Network Working Group
Request for Comments: 1347

Ross Callon
DEC

June 1992

TCP and UDP with Bigger Addresses (TUBA), A Simple Proposal for Internet Addressing and Routing

Status of the Memo

This memo provides information for the Internet community. It does not specify an Internet standard. Distribution of this memo is unlimited.

1 Summary

The Internet is approaching a situation in which the current IP address space is no longer adequate for global addressing and routing. This is causing problems including: (i) Internet backbones and regionals are suffering from the need to maintain large amounts of routing information which is growing rapidly in size (approximately doubling each year); (ii) The Internet is running out of IP network numbers to assign. There is an urgent need to develop and deploy an approach to addressing and routing which solves these problems and allows scaling to several orders of magnitude larger than the existing Internet. However, it is necessary for any change to be deployed in an incremental manner, allowing graceful transition from the current Internet without disruption of service. [1]

This paper describes a simple proposal which provides a long-term solution to Internet addressing, routing, and scaling. This involves a gradual migration from the current Internet Suite (which is based on Internet applications, running over TCP or UDP, running over IP) to an updated suite (based on the same Internet applications, running over TCP or UDP, running over CLNP [2]). This approach is known as "TUBA" (TCP & UDP with Bigger Addresses).

This paper describes a proposal for how transition may be accomplished. Description of the manner in which use of CLNP, NSAP addresses, and related network/Internet layer protocols (ES-IS, IS-IS, and IDRP) allow scaling to a very large ubiquitous worldwide Internet is outside of the scope of this paper.

Originally, it was thought that any practical proposal needed to address the immediate short-term problem of routing information explosion (in addition to the long-term problem of scaling to a worldwide Internet). Given the current problems caused by excessive routing information in IP backbones, this could require older IP-based systems to talk to other older IP-based systems over intervening Internet backbones which did not support IP. This in turn would require either translation of IP packets into CLNP packets and vice versa, or encapsulation of IP packets inside CLNP packets. However, other shorter-term techniques (for example [3]) have been proposed which will allow the Internet to operate successfully for several years using the current IP address space. This in turn allows more time for IP-to-CLNP migration, which in turn allows for a much simpler migration technique.

The TUBA proposal therefore makes use of a simple long-term migration proposal based on a gradual update of Internet Hosts (to run Internet applications over CLNP) and DNS servers (to return larger addresses). This proposal requires routers to be updated to support forwarding of CLNP (in addition to IP). However, this proposal does **not** require encapsulation **nor** translation of packets **nor** address mapping. IP addresses and NSAP addresses may be assigned and used independently during the migration period. Routing and forwarding of IP and CLNP packets may be done independently.

This paper provides a draft overview of TUBA. The detailed operation of TUBA has been left for further study.

2 Long-Term Goal of TUBA

This proposal seeks to take advantage of the success of the Internet Suite, the greatest part of which is probably the use of IP itself. IP offers a ubiquitous network service, based on datagram (connectionless) operation, and on globally significant IP addresses which are structured to aid routing. Unfortunately, the limited 32-bit IP address is gradually becoming inadequate for routing and addressing in a global Internet. Scaling to the anticipated future size of the worldwide Internet requires much larger addresses allowing a multi-level hierarchical address assignment.

If we had the luxury of starting over from scratch, most likely we would base the Internet on a new datagram internet protocol with much larger multi-level addresses. In principle, there are many choices available for a new datagram internet protocol. For example, the current IP could be augmented by addition of larger addresses, or a new protocol could be developed. However, the development, standardization, implementation, testing, debugging and deployment of a new protocol (as well as associated routing and host-to-router protocols) would take a very large amount of time and energy, and is not guaranteed to lead to success. In addition, there is already such a protocol available. In particular, the ConnectionLess Network Protocol (CLNP [1]) is very similar to IP, and offers the required datagram service and address flexibility. CLNP is currently being deployed in the Internet backbones and regionals, and is available in vendor products. This proposal does not actually require use of CLNP (the main content of this proposal is a graceful migration path from the current IP to a new IP offering a larger address space), but use of CLNP will be assumed.

This proposal seeks to minimize the risk associated with migration to a new IP address space. In addition, this proposal is motivated by the requirement to allow the Internet to scale, which implies use of Internet applications in a **very** large ubiquitous worldwide Internet. It is therefore proposed that existing Internet transport and application protocols continue to operate unchanged, except for the replacement of 32-bit IP addresses with larger addresses. The use of larger addresses will have some effect on applications, particularly on the Domain Name Service. TUBA does not mean having to move over to OSI completely. It would mean only replacing IP with CLNP. TCP, UDP, and the traditional TCP/IP applications would run on top of CLNP.

The long term goal of the TUBA proposal involves transition to a worldwide Internet which operates much as the current Internet, but with CLNP replacing IP and with NSAP addresses replacing IP addresses. Operation of this updated protocol suite will be very similar to the current operation. For example, in order to initiate communication with another host, a host will obtain a internet address in the same manner that it normally does, except that the address would be larger. In many or most cases, this implies that the host would contact the DNS server, obtain a mapping from the known DNS name to an internet address, and send application packets encapsulated in TCP or UDP, which are in turn encapsulated in CLNP. This long term goal requires a specification for how TCP and UDP are run over CLNP. Similarly, DNS servers need to be updated to deal with NSAP addresses, and routers need to be updated to forward CLNP packets. This proposal does not involve any wider-spread migration to OSI protocols.

TUBA does not actually depend upon DNS for its operation. Any method that is used for obtaining Internet addresses may be updated to be able to return larger (NSAP) addresses, and then can be used with TUBA.

3 Migration

Figure 1 illustrates the basic operation of TUBA. Illustrated is a single Internet Routing Domain, which is also interconnected with Internet backbones and/or regionals. Illustrated are two "updated" Internet Hosts N1 and N2, as well as two older hosts H1 and H2, plus a DNS server and two border routers. It is assumed that the routers internal to the routing domain are capable of forwarding both IP and CLNP traffic (this could be done either by using multi-protocol routers which can forward both protocol suites, or by using a different set of routers for each suite).

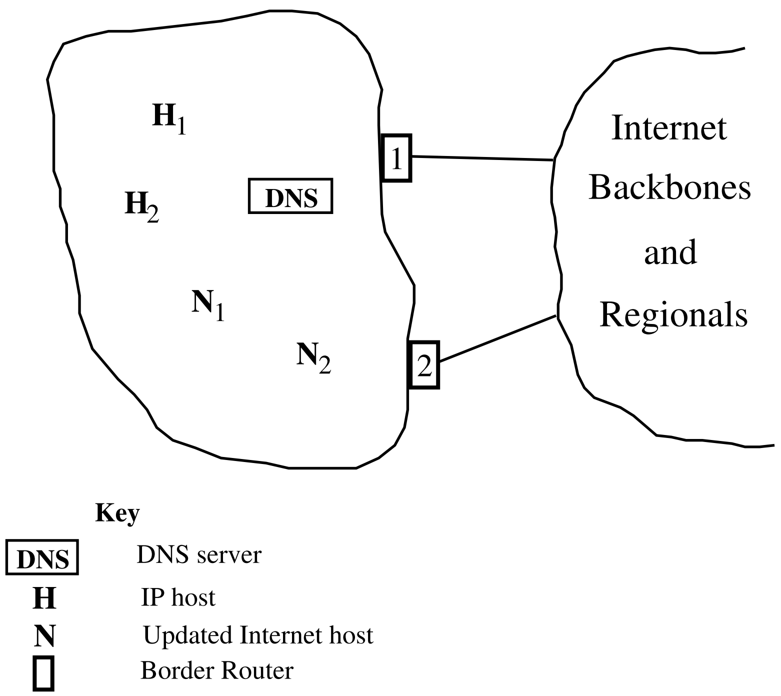

Figure 1 - Overview of TUBA

Updated Internet hosts talk to old Internet hosts using the current Internet suite unchanged. Updated Internet hosts talk to other updated Internet hosts using (TCP or UDP over) CLNP. This implies that updated Internet hosts must be able to send either old-style packets (using IP), or new style packet (using CLNP). Which to send is determined via the normal name-to-address lookup.

Thus, suppose that host N1 wants to communicate with host H1. In this case, N1 asks its local DNS server for the address associated with H1. In this case, since H1 is a older (not-updated) host, the address available for H1 is an IP address, and thus the DNS response returned to N1 specifies an IP address. This allows N1 to know that it needs to send a normal old-style Internet suite packet (encapsulated in IP) to H1.

Suppose that host N1 wants to communicate with host N2. In this case, again N1 contacts the DNS server. If the routers in the domain have not been updated (to forward CLNP), or if the DNS resource record for N2 has not been updated, then the DNS server will respond with a normal IP address, and the communication between N1 and N2 will use IP (updated hosts in environments where the local routers do not handle CLNP are discussed in section 6.3). However, assuming that the routers in the domain have been updated (to forward CLNP), that the DNS server has been updated (to be able to return NSAP addresses), and that the appropriate resource records for NSAP addresses have been configured into the DNS server, then the DNS server will respond to N1 with the NSAP address for N2, allowing N1 to know to use CLNP (instead of IP) for communication with N2.

A new resource record type will be defined for NSAP addresses. New hosts ask for both the new and old (IP address) resource records. Older DNS servers will not have the new resource record type, and will therefore respond with only IP address information. Updated DNS servers will have the new resource record information for the requested DNS name only if the associated host has been updated (otherwise the updated DNS server again will respond with an IP address).

Hosts and/or applications which do not use DNS operate in a similar method. For example, suppose that local name to address records are maintained in host table entries on each local workstation. When a workstation is updated to be able to run Internet applications over CLNP, then the host table on the host may also be updated to contain updated NSAP addresses for other hosts which have also been updated. The associated entries for non-updated hosts would continue to contain IP addresses. Thus, again when an updated host wants to initiate communication with another host, it would look up the associated Internet address in the normal manner. If the address returned is a normal 32-bit IP address, then the host would initiate a request using an Internet application over TCP (or UDP) over IP (as at present). If the returned address is a longer NSAP address, then the host would initiate a request using an Internet application over TCP (or UDP) over CLNP.

4 Running TCP and UDP Over CLNP

TCP is run directly on top of CLNP (i.e., the TCP packet is encapsulated directly inside a CLNP packet — the TCP header occurs directly following the CLNP header). Use of TCP over CLNP is straightforward, with the only non-trivial issue being how to generate the TCP pseudo-header (for use in generating the TCP checksum).

Note that TUBA runs TCP over CLNP on an end-to-end basis (for example, there is no intention to translate CLNP packets into IP packets). This implies that only "consenting updated systems" will be running TCP over CLNP; which in turn implies that, for purposes of generating the TCP pseudoheader from the CLNP header, backward compatibility with existing systems is not an issue. There are therefore several options available for how to generate the pseudoheader. The pseudoheader could be set to all zeros (implying that the TCP header checksum would only be covering the TCP header). Alternatively, the pseudoheader could be calculated from the CLNP header. For example, the "source address" in the TCP pseudoheader could be replaced with two bytes of zero plus a two byte checksum run on the source NSAP address length and address (and similarly for the destination address); the "protocol" could be replaced by the destination address selector value; and the "TCP Length" could be calculated from the CLNP packet in the same manner that it is currently calculated from the IP packet. The details of how the pseudoheader is composed is for further study.

UDP is transmitted over CLNP in the same manner. In particular, the UDP packet is encapsulated directly inside a CLNP packet. Similarly, the same options are available for the UDP pseudoheader as for the TCP pseudoheader.

5 Updates to the Domain Name Service

TUBA requires that a new DNS resource record entry type ("long-address") be defined, to store longer Internet (i.e., NSAP) addresses. This resource record allows mapping from DNS names to NSAP addresses, and will contain entries for systems which are able to run Internet applications, over TCP or UDP, over CLNP.

The presence of a "long-address" resource record for mapping a particular DNS name to a particular NSAP address can be used to imply that the associated system is an updated Internet host. This specifically does **not** imply that the system is capable of running OSI protocols for any other purpose. Also, the NSAP address used for running Internet applications (over TCP or UDP over CLNP) does not need to have any relationship with other NSAP addresses which may be assigned to the same host. For example, a "dual stack" host may be able to run Internet applications over TCP over CLNP, and may also be able to run OSI applications over TP4 over CLNP. Such a host may have a single NSAP address assigned (which is used for both purposes), or may have separate NSAP addresses assigned for the two protocol stacks. The "long-address" resource record, if present, may be assumed to contain the correct NSAP address for running Internet applications over CLNP, but may not be assumed to contain the correct NSAP address for any other purpose.

The backward translation (from NSAP address to DNS name) is facilitated by definition of an associated resource record. This resource record is known as "long-in-addr.arpa", and is used in a manner analogous to the existing "in-addr.arpa".

Updated Internet hosts, when initiating communication with another host, need to know whether that host has been updated. The host will request the address-class "internet address", entry-type "long-address" from its local DNS server. If the local DNS server has not yet been updated, then the long address resource record will not be available, and an error response will be returned. In this case, the updated hosts must then ask for the regular Internet address. This allows updated hosts to be deployed in environments in which the DNS servers have not yet been updated.

An updated DNS server, if asked for the long-address corresponding to a particular DNS name, does a normal DNS search to obtain the information. If the long-address corresponding to that name is not available, then the updated DNS server will return the resource record type containing the normal 32-bit IP address (if available). This allows more efficient operation between updated hosts and old hosts in an environment in which the DNS servers have been updated.

Interactions between DNS servers can be done over either IP or CLNP, in a manner analogous to interactions between hosts. DNS servers currently maintain entries in their databases which allow them to find IP addresses of other DNS servers. These can be updated to include a combination of IP addresses and NSAP addresses of other servers. If an NSAP address is available, then the communication with the other DNS server can use CLNP, otherwise the interaction between DNS servers uses IP. Initially, it is likely that all communication between DNS servers will use IP (as at present). During the migration process, the DNS servers can be updated to communicate with each other using CLNP.

6 Other Technical Details

6.1 When 32-Bit IP Addresses Fail

Eventually, the IP address space will become inadequate for global routing and addressing. At this point, the remaining older (not yet updated) IP hosts will not be able to interoperate directly over the global Internet. This time can be postponed by careful allocation of IP addresses and use of "Classless Inter-Domain Routing" (CIDR [3]), and if necessary by encapsulation (either of IP in IP, or IP in CLNP). In addition, the number of hosts affected by this can be minimized by aggressive deployment of updated software based on TUBA.

When the IP address space becomes inadequate for global routing and addressing, for purposes of IP addressing the Internet will need to be split into "IP address domains". 32-bit IP addresses will be meaningful only within an address domain, allowing the old IP hosts to continue to be used locally. For communications between domains, there are two possibilities: (i) The user at an old system can use application layer relays (such as mail relays for 822 mail, or by Telnetting to an updated system in order to allow Telnet or FTP to a remote system in another domain); or (ii) Network Address Translation can be used [4].

6.2 Applications which use IP Addresses Internally

There are some application protocols (such as FTP and NFS) which pass around and use IP addresses internally. Migration to a larger address space (whether based on CLNP or other protocol) will require either that these applications be limited to local use (within an "IP address domain" in which 32-bit IP addresses are meaningful) or be updated to either: (i) Use larger network addresses instead of 32-bit IP addresses; or (ii) Use some other globally-significant identifiers, such as DNS names.

6.3 Updated Hosts in IP-Only Environments

There may be some updated Internet hosts which are deployed in networks that do not yet have CLNP service, or where CLNP service is available locally, but not to the global Internet. In these cases, it will be necessary for the updated Internet hosts to know to initially send all Internet traffic (or all non-local traffic) using IP, even when the remote system also has been updated. There are several ways that this can be accomplished, such as: (i) The host could contains a manual configuration parameter controlling whether to always use IP, or to use IP or CLNP depending upon remote address; (ii) The DNS resolver on the host could be "lied to" to believe that all remote requests are supposed to go to some particular server, and that server could intervene and change all remote requests for long-addresses into requests for normal IP addresses.

6.4 Local Network Address Translation

Network Address Translation (NAT [4]) has been proposed as a means to allow global communication between hosts which use locally-significant IP addresses. NAT requires that IP addresses be mapped at address domain boundaries, either to globally significant addresses, or to local addresses meaningful in the next address domain along the packet's path. It is possible to define a version of NAT which is "local" to an addressing domain, in the sense that (locally significant) IP packets are mapped to globally significant CLNP packets before exiting a domain, in a manner which is transparent to systems outside of the domain.

NAT allows old systems to continue to be used globally without application gateways, at the cost of significant additional complexity and possibly performance costs (associated with translation or encapsulation of network packets at IP address domain boundaries). NAT does not address the problem of applications which pass around and use IP addresses internally.

The details of Network Address Translation is outside of the scope of this document.

6.5 Streamlining Operation of CLNP

CLNP contains a number of optional and/or variable length fields. For example, CLNP allows addresses to be any integral number of bytes up to 20 bytes in length. It is proposed to "profile" CLNP in order to allow streamlining of router operation. For example, this might involve specifying that all Internet hosts will use an NSAP address of precisely 20 bytes in length, and may specify which optional fields (if any) will be present in all CLNP packets. This can allow all CLNP packets transmitted by Internet hosts to use a constant header format, in order to speed up header parsing in routers. The details of the Internet CLNP profile is for further study.

7 References

[1] "The IAB Routing and Addressing Task Force: Summary Report", work in progress.

[2] "Protocol for Providing the Connectionless-Mode Network Service", ISO 8473, 1988.

[3] "Supernetting: An Address Assignment and Aggregation Strategy", V.Fuller, T.Li, J.Yu, and K.Varadhan, March 1992.

[4] "Extending the IP Internet Through Address Reuse", Paul Tsuchiya, December 1991.

8 Security Considerations

Security issues are not discussed in this memo.

9 Author's Address

Ross Callon
Digital Equipment Corporation
550 King Street, LKG 1-2/A19
Littleton, MA 01460-1289

Phone: 508-486-5009

Email: Callon@bigfut.lkg.dec.com