

Hitchhiker's Guide to the Internet

Ed Krol

The background of the lower half of the page is a vibrant cyan color. Overlaid on this is a complex, abstract pattern of thick, purple lines. These lines form various geometric shapes, including circles, semi-circles, rectangles, and triangles. Some lines are solid, while others are broken or dashed, creating a sense of movement and connectivity. A prominent feature is a large, downward-pointing triangle on the right side, and a large, semi-circular shape on the left side. The overall effect is a modern, digital aesthetic.

Project Gutenberg

The Project Gutenberg EBook of Hitchhiker's Guide to the Internet, by Ed Krol

This eBook is for the use of anyone anywhere at no cost and with almost no restrictions whatsoever. You may copy it, give it away or re-use it under the terms of the Project Gutenberg License included with this eBook or online at www.gutenberg.org

** This is a COPYRIGHTED Project Gutenberg eBook, Details Below ** ** Please follow the copyright guidelines in this file. **

Title: Hitchhiker's Guide to the Internet

Author: Ed Krol

Posting Date: December 17, 2011 [EBook #39] Release Date: September, 1992

Language: English

*** START OF THIS PROJECT GUTENBERG EBOOK
HITCHHIKER'S GUIDE TO THE INTERNET ***

The Hitchhikers Guide to the Internet
25 August 1987

Ed Krol
krol@uxc.cso.uiuc.edu

This document was produced through funding of the National Science Foundation.

Copyright (C) 1987, by the Board of Trustees of The University of Illinois. Permission to duplicate this document, in whole or part, is granted provided reference is made to the source and this copyright is included in whole copies.

This document assumes that one is familiar with the workings of a non-connected simple IP network (e.g. a few 4.2 BSD systems on an Ethernet not connected to anywhere else). Appendix A contains remedial information to get one to this point. Its purpose is to get that person, familiar with a simple net, versed in the "oral tradition" of the Internet to the point that that net can be connected to the Internet with little danger to either. It is not a tutorial, it consists of pointers to other places, literature, and hints which are not normally documented. Since the Internet is a dynamic environment, changes to this document will be

made regularly. The author welcomes comments and suggestions. This is especially true of terms for the glossary (definitions are not necessary).

In the beginning there was the ARPAnet, a wide area experimental network connecting hosts and terminal servers together. Procedures were set up to regulate the allocation of addresses and to create voluntary standards for the network. As local area networks became more pervasive, many hosts became gateways to local networks. A network layer to allow the interoperation of these networks was developed and called IP (Internet Protocol). Over time other groups created long haul IP based networks (NASA, NSF, states...). These nets, too, interoperate because of IP. The collection of all of these interoperating networks is the Internet.

Two groups do much of the research and information work of the Internet (ISI and SRI). ISI (the Informational Sciences Institute) does much of the research, standardization, and allocation work of the Internet. SRI International provides information services for the Internet. In fact, after you are connected to the Internet most of the information in this document can be retrieved from the Network Information Center (NIC) run by SRI.

Operating the Internet

Each network, be it the ARPAnet, NSFnet or a regional network, has its own operations center. The ARPAnet is run by BBN, Inc. under contract from DARPA. Their facility is called the Network Operations Center or NOC. Cornell University temporarily operates NSFnet (called the Network Information Service Center, NISC). It goes on to the

-2-

regionals having similar facilities to monitor and keep watch over the goings on of their portion of the Internet. In addition, they all should have some knowledge of what is happening to the Internet in total. If a problem comes up, it is suggested that a campus network liaison should contact the network operator to which he is directly connected. That is, if you are connected to a regional network (which is gatewayed to the NSFnet, which is connected to the ARPAnet...) and have a problem, you should contact your regional network operations center.

RFCs

The internal workings of the Internet are defined by a set of documents called RFCs (Request for Comments). The general process for creating an RFC is for someone wanting something formalized to write a document describing the issue and mailing it to Jon Postel (postel@isi.edu). He acts as a referee for the proposal. It is then commented upon by all those wishing to take part in the discussion (electronically of course). It may go through multiple revisions. Should it be generally accepted as a

good idea, it will be assigned a number and filed with the RFCs.

The RFCs can be divided into five groups: required, suggested, directional, informational and obsolete. Required RFC's (e.g. RFC-791, The Internet Protocol) must be implemented on any host connected to the Internet. Suggested RFCs are generally implemented by network hosts. Lack of them does not preclude access to the Internet, but may impact its usability. RFC-793 (Transmission Control Protocol) is a suggested RFC. Directional RFCs were discussed and agreed to, but their application has never come into wide use. This may be due to the lack of wide need for the specific application (RFC-937 The Post Office Protocol) or that, although technically superior, ran against other pervasive approaches (RFC-891 Hello). It is suggested that should the facility be required by a particular site, an implementation be done in accordance with the RFC. This insures that, should the idea be one whose time has come, the implementation will be in accordance with some standard and will be generally usable. Informational RFCs contain factual information about the Internet and its operation (RFC-990, Assigned Numbers). Finally, as the Internet and technology have grown, some RFCs have become unnecessary. These obsolete RFCs cannot be ignored, however. Frequently when a change is made to some RFC that causes a new one to be issued obsoleting others, the new RFC only contains explanations and motivations for the change. Understanding the model on which the whole facility is

based may involve reading the original and subsequent RFCs on the topic.

-3-

(Appendix B contains a list of what are considered to be the major RFCs necessary for understanding the Internet).

The Network Information Center

The NIC is a facility available to all Internet users which provides information to the community. There are three means of NIC contact: network, telephone, and mail. The network accesses are the most prevalent. Interactive access is frequently used to do queries of NIC service overviews, look up user and host names, and scan lists of NIC documents. It is available by using

```
%telnet sri-nic.arpa
```

on a BSD system and following the directions provided by a user friendly prompter. From poking around in the databases provided one might decide that a document named NETINFO:NUG.DOC (The Users Guide to the ARPAnet) would be worth having. It could be retrieved via an anonymous FTP. An anonymous FTP would proceed something like the following. (The dialogue may vary

slightly depending on the implementation of FTP you are using).

```
%ftp sri-nic.arpa
Connected to sri-nic.arpa.
220 SRI_NIC.ARPA FTP Server Process 5Z(47)-6 at Wed
17-Jun-87 12:00 PDT
Name (sri-nic.arpa:myname): anonymous
331 ANONYMOUS user ok, send real ident as password.
Password: myname
230 User ANONYMOUS logged in at Wed 17-Jun-87 12:01
PDT,
job 15.
ftp> get netinfo:nug.doc
200 Port 18.144 at host 128.174.5.50 accepted.
150 ASCII retrieve of <NETINFO>NUG.DOC.11 started.
226 Transfer Completed 157675 (8) bytes transferred
local: netinfo:nug.doc remote:netinfo:nug.doc
157675 bytes in 4.5e+02 seconds (0.34 Kbytes/s)
ftp> quit
221 QUIT command received. Goodbye.
```

(Another good initial document to fetch is
NETINFO:WHAT-THE-NIC-DOES.TXT)!

Questions of the NIC or problems with services can be
asked
of or reported to using electronic mail. The following
addresses can be used:

NIC@SRI-NIC.ARPA General user assistance, document
requests

REGISTRAR@SRI-NIC.ARPA User registration and WHOIS updates

HOSTMASTER@SRI-NIC.ARPA Hostname and domain changes and updates

ACTION@SRI-NIC.ARPA SRI-NIC computer operations

SUGGESTIONS@SRI-NIC.ARPA Comments on NIC publications and services

-4-

For people without network access, or if the number of documents is large, many of the NIC documents are available in printed form for a small charge. One frequently ordered document for starting sites is a compendium of major RFCs. Telephone access is used primarily for questions or problems with network access. (See appendix B for mail/telephone contact numbers).

The NSFnet Network Service Center

The NSFnet Network Service Center (NNSC) is funded by NSF to provide a first level of aid to users of NSFnet should they have questions or encounter problems traversing the network. It is run by BBN Inc. Karen Roubicek (roubicek@nnsf.net) is the NNSC user liaison.

The NNSC, which currently has information and documents online and in printed form, plans to distribute news through network mailing lists, bulletins, newsletters, and online reports. The NNSC also maintains a database of contact points and sources of additional information about NSFnet component networks and supercomputer centers.

Prospective or current users who do not know whom to call concerning questions about NSFnet use, should contact the NNSC. The NNSC will answer general questions, and, for detailed information relating to specific components of the Internet, will help users find the appropriate contact for further assistance. (Appendix B)

Mail Reflectors

The way most people keep up to date on network news is through subscription to a number of mail reflectors. Mail reflectors are special electronic mailboxes which, when they receive a message, resend it to a list of other mailboxes. This in effect creates a discussion group on a particular topic. Each subscriber sees all the mail forwarded by the reflector, and if one wants to put his "two cents" in sends a message with the comments to the reflector....

The general format to subscribe to a mail list is to find the address reflector and append the string -REQUEST to the mailbox name (not the host name). For example, if you wanted to take part in the mailing list for NSFnet reflected by NSFNET@NNSC.NSF.NET, one sends a request to

-5-

NSFNET-REQUEST@NNSC.NSF.NET. This may be a wonderful scheme, but the problem is that you must know the list exists in the first place. It is suggested that, if you are interested, you read the mail from one list (like NSFNET) and you will probably become familiar with the existence of others. A registration service for mail reflectors is provided by the NIC in the files NETINFO:INTEREST-GROUPS-1.TXT, NETINFO:INTEREST-GROUPS-2.TXT, and NETINFO:INTEREST-GROUPS-3.TXT.

The NSFNET mail reflector is targeted at those people who have a day to day interest in the news of the NSFnet (the backbone, regional network, and Internet inter-connection site workers). The messages are reflected by a central location and are sent as separate messages to each subscriber. This creates hundreds of messages on the wide area networks where bandwidth is the scarcest.

There are two ways in which a campus could spread the news and not cause these messages to inundate the wide area networks. One is to re-reflect the message on the campus. That is, set up a reflector on a local machine which forwards the message to a campus distribution list. The other is to create an alias on a campus machine which

places the messages into a notesfile on the topic. Campus users who want the information could access the notesfile and see the messages that have been sent since their last access. One might also elect to have the campus wide area network liaison screen the messages in either case and only forward those which are considered of merit. Either of these schemes allows one message to be sent to the campus, while allowing wide distribution within.

Address Allocation

Before a local network can be connected to the Internet it must be allocated a unique IP address. These addresses are allocated by ISI. The allocation process consists of getting an application form received from ISI. (Send a message to `hostmaster@sri-nic.arpa` and ask for the template for a connected address). This template is filled out and mailed back to hostmaster. An address is allocated and e-mailed back to you. This can also be done by postal mail (Appendix B).

IP addresses are 32 bits long. It is usually written as four decimal numbers separated by periods (e.g., 192.17.5.100). Each number is the value of an octet of the 32 bits. It was seen from the beginning that some networks might choose to organize themselves as very flat (one net with a lot of nodes) and some might organize hierarchically

-6-

(many interconnected nets with fewer nodes each and a backbone). To provide for these cases, addresses were differentiated into class A, B, and C networks. This classification had to do with the interpretation of the octets. Class A networks have the first octet as a network address and the remaining three as a host address on that network. Class C addresses have three octets of network address and one of host. Class B is split two and two. Therefore, there is an address space for a few large nets, a reasonable number of medium nets and a large number of small nets. The top two bits in the first octet are coded to tell the address format. All of the class A nets have been allocated. So one has to choose between Class B and Class C when placing an order. (There are also class D (Multicast) and E (Experimental) formats. Multicast addresses will likely come into greater use in the near future, but are not frequently used now).

In the past sites requiring multiple network addresses requested multiple discrete addresses (usually Class C). This was done because much of the software available (notably 4.2BSD) could not deal with subnetted addresses. Information on how to reach a particular network (routing

information) must be stored in Internet gateways and packet switches. Some of these nodes have a limited capability to store and exchange routing information (limited to about 300 networks). Therefore, it is suggested that any campus announce (make known to the Internet) no more than two discrete network numbers.

If a campus expects to be constrained by this, it should consider subnetting. Subnetting (RFC-932) allows one to announce one address to the Internet and use a set of addresses on the campus. Basically, one defines a mask which allows the network to differentiate between the network portion and host portion of the address. By using a different mask on the Internet and the campus, the address can be interpreted in multiple ways. For example, if a campus requires two networks internally and has the 32,000 addresses beginning 128.174.X.X (a Class B address) allocated to it, the campus could allocate 128.174.5.X to one part of campus and 128.174.10.X to another. By advertising 128.174 to the Internet with a subnet mask of FF.FF.00.00, the Internet would treat these two addresses as one. Within the campus a mask of FF.FF.FF.00 would be used, allowing the campus to treat the addresses as separate entities. (In reality you don't pass the subnet mask of FF.FF.00.00 to the Internet, the octet meaning is implicit in its being a class B address). A word of warning is necessary. Not all systems know how to do subnetting. Some 4.2BSD systems require additional software. 4.3BSD systems subnet as released. Other devices

-7-

and operating systems vary in the problems they have dealing with subnets. Frequently these machines can be used as a leaf on a network but not as a gateway within the subnetted portion of the network. As time passes and more systems become 4.3BSD based, these problems should disappear.

There has been some confusion in the past over the format of an IP broadcast address. Some machines used an address of all zeros to mean broadcast and some all ones. This was confusing when machines of both type were connected to the same network. The broadcast address of all ones has been adopted to end the grief. Some systems (e.g. 4.2 BSD) allow one to choose the format of the broadcast address. If a system does allow this choice, care should be taken that the all ones format is chosen. (This is explained in RFC-1009 and RFC-1010).

Internet Problems

There are a number of problems with the Internet. Solutions to the problems range from software changes to

long term research projects. Some of the major ones are detailed below:

Number of Networks

When the Internet was designed it was to have about 50 connected networks. With the explosion of networking, the number is now approaching 300. The software in a group of critical gateways (called the core gateways of the ARPAnet) are not able to pass or store much more than that number. In the short term, core reallocation and recoding has raised the number slightly. By the summer of '88 the current PDP-11 core gateways will be replaced with BBN Butterfly gateways which will solve the problem.

Routing Issues

Along with sheer mass of the data necessary to route packets to a large number of networks, there are many problems with the updating, stability, and optimality of the routing algorithms. Much research is being done in the area, but the optimal solution to these routing problems is still years away. In most cases the the routing we have today works, but sub-optimally and sometimes unpredictably.

Trust Issues

Gateways exchange network routing information. Currently, most gateways accept on faith that the information provided about the state of the network is correct. In the past this was not a big problem since most of the gateways belonged to a single administrative entity (DARPA). Now with multiple wide area networks under different administrations, a rogue gateway somewhere in the net could cripple the Internet. There is design work going on to solve both the problem of a gateway doing unreasonable things and providing enough information to reasonably route data between multiply connected networks (multi-homed networks).

Capacity & Congestion

Many portions of the ARPAnet are very congested during the busy part of the day. Additional links are planned to alleviate this congestion, but the implementation will take a few months.

These problems and the future direction of the Internet are determined by the Internet Architect (Dave Clark of MIT) being advised by the Internet Activities Board (IAB). This board is composed of chairmen of a number of committees with responsibility for various specialized areas of the Internet. The committees composing the IAB and their chairmen are:

Committee Chair

Autonomous Networks Deborah Estrin
End-to-End Services Bob Braden
Internet Architecture Dave Mills
Internet Engineering Phil Gross
EGP2 Mike Petry
Name Domain Planning Doug Kingston
Gateway Monitoring Craig Partridge
Internic Jake Feinler
Performance & Congestion Control Robert Stine
NSF Routing Chuck Hedrick
Misc. MilSup Issues Mike St. Johns
Privacy Steve Kent
IRINET Requirements Vint Cerf
Robustness & Survivability Jim Mathis
Scientific Requirements Barry Leiner

Note that under Internet Engineering, there are a set of task forces and chairs to look at short term concerns. The chairs of these task forces are not part of the IAB.

-9-

Routing

Routing is the algorithm by which a network directs a packet from its source to its destination. To appreciate the problem, watch a small child trying to find a table in a restaurant. From the adult point of view the structure of the dining room is seen and an optimal route easily chosen. The child, however, is presented with a set of paths between tables where a good path, let alone the optimal one to the goal is not discernible.***

A little more background might be appropriate. IP gateways (more correctly routers) are boxes which have connections to multiple networks and pass traffic between these nets. They decide how the packet is to be sent based on the information in the IP header of the packet and the state of the network. Each interface on a router has a unique address appropriate to the network to which it is connected. The information in the IP header which is used is primarily the destination address. Other information (e.g. type of service) is largely ignored at this time. The state of the network is determined by the routers passing information among themselves. The distribution of the database (what each node knows), the form of the updates, and metrics used to measure the value of a connection, are the parameters which determine the characteristics of a routing protocol.

Under some algorithms each node in the network has complete knowledge of the state of the network (the adult algorithm). This implies the nodes must have larger amounts of local storage and enough CPU to search the large tables in a short enough time (remember this must

be done for each packet). Also, routing updates usually contain only changes to the existing information (or you spend a large amount of the network capacity passing around megabyte routing updates). This type of algorithm has several problems. Since the only way the routing information can be passed around is across the network and the propagation time is non-trivial, the view of the network at each node is a correct historical view of the network at varying times in the past. (The adult algorithm, but rather than looking directly at the dining area, looking at a photograph of the dining room. One is likely to pick the optimal route and find a bus-cart has moved in to block the path after the photo was taken). These inconsistencies can cause circular routes (called routing loops) where once a packet enters it is routed in a closed path until its time to live (TTL) field expires and it is discarded.

Other algorithms may know about only a subset of the network. To prevent loops in these protocols, they are usually used in a hierarchical network. They know completely about their own area, but to leave that area they go to one particular place (the default gateway). Typically these are used in smaller networks (campus, regional...).

-10-

Routing protocols in current use:

Static (no protocol-table/default routing)

Don't laugh. It is probably the most reliable, easiest to implement, and least likely to get one into trouble for a small network or a leaf on the Internet. This is, also, the only method available on some CPU-operating system combinations. If a host is connected to an Ethernet which has only one gateway off of it, one should make that the default gateway for the host and do no other routing. (Of course that gateway may pass the reachability information somehow on the other side of itself).

One word of warning, it is only with extreme caution that one should use static routes in the middle of a network which is also using dynamic routing. The routers passing dynamic information are sometimes confused by conflicting dynamic and static routes. If your host is on an ethernet with multiple routers to other networks on it and the routers are doing dynamic routing among

themselves, it is usually better to take part in the dynamic routing than to use static routes.

RIP

RIP is a routing protocol based on XNS (Xerox Network System) adapted for IP networks. It is used by many routers (Proteon, cisco, UB...) and many BSD Unix systems BSD systems typically run a program called "routed" to exchange information with other systems running RIP. RIP works best for nets of small diameter where the links are of equal speed. The reason for this is that the metric used to determine which path is best is the hop-count. A hop is a traversal across a gateway. So, all machines on the same Ethernet are zero hops away. If a router connects connects two net- works directly, a machine on the other side of the router is one hop away.... As the routing information is passed through a gateway, the gateway adds one to the hop counts to keep them consistent across the net- work. The diameter of a network is defined as the largest hop-count possible within a network. Unfor- tunately, a hop count of 16 is defined as infinity in RIP meaning the link is down. Therefore, RIP will not allow hosts separated by more than 15 gateways in the RIP space to communicate.

The other problem with hop-count metrics is that if links have different speeds, that difference is not

-11-

reflected in the hop-count. So a one hop satellite link (with a .5 sec delay) at 56kb would be used instead of a two hop T1 connection. Congestion can be viewed as a decrease in the efficacy of a link. So, as a link gets more congested, RIP will still know it is the best hop-count route and congest it even more by throwing more packets on the queue for that link.

The protocol is not well documented. A group of people are working on producing an RFC to both define the current RIP and to do some extensions to it to allow it to better cope with larger networks. Currently, the best documentation for RIP appears to be the code to BSD "routed".

Routed

The ROUTED program, which does RIP for 4.2BSD systems, has many options. One of the most frequently used is: "routed -q" (quiet mode) which means listen to RIP information but never broadcast it. This would be used by a machine on

a network with multiple RIP speaking gateways. It allows the host to determine which gateway is best (hopwise) to use to reach a distant network. (Of course you might want to have a default gateway to prevent having to pass all the addresses known to the Internet around with RIP).

There are two ways to insert static routes into "routed", the "/etc/gateways" file and the "route add" command. Static routes are useful if you know how to reach a distant network, but you are not receiving that route using RIP. For the most part the "route add" command is preferable to use. The reason for this is that the command adds the route to that machine's routing table but does not export it through RIP. The "/etc/gateways" file takes precedence over any routing information received through a RIP update. It is also broadcast as fact in RIP updates produced by the host without question, so if a mistake is made in the "/etc/gateways" file, that mistake will soon permeate the RIP space and may bring the network to its knees.

One of the problems with "routed" is that you have very little control over what gets broadcast and what doesn't. Many times in larger networks where various parts of the network are under different administrative controls, you would like to pass on through RIP only nets which you receive from RIP and you know are reasonable. This prevents people from adding IP addresses to the network which may be illegal and you being

responsible for passing them on to the Internet.
This

-12-

type of reasonability checks are not available with "routed" and leave it usable, but inadequate for large networks.

Hello (RFC-891)

Hello is a routing protocol which was designed and implemented in a experimental software router called a "Fuzzball" which runs on a PDP-11. It does not have wide usage, but is the routing protocol currently used on the NSFnet backbone. The data transferred between nodes is similar to RIP (a list of networks and their metrics). The metric, however, is milliseconds of delay. This allows Hello to be used over nets of various link speeds and performs better in congestive situations.

One of the most interesting side effects of Hello based networks is their great timekeeping ability. If you consider the problem of measuring delay on a link for the metric, you find that it is not an easy thing to do. You cannot measure round trip time

since the return link may be more congested, of a different speed, or even not there. It is not really feasible for each node on the network to have a builtin WWV (nationwide radio time standard) receiver. So, you must design an algorithm to pass around time between nodes over the network links where the delay in transmission can only be approximated. Hello routers do this and in a nationwide network maintain synchronized time within milliseconds.

Exterior Gateway Protocol (EGP RFC-904)

EGP is not strictly a routing protocol, it is a reachability protocol. It tells only if nets can be reached through a particular gateway, not how good the connection is. It is the standard by which gateways to local nets inform the ARPAnet of the nets they can reach. There is a metric passed around by EGP but its usage is not standardized formally. Its typical value is value is 1 to 8 which are arbitrary goodness of link values understood by the internal DDN gateways. The smaller the value the better and a value of 8 being unreachable. A quirk of the protocol prevents distinguishing between 1 and 2, 3 and 4..., so the usability of this as a metric is as three values and unreachable. Within NSFnet the values used are 1, 3, and unreachable. Many routers talk EGP so they can be used for ARPAnet gateways.

-13-

Gated

So we have regional and campus networks talking RIP among themselves, the NSFnet backbone talking Hello, and the DDN speaking EGP. How do they interoperate? In the beginning there was static routing, assembled into the Fuzzball software configured for each site. The problem with doing static routing in the middle of the network is that it is broadcast to the Internet whether it is usable or not. Therefore, if a net becomes unreachable and you try to get there, dynamic routing will immediately issue a net unreachable to you. Under static routing the routers would think the net could be reached and would continue trying until the application gave up (in 2 or more minutes). Mark Fedor of Cornell (fedor@devvax.tn.cornell.edu) attempted to solve these problems with a replacement for "routed" called "gated".

"Gated" talks RIP to RIP speaking hosts, EGP to EGP speakers, and Hello to Hello'ers. These speakers frequently all live on one Ethernet, but

luckily (or unluckily) cannot understand each others ruminations. In addition, under configuration file control it can filter the conversion. For example, one can produce a configuration saying announce RIP nets via Hello only if they are specified in a list and are reachable by way of a RIP broadcast as well. This means that if a rogue network appears in your local site's RIP space, it won't be passed through to the Hello side of the world. There are also configuration options to do static routing and name trusted gateways.

This may sound like the greatest thing since sliced bread, but there is a catch called metric conversion. You have RIP measuring in hops, Hello measuring in milliseconds, and EGP using arbitrary small numbers. The big questions is how many hops to a millisecond, how many milliseconds in the EGP number 3.... Also, remember that infinity (unreachability) is 16 to RIP, 30000 or so to Hello, and 8 to the DDN with EGP. Getting all these metrics to work well together is no small feat. If done incorrectly and you translate an RIP of 16 into an EGP of 6, everyone in the ARPAnet will still think your gateway can reach the unreachable and will send every packet in the world your way. For these reasons, Mark requests that you consult closely with him when configuring and using "gated".

-14-

"Names"

All routing across the network is done by means of the IP address associated with a packet. Since humans find it difficult to remember addresses like 128.174.5.50, a symbolic name register was set up at the NIC where people would say "I would like my host to be named 'uiucuxc'". Machines connected to the Internet across the nation would connect to the NIC in the middle of the night, check modification dates on the hosts file, and if modified move it to their local machine. With the advent of workstations and micros, changes to the host file would have to be made nightly. It would also be very labor intensive and consume a lot of network bandwidth. RFC-882 and a number of others describe domain name service, a distributed data base system for mapping names into addresses.

We must look a little more closely into what's in a name. First, note that an address specifies a particular connection on a specific network. If the machine moves, the address changes. Second, a machine can have one or more names and one or more network addresses (connections) to different networks. Names point to a something which does useful work (i.e. the machine) and

IP addresses point to an interface on that provider. A name is a purely symbolic representation of a list of addresses on the network. If a machine moves to a different network, the addresses will change but the name could remain the same.

Domain names are tree structured names with the root of the tree at the right. For example:

uxc.cso.uiuc.edu

is a machine called 'uxc' (purely arbitrary), within the subdomains method of allocation of the U of I) and 'uiuc' (the University of Illinois at Urbana), registered with 'edu' (the set of educational institutions).

A simplified model of how a name is resolved is that on the user's machine there is a resolver. The resolver knows how to contact across the network a root name server. Root servers are the base of the tree structured data retrieval system. They know who is responsible for handling first level domains (e.g. 'edu'). What root servers to use is an installation parameter. From the root server the resolver finds out who provides 'edu' service. It contacts the 'edu' name server which supplies it with a list of addresses of servers for the subdomains (like 'uiuc'). This action is repeated with the subdomain servers until the final sub-domain returns a list of addresses of interfaces on the host in question. The user's machine then has its choice of which of these addresses to use for communication.

-15-

A group may apply for its own domain name (like 'uiuc' above). This is done in a manner similar to the IP address allocation. The only requirements are that the requestor have two machines reachable from the Internet, which will act as name servers for that domain. Those servers could also act as servers for subdomains or other servers could be designated as such. Note that the servers need not be located in any particular place, as long as they are reachable for name resolution. (U of I could ask Michigan State to act on its behalf and that would be fine). The biggest problem is that someone must do maintenance on the database. If the machine is not convenient, that might not be done in a timely fashion. The other thing to note is that once the domain is allocated to an administrative entity, that entity can freely allocate subdomains using whatever manner it sees fit.

The Berkeley Internet Name Domain (BIND) Server implements the Internet name server for UNIX systems. The name server is a distributed data base system that allows clients to name resources and to share that information with other network hosts. BIND is integrated with 4.3BSD and is used to lookup and store host names,

addresses, mail agents, host information, and more. It replaces the "/etc/hosts" file for host name lookup. BIND is still an evolving program. To keep up with reports on operational problems, future design decisions, etc, join the BIND mailing list by sending a request to "bind-request@ucbarp.Berkeley.EDU". BIND can also be obtained via anonymous FTP from ucbarpa.berkeley.edu.

There are several advantages in using BIND. One of the most important is that it frees a host from relying on "/etc/hosts" being up to date and complete. Within the .uiuc.edu domain, only a few hosts are included in the host table distributed by SRI. The remainder are listed locally within the BIND tables on uxc.cso.uiuc.edu (the server machine for most of the .uiuc.edu domain). All are equally reachable from any other Internet host running BIND.

BIND can also provide mail forwarding information for interior hosts not directly reachable from the Internet. These hosts can either be on non-advertised networks, or not connected to a network at all, as in the case of UUCP-reachable hosts. More information on BIND is available in the "Name Server Operations Guide for BIND" in "UNIX System Manager's Manual", 4.3BSD release.

There are a few special domains on the network, like SRI-NIC.ARPA. The 'arpa' domain is historical, referring to hosts registered in the old hosts database at the NIC. There are others of the form NNSC.NSF.NET. These special

domains are used sparingly and require ample justification. They refer to servers under the administrative control of

-16-

the network rather than any single organization. This allows for the actual server to be moved around the net while the user interface to that machine remains constant. That is, should BBN relinquish control of the NNSC, the new provider would be pointed to by that name.

In actuality, the domain system is a much more general and complex system than has been described. Resolvers and some servers cache information to allow steps in the resolution to be skipped. Information provided by the servers can be arbitrary, not merely IP addresses. This allows the system to be used both by non-IP networks and for mail, where it may be necessary to give information on intermediate mail bridges.

What's wrong with Berkeley Unix

University of California at Berkeley has been funded by DARPA to modify the Unix system in a number of ways. Included in these modifications is support for the Internet protocols. In earlier versions (e.g. BSD 4.2) there was good support for the basic Internet protocols (TCP, IP, SMTP, ARP) which allowed it to perform nicely on IP ether- nets

and smaller Internets. There were deficiencies, however, when it was connected to complicated networks. Most of these problems have been resolved under the newest release (BSD 4.3). Since it is the springboard from which many vendors have launched Unix implementations (either by porting the existing code or by using it as a model), many implementations (e.g. Ultrix) are still based on BSD 4.2. Therefore, many implementations still exist with the BSD 4.2 problems. As time goes on, when BSD 4.3 trickles through vendors as new release, many of the problems will be resolved. Following is a list of some problem scenarios and their handling under each of these releases.

ICMP redirects

Under the Internet model, all a system needs to know to get anywhere in the Internet is its own address, the address of where it wants to go, and how to reach a gateway which knows about the Internet. It doesn't have to be the best gateway. If the system is on a network with multiple gateways, and a host sends a packet for delivery to a gateway which feels another directly connected gateway is more appropriate, the gateway sends the sender a message. This message is an ICMP redirect, which politely says "I'll deliver this message for you, but you really ought to use that gateway over there to reach this host". BSD 4.2 ignores these messages. This creates more stress on the gateways and the local network, since for every packet

-17-

sent, the gateway sends a packet to the originator. BSD 4.3 uses the redirect to update its routing tables, will use the route until it times out, then revert to the use of the route it thinks is should use. The whole process then repeats, but it is far better than one per packet.

Trailers

An application (like FTP) sends a string of octets to TCP which breaks it into chunks, and adds a TCP header. TCP then sends blocks of data to IP which adds its own headers and ships the packets over the network. All this prepending of the data with headers causes memory moves in both the sending and the receiving machines. Someone got the bright idea that if packets were long and they stuck the headers on the end (they became trailers), the receiving machine could put the packet on the beginning of a page boundary and if the trailer was OK merely delete it and transfer control of the page with no memory moves involved. The problem is that trailers were never standardized and most gateways don't know to

look for the routing information at the end of the block. When trailers are used, the machine typically works fine on the local network (no gateways involved) and for short blocks through gateways (on which trailers aren't used). So TELNET and FTP's of very short files work just fine and FTP's of long files seem to hang. On BSD 4.2 trailers are a boot option and one should make sure they are off when using the Internet. BSD 4.3 negotiates trailers, so it uses them on its local net and doesn't use them when going across the network.

Retransmissions

TCP fires off blocks to its partner at the far end of the connection. If it doesn't receive an acknowledge- ment in a reasonable amount of time it retransmits the blocks. The determination of what is reasonable is done by TCP's retransmission algorithm. There is no correct algorithm but some are better than others, where better is measured by the number of retransmissions done unnecessarily. BSD 4.2 had a retransmission algorithm which retransmitted quickly and often. This is exactly what you would want if you had a bunch of machines on an ethernet (a low delay network of large bandwidth). If you have a network of relatively longer delay and scarce bandwidth (e.g. 56kb lines), it tends to retransmit too aggressively. Therefore, it makes

the networks and gateways pass more traffic than is really necessary for a given conversation. Retransmission algorithms do adapt to the delay of the network

-18-

after a few packets, but 4.2's adapts slowly in delay situations. BSD 4.3 does a lot better and tries to do the best for both worlds. It fires off a few retransmissions really quickly assuming it is on a low delay network, and then backs off very quickly. It also allows the delay to be about 4 minutes before it gives up and declares the connection broken.

-19-

Appendix A

References to Remedial Information

Quaterman and Hoskins, "Notable Computer Networks", Communications of the ACM, Vol 29, #10, pp. 932-971 (October, 1986).

Tannenbaum, Andrew S., Computer Networks, Prentice Hall, 1981.

Hedrick, Chuck, Introduction to the Internet Protocols, Anonymous FTP from topaz.rutgers.edu, directory pub/tcp-ip-docs, file tcp-ip-intro.doc.

Appendix B
List of Major RFCs

- RFC-768 User Datagram Protocol (UDP)
- RFC-791 Internet Protocol (IP)
- RFC-792 Internet Control Message Protocol (ICMP)
- RFC-793 Transmission Control Protocol (TCP)
- RFC-821 Simple Mail Transfer Protocol (SMTP)
- RFC-822 Standard for the Format of ARPA Internet Text Messages
- RFC-854 Telnet Protocol
- RFC-917 * Internet Subnets
- RFC-919 * Broadcasting Internet Datagrams
- RFC-922 * Broadcasting Internet Datagrams in the Presence of Subnets
- RFC-940 * Toward an Internet Standard Scheme for Subnetting
- RFC-947 * Multi-network Broadcasting within the Internet
- RFC-950 * Internet Standard Subnetting Procedure
- RFC-959 File Transfer Protocol (FTP)
- RFC-966 * Host Groups: A Multicast Extension to the Internet Protocol
- RFC-988 * Host Extensions for IP Multicasting

RFC-997 * Internet Numbers
RFC-1010 * Assigned Numbers
RFC-1011 * Official ARPA-Internet Protocols

RFC's marked with the asterisk (*) are not included in the 1985 DDN Protocol Handbook.

Note: This list is a portion of a list of RFC's by topic retrieved from the NIC under NETINFO:RFC-SETS.TXT (anonymous FTP of course).

The following list is not necessary for connection to the Internet, but is useful in understanding the domain system, mail system, and gateways:

RFC-882 Domain Names - Concepts and Facilities
RFC-883 Domain Names - Implementation
RFC-973 Domain System Changes and Observations
RFC-974 Mail Routing and the Domain System
RFC-1009 Requirements for Internet Gateways

-21-

Appendix C
Contact Points for Network Information

Network Information Center (NIC)

DDN Network Information Center
SRI International, Room EJ291
333 Ravenswood Avenue
Menlo Park, CA 94025
(800) 235-3155 or (415) 859-3695
NIC@SRI-NIC.ARPA

NSF Network Service Center (NNSC)

NNSC
BBN Laboratories Inc.
10 Moulton St.
Cambridge, MA 02238
(617) 497-3400
NNSC@NNSC.NSF.NET

Glossary

core gateway

The innermost gateways of the ARPAnet. These gateways have a total picture of the reachability to all networks known to the ARPAnet with EGP. They then redistribute reachability information to all those gateways speaking EGP. It is from them your EGP agent (there is one acting for you somewhere if you can reach the ARPAnet) finds out it can reach all the nets on the ARPAnet. Which is then passed to you via Hello, gated, RIP....

count to infinity

The symptom of a routing problem where routing information is passed in a circular manner through multiple gateways. Each gateway increments the metric appropriately and passes it on. As the metric is passed around the loop, it increments to ever increasing values til it reaches the maximum for the routing protocol being used, which typically denotes a link outage.

hold down

When a router discovers a path in the network has gone down announcing that that path is down for a minimum amount of time (usually at least two minutes). This allows for the propagation of the routing information across the network and prevents the formation of routing loops.

split horizon

When a router (or group of routers working in consort) accept routing information from multiple external networks, but do not pass on information learned from one external network to any others. This is an attempt to prevent bogus routes to a network from being propagated because of gossip or counting to infinity.

End of Project Gutenberg's Hitchhiker's Guide to the Internet, by Ed Krol

*** END OF THIS PROJECT GUTENBERG EBOOK
HITCHHIKER'S GUIDE TO THE INTERNET ***

***** This file should be named 39.txt or 39.zip ***** This
and all associated files of various formats will be found in:
<http://www.gutenberg.org/3/39/>

Updated editions will replace the previous one—the old editions will be renamed.

Creating the works from public domain print editions means that no one owns a United States copyright in these works, so the Foundation (and you!) can copy and distribute it in the United States without permission and without paying copyright royalties. Special rules, set forth in the General Terms of Use part of this license, apply to copying

and distributing Project Gutenberg-tm electronic works to protect the PROJECT GUTENBERG-tm concept and trademark. Project Gutenberg is a registered trademark, and may not be used if you charge for the eBooks, unless you receive specific permission. If you do not charge anything for copies of this eBook, complying with the rules is very easy. You may use this eBook for nearly any purpose such as creation of derivative works, reports, performances and research. They may be modified and printed and given away—you may do practically ANYTHING with public domain eBooks. Redistribution is subject to the trademark license, especially commercial redistribution.

*** START: FULL LICENSE ***

THE FULL PROJECT GUTENBERG LICENSE PLEASE READ THIS BEFORE YOU DISTRIBUTE OR USE THIS WORK

To protect the Project Gutenberg-tm mission of promoting the free distribution of electronic works, by using or distributing this work (or any other work associated in any way with the phrase "Project Gutenberg"), you agree to comply with all the terms of the Full Project Gutenberg-tm License (available with this file or online at <http://www.gutenberg.org/license>).

Section 1. General Terms of Use and Redistributing Project Gutenberg-tm electronic works

1.A. By reading or using any part of this Project Gutenberg-tm electronic work, you indicate that you have read, understand, agree to and accept all the terms of this license and intellectual property (trademark/copyright) agreement. If you do not agree to abide by all the terms of this agreement, you must cease using and return or destroy all copies of Project Gutenberg-tm electronic works in your possession. If you paid a fee for obtaining a copy of or access to a Project Gutenberg-tm electronic work and you do not agree to be bound by the terms of this agreement, you may obtain a refund from the person or entity to whom you paid the fee as set forth in paragraph 1.E.8.

1.B. "Project Gutenberg" is a registered trademark. It may only be used on or associated in any way with an electronic work by people who agree to be bound by the terms of this agreement. There are a few things that you can do with most Project Gutenberg-tm electronic works even without complying with the full terms of this agreement. See paragraph 1.C below. There are a lot of things you can do with Project Gutenberg-tm electronic works if you follow the terms of this agreement and help preserve free future access to Project Gutenberg-tm electronic works. See paragraph 1.E below.

1.C. The Project Gutenberg Literary Archive Foundation ("the Foundation" or PGLAF), owns a compilation copyright in the collection of Project Gutenberg-tm electronic works. Nearly all the individual works in the collection are in the public domain in the United States. If an individual work is in the public domain in the United States and you are located

in the United States, we do not claim a right to prevent you from copying, distributing, performing, displaying or creating derivative works based on the work as long as all references to Project Gutenberg are removed. Of course, we hope that you will support the Project Gutenberg-tm mission of promoting free access to electronic works by freely sharing Project Gutenberg-tm works in compliance with the terms of this agreement for keeping the Project Gutenberg-tm name associated with the work. You can easily comply with the terms of this agreement by keeping this work in the same format with its attached full Project Gutenberg-tm License when you share it without charge with others. This particular work is one of the few copyrighted individual works included with the permission of the copyright holder. Information on the copyright owner for this particular work and the terms of use imposed by the copyright holder on this work are set forth at the beginning of this work.

1.D. The copyright laws of the place where you are located also govern what you can do with this work. Copyright laws in most countries are in a constant state of change. If you are outside the United States, check the laws of your country in addition to the terms of this agreement before downloading, copying, displaying, performing, distributing or creating derivative works based on this work or any other Project Gutenberg-tm work. The Foundation makes no representations concerning the copyright status of any work in any country outside the United States.

1.E. Unless you have removed all references to Project Gutenberg:

1.E.1. The following sentence, with active links to, or other immediate access to, the full Project Gutenberg-tm License must appear prominently whenever any copy of a Project Gutenberg-tm work (any work on which the phrase "Project Gutenberg" appears, or with which the phrase "Project Gutenberg" is associated) is accessed, displayed, performed, viewed, copied or distributed:

This eBook is for the use of anyone anywhere at no cost and with almost no restrictions whatsoever. You may copy it, give it away or re-use it under the terms of the Project Gutenberg License included with this eBook or online at www.gutenberg.org

1.E.2. If an individual Project Gutenberg-tm electronic work is derived from the public domain (does not contain a notice indicating that it is posted with permission of the copyright holder), the work can be copied and distributed to anyone in the United States without paying any fees or charges. If you are redistributing or providing access to a work with the phrase "Project Gutenberg" associated with or appearing on the work, you must comply either with the requirements of paragraphs 1.E.1 through 1.E.7 or obtain permission for the use of the work and the Project Gutenberg-tm trademark as set forth in paragraphs 1.E.8 or 1.E.9.

1.E.3. If an individual Project Gutenberg-tm electronic work is posted with the permission of the copyright holder, your use and distribution must comply with both paragraphs 1.E.1 through 1.E.7 and any additional terms imposed by

the copyright holder. Additional terms will be linked to the Project Gutenberg-tm License for all works posted with the permission of the copyright holder found at the beginning of this work.

1.E.4. Do not unlink or detach or remove the full Project Gutenberg-tm License terms from this work, or any files containing a part of this work or any other work associated with Project Gutenberg-tm.

1.E.5. Do not copy, display, perform, distribute or redistribute this electronic work, or any part of this electronic work, without prominently displaying the sentence set forth in paragraph 1.E.1 with active links or immediate access to the full terms of the Project Gutenberg-tm License.

1.E.6. You may convert to and distribute this work in any binary, compressed, marked up, nonproprietary or proprietary form, including any word processing or hypertext form. However, if you provide access to or distribute copies of a Project Gutenberg-tm work in a format other than "Plain Vanilla ASCII" or other format used in the official version posted on the official Project Gutenberg-tm web site (www.gutenberg.org), you must, at no additional cost, fee or expense to the user, provide a copy, a means of exporting a copy, or a means of obtaining a copy upon request, of the work in its original "Plain Vanilla ASCII" or other form. Any alternate format must include the full Project Gutenberg-tm License as specified in paragraph 1.E.1.

1.E.7. Do not charge a fee for access to, viewing, displaying, performing, copying or distributing any Project Gutenberg-tm works unless you comply with paragraph 1.E.8 or 1.E.9.

1.E.8. You may charge a reasonable fee for copies of or providing access to or distributing Project Gutenberg-tm electronic works provided that

- You pay a royalty fee of 20% of the gross profits you derive from the use of Project Gutenberg-tm works calculated using the method you already use to calculate your applicable taxes. The fee is owed to the owner of the Project Gutenberg-tm trademark, but he has agreed to donate royalties under this paragraph to the Project Gutenberg Literary Archive Foundation. Royalty payments must be paid within 60 days following each date on which you prepare (or are legally required to prepare) your periodic tax returns. Royalty payments should be clearly marked as such and sent to the Project Gutenberg Literary Archive Foundation at the address specified in Section 4, "Information about donations to the Project Gutenberg Literary Archive Foundation."

- You provide a full refund of any money paid by a user who notifies you in writing (or by e-mail) within 30 days of receipt that s/he does not agree to the terms of the full Project Gutenberg-tm License. You must require such a user to return or destroy all copies of the works possessed in a physical medium and discontinue all use of and all access to other copies of Project Gutenberg-tm works.

- You provide, in accordance with paragraph 1.F.3, a full refund of any money paid for a work or a replacement copy, if a defect in the electronic work is discovered and reported to you within 90 days of receipt of the work.

- You comply with all other terms of this agreement for free distribution of Project Gutenberg-tm works.

1.E.9. If you wish to charge a fee or distribute a Project Gutenberg-tm electronic work or group of works on different terms than are set forth in this agreement, you must obtain permission in writing from both the Project Gutenberg Literary Archive Foundation and Michael Hart, the owner of the Project Gutenberg-tm trademark. Contact the Foundation as set forth in Section 3 below.

1.F.

1.F.1. Project Gutenberg volunteers and employees expend considerable effort to identify, do copyright research on, transcribe and proofread public domain works in creating the Project Gutenberg-tm collection. Despite these efforts, Project Gutenberg-tm electronic works, and the medium on which they may be stored, may contain "Defects," such as, but not limited to, incomplete, inaccurate or corrupt data, transcription errors, a copyright or other intellectual property infringement, a defective or damaged disk or other medium, a computer virus, or computer codes that damage or cannot be read by your equipment.

1.F.2. LIMITED WARRANTY, DISCLAIMER OF DAMAGES - Except for the "Right of Replacement or Refund" described in paragraph 1.F.3, the Project Gutenberg Literary Archive Foundation, the owner of the Project Gutenberg-tm trademark, and any other party distributing a Project Gutenberg-tm electronic work under this agreement, disclaim all liability to you for damages, costs and expenses, including legal fees. YOU AGREE THAT YOU HAVE NO REMEDIES FOR NEGLIGENCE, STRICT LIABILITY, BREACH OF WARRANTY OR BREACH OF CONTRACT EXCEPT THOSE PROVIDED IN PARAGRAPH 1.F.3. YOU AGREE THAT THE FOUNDATION, THE TRADEMARK OWNER, AND ANY DISTRIBUTOR UNDER THIS AGREEMENT WILL NOT BE LIABLE TO YOU FOR ACTUAL, DIRECT, INDIRECT, CONSEQUENTIAL, PUNITIVE OR INCIDENTAL DAMAGES EVEN IF YOU GIVE NOTICE OF THE POSSIBILITY OF SUCH DAMAGE.

1.F.3. LIMITED RIGHT OF REPLACEMENT OR REFUND - If you discover a defect in this electronic work within 90 days of receiving it, you can receive a refund of the money (if any) you paid for it by sending a written explanation to the person you received the work from. If you received the work on a physical medium, you must return the medium with your written explanation. The person or entity that provided you with the defective work may elect to provide a replacement copy in lieu of a refund. If you received the work electronically, the person or entity providing it to you may choose to give you a second opportunity to receive the work electronically in lieu of a refund. If the second copy is also defective, you may demand a refund in writing without further opportunities to fix the problem.

1.F.4. Except for the limited right of replacement or refund set forth in paragraph 1.F.3, this work is provided to you 'AS-IS,' WITH NO OTHER WARRANTIES OF ANY KIND, EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO WARRANTIES OF MERCHANTABILITY OR FITNESS FOR ANY PURPOSE.

1.F.5. Some states do not allow disclaimers of certain implied warranties or the exclusion or limitation of certain types of damages. If any disclaimer or limitation set forth in this agreement violates the law of the state applicable to this agreement, the agreement shall be interpreted to make the maximum disclaimer or limitation permitted by the applicable state law. The invalidity or unenforceability of any provision of this agreement shall not void the remaining provisions.

1.F.6. INDEMNITY - You agree to indemnify and hold the Foundation, the trademark owner, any agent or employee of the Foundation, anyone providing copies of Project Gutenberg-tm electronic works in accordance with this agreement, and any volunteers associated with the production, promotion and distribution of Project Gutenberg-tm electronic works, harmless from all liability, costs and expenses, including legal fees, that arise directly or indirectly from any of the following which you do or cause to occur: (a) distribution of this or any Project Gutenberg-tm work, (b) alteration, modification, or additions or deletions to any Project Gutenberg-tm work, and (c) any Defect you cause.

Section 2. Information about the Mission of Project Gutenberg-tm

Project Gutenberg-tm is synonymous with the free distribution of electronic works in formats readable by the widest variety of computers including obsolete, old, middle-aged and new computers. It exists because of the efforts of hundreds of volunteers and donations from people in all walks of life.

Volunteers and financial support to provide volunteers with the assistance they need are critical to reaching Project Gutenberg-tm's goals and ensuring that the Project Gutenberg-tm collection will remain freely available for generations to come. In 2001, the Project Gutenberg Literary Archive Foundation was created to provide a secure and permanent future for Project Gutenberg-tm and future generations. To learn more about the Project Gutenberg Literary Archive Foundation and how your efforts and donations can help, see Sections 3 and 4 and the Foundation web page at <http://www.pglaaf.org>.

Section 3. Information about the Project Gutenberg Literary Archive Foundation

The Project Gutenberg Literary Archive Foundation is a non profit 501(c)(3) educational corporation organized under the laws of the state of Mississippi and granted tax exempt status by the Internal Revenue Service. The Foundation's EIN or federal tax identification number is 64-6221541. Its 501(c)(3) letter is posted at <http://pglaf.org/fundraising>. Contributions to the Project Gutenberg Literary Archive Foundation are tax deductible to the full extent permitted by U.S. federal laws and your state's laws.

The Foundation's principal office is located at 4557 Melan Dr. S. Fairbanks, AK, 99712., but its volunteers and employees are scattered throughout numerous locations. Its business office is located at 809 North 1500 West, Salt Lake City, UT 84116, (801) 596-1887, email business@pglaf.org. Email contact links and up to date contact information can be found at the Foundation's web site and official page at <http://pglaf.org>

For additional contact information:

Dr. Gregory B. Newby
Chief Executive and Director
gbnewby@pglaf.org

Section 4. Information about Donations to the Project Gutenberg Literary Archive Foundation

Project Gutenberg-tm depends upon and cannot survive without wide spread public support and donations to carry out its mission of increasing the number of public domain

and licensed works that can be freely distributed in machine readable form accessible by the widest array of equipment including outdated equipment. Many small donations (\$1 to \$5,000) are particularly important to maintaining tax exempt status with the IRS.

The Foundation is committed to complying with the laws regulating charities and charitable donations in all 50 states of the United States. Compliance requirements are not uniform and it takes a considerable effort, much paperwork and many fees to meet and keep up with these requirements. We do not solicit donations in locations where we have not received written confirmation of compliance. To SEND DONATIONS or determine the status of compliance for any particular state visit <http://pglaf.org>

While we cannot and do not solicit contributions from states where we have not met the solicitation requirements, we know of no prohibition against accepting unsolicited donations from donors in such states who approach us with offers to donate.

International donations are gratefully accepted, but we cannot make any statements concerning tax treatment of donations received from outside the United States. U.S. laws alone swamp our small staff.

Please check the Project Gutenberg Web pages for current donation methods and addresses. Donations are accepted in a number of other ways including checks, online payments and credit card donations. To donate, please visit: <http://pglaf.org/donate>

Section 5. General Information About Project Gutenberg-tm electronic works.

Professor Michael S. Hart is the originator of the Project Gutenberg-tm concept of a library of electronic works that could be freely shared with anyone. For thirty years, he produced and distributed Project Gutenberg-tm eBooks with only a loose network of volunteer support.

Project Gutenberg-tm eBooks are often created from several printed editions, all of which are confirmed as Public Domain in the U.S. unless a copyright notice is included. Thus, we do not necessarily keep eBooks in compliance with any particular paper edition.

Each eBook is in a subdirectory of the same number as the eBook's eBook number, often in several formats including plain vanilla ASCII, compressed (zipped), HTML and others.

Corrected EDITIONS of our eBooks replace the old file and take over the old filename and etext number. The replaced older file is renamed. VERSIONS based on separate sources are treated as new eBooks receiving new filenames and etext numbers.

Most people start at our Web site which has the main PG search facility:

<http://www.gutenberg.org>

This Web site includes information about Project Gutenberg-tm, including how to make donations to the Project Gutenberg Literary Archive Foundation, how to help produce our new eBooks, and how to subscribe to our email newsletter to hear about new eBooks.

EBooks posted prior to November 2003, with eBook numbers BELOW #10000, are filed in directories based on their release date. If you want to download any of these eBooks directly, rather than using the regular search system you may utilize the following addresses and just download by the etext year.

<http://www.ibiblio.org/gutenberg/etext06>

(Or /etext 05, 04, 03, 02, 01, 00, 99, 98, 97, 96, 95, 94, 93, 92, 91 or 90)

EBooks posted since November 2003, with etext numbers OVER #10000, are filed in a different way. The year of a release date is no longer part of the directory path. The path is based on the etext number (which is identical to the filename). The path to the file is made up of single digits corresponding to all but the last digit in the filename. For example an eBook of filename 10234 would be found at:

<http://www.gutenberg.org/1/0/2/3/10234>

or filename 24689 would be found at:
<http://www.gutenberg.org/2/4/6/8/24689>

An alternative method of locating eBooks:
<http://www.gutenberg.org/GUTINDEX.ALL>

***** END: FULL LICENSE *****