A Self-Configuring and Self-Administering Name System with Dynamic Address Assignment

PAUL HUCK
Oracle Corporation
MICHAEL BUTLER
MITRE Corporation
AMAR GUPTA
Massachusetts Institute of Technology
and
MICHAEL FENG
Oracle Corporation

In this article we present a distributed system that stores name-to-address bindings and provides name resolution to a network of computers. This name system consists of a network of name services that are individually self-configuring and self-administering. The name service consists of an agent program that works in conjunction with the current implementation of the Domain Name System (DNS) program. The DNS agent program automatically configures the Berkeley Internet Name Domain (BIND) process during start-up and dynamically reconfigures and administers the BIND process based on the changing state of the network. The proposed name system offers high scalability and fault-tolerance capabilities and communicates using standard Internet protocols.

Categories and Subject Descriptors: C.2.1 [**Computer Communication Networks**]: Network Architecture and Design—*Distributed networks*; *Network communications*; C.2.3 [**Computer Communication Networks**]: Network Operations—*Network management*; *network monitoring*

General Terms: Algorithms, Design, Performance

Additional Key Words and Phrases: Berkeley Internet Name Domain, dynamic reconfiguration, name-to-name address binding, self-administering systems, self-configuring systems

Authors' addresses: A. Gupta, Sloan School of Management, Rm. E60-309, MIT, Cambridge, MA, 02139; email: agupta@mit.edu; P. Huck, Oracle Corp.; M. Butler, MITRE Corp.; M. Feng, Oracle Corp.

1. INTRODUCTION

As we approach the close of its second decade, the underlying technology of the Internet, specifically the TCP/IP protocol suite, has shown a tremendous resilience that attests to the framers' foresight. Their achievement is even more remarkable when we consider that this architecture was conceived at a time when there were remarkably few network hosts, very low bandwidth, there were relatively few types of digital transactions, computer mobility was unheard of, and corporations and campuses were generally centralized at a single geographic location.

The Internet today is rapidly approaching a billion hosts; link bandwidths vary from a paltry hundred bits per second to gigabits per second; links carry realtime voice, video, and data traffic including financial, medical, and criminal records, which place demands on privacy as well as performance; host mobility is commonplace; and organizations are now geographically distributed but are connected by virtual private networks. With the advent of "Internet ready" embedded processors, we can expect these increases in the Internet's size, diversity, and fluidity to continue.

The Internet's underlying architectural concepts are as sound as ever, but weaknesses (growing pains?) are to be expected as the Internet continues to evolve. One such area is address space management in the presence of host mobility, micronetworks (e.g., home networks), and network reconfiguration. We begin with a brief historical perspective on address management, recap the state of things today, and explore an alternative for the future.

In the infancy of the Internet, address space management was deceptively simple. It was envisioned that all interfaces on all hosts would have unique, permanently assigned IP addresses.[1] Thus, large blocks of addresses were assigned to an institution or a corporation to administer as it saw fit. The size of the allocated block was determined by the organization's anticipated needs and fell into three classes, A, B, and C. This approach worked exceptionally well, since it distributed the management of the four billion available addresses among the participating organizations, giving them the responsibility and authority to manage a portion of the address space. Although such allocations were wasteful,[2] the perception was that the available 32-bit address space was inconceivably large.

A few years ago this inefficient management of address space led to a short-term crisis; it appeared that the then "inconceivably large" address space was being rapidly exhausted. One of the solutions, IPv6, offered a currently "inconceivably large" 128-bit address space as an alternative. Although several IPv6 implementations are now available, IPv6 has not—and in fact might never—displace IPv4 and its paltry address space.

One of the ways that the "address space crisis" in IPv4 was averted was rethinking one of IP's initial assumptions—the desire for a globally unique, permanent address assignment for every host interface. It is arguable whether this change in assumptions was a conscious effort or occurred accidentally, but

[1]Excepting experimental subnets in the 192.168.x.y address space.
[2]A Class A address allocation contains 16 million host addresses.

address space management changed fundamentally with the introduction of DHCP, masquerading firewalls, and VPNs. We now routinely violate the permanency, and even uniqueness, of address assignment, by recycling IP addresses both temporally and spatially.

Such transient assignments have, however, exposed some undesirable artifacts in other protocols; specifically, TCP's concept of a connection is integrally coupled to the invariance of source and destination address as presented in the packet header. Hence, the loss of a DHCP lease (e.g., by a dropped PPP connection) effectively severs any open TCP connections. Mobile-IP is, in large part, a workaround for TCP's dependence on address permanence. It offers TCP the illusion of a permanent IP address, while simultaneously providing routing changes based on the changes in IP address. (With MobileIP, mobile hosts retain their permanent addresses, but their packets are transiently tunneled to the "care of" address of a mobile agent at their point of attachment to the network.)

Dynamic address assignment is, however, an exceedingly powerful technique that offers solutions not only to the shrinking address space problem, but to the larger address space management and network configuration problems as well. Route aggregation, for example, could be optimized if address reassignment were possible. ATM (Asynchronous Transfer Mode) protocols took such dynamic address assignment to its logical conclusion, making all interface addresses locally assigned and explicitly supporting the concept of trunking, thereby avoiding the address space problem in its entirety.

We might ask why both the concepts of interface addresses and host names are preserved. The two techniques are complementary and address different problems. Internet addresses the 32-bit values that appear in IP packets, provide a convenient, fixed-size, shorthand indication of a packet's destination, and so prescribe its migration through the network. When the address space is carefully administered, addresses also provide sufficient information for route aggregation or "trunking," routing traffic based on only a few bits of the address which greatly reduces the size of routing tables in both hosts and routers. This shorthand representation is compact and very convenient for machine processing.

However, numeric addresses, whether permanently or transiently assigned, are an exceptionally inconvenient representation for humans. Dynamically assigned addresses only exacerbate the problem. Host names provide much better mnemonics—as well as a much larger total address space and the opportunity for many-to-many name-to-address mappings as well as dynamic address assignment.

In the Internet, dynamic address assignment is currently primarily applied to end-user machines using either DHCP or Mobile IP, and IP masquerading is used in firewalls at the enterprise boundaries. Virtual servers (where a single virtual URL maps to one of several physical hosts) offer another example of dynamic address binding, though the implementation differs in details. Although such uses have shown the power of dynamic address binding, we still do not exploit dynamic address binding to obviate address management in network design.

Serendipitously, the early inclusion of this mnemonic "crutch," which allowed named hosts, also laid the cornerstone for dynamic address assignment. The mapping between the large and mnemonic name space and the smaller and compact addresses space is universally accomplished by the hierarchically organized DNS servers. In order to reap the benefits of a more dynamically assigned address space; we will need a more fluid name-to-address mapping than presently available.

The remainder of this article addresses the key issues associated with modifying DNS to enable more dynamic address assignment.

2. OVERVIEW

The Autonomous Network Management initiative attempts to develop a collection of protocols and software tools that will enable networks to be configured rapidly and maintained by untrained personnel. Industry efforts such as DHCP and the IETF's ZeroConf Working Group have attempted to address some issues related to automation and configuration, but these are focused exclusively on providing a solution for endhosts and ignore other network components. The approach we take to automate the network planning and management task splits the problem into four interrelated subtasks: address assignment; host name registration; network augmentation; and resource redistribution for load balancing. In this article, we focus on the first two aspects. Our approach concentrates on providing dynamic interface address assignment for all network assets, so that the network components can themselves efficiently renegotiate addresses each time the network is extended. To facilitate the latter, a more flexible DNS hostname-to-address binding scheme must be incorporated.

For more than a decade, the Berkeley Internet Name Domain (BIND) has been the de facto standard for DNS implementations [Shadow IPserver 2001]. Nearly all modern networks use the public domain BIND implementation or commercial products derived from BIND. In fact, BIND has become virtually synonymous with the industry standards that define the DNS architecture. The DNS specification was written to allow product interoperability, and so permit and encourage flexibility in each DNS implementation. Thus, it is possible to develop new DNS implementations with added features, while remaining compliant with existing standards.

In particular, it would be desirable if a DNS service could be run with no lengthy configuration process and no need for user intervention after it has started. In order to function properly, the current implementation of BIND requires a network administrator to write several configuration files. These configuration files tell BIND the zone it is responsible for, the location of other DNS servers on the network, and the location of the root name server. Setting up this configuration can be a tedious process, as every time a new host is added or removed from the network, these files must be updated. In addition, if the network topology changes (such as by adding or removing name servers or by merging two networks), significant changes are required to the configuration of BIND as the DNS zones and ownership of the zones may have changed too. This is certainly not ideal for home networks or networks consisting of mobile or wireless

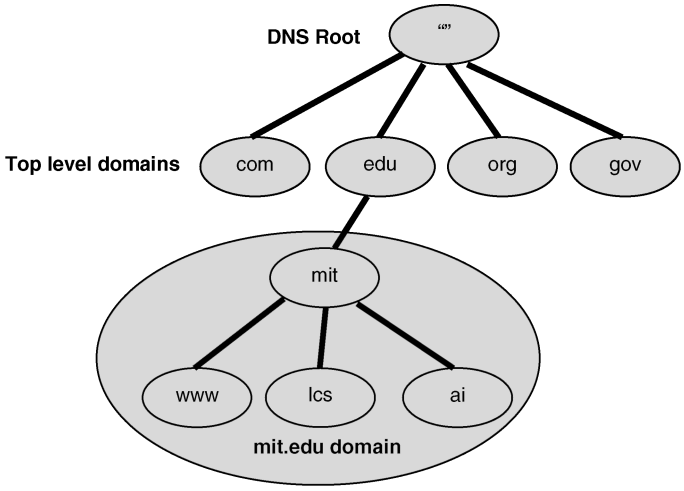

Fig. 1. DNS hierarchy.

clients. In a home networked environment, hosts are frequently leaving and entering the network as appliances are turned on and off, or as new hosts are added. Mobile clients also will often join and leave networks as they are transported across network boundaries. In an ideal environment, hosts should be free to join and leave the network and should immediately gain access to the services; this requires that the DNS be able to configure itself on a dynamic basis.

3. DNS

This section describes the hierarchical naming system and distributed architecture of the current DNS standard. It also examines the operations involved for performing updates to the network.

DNS was developed in 1984 by Paul Mockapetris as a distributed database that could resolve the address of any computer on the network [Mockapetris 1987]. It was created to replace the ASCII host files that associated host names with their respective IP addresses. These host files resided on every host in the network, and if a name was not listed in the file, that host could not be reached. As the Internet grew to become a worldwide network, the process of maintaining the host files became increasingly unwieldy, leading to a growing need for the DNS.

The DNS directory can be thought of as a tree, with each node on the tree representing a “domain” [Albitz and Liu 1998]. Each domain is identified and located in the tree by its domain name, which uses the familiar dotted notation (www.mit.edu, for example.). As we read the domain name from right to left, we can follow the path down the DNS tree to the correct node (see Figure 1). The “edu” domain is one of many top-level domains; others include “com”, “gov”, “org”, and “uk”. The full and unique domain name of any node in the tree is the sequence of labels on the path from that node to the root.

The hierarchical tree of domain names can be referred to as the domain name space. Each domain in the space has host data associated with it. These host

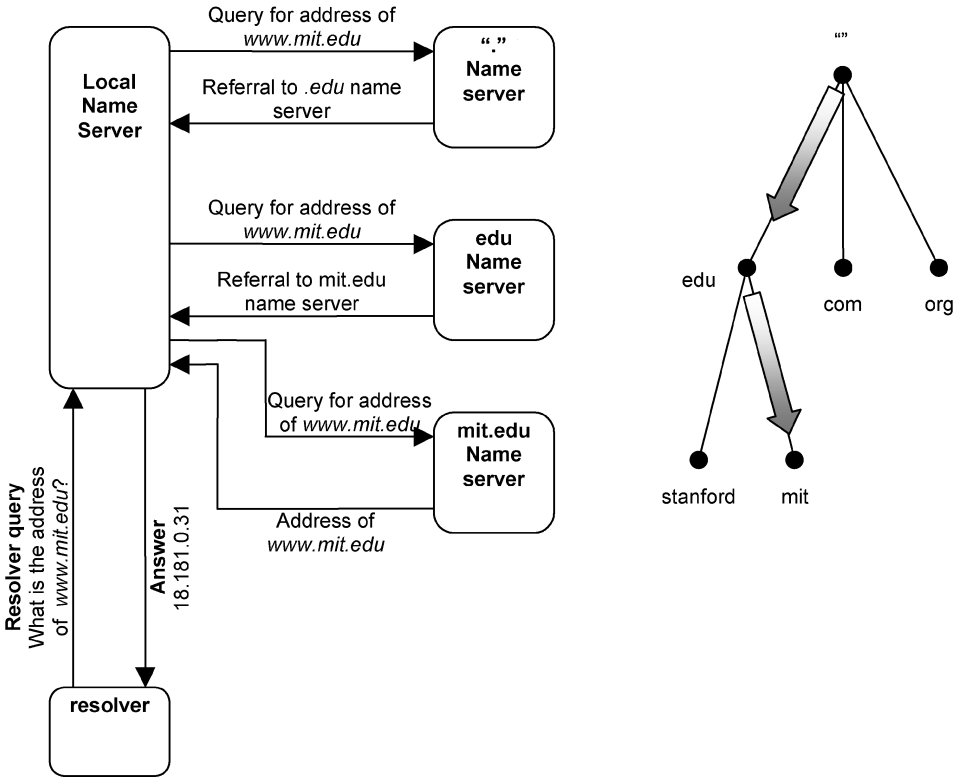

Fig. 2. Resolution of www.mit.edu through DNS (from [Feng 2001]).

data are defined in resource records. There are many different types of resource records—the most common is the address-record (A-record), which associates the domain name with an IP address.

The other distinguishing characteristic of the DNS architecture is its distributed implementation. DNS servers may be operated by any organization that owns a domain. Each DNS server is given authority for one or more "zones." A zone is a collection of one or more DNS domains that share the same parent domain, along with the associated resource records. A DNS server receives authority over a zone when the network manager responsible for the domain that contains that zone delegates the authority to that particular DNS server. Hence the DNS infrastructure is distributed both geographically and administratively [Shadow IPserver 2001].

3.1 Name Servers

Each zone in the domain name space has at least two name servers authoritative for it, a primary one and a secondary one. Authoritative name servers are the only name servers guaranteed to contain the correct resource records for their zone. When querying a name on the Internet, a resolver can only be assured that the address is correct if the answer comes from an authoritative name server for the zone that is being queried (Figure 2). To improve

performance, other name servers may cache resource records, but each cached entry has time-to-live to prevent staleness of data. A name server stores all the resource records for the zone for which it is authoritative. The primary name server is an authoritative server for its zone(s) and reads its zone data from the zone files [Albitz and Liu 1998]. When changes are made to the zone's resource records, they must be made to the primary's zone file. This data is sent to secondary name servers upon request. A secondary name server is also authoritative for its zone(s); however, it obtains its zone data via data transfers from other name servers. A secondary name server periodically interrogates the primary or other secondary servers to determine if its zone's data have changed. The period of this interrogation can be set by a network administrator. A single name server can act as both the primary and secondary server for two or more different zones.

3.2 Dynamic DNS and Zone Transfers

The original DNS specification was written with static networks in mind. It was assumed that hosts would join and leave the network infrequently. However, in modern networks, computers are free to join and leave the network, and many new devices are being connected. In order to deal with dynamically changing networks, several extensions to DNS have been implemented. Specifically, RFC 2136 [Vixie et al. 1997] defines the DNS Update protocol that allows for dynamic updates to the DNS. A dynamic update is a deletion or addition of resource records to a zone. It can be sent by either a DNS client, DNS server, or a DHCP server (see Section 4). The *update* signal is sent to the primary name server, which receives the signal and permanently updates its zone file. The signal may be sent to the primary server directly, or passed through one or more secondary servers, until it reaches the primary. When a primary server fulfills an *update* request, it can use the *notify* signal to inform its secondary servers that the zone information has changed.

In order for multiple name servers to maintain consistency of their records, zone transfers are performed on a periodic basis. A zone transfer is the transfer of resource records from a primary name server to a secondary name server. A full zone transfer occurs when all resource records are sent. Instead of sending all the resource records, it is possible for the primary name server to perform an incremental zone transfer. This will transfer only those records updated since the last zone transfer. A secondary name server requests an incremental zone transfer and a primary server chooses whether it will perform a full zone transfer or an incremental one. It is recommended that full zone transfers be performed no more than once every 15 minutes and at least once every 24 hours [BIND 2001].

The following is a brief example of how the current version of BIND propagates changes to the resource records to all authoritative name servers. The process is illustrated in Figure 3.

(1) A request for the addition or deletion of a host is received by the primary server. This may come from administrator manually editing the zone file, or through an *update* message received. If the *update* message is received

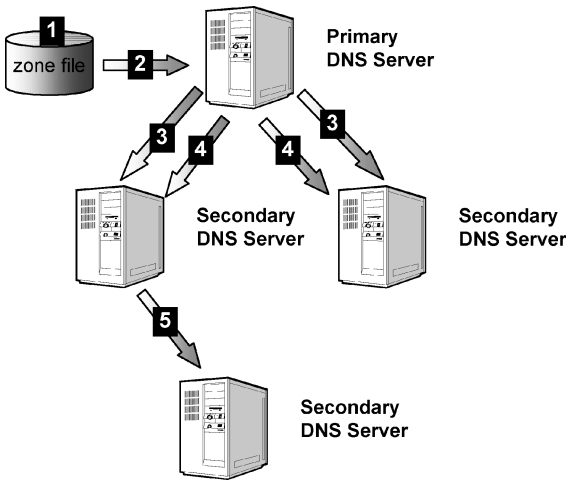

Fig. 3. BIND updates (from [Shadow IPserver 2001]).

at a secondary server, it will pass it to the primary server if it knows its location, or to another secondary server. This is continued until the primary server receives the *update* signal. Once the zone file is changed, the serial number of the file is incremented.

(2) The primary name server reads the edited zone file. The frequency at which the server rereads its zone file and checks for zone changes is a configurable parameter of BIND.

(3) The primary server will send a *notify* message to all known secondary servers. The primary server will wait some time between sending each *notify* to reduce the chance of multiple secondary servers requesting zone transfers at the same time.

(4) If the secondary server(s) support(s) the *notify* signal, a zone transfer is immediately initiated. Otherwise, the secondary server will discard the *notify* and wait until the next scheduled zone transfer time.

(5) The secondary server then notifies any other secondary servers that may be dependent on it for zone transfers. This multi-level transfer is continued until all secondary servers have received the changed records.

While the dynamic update system may lessen the amount of administrative work for the name servers, it does not make them administrator-free. Each name server in the network must be configured with the address of either the primary DNS server or other secondary DNS servers for its zone. Name servers are not free to join and leave the network. If a secondary name server is added to the network, either the primary DNS server or other secondary DNS must be configured to recognize the presence of the new secondary name server so that the primary DNS server can receive *updates* properly. Also, if the primary name server is removed and another added in its place, an administrator must manually change the configuration of every secondary server so that it is informed of the new primary DNS server. In addition, the current DNS system

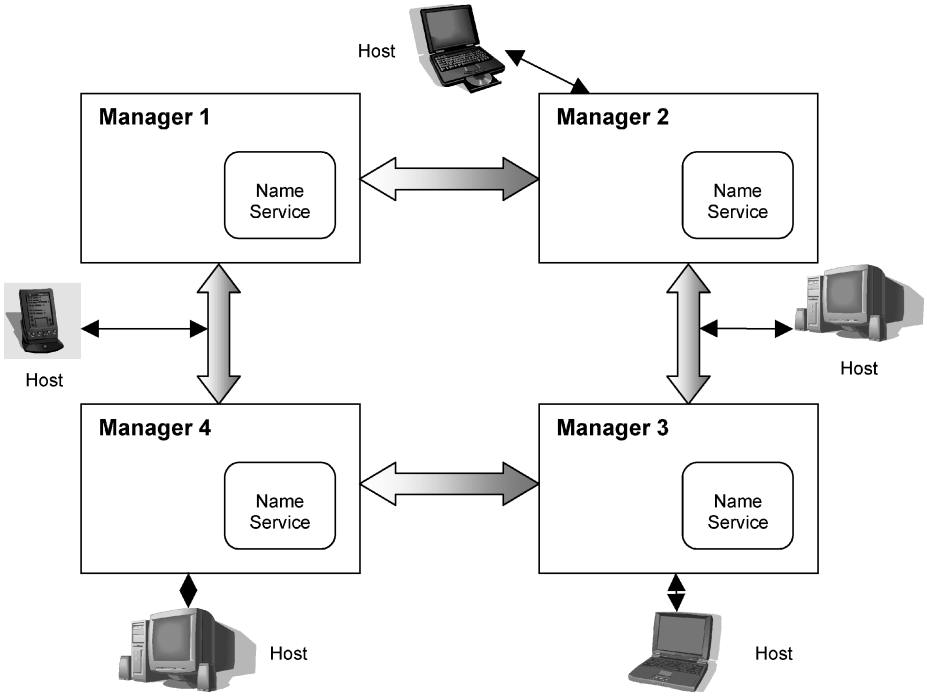

Fig. 4. Sample network.

requires extensive initialization effort to ensure proper operation. Zones must be properly allocated with specific primary and secondary servers and a clear domain hierarchy must be defined. These tasks all require extensive knowledge and effort from an administrator.

The above problems are addressed by the system described in the next section. It uses a combination of an agent program monitoring BIND and dynamic DNS protocols to maintain a true self-configuring and self-administrating naming system.

4. SELF-CONfiGURING AND SELF-ADMINISTERING NAME SYSTEM

4.1 Network Overview

The naming system described in this section was conceived as part of a larger network system that is under development at a company codenamed Research Corporation in this article. The proposed design provides a full array of network services to the hosts on the network. This is done with virtually no manual configuration and no administration. The network design allows for a group of computers to be physically connected (through Ethernet or other network media), and after a short time, they will all be properly configured in a working network. On this network there are two types of nodes: managers and hosts (Figures 3 and 4). Managers form the backbone of the network and provide services, such as name-to-address resolution, to the hosts. Besides the

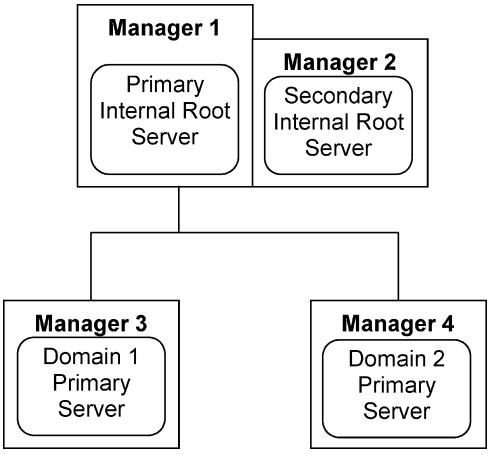

Fig. 5. Name space view of systems in Figure 4.

name-to-address resolution service, the managers also provide dynamic IP configuration for new hosts, discovery mechanisms for new managers, packet routing, and a database of all hosts in the system. The services directly related to the naming service are automatic IP configuration and manager discovery—as these will be the services that the name system will directly interact with. The goal of the network system is to allow any computer, either a host or a manager, the freedom to join or leave the network, without the need for any external administration. The network should be able to detect the presence of a new node or the absence of an existing one and deal with either type of change in an appropriate manner. In addition, the network system, and in particular the name system, should deal gracefully with the merging of two networks. Conflicts should be detected and resolved quickly.

In order to provide the self-configuring and self-administrating name service, each manager in the system is designed to run three types of services: the IP configuration service, the manager-discovery service, and the agent-based name service. The IP configuration service is handled by a standard implementation of the dynamic host configuration protocol (DHCP). DHCP is a network protocol specified by the IETF that provides hosts with configuration parameters for their IP interface [Droms 1997]. This includes an IP address, a domain name, a subnet mask, a default gateway, and a location of DNS server. DHCP also allows a host to retain a previous configured domain name while receiving an IP address. After receiving the necessary information from a DHCP server, a previously unconfigured host's IP interface has all the necessary parameters in order to begin transmitting and receiving on the network. DHCP requires no prior configuration; as a host locates a DHCP server by broadcasting discover messages on the local network. All modern operating systems, and even most embedded devices, support DHCP. Overall, DHCP meets the goals and design objectives of the desired self-configuring and self-administrating network.

The process of manager discovery is having the manager discovery process periodically broadcast "discover packets" to each interface. These packets

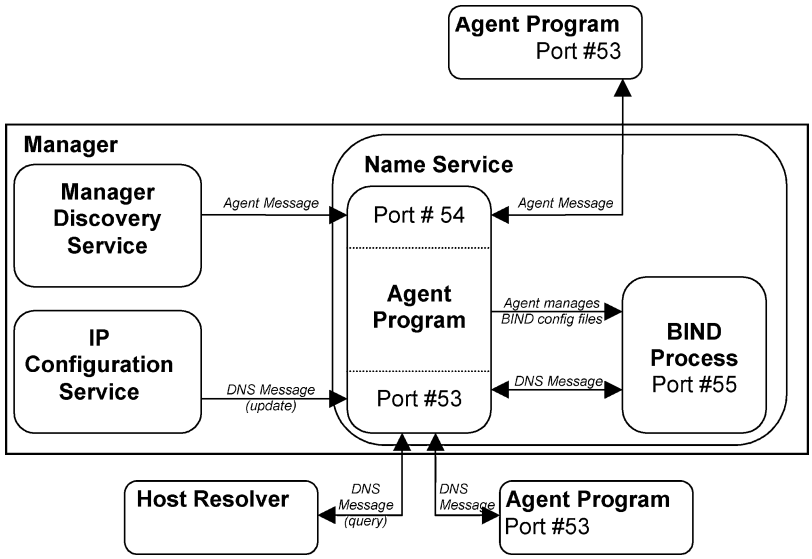

Fig. 6. Relationship among agent, BIND, and other processes (from [Feng 2001]).

contain the source address and a unique network number where the manager resides. All other manager discovery processes will receive the packet, and if the network number is known, the packet will be dropped. If, however, the network number is unknown, that manager will respond to the source of the "discover" with a "discover reply." This reply includes the network number for the new manager, allowing the manager discovery process to inform the local name service of any new managers that appear on the network.

The design and the operation of the name service are discussed in the next section.

4.2 Name Service Interface

The name service runs on every manager in the system and consists of two concurrently running processes. The first is the BIND implementation of DNS. As stated before, BIND is currently used in the overwhelming majority of name servers on the Internet. It provides a scalable, stable, and fault-tolerant basis for the name service. The second process is an agent program that reacts to network conditions and configures BIND automatically. Figure 4 shows the relationship and manner of communication among BIND, the agent program, the other local manager processes, as well as other managers in the network. The agent program uses Berkeley UDP sockets to listen for two different message formats. On port 53, the standard DNS port, the agent program listens for DNS messages. It acts as a filter for the DNS messages, sending queries directly to the BIND process, and processing update messages from the IP configuration process. On port 54, the agent program listens for any agent messages coming from either the manager discovery service or other agents in the system. The agent messages allow an agent process to gain knowledge of other agents and offer a method of communication between agents. The IETF has designated

port 54 to be used for XNS Clearinghouse [Reynolds and Postel 1994]. We chose to use port 54 because XNS is uncommon on modern servers. If this service is needed, another port may be specified for agent communication. The details of both DNS and agent message processing are discussed in the following sections.

4.3 Internal Root Servers

The name servers in the system are only authoritative for the IP addresses of the subset of hosts configured by the naming system. If the network is connected to a larger network, such as the Internet or other networks serviced by the naming system, DNS servers outside the local network must be queried. With this in mind, the name system is implemented using the idea of internal root servers. Certain name services in the manager network act as the internal root servers and are authoritative for the entire domain that is serviced by the self-configuring naming system. These are not true DNS root servers because they are not authoritative for every domain name on the internet, but only for those existing inside the domain serviced by the self-configuring naming service. So as long as a name service knows the location of at least one internal root server, the name service will be able to provide name resolution for the network's name space. For example, if the entire mit.edu domain is using the self-configuring naming system, a host in the lcs.mit.edu subdomain may wish to know the IP address of a computer in another subdomain serviced by the naming system (such as media.mit.edu). This host will query any local name system, and since the manager is not authoritative for the media.mit.edu domain, it will query the internal root server of the mit.edu domain. This server may resolve any domain name in the mit.edu domain by directing the resolver to the proper media.mit.edu authoritative name server. For resolving names outside the domain, such as Internet host names, the Internet root servers may be queried.

4.4 DNS Messages

DNS updates are the only DNS messages that the agent program processes directly. All other messages are simply directed to the local BIND process on port 55. These messages will be DNS queries, and BIND will process the query, return an answer to the agent process, which in turn will forward the response to the inquirer. The agent process is the only entity with access to the BIND process. To all hosts on the network, it appears as though BIND is running on the normal port.

4.4.1 *DNS Update.* DNS updates, however, require special processing by the agent. DNS updates only occur when the IP configuration service or manager discovery service detects the addition or deletion of a host or manager on the network. The respective local manager service then assembles the appropriate DNS update message containing the resource record that must be added, modified, or deleted and sends it to the agent program. Although DNS updates may be sent to any DNS server that is authoritative for a zone, they will always end up being processed by the primary master for the zone, since that is the

Table I. Agent Message Header Fields

Header Field	Description
Message ID	The unique message ID for the message.
Opcode	The operation code for message. Can have one of six values; see Table II.
S/R flag	Send/Response Flag. A flag indicating whether the message is a send request or response to an earlier message.

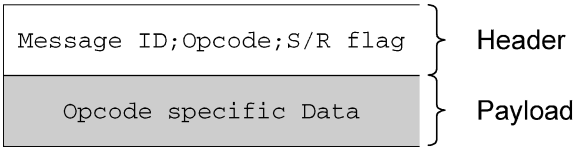

Fig. 7. Agent message format.

only server that has the definitive set of resource records for a particular zone. As stated earlier, standard DNS implementations specify that any secondary server that receives a DNS update message must forward it directly to the primary master if its location is known, or alternatively through the secondary server chain until it reaches the primary master [Vixie et al. 1997]. The secondary servers will only learn of the update through *notify* signals originating from the primary master.

To try to improve the efficiency of this mechanism, the agent program will examine the update message, and if it applies to RR in the zone where it is authoritative, will send the primary master for that zone directly to it. Each agent process has knowledge of its primary master, as this is part of its state information described in Section 4.6. When the agent program is not authoritative for the zone receiving the update, it must search for the primary master for that zone on the network. To do this, the agent program sends a Start of Authority (SOA) DNS query to the local BIND process. The SOA asks the DNS for the address of the primary master for a particular zone. Upon receipt of the SOA, BIND performs the standard DNS resolution to find the information on the zone name. If the zone exists, BIND returns to the agent process the IP address of the primary master for the requested zone, and the agent process forwards the update to that address. If, however, the DNS update message applies to a zone that does not exist on the network, BIND returns the nonexistent domain error flag (NX_DOMAIN), and the agent program configures the local BIND process to make it authoritative for the new zone. The reconfigured BIND process will then process the update request. This allows for dynamic creation of zones, and is especially useful in resolving all naming conflicts when two networks are merged.

4.5 Agent Messages

Agent messages are used for interagent communication and as a mechanism for agent discovery. Figure 7 shows the structure of an agent message. It consists of a three-field header followed by a payload section. The header fields are explained in Table I. The payload holds the data from the message. The data are

Table II. Opcode Values

Opcode	Brief Description
Discover	Informs agent of new managers.
Leave	Informs agent of lost managers.
Getroot	Used to request the location of internal root server.
Becomeslave	Sent to an agent to make it a slave to the sender.
Negotiatemaster	Used to resolve primary master conflicts.
Forceslave	Sent by the winner of two conflicting servers, to force the loser into becoming a slave.

specific to each Opcode. The data sent with each agent message are explained in more detail in the following sections.

The Opcode may take on one of six values, which are described in Table II. Currently, the header is formatted as ASCII text separated by semicolons. A sample agent message header is shown here:

3;getroot;response

The total length of this header is 18 bytes (one byte for each character). However, it is possible to represent each field with a more efficient binary representation. If the message ID were encoded as a 12-bit binary number (allowing for 2^{12} unique IDs), the opcode as a 3-bit number, and the S/R flag as a one-bit field, the header size could be reduced to two bytes. The case of ASCII text was chosen to simplify the implementation and debugging of the system.

4.5.1 *Agent Message Behavior.* The manager discovery process sends the local agent process *discover* and *leave* messages whenever it discovers that a manager has either joined or left the system. These messages include information about the new or lost server such as the zone it is authoritative for and whether it was a master or a slave for that zone. They allow an agent to gain a current view of network topology and can also help warn against conflicts and errors arising in the network. For example, if a manager that was a primary master for a zone disappeared, a slave for that zone would receive the appropriate *leave* message and negotiate with its peers to elect a new primary server for that zone.

The *negotiatemaster* and *forceslave* messages are designed to be in such situations where the *negotiatemaster* is used when two managers discover that they are primary masters for the same zone. This may happen when two previously unconnected networks are physically joined. The two agents will exchange *negotiatemaster* messages and elect a new master. The *negotiatemaster* message includes metrics to help determine the optimal master, such as the current servers number of slaves. The winner of the election then sends a *forceslave* to the loser and requests the latter's zone data, so that such data may be merged with the existing data. The merge is accomplished via a zone transfer from the loser to the winner. If any conflicts are detected during the merge (such as two different hosts with the same name), a new subdomain is automatically created for the loser, and every host listed in the loser's resource records is placed in that subdomain. An example is shown in Figures 8 (a) thru (f).

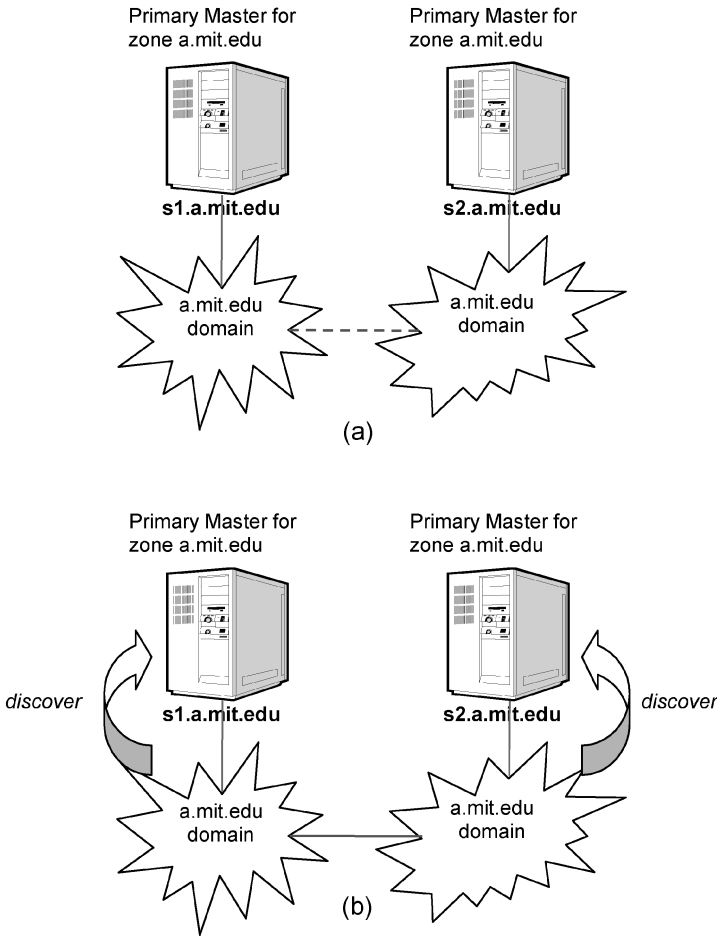

Fig. 8. (a) Merging two previously unconnected networks with name conflicts begins with a new physical connection joining the two; (b) agent programs in both networks receive *discover* messages informing them of the newly discovered managers.

The *becomeslave* message can also make an agent a slave for a zone. However, while the *forceslave* message is only used between two competing primary servers, the *becomeslave* can be sent to any agent in the system. It is not used for elections, but only to notify agents of new primary servers. If a new primary server for a zone leaves the network and a new one is elected, the *becomeslave* message is passed to all the old slaves of the former primary server and to the slaves of the loser in the election. The *becomeslave* message informs them of the new primary server and allows the agents to reconfigure themselves accordingly.

The *getroot* message is used by agents to share the internal root server information. This is useful when a new unconfigured agent is introduced to the system and wishes to know the internal root server. The *getroot* function operates on the basis that any configured DNS server queried with the *getroot*

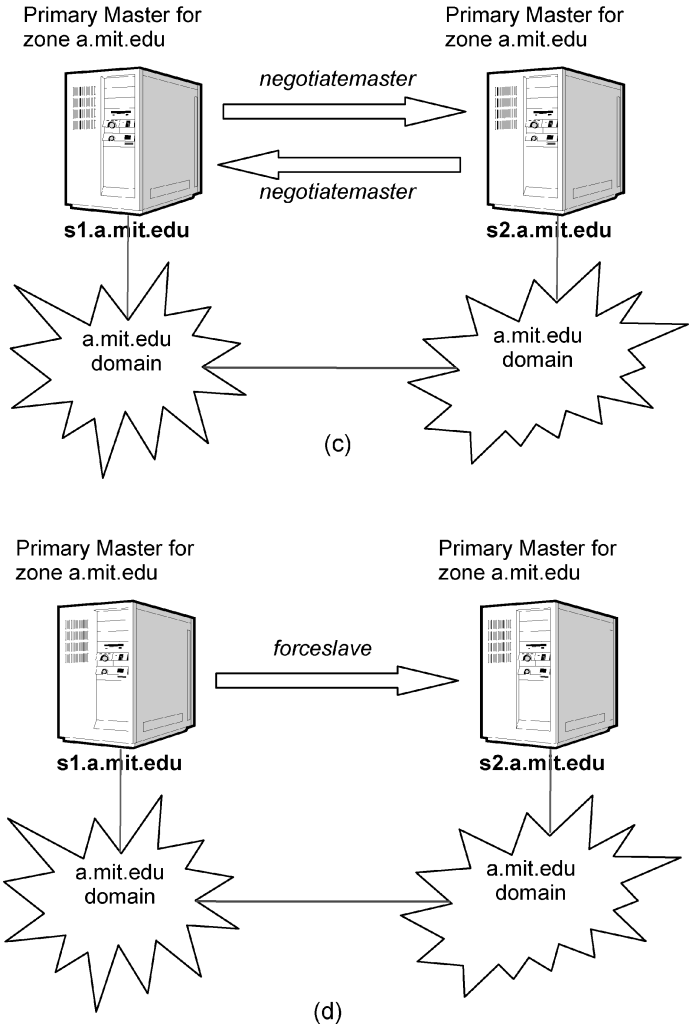

Fig. 8. (c) the conflicting primary master servers send each other *negotiatemaster* messages; (d) the winner of the election sends a *forceslave* message to the loser.

message will have a reference to a root server; this root server can be another server or the root server itself. This assumption is valid because the server is either authoritative or not for the particular root. In the former case, the authoritative root server will simply return its own information to the querying server; in the latter case, it will have a root server reference that will be used to query for the relevant zone information.

4.6 State Information

To assist with the agent tasks, each agent stores state information about itself and other managers in the network, as well as a log of all messages it has received. In particular, the agent keeps a record for all managers that it knows exist on the network. This *server record* contains the name, IP address, a unique

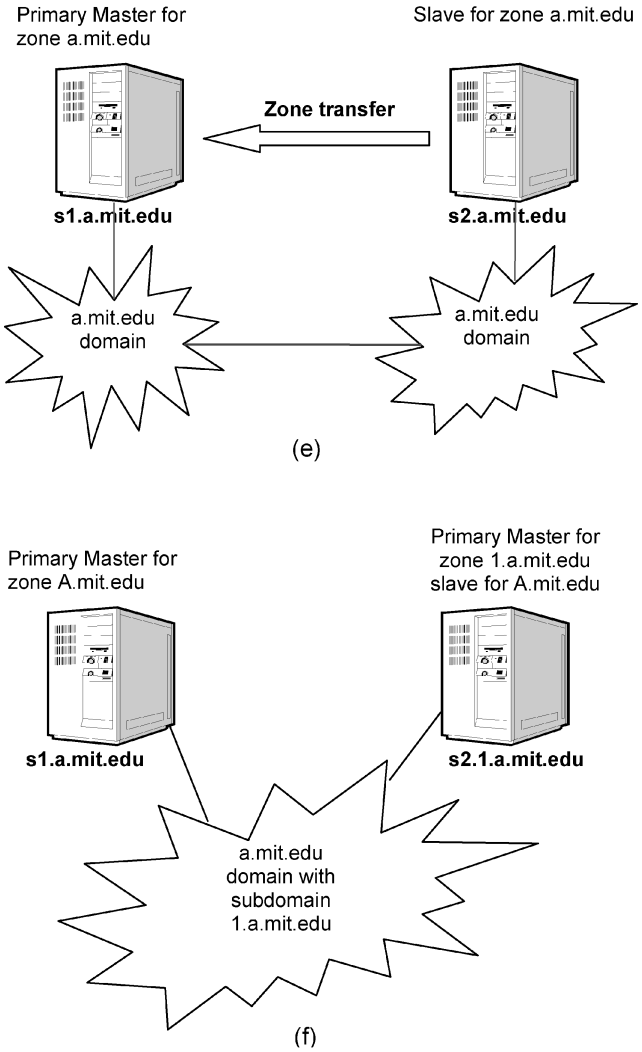

Fig. 8. (e) The new slave performs a zone transfer with the master; (f) because there are name conflicts, a new subdomain is automatically created for every host in the loser's network. Every host in that network is renamed to be a part of the subdomain.

server ID, and a status flag that tells if the server is configured or just starting up (Table III). Each server record is placed into one or more of the following categories:

—*own servers*: contains the agent program's own server information;

—*known servers*: list of server records for every manager on the network;

—*root servers*: server records for every domain root server on the network; and

—*newly discovered servers*: list of recently discovered managers. Once the agent processes the *discover* message, the server record is moved to one of the above categories.

Table III. Server Record Fields

Server Record Field	Description
Name	Name of the manager the name service resides on.
IP Address	IP address of the manager the name service resides on.
Server ID	Unique ID of the manager the name service resides on. The unique ID may be derived from the MAC address of the network interface in use by the manager.
Status Flag	*Configured*: name service is configured with the location of an internal root server. *Start-up*: name service is unaware of a location of an internal root server.

In addition to keeping records for every manager, each agent also keeps information on every zone where it is authoritative. If the server is a slave in that zone, the zone information includes the server record of the primary master for the zone. If the zone is the master for the zone, then the zone information includes the number of slaves for the zone, the *server records* of all the slaves, and the number of slaves required. If the number of slaves is less than the number required, the agent will attempt to send *becomeslave* messages to other agents.

4.7 Operation

At any time, the agent program can be in two different states: the *configured state* and the *start-up state*. The state of the agent is stored in the in the *server record* information for the agent, as described in Section 4.6.

4.7.1 *Initial Configuration.* The key piece of information any agent needs to operate is the location of an internal root server. On start-up, the primary goal of any new agent is to find the location of an internal root server. Once an internal root is found, the agent is put into the configured state. An agent can only be in the configured state if it knows the location of the internal root server, otherwise it is in the unconfigured state.

Figure 9 depicts the start-up scenario and the states of the transition between unconfigured and configured. When a new agent is started, it will wait for *discover* messages from the discovery process on the local manager. When at least one *discover* message is received, the new agent will send one *getroot* message to one of the discovered agents. If a specified timeout period expires and no other manager has been located, the agent will assume it is the only running manager on the system and configure itself to be the internal root. If the new agent hears only *discover* messages from unconfigured managers, then they will compare Server IDs and elect the server with the highest ID to be the internal root server. This may happen when two or more managers join the network at the same time.

4.7.2 *Configured.* Once an agent has entered the configured state, it is ready to handle any name resolution query on the network. Even if it does not have the requested name in its database, it can query the internal root server and find an answer by working recursively down the DNS tree. When

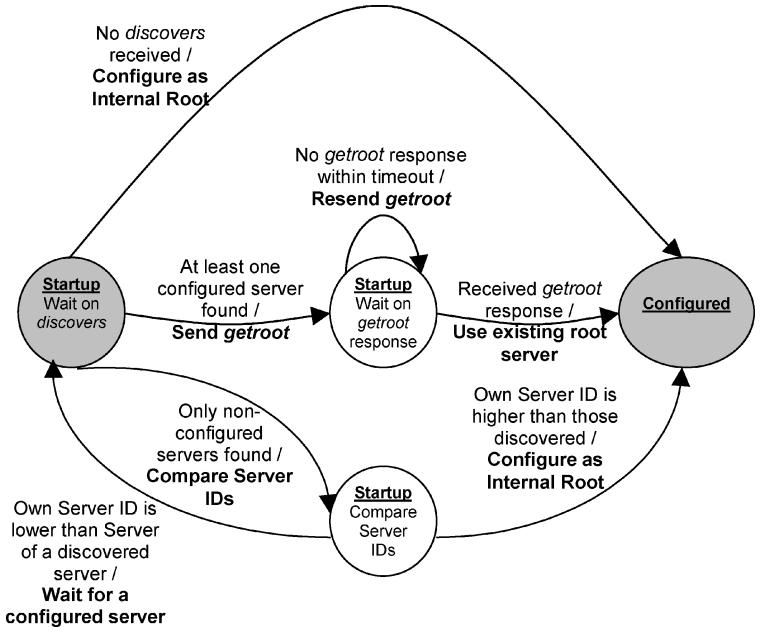

Fig. 9. Start-up scenario.

the agent is in the configured state, it listens for queries, DNS updates, *discover* and *leave* messages, and (if it is a primary master for a zone) periodically runs the getSlave function. The getSlave function is used to find more slaves for a primary master. For every zone where the agent is a primary master, the getSlave function checks the zone information to see if more slaves are needed. If more slaves are needed, the agent picks a random server out of the known servers list that is not already a slave for the zone. It then sends a *becomeslave* message to that server.

The behavior of the agent in response to both DNS queries and updates is described in Section 4.4. The arrival of a *discover* or *leave* message signals a change in network topology because managers have either joined or left the network. In the case of a *discover* message, there are two options: the discovery of a configured manager or the discovery of an unconfigured manager. The case of the discovered unconfigured manager is rather simple, as an agent need only record the *server record* of the new manager and respond to any *getroot* messages it may receive. When a configured manager is discovered, the process is slightly more complex.

Configured managers are discovered when two configured networks are joined. Network unions are revealed by the manager discovery process, as two managers will have different network IDs previously unknown to the other manager. In this case, it is possible to have a conflict where two managers are primary masters for the same zone. Therefore, an agent needs to have a mechanism to detect and resolve this conflict.

Figure 10 illustrates the procedure that is initiated when a configured agent receives *discover* messages. When an agent is a primary master for a zone and

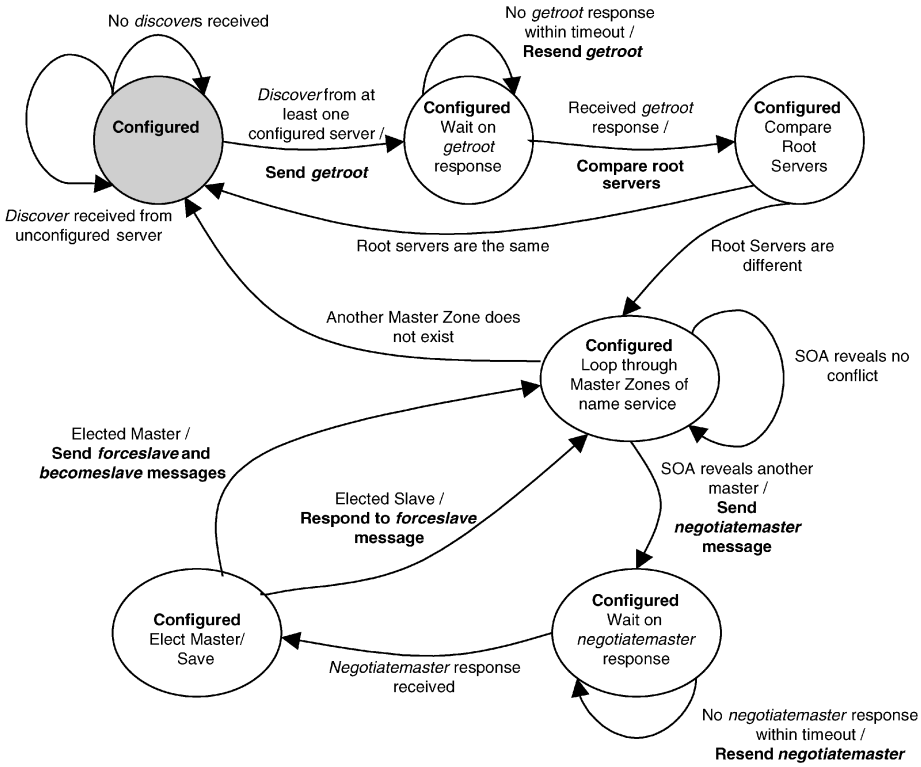

Fig. 10. *Discover* message-handling in a configured state.

receives a *discover* message with information on new configured managers, the agent will send a *getroot* message to the newly discovered server with highest ID (the server ID is used so that only one server is contacted, any other metric could be used as well). Only primary masters need to participate in the conflict-detection scheme, as any slave's master will detect the conflict and transfer the new network information in a future zone transfer.

Zone conflicts can only occur if the two networks have different internal root servers. If both networks shared the same internal root server, it would be impossible for zone conflicts to occur, since a single DNS root tree will not allow the same zone to have two different primary masters. When the two internal root servers are different, an agent must make sure that it is the only primary master for its zone. Therefore, the agent will send a SOA query to the differing root server for every zone that it serves as primary master. The differing root server will then respond with the address of the manager that it considers the primary master for that zone, or with an error message indicating that it is not aware of the zone. If the SOA response is an error message or if the address matches the querying agent, then no conflict exists. If the SOA message is an error, this indicates the other root server is not aware of the zone on the network. However, it can still resolve names in that zone by recursing down the DNS tree from the top node.

If, however, the differing root server believes another manager to be the primary master for the zone, a conflict exists and must be resolved by the *negotiatemaster* master message described in Section 4.5.

Arrival of a *leave* message may also require an agent to reconfigure itself. If a master server loses a slave, it will delete the slave's server record from its state information and send an *update* to the BIND process informing it of the lost name server. If an agent is a slave, it only needs to be concerned with *leave* messages that inform it that its master has left the network. All other *leave* messages will be processed by its master. When a primary master has left the network, all the slaves will become aware of this change and will need to elect a new master for their zone. The slave with the highest server ID becomes the new master. The highest server ID can be calculated by looking at the known servers state information (see Section 4.6). This new primary master will then send *becomeslave* messages to every other slave informing them of their new master.

4.8 Controling the BIND Process

To start the BIND process, the command *named* is run. Before *named* can be run, one configuration file, *named.conf*, and a set of database files need to be configured. The *named.conf* file lists all the zones where the name server is authoritative and whether the name server is a slave or master for that zone. A database file contains all the name information for a zone, and there is a separate database file for each zone in the *named.conf* file. If the *named.conf* file or the database files are changed in any way, a SIGHUP signal can be sent to the *named* process. This signal tells the *named* process to reread its configuration and database files.

The agent program controls BIND with the *named.conf* and database files. During start-up, the agent program uses a Perl script to generate a *named.conf* file and the appropriate database files. Once these files are generated, the agent program automatically start the *named* process. BIND is then ready to answer queries and to handle updates to the name database. Once the *named* process has started, the possibility exists that the name service will need to handle another zone. The agent program will append that zone information to the *named.conf* file, create the appropriate database file, and send a SIGHUP signal to the *named* process. With the agent program in control, BIND will always be configured properly to store names and to handle all the network name resolutions.

4.9 Implementation Notes

The agent program was implemented using Gnu C++ on the Unix operating system. The program consists of a set of classes that control the state of the agent and store network information. The main class for this agent program is the Agent class. The Agent class controls the flow of the program and handles the input and output of agent and DNS messages. The UDP protocol was used to send and receive agent and DNS messages, and the standard UNIX Berkeley socket library was used to create the necessary UDP sockets. The

Agent class retrieved messages by using a select call on each socket to check the availability of data on the socket. Also, callback functions were used to perform time-dependent and periodic events.

The Agent class uses a number of storage classes that store the state of the network. These storage classes include the servers class, the domains class, and the MsgLog class. The servers class helps store the other managers and name services that exist in the network. The domains class helps to store the zones where the agent is authoritative. The MsgLog class stores information on the messages received by the agent program. The C++ standard template library (STL) was used extensively in the implementation of these storage classes. In particular, the map, the list, and the vector template classes proved to be very useful in storing information.

Embedded Perl was used to generate the configuration and database files for *named*. Embedded Perl allows a C++ program to call Perl commands and subroutines. All the Perl commands and subroutines required to generate the *named* files were placed in a file called *agent.pl*, and the agent program used embedded Perl to call the subroutines in *agent.pl* every time the *named* configuration files needed to be changed.

4.10 Security Concerns

The current implementation of the system has no built-in security mechanisms. Securing a network designed to be configuration-free and administrator-free creates a conflict, as any meaningful security mechanism requires some sort of configuration [Williams 2000]. The best that anyone can hope is to accomplish minimize the configuration overhead necessary to keep the system secure. For a closed environment, such as a corporate intranet, perhaps the easiest security model is to have none at all but to simply rely on physical security mechanisms to control access to the machines [Toiranen 2001]. This will work well for an environment that is capably controled and where one can physically control access to the network. If a corporation only allows approved machines and approved users to plug into the Ethernet, it is impossible for an outside intruder to gain access to the network.

However, as wireless networks become increasingly popular, controling physical access to networks becomes impossible [Guttman 2000]. Insecure wireless networks allow anyone in range to send and receive data. This allows theft of services (e.g., an individual using a neighbor's Internet connection), as well as the potential for malicious hosts to be introduced to the system.

Another reason to incorporate security capabilities is When a network managed by the agent-naming system is connected to another larger network such as the Internet. If the system is left insecure, any host in the world could attempt to violate the network's security mechanisms.

When a network is managed by the agent naming system, it should be *at least* as secure as a standard IP network. That is, the agent-naming system should not open up any new holes in the DNS system. Currently, most DNS implementations are completely insecure, relying on redundancy and physical security only [Wellington 2000]. DNS servers are kept in a trusted, safe

location and distributed throughout the network. Normal DNS servers allow administrators at the physical machine only to change the configuration of the DNS server. However, because the agent-naming system allows other agents to change its configuration, it is prone to attacks from outside. If a malicious user could introduce a "rogue agent" into the system, it could do significant damage.

The managers in the agent-naming system have two main functions. First, they provide IP configuration information to new hosts that join the network and second they provide the name to address the resolution service for the network. Each of these services is vulnerable to attack, and for a fully secure network, both must be secured.

There are two possible methodologies for securing the managers: the first is to use IPSec to provide network layer security for all traffic [Kent and Atkinson 1998]; the second is to secure the individual protocols that provide IP configuration and name to address resolution [Droms 2001; Hornstein et al. 2000; Vixie et al. 2000; Wellington 2000].

Using IPSec, all traffic between managers is authenticated and optionally encrypted [Kent and Atkinson 1998]; this ensures that no malicious user can place a rogue manager into the network, for it will not be able to authenticate itself. In addition, hosts can authenticate managers, so that when they receive their IP configuration, they can make sure it is correct and that it came from a known manager.

The other option is to utilize secure versions of DNS and DHCP. RFC 2535 defines DNS extensions that allow for host and DNS server authentication [Eastlake 1999], while RFC 3007 specifies extensions to Eastlake [1999] that will provide authentication for *update* messages [Wellington 2000]. Authentication for DHCP messages is currently being proposed by the IETF [Droms 2001]. These protocols may prove to be lighter weight than IPSec; however, they do not provide confidentiality. This is a deliberate design decision by the IETF because it believes DNS data to be public [Eastlake 1999]. If confidentiality of data is desired, IPSec must be used to encrypt all data. IPSec may also be easier to implement, since it is a broad solution covering all network traffic. The only requirement is that the TCP/IP stacks of both the host and manager support IPSec. When IPSec is not used for security, any additional services added to the system must provide their own mechanism and protocol for authentication.

5. RELATED WORK

5.1 ZEROCONF

The Internet Engineering Task Force's (IETF) ZEROCONF Working Group is currently proposing a protocol to enable networking in the absence of configuration and administration [ZEROCONF 2001]. Although it has yet to propose a specific implementation or specification of the protocols, the WG has set a list of requirements for ZEROCONF protocols [Hattig 2001]. The goal is to allow hosts to communicate using IP without requiring any prior configuration or the presence of network services. Of particular relevance to this article is the name-to-address resolution problem. The ZEROCONF requirements specifically state

that there should be no need for preconfigured DNS or DHCP servers. In fact, the ZEROCONF protocols are required to work in the absence of DNS servers. However, when these services are present, the ZEROCONF requirements direct that hosts should use them. Thus, the ZEROCONF requirements offer temporary and inferior solutions to the name-resolution problem until a complete name resolver such as a DNS server is located.

ZEROCONF protocols face two challenges when determining name-to-address bindings. The first is obtaining a unique address and/or hostname on the network. This is handled extremely well in modern networks by a DHCP server; however, ZERCONF protocols must not rely on the presence of DHCP server. So, the WG recommends using either IPv6 or IPv4 auto-configuration mechanisms [Guttman 2000]. IPv6 auto configuration is vastly superior because it allows hosts to obtain a link-local address (useful only on a single network) using address auto configuration [Guttman 2000] and a routable address by discovering routers using Neighbor Discovery [Narton et al. 1998]. IPv4 auto configuration is still in the research state (with the IETF), but the initial specifications allow a host to get only a link-local address [IETF 2001]. This will prevent the host from communicating with any device not directly on the same network. Also, while IPv6 provides nearly all necessary network parameters such an address, domain name, default router, and DNS server location (if present), IPv4 provides only an address. Thus, if a host configured with IPv4 auto configuration leaves a network and rejoins, it may have a new address, while a host configured by IPv6 will have a permanent method of contact, its domain name. While IPv6 clearly offers more advantages, it is expected that IPv4 will dominate for some time [Cheshire and Aboba 2001], because IPv6 is a relatively new standard and the lack of fully complaint IPv6 routers and hosts on most networks has prevented it from gaining widespread acceptance.

Once a ZEROCONF host has obtained an address on the network, it still needs to discover other hosts and resolve domain names. The ZEROCONF requirements state that hosts should use multicast to resolve names in the absence of a DNS server. In order to support this requirement, an IP host also needs to listen for such requests and to respond when the request corresponds to the host's own name. The IETF currently has two ongoing activities in this area: multicast DNS [Esibov et al. 2000] and "IPv6 Node Information Queries" [Crawford 2000]. In each case, all hosts run a "stub" name service that responds only when it fields a request for its hostname. The stub service does not provide any name to address resolution for other hosts on the network.

While the naming service proposed in this article may fit into the broad goals of the Zero Configuration Working Group, it has several key differences. The most obvious is that the ZEROCONF requirements are designed to work with a network composed entirely of client devices, with no service providers or managers in the network. By design, the agent program is run on a network manager, and provides DNS services for the entire network. While the ZEROCONF requirements specify that hosts need no previous configuration, they do rely on more complete solutions such as DNS and DHCP for long-term operation and scalability. However, the ZEROCONF Working Group does not require that these servers be self-configuring and self-administrating. This is

precisely the problem that the agent program attempts to solve. In a network managed by the self-configuring naming service described in this article, both hosts *and* managers are administrator-free and may join and leave the network freely. Hence it is possible to have a network that runs the agent program on the DNS servers and also meets the ZEROCONF requirements. Hosts can use the ZEROCONF protocols to obtain an address until the discovery of a manager. Once a manager is found, the host is free to resolve any name on the network using standard DNS calls.

5.2 Non-IP Networks

The requirements laid out by the ZEROCONF Working Group stress the importance that computers networked together "just work" in the absence of service providers such as DNS and DHCP [Hattig 2001]. Two protocols in use today provide this level of functionality. The AppleTalk suite of protocols is simple to operate in small networks. Plugging a group of Macs into an Ethernet hub will get a working AppleTalk network without the need to set-up specialized servers like DNS [Guttman 2000]. As a consequence, AppleTalk networks can be used in homes, schoolrooms, and small offices—environments where IP networks have been absent because they are too complicated and costly to administer. NetBIOS provides similar functionality and ease of use on Microsoft Windows machines.

However, because nearly all computers today are connected to the Internet, they also require TCP/IP to be configured, as this is the protocol of the Internet. Therefore, the benefits of AppleTalk and NetBIOS are overshadowed because application developers need to support two protocols: TCP/IP to access the Internet and either AppleTalk or NetBIOS to access the local network. This is the motivation behind our research, as well as that of the ZEROCONF Working Group and other efforts to make the IP suite of protocols simple to configure. Allowing developers to concentrate on one protocol for all communications will make application development simpler and more efficient. In addition, both protocols have problems when scaling to networks the size of the Internet. AppleTalk relies on broadcast packets to name resolution, thus creating a large number of unnecessary packets being sent to each host. A properly configured NetBIOS network can be designed such that no broadcasts are necessary. However, NetBIOS relies on a flat 16-character alphanumeric name space, which will undoubtedly lead to naming conflicts, particularly when networks are merged.

5.3 Easing DNS Administration

The above solutions are all replacements for a full DNS system; either by using other methods of name-to-address mapping over IP or by using different protocols altogether. However, because IP networks and DNS were integrated into nearly all modern network systems, there have been other efforts to ease the configuration and administration of DNS.

Researchers in Japan have attempted to tackle DNS administration problems by simplifying configuration tasks and eliminating repetitive tasks through automation [Giap et al. 1998]. They developed a program with a

graphical interface that reduced the work of a DNS administrator. Their tool, called *nssetup*, automates repetitive tasks such as generating a DNS database file from the machine's host file, keeping the root cache file up-to-date and maintaining the reverse address look-up table. To check the correctness of the configuration, *nssetup* contains a feature that checks if the name server is up and running. In addition, *nssetup* provides a graphical user interface for configuring the resolver and adding new hosts into the database. The *nssetup* developers showed that it is considerably faster to configure a name server with *nssetup* than without it. They state they have reduced the configuration time from two hours to three minutes.

However, *nssetup* does not truly reach the goal of zero administration. It simply provides a nice user interface and a good set of default configuration values for BIND. Every time new DNS servers are added, an administrator must configure them as well as every DNS server already running on the network, so that they are properly integrated into the network. The agent program requires no prior configuration when DNS servers are added. Simply starting the agent program is all that is needed; it will locate other managers on the network and they will configure themselves accordingly, all without user intervention. The time it takes the agent program to configure a name service is less than the time using *nssetup*, and does not require a human administrator.

5.4 Commercial Solutions

There are several commercial products available today that aim to simplify the operation and administration of DNS servers. Most of them provide solutions that include both DHCP and dynamic DNS in a single software package. These products are aimed at corporations that run a large intranet with their own DNS servers. By including both dynamic DNS and DHCP, they allow hosts to freely join and leave a network with no manual configuration. DHCP will assign the new host an IP address and send DNS *update* signals to notify the DNS server of the new host, so that it may be located on the network. Figure 11 shows the integration of DNS and DHCP in the commercial solutions with the addition of remote administration. Some popular products include Lucent QIP Enterprise 5.0, Check Point Meta IP 4.1, and NTS IP Server.

It is important to note that Meta IP includes some changes to the standard DNS protocol to allow for additional functionality. Specifically, the developers of Meta IP introduced a proprietary way for primary and secondary servers to notify other slaves of zone updates. However, recent BIND implementations support the *notify* message, which provides similar functionality [Vixie 1996]. Because the agent name service runs BIND, it also supports this feature.

The NTS IP Server is built from the ground up, and is not based on any previous implementation of BIND. This allowed the developers to implement two proprietary extensions the DNS protocol. The first DNS extension developed by NTS allows servers to have "peer back-ups," which are essentially duplicates of the server, and unlike secondary servers, are always consistent with the main server. Essentially, "peer back-up" allows for more than one primary server. Thus, if one fails, the other is available for updates. The other DNS extension allows for zone "coserving." A zone can be split into a number of pieces, each

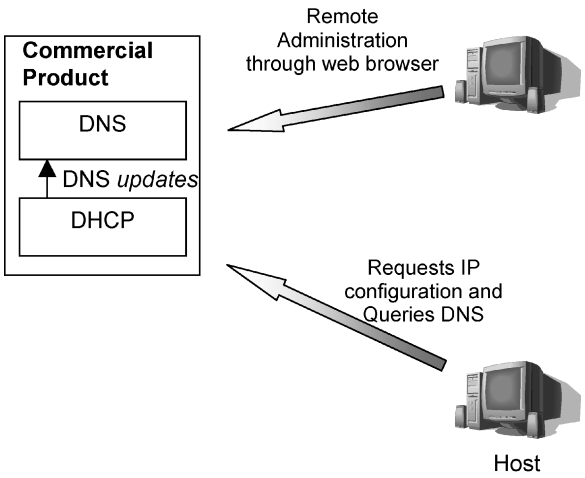

Fig. 11. Integrating DHCP and DNS.

served by a different server. However, any of the servers may be queried and may appear to be primary masters for the zone. They communicate among each other and return the correct answer to the resolver. For larger zones, this may lead to improved performance, as the load can be balanced among a number of different servers. However, these extensions are not Internet standards, and only those servers running the NTS IP Server will be able to use them.

5.4.1 *Comparisons to the Agent Name System.* In addition to the products described above, there exist other similar solutions to ease the administration and operation of DNS [JOIN DDNS 2001; Optivity NetID 2001]. Nearly all of these vendor products offer features that appeal to the corporate environment, including user profile support, directory services, and remote administration. While these features are certainly advantageous in the corporate environment, they do not pertain directly to the name system. In addition, the goal of these commercial solutions is to provide a central point of administration for the entire network.

In contrast, the agent name system takes a decentralized approach to the administration of the network. No one manager is in charge of the entire management, and most network operations, such as creating a new zone and electing a new master, require communication among multiple managers. The decentralized approach allows the relevant portions of the network into the decision-making process.

Another key difference between the agent name system and products described in this section is the procedure for introducing new name servers into the system. The commercial systems and the agent name system both require no administrative input when hosts join and leave the network. However, only the agent name system allows servers to do the same. In all of the above products, extensive reconfiguration is required when a server exits or joins the network.

While most of the commercial products provide a method to handle *update* messages while the primary server is down, none provide a permanent solution

like the agent name system. The commercial products simply store the *updates* until the primary comes back online or allow the secondary server to process them. However, if the secondary also fails, the updates will be lost. In contrast, the agent name system quickly promotes a new server to primary—allowing all *updates* to be recorded permanently. Networks managed by commercial products must be carefully planned, so that the master/slave server relationship exists for all zones. Servers must be configured so that they are aware of the locations of the other servers. In the best case, the commercial products provide a human administrator a nice graphical user interface for performing the necessary configuration. In the worst case, manual editing of the configuration files is needed. Whatever the case, the agent name system provides a simpler solution. No human administration is necessary when name servers exit or join the network, since the managers will communicate among themselves and adapt dynamically to the changing network topology. In addition, the agent name system solution usually works with existing DNS standards and implementations, ensuring compatibility and taking advantage of the proven stability of the existing implementations.

5.5 Performance Comparison

The agent program is able to autonomously configure a DNS process without any human intervention. To accomplish name service configuration, multiple agent programs communicate with each other and exchange information, and based on information from other agents, an agent will generate the *named.conf* file and call the *named* command. No human is required to configure the name services, and configuration is done in a decentralized manner. This is different from existing configuration solutions in industry and research.

Name service solutions in industry require humans to manually install the server software and configure the server as a name server. For example, there is much planning and modeling in the configuration process of the Lucent QIP Enterprise system [Lucent 2001]. In order for the QIP system to work effectively, physical and logical relationships between network objects and users have to be determined. This can be a time-consuming process for a network administrator. With the agent program, the only thing to do is to run the agent program.

5.5.1 *Memory Usage of a Name System.* The two components of the name service, the agent program and the BIND process, took up approximately 2.4% of total machine memory for testing the name system. The machines used for testing all had Intel Pentium 450 MHz processors with 128 Megabytes of RAM. To show program memory usage the agent program was run on a machine and the program *top* was run to give a memory and CPU usage report. The output of the *top* program is shown in Table IV. As shown in table, the agent program occupies 1.7 megabytes of memory and the BIND process, *named*, occupies 1.4 megabytes of memory.

Both processes were idle when *top* was run, and so the CPU usage is zero for both processes. In times of heavy agent and DNS message traffic, the CPU usage is around 5%. During zone transfers, memory usage may go up to two

Table IV. Name Service Memory Usage: *top* Output

Size	RSS	Share	% CPU	% MEM	
1784	1784	1456	0.0	1.3	agent
1448	1448	708	0.0	1.1	named

or three times the idle usage because *named* forks off from new *named* processes to handle zone transfers; but 128 megabytes of RAM should be sufficient to support this extra load. These numbers support the assertion that the name service does not place a large load on a machine. Standard computers should be able to run the agent program and the BIND process without many problems.

6. CONCLUSION

The implementation of the agent program successfully solves the problem of providing a scalable and fault-tolerant way to store name-to-address mappings and to handle name-to-address resolutions in the absence of human configuration or administration. The agent program handles the configuration and administration of a DNS name server implementation called BIND, and BIND stores name information and answers name queries with scalability and fault-tolerance in mind. In comparison to proposals by other researchers, the agent name system offers superior functionality for large networks, and can also be more easily integrated with existing Internet solutions. The name system offers solutions for many of today's networks. It offers a simple solution for homes and small businesses that cannot afford a permanent IT staff to administer their network, while being scalable enough to handle networks of much larger size and complexity.

7. APPENDIX

7.1 Name Service Information

The agent program stores information about all the name services in the system. The agent obtains this information through *discover* messages from the manager and by asking other agents in the system. The information stored for an individual name service is shown below.

Server Record Field	Description
Name	The name of the manager on which the name service resides.
IP Address	The IP address of the manager on which the name service resides.
Server ID	The unique ID of the manager on which the name service resides. The unique ID will be the MAC address of the network interface used by the name service.
Status Flag	A status flag indicating whether or not the name service has configured its root servers. If the name service has configured its root servers, the status flag will have the value: *srv_rs_configured*. If not, then the status flag will have the value: *srv_start_up*.

These four fields make up a *server record*, and this name service information comes from the manager on which the name service resides. For usability, the agent program stores the network's name service information in four separate

groups of server records. The agent program uses these four server groups to quickly look up information about the other servers in the network and to exchange information with other agent programs. The four groups can have overlapping server records, as shown below.

Server Record List Type	Description
Own Servers	This group of server records contains the agent program's own server information.
Known Servers	This group of server records contains the server information for all the other name servers in the network.
Root Servers	This group of server records contains the server information for all the root servers in the network. This information is useful when configuring BIND and sharing root server information among agent programs.
Newly Discovered Servers	This group of server records contains server information for all the name servers that have just been discovered by the discovery portion of the manager. Every time a new name server enters the system, the agent will be notified, and it is here that the new server information is stored. Once the agent processes this newly discovered information, the server record will be deleted from the newly discovered servers and added onto the known servers.

7.2 Zone Information

The agent program controls the zones where the name service is authoritative, and this name service zone information is stored as a data structure in the agent program. The following information about the zone is stored:

1) **Zone Name**—The name of the zone where the name service is authoritative.
2) **Master/Slave Flag**—Indicates the name service is a master or slave for the zone
   a. If the name service is a *master*, the agent also stores:
      i **Slave Information**—the server record of the slaves for the zone;
      ii **Number of Slaves**—the number of slaves that exist for the zone;
      iii **Number Slaves Required**—the number of slaves required for the zone;
      iv **Need More Slaves**—Boolean flag indicating whether or not any more slaves are needed for the zone:
   b. If the name service is a *slave*, the agent also stores:
      i **Master Information**—the server record of the primary master for the zone.
3) If available, the **Parent Zone Name** and **IP address** of the parent zone's name service.

7.3 Agent Messages

The agent message consists of a header section and a payload field. The header section has a message ID, Opcode, and a Send/Response Flag. These three fields are delimited by a semicolon and are in plain text format. The payload section can be of variable size, and each field in the payload section is delimited by a semicolon.

Server Record Format

Name
IP Address
Server ID
Status Flag

Discover Message

Msg_Id	number
Opcode	discover
Send/Response Flag	***Send***
Payload	Server Records of any newly discovered servers.

Leave Message

Msg_Id	number
Opcode	leave
Send/Response Flag	***Send***
Payload	Server Records of any lost servers.

Getroot Message

Msg_Id	number
Opcode	getroot
Send/Response Flag	***Send***
Payload	Empty

Msg_Id	number
Opcode	getroot
Send/Response Flag	***Response***
Payload	Server Records for the root servers in the system.

Becomeslave Message

Msg_Id	number
Opcode	becomeslave
Send/Response Flag	***Send***
Payload	Zone Name the agent should be a slave for.
	Server Record for the master name service.
	Parent Name Service Information: Name and IP Address of Parent Name Service.

Msg_Id	number
Opcode	becomeslave
Send/Response Flag	***Response***
Payload	Zone Name the agent should be a slave for.
	Server Record for the slave name service.

Negotiatemaster Message

Msg_Id	number
Opcode	negotiatemaster
Send/Response Flag	***Send***
Payload	Zone name in conflict
	Number of Slaves the agent has for the zone.
	Server Record of the responding agent.

Msg_Id	number
Opcode	negotiatemaster
Send/Response Flag	**Response**
Payload	Zone name in conflict
	Number of Slaves the agent has for the zone.
	Server Record of the responding agent.

Forceslave Message

Msg_Id	number
Opcode	forceslave
Send/Response Flag	***Send***
Payload	Zone Name the agent is forced be a slave for.
	Server Record for the master name service.
	Parent Name Service Information: Name and IP Address of Parent Name Service.

Msg_Id	number
Opcode	forceslave
Send/Response Flag	***Response***
Payload	Zone Name the agent should be a slave for.
	Server Record for the master name service.
	Server Records for all the slave servers the agent used to be the master for.

ACKNOWLEDGMENTS

We would like to thank Ralph Preston, Sam Wiebenson, Jeffrey Yu, and Kevin Grace for their help and assistance with this paper.

REFERENCES

ALBITZ, P. AND LIU, C. 1998. *DNS and Bind*. 3rd ed., O'Reilly.

AppleTalk Network System Overview. 1990. Addison-Wesley, Reading, MA.

BIND. 2001. *BIND Administrator Reference Manual*, Nominum BIND Development Team, Jan.

CHESHIRE, S. AND ABOBA, B. 2001. Dynamic configuration of IPv4 link-local addresses. Draft-ietf-zeroconf-ipv4-linklocal-02.txt, March. (work in progress).

CRAWFORD, M. 2000. IPv6 node information queries. Draft-ietf-ipngwg-icmp-name-lookups-07.txt. Aug.

DROMS, R. 1997. Dynamic host configuration protocol. RFC 2131, March.

DROMS, R., ED. 2001. Authentication for DHCP messages. Draft-ietf-dhc-authentication-16.txt, Jan. (work in progress).

EASTLAKE, D. 1999a. DNS security operational considerations. RFC 2541, March.

EASTLAKE, D. 1999b. Domain name system security extensions. RFC 2535, March.

ESIBOV, L., ABOBA, B., AND THALER, D. 2000. Multicast DNS. Draft-ietf-dnsext-mdns-00.txt, Nov.

FENG, M. 2001. A self-configuring and self-administering name system. Master's thesis, MIT, Cambridge, MA.

GIAP, C. S., KADOBAYASHI, Y., AND YAMAGUCHI, S. 1998. Zero internet administration approach: The case of DNS. In *Proceedings of the 12th International Conference on Information Networking* (*ICOIN*), 350–355.

GUTTMAN, E. 2000. Zero configuration networking. In *Proceedings of INET 2000*.

HATTIG, M., ED. 2001. *Zeroconf requirements*. Draft-ietf-zeroconf-reqts-07.txt, March (work in progress).

HORNSTEIN, K., ET AL. 2000. DHCP authentication via Kerberos V. Draft-hornstein-dhe-Kerbauth-06.txt, Oct. 2001.

IETF. 2001. Dynamic DNS, infrastructure essential for today's intranets. http://www.nts.com/collateral/ddnstechpaper.pdf.

JOIN DDNS. 2001. http://www.join.com/ddns.html.
KENT, S. AND ATKINSON, R. 1998. Security architecture for the Internet protocol. RFC 2401, Nov.
LUCENT 2001. Lucent QIP Enterprise 5.0: Automating IP services management. http://www.qip.lucent.com/products/qipent_6093.pdf.
META 2001. Meta IP technical white paper. http://www.checkpoint.com/products/metaip/whitepaper.pdf.
MOCKAPETRIS, P. V. 1987a. Domain names—concepts and facilities. RFC 1034, Nov.
MOCKAPETRIS, P. V. 1987b. Domains names—implementation and specification. RFC 1035 Nov.
NARTON, T., NORDMARK, E., AND SIMPSON, W. 1998. Neighbor discovery for IP version 6 (IPv6). RFC 2461, Dec.
Optivity NetID. 2001. http://www.nortelnetworks.com/products/01/unifiedmanagement/collateral/onetid_brief.pdf.
Protocol. 1987. Protocol standard for a NetBIOS service on a TCP/UDP transport: Concepts and methods. RFC 1001, March.
REYNOLDS, J. AND POSTEL, J. 1994. Assigned numbers. RFC 1700, Oct.
SHADOW IPSERVER. 2001. http://www.nts.com/collateral/ipserverdatasheet.pdf.
SIDHU, G. S., ANDREWS, R. F., AND OPPENHEIMER, A. B. 1990. *Inside Appletalk*. 2nd ed., Addison-Wesley, Reading, MA.
STAPP, M. AND REKHTER, Y. 2001. The DHCP client FQDN option. Draft-ietf-dhc-fqdn-option-01.txt, March. (work in progress).
THAYER, R., DORASWAMY, N., AND GLENN, R. 1998. IP security: Document roadmap. RFC 2411, Nov.
TOIVANEN, H. 2001. Secure zero configuration. http://citeseer.nj.nec.com/401221.html.
VIXIE, P. 1996. A mechanism for prompt notification of zone changes (DNS NOTIFY). RFC 1996, Aug.
VIXIE, P., ET AL. 1997. Dynamic updates in the domain name system (DNS update). RFC 2136, April.
VIXIE, P., ET AL. 2000. Secret key transaction authentication for DNS (TSIG). RFC 2845, May.
WELLINGTON, B. 2000. Secure domain name system (DNS) dynamic update. RFC 3007, Nov.
WILLIAMS, A. 2000. Securing zeroconf networks. Draft-williams-zeroconf-security-00.txt. Nov. (work in progress).
ZEROCONF 2001. Zero configuration networking (zeroconf) charter. http://www.ietf.org/html.charters/zeroconf-charter.html.
HUCK, P. 2001. Zero configuration name services for IP networks. Master's thesis, MIT, Cambridge, MA.

Received June 2001; revised September 2001; accepted October 2001