Network Working Group
Request for Comments: 1726
Category: Informational

C. Partridge
BBN Systems and Technologies
F. Kastenholz
FTP Software

December 1994

Technical Criteria for Choosing IP The Next Generation (IPng)

Status of this Memo

Abstract

This document was submitted to the IPng Area in response to RFC 1550. Publication of this document does not imply acceptance by the IPng Area of any ideas expressed within. Comments should be submitted to the big-internet@munnari.oz.au mailing list. This RFC specifies criteria related to mobility for consideration in design and selection of the Next Generation of IP.

Table of Contents

1. Introduction

This version of this memo was commissioned by the IPng area of the IETF in order to define a set of criteria to be used in evaluating the protocols being proposed for adoption as the next generation of IP.

The criteria presented here were culled from several sources, including "IP Version 7" [1], "IESG Deliberations on Routing and Addressing" [2], "Towards the Future Internet Architecture" [3], the IPng Requirements BOF held at the Washington D.C. IETF Meeting in December of 1992, the IPng Working Group meeting at the Seattle IETF meeting in March 1994, the discussions held on the Big-Internet mailing list (big-internet@munnari.oz.au, send requests to join to big-internet-request@munnari.oz.au), discussions with the IPng Area Directors and Directorate, and the mailing lists devoted to the individual IPng efforts.

This document presumes that a new IP-layer protocol is actually desired. There is some discussion in the community as to whether we can extend the life of IPv4 for a significant amount of time by

better engineering of, e.g., routing protocols, or we should develop IPng now. This question is not addressed in this document.

We would like to gratefully acknowledge the assistance of literally hundreds of people who shared their views and insights with us. However, this memo is solely the personal opinion of the authors and in no way represents, nor should it be construed as representing, the opinion of the ISOC, the IAB, the IRTF, the IESG, the IETF, the Internet community as a whole, nor the authors' respective employers.

2. Goals

We believe that by developing a list of criteria for evaluating proposals for IP The Next Generation (IPng), the IETF will make it easier for developers of proposals to prioritize their work and efforts and make reasoned choices as to where they should spend relatively more and less time. Furthermore, a list of criteria may help the IETF community determine which proposals are serious contenders for a next generation IP, and which proposals are insufficient to the task. Note that these criteria are probably not sufficient to make final decisions about which proposal is best. Questions such as whether to trade a little performance (e.g., packets per second routed) for slightly more functionality (e.g., more flexible routing) cannot be easily addressed by a simple list of criteria. However, at minimum, we believe that protocols that meet these criteria are capable of serving as the future IPng.

This set of criteria originally began as an ordered list, with the goal of ranking the importance of various criteria. Eventually, the layout evolved into the current form, where each criterion was presented without weighting, but a time frame, indicating approximately when a specific criterion, or feature of a criterion, should be available was added to the specification.

We have attempted to state the criteria in the form of goals or requirements and not demand specific engineering solutions. For example, there has been talk in the community of making route aggregation a requirement. We believe that route aggregation is not, in and of itself, a requirement but rather one part of a solution to the real problem of scaling to some very large, complex topology. Therefore, route aggregation is NOT listed as a requirement; instead, the more general functional goal of having the routing scale is listed instead of the particular mechanism of route aggregation.

In determining the relative timing of the various criteria, we have had two guiding principles. First, IPng must offer an internetwork service akin to that of IPv4, but improved to handle the well-known and widely-understood problems of scaling the Internet architecture

to more end-points and an ever increasing range of bandwidths. Second, it must be desirable for users and network managers to upgrade their equipment to support IPng. At a minimum, this second point implies that there must be a straightforward way to transition systems from IPv4 to IPng. But it also strongly suggests that IPng should offer features that IPv4 does not; new features provide a motivation to deploy IPng more quickly.

3. Note on Terminology

The existing proposals tend distinguish between end-point identification of, e.g., individual hosts, and topological addresses of network attachment points. In this memo we do not make that distinction. We use the term "address" as it is currently used in IPv4; i.e., for both the identification of a particular endpoint or host AND as the topological address of a point on the network. We presume that if the endpoint/ address split remains, the proposals will make the proper distinctions with respect to the criteria enumerated below.

4. General Principles

4.1 Architectural Simplicity

> In anything at all, perfection is finally attained not when there is no longer anything to add, but when there is no longer anything to take away.
>
> Antoine de Saint-Exupery

We believe that many communications functions are more appropriately performed at protocol layers other than the IP layer. We see protocol stacks as hourglass-shaped, with IPng in the middle, or waist, of the hourglass. As such, essentially all higher-layer protocols make use of and rely upon IPng. Similarly IPng, by virtue of its position in the "protocol hourglass" encompasses a wide variety of lower-layer protocols. When IPng does not perform a particular function or provide a certain service, it should not get in the way of the other elements of the protocol stack which may well wish to perform the function.

4.2 One Protocol to Bind Them All

One of the most important aspects of The Internet is that it provides global IP-layer connectivity. The IP layer provides the point of commonality among all of the nodes on the Internet. In effect, the main goal of the Internet is to provide an IP Connectivity Service to all who wish it.

This does NOT say that the Internet is a One-Protocol Internet. The Internet is today, and shall remain in the future, a Multi-Protocol Internet. Multi-Protocol operations are required to allow for continued testing, experimentation, and development and because service providers' customers clearly want to be able to run protocols such as CLNP, DECNET, and Novell over their Internet connections.

4.3 Live Long

It is very difficult to change a protocol as central to the workings of the Internet as IP. Even more problematic is changing such a protocol frequently. This simply can not be done. We believe that it is impossible to expect the community to make significant, non-backward compatible changes to the IP layer more often than once every 10-15 years. In order to be conservative, we strongly urge protocol developers to consider what the Internet will look like in 20 years and design their protocols to fit that vision.

As a data point, the SNMP community has had great difficulty moving from SNMPv1 to SNMPv2. Frequent changes in software are hard.

4.4 Live Long AND Prosper

We believe that simply allowing for bigger addresses and more efficient routing is not enough of a benefit to encourage vendors, service providers, and users to switch to IPng, with its attendant disruptions of service, etc. These problems can be solved much more simply with faster routers, balkanization of the Internet address space, and other hacks.

We believe that there must be positive functional or operational benefits to switching to IPng.

In other words, IPng must be able to live for a long time AND it must allow the Internet to prosper and to grow to serve new applications and user needs.

4.5 Co-operative Anarchy

A major contributor to the Internet's success is the fact that there is no single, centralized, point of control or promulgator of policy for the entire network. This allows individual constituents of the network to tailor their own networks, environments, and policies to suit their own needs. The individual constituents must cooperate only to the degree necessary to ensure that they interoperate.

We believe that this decentralized and decoupled nature of the Internet must be preserved. Only a minimum amount of centralization or forced cooperation will be tolerated by the community as a whole.

We also believe that there are some tangible benefits to this decoupled nature. For example,

* It is easier to experiment with new protocols and services and then roll out intermediate and final results in a controlled fashion.
* By eliminating a single point of control, a single point of failure is also eliminated, making it much less likely that the entire network will fail.
* It allows the administrative tasks for the network to be more widely distributed.

An example of the benefits of this "Cooperative Anarchy" can be seen in the benefits derived from using the Domain Naming System over the original HOSTS.TXT system.

5. Criteria

This section enumerates the criteria against which we suggest the IP The Next Generation proposals be evaluated.

Each criterion is presented in its own section. The first paragraph of each section is a short, one or two sentence statement of the criterion. Additional paragraphs then explain the criterion in more detail, clarify what it does and does not say and provide some indication of its relative importance.

Also, each criterion includes a subsection called "Time Frame". This is intended to give a rough indication of when the authors believe that the particular criterion will become "important". We believe that if an element of technology is significant enough to include in this document then we probably understand the technology enough to predict how important that technology will be. In general, these time frames indicate that, within the desired time frame, we should be able to get an understanding of how the feature will be added to a protocol, perhaps after discussions with the engineers doing the development. Time Frame is not a deployment schedule since deployment schedules depend on non-technical issues, such as vendors determining whether a market exists, users fitting new releases into their systems, and so on.

5.1 Scale

CRITERION

The IPng Protocol must scale to allow the identification and addressing of at least 10**12 end systems (and preferably much more). The IPng Protocol, and its associated routing protocols and architecture must allow for at least 10**9 individual networks (and preferably more). The routing schemes must scale at a rate that is less than the square root of the number of constituent networks [10].

DISCUSSION

The initial, motivating, purpose of the IPng effort is to allow the Internet to grow beyond the size constraints imposed by the current IPv4 addressing and routing technologies.

Both aspects of scaling are important. If we can't route then connecting all these hosts is worthless, but without connected hosts, there's no point in routing, so we must scale in both directions.

In any proposal, particular attention must be paid to describing the routing hierarchy, how the routing and addressing will be organized, how different layers of the routing interact, and the relationship between addressing and routing.

Particular attention must be paid to describing what happens when the size of the network approaches these limits. How are network, forwarding, and routing performance affected? Does performance fall off or does the network simply stop as the limit is neared.

This criterion is the essential problem motivating the transition to IPng. If the proposed protocol does not satisfy this criteria, there is no point in considering it.

We note that one of the white papers solicited for the IPng process [5] indicates that 10**12 end nodes is a reasonable estimate based on the expected number of homes in the world and adding two orders of magnitude for "safety". However, this white paper treats each home in the world as an end-node of a world-wide Internet. We believe that each home in the world will in fact be a network of the world-wide Internet. Therefore, if we take [5]'s derivation of 10**12 as accurate, and change their assumption that a home will be an end-node to a home being a network, we may expect that there will be the need to support at least 10**12 networks, with the possibility of supporting up to 10**15 end-nodes.

Time Frame

Any IPng proposal should be able to show immediately that it has an architecture for the needed routing protocols, addressing schemes, abstraction techniques, algorithms, data structures, and so on that can support growth to the required scales.

Actual development, specification, and deployment of the needed protocols can be deferred until IPng deployment has extended far enough to require such protocols. A proposed IPng should be able to demonstrate ahead of time that it can scale as needed.

5.2 Topological Flexibility

CRITERION

The routing architecture and protocols of IPng must allow for many different network topologies. The routing architecture and protocols must not assume that the network's physical structure is a tree.

DISCUSSION

As the Internet becomes ever more global and ubiquitous, it will develop new and different topologies. We already see cases where the network hierarchy is very "broad" with many subnetworks, each with only a few hosts and where it is very "narrow", with few subnetworks each with many hosts. We can expect these and other topological forms in the future. Furthermore, since we expect that IPng will allow for many more levels of hierarchy than are allowed under IPv4, we can expect very "tall" and very "short" topologies as well.

Constituent organizations of the Internet should be allowed to structure their internal topologies in any manner they see fit. Within reasonable implementation limits, organizations should be allowed to structure their addressing in any manner. We specifically wish to point out that if the network's topology or addressing is hierarchical, constituent organizations should be able to allocate to themselves as many levels of hierarchy as they wish.

It is very possible that the diameter of the Internet will grow to be extremely large; perhaps larger than 256 hops.

Neither the current, nor the future, Internet will be physically structured as a tree, nor can we assume that connectivity can occur only between certain points in the graph. The routing and addressing architectures must allow for multiply connected networks and be able to utilize multiple paths for any reason, including redundancy, load sharing, type- and quality-of-service

differentiation.

Time Frame
We believe that Topological Flexibility is an inherent element of a protocol and therefore should be immediately demonstrable in an IPng proposal.

5.3 Performance

CRITERION
A state of the art, commercial grade router must be able to process and forward IPng traffic at speeds capable of fully utilizing common, commercially available, high-speed media at the time. Furthermore, at a minimum, a host must be able to achieve data transfer rates with IPng comparable to the rates achieved with IPv4, using similar levels of host resources.

DISCUSSION
Network media speeds are constantly increasing. It is essential that the Internet's switching elements (routers) be able to keep up with the media speeds.

We limit this requirement to commercially available routers and media. If some network site can obtain a particular media technology "off the shelf", then it should also be able to obtain the needed routing technology "off the shelf." One can always go into some laboratory or research center and find newer, faster, technologies for network media and for routing. We do believe, however, that IPng should be routable at a speed sufficient to fully utilize the fastest available media, though that might require specially built, custom, devices.

We expect that more and more services will be available over the Internet. It is not unreasonable, therefore, to expect that the ratio of "local" traffic (i.e., the traffic that stays on one's local network) to "export" traffic (i.e., traffic destined to or sourced from a network other than one's own local network) will change, and the percent of export traffic will increase.

We note that the host performance requirement should not be taken to imply that IPng need only be as good as IPv4. If an IPng candidate can achieve better performance with equivalent resources (or equivalent transfer rates with fewer resources) vis-a-vis IPv4 then so much the better. We also observe that many researchers believe that a proper IPng router should be capable of routing IPng traffic over links at speeds that are capable of fully utilizing an ATM switch on the link.

Some developments indicate that the use of very high speed point-to-point connections may become commonplace. In particular, [5] indicates that OC-3 speeds may be widely used in the Cable TV Industry and there may be many OC-3 speed lines connecting to central switching elements.

Processing of the IPng header, and subsequent headers (such as the transport header), can be made more efficient by aligning fields on their natural boundaries and making header lengths integral multiples of typical word lengths (32, 64, and 128 bits have been suggested) in order to preserve alignment in following headers.

We point out that optimizing the header's fields and lengths only to today's processors may not be sufficient for the long term. Processor word and cache-line lengths, and memory widths are constantly increasing. In doing header optimizations, the designer should predict word-widths one or two CPU generations into the future and optimize accordingly. If IPv4 and TCP had been optimized for processors common when they were designed, they would be very efficient for 6502s and Z-80s.

Time Frame

An IPng proposal must provide a plausible argument of how it will scale up in performance. (Obviously no one can completely predict the future, but the idea is to illustrate that if technology trends in processor performance and memory performance continue, and perhaps using techniques like parallelism, there is reason to believe the proposed IPng will scale as technology scales).

5.4 Robust Service

CRITERION

The network service and its associated routing and control protocols must be robust.

DISCUSSION

Murphy's Law applies to networking. Any proposed IPng protocol must be well-behaved in the face of malformed packets, mis-information, and occasional failures of links, routers and hosts. IPng should perform gracefully in response to willful management and configuration mistakes (i.e., service outages should be minimized).

Putting this requirement another way, IPng must make it possible to continue the Internet tradition of being conservative in what is sent, but liberal in what one is willing to receive.

We note that IPv4 is reasonably robust and any proposed IPng must be at least as robust as IPv4.

Hostile attacks on the network layer and Byzantine failure modes must be dealt with in a safe and graceful manner.

We note that Robust Service is, in some form, a part of security and vice-versa.

The detrimental effects of failures, errors, buggy implementations, and misconfigurations must be localized as much as possible. For example, misconfiguring a workstation's IP Address should not break the routing protocols. in the event of misconfigurations, IPng must to be able to detect and at least warn, if not work around, any misconfigurations.

Due to its size, complexity, decentralized administration, error-prone users and administrators, and so on, The Internet is a very hostile environment. If a protocol can not be used in such a hostile environment then it is not suitable for use in the Internet.

Some predictions have been made that, as the Internet grows and as more and more technically less-sophisticated sites get connected, there will be more failures in the network. These failures may be a combination of simple size; if the size of the network goes up by a factor of n, then the total number of failures in the network can be expected to increase by some function of n. Also, as the network's users become less sophisticated, it can be assumed that they will make more, innocent and well meaning, mistakes, either in configuration or use of their systems.

The IPng protocols should be able to continue operating in an environment that suffers more, total, outages than we are currently used to. Similarly, the protocols must protect the general population from errors (either of omission or commission) made by individual users and sites.

Time Frame

We believe that the elements of Robust Service should be available immediately in the protocol with two exceptions.

The security aspects of Robust Service are, in fact, described elsewhere in this document.

Protection against Byzantine failure modes is not needed immediately. A proposed architecture for it should be done immediately. Prototype development should be completed in 12-18 months, with final deployment as needed.

5.5 Transition

CRITERION

The protocol must have a straightforward transition plan from the current IPv4.

DISCUSSION

A smooth, orderly, transition from IPv4 to IPng is needed. If we can't transition to the new protocol, then no matter how wonderful it is, we'll never get to it.

We believe that it is not possible to have a "flag-day" form of transition in which all hosts and routers must change over at once. The size, complexity, and distributed administration of the Internet make such a cutover impossible.

Rather, IPng will need to co-exist with IPv4 for some period of time. There are a number of ways to achieve this co-existence such as requiring hosts to support two stacks, converting between protocols, or using backward compatible extensions to IPv4. Each scheme has its strengths and weaknesses, which have to be weighed.

Furthermore, we note that, in all probability, there will be IPv4 hosts on the Internet effectively forever. IPng must provide mechanisms to allow these hosts to communicate, even after IPng has become the dominant network layer protocol in the Internet.

The absence of a rational and well-defined transition plan is not acceptable. Indeed, the difficulty of running a network that is transitioning from IPv4 to IPng must be minimized. (A good target is that running a mixed IPv4-IPng network should be no more and preferably less difficult than running IPv4 in parallel with existing non-IP protocols).

Furthermore, a network in transition must still be robust. IPng schemes which maximize stability and connectivity in mixed IPv4-IPng networks are preferred.

Finally, IPng is expected to evolve over time and therefore, it must be possible to have multiple versions of IPng, some in production use, some in experimental, developmental, or evaluation use, to coexist on the network. Transition plans must address this issue.

The transition plan must address the following general areas of the Internet's infrastructure:

- Host Protocols and Software
- Router Protocols and Software
- Security and Authentication
- Domain Name System
- Network Management
- Operations Tools (e.g., Ping and Traceroute)
- Operations and Administration procedures

The impact on protocols which use IP addresses as data (e.g., DNS, distributed file systems, SNMP and FTP) must be specifically addressed.

The transition plan should address the issue of cost distribution. That is, it should identify what tasks are required of the service providers, of the end users, of the backbones and so on.

Time Frame

A transition plan is required immediately.

5.6 Media Independence

CRITERION

The protocol must work across an internetwork of many different LAN, MAN, and WAN media, with individual link speeds ranging from a ones-of-bits per second to hundreds of gigabits per second. Multiple-access and point-to-point media must be supported, as must media supporting both switched and permanent circuits.

DISCUSSION

The joy of IP is that it works over just about anything. This generality must be preserved. The ease of adding new technologies, and ability to continue operating with old technologies must be maintained.

We believe this range of speed is right for the next twenty years, though we may wish to require terabit performance at the high-end. We believe that, at a minimum, media running at 500 gigabits per second will be commonly available within 10 years. The low end of the link-speed range is based on the speed of systems like pagers and ELF (ELF connects to submerged submarines and has a "speed" on the order of <10 characters per second).

By switched circuits we mean both "permanent" connections such as X.25 and Frame Relay services AND "temporary" types of dialup connections similar to today's SLIP and dialup PPP services, and

perhaps, ATM SVCs. The latter form of connection implies that dynamic network access (i.e., the ability to unplug a machine, move it to a different point on the network topology, and plug it back in, possibly with a changed IPng address) is required. We note that this is an aspect of mobility.

By work, we mean we have hopes that a stream of IPng datagrams (whether from one source, or many) can come close to filling the link at high speeds, but also scales gracefully to low speeds.

Many network media are multi-protocol. It is essential that IPng be able to peacefully co-exist on such media with other protocols. Routers and hosts must be able to discriminate among the protocols that might be present on such a medium. For example, on an Ethernet, a specific, IPng Ethernet Type value might be called for; or the old IPv4 Ethernet type is used and the first four (version number in the old IPv4 header) bits would distinguish between IPv4 and IPng.

Different media have different MAC address formats and schemes. It must be possible for a node to dynamically determine the MAC address of a node given that node's IP address. We explicitly prohibit using static, manually configured mappings as the standard approach.

Another aspect of this criterion is that many different MTUs will be found in an IPng internetwork. An IPng must be able to operate in such a multi-MTU environment. It must be able to adapt to the MTUs of the physical media over which it operates. Two possible techniques for dealing with this are path MTU discovery and fragmentation and reassembly; other techniques might certainly be developed.

We note that, as of this writing (mid 1994), ATM seems to be set to become a major network media technology. Any IPng should be designed to operate over ATM. However, IPng still must be able to operate over other, more "traditional" network media. Furthermore, a host on an ATM network must be able to interoperate with a host on another, non-ATM, medium, with no more difficulty or complexity than hosts on different media can interoperate today using IPv4.

IPng must be able to deal both with "dumb" media, such as we have today, and newer, more intelligent, media. In particular, IPng functions must be able to exist harmoniously with lower-layer realizations of the same, or similar, functions. Routing and resource management are two areas where designers should pay particular attention. Some subnetwork technologies may include

integral accounting and billing capabilities, and IPng must provide the correct control information to such subnetworks.

Time Frame

Specifications for current media encapsulations (i.e., all encapsulations that are currently Proposed standards, or higher, in the IETF) are required immediately. These specifications must include any auxiliary protocols needed (such as an address resolution mechanism for Ethernet or the link control protocol for PPP). A general 'guide' should also be available immediately to help others develop additional media encapsulations. Other, newer, encapsulations can be developed as the need becomes apparent.

Van Jacobson-like header compression should be shown immediately, as should any other aspects of very-low-speed media. Similarly, any specific aspects of high-speed media should be shown immediately.

5.7 Unreliable Datagram Service

CRITERION

The protocol must support an unreliable datagram delivery service.

DISCUSSION

We like IP's datagram service and it seems to work very well. So we must keep it. In particular, the ability, within IPv4, to send an independent datagram, without prearrangement, is extremely valuable (in fact, may be required for some applications such as SNMP) and must be retained.

Furthermore, the design principle that says that we can take any datagram and throw it away with no warning or other action, or take any router and turn it off with no warning, and have datagram traffic still work, is very powerful. This vastly enhances the robustness of the network and vastly eases administration and maintenance of the network. It also vastly simplifies the design and implementation of software [14].

Furthermore, the Unreliable Datagram Service should support some minimal level of service; something that is approximately equivalent to IPv4 service. This has two functions; it eases the task of IPv4/IPng translating systems in mapping IPv4 traffic to IPng and vice versa, and it simplifies the task of fitting IPng into small, limited environments such as boot ROMs.

Time Frame

Unreliable Datagram Service must be available immediately.

5.8 Configuration, Administration, and Operation

CRITERION

The protocol must permit easy and largely distributed configuration and operation. Automatic configuration of hosts and routers is required.

DISCUSSION

People complain that IP is hard to manage. We cannot plug and play. We must fix that problem.

We do note that fully automated configuration, especially for large, complex networks, is still a topic of research. Our concern is mostly for small and medium sized, less complex, networks; places where the essential knowledge and skills would not be as readily available.

In dealing with this criterion, address assignment and delegation procedures and restrictions should be addressed by the proposal. Furthermore, "ownership" of addresses (e.g., user or service provider) has recently become a concern and the issue should be addressed.

We require that a node be able to dynamically obtain all of its operational, IP-level parameters at boot time via a dynamic configuration mechanism.

A host must be able to dynamically discover routers on the host's local network. Ideally, the information which a host learns via this mechanism would also allow the host to make a rational selection of which first-hop router to send any given packet to. IPng must not mandate that users or administrators manually configure first-hop routers into hosts.

Also, a strength of IPv4 has been its ability to be used on isolated subnets. IPng hosts must be able to work on networks without routers present.

Additional elements of this criterion are:

* Ease of address allocation.
* Ease of changing the topology of the network within a particular routing domain.
* Ease of changing network provider.
* Ease of (re)configuring host/endpoint parameters such as addressing and identification.
* Ease of (re)configuring router parameters such as addressing and identification.

* Address allocation and assignment authority must be delegated as far 'down' the administrative hierarchy as possible.

The requirements of this section apply only to IPng and its supporting protocols (such as for routing, address resolution, and network-layer control). Specifically, as far as IPng is concerned, we are concerned only with how routers and hosts get their configuration information.

We note that in general, automatic configuration of hosts is a large and complex problem and getting the configuration information into hosts and routers is only one, small, piece of the problem. A large amount of additional, non-Internet-layer work is needed in order to be able to do "plug-and-play" networking. Other aspects of "plug-and-play" networking include things like: Autoregistration of new nodes with DNS, configuring security service systems (e.g., Kerberos), setting up email relays and mail servers, locating network resources, adding entries to NFS export files, and so on. To a large degree, these capabilities do not have any dependence on the IPng protocol (other than, perhaps, the format of addresses).

We require that any IPng proposal not impede or prevent, in any way, the development of "plug-and-play" network configuration technologies.

Automatic configuration of network nodes must not prevent users or administrators from also being able to manually configure their systems.

Time Frame

A method for plug and play on small subnets is immediately required.

We believe that this is an extremely critical area for any IPng as a major complaint of the IP community as a whole is the difficulty in administering large IP networks. Furthermore, ease of installation is likely to speed the deployment of IPng.

5.9 Secure Operation

CRITERION

IPng must provide a secure network layer.

DISCUSSION

We need to be sure that we have not created a network that is a cracker's playground.

In order to meet the Robustness criterion, some elements of what is commonly shrugged off as "security" are needed; e.g., to prevent a villain from injecting bogus routing packets, and destroying the routing system within the network. This criterion covers those aspects of security that are not needed to provide the Robustness criterion.

Another aspect of security is non-repudiation of origin. In order to adequately support the expected need for simple accounting, we believe that this is a necessary feature.

In order to safely support requirements of the commercial world, IPng-level security must have capabilities to prevent eavesdroppers from monitoring traffic and deducing traffic patterns. This is particularly important in multi-access networks such as cable TV networks [5].

Aspects of security at the IP level to be considered include:

* Denial of service protections [6],
* Continuity of operations [6],
* Precedence and preemption [6],
* Ability to allow rule-based access control technologies [6]
* Protection of routing and control-protocol operations [9],
* Authentication of routing information exchanges, packets, data, and sources (e.g., make sure that the routing packet came from a router) [9],
* QOS security (i.e., protection against improper use of network-layer resources, functions, and capabilities),
* Auto-configuration protocol operations in that the host must be assured that it is getting its information from proper sources,
* Traffic pattern confidentiality is strongly desired by several communities [9] and [5].

Time Frame

Security should be an integral component of any IPng from the beginning.

5.10 Unique Naming

CRITERION

IPng must assign all IP-Layer objects in the global, ubiquitous, Internet unique names. These names may or may not have any location, topology, or routing significance.

DISCUSSION

We use the term "Name" in this criterion synonymously with the term "End Point Identifier" as used in the NIMROD proposal, or as

IP Addresses uniquely identify interfaces/hosts in IPv4. These names may or may not carry any routing or topology information. See [11] for more discussion on this topic.

IPng must provide identifiers which are suitable for use as globally unique, unambiguous, and ubiquitous names for endpoints, nodes, interfaces, and the like. Every datagram must carry the identifier of both its source and its destination (or some method must be available to determine these identifiers, given a datagram). We believe that this is required in order to support certain accounting functions.

Other functions and uses of unique names are:

* To uniquely identify endpoints (thus if the unique name and address are not the same, the TCP pseudo-header should include the unique name rather than the address)
* To allow endpoints to change topological location on the network (e.g., migrate) without changing their unique names.
* To give one or more unique names to a node on the network (i.e., one node may have multiple unique names)

An identifier must refer to one and only one object while that object is in existence. Furthermore, after an object ceases to exist, the identifier should be kept unused long enough to ensure that any packets containing the identifier have drained out of the Internet system, and that other references to the identifier have probably been lost. We note that the term "existence" is as much an administrative issue as a technical one. For example, if a workstation is reassigned, given a new IP address and node name, and attached to a new subnetwork, is it the same object or not. This does argue for a namespace that is relatively large and relatively stable.

Time Frame

We see this as a fundamental element of the IP layer and it should be in the protocol from the beginning.

5.11 Access

CRITERION

The protocols that define IPng, its associated protocols (similar to ARP and ICMP in IPv4) and the routing protocols (as in OSPF, BGP, and RIP for IPv4) must be published as standards track RFCs and must satisfy the requirements specified in RFC1310. These documents should be as freely available and redistributable as the IPv4 and related RFCs. There must be no specification-related licensing fees for implementing or selling IPng software.

DISCUSSION

An essential aspect of the development of the Internet and its protocols has been the fact that the protocol specifications are freely available to anyone who wishes a copy. Beyond simply minimizing the cost of learning about the technology, the free access to specifications has made it easy for researchers and developers to easily incorporate portions of old protocol specifications in the revised specifications. This type of easy access to the standards documents is required for IPng.

Time Frame

An IPng and its related protocols must meet these standards for openness before an IPng can be approved.

5.12 Multicast

CRITERION

The protocol must support both unicast and multicast packet transmission. Part of the multicast capability is a requirement to be able to send to "all IP hosts on a given subnetwork". Dynamic and automatic routing of multicasts is also required.

DISCUSSION

IPv4 has made heavy use of the ability to multicast requests to all IPv4 hosts on a subnet, especially for autoconfiguration. This ability must be retained in IPng.

Unfortunately, IPv4 currently uses the local media broadcast address to multicast to all IP hosts. This behavior is anti-social in mixed-protocol networks and should be fixed in IPng. There’s no good reason for IPng to send to all hosts on a subnet when it only wishes to send to all IPng hosts. The protocol must make allowances for media that do not support true multicasting.

In the past few years, we have begun to deploy support for wide-area multicast addressing in the Internet, and it has proved valuable. This capability must not be lost in the transition to IPng.

The ability to restrict the range of a multicast to specific networks is also important. Furthermore, it must be possible to "selectively" multicast packets. That is, it must be possible to send a multicast to a remote, specific portion or area of the Internet (such as a specific network or subnetwork) and then have that multicast limited to just that specific area. Furthermore, any given network or subnetwork should be capable of supporting 2**16 "local" multicast groups, i.e., groups that are not propagated to other networks. See [8].

It should be noted that addressing -- specifically the syntax and semantics of addresses -- has a great impact on the scalability of the architecture.

Currently, large-scale multicasts are routed manually through the Internet. While this is fine for experiments, a "production" system requires that multicast-routing be dynamic and automatic. Multicast groups must be able to be created and destroyed, hosts must be able to join and leave multicast groups and the network routing infrastructure must be able to locate new multicast groups and destinations and route traffic to those destinations all without manual intervention.

Large, topologically dispersed, multicast groups (with up to 10**6 participants) must be supported. Some applications are given in [8].

Time Frame

Obviously, address formats, algorithms for processing and interpreting the multicast addresses must be immediately available in IPng. Broadcast and Multicast transmission/reception of packets are required immediately. Dynamic routing of multicast packets must be available within 18 months.

We believe that Multicast Addressing is vital to support future applications such as remote conferencing. It is also used quite heavily in the current Internet for things like service location and routing.

5.13 Extensibility

CRITERION

The protocol must be extensible; it must be able to evolve to meet the future service needs of the Internet. This evolution must be achievable without requiring network-wide software upgrades. IPng is expected to evolve over time. As it evolves, it must be able to allow different versions to coexist on the same network.

DISCUSSION

We do not today know all of the things that we will want the Internet to be able to do 10 years from now. At the same time, it is not reasonable to ask users to transition to a new protocol with each passing decade. Thus, we believe that it must be possible to extend IPng to support new services and facilities. Furthermore, it is essential that any extensions can be incrementally deployed to only those systems which desire to use them. Systems upgraded in this fashion must still be able to communicate with systems which have not been so upgraded.

There are several aspects to extensibility:

5.13.1 Algorithms
The algorithms used in processing IPng information should be decoupled from the protocol itself. It should be possible to change these algorithms without necessarily requiring protocol, datastructure, or header changes.

5.13.2 Headers
The content of packet headers should be extensible. As more features and functions are required of IPng, it may be necessary to add more information to the IPng headers. We note that for IPv4, the use of options has proven less than entirely satisfactory since options have tended to be inefficient to process.

5.13.3 Data Structures
The fundamental data structures of IPng should not be bound with the other elements of the protocol. E.g., things like address formats should not be intimately tied with the routing and forwarding algorithms in the way that the IPv4 address class mechanism was tied to IPv4 routing and forwarding.

5.13.4 Packets
It should be possible to add additional packet-types to IPng. These could be for, e.g., new control and/or monitoring operations.

We note that, everything else being equal, having larger, oversized, number spaces is preferable to having number spaces that are "just large enough". Larger spaces afford more flexibility on the part of network designers and operators and allow for further experimentation on the part of the scientists, engineers, and developers. See [7].

Time Frame
A framework showing mechanisms for extending the protocol must be provided immediately.

5.14 Network Service

CRITERION
The protocol must allow the network (routers, intelligent media, hosts, and so on) to associate packets with particular service classes and provide them with the services specified by those classes.

DISCUSSION

For many reasons, such as accounting, security and multimedia, it is desirable to treat different packets differently in the network.

For example, multimedia is now on our desktop and will be an essential part of future networking. So we have to find ways to support it; and a failure to support it may mean users choose to use protocols other than IPng.

The IETF multicasts have shown that we can currently support multimedia over internetworks with some hitches. If the network can be guaranteed to provide the necessary service levels for this traffic, we will dramatically increase its success.

This criterion includes features such as policy-based routing, flows, resource reservation, network service technologies, type-of-service and quality-of-service and so on.

In order to properly support commercial provision and use of Internetwork service, and account for the use of these services (i.e., support the economic principle of "value paid for value received") it must be possible to obtain guarantees of service levels. Similarly, if the network can not support a previously guaranteed service level, it must report this to those to whom it guaranteed the service.

Network service provisions must be secure. The network-layer security must generally prevent one host from surreptitiously obtaining or disrupting the use of resources which another host has validly acquired. (Some security failures are acceptable, but the failure rate must be very low and the rate should be quantifiable).

One of the parameters of network service that may be requested must be cost-based.

As far as possible, given the limitations of underlying media and IP’s model of a robust internet datagram service, real-time, mission-critical applications must be supported by IPng [6].

Users must be able to confirm that they are, in fact, getting the services that they have requested.

Time Frame

This should be available within 24 months.

5.15 Support for Mobility

CRITERION

The protocol must support mobile hosts, networks and internetworks.

DISCUSSION

Again, mobility is becoming increasingly important. Look at the portables that everyone is carrying. Note the strength of the Apple commercial showing someone automatically connecting up her Powerbook to her computer back in the office. There have been a number of pilot projects showing ways to support mobility in IPv4. All have some drawbacks. But like network service grades, if we can support mobility, IPng will have features that will encourage transition.

We use an encompassing definition of "mobility" here. Mobility typically means one of two things to people: 1) Hosts that physically move and remain connected (via some wireless datalink) with sessions and transport-layer connections remaining ’open’ or ’active’ and 2) Disconnecting a host from one spot in the network, connecting it back in another arbitrary spot and continuing to work. Both forms are required.

Reference [6] discusses possible future use of IP-based networks in the US Navy’s ships, planes, and shore installations. Their basic model is that each ship, plane and shore installation represents at least one IP network. The ship- and plane-based networks, obviously, are mobile as these craft move around the world. Furthermore, most, if not all, Naval surface combatants carry some aircraft (at a minimum, a helicopter or two). So, not only must there be mobile networks (the ships that move around), but there must be mobile internetworks: the ships carrying the aircraft where each aircraft has its own network, which is connected to the ship’s network and the whole thing is moving.

There is also the requirement for dynamic mobility; a plane might take off from aircraft carrier A and land on carrier B so it obviously would want to "connect" to B’s network. This situation might be even more complex since the plane might wish to retain connectivity to its "home" network; that is, the plane might remain connected to the ship-borne networks of both aircraft carriers, A and B.

These requirements are not limited to just the navy. They apply to the civilian and commercial worlds as well. For example, in civil airliners, commercial cargo and passenger ships, trains, cars and so on.

Time Frame
The mobility algorithms are stabilizing and we would hope to see an IPng mobility framework within a year.

5.16 Control Protocol

CRITERION
The protocol must include elementary support for testing and debugging networks.

DISCUSSION
An important feature of IPv4 is the ICMP and its debugging, support, and control features. Specific ICMP messages that have proven extraordinarily useful within IPv4 are Echo Request/Reply (a.k.a ping), Destination Unreachable and Redirect. Functions similar to these should be in IPng.

This criterion explicitly does not concern itself with configuration related messages of ICMP. We believe that these are adequately covered by the configuration criterion in this memo.

One limitation of today's ICMP that should be fixed in IPng's control protocol is that more than just the IPng header plus 64 bits of a failed datagram should be returned in the error message. In some situations, this is too little to carry all the critical protocol information that indicates why a datagram failed. At minimum, any IPng control protocol should return the entire IPng and transport headers (including options or nested headers).

Time Frame
Support for these functions is required immediately.

5.17 Private Networks

CRITERION
IPng must allow users to build private internetworks on top of the basic Internet Infrastructure. Both private IP-based internetworks and private non-IP-based (e.g., CLNP or AppleTalk) internetworks must be supported.

DISCUSSION
In the current Internet, these capabilities are used by the research community to develop new IP services and capabilities (e.g., the MBone) and by users to interconnect non-IP islands over the Internet (e.g., CLNP and DecNet use in the UK).

The capability of building networks on top of the Internet have been shown to be useful. Private networks allow the Internet to

be extended and modified in ways that 1) were not foreseen by the original builders and 2) do not disrupt the day-to-day operations of other users.

We note that, today in the IPv4 Internet, tunneling is widely used to provide these capabilities.

Finally, we note that there might not be any features that specifically need to be added to IPng in order to support the desired functions (i.e., one might treat a private network protocol simply as another IP client protocol, just like TCP or UDP). If this is the case, then IPng must not prevent these functions from being performed.

Time Frame

Some of these capabilities may be required to support other criteria (e.g., transition) and as such, the timing of the specifications is governed by the other criteria (e.g., immediately in the case of transition). Others may be produced as desired.

6. Things We Chose Not to Require

This section contains items which we felt should not impact the choice of an IPng. Listing an item here does not mean that a protocol MUST NOT do something. It means that the authors do not believe that it matters whether the feature is in the protocol or not. If a protocol includes one of the items listed here, that's cool. If it doesn't; that's cool too. A feature might be necessary in order to meet some other criterion. Our point is merely that the feature need not be required for its own sake.

6.1 Fragmentation

The technology exists for path MTU discovery. Presumably, IPng will continue to provide this technology. Therefore, we believe that IPng Fragmentation and Reassembly, as provided in IPv4, is not necessary. We note that fragmentation has been shown to be detrimental to network performance and strongly recommend that it be avoided.

6.2 IP Header Checksum

There has been discussion indicating that the IP Checksum does not provide enough error protection to warrant its performance impact. The argument states that there is almost always a stronger datalink level CRC, and that end-to-end protection is provided by the TCP checksum. Therefore we believe that an IPng checksum is not required per-se.

6.3 Firewalls

Some have requested that IPng include support for firewalls. The authors believe that firewalls are one particular solution to the problem of security and, as such, do not consider that support for firewalls is a valid requirement for IPng. (At the same time, we would hope that no IPng is hostile to firewalls without offering some equivalent security solution).

6.4 Network Management

Network Management properly is a task to be carried out by additional protocols and standards, such as SNMP and its MIBs. We believe that network management, per se, is not an attribute of the IPng protocol. Furthermore, network management is viewed as a support, or service, function. Network management should be developed to fit IPng and not the other way round.

6.5 Accounting

We believe that accounting, like network management, must be designed to fit the IPng protocol, and not the other way round. Therefore, accounting, in and of itself, is not a requirement of IPng. However, there are some facets of the protocol that have been specified to make accounting easier, such as non-repudiation of origin under security, and the unique naming requirement for sorting datagrams into classes. Note that a parameter of network service that IPng must support is cost.

6.6 Routing

Routing is a very critical part of the Internet. In fact, the Internet Engineering Task Force has a separate Area which is chartered to deal only with routing issues. This Area is separate from the more general Internet Area.

We see that routing is also a critical component of IPng. There are several criteria, such as Scaling, Addressing, and Network Services, which are intimately entwined with routing. In order to stress the critical nature and importance of routing, we have chosen to devote a separate chapter to specifically enumerating some of the requirements and issues that IPng routing must address. All of these issues, we believe, fall out of the general criteria presented in the previous chapter.

6.6.1 Scale

First and foremost, the routing architecture must scale to support a very large Internet. Current expectations are for an Internet of about 10**9 to 10**12 networks. The routing architecture must be able to deal with networks of this size. Furthermore, the routing architecture must be able to deal with this size without requiring massive, global databases and algorithms. Such databases or algorithms would, in effect, be single points of failure in the architecture (which is not robust), and because of the nature of Internet administration (cooperative anarchy), it would be impossible to maintain the needed consistency.

6.6.2 Policy

Networks (both transit and non-transit) must be able to set their own policies for the types of traffic that they will admit. The routing architecture must make these policies available to the network as a whole. Furthermore, nodes must be able to select routes for their traffic based on the advertised policies.

6.6.3 QOS

A key element of the network service criteria is that differing applications wish to acquire differing grades of network service. It is essential that this service information be propagated around the network.

6.6.4 Feedback

As users select specific routes over which to send their traffic, they must be provided feedback from the routing architecture. This feedback should allow the user to determine whether the desired routes are actually available or not, whether the desired services are being provided, and so forth.

This would allow users to modify their service requirements or even change their routes, as needed.

6.6.5 Stability

With the addition of additional data into the routing system (i.e., routes are based not only on connectivity, as in IPv4, but also on policies, service grades, and so on), the stability of the routes may suffer. We offer as evidence the early ARPANET which experimented with load-based routing. Routes would remain in flux, changing from one saturated link, to another, unused, link.

This must not be allowed to happen. If anything, routes should be even more stable under IPng's routing architecture than under the current architecture.

6.6.6 Multicast

Multicast will be more important in IPng than it is today in IPv4. Multicast groups may be very large and very distributed. Membership in multicast groups will be very dynamic. The routing architecture must be able to cope with this.

Furthermore, the routing architecture must be able to build multicast routes dynamically, based on factors such as group membership, member location, requested and available qualities of service, and so on.

7. References

[1] Internet Architecture Board, "IP Version 7", Draft 8, Work in Progress, July, 1992.

[2] Gross, P., and P. Almquist, "IESG Deliberations on Routing and Addressing", RFC 1380, IESG Chair, IESG Internet AD, November 1992.

[3] Clark, D., Chapin, L., Cerf, V., Braden, R., and R. Hobby, "Toward the Future Internet Architecture", RFC 1287, MIT, BBN, CNRI, USC/Information Sciences Institute, UC Davis, December 1991.

[4] Dave Clark's paper at SIGCOMM '88 where he pointed out that the design of TCP/IP was guided, in large part, by an ordered list of requirements.

[5] Vecchi, M., "IPng Requirements: A Cable Television Industry Viewpoint", RFC 1686, Time Warner Cable, August 1994.

[6] Green, D., Irey, P., Marlow, D. and K. O'Donoghue, "HPN Working Group Input to the IPng Requirements Solicitation, RFC 1679, NSWC-DD, August 1994.

[7] Bellovin, S., "On Many Addresses per Host", RFC 1681, AT&T Bell Laboratories, August 1994.

[8] Symington, S., Wood, D., and J. Pullen, "Modelling and Simulation Requirements for IPng", RFC 1667, Mitre Corporation and George Mason University, August 1994.

[9] Internet Architecture Board, "Report of the IAB Workshop on Security in the Internet Architecture, RFC 1636, IAB, June 1994.

[10] Private EMAIL from Tony Li to IPNG Directorate Mailing List, 18 April 1994 18:42:05.

[11] Saltzer, J., On the Naming and Binding of Network Destinations", RFC 1498, M.I.T. Laboratory for Computer Science, August 1993.

[12] Postel, J., "Transmission Control Protocol - DARPA Internet Program Protocol Specification", STD 7, RFC 793, DARPA, September 1981.

[13] EMAIL from Robert Elz to the Big Internet mailing list, approximately 4 May 1994.

[14] Chiappa, N., "Nimrod and IPng Technical Requirements", Work in Progress.

8. Security Considerations

Security is not directly addressed by this memo. However, as this memo codifies goals for a new generation of network layer protocol, the security provided by such a protocol is addressed. Security has been raised as an issue in several of the requirements stated in this memo. Furthermore, a specific requirement for security has been made.

9. Acknowledgements

The authors gratefully acknowledge the assistance and input provided by the many people who have reviewed and commented upon this document.

10. Authors' Addresses

Craig Partridge
BBN Systems and Technologies
10 Moulton St.
Cambridge, MA 02138

EMail: craig@aland.bbn.com

Frank Kastenholz
FTP Software, Inc.
2 High St.
North Andover, MA, 01845-2620 USA

EMail: kasten@ftp.com