Network Working Group
Request for Comments: 3675
Category: Informational

D. Eastlake 3rd
Motorola Laboratories

February 2004

.sex Considered Dangerous

Status of this Memo

Copyright Notice

Abstract

Periodically there are proposals to mandate the use of a special top level name or an IP address bit to flag "adult" or "unsafe" material or the like. This document explains why this is an ill considered idea from the legal, philosophical, and particularly, the technical points of view.

Table of Contents

1. Introduction

Periodically there are proposals to mandate the use of a special top level name or an IP address bit to flag "adult" or "unsafe" material or the like. This document explains why this is an ill considered idea from the legal, philosophical, and the technical points of view.

2. Background

The concept of a .sex, .xxx, .adult, or similar top-level domain in which it would be mandatory to locate salacious or similar material is periodically suggested by some politicians and commentators. Other proposals have included a domain reserved exclusively for material viewed as appropriate for minors, or using IP address bits or ranges to segregate content.

In an October 1998 report accompanying the Child Online Protection Act, the House Commerce committee said, "there are no technical barriers to creating an adult domain, and it would be very easy to block all websites within an adult domain". The report also said that the committee was wary of regulating the computer industry and that any decision by the U.S. government "will have international consequences" [HOUSEREPORT].

British Telecom has backed adult top-level domains, saying in a 1998 letter to the U.S. Department of Commerce that it "strongly supported" that plan. The reason: "Sexually explicit services could then be legally required to operate with domain names in this gTLD [that] would make it much simpler and easier to control access to such sites..." [BT]. One of ICANN's progenitors, the GTLD-MOU committee, suggested a "red-light-zone" top-level domain in a September 1997 request for comment [GTLD-MOU].

Some adult industry executives have endorsed the concept. In 1998, Seth Warshavsky, president of the Internet Entertainment Group, told the U.S. Senate Commerce committee that he would like to see a .adult domain. "We're suggesting the creation of a new top-level domain called '.adult' where all sexually explicit material on the Net would reside," Warshavsky said in an interview at the time [WARSHAVSKY].

More recently, other entrepreneurs in the industry have said that they do not necessarily object to the creation of an adult domain as long as they may continue to use .com.

Conservative groups in the U.S. say they are not eager for such a domain, and prefer criminal laws directed at publishers and distributors of sexually-explicit material. The National Law Center for Children and Families in Fairfax, Virginia, said in February 2001 that it did not favor any such proposal. For different reasons, the American Civil Liberties Union and other civil liberties groups also oppose it.

Sen. Joseph Lieberman, the U.S. Democratic Party's vice presidential nominee, endorsed the idea at a June 2000 meeting of the federal Commission on Child Online Protection. Lieberman said in a prepared statement that "we would ask the arbiters of the Internet to simply abide by the same standard as the proprietor of an X-rated movie theater or the owner of a convenience store who sells sexually-explicit magazines" [LIEBERMAN].

In the 1998 law creating this commission, the U.S. Congress required the members to investigate "the establishment of a domain name for posting of any material that is harmful to minors", The commission devoted a section of its October 2000 report to that topic. It concluded that both a .xxx and a .kids domain are technically possible, but would require action by ICANN. The report said that an adult domain might be only "moderately effective" and raises privacy and free speech concerns [COPAREPORT].

The commission also explored the creation of a so-called red zone or green zone for content by means of allocation of a new set of IP addresses under IPv6. Any material not in one of those two zones would be viewed as in a gray zone and not necessarily appropriate or inappropriate for minors. Comments from commissioners were largely negative: "Effectiveness would require substantial effort to attach content to specific IP numbers. This approach could potentially reduce flexibility and impede optimal network performance. It would not be effective at blocking access to chat, newsgroups, or instant messaging".

In October 2000, ICANN rejected a .xxx domain during its initial round of approving additional top-level domains. The reasons are not entirely clear, but former ICANN Chairwoman Esther Dyson said that the adult industry did not entirely agree that such a domain would be appropriate. One .xxx hopeful, ICM Registry of Ontario, Canada, in December 2000 asked ICANN to reconsider its decision [ICM-REGISTRY].

In 2002, the U.S. Congress mandated the creation of a kids.us domain for "child safe" material. This was after being convinced that for reasons, some of which are described in the following section, trying to legislate standards for the whole world with a .kids domain was inappropriate.

3. Legal and Philosophical Problems

When it comes to sexually-explicit material, every person, court, and government has a different view of what's acceptable and what is not. Attitudes change over time, and what is viewed as appropriate in one town or year may spark protests in the next. When faced with the slippery nature of what depictions of sexual activity should be illegal or not, one U.S. Supreme Court justice blithely defined obscenity as: "I know it when I see it".

In the U.S.A., obscenity is defined as explicit sexual material that, among other things, violates "contemporary community standards" -- in other words, even at the national level, there is no agreed-upon rule governing what is illegal and what is not. Making matters more knotty is that there are over 200 United Nations country codes, and in most of them, political subdivisions can impose their own restrictions. Even for legal nude modeling, age restrictions differ. They're commonly 18 years of age, but only 17 years of age in one Scandinavian country. A photographer there conducting what's viewed as a legal and proper photo shoot would be branded a felon and child pornographer in the U.S.A. In yet other countries and groups, the entire concept of nude photography or even any photography of a person in any form may be religiously unacceptable.

Saudi Arabia, Iran, Northern Nigeria, and China are not likely to have the same liberal views as, say, the Netherlands or Denmark. Saudi Arabia and China, like some other nations, extensively filter their Internet connection and have created government agencies to protect their society from web sites that officials view as immoral. Their views on what should be included in a .sex domain would hardly be identical to those in liberal western nations.

Those wildly different opinions on sexual material make it inconceivable that a global consensus can ever be reached on what is appropriate or inappropriate for a .sex or .adult top-level domain. Moreover, the existence of such a domain would create an irresistible temptation on the part of conservative legislators to require controversial publishers to move to that domain and punish those who do not.

Some conservative politicians already have complained that ICANN did not approve .xxx in its October 2000 meeting. During a February 2001 hearing in the U.S. House of Representatives, legislators warned that they "want to explore ICANN's rationale for not approving two particular top level domain names -- .kids and .xxx -- as a means to protect kids from the awful smut which is so widespread on the Internet".

It seems plausible that only a few adult publishers, and not those who have invested resources in building a brand around a .com site, would voluntarily abandon their current domain name. Instead, they'd likely add a .xxx variant and keep their original address. The existence of .xxx could propel legislators in the U.S. and other countries to require them to publish exclusively from an adult domain, a move that would invite ongoing political interference with Internet governance, and raise concerns about forced speech and self-labeling.

In fact, the ultimate arbiter of generic top-level domain names -- at least currently -- is not ICANN, but the U.S. government. The U.S. Congress' General Accounting Office in July 2000 reported that the Commerce Department continues to be responsible for domain names allowed by the authoritative root [GAO]. The GAO's auditors concluded it was unclear whether the Commerce Department has the "requisite authority" under current law to transfer that responsibility to ICANN.

The American Civil Liberties Union -- and other members of the international Global Internet Liberty Campaign -- caution that publishers speaking frankly about birth control, AIDS prevention, gay and lesbian sex, the social problem of prison rape, etc., could be coerced into moving to an adult domain. Once there, they would be stigmatized and easily blocked by schools, libraries, companies, and other groups using filtering software. Publishers of such information, who do not view themselves as pornographers and retain their existing addresses, could be targeted for prosecution.

The existence of an adult top-level domain would likely open the door for related efforts, either policy or legislative. There are many different axes through which offensive material can be defined: Sex, violence, hate, heresy, subversion, blasphemy, illegal drugs, profanity, political correctness, glorification of crime, incitement to break the law, and so on. Such suggestions invite the ongoing lobbying of ICANN, the U.S. government, and other policy-making bodies by special-interest groups that are not concerned with the technical feasibility or practicality of their advice.

An adult top-level domain could have negative legal repercussions by endangering free expression. U.S. Supreme Court Justice Sandra Day O'Connor has suggested that the presence of "adult zones" on the Internet would make a future Communications Decency Act (CDA) more likely to be viewed as constitutional. In her partial dissent to the Supreme Court's rejection of the CDA in 1997 [CDA], O'Connor said that "the prospects for the eventual zoning of the Internet appear promising". (The Supreme Court ruled that the CDA violated free speech rights by making it a crime to distribute "indecent" or "patently offensive" material online.)

Privacy could be harmed by such a proposal. It would become easier for repressive governments and other institutions to track visits to sites in a domain labeled as adult and record personally-identifiable information about the visitor. Repressive governments would instantly have more power to monitor naive users and prosecute them for their activities. It's also implausible that a top-level domain would be effective in controlling access to chat, email, newsgroups, instant messaging, and new services as yet to be invented.

4. Technical Difficulties

Even ignoring the philosophical and legal difficulties outlined above, there are substantial technical difficulties in attempting to impose content classification by domain names or IP addresses. Mandatory content labeling is usually advanced with the idea of using a top level domain name, discussed in section 4.1., but we also discuss the possibility of using IP address bits or ranges in section 4.2.

In section 4.1.4., difficulties with a few particular higher level protocols are discussed. In some cases, these protocols use different name spaces. It should be kept in mind that additional future protocols may be devised with as yet undreamed of naming characteristics.

We also discuss PICS labels [PICS] as an alternative technology in section 4.3.

Only a limited technical background is assumed, so some basic information is included below. In some cases, descriptions are simplified and details omitted.

This technical discussion minimizes the definitional problems. However, it is still necessary for evaluating some technical considerations to have some estimate of the amount of categorization that would be necessary for a realistic global censorship system. There is no hope of agreement on this point. For our purposes, we

will arbitrarily assume that the world's population consists of approximately 90,000 overlapping communities, each of which would have a different categorization of interest. Further, we arbitrarily assume that some unspecified but clever encoding scheme enables a proper global categorization of all information by a 300 bit label. Some would say a 300 bit label is too large, others that it is too small. Regardless, we will use it for some technical evaluations.

4.1. Content Filtering Using Names

The most prominent user visible part of Internet naming and addressing is the domain name system [RFC 1034, 1035]. Domain Names are dotted sequences of labels, such as aol.com, world.std.com, www.rosslynchapel.org.uk, or ftp.gnu.lcs.mit.edu [RFC 1035, 1591, 2606]. Domain Names form an important part of most World Wide Web addresses or URLs [RFC 2396], commonly appearing after "//". Security for the domain name system is being standardized [RFC 2535], but has not been deployed to any significant extent.

Domain names designate nodes in a globally distributed hierarchically delegated database. A wide variety of information can be stored at these nodes, including IP addresses of machines on the network (see section 4.2. below), mail delivery information, and other types of information. Thus, the data stored at foo.example.com could be the numeric information for sending data to a particular machine, which would be used if you tried to browse <http://foo.example.com>, the name of a computer (say mailhost.example.com) to handle mail addressed to anyone "@foo.example.com", and/or other information.

There are also other naming systems in use, such as news group names and Internet Relay Chat (IRC) channel names.

The usual labeling idea presented is to reserve a top level name, such as .sex or .xxx for "adult" material and/or .kids for "safe" material or the like. The technical and linguistic problems with this are described in the subsections below.

4.1.1. Linguistic Problems

When using name labeling, the first problem is from whose language do you take the names to impose? Words and acronyms can have very different meanings in different languages and the probability of confusion is multiplied when phonetic collisions are considered.

As an example of possible problems, note that for several years the government of Turkmenistan suspended new registrations in ".tm", which had previously been a source of revenue, because some of the

registered second level domain names may have been problematic. In particular, their web home page at <http://www.nic.tm> said:

> Statement from the .TM NIC
>
> > "The response to the .TM registry has been overwhelming. Thousands of names have been registered from all over the world. Some of the names registered, however, may be legally obscene in Turkmenistan, and as a result the .TM NIC registry is reviewing its naming policy for future registrations. The .TM NIC has suspended registrations until a new policy can be implemented. We hope to be live again shortly."

There are approximately 6,000 languages in use in the world today, although this is expected to decline to around 3,000 by the year 2100.

4.1.2. Explosion of Top Level Domain Names (TLDs)

An important aspect of the design of the Domain Name System (DNS) is the hierarchical delegation of data maintenance. The DNS really only works, and has been able to scale over the five orders of magnitude it has grown since its initial deployment, due to this delegation.

The first problem is that one would expect most computers or web sites to have a mix of material, only some of which should be specially classified. Using special top level domain names (TLDs) multiplies the number of DNS zones the site has to worry about. For example, assume the site has somehow already sorted its material into "kids", "normal", and "adult" piles. Without special TLD labels, it can store them under kids.example.net, adult.example.net, and other.example.net, for instance. This would require only the maintenance of the single example.net zone of database entries. With special TLD labeling, at least example.net (for normal stuff), example.net.sex, and example.net.kids would need to be maintained, which are in three separate zones, in different parts of the DNS tree, under three separate delegations.

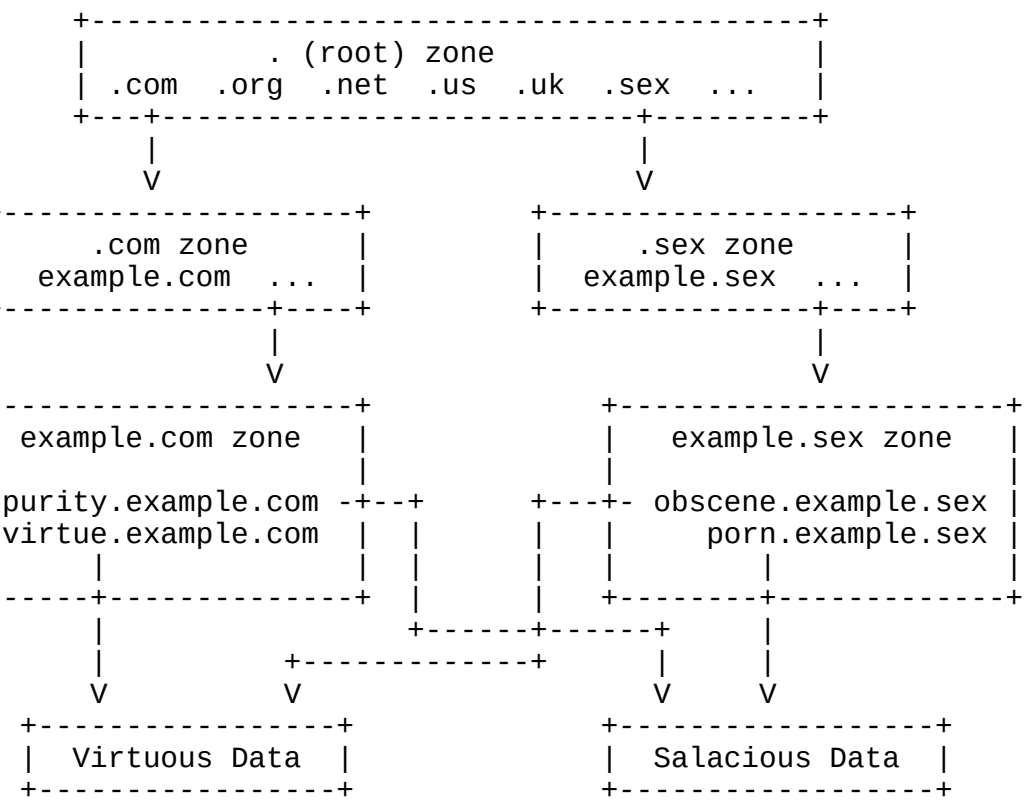
As the number of categories expands, the number of category combinations explodes, and this quickly becomes completely unmanageable. If 300 bits worth of labeling is required, the system could, in theory, need 2**300 name categories, an impossibility. No individual site would need to use all categories and the category domain names would not all have to be top level names. But it would still be an unmanageable nightmare.

4.1.3. You Can’t Control What Names Point At You!

Providers of data on the Internet cannot stop anyone from creating names pointing to their computer’s IP address with misleading domain names.

The DNS system works as a database. It associates certain data, called resource records, or RRs, with domain names. In particular, it can associate IP address resource records with domain names. For example, when you browse a URL, most commonly a domain name within that URL is looked up in the DNS. The resulting address is then used to address the packets sent from your web browser or other software to the server or peer.

Remember what we said in Section 4.1.1. about hierarchical delegation? Control is delegated and anyone controlling a DNS zone of data, say example.com, can insert data at that name or any deeper name (except to the extent that they delegate some of the deeper namespace to yet others). So the controller of example.com can insert data so that purity.example.com has, associated with it, the same computer address, which is associated with www.obscene.example.sex. This directs any reference to purity.example.com to use the associated IP address which is the same as the www.obscene.example.sex web site. The manager of that hypothetical web site, who controls the obscene.example.xxx zone, has no control over the example.com DNS zone. They are technically incapable of causing it to conform to any ".sex" labeling law. In the alternative, someone could create a name conforming to an adult labeling requirement, such as foo.stuff.sex, that actually pointed to someone else’s entirely unobjectionable site, perhaps for the purpose of polluting the labeling. See diagram below. Each "zone" could be hosted on a different set of physical computers.



4.1.4. Particular Protocol Difficulties

There are additional considerations related to particular protocols. We consider only a few here. The first two, electronic mail and the World Wide Web, use domain name addressing. The second two, net news and IRC, use different name spaces and illustrate further technical problems with name based labeling.

4.1.4.1. Electronic Mail (SMTP)

Standard Internet tools provide no way to stop users from putting arbitrary domain names inside email headers.

The standard Internet electronic mail protocol separates "envelope" information from content [RFC 2821, 2822]. The envelope information indicates where a message claims to have originated and to whom it should be delivered. The content has fields starting with labels like "From:" and "To:", but these content fields actually have no effect and can be arbitrarily forged using simple, normally available software, such a telnetting to the SMTP port on a mail server. Content fields are not compared with envelope fields. To require them to be the same would be like requiring that postal letters

deposited in a mail box list that mail box as their return address and only allowing residence or business return addresses on mail picked up by the post office from that residence or business.

While different mail clients display envelope information and headers from the content of email differently, generally the principle content fields are given prominence. Thus, while not exactly the same as content labeling, it should be noted that it is trivial to send mail to anyone with arbitrary domain names in the email addresses appearing in the From and To headers, etc.

It is also easy up set up a host to forward mail to an email address or mailing list. Mail sent with normal mail tools to this forwarder will automatically have content headers reflecting the forwarder's name, but the forwarder will change the envelope information and cause the mail to be actually sent to the forwarding destination mail address.

For example, (with names disguised) there is a social mailing list innocuous@foo.example.org, and someone set up a forwarder at cat-torturers@other.example. Mail sent to the forwarder is forwarded and appears on the innocuous mailing list but with a "To: cat-torturers@other.example" header in its body, instead of the usual "To: innocuous@foo.example.org" content header. Mail reader software then displays the cat-torturers header. Similar things can be done using the "bcc" or "blind courtesy copy" feature of Internet mail.

There is work proceeding on securing email; however, such efforts at present only allow you to verify whether or not a particular entity was the actual author of the mail. When providing authentication, they add yet a third type of "From" address to the envelope and content "From" addresses, but they do not relate to controlling or authenticating domain names in the content of the mail.

4.1.4.2. Web Access (HTTP)

With modern web servers and browsers supporting HTTP 1.1 [RFC 2616], the domain name used to access the site is available. Thus, web sites with different domain names can be accessed even if they are on the same machine at the same IP address. This is a small plus for name-based labeling since different categories of information on the same computer can be set up to be accessed via different domain names. But for a computer with any reasonable variety of data, the explosion of trying to differently name all types of data would require an unmanageable number of names.

With earlier HTTP 1.0 [RFC 1945], when a web request was sent to a server machine, the original domain name used in the URI was not included.

On the other hand, the web has automatic forwarding. Thus, when one tries to access data at a particular domain name, the server there can re-direct your browser, temporarily or permanently, to a different name, or it can re-direct you to a numeric IP address so as to by-pass name filtering.

4.1.4.3. News (NNTP)

Net news [RFC 977, 2980] uses hierarchically structured newsgroup names that are similar in appearance to domain names, except that the most significant label is on the left and the least on the right, the opposite of domain names. However, while the names are structured hierarchically, there is no central control. Instead, news servers periodically connect to other news servers that have agreed to exchange messages with them and they update each other on messages only in those newsgroups in which they wish to exchange messages.

Although hierarchical zones in the domain name system are locally managed, they need to be reachable starting at the top level root servers which are in turn more or less controlled by ICANN and the US Department of Commerce. With no such central point or points in the net news world, any pair or larger set of news servers anywhere in the world can agree to exchange news messages under any news group names they like, including duplicates of those used elsewhere in the net, making central control or even influence virtually impossible. In fact, within some parts of the news group namespace on some servers, anyone can create new newsgroups with arbitrary names.

Even if news group names could be controlled, the contents of the messages are determined by posters. While some groups are moderated, most are not. "Cancel" messages can be sent out for news messages, but that mechanism is subject to abuse, so some servers are configured to ignore cancels. In any case, the message may have been distributed to a huge number of computers world wide before any cancel is sent out.

And of course, fitting 300 bits worth of labeling into news group names is just as impossible as it is to fit into domain names.

4.1.4.4. Internet Relay Chat (IRC)

Internet Relay Chat [RFC 2810-2813] is another example of a service which uses a different name space. It uses a single level space of "channel names" that are meaningful within a particular network of IRC servers. Because it is not hierarchical, each server must know about all names, which limits the size of a network of servers.

As with newsgroup names, the fact that IRC channel names are local decisions, not subject to or reachable from any global "root", makes centralized political control virtually impossible.

4.2. Content Filtering Using IP Addressing

A key characteristic of the Internet Protocol (IP) on which the Internet is based is that it breaks data up into "packets". These packets are individually handled and routed from source to destination. Each packet carries a numeric address for the destination point to which the Internet will try to deliver the packet.

(End users do not normally see these numeric addresses but instead deal with "domain names" as described in section 4.1. above.)

The predominant numeric address system now in use is called IPv4, or Internet Protocol Version 4, which provides for 32 bit addresses [RFC 791]. There is increasing migration to the newer IPv6 [RFC 2460], which provides for 128 bit addresses [RFC 2373, 2374].

Packets can be modified maliciously in transit but the most common result of this is denial of service.

One problem in using addressing for content filtering is that this is a very coarse technique. IP addresses refer to network interfaces, which usually correspond to entire computer systems which could house multiple web pages, sets of files, etc., only a small part of which it was desired to block or enable. Increasingly, a single IP address may correspond to a NAT (Network Address Translation) box [RFC 2663] which hides multiple computers behind it, although in that case, these computers are usually not servers.

However, even beyond this problem of coarse granularity, the practical constraints of hierarchical routing make the allocation of even a single IPv4 address bit or a significant number of IPv6 address bits impossible.

4.2.1. Hierarchical Routing

IP addresses are technically inappropriate for content filtering because their assignment is intimately tied to network routing and topology.

As packets of data flow through the Internet, decisions must be made as to how to forward them "towards" their destination. This is done by comparing the initial bits of the packet destination address to entries in a "routing table" and forwarding the packets as indicated by the table entry with the longest prefix match.

While the Internet is actually a mesh, if, for simplicity, we consider it to have a central backbone at the "top", a packet is typically routed as follows:

The local networking code looks at its routing table to determine if the packet should be sent directly to another computer on the "local" network, to a router to specially forward it to another nearby network, or routed "up" to a "default" router to forward it to a higher level service provider's network. If the packet's destination is "far enough away", it will eventually get forwarded up to a router on the backbone. Such a router cannot send the packet "up" since it is at the top, or "default free" zone, and must have a complete table of other top level routers in which to send the packet. Currently, such top level routers are very large and expensive devices. They must be able to maintain tables of tens of thousands of routes. When the packet gets to the top level router of the part of the network within which its destination lies, it gets forwarded "down" to successive routers which are more and more specific and local until eventually it gets to a router on the local network where its destination address lies. This local router sends the packet directly to the destination computer.

Because all of these routing decisions are made on a longest prefix match basis, it can be seen that IP addresses are not general names or labels, but are critically and intimately associated with the actual topology and routing structure of the network. If they were assigned at random, routers would be required to remember so many specific routes for specific addresses that it would far exceed the current technical capabilities for router design. The Internet would be fatally disrupted and would not work.

It should also be noted that there is some inefficiency in allocation at each level of hierarchy [RFC 1715]. Generally, allocations are of a power of two addresses and as requirements grow and/or shrink, it is not practical to use every address.

(The above simplified description ignores multi-homing and many other details.)

4.2.2. IP Version 4 Addresses

There just isn't any practical way to reallocate even one bit of IPv4 global Internet Addresses for content filtering use. Such addresses are in short supply. Such an allocation would, in effect, cut the number of available addresses in half. There just aren't enough addresses, even without the inefficiency of hierarchical allocation [RFC 1715] and routing, to do this. Even if there were, current numbers have not been allocated with this in mind so that renumbering by every organization with hosts on the Internet would be required, a Herculean task costing in the billions of dollars.

Even if these problems were overcome, the allocation of even a single bit near the top of the address bits would likely double the number of routes in the default free zone. This would exceed the capacity of current routers and require the upgrade of thousands of them to new routers that do not exist yet at a gargantuan cost. The allocation of a bit near the bottom of the address bits would require world-wide local reconfiguration which would be impractical to require or enforce, even if the bit were available.

And all this is if only a single bit is allocated to content labeling, let alone more than one. And we are assuming you would actually need 300 bits, more than there are!

Basically, the idea is a non-starter.

4.2.3. IP Version 6 Addresses

IPv6 provides 128 bit address fields [RFC 2373, 2374]. Furthermore, allocation of IPv6 addresses is in its infancy. Thus, the allocation of say, one bit of IPv6 address for labeling is conceivable.

However, as discussed above (section 4.2.1.), every high bit allocated for labeling doubles the cost imposed on the routing system. Allocating one bit would generally double the size of routing tables.

Allocating two bits would multiply them by four. Allocating the 300 bits we assume necessary for realistic world wide labeling is logically impossible for IPv6, 300 being a lot larger than 128, and if it were, would result in technically unachievable routing table sizes. Even allocating, say, 20 bits, if that were possible, would impossibly multiply table sizes by a million.

Allocating low bits also has problems. There are technical proposals that use the bottom 64 bits in a manner incompatible with their use for labels [RFC 2374]. So it would probably have to be "middle bits" (actually low bits of the upper half). As with IPv4, it would be impossible to enforce this world wide. If it were possible, one or two bits could be allocated there, which would be clearly inadequate.

4.3. PICS Labels

PICS Labels (Platform for Internet Content Selection) is a generalized system for providing "ratings" for Internet accessible material. The PICS documents [PICS] should be consulted for details. In general, PICS assumes an arbitrarily large number of rating services and rating systems. Each service and system is identified by a URL.

It would be quite reasonable to have multiple PICS services that, in the aggregate, provided 300 bits of label information or more. There could be a PICS service for every community of interest. This sort of technology is really the only reasonable way to make categorizations or labelings of material available in a diverse and dynamic world.

While such PICS label services could be used to distribute government promulgated censorship categories, for example, it is not clear how this is any worse than government censorship via national firewalls.

A PICS rating system is essentially a definition of one or more dimensions and the numeric range of the values that can be assigned in each dimension to a rated object. A service is a source of labels where a label includes actual ratings. Ratings are either specific or generic. A specific rating applies only to the material at a particular URL [RFC 2396] and does not cover anything referenced from it, even included image files. A generic rating applies to the specified URL and to all URLs for which the stated URL is a prefix.

A simplified example label might look like the following:

```
(PICS-1.1 "http://movie-rating-service.example.net"
   labels for "ftp://movies.example.sex/raunchy-movie"
   ratings (sex 6 violence 1 language 8 drugs 2 Satanism 0))
```

Machine readable rating system descriptions include the range of values and set of dimensions provided. Additional information, such as beginning and ending time of validity, can be incorporated into labels.

Labels can currently be made available in three ways: (1) embedded in HTML, (2) provided with data in an HTTP response, and (3) separately from a third party. If content is required to have labels embedded in it or transmitted by the source when data is returned, as in the first two ways listed above, it raises the problems of categorization granularity and forced speech. However, if used in the third way whereby a separate party determines and provides labels for content, and users are free to select whatever such third party or parties they wish to consult, it can support a myriad of categories, editors, and evaluators to exist in parallel.

Digital signatures are available to secure PICS Labels [PICS].

5. Security Considerations

Any labeling or categorization scheme must assume that there will be deliberate attempts to cause data to be incorrectly labeled and incorrectly categorized. This might be due to some perceived advantage of particular labeling or merely to disrupt the system. After all, if sources would always accurately and conveniently label sent information, security would be much easier [RFC 3514]. Such enforceability considerations are discussed in conjunction with the various mechanisms mentioned in this document.

6. Conclusions

The concept that a single top level domain name, such as .sex, or a single IP address bit, could be allocated and become the mandatory home of "adult" or "offensive" material world wide is legal and technical nonsense.

Global agreement on what sort of material should be in such a ghetto is impossible. In the world wide context, the use of a single category or small number of categories is absurd. The implementation of a reasonable size label that could encompass the criterion of the many communities of the world, such as 300 bits, is technically impossible at the domain name or IP address level and will remain so for the foreseeable future. Besides technical impossibility, such a mandate would be an illegal forcing of speech in some jurisdictions, as well as cause severe linguistic problems for domain or other character string names.

However, the concept of a plethora of independent reviewers, some of which might be governmental agencies, and the ability of those accessing information to select and utilize ratings assigned by such reviewers, is possible.

7. References

7.1. Normative References

[PICS] Platform for Internet Content Selection PICS 1.1 Rating Services and Rating Systems -- and Their Machine Readable Descriptions <http://www.w3.org/TR/REC-PICS-services>, October 1996.

PICS 1.1 Label Distribution -- Label Syntax and Communication Protocols <http://www.w3.org/TR/REC-PICS-labels>, October 1996.

PICSRules 1.1 Specification <http://www.w3.org/TR/REC-PICSRules>, December 1997.

PICS Signed Labels (DSIG) 1.0 Specification <http://www.w3.org/TR/REC-DSig-label/>, May 1998.

[RFC 791] Postel, J., "Internet Protocol", STD 5, RFC 791, September 1981.

[RFC 977] Kantor, B. and P. Lapsley, "Network News Transfer Protocol", RFC 977, February 1986.

[RFC 1035] Mockapetris, P., "Domain Names - Implementation and Specifications", STD 13, RFC 1035, November 1987.

[RFC 1591] Postel, J., "Domain Name System Structure and Delegation", RFC 1591, March 1994.

[RFC 1945] Berners-Lee, T., Fielding, R. and H. Frystyk, "Hypertext Transfer Protocol -- HTTP/1.0", RFC 1945, May 1996.

[RFC 2373] Hinden, R. and S. Deering, "IP Version 6 Addressing Architecture", RFC 2373, July 1998.

[RFC 2374] Hinden, R., O'Dell, M. and S. Deering, "An IPv6 Aggregatable Global Unicast Address Format", RFC 2374, July 1998.

[RFC 2616] Fielding, R., Gettys, J., Mogul, J., Frystyk, H., Masinter, L., Leach, P. and T. Berners-Lee, "Hypertext Transfer Protocol -- HTTP/1.1", RFC 2616, June 1999.

[RFC 2663] Srisuresh, P. and M. Holdrege, "IP Network Address Translator (NAT) Terminology and Considerations", RFC 2663, August 1999.

[RFC 2810] Kalt, C., "Internet Relay Chat: Architecture", RFC 2810, April 2000.

[RFC 2821] Klensin, J., Ed., "Simple Mail Transfer Protocol", RFC 2821, April 2001.

[RFC 2822] Resnick, P., Ed., "Internet Message Format", RFC 2822, April 2001.

[RFC 2980] Barber, S., "Common NNTP Extensions", RFC 2980, October 2000.

7.2. Informative References

[BT] "British Telecom comments to U.S. Commerce Department", February 20, 1998, <http://www.ntia.doc.gov/ntiahome/domainname/130dftmail/BT.htm>

[CDA] "Reno v. American Civil Liberties Union", 117 S.Ct. 2329, June 26, 1997,

[COPAREPORT] "Final Report of the COPA Commission to the U.S. Congress", October 20, 2000, <http://www.copacommission.org/report/newtopleveldomain.shtml>

[GAO] "GAO Report OGC-00-33R", July 7, 2000, <http://www.gao.gov/new.items/og00033r.pdf>

[GTLD-MOU] "GTLD-MOU Policy Oversight committee RFC 97-02", September 13, 1997, <http://www.gtld-mou.org/docs/notice-97-02.html>

[HOUSEREPORT] "U.S. House Commerce Committee report", 105th Congress, October 5, 1998. <http://www.epic.org/free_speech/censorship/hr3783-report.html>

[ICM-REGISTRY] "Request for reconsideration from ICM Registry to ICANN", December 15, 2000, <http://www.icann.org/committees/reconsideration/icm-request-16dec00.htm>

[LIEBERMAN] "Testimony of Senator Joe Lieberman before Children's Online Protection Act Commission", June 8, 2000, <http://www.senate.gov/~lieberman/press/00/06/2000608958.html>

[RFC 1034] Mockapetris, P., "Domain Names - Concepts and Facilities", STD 13, RFC 1034, November 1987.

[RFC 1715] Huitema, C., "The H Ratio for Address Assignment Efficiency", RFC 1715, November 1994.

[RFC 2396] Berners-Lee, T., Fielding, R. and L. Masinter, "Uniform Resource Identifiers (URI): Generic Syntax", RFC 2396, August 1998.

[RFC 2460] Deering, S. and R. Hinden, "Internet Protocol, Version 6 (IPv6) Specification", RFC 2460, December 1998.

[RFC 2535] Eastlake, 3rd, D., "Domain Name System Security Extensions", RFC 2535, March 1999.

[RFC 2606] Eastlake, 3rd, D. and A. Panitz, "Reserved Top Level DNS Names", BCP 32, RFC 2606, June 1999.

[RFC 2811] Kalt, C., "Internet Relay Chat: Channel Management", RFC 2811, April 2000.

[RFC 2812] Kalt, C., "Internet Relay Chat: Client Protocol", RFC 2812, April 2000.

[RFC 2813] Kalt, C., "Internet Relay Chat: Server Protocol", RFC 2813, April 2000.

[RFC 2854] Connelly, D. and L. Masinter, "The 'text/html' Media Type", RFC 2854, June 2000.

[RFC 3513] Hinden, R. and S. Deering, "Internet Protocol Version 6 (IPv6) Addressing Architecture", RFC 3513, April 2003.

[RFC 3514] Bellovin, S., "The Security Flag in the IPv4 Header", 1 April 2003.

[WARSHAVSKY] Congress weighs Net porn bills," CNET article, February 10, 1998, <http://news.cnet.com/news/0-1005-200-326435.html>

8. Acknowledgement

The contribution and efforts of Declan McCullagh, who wrote substantially all of sections 2 and 3 of this document, are gratefully acknowledged.

9. Authors' Addresses

Donald E. Eastlake 3rd
Motorola Laboratories
155 Beaver Street
Milford, MA 01757 USA

Phone: +1-508-786-7554 (w)
+1-508-634-2066 (h)
EMail: dee3@torque.pothole.com

10. Full Copyright Statement

Acknowledgement

Funding for the RFC Editor function is currently provided by the Internet Society.