1 Why a new Internet Protocol?

The Internet Protocol (IP) ([1]), forms the common protocol foundation for the world-wide Internet. By definition, those hosts that can communicate directly via IP packets are said to be part of the *core* Internet, as opposed to those who are only reachable by electronic mail, say. Just about any data link layer has been specified to run over IP, ranging in speed from HIPPI and ATM to Ethernet, ISDN, modems and carrier pigeons. In turn, a number of transport protocols can run over IP. IP and the stack below is the only data protocol visible to routers inside the Internet. Besides its use within the global, interconnected Internet (capital "I"), it is also used in a large number of private internets (lower-case "i'). Since the original Berkeley Unix operating system, almost all workstations have been equipped with it as part of the manufacturer-supplied operating system. PCs have tended to use Novell's proprietary IPX instead, but this appears to be changing primarily due to the desire to access Internet information resources.

The internet protocol is a datagram protocol, with each packet being independently routed from source to destination. This is also the mode of operation of most local area networks, in contrast to connection-oriented networks like X.25, frame relay or ATM. Each packet needs to be self-contained; in particular, it needs to contain a globally unique address.

2 A bit of history

ARPAnet, the progenitor of today's Internet, had only a dozen or so sites. Thus, the original designers considered a 32-bit address, with its ability to enumerate around 4 billion hosts (end systems or computers), as more than sufficient. Like today's Internet, the early Internet connected independently administered local area networks (LANs). Thus, an IP address consists of a network part, enumerating these local networks, and a host part, designating the individual computer within a LAN. Routers outside the LAN only need to maintain topology information about the networks, rather than each individual host. Since these LANs differed in size, the 32-bit address space was divided into three classes, with up to 128 networks with 16 million hosts (class A), 16,384 networks with 65,536 hosts each (class B) and 2 million networks with 256 hosts each (class C). These addresses were assigned permanently on a first-come, first-serve basis, without regard for network topology. Class B proved to be the most appropriate for most sites, caused in large part by the replacement of central timesharing systems by personal computers. The class-based structure and the division into network and host part meant that the utilization of the address space would remain fairly low, so that far fewer than the theoretically possible 4 billion hosts within 2 million networks would "fit".

In the late 80s, the Internet had grown to 200,000 hosts. It became clear that with continued growth, particularly class B addresses would become scarce. In addition, the number of networks which the backbone network had to know about were beginning to tax the capacity of these routers. After a workshop in Kyoto, Japan, in spring of 1992, a proposal was made to replace IP by the OSI equivalent, the connectionless network protocol (CLNP). It had been in limited use within DECnet and offered variable-length addresses, so-called NSAPs (networks service address points). Typical NSAPs are 20 bytes long, allocated by a hierarchy of administrative domains. The Kyoto proposal caused a near-revolt within the Internet Engineering Task Force (IETF), the standardization body overseeing development of Internet-related protocols. Instead, a kind of design competition was set up, with several groups specifying replacement Internet protocols, with names such as TUBA, which continued the CLNP work, SIPP, and CATNIP [1]. The effort was labeled as IPng (IP next generation), with a nod towards the StarTrek movie series.

Since the deployment of a new internet protocol would take a number of years, an intermediate measure called Classless Interdomain Routing (CIDR) was specified. It abolished the traditional class-based address structure in favor of assigning topologically based groups of class-C addresses. Ideally, all networks attached to a local provider would share a common prefix, so that the local

[1] These are the ones remaining after a number of mergers.

provider would only have to advertise a single prefix for all its attached networks rather than entries for each individual network. CIDR slowed the growth of globally visible network addresses and, by assigning plentiful class-C instead of scarce class-B addresses, pushed the address reserves well into the next century.

In the summer of 1994, a body of "jurors" chosen for the purpose picked a slightly modified version of SIPP, as had been largely expected by the IETF. Since then, protocol details are in the process of being nailed down, with proposed standards, the first stage of the Internet standardization process, due this spring. The new protocol is now officially designated as IPv6 (IP version 6), while the current version is designated as IPv4.

3 While we're at it...

The primary motivator for replacing IPv4 was the perceived exhaustion of the address space. However, once it became clear that any new protocol changes would entail significant transition costs in hosts and routers, a new packet format also offered opportunities for simplification and additional functionality.

The IPv6 packet format is rather simple, as shown in Fig. ??. It consists of a version identifier that distinguishes IPv6 packets from IPv4 (and other) network layer packets. A flow label, with semantics still largely undefined, allows to group packets for common handling, e.g., to ensure priority scheduling of all video packets. The next-header field indicates whether the fixed header is followed by an IP-layer option or the transport protocol payload, combining the functions of the IPv4 next-protocol field and the option length indicator. The hop limit field, replacing the time-to-live field in IPv4, is initialized to the maximum number of routers the packet is allowed to traverse before being discarded. It avoids packets looping within the network forever. In IPv4, this field was also supposed to limit the maximum time a packet could spend in the network; this functionality was dropped since it is expensive to implement and was rarely implemented properly. Source and destination addresses are 16 bytes (128 bits) long, allowing to theoretically address $3 \cdot 10^3$8 end systems, or roughly one for every cell of every living human being on the planet.

The IPv4 header checksum was omitted, as the combination of link-layer and transport-layer checksum affords enough protection and avoids the recomputation at each router due to the changing hop limit field. The IPv4 type-of-service indicator has not been widely used; its function can be subsumed by the flow label.

In IPv4, routers would break packets into fragments if they would not fit within the maximum packet size of the link layer. In IPv6, only end systems perform fragmentation, and the necessary information (mainly a fragment offset and a packet identifier) is carried in an option rather than the fixed header.

By omitting these fields, the overall header length has "only" doubled to 40 bytes, even though the address length has quadrupled. An important design consideration was efficient implementation in current 64-bit RISC processors, which impose address-alignment restrictions. All fields are aligned on their "natural" width. The hop limit can be adjusted simply by decrementing a 64-bit word. All options that need to be handled by each router are grouped together so that they can be quickly found.

All compliant IPv6 implementation must offer integrity (authentication) and privacy through encryption, greatly reducing the risk of intrusion and data theft.

4 Addressing

Probably the most important function of the network layer is to get packets through a globally connected network. IPv6 addresses both identify and localize an end system, i.e., they are globally unique and are assigned in a manner reflecting network topology. Thus, they appear to be very similar to telephone numbers. However, unlike the current telephone network, a single customer may be connected to several different Internet service providers at widely separated geographical locations, which may in turn not be directly connected. As an example, consider a large company with a private network spanning continents. In some cases, the traditional distinction of providers and subscribers may not hold either, e.g., if an automotive company provides network services between its suppliers. Traditional geographic boundaries like oceans may be far less important for network connectivity than regulatory or political hurdles. It is well known that many intra-European packets travel by way of the U.S., for example. For a datagram network with dynamic routing, it is also not feasible to have a "flat" address space like the North American 800 numbers or the German 0130 numbers, with a translation to a "real" telephone number. (Note that a typical large long-distance telephone service provider only has about one hundred nodes to route between.) There appears to be no address assignment mechanism that can efficiently operate under all of these constraints, so compromises must be made.

The initial addressing architecture proposed for IPv6 is based on a hierarchy continent – network provider – subscriber (network) – host, possibly with an additional national registry level above the provider. At each level, variable length prefixes would be allocated according to the needs of the next lower hierarchy. Packets are routed longest-match-first, i.e., to the destination which shares the most common high-order bits with the packet destination address. For example, if all European hosts share the same prefix, a router in the U.S. might carry a separate, longer prefix for hosts in Iceland if they travel on a separate fiber. Another architecture under consideration is geographic addressing, very similar to the current assignment of German telephone numbers. However, it requires that providers are interconnected in each city, say.

To avoid locking in subscribers to their current provider and to reflect long-term changes in network topology, the network part of the address should be changeable, without having to tear down existing connections, reconfigure and reboot each and every host. Autoconfiguration protocols under consideration allow a host to learn its network address from its local router and allow routers to notify hosts of changing network addresses. The IPv6 address space is large enough that an organization can decide to use the 48-bit Ethernet addresses as the local address, avoiding the need to manage a local address space.

5 Getting from here to there

Clearly, there can be no "flag day" where all Internet hosts suddenly stop using the old IP and switch to the new. For one, some systems like printers have their software in ROM and will never speak IPv6. Others have long been orphaned by their manufacturers or have to run old operating system versions for one reason or another. Systems can maintain their old 32-bit IP address and still communicate using IPv6.

6 The future

The IPng working groups are currently putting the final touches on the core documents, with implementations expected later this year. A large number of other Internet standards will have to be changed, usually simply in wording or by allowing larger addresses. However, a number of technical issues remain. For example, key distribution for authentication and encryption has not yet been solved satisfactorily, partially because necessary algorithms are subject to U.S. patent protection. The larger question of the role of IPv6 in an ATM-dominated network, should this come to pass, remains an intriguing one.

References

[1] W. R. Stevens, *TCP/IP illustrated: the protocols*, vol. 1. Reading, Massachusetts: Addison-Wesley, 1994.