine how many domains were acting as Internet isolation barriers (which indicates how many institutions are using this domain isolation mechanism); and how many domains were being isolated from the Internet using this mechanism. The results of these measurements are shown in Table 4. The percentage figures shown in parentheses on the "Total" line indicate the percentage of domains from the estimated set of 3,134 Internet domains in the August 1990 seed database, and 3,853 domains in the Internet February 1991 seed database. (See Section 2.2 for the origin of these full domain list counts.)

From this table, it is clear that a non-trivial number of domains are making indirect access the only way to communicate with their machines, but that relative to the number of sites in each of these domains, the practice is not "rampant" yet. In fact, there was a slight percentage decrease in both isolating and isolated domains from August to February, indicating that the Internet is growing faster than the isolation mechanism is being applied. The only domain type in which significant numerical levels of change occurred were a decrease in U.S. educational institutions' use of this mechanism, and an increase in U.S. commercial institutions' use of this mechanism. Note that in some cases domains might have previously been isolated because they had not yet finished preparations for connecting to the Internet. This is a likely explanation for U.S. educational institutions' decreased presence in Table 4 from August to February. In contrast, commercial institutions are more likely to isolate their internal networks because of security concerns, which would account for the increase in commercial use of this mechanism from August to February.

August 1990 Results				February 1991 Results			
Top-Level Domain Name	Isolating Sub-Domains	Top-Level Domain Name	Isolated Sub-Domains	Top-Level Domain Name	Isolating Sub-Domains	Top-Level Domain Name	Isolated Sub-Domains
edu	58	com	115	com	48	com	229
com	35	uk	110	edu	47	edu	73
mil	7	edu	103	ca	8	jp	69
ca	6	mil	42	mil	7	au	55
jp	4	au	37	jp	5	mil	34
gov	4	jp	29	gov	5	ca	15
au	3	ca	13	fi	5	gov	11
uk	2	gov	9	au	3	nz	8
nl	2	nl	2	us	2	ch	6
net	2	net	2	nz	2	uk	4
fi	2	fi	2	no	2	fi	4
us	1	us	1	nl	2	us	3
se	1	se	1	ch	2	no	2
org	1	org	1	at	2	nl	2
no	1	no	1	uk	1	dk	1
dk	1	dk	1	dk	1	de	1
				de	1	at	1
Total	130 (4.1%)	**Total**	469 (15.0%)	**Total**	143 (3.7%)	**Total**	518 (13.4%)

Table 4: Breakdown of Isolating and Isolated Domains

To provide a more fine grained measurement of the extent to which sites are making their machines only indirectly accessible via the internet, we would also need to measure how many machines could not be reached in a domain where some machine was reachable. For example, the machine jade.bellcore.com is not reachable by SMTP, but the machine thumper.bellcore.com is reachable by SMTP, and acts to isolate the internal network from the Internet. Nonetheless, the measurements we did provide a reasonable first approximation, at lower Internet load (since fewer machines need to be probed to determine the information we have presented).

4. Conclusions and Future Work

It is with some sense of relief that we are able to report that "The Great Disconnection" is not happening to a significant extent yet. For the time being relatively few sites have disconnected or otherwise distanced themselves from the Internet, and Internet upper layer service reachability continues to grow in leaps and bounds. Currently, the number of service reachable sites on the Internet is growing significantly faster than the rate at which sites are disconnecting or isolating themselves from the Internet.

It is our hope that sites around the Internet will seek less extreme ways to enhance security than disconnecting or isolating themselves from the Internet. The world has much to lose in terms of potential future wide area network services if sites distance themselves from the Internet.

There are at least three directions in which we plan to extend the work presented in this paper. First, we plan to continue running these measurements regularly for some time into the future, to follow the growth and changes in the Internet. Second, we may run measurement experiments to characterize service reachability at a finer grain, for example by breaking down the services measured into telnet/FTP/finger/mail. Third, we are currently formulating a mathematical model of network growth, based on observations of the statistical nature of the data in Table 1.

Acknowledgements

We would like to thank Phil Karn for mentioning Dave Clark's Interop talk, which introduced the term from which the title of this paper was taken. The author was not present at that conference. We also thank David Wood for providing comments on an earlier draft of this paper.

This material is based upon work supported in part by NSF cooperative agreement DCR-8420944, and by a grant from AT&T Bell Laboratories.

5. Bibliography

[CCITT 1988]
CCITT. The Directory, Part 1: Overview of Concepts, Models and Services. ISO DIS 9594-1, CCITT, Gloucester, England, Dec. 1988. Draft Recommendation X.500.

[Cheswick 1990]
B. Cheswick. The Design of A Secure Internet Gateway. *Proc. USENIX Summer Conf.*, June 1990.

[Feinler et al. 1982]
E. Feinler, K. Harrenstien, Z. Su and V. White. DoD Internet Host Table Specification. Req. For Com. 810, Network Information Center, SRI Int., Mar. 1982.

[Harrenstien 1977]
K. Harrenstien. Name/Finger. Req. For Com. 742, SRI Int., Dec. 1977.

[Harrenstien, Stahl & Feinler 1985]
K. Harrenstien, M. Stahl and E. Feinler. NICName/Whois. Req. For Com. 954, Oct. 1985.

[Lottor 1990]
M. Lottor. Personal Communication. Discussion of measured number of domains and machines in the Internet. Apr. 1990.

[Mockapetris 1987]
P. Mockapetris. Domain Names - Concepts and Facilities. Req. For Com. 1034, USC Information Sci. Institute, Nov. 1987.

[National Research Council 1991]
National Research Council. *Computers at Risk: Safe Computing in the Information Age.* National Academy Press, Washington, D.C., 1991.

[Postel 1982]
J. B. Postel. Simple Mail Transfer Protocol. Req. For Com. 821, USC Information Sci. Institute, Aug. 1982.

[Postel & Reynolds 1983]
J. Postel and J. Reynolds. Telnet Protocol Specification. Req. For Com. 854, USC Information Sci. Institute, May 1983.

[Postel & Reynolds 1985]
J. Postel and J. Reynolds. File Transfer Protocol (FTP). Req. For Com. 959, USC Information Sci. Institute, Oct. 1985.

[Quarterman & Hoskins 1986]
J. S. Quarterman and J. C. Hoskins. Notable Computer Networks. *Commun. ACM*, 23(10), pp. 932-971, Oct. 1986.

[Schwartz & Tsirigotis 1991a]
M. F. Schwartz and P. G. Tsirigotis. Experience with a Semantically Cognizant Internet White Pages Directory Tool. To appear, *J. Internetworking: Research and Experience*, 1991.

[Schwartz & Tsirigotis 1991b]
M. F. Schwartz and P. G. Tsirigotis. Techniques for Supporting Wide Area Distributed Applications. Tech. Rep. CU-CS-519-91, Dept. Comput. Sci., Univ. Colorado, Boulder, CO, Feb. 1991. Submitted for publication.

[Schwartz 1991]
M. F. Schwartz. Resource Discovery and Related Research at the University of Colorado. *To appear*, ConneXions - The Interoperability Report, *Interop, Inc., May 1991.*

[Spafford 1989]
E. H. Spafford. The Internet Worm: Crisis and Aftermath. *Commun. ACM*, 32(6), pp. 678-687, June 1989.

[Stoll 1988] C. Stoll. Stalking the Wiley Hacker. *Commun. ACM*, 31(5), pp. 484-497, May 1988.

[Zimmerman 1990]
D. Zimmerman. The Finger User Information Protocol. Req. For Com. 1194, Center for Discrete Mathematics and Theoretical Computer Science, Nov. 1990.