

# IP Address Translation and Routing

By Mark E. Donaldson

## Name Resolution

While IP is designed to work with the 32-bit IP addresses of the source and the destination hosts, computer users are much better at using and remembering names than IP addresses. If a name is used as an alias for the IP address, a mechanism must exist for assigning that name to the appropriate IP node to ensure its uniqueness and resolving it to its IP address.

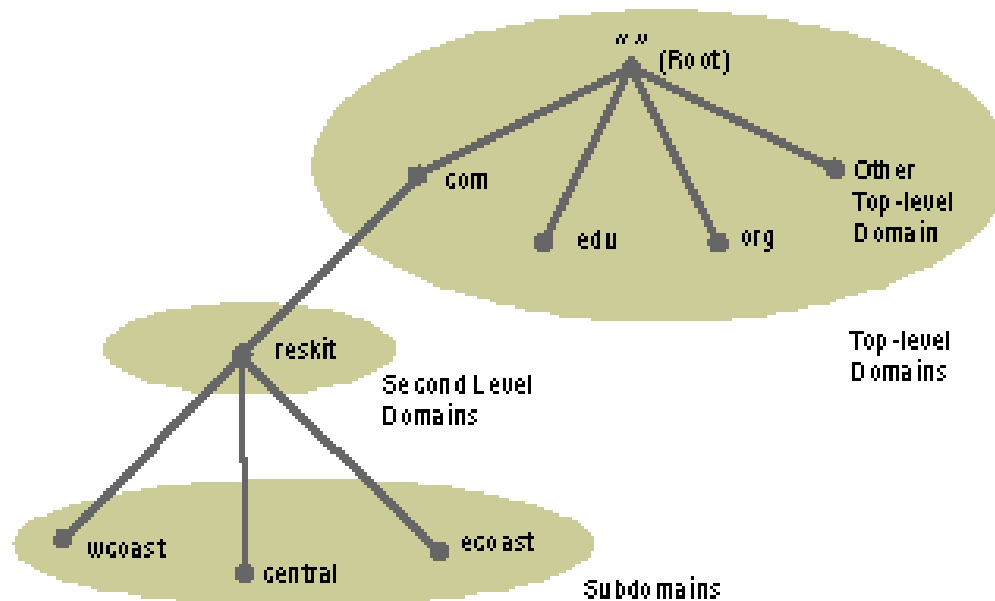
## Host Name Resolution

A *host name* is an alias assigned to an IP node to identify it as a TCP/IP host. The host name can be up to 255 characters long and can contain alphabetic and numeric characters and the “-” and “.” characters. Multiple host names can be assigned to the same host. For Windows 2000–based computers, the host name does not have to match the Windows 2000 computer name.

Host names can take various forms. The two most common forms are a nickname and a domain name. A nickname is an alias to an IP address that individual people can assign and use. A *domain name* is a structured name that follows Internet conventions.

## Domain Names

The InterNIC has created and maintains a hierarchical namespace called the *Domain Name System* (DNS). DNS is a naming scheme that looks similar to the directory structure for files on a disk. However, instead of tracing a file from the root directory through subdirectories to its final location and its file name, a host name is traced from its final location through its parent domains back up to the root. The unique name of the host, representing its position in the hierarchy, is called its *Fully Qualified Domain Name* (FQDN). The top-level domain namespace is shown in the figure below with example second-level and subdomains.



# IP Address Translation and Routing

By Mark E. Donaldson

The domain namespace consists of:

- The root domain, representing the root of the namespace and indicated with a "" (null).
- Top-level domains, those directly below the root, indicating a type of organization. On the Internet, the InterNIC is responsible for the maintenance of top-level domain names.
- Second-level domains, below the top level domains, identifying a specific organization within its top-level domain. On the Internet, the InterNIC is responsible for the maintenance of second-level domain names and ensuring their uniqueness.
- Subdomains of the organization, below the second-level domain. The individual organization is responsible for the creation and maintenance of subdomains.

Domain names are not case sensitive. Organizations not connected to the Internet can implement whatever top and second-level domain names they want. However, typical implementations do adhere to the InterNIC specification so that eventual participation in the Internet will not require a renaming process.

## Host Name Resolution Using a Hosts File

One common way to resolve a host name to an IP address is to use a locally stored database file that contains IP-address-to-host-name mappings. On most UNIX systems, this file is `/etc/hosts`. On Windows 2000 systems, it is the Hosts file in the `\%SystemRoot%\system32\drivers\etc` directory. Following is an example of the contents of the Hosts file:

```
#
# Table of IP addresses and host names
#
127.0.0.1      localhost
139.41.34.1   router
167.91.45.121 server1.central.slate.com s1
```

Within the Hosts file:

- Multiple host names can be assigned to the same IP address. This allows the user at this computer to refer to this server using the nickname rather than typing the entire FQDN.
- Entries can be case sensitive depending on the platform. Entries in the Hosts file for UNIX computers are case sensitive. Entries in the Hosts file for Windows 2000 and Windows NT-based computers are not case sensitive.

The advantage of using a Hosts file is that it is customizable for the user. Each user can create whatever entries they want, including easy-to-remember nicknames for frequently accessed resources. However, the individual maintenance of the Hosts file does not scale well to storing large numbers of FQDN mappings.

## Host Name Resolution Using a DNS Server

To make host name resolution scalable and centrally manageable, IP address mappings for FQDNs are stored on *DNS servers*, computers that stores FQDN-to-IP-address mappings. To enable the querying of a DNS server by a host computer, a component called the *DNS resolver* is enabled and

# IP Address Translation and Routing

By Mark E. Donaldson

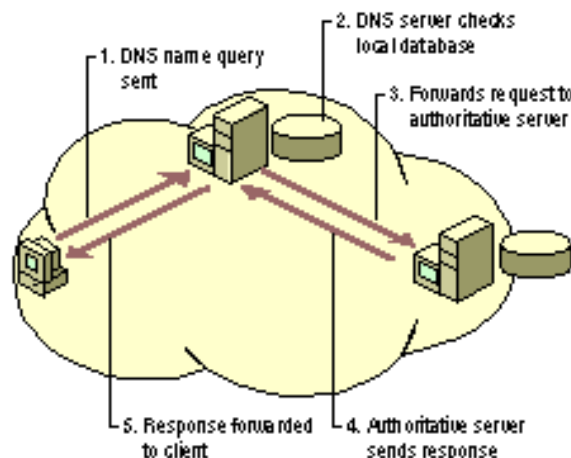
configured with the IP address of the DNS server. The DNS resolver is a built-in component of TCP/IP protocol stacks supplied with most network operating systems, including Windows 2000.

When a Windows Sockets application is given an FQDN as the destination location, the application calls a Windows Sockets function to resolve the name to an IP address. The request is passed to the DNS resolver component in the TCP/IP protocol. The DNS resolver packages the FQDN request as a DNS Name Query packet and sends it to the DNS server.

DNS is a distributed naming system. Rather than storing all the records for the entire namespace on each DNS server, each DNS server only stores the records for a specific portion of the namespace. The DNS server is authoritative for the portion of the namespace that corresponds to records stored on that DNS server. In the case of the Internet, hundreds of DNS servers store various portions of the Internet namespace. To facilitate the resolution of any valid domain name by any DNS server, DNS servers are also configured with pointer records to other DNS servers.

The following process outlines what happens when the DNS resolver component on a host sends a DNS query to a DNS server. This process is shown in Figure 1.12 and is simplified so that you can gain a basic understanding of the DNS resolution process.

1. The DNS resolver component of the DNS client formats a DNS Name Query containing the FQDN and sends it to the configured DNS server.
2. The DNS server checks the FQDN in the DNS Name Query against locally stored address records. If a record is found, the IP address corresponding to the requested FQDN is sent back to the client.
3. If the FQDN is not found, the DNS server forwards the request to a DNS server that is authoritative for the FQDN.
4. The authoritative DNS server returns the reply, containing the resolved IP address, back to the original DNS server.
5. The original DNS server sends the IP address mapping information to the client.



# IP Address Translation and Routing

By Mark E. Donaldson

To obtain the IP address of a server that is authoritative for the FQDN, DNS servers on the Internet go through an iterative process of querying multiple DNS servers until the authoritative server is found.

## Combining a Local Database File with DNS

TCP/IP implementations, including Windows 2000, allow the use of both a local database file and a DNS server to resolve host names. When a user specifies a host name in a TCP/IP command or utility:

1. TCP/IP checks the local database file (the Hosts file) for a matching name.
2. If a matching name is not found in the local database file, the host name is packaged as a DNS Name Query and sent to the configured DNS server.

## IP Routing

After the host name is resolved to an IP address, the IP packet must be sent by the sending host to the resolved IP address. *Routing* is the process of forwarding a packet based on the destination IP address. Routing occurs at a sending TCP/IP host and at an IP router. A *router* is a device that forwards the packets from one network to another. Routers are also commonly referred to as *gateways*. In both cases, sending host and router, a decision has to be made about where the packet is forwarded.

To make these decisions, the IP layer consults a routing table stored in memory. Routing table entries are created by default when TCP/IP initializes and additional entries are added either manually by a system administrator or automatically through communication with routers.

## Direct and Indirect Delivery

Forwarded IP packets use at least one of two types of delivery based on whether the IP packet is forwarded to the final destination or whether it is forwarded to an IP router. These two types of delivery are known as direct and indirect delivery.

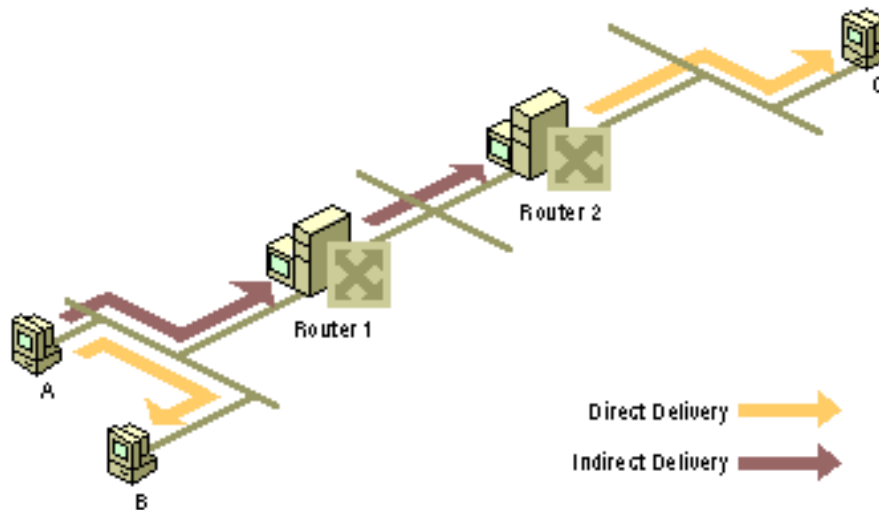
*Direct delivery* occurs when the IP node (either the sending node or an IP router) forwards a packet to the final destination on a directly attached network. The IP node encapsulates the IP datagram in a frame format for the Network Interface layer (such as Ethernet or Token Ring) addressed to the destination's physical address.

*Indirect delivery* occurs when the IP node (either the sending node or an IP router) forwards a packet to an intermediate node (an IP router) because the final destination is not on a directly attached network. The IP node encapsulates the IP datagram in a frame format, addressed to the IP router's physical address, for the Network Interface layer (such as Ethernet or Token Ring).

IP routing is a combination of direct and indirect deliveries. In the figure below Figure 2 when sending packets to node B, node A performs a direct delivery. When sending packets to node C, node A performs an indirect delivery to Router 1. Router 1 performs an indirect delivery to Router 2. Router 2 performs a direct delivery to node C.

# IP Address Translation and Routing

By Mark E. Donaldson



## IP Routing Table

A routing table is present on all IP nodes. The routing table stores information about IP networks and how they can be reached (either directly or indirectly). Because all IP nodes perform some form of IP routing, routing tables are not exclusive to IP routers. Any node loading the TCP/IP protocol has a routing table. There are a series of default entries according to the configuration of the node and additional entries can be entered either manually through TCP/IP utilities or dynamically through interaction with routers.

When an IP packet is to be forwarded, the routing table is used to determine:

1. The forwarding or next-hop IP address:

For a direct delivery, the forwarding IP address is the destination IP address in the IP packet. For an indirect delivery, the forwarding IP address is the IP address of a router.

2. The interface to be used for the forwarding:

The interface identifies the physical or logical interface such as a network adapter that is used to forward the packet to either its destination or the next router.

## IP Routing Table Entry Types

An entry in the IP routing table contains the following information in the order presented:

*Network ID.* The network ID or destination corresponding to the route. The network ID can be class-based, subnet, or supernet network ID, or an IP address for a host route.

*Network Mask.* The mask that is used to match a destination IP address to the network ID.

*Next Hop.* The IP address of the next hop.

*Interface.* An indication of which network interface is used to forward the IP packet.

# IP Address Translation and Routing

By Mark E. Donaldson

*Metric.* A number used to indicate the cost of the route so the best route among possible multiple routes to the same destination can be selected. A common use of the metric is to indicate the number of hops (routers crossed) to the network ID.

Routing table entries can be used to store the following types of routes:

*Directly Attached Network IDs.* Routes for network IDs that are directly attached. For directly attached networks, the Next Hop field can be blank or contain the IP address of the interface on that network.

*Remote Network IDs.* Routes for network IDs that are not directly attached but are available across other routers. For remote networks, the Next Hop field is the IP address of a local router in between the forwarding node and the remote network.

*Host Routes.* A route to a specific IP address. Host routes allow routing to occur on a per-IP address basis. For host routes, the network ID is the IP address of the specified host and the network mask is 255.255.255.255.

*Default Route.* The default route is designed to be used when a more specific network ID or host route is not found. The default route network ID is 0.0.0.0 with the network mask of 0.0.0.0.

## Route Determination Process

To determine which routing table entry is used for the forwarding decision, IP uses the following process:

- For each entry in a routing table, perform a bit-wise logical AND between the destination IP address and the network mask. Compare the result with the network ID of the entry for a match.
- The list of matching routes is compiled. The route that has the longest match (the route that matched the most amount of bits with the destination IP address) is chosen. The longest matching route is the most specific route to the destination IP address. If multiple entries with the longest match are found (multiple routes to the same network ID, for example), the router uses the lowest metric to select the best route. If multiple entries exist that are the longest match and the lowest metric, the router is free to choose which routing table entry to use.

The end result of the route determination process is the choice of a single route in the routing table. The route chosen yields a forwarding IP address (the next hop IP address) and an interface (the port). If the route determination process fails to find a route, IP declares a routing error. For the sending host, an IP routing error is internally indicated to the upper layer protocol such as TCP or UDP. For a router, an ICMP Destination Unreachable-Host Unreachable message is sent to the source host.

## Example Routing Table for Windows 2000

Table 1.28 shows the default routing table for a Windows 2000-based host (not a router). The host has a single network adapter and has the IP address 157.55.27.90, subnet mask 255.255.240.0 (/20), and default gateway of 157.55.16.1.

# IP Address Translation and Routing

By Mark E. Donaldson

## Windows 2000 Routing Table

Network Destination	Netmask	Gateway	Interface	Metric	Purpose
0.0.0.0	0.0.0.0	157.55.16.1	157.55.27.90	1	Default Route
127.0.0.0	255.0.0.0	127.0.0.1	127.0.0.1	1	Loopback Network
157.55.16.0	255.255.240.0	157.55.27.90	157.55.27.90	1	Directly Attached Network
157.55.27.90	255.255.255.255	127.0.0.1	127.0.0.1	1	Local Host
157.55.255.255	255.255.255.255	157.55.27.90	157.55.27.90	1	Network Broadcast
224.0.0.0	224.0.0.0	157.55.27.90	157.55.27.90	1	Multicast Address
255.255.255.255	255.255.255.255	157.55.27.90	157.55.27.90	1	Limited Broadcast

### Default Route

The entry corresponding to the default gateway configuration is a network destination of 0.0.0.0 with a network mask (netmask) of 0.0.0.0. Any destination IP address joined with 0.0.0.0 by a logical AND results in 0.0.0.0. Therefore, for any IP address, the default route produces a match. If the default route is chosen because no better routes were found, the IP packet is forwarded to the IP address in the Gateway column using the interface corresponding to the IP address in the Interface column.

### Loopback Network

The loopback network entry is designed to take any IP address of the form 127.x.y.z and forward it to the special loopback address of 127.0.0.1.

### Directly Attached Network

The local network entry corresponds to the directly attached network. IP packets destined for the directly attached network are not forwarded to a router but sent directly to the destination. Note that the Gateway and Interface columns match the IP address of the node. This indicates that the packet is sent from the network adapter corresponding to the node's IP address.

### Local Host

The local host entry is a host route (network mask of 255.255.255.255) corresponding to the IP address of the host. All IP datagrams to the IP address of the host are forwarded to the loopback address.

### Network Broadcast

The network broadcast entry is a host route (network mask of 255.255.255.255) corresponding to the all-subnets directed broadcast address (all subnets of class B network ID 157.55.0.0). Packets addressed to the all-subnets directed broadcast are sent from the network adapter corresponding to the node's IP address.

### Multicast Address

The multicast address, with its class D network mask, is used to route any multicast IP packets from the network adapter corresponding to the node's IP address.

# IP Address Translation and Routing

By Mark E. Donaldson

## Limited Broadcast

The *limited broadcast address* is a host route (network mask of 255.255.255.255). Packets addressed to the limited broadcast are sent from the network adapter corresponding to the node's IP address.

To view the IP routing table on a Windows 2000-based computer, type **route print** at a Windows 2000 command prompt. When determining the forwarding or next-hop IP address from a route in the routing table:

- If the gateway address is the same as the interface address, the forwarding IP address is set to the destination IP address of the IP packet.
- If the gateway address is not the same as the interface address, the forwarding IP address is set to the gateway address.

For example, when traffic is sent to 157.55.16.48, the most specific route is the route for the directly attached network (157.55.16.0/20). The forwarding IP address is set to destination IP address (157.55.16.48) and the interface is the network adapter, which has been assigned the IP address 157.55.27.90.

When sending traffic to 157.20.0.79, the most specific route is the default route (0.0.0.0/0). The forwarding IP address is set to the gateway address (157.20.16.1) and the interface is the network adapter, which has been assigned the IP address 157.55.27.90.

## Routing Processes

The IP routing processes on all nodes involved in the delivery of an IP packet includes: the sending host, the intermediate routers, and the destination host.

### IP on the Sending Host

When a packet is sent by a sending host, the packet is handed from an upper layer protocol (TCP, UDP, or ICMP) to IP. IP on the sending host does the following:

1. Sets the Time-to-Live (TTL) value to either a default or application-specified value.
2. IP checks its routing table for the best route to the destination IP address. If no route is found, IP indicates a routing error to the upper layer protocol (TCP, UDP, or ICMP).
3. Based on the most specific route, IP determines the forwarding IP address and the interface to be used for forwarding the packet.
4. IP hands the packet, the forwarding IP address, and the interface to Address Resolution Protocol (ARP), and then ARP resolves the forwarding IP address to its media access control (MAC) address and forwards the packet.

### IP on the Router

When a packet is received at a router, the packet is passed to IP. IP on the router does the following:

1. IP verifies the IP header checksum. If the IP header checksum fails, the IP packet is discarded without notification to the user. This is known as a silent discard.

# IP Address Translation and Routing

By Mark E. Donaldson

2. IP verifies whether the destination IP address in the IP datagram corresponds to an IP address assigned to a router interface. If so, the router processes the IP datagram as the destination host (see step 3 in the following “IP on the Destination Host” section).
3. If the destination IP address is not the router, IP decreases the time-to-live (TTL) by 1. If the TTL is 0, the router discards the packet and sends an ICMP Time Expired-TTL Expired message to the sender.
4. If the TTL is 1 or greater, IP updates the TTL field and calculates a new IP header checksum.
5. IP checks its routing table for the best route to the destination IP address in the IP datagram. If no route is found, the router discards the packet and sends an ICMP Destination Unreachable-Network Unreachable message to the sender.
6. Based on the best route found, IP determines the forwarding IP address and the interface to be used for forwarding the packet.
7. IP hands the packet, the forwarding IP address, and the interface to ARP, and then ARP forwards the packet to the appropriate MAC address.

This entire process is repeated at each router in the path between the source and destination host.

## IP on the Destination Host

When a packet is received at the destination host, it is passed up to IP. IP on the destination host does the following:

1. IP verifies the IP header checksum. If the IP header checksum fails, the IP packet is silently discarded.
2. IP verifies that the destination IP address in the IP datagram corresponds to an IP address assigned to the host. If the destination IP address is not assigned to the host, the IP packet is silently discarded.
3. Based on the IP protocol field, IP passes the IP datagram without the IP header to the appropriate upper-level protocol. If the protocol does not exist, ICMP sends a Destination Unreachable-Protocol Unreachable message back to the sender.
4. For TCP and UDP packets, the destination port is checked and the TCP segment or UDP header is processed. If no application exists for the UDP port number, ICMP sends a Destination Unreachable-Port Unreachable message back to the sender. If no application exists for the TCP port number, TCP sends a Connection Reset segment back to the sender.

## Static and Dynamic IP Routers

For IP routing between routers to occur efficiently in the IP internetwork, routers must have explicit knowledge of remote network IDs or be properly configured with a default route. On large IP internetworks, one of the challenges faced by network administrators is how to maintain the routing tables on their IP routers so that IP traffic flow is traveling the best path and is fault tolerant.

There are two ways of maintaining routing table entries on IP routers:

# IP Address Translation and Routing

By Mark E. Donaldson

- Manually - Static IP routers have routing tables that do not change unless manually changed by a network administrator. Static routing relies on the manual administration of the routing table. Remote network IDs are not discovered by static routers and must be manually configured. Static routers are not fault tolerant. If a static router goes down, neighboring routers do not sense the fault and inform other routers.
- Automatically - Dynamic IP routers have routing tables that change automatically based on the communication of routing information with other routers. Dynamic routing employs the use of routing protocols, such as Routing Information Protocol (RIP) and Open Shortest Path First (OSPF), to dynamically update the routing table through the exchange of routing information between routers. Remote network IDs are discovered by dynamic routers and automatically entered into the routing table. Dynamic routers are fault tolerant. If a dynamic router goes down, the fault is sensed by neighboring routers who propagate the changed routing information to the other routers in the internetwork.

## **Physical Address Resolution**

Based on the destination IP address and the route determination process, IP determines the forwarding IP address and interface to be used to forward the packet. IP then hands the IP packet, the forwarding IP address, and the interface, to ARP.

If the forwarding IP address is the same as the destination IP address, then ARP performs a direct delivery. In a direct delivery, the MAC address corresponding to the destination IP address must be resolved.

If the forwarding IP address is not the same as the destination IP address, then ARP performs an indirect delivery. The forwarding IP address is the IP address of a router between the current IP node and the final destination. In an indirect delivery, the MAC address corresponding to the IP address of the router must be resolved.

To resolve a forwarding IP address to its MAC address, ARP uses the broadcasting facility on shared access networking technologies (such as Ethernet or Token Ring) to send out a broadcasted ARP Request frame. An ARP Reply, containing the MAC address corresponding to the requested forwarding IP address, is sent back to the sender of the ARP Request.

## **ARP Cache**

To keep the number of broadcasted ARP Request frames to a minimum, many TCP/IP protocol stacks incorporate an *ARP cache*, a table of recently resolved IP addresses and their corresponding MAC addresses. The ARP cache is checked first before sending an ARP Request frame. Each interface has its own ARP cache.

Depending on the vendor implementation, the ARP cache can have the following qualities:

- ARP cache entries can be dynamic (based on ARP Replies) or static. Static ARP entries are permanent and are manually added using a TCP/IP utility such as the ARP utility provided with Windows 2000. Static ARP cache entries are used to prevent ARP Requests for commonly-used local IP addresses, such as routers and servers. The problem with static ARP entries is that they have to be manually updated when network interface equipment changes.

# IP Address Translation and Routing

By Mark E. Donaldson

- Dynamic ARP cache entries have a time-out value associated with them to remove entries in the cache after a specified period of time. Dynamic ARP cache entries for Windows 2000 TCP/IP are given a maximum time of 10 minutes before being removed.

To view the ARP cache on a Windows 2000–based computer, type **arp -a** at a Windows 2000 command prompt.

## ARP Process

IP sends information to ARP. ARP receives the IP packet, the forwarding IP address, and the interface to be used to forward the packet. Whether performing a direct or indirect delivery, ARP performs the following process, as displayed in Figure 1.15:

- Based on the interface and the forwarding IP address, ARP consults the appropriate ARP cache for an entry for the forwarding IP address. If an entry is found, ARP skips to step 6.
- If the entry is not found, ARP builds an ARP Request frame containing the MAC address of the interface sending the ARP Request, the IP address of the interface sending the ARP Request, and the forwarding IP address. ARP then broadcasts the ARP Request using the appropriate interface.
- All hosts receive the broadcasted frame and the ARP Request is processed. If the receiving host's IP address matches the requested IP address (the forwarding IP address), its ARP cache is updated with the address mapping of the sender of the ARP Request. If the receiving host's IP address does not match the requested IP address, the ARP Request is silently discarded.
- The receiving host formulates an ARP Reply containing the requested MAC address and sends it directly to the sender of the ARP Request.
- When the ARP Reply is received by the sender of the ARP Request, it updates its ARP cache with the address mapping. Between the ARP Request and the ARP Reply, both hosts have each other's address mappings in their ARP caches.
- The IP packet is sent to the forwarding host by addressing it to the resolved MAC address.

