Network Working Group
Request for Comments: 1812
Obsoletes: 1716, 1009
Category: Standards Track

F. Baker, Editor
Cisco Systems
June 1995

# Requirements for IP Version 4 Routers

Status of this Memo

This document specifies an Internet standards track protocol for the Internet community, and requests discussion and suggestions for improvements. Please refer to the current edition of the "Internet Official Protocol Standards" (STD 1) for the standardization state and status of this protocol. Distribution of this memo is unlimited.

PREFACE

This document is an updated version of RFC 1716, the historical Router Requirements document. That RFC preserved the significant work that went into the working group, but failed to adequately describe current technology for the IESG to consider it a current standard.

The current editor had been asked to bring the document up to date, so that it is useful as a procurement specification and a guide to implementors. In this, he stands squarely on the shoulders of those who have gone before him, and depends largely on expert contributors for text. Any credit is theirs; the errors are his.

The content and form of this document are due, in large part, to the working group's chair, and document's original editor and author: Philip Almquist. It is also largely due to the efforts of its previous editor, Frank Kastenholz. Without their efforts, this document would not exist.

## 2. INTERNET ARCHITECTURE

This chapter does not contain any requirements. However, it does contain useful background information on the general architecture of the Internet and of routers.

General background and discussion on the Internet architecture and supporting protocol suite can be found in the DDN Protocol Handbook [ARCH:1]; for background see for example [ARCH:2], [ARCH:3], and [ARCH:4]. The Internet architecture and protocols are also covered in an ever-growing number of textbooks, such as [ARCH:5] and [ARCH:6].

## 2.1 Introduction

The Internet system consists of a number of interconnected packet networks supporting communication among host computers using the Internet protocols. These protocols include the Internet Protocol (IP), the Internet Control Message Protocol (ICMP), the Internet Group Management Protocol (IGMP), and a variety transport and application protocols that depend upon them. As was described in Section [1.2], the Internet Engineering Steering Group periodically releases an Official Protocols memo listing all the Internet protocols.

All Internet protocols use IP as the basic data transport mechanism. IP is a datagram, or connectionless, internetwork service and includes provision for addressing, type-of-service specification, fragmentation and reassembly, and security. ICMP and IGMP are considered integral parts of IP, although they are architecturally layered upon IP. ICMP provides error reporting, flow control, first-hop router redirection, and other maintenance and control functions. IGMP provides the mechanisms by which hosts and routers can join and leave IP multicast groups.

Reliable data delivery is provided in the Internet protocol suite by Transport Layer protocols such as the Transmission Control Protocol (TCP), which provides end-end retransmission, resequencing and connection control. Transport Layer connectionless service is provided by the User Datagram Protocol (UDP).

## 2.2 Elements of the Architecture

### 2.2.1 Protocol Layering

To communicate using the Internet system, a host must implement the layered set of protocols comprising the Internet protocol suite. A host typically must implement at least one protocol from each layer.

The protocol layers used in the Internet architecture are as follows [ARCH:7]:

Application Layer

> The Application Layer is the top layer of the Internet protocol suite. The Internet suite does not further subdivide the Application Layer, although some application layer protocols do contain some internal sub-layering. The application layer of the Internet suite essentially combines the functions of the top two layers - Presentation and Application - of the OSI Reference Model [ARCH:8]. The Application Layer in the Internet protocol suite also includes some of the function relegated to the Session Layer in the OSI Reference Model.

We distinguish two categories of application layer protocols: user protocols that provide service directly to users, and support protocols that provide common system functions. The most common Internet user protocols are:

- Telnet (remote login)
- FTP (file transfer)
- SMTP (electronic mail delivery)

There are a number of other standardized user protocols and many private user protocols.

Support protocols, used for host name mapping, booting, and management include SNMP, BOOTP, TFTP, the Domain Name System (DNS) protocol, and a variety of routing protocols.

Application Layer protocols relevant to routers are discussed in chapters 7, 8, and 9 of this memo.

Transport Layer

The Transport Layer provides end-to-end communication services. This layer is roughly equivalent to the Transport Layer in the OSI Reference Model, except that it also incorporates some of OSI's Session Layer establishment and destruction functions.

There are two primary Transport Layer protocols at present:

- Transmission Control Protocol (TCP)
- User Datagram Protocol (UDP)

TCP is a reliable connection-oriented transport service that provides end-to-end reliability, resequencing, and flow control. UDP is a connectionless (datagram) transport service. Other transport protocols have been developed by the research community, and the set of official Internet transport protocols may be expanded in the future.

Transport Layer protocols relevant to routers are discussed in Chapter 6.

Internet Layer

All Internet transport protocols use the Internet Protocol (IP) to carry data from source host to destination host. IP is a connectionless or datagram internetwork service, providing no end-to-end delivery guarantees. IP datagrams may arrive at the destination host damaged, duplicated, out of order, or not at all. The layers above IP are responsible for reliable delivery service when it is required. The IP protocol includes provision for addressing, type-of-service specification, fragmentation and reassembly, and security.

The datagram or connectionless nature of IP is a fundamental and characteristic feature of the Internet architecture.

The Internet Control Message Protocol (ICMP) is a control protocol that is considered to be an integral part of IP, although it is architecturally layered upon IP - it uses IP to carry its data end-to-end. ICMP provides error reporting, congestion reporting, and first-hop router redirection.

The Internet Group Management Protocol (IGMP) is an Internet layer protocol used for establishing dynamic host groups for IP multicasting.

The Internet layer protocols IP, ICMP, and IGMP are discussed in chapter 4.

Link Layer

To communicate on a directly connected network, a host must implement the communication protocol used to interface to that network. We call this a Link Layer protocol.

Some older Internet documents refer to this layer as the Network Layer, but it is not the same as the Network Layer in the OSI Reference Model.

This layer contains everything below the Internet Layer and above the Physical Layer (which is the media connectivity, normally electrical or optical, which encodes and transports messages). Its responsibility is the correct delivery of messages, among which it does not differentiate.

Protocols in this Layer are generally outside the scope of Internet standardization; the Internet (intentionally) uses existing standards whenever possible. Thus, Internet Link Layer standards usually address only address resolution and rules for transmitting IP packets over specific Link Layer protocols. Internet Link Layer standards are discussed in chapter 3.

### 2.2.2 Networks

The constituent networks of the Internet system are required to provide only packet (connectionless) transport. According to the IP service specification, datagrams can be delivered out of order, be lost or duplicated, and/or contain errors.

For reasonable performance of the protocols that use IP (e.g., TCP), the loss rate of the network should be very low. In networks providing connection-oriented service, the extra reliability provided by virtual circuits enhances the end-end robustness of the system, but is not necessary for Internet operation.

Constituent networks may generally be divided into two classes:

- Local-Area Networks (LANs) LANs may have a variety of designs. LANs normally cover a small geographical area (e.g., a single building or plant site) and provide high bandwidth with low delays. LANs may be passive (similar to Ethernet) or they may be active (such as ATM).
- Wide-Area Networks (WANs) Geographically dispersed hosts and LANs are interconnected by wide-area networks, also called long-haul networks. These networks may have a complex internal structure of lines and packet-switches, or they may be as simple as point-to-point lines.

### 2.2.3 Routers

In the Internet model, constituent networks are connected together by IP datagram forwarders which are called routers or IP routers. In this document, every use of the term router is equivalent to IP router. Many older Internet documents refer to routers as gateways.

Historically, routers have been realized with packet-switching software executing on a general-purpose CPU. However, as custom hardware development becomes cheaper and as higher throughput is

required, special purpose hardware is becoming increasingly common. This specification applies to routers regardless of how they are implemented.

A router connects to two or more logical interfaces, represented by IP subnets or unnumbered point to point lines (discussed in section [2.2.7]). Thus, it has at least one physical interface. Forwarding an IP datagram generally requires the router to choose the address and relevant interface of the next-hop router or (for the final hop) the destination host. This choice, called relaying or forwarding depends upon a route database within the router. The route database is also called a routing table or forwarding table. The term "router" derives from the process of building this route database; routing protocols and configuration interact in a process called routing.

The routing database should be maintained dynamically to reflect the current topology of the Internet system. A router normally accomplishes this by participating in distributed routing and reachability algorithms with other routers.

Routers provide datagram transport only, and they seek to minimize the state information necessary to sustain this service in the interest of routing flexibility and robustness.

Packet switching devices may also operate at the Link Layer; such devices are usually called bridges. Network segments that are connected by bridges share the same IP network prefix forming a single IP subnet. These other devices are outside the scope of this document.

### 2.2.4 Autonomous Systems

An Autonomous System (AS) is a connected segment of a network topology that consists of a collection of subnetworks (with hosts attached) interconnected by a set of routes. The subnetworks and the routers are expected to be under the control of a single operations and maintenance (O&M) organization. Within an AS routers may use one or more interior routing protocols, and sometimes several sets of metrics. An AS is expected to present to other ASs an appearence of a coherent interior routing plan, and a consistent picture of the destinations reachable through the AS. An AS is identified by an Autonomous System number.

The concept of an AS plays an important role in the Internet routing (see Section 7.1).

### 2.2.5 Addressing Architecture

An IP datagram carries 32-bit source and destination addresses, each of which is partitioned into two parts - a constituent network prefix and a host number on that network. Symbolically:

```
IP-address ::= { <Network-prefix>, <Host-number> }
```

To finally deliver the datagram, the last router in its path must map the Host-number (or rest) part of an IP address to the host's Link Layer address.

#### 2.2.5.1 Classical IP Addressing Architecture

Although well documented elsewhere [INTERNET:2], it is useful to describe the historical use of the network prefix. The language developed to describe it is used in this and other documents and permeates the thinking behind many protocols.

The simplest classical network prefix is the Class A, B, C, D, or E network prefix. These address ranges are discriminated by observing the values of the most significant bits of the address, and break the address into simple prefix and host number fields. This is described in [INTERNET:18]. In short, the classification is:

| | | |
|---|---|---|
| 0xxx | Class A | general purpose unicast addresses with standard 8 bit prefix |
| 10xx | Class B | general purpose unicast addresses with standard 16 bit prefix |
| 110x | Class C | general purpose unicast addresses with standard 24 bit prefix |
| 1110 | Class D | IP Multicast Addresses - 28 bit prefix, non- aggregatable |
| 1111 | Class E | reserved for experimental use |

This simple notion has been extended by the concept of subnets. These were introduced to allow arbitrary complexity of interconnected LAN structures within an organization, while insulating the Internet system against explosive growth in assigned network prefixes and routing complexity. Subnets provide a multi-level hierarchical routing structure for the Internet system. The subnet extension, described in [INTERNET:2], is a required part of the Internet architecture. The basic idea is to partition the <Host-number> field into two parts: a subnet number, and a true host number on that subnet:

```
IP-address ::=
  { <Network-number>, <Subnet-number>, <Host-number> }
```

The interconnected physical networks within an organization use the same network prefix but different subnet numbers. The distinction between the subnets of such a subnetted network is not normally visible outside of that network. Thus, routing in the rest of the Internet uses only the <Network-prefix> part of the IP destination address. Routers outside the network treat <Network-prefix> and <Host-number> together as an uninterpreted rest part of the 32-bit IP address. Within the subnetted network, the routers use the extended network prefix:

```
{ <Network-number>, <Subnet-number> }
```

The bit positions containing this extended network number have historically been indicated by a 32-bit mask called the subnet mask. The <Subnet-number> bits SHOULD be contiguous and fall between the <Network-number> and the <Host-number> fields. More up to date protocols do not refer to a subnet mask, but to a prefix length; the "prefix" portion of an address is that which would be selected by a subnet mask whose most significant bits are all ones and the rest are zeroes. The length of the prefix equals the number of ones in the subnet mask. This document assumes that all subnet masks are expressible as prefix lengths.

The inventors of the subnet mechanism presumed that each piece of an organization's network would have only a single subnet number. In practice, it has often proven necessary or useful to have several subnets share a single physical cable. For this reason, routers should be capable of configuring multiple subnets on the same physical interfaces, and treat them (from a routing or forwarding perspective) as though they were distinct physical interfaces.

### 2.2.5.2 Classless Inter Domain Routing (CIDR)

The explosive growth of the Internet has forced a review of address assignment policies. The traditional uses of general purpose (Class A, B, and C) networks have been modified to achieve better use of IP's

32-bit address space. Classless Inter Domain Routing (CIDR) [INTERNET:15] is a method currently being deployed in the Internet backbones to achieve this added efficiency. CIDR depends on deploying and routing to arbitrarily sized networks. In this model, hosts and routers make no assumptions about the use of addressing in the internet. The Class D (IP Multicast) and Class E (Experimental) address spaces are preserved, although this is primarily an assignment policy.

By definition, CIDR comprises three elements:

- topologically significant address assignment,
- routing protocols that are capable of aggregating network layer reachability information, and
- consistent forwarding algorithm ("longest match").

The use of networks and subnets is now historical, although the language used to describe them remains in current use. They have been replaced by the more tractable concept of a network prefix. A network prefix is, by definition, a contiguous set of bits at the more significant end of the address that defines a set of systems; host numbers select among those systems. There is no requirement that all the internet use network prefixes uniformly. To collapse routing information, it is useful to divide the internet into addressing domains. Within such a domain, detailed information is available about constituent networks; outside it, only the common network prefix is advertised.

The classical IP addressing architecture used addresses and subnet masks to discriminate the host number from the network prefix. With network prefixes, it is sufficient to indicate the number of bits in the prefix. Both representations are in common use. Architecturally correct subnet masks are capable of being represented using the prefix length description. They comprise that subset of all possible bits patterns that have

- a contiguous string of ones at the more significant end,
- a contiguous string of zeros at the less significant end, and
- no intervening bits.

Routers SHOULD always treat a route as a network prefix, and SHOULD reject configuration and routing information inconsistent with that model.

```
IP-address ::= { <Network-prefix>, <Host-number> }
```

An effect of the use of CIDR is that the set of destinations associated with address prefixes in the routing table may exhibit subset relationship. A route describing a smaller set of destinations (a longer prefix) is said to be more specific than a route describing a larger set of destinations (a shorter prefix); similarly, a route describing a larger set of destinations (a shorter prefix) is said to be less specific than a route describing a smaller set of destinations (a longer prefix). Routers must use the most specific matching route (the longest matching network prefix) when forwarding traffic.

### 2.2.6 IP Multicasting

IP multicasting is an extension of Link Layer multicast to IP internets. Using IP multicasts, a single datagram can be addressed to multiple hosts without sending it to all. In the extended case, these hosts may reside in different address domains. This collection of hosts is called a multicast group. Each multicast group is represented as a Class D IP address. An IP datagram sent to the group is to be

delivered to each group member with the same best- effort delivery as that provided for unicast IP traffic. The sender of the datagram does not itself need to be a member of the destination group.

The semantics of IP multicast group membership are defined in [INTERNET:4]. That document describes how hosts and routers join and leave multicast groups. It also defines a protocol, the Internet Group Management Protocol (IGMP), that monitors IP multicast group membership.

Forwarding of IP multicast datagrams is accomplished either through static routing information or via a multicast routing protocol. Devices that forward IP multicast datagrams are called multicast routers. They may or may not also forward IP unicasts. Multicast datagrams are forwarded on the basis of both their source and destination addresses. Forwarding of IP multicast packets is described in more detail in Section [5.2.1]. Appendix D discusses multicast routing protocols.

### 2.2.7 Unnumbered Lines and Networks Prefixes

Traditionally, each network interface on an IP host or router has its own IP address. This can cause inefficient use of the scarce IP address space, since it forces allocation of an IP network prefix to every point-to-point link.

To solve this problem, a number of people have proposed and implemented the concept of unnumbered point to point lines. An unnumbered point to point line does not have any network prefix associated with it. As a consequence, the network interfaces connected to an unnumbered point to point line do not have IP addresses.

Because the IP architecture has traditionally assumed that all interfaces had IP addresses, these unnumbered interfaces cause some interesting dilemmas. For example, some IP options (e.g., Record Route) specify that a router must insert the interface address into the option, but an unnumbered interface has no IP address. Even more fundamental (as we shall see in chapter 5) is that routes contain the IP address of the next hop router. A router expects that this IP address will be on an IP (sub)net to which the router is connected. That assumption is of course violated if the only connection is an unnumbered point to point line.

To get around these difficulties, two schemes have been conceived. The first scheme says that two routers connected by an unnumbered point to point line are not really two routers at all, but rather two half-routers that together make up a single virtual router. The unnumbered point to point line is essentially considered to be an internal bus in the virtual router. The two halves of the virtual router must coordinate their activities in such a way that they act exactly like a single router.

This scheme fits in well with the IP architecture, but suffers from two important drawbacks. The first is that, although it handles the common case of a single unnumbered point to point line, it is not readily extensible to handle the case of a mesh of routers and unnumbered point to point lines. The second drawback is that the interactions between the half routers are necessarily complex and are not standardized, effectively precluding the connection of equipment from different vendors using unnumbered point to point lines.

Because of these drawbacks, this memo has adopted an alternate scheme, which has been invented multiple times but which is probably originally attributable to Phil Karn. In this scheme, a router that has unnumbered point to point lines also has a special IP address, called a router-id in this memo. The

router-id is one of the router's IP addresses (a router is required to have at least one IP address). This router-id is used as if it is the IP address of all unnumbered interfaces.

### 2.2.8 Notable Oddities

#### 2.2.8.1 Embedded Routers

A router may be a stand-alone computer system, dedicated to its IP router functions. Alternatively, it is possible to embed router functions within a host operating system that supports connections to two or more networks. The best-known example of an operating system with embedded router code is the Berkeley BSD system. The embedded router feature seems to make building a network easy, but it has a number of hidden pitfalls:

1. If a host has only a single constituent-network interface, it should not act as a router.

   For example, hosts with embedded router code that gratuitously forward broadcast packets or datagrams on the same net often cause packet avalanches.

2. If a (multihomed) host acts as a router, it is subject to the requirements for routers contained in this document.

   For example, the routing protocol issues and the router control and monitoring problems are as hard and important for embedded routers as for stand-alone routers.

   Internet router requirements and specifications may change independently of operating system changes. An administration that operates an embedded router in the Internet is strongly advised to maintain and update the router code. This might require router source code.

3. When a host executes embedded router code, it becomes part of the Internet infrastructure. Thus, errors in software or configuration can hinder communication between other hosts. As a consequence, the host administrator must lose some autonomy.

   In many circumstances, a host administrator will need to disable router code embedded in the operating system. For this reason, it should be straightforward to disable embedded router functionality.

4. When a host running embedded router code is concurrently used for other services, the Operation and Maintenance requirements for the two modes of use may conflict.

   For example, router O&M will in many cases be performed remotely by an operations center; this may require privileged system access that the host administrator would not normally want to distribute.

#### 2.2.8.2 Transparent Routers

There are two basic models for interconnecting local-area networks and wide-area (or long-haul) networks in the Internet. In the first, the local-area network is assigned a network prefix and all routers in the Internet must know how to route to that network. In the second, the local-area network shares (a small part of) the address space of the wide-area network. Routers that support this second model are

called address sharing routers or transparent routers. The focus of this memo is on routers that support the first model, but this is not intended to exclude the use of transparent routers.

The basic idea of a transparent router is that the hosts on the local-area network behind such a router share the address space of the wide-area network in front of the router. In certain situations this is a very useful approach and the limitations do not present significant drawbacks.

The words in front and behind indicate one of the limitations of this approach: this model of interconnection is suitable only for a geographically (and topologically) limited stub environment. It requires that there be some form of logical addressing in the network level addressing of the wide-area network. IP addresses in the local environment map to a few (usually one) physical address in the wide-area network. This mapping occurs in a way consistent with the { IP address <-> network address } mapping used throughout the wide-area network.

Multihoming is possible on one wide-area network, but may present routing problems if the interfaces are geographically or topologically separated. Multihoming on two (or more) wide-area networks is a problem due to the confusion of addresses.

The behavior that hosts see from other hosts in what is apparently the same network may differ if the transparent router cannot fully emulate the normal wide-area network service. For example, the ARPANET used a Link Layer protocol that provided a Destination Dead indication in response to an attempt to send to a host that was off- line. However, if there were a transparent router between the ARPANET and an Ethernet, a host on the ARPANET would not receive a Destination Dead indication for Ethernet hosts.

## 2.3 Router Characteristics

An Internet router performs the following functions:

1. Conforms to specific Internet protocols specified in this document, including the Internet Protocol (IP), Internet Control Message Protocol (ICMP), and others as necessary.
2. Interfaces to two or more packet networks. For each connected network the router must implement the functions required by that network. These functions typically include:
    - Encapsulating and decapsulating the IP datagrams with the connected network framing (e.g., an Ethernet header and checksum),
    - Sending and receiving IP datagrams up to the maximum size supported by that network, this size is the network's Maximum Transmission Unit or MTU,
    - Translating the IP destination address into an appropriate network-level address for the connected network (e.g., an Ethernet hardware address), if needed, and
    - Responding to network flow control and error indications, if any.

    See chapter 3 (Link Layer).

3. Receives and forwards Internet datagrams. Important issues in this process are buffer management, congestion control, and fairness.
    - Recognizes error conditions and generates ICMP error and information messages as required.
    - Drops datagrams whose time-to-live fields have reached zero.
    - Fragments datagrams when necessary to fit into the MTU of the next network.

See chapter 4 (Internet Layer - Protocols) and chapter 5 (Internet Layer - Forwarding) for more information.

4. Chooses a next-hop destination for each IP datagram, based on the information in its routing database. See chapter 5 (Internet Layer - Forwarding) for more information.
5. (Usually) supports an interior gateway protocol (IGP) to carry out distributed routing and reachability algorithms with the other routers in the same autonomous system. In addition, some routers will need to support an exterior gateway protocol (EGP) to exchange topological information with other autonomous systems. See chapter 7 (Application Layer - Routing Protocols) for more information.
6. Provides network management and system support facilities, including loading, debugging, status reporting, exception reporting and control. See chapter 8 (Application Layer - Network Management Protocols) and chapter 10 (Operation and Maintenance) for more information.

A router vendor will have many choices on power, complexity, and features for a particular router product. It may be helpful to observe that the Internet system is neither homogeneous nor fully connected. For reasons of technology and geography it is growing into a global interconnect system plus a fringe of LANs around the edge. More and more these fringe LANs are becoming richly interconnected, thus making them less out on the fringe and more demanding on router requirements.

- The global interconnect system is composed of a number of wide-area networks to which are attached routers of several Autonomous Systems (AS); there are relatively few hosts connected directly to the system.
- Most hosts are connected to LANs. Many organizations have clusters of LANs interconnected by local routers. Each such cluster is connected by routers at one or more points into the global interconnect system. If it is connected at only one point, a LAN is known as a stub network.

Routers in the global interconnect system generally require:

- Advanced Routing and Forwarding Algorithms

  These routers need routing algorithms that are highly dynamic, impose minimal processing and communication burdens, and offer type-of-service routing. Congestion is still not a completely resolved issue (see Section [5.3.6]). Improvements in these areas are expected, as the research community is actively working on these issues.

- High Availability

  These routers need to be highly reliable, providing 24 hours a day, 7 days a week service. Equipment and software faults can have a wide-spread (sometimes global) effect. In case of failure, they must recover quickly. In any environment, a router must be highly robust and able to operate, possibly in a degraded state, under conditions of extreme congestion or failure of network resources.

- Advanced O&M Features

  Internet routers normally operate in an unattended mode. They will typically be operated remotely from a centralized monitoring center. They need to provide sophisticated means for monitoring and measuring traffic and other events and for diagnosing faults.

- High Performance

  Long-haul lines in the Internet today are most frequently full duplex 56 KBPS, DS1 (1.544 Mbps), or DS3 (45 Mbps) speeds. LANs, which are half duplex multiaccess media, are typically Ethernet (10Mbps) and, to a lesser degree, FDDI (100Mbps). However, network media technology is constantly advancing and higher speeds are likely in the future.

The requirements for routers used in the LAN fringe (e.g., campus networks) depend greatly on the demands of the local networks. These may be high or medium-performance devices, probably competitively procured from several different vendors and operated by an internal organization (e.g., a campus computing center). The design of these routers should emphasize low average latency and good burst performance, together with delay and type-of-service sensitive resource management. In this environment there may be less formal O&M but it will not be less important. The need for the routing mechanism to be highly dynamic will become more important as networks become more complex and interconnected. Users will demand more out of their local connections because of the speed of the global interconnects.

As networks have grown, and as more networks have become old enough that they are phasing out older equipment, it has become increasingly imperative that routers interoperate with routers from other vendors.

Even though the Internet system is not fully interconnected, many parts of the system need to have redundant connectivity. Rich connectivity allows reliable service despite failures of communication lines and routers, and it can also improve service by shortening Internet paths and by providing additional capacity. Unfortunately, this richer topology can make it much more difficult to choose the best path to a particular destination.

## 2.4 Architectural Assumptions

The current Internet architecture is based on a set of assumptions about the communication system. The assumptions most relevant to routers are as follows:

- The Internet is a network of networks.

  Each host is directly connected to some particular network(s); its connection to the Internet is only conceptual. Two hosts on the same network communicate with each other using the same set of protocols that they would use to communicate with hosts on distant networks.

- Routers do not keep connection state information.

  To improve the robustness of the communication system, routers are designed to be stateless, forwarding each IP packet independently of other packets. As a result, redundant paths can be exploited to provide robust service in spite of failures of intervening routers and networks.

  All state information required for end-to-end flow control and reliability is implemented in the hosts, in the transport layer or in application programs. All connection control information is thus co-located with the end points of the communication, so it will be lost only if an end point fails. Routers control message flow only indirectly, by dropping packets or increasing network delay.

Note that future protocol developments may well end up putting some more state into routers. This is especially likely for multicast routing, resource reservation, and flow based forwarding.

- Routing complexity should be in the routers.

  Routing is a complex and difficult problem, and ought to be performed by the routers, not the hosts. An important objective is to insulate host software from changes caused by the inevitable evolution of the Internet routing architecture.

- The system must tolerate wide network variation.

  A basic objective of the Internet design is to tolerate a wide range of network characteristics - e.g., bandwidth, delay, packet loss, packet reordering, and maximum packet size. Another objective is robustness against failure of individual networks, routers, and hosts, using whatever bandwidth is still available. Finally, the goal is full open system interconnection: an Internet router must be able to interoperate robustly and effectively with any other router or Internet host, across diverse Internet paths.

  Sometimes implementors have designed for less ambitious goals. For example, the LAN environment is typically much more benign than the Internet as a whole; LANs have low packet loss and delay and do not reorder packets. Some vendors have fielded implementations that are adequate for a simple LAN environment, but work badly for general interoperation. The vendor justifies such a product as being economical within the restricted LAN market. However, isolated LANs seldom stay isolated for long. They are soon connected to each other, to organization-wide internets, and eventually to the global Internet system. In the end, neither the customer nor the vendor is served by incomplete or substandard routers.

  The requirements in this document are designed for a full-function router. It is intended that fully compliant routers will be usable in almost any part of the Internet.