

# The Internet Protocol *Journal*

September 2001

Volume 4, Number 3

*A Quarterly Technical Publication for  
Internet and Intranet Professionals*

## In This Issue

From the Editor .....	1
MPLS.....	2
A Unique Root.....	15
Book Review.....	29
Call for Papers .....	31
Fragments .....	32

## FROM THE EDITOR

*Multiprotocol Label Switching* (MPLS) is a technology that has received a great deal of attention in recent years. The IETF alone has produced over 300 Internet Drafts and numerous RFCs related to MPLS and continues its work on refining the standards. So, what is MPLS all about? We asked Bill Stallings to give us a basic tutorial.

The tragic events of September 11, 2001 have focused attention on the stability and robustness of the Internet. The Internet played an important role in the aftermath of the terrorist attacks. While popular news Web sites initially appeared overloaded, a great deal of private traffic in the form of instant messaging and e-mail took place. Companies directly or indirectly affected by the events in New York and Washington were quick to use the Web as a way to disseminate important information to their clients as well as to their employees. In many cases, the Internet was used in place of an overloaded telephone network. With this in mind, The *Internet Corporation for Assigned Names and Numbers* (ICANN) has decided to re-focus its next meeting to address issues of Internet stability and security, particularly with regard to naming and addressing. (See "Fragments," page 32.) To provide some background information, we bring you the article "A Unique, Authoritative Root for the DNS," by M. Stuart Lynn, the president and CEO of ICANN. Since this article has been posted for public comment, you are encouraged to address your feedback to: [comments@icann.org](mailto:comments@icann.org)

We would like to remind our readers to send us postal address updates. The computer-communications industry is one where people change jobs and locations often. While we do receive some address changes automatically when mail is returned to us, it is much more reliable to send us e-mail with the new information. In the near future, readers will be able to make address changes and select delivery options through a Web interface which will be deployed at <http://www.cisco.com/ipj>. Until then, please send your updates to [ipj@cisco.com](mailto:ipj@cisco.com)

—Ole J. Jacobsen, Editor and Publisher  
[ole@cisco.com](mailto:ole@cisco.com)

You can download IPJ  
back issues and find  
subscription information at:  
[www.cisco.com/ipj](http://www.cisco.com/ipj)

# MPLS

by William Stallings

**M**ultiprotocol Label Switching (MPLS) is a promising effort to provide the kind of traffic management and connection-oriented *Quality of Service* (QoS) support found in *Asynchronous Transfer Mode* (ATM) networks, to speed up the IP packet-forwarding process, and to retain the flexibility of an IP-based networking approach.

## Background

The roots of MPLS go back to numerous efforts in the mid-1990s to combine IP and ATM technologies. The first such effort to reach the marketplace was IP switching, developed by Ipsilon. To compete with this offering, numerous other companies announced their own products, notably Cisco Systems (Tag Switching), IBM (aggregate route-based IP switching), and Cascade (IP Navigator). The goal of all these products was to improve the throughput and delay performance of IP, and all took the same basic approach: Use a standard routing protocol such as *Open Shortest Path First* (OSPF) to define paths between endpoints; assign packets to these paths as they enter the network; and use ATM switches to move packets along the paths. When these products came out, ATM switches were much faster than IP routers, and the intent was to improve performance by pushing as much of the traffic as possible down to the ATM level and using ATM switching hardware.

In response to these proprietary initiatives, the *Internet Engineering Task Force* (IETF) set up the MPLS working group in 1997 to develop a common, standardized approach. The working group issued its first set of Proposed Standards in 2001. Meanwhile, however, the market did not stand still. The late 1990s saw the introduction of many routers that are as fast as ATM switches, eliminating the need to provide both ATM and IP technology in the same network.

Nevertheless, MPLS has a strong role to play. MPLS reduces the amount of per-packet processing required at each router in an IP-based network, enhancing router performance even more. More significantly, MPLS provides significant new capabilities in four areas that have ensured its popularity: QoS support, traffic engineering, *Virtual Private Networks* (VPNs), and multiprotocol support. Before turning to the details of MPLS, we briefly examine each of these.

### Connection-Oriented QoS Support

Network managers and users require increasingly sophisticated QoS support for numerous reasons. The following are key requirements:

- Guarantee a fixed amount of capacity for specific applications, such as audio/video conference
- Control latency and jitter and ensure capacity for voice
- Provide very specific, guaranteed, and quantifiable service-level agreements, or traffic contracts
- Configure varying degrees of QoS for multiple network customers

A connectionless network, such as in IP-based internetwork, cannot provide truly firm QoS commitments. A *Differentiated Service* (DS) framework works in only a general way and upon aggregates of traffic from numerous sources. An *Integrated Services* (IS) framework, using the *Resource Reservation Protocol* (RSVP), has some of the flavor of a connection-oriented approach, but is nevertheless limited in terms of its flexibility and scalability. For services such as voice and video that require a network with high predictability, the DS and IS approaches, by themselves, may prove inadequate on a heavily loaded network. By contrast, a connection-oriented network has powerful traffic-management and QoS capabilities. MPLS imposes a connection-oriented framework on an IP-based internet and thus provides the foundation for sophisticated and reliable QoS traffic contracts.

### Traffic Engineering

MPLS makes it easy to commit network resources in such a way as to balance the load in the face of a given demand and to commit to differential levels of support to meet various user traffic requirements. The ability to dynamically define routes, plan resource commitments on the basis of known demand, and optimize network utilization is referred to as *traffic engineering*.

With the basic IP mechanism, there is a primitive form of automated traffic engineering. Specifically, routing protocols such as OSPF enable routers to dynamically change the route to a given destination on a packet-by-packet basis to try to balance load. But such dynamic routing reacts in a very simple manner to congestion and does not provide a way to support QoS. All traffic between two endpoints follows the same route, which may be changed when congestion occurs. MPLS, on the other hand, is aware of not just individual packets, but flows of packets in which each flow has certain QoS requirements and a predictable traffic demand. With MPLS, it is possible to set up routes on the basis of these individual flows, with two different flows between the same endpoints perhaps following different routers. Further, when congestion threatens, MPLS paths can be rerouted intelligently. That is, instead of simply changing the route on a packet-by-packet basis, with MPLS, the routes are changed on a flow-by-flow basis, taking advantage of the known traffic demands of each flow. Effective use of traffic engineering can substantially increase usable network capacity.

### VPN Support

MPLS provides an efficient mechanism for supporting VPNs. With a VPN, the traffic of a given enterprise or group passes transparently through an internet in a way that effectively segregates that traffic from other packets on the internet, providing performance guarantees and security.

### Multiprotocol Support

MPLS, which can be used on many networking technologies, is an enhancement to the way a connectionless IP-based internet is operated, requiring an upgrade to IP routers to support the MPLS features. MPLS-enabled routers can coexist with ordinary IP routers, facilitating the introduction of evolution to MPLS schemes. MPLS is also designed to work in ATM and Frame Relay networks. Again, MPLS-enabled ATM switches and MPLS-enabled Frame Relay switches can be configured to coexist with ordinary switches. Furthermore, MPLS can be used in a pure IP-based internet, a pure ATM network, a pure Frame Relay network, or an internet that includes two or even all three technologies. This universal nature of MPLS should appeal to users who currently have mixed network technologies and seek ways to optimize resources and expand QoS support.

For the remainder of this discussion, we focus on the use of MPLS in IP-based internets, with brief comments about formatting issues for ATM and Frame Relay networks.

### MPLS Operation

An MPLS network or internet consists of a set of nodes, called *Label Switched Routers* (LSRs), that are capable of switching and routing packets on the basis of a label which has been appended to each packet. Labels define a flow of packets between two endpoints or, in the case of multicast, between a source endpoint and a multicast group of destination endpoints. For each distinct flow, called a *Forwarding Equivalence Class* (FEC), a specific path through the network of LSRs is defined. Thus, MPLS is a connection-oriented technology. Associated with each FEC is a traffic characterization that defines the QoS requirements for that flow. The LSRs do not need to examine or process the IP header, but rather simply forward each packet based on its label value. Therefore, the forwarding process is simpler than with an IP router.

Figure 1: MPLS Operation

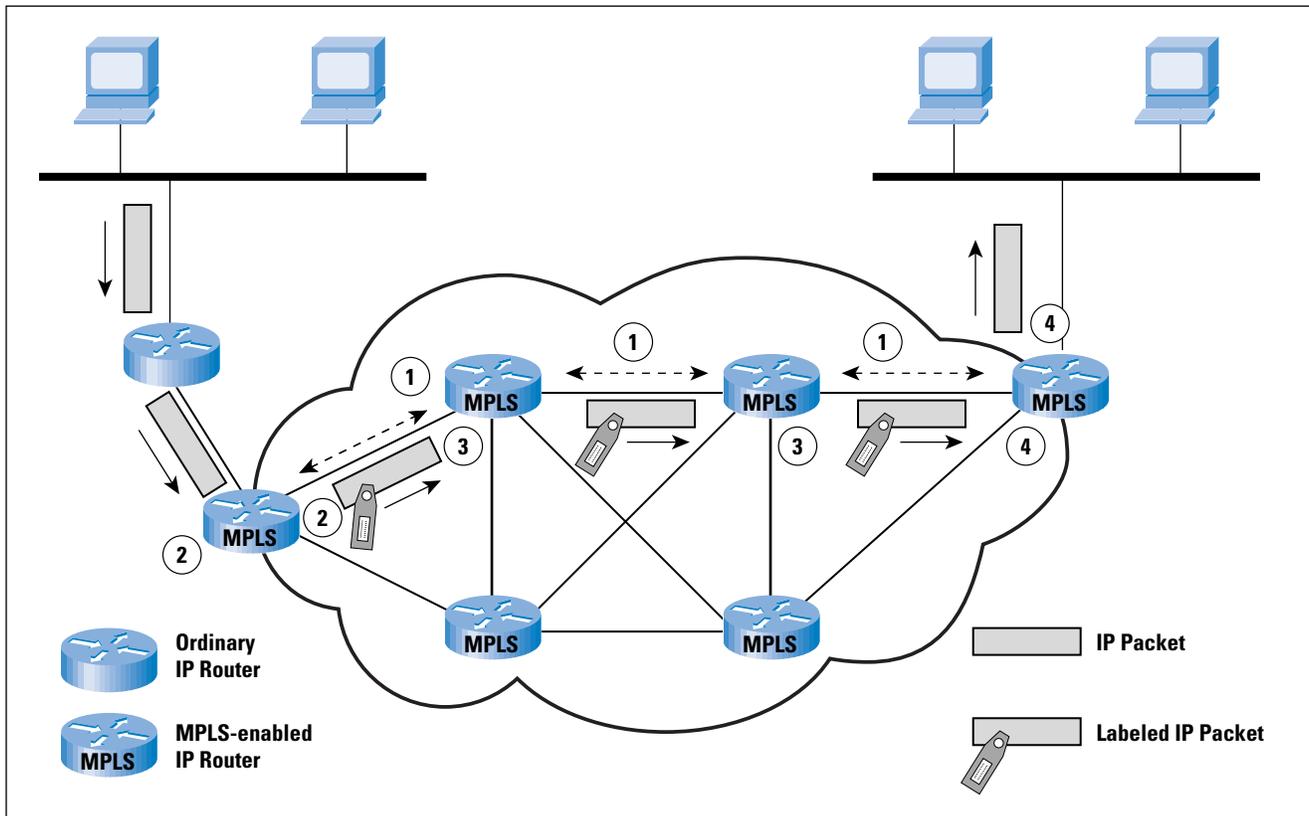


Figure 1, based on one in<sup>[4]</sup>, depicts the operation of MPLS within a domain of MPLS-enabled routers. The following are key elements of the operation.

1. Prior to the routing and delivery of packets in a given FEC, a path through the network, known as a *Label Switched Path (LSP)*, must be defined and the QoS parameters along that path must be established. The QoS parameters determine (1) how many resources to commit to the path, and (2) what queuing and discarding policy to establish at each LSR for packets in this FEC. To accomplish these tasks, two protocols are used to exchange the necessary information among routers:
  - (a) An interior routing protocol, such as OSPF, is used to exchange reachability and routing information.
  - (b) Labels must be assigned to the packets for a particular FEC. Because the use of globally unique labels would impose a management burden and limit the number of usable labels, labels have local significance only, as discussed subsequently. A network operator can specify explicit routes manually and assign the appropriate label values. Alternatively, a protocol is used to determine the route and establish label values between adjacent LSRs. Either of two protocols can be used for this purpose: the *Label Distribution Protocol (LDP)* or an enhanced version of RSVP.

2. A packet enters an MPLS domain through an ingress edge LSR where it is processed to determine which network-layer services it requires, defining its QoS. The LSR assigns this packet to a particular FEC, and therefore a particular LSP, appends the appropriate label to the packet, and forwards the packet. If no LSP yet exists for this FEC, the edge LSR must cooperate with the other LSRs in defining a new LSP.
3. Within the MPLS domain, as each LSR receives a labeled packet, it:
  - (a) Removes the incoming label and attaches the appropriate outgoing label to the packet.
  - (b) Forwards the packet to the next LSR along the LSP.
4. The egress edge LSR strips the label, reads the IP packet header, and forwards the packet to its final destination.

Several key features of MLSP operation can be noted at this point:

1. An MPLS domain consists of a contiguous, or connected, set of MPLS-enabled routers. Traffic can enter or exit the domain from an endpoint on a directly connected network, as shown in the upper-right corner of Figure 1. Traffic may also arrive from an ordinary router that connects to a portion of the internet not using MPLS, as shown in the upper-left corner of Figure 1.
2. The FEC for a packet can be determined by one or more of a number of parameters, as specified by the network manager. Among the possible parameters:
  - Source or destination IP addresses or IP network addresses
  - Source or destination port numbers
  - IP protocol ID
  - Differentiated services codepoint
  - IPv6 flow label
3. Forwarding is achieved by doing a simple lookup in a predefined table that maps label values to next-hop addresses. There is no need to examine or process the IP header or to make a routing decision based on destination IP address.
4. A particular *Per-Hop Behavior* (PHB) can be defined at an LSR for a given FEC. The PHB defines the queuing priority of the packets for this FEC and the discard policy.
5. Packets sent between the same endpoints may belong to different FECs. Thus, they will be labeled differently, will experience different PHB at each LSR, and may follow different paths through the network.

Figure 2: MPLS Packet Forwarding

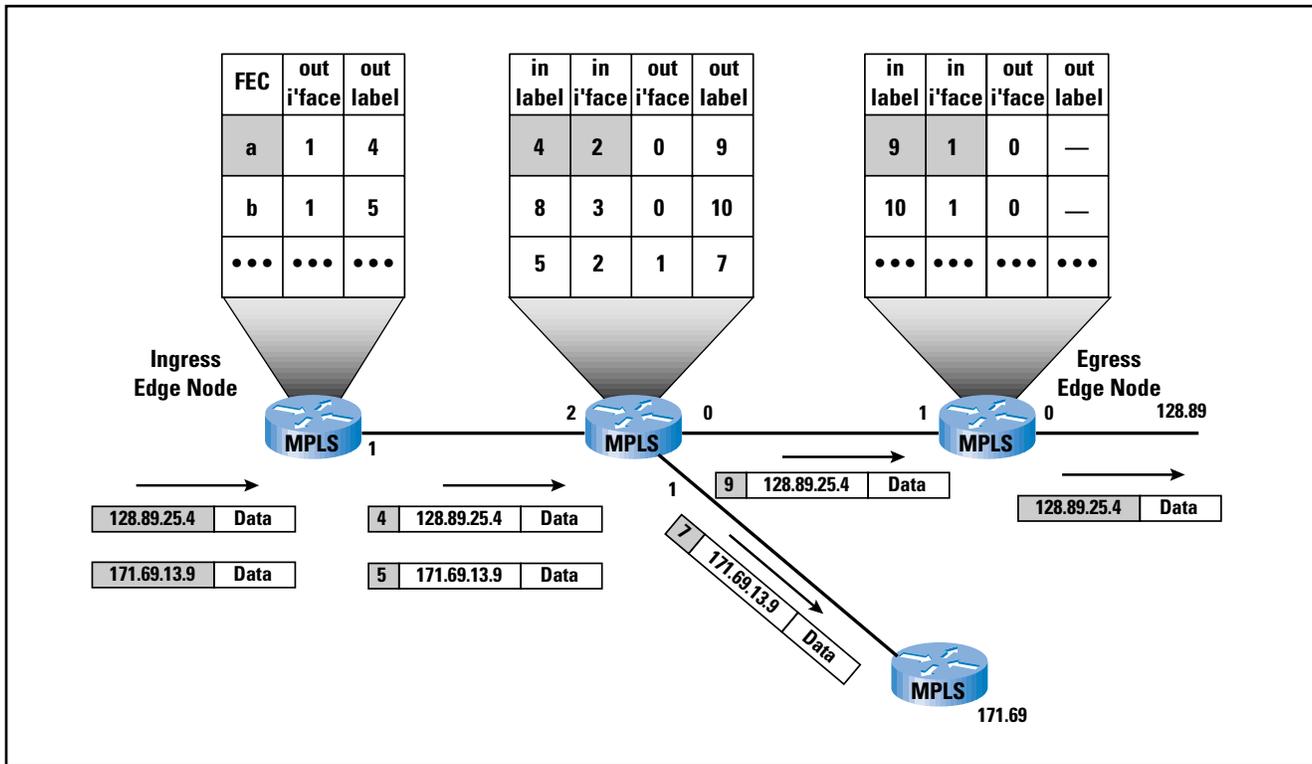


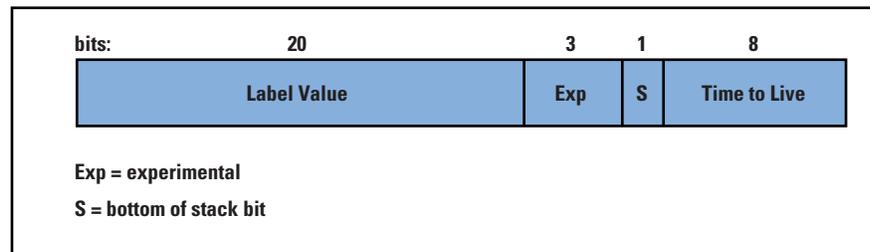
Figure 2 shows the label-handling and label-forwarding operation in more detail. Each LSR maintains a forwarding table for each LSP passing through the LSR. When a labeled packet arrives, the LSR indexes the forwarding table to determine the next hop. For scalability, as was mentioned, labels have local significance only. Thus, the LSR removes the incoming label from the packet and attaches the matching outgoing label before forwarding the packet. The ingress-edge LSR determines the FEC for each incoming unlabeled packet and, on the basis of the FEC, assigns the packet to a particular LSP, attaches the corresponding label, and forwards the packet.

### Label Stacking

One of the most powerful features of MPLS is *label stacking*. A labeled packet may carry many labels, organized as a last-in-first-out stack. Processing is always based on the top label. At any LSR, a label may be added to the stack (push operation) or removed from the stack (pop operation). Label stacking allows the aggregation of LSPs into a single LSP for a portion of the route through a network, creating a *tunnel*. At the beginning of the tunnel, an LSR assigns the same label to packets from a number of LSPs by pushing the label onto the stack of each packet. At the end of the tunnel, another LSR pops the top element from the label stack, revealing the inner label. This is similar to ATM, which has one level of stacking (virtual channels inside virtual paths), but MPLS supports unlimited stacking.

Label stacking provides considerable flexibility. An enterprise could establish MPLS-enabled networks at various sites and establish numerous LSPs at each site. The enterprise could then use label stacking to aggregate multiple flows of its own traffic before handing it to an access provider. The access provider could aggregate traffic from multiple enterprises before handing it to a larger service provider. Service providers could aggregate many LSPs into a relatively small number of tunnels between points of presence. Fewer tunnels means smaller tables, making it easier for a provider to scale the network core.

Figure 3: MPLS Label Format



### Label Format and Placement

An MPLS label is a 32-bit field consisting of the following elements (Figure 3):

- *Label value*: locally significant 20-bit label
- *Exp*: 3 bits reserved for experimental use; for example, these bits could communicate DS information or PHB guidance
- *S*: set to one for the oldest entry in the stack, and zero for all other entries
- *Time To Live (TTL)*: 8 bits used to encode a hop count, or time to live, value

### Time-to-Live Processing

A key field in the IP packet header is the TTL field (IPv4), or Hop Limit (IPv6). In an ordinary IP-based internet, this field is decremented at each router and the packet is dropped if the count falls to zero. This is done to avoid looping or having the packet remain too long in the internet because of faulty routing. Because an LSR does not examine the IP header, the TTL field is included in the label so that the TTL function is still supported. The rules for processing the TTL field in the label are as follows:

1. When an IP packet arrives at an ingress edge LSR of an MPLS domain, a single label stack entry is added to the packet. The TTL value of this label stack entry is set to the value of the IP TTL value. If the IP TTL field needs to be decremented, as part of the IP processing, it is assumed that this has already been done.

When an MPLS packet arrives at an internal LSR of an MPLS domain, the TTL value in the top label stack entry is decremented.

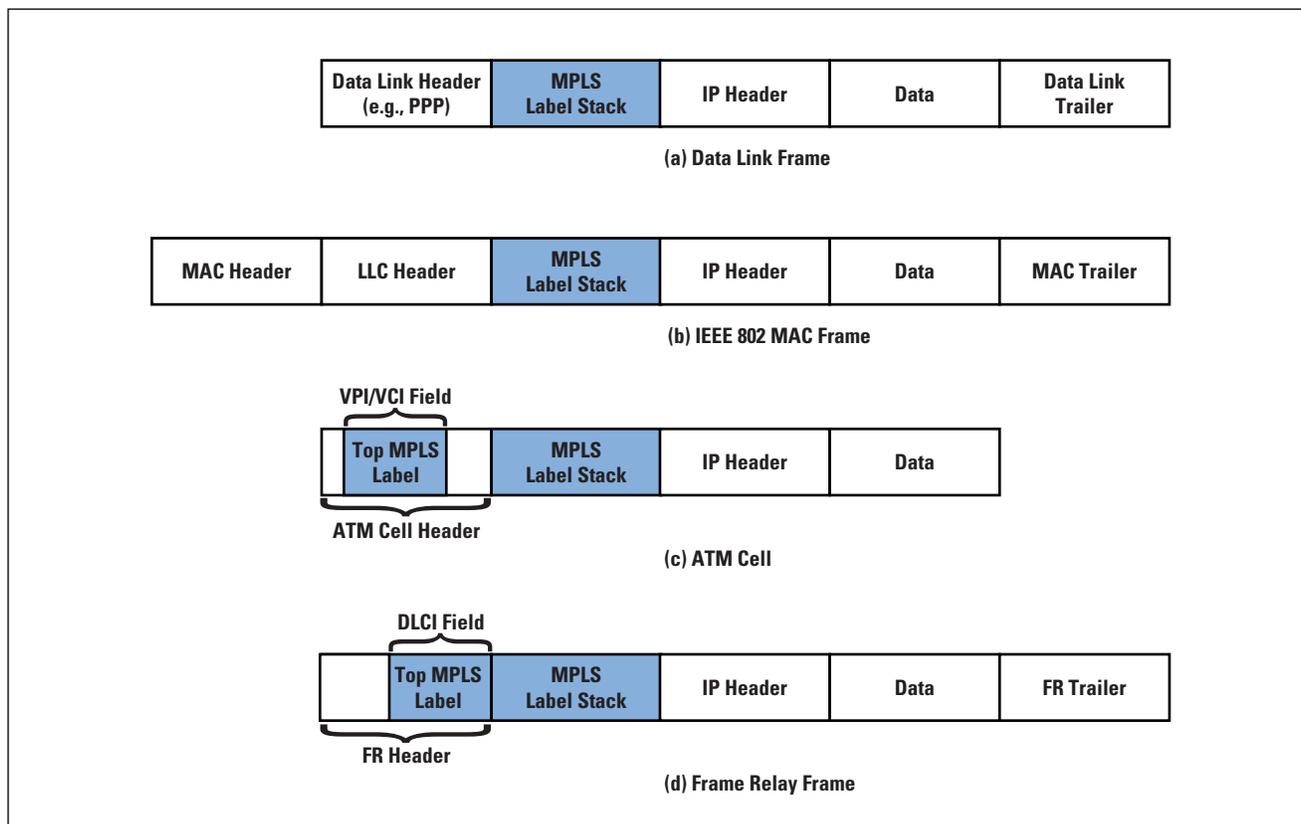
Then:

- (a) If this value is zero, the MPLS packet is not forwarded. Depending on the label value in the label stack entry, the packet may be simply discarded, or it may be passed to the appropriate “ordinary” network layer for error processing (for example, for the generation of an *Internet Control Message Protocol* [ICMP] error message).
  - (b) If this value is positive, it is placed in the TTL field of the top label stack entry for the outgoing MPLS packet, and the packet is forwarded. The outgoing TTL value is a function solely of the incoming TTL value, and is independent of whether any labels are pushed or popped before forwarding. There is no significance to the value of the TTL field in any label stack entry that is not at the top of the stack.
2. When an MPLS packet arrives at an egress edge LSR of an MPLS domain, the TTL value in the single label stack entry is decremented and the label is popped, resulting in an empty label stack. Then:
- (a) If this value is zero, the IP packet is not forwarded. Depending on the label value in the label stack entry, the packet may be simply discarded, or it may be passed to the appropriate “ordinary” network layer for error processing.
  - (b) If this value is positive, it is placed in the TTL field of the IP header, and the IP packet is forwarded using ordinary IP routing. Note that the IP header checksum must be modified prior to forwarding.

### **Label Stack**

The label stack entries appear after the data link layer headers, but before any network layer headers. The top of the label stack appears earliest in the packet (closest to the network layer header), and the bottom appears latest (closest to the data link header). The network layer packet immediately follows the label stack entry that has the *S* bit set. In a data link frame, such as for the *Point-to-Point Protocol* (PPP), the label stack appears between the IP header and the data link header (Figure 4a). For an IEEE 802 frame, the label stack appears between the IP header and the *Logical Link Control* (LLC) header (Figure 4b).

Figure 4: Position of MPLS Label



If MPLS is used over a connection-oriented network service, a slightly different approach may be taken, as shown in Figure 4c and d. For ATM cells, the label value in the topmost label is placed in the *Virtual Path/Channel Identifier* (VPI/VCI) field in the ATM cell header. The entire top label remains at the top of the label stack, which is inserted between the cell header and the IP header. Placing the label value in the ATM cell header facilitates switching by an ATM switch, which would, as usual, need to look only at the cell header. Similarly, the topmost label value can be placed in the *Data Link Connection Identifier* (DLCI) field of a Frame Relay header. Note that in both these cases, the TTL field is not visible to the switch and so is not decremented. The reader should consult the MPLS specifications for the details of the way this situation is handled.

#### FECs, LSPs, and Labels

To understand MPLS, it is necessary to understand the operational relationship among FECs, LSPs, and labels. The specifications covering all the ramifications of this relationship are lengthy. In the remainder of this section, we provide a summary.

The essence of MPLS functionality is that traffic is grouped into FECs. The traffic in an FEC transits an MPLS domain along an LSP. Individual packets in an FEC are uniquely identified as being part of a given FEC by means of a *locally significant label*.

At each LSR, each labeled packet is forwarded on the basis of its label value, with the LSR replacing the incoming label value with an outgoing label value.

The overall scheme described in the previous paragraph imposes numerous requirements. Specifically:

1. Traffic must be assigned to a particular FEC.
2. A routing protocol is needed to determine the topology and current conditions in the domain so that a particular LSP can be assigned to an FEC. The routing protocol must be able to gather and use information to support the QoS requirements of the FEC.
3. Individual LSRs must become aware of the LSP for a given FEC, must assign an incoming label to the LSP, and must communicate that label to any other LSR that may send it packets for this FEC.

The first requirement is outside the scope of the MPLS specifications. The assignment needs to be done either by manual configuration, by means of some signaling protocol, or by an analysis of incoming packets at ingress LSRs. Before looking at the other two requirements, let us consider the topology of LSPs. We can classify these in the following manner:

- *Unique ingress and egress LSR*: In this case a single path through the MPLS domain is needed.
- *Unique egress LSR, multiple ingress LSRs*: If traffic assigned to a single FEC can arise from different sources that enter the network at different ingress LSRs, then this situation occurs. An example is an enterprise intranet at a single location but with access to an MPLS domain through multiple MPLS ingress LSRs. This situation would call for multiple paths through the MPLS domain, probably sharing a final few hops.
- *Multiple egress LSRs for unicast traffic*: RFC 3031 states that most commonly, a packet is assigned to a FEC based (completely or partially) on its network layer destination address. If not, then it is possible that the FEC would require paths to multiple distinct egress LSRs. However, more likely, there would be a cluster of destination networks, all of which are reached via the same MPLS egress LSR.
- *Multicast*: RFC 3031 lists multicast as a subject for further study.

### Route Selection

Route selection refers to the selection of an LSP for a particular FEC. The MPLS architecture supports two options: hop-by-hop routing and explicit routing.

With *hop-by-hop routing*, each LSR independently chooses the next hop for each FEC. The RFC implies that this option makes use of an ordinary routing protocol, such as OSPF.

This option provides some of the advantages of MPLS, including rapid switching by labels, the ability to use label stacking, and differential treatment of packets from different FECs following the same route. However, because of the limited use of performance metrics in typical routing protocols, hop-by-hop routing does not readily support traffic engineering or policy routing (defining routes based on some policy related to QoS, security, or some other consideration).

With *explicit routing*, a single LSR, usually the ingress or egress LSR, specifies some or all of the LSRs in the LSP for a given FEC. For strict explicit routing, an LSR specifies all of the LSRs on an LSP. For loose explicit routing, only some of the LSRs are specified. Explicit routing provides all the benefits of MPLS, including the ability to do traffic engineering and policy routing.

Explicit routes can be selected by configuration, that is, set up ahead of time, or dynamically. Dynamic explicit routing would provide the best scope for traffic engineering. For dynamic explicit routing, the LSR setting up the LSP would need information about the topology of the MPLS domain as well as QoS-related information about that domain. An MPLS traffic engineering specification<sup>[2]</sup> suggests that the QoS-related information falls into two categories:

- A set of attributes associated with an FEC or a collection of similar FECs that collectively specify their behavioral characteristics
- A set of attributes associated with resources (nodes, links) that constrain the placement of LSPs through them

A routing algorithm that accounts for the traffic requirements of various flows and the resources available along various hops and through various nodes is referred to as a *constraint-based routing algorithm*. In essence, a network that uses a constraint-based routing algorithm is aware of current utilization, existing capacity, and committed services at all times. Traditional routing algorithms, such as OSPF and the *Border Gateway Protocol* (BGP), do not employ a sufficient array of cost metrics in their algorithms to qualify as constraint-based.

Furthermore, for any given route calculation, only a single cost metric (for instance, number of hops, delay) can be used. For MPLS, it is necessary either to augment an existing routing protocol or to deploy a new one. For example, an enhanced version of OSPF has been defined<sup>[1]</sup> that provides at least some of the support required for MPLS. Examples of metrics that would be useful to constraint-based routing include the following:

- Maximum link data rate
- Current capacity reservation
- Packet loss ratio
- Link propagation delay

### Label Distribution

Route selection consists of defining an LSP for an FEC. A separate function is the actual setting up of the LSP. For this purpose, each LSR on the LSP must:

1. Assign a label to the LSP to be used to recognize incoming packets that belong to the corresponding FEC.
2. Inform all potential upstream nodes (nodes that will send packets for this FEC to this LSR) of the label assigned by this LSR to this FEC, so that these nodes can properly label packets to be sent to this LSR.
3. Learn the next hop for this LSP and learn the label that the downstream node (LSR that is the next hop) has assigned to this FEC. This process will enable this LSR to map an incoming label to an outgoing label.

The first item in the preceding list is a local function. Items 2 and 3 must be done either by manual configuration or by using some sort of label distribution protocol. Thus, the essence of a label distribution protocol is that it enables one LSR to inform others of the label/FEC bindings it has made. In addition, a label distribution protocol enables two LSRs to learn each other's MPLS capabilities. The MPLS architecture does not assume a single label distribution protocol but allows for multiple such protocols. Specifically, RFC 3031 refers to a new label distribution protocol and to enhancements to existing protocols, such as RSVP and BGP, to serve the purpose.

The relationship between label distribution and route selection is complex. It is best to look at in the context of the two types of route selection.

With hop-by-hop route selection, no specific attention is paid to traffic engineering or policy routing concerns, as we have seen. In such a case, an ordinary routing protocol such as OSPF is used to determine the next hop by each LSR. A relatively straightforward label distribution protocol can operate using the routing protocol to design routes.

With explicit route selection, a more sophisticated routing algorithm must be implemented, one that does not employ a single metric to design a route. In this case, a label distribution protocol could make use of a separate route selection protocol, such as an enhanced OSPF, or incorporate a routing algorithm into a more complex label distribution protocol.

## References

The two most important defining documents for MPLS are [5] and [6]. Reference [3] provides a thorough treatment of MPLS; [8] covers not only MPLS but other Internet QoS concepts; it includes an excellent chapter on MPLS traffic engineering. Reference [7] includes a concise overview of the MPLS architecture and describes the various proprietary efforts that preceded MPLS.

- [1] Apostolopoulos, G., et al., “QoS Routing Mechanisms and OSPF Extensions,” RFC 2676, August 1999.
- [2] Awduche, D., et al. “Requirements for Traffic Engineering over MPLS,” RFC 2702, September 1999.
- [3] Black, U., *MPLS and Label Switching Networks*, ISBN 0130158232, Prentice Hall, 2001.
- [4] Redford, R., “Enabling Business IP Services with Multiprotocol Label Switching,” Cisco White Paper, July 2000 ([www.cisco.com](http://www.cisco.com)).
- [5] Rosen, E., et al. “Multiprotocol Label Switching Architecture,” RFC 3031, January 2001.
- [6] Rosen, E., et al. “MPLS Label Stack Encoding,” RFC 3032, January 2001.
- [7] Viswanathan, A., et al., “Evolution of Multiprotocol Label Switching,” *IEEE Communications Magazine*, May 1998.
- [8] Wang, Z., *Internet QoS: Architectures and Mechanisms for Quality of Service*, ISBN 1558606084, Morgan Kaufmann, 2001.

## Useful Web Sites

- *MPLS Forum*: An industry forum to promote MPLS:  
<http://www.mplsforum.org/>
- *MPLS Resource Center*: Clearinghouse for information on MPLS:  
<http://www.mplsrc.com/>
- *MPLS Working Group*: Chartered by IETF to develop standards related to MPLS. The Web site includes all relevant RFCs and Internet Drafts:  
<http://www.ietf.org/html.charters/mpls-charter.html>

WILLIAM STALLINGS is a consultant, lecturer, and author of over a dozen books on data communications and computer networking. He also maintains a computer science resource site for CS students and professionals at [WilliamStallings.com/StudentSupport.html](http://WilliamStallings.com/StudentSupport.html). He has a PhD in computer science from M.I.T. His latest book is *Wireless Communications and Networks* (Prentice Hall, 2001). His home in cyberspace is [WilliamStallings.com](http://WilliamStallings.com) and he can be reached at [ws@shore.net](mailto:ws@shore.net)

# A Unique, Authoritative Root for the DNS

by M. Stuart Lynn, ICANN

The following *Internet Coordination Policy* (ICP) is being posted for the information of the Internet community by the *Internet Corporation for Assigned Names and Numbers* (ICANN) and is a statement of policy currently followed in administering the authoritative root of the Domain Name System. Comments on this article are welcome and should be directed to [comments@icann.org](mailto:comments@icann.org)

## Abstract

This article reaffirms ICANN's commitment to a single, authoritative public root for the Internet *Domain Name System* (DNS) and to the management of that unique root in the public interest according to policies developed through community processes. This commitment is founded on the technical and other advice of the community and is embodied in existing ICANN policy.

The DNS is intended to provide a convenient means of referring to sites available on the Internet. By offering users an easy-to-use and reliable means of unambiguously referring to Web sites, e-mail servers, and the Internet's many other services, the DNS has helped the Internet achieve its promise as a global communications medium for commerce, research, education, and cultural and other expressive activities.

The DNS is a globally distributed database of domain name (and other) information. One of its core design goals is that it reliably provides the same answers to the same queries from any source on the public Internet, thereby supporting predictable routing of Internet communications. Achievement of that design goal requires a globally unique public name space derived from a single, globally unique DNS *root*.

Although the Internet allows a high degree of decentralized activities, coordination of the assignment function by a single authority is necessary where unique parameter values are technically required. Because of the uniqueness requirement, the content and operation of the DNS root must be coordinated by a central entity.

Where central coordination is necessary, it should be performed by an organization dedicated to serving the public interest and that acts according to policies developed through processes that are developed through the participation of affected stakeholders. Traditionally, the responsibility for performing the central coordinating functions of the global Internet for the public good, including management of the unique public DNS root, has been carried out by the *Internet Assigned Numbers Authority* (IANA)<sup>[12]</sup>. ICANN's core mission is to continue the work of the IANA in a more formalized and globally representative framework, to ensure the views of all the Internet's stakeholders are taken into account in carrying out this public trust.

Over the past several years, some private organizations have established DNS roots as alternates to the authoritative root. Some uses of these alternate roots do not jeopardize the stability of the DNS. For example, some are purely private roots operating inside institutions and are carefully insulated from the DNS. Others are purely experimental in the best traditions of the Internet and are carefully managed so as not to interfere with the operation of the DNS. These both operate within community-established norms.

Frequently, however, these alternate roots have been established to support top-level or pseudo-top-level domain name registries that are operated for profit. Yet other alternate roots have been established by certain individuals to protest the policies developed by the broader community processes for management of the authoritative root, or to express their disinterest in participating in those processes. These alternate roots have not been launched through any ICANN consensus processes, so they have not been entered into the authoritative root managed by the IANA or ICANN.

These alternate roots typically substitute insular concerns in place of the community-based processes that govern the management of the authoritative root. Their operators decide to include particular top-level domains in these alternate roots that have not been subjected to the tests of community support and conformance with consensus processes—coordinated by ICANN—that would allow their inclusion in the authoritative root. These decisions of the alternate-root operators have been made without any apparent regard for the fundamental public-interest concern of Internet stability. The widespread use of active domain names in these alternate roots could in fact impair the uniqueness of the authoritative name-resolution mechanism and hence the stability of the DNS.

ICANN's mandate to preserve stability of the DNS requires that it avoid encouraging the proliferation of these alternate roots that could cause conflicts and instability. This means that ICANN continues to adhere to community-based processes in its decisions regarding the content of the authoritative root. Within its current policy framework, ICANN can give no preference to those who choose to work outside of these processes and outside of the policies engendered by this public trust.

None of this precludes experimentation done in a manner that does not threaten the stability of name resolution in the authoritative DNS. Responsible experimentation is essential to the vitality of the Internet. Nor does it preclude the ultimate introduction of new architectures that may ultimately obviate the need for a unique, authoritative root. But the translation of experiments into production and the introduction of new architectures require community-based approaches, and are not compatible with individual efforts to gain proprietary advantage.

### The Technical Need for a Single Authoritative Root

The DNS was originally deployed in the mid-1980s<sup>[13]</sup> as an improved means of mapping easy-to-remember names (i.e., **example.com**) to the IP addresses (i.e., **128.9.176.32**) by which packets are routed on the Internet. It is a distributed database that holds this mapping information (as well as various other types of technical information regarding computers on the Internet) in *resource records*. The DNS provides these resource records in response to queries it receives from programs called *resolvers* on individual computers throughout the Internet. The resolvers translate domain names into the corresponding IP addresses.

From the inception of the DNS, its most fundamental design goal has been to provide the same answers to the same queries issued from any place on the Internet. As stated in RFC 1034, the basic specification of the DNS's "Concepts and Facilities,"<sup>[16]</sup> "The primary (design) goal is a consistent name space which will be used for referring to resources." And as reiterated in RFC 2535, "Domain Name System Security Extensions,"<sup>[15]</sup> "It is part of the design philosophy of the DNS that the data in it is public and that the DNS gives the same answers to all inquirers."

The DNS is hierarchical. By design, the hierarchy begins with a group of *root nameservers* (often called simply *root servers*), which are specially-designated computers operated under common coordination that provide information about which other computers are authoritative regarding the top-level domains in the DNS naming structure. These set of root servers house the *authoritative root*. Thus, a resolver seeking information concerning a domain name such as **www.example.com** obtains one of the root servers' resource records about **.com**, which tells the resolver which computers have authoritative information about names within the **.com** top-level domain. The resolver then queries one of those authoritative **.com** nameservers about **example.com**, to locate the nameservers for **example.com**. A query is then made to one of those nameservers obtain the IP address of the computer designated by the name **www.example.com**.

The principal advantage of this hierarchical structure is that it allows different parts of the naming database to be maintained by different entities. According to the DNS's design, each domain was intended to be administered by a single entity.<sup>[19]</sup>

When the DNS was deployed in the mid-1980s, a set of root nameservers was designated and several top-level domains were established. These root nameservers (there are now 13 of them distributed around the world) are intended to provide authoritative information about which nameservers hold the naming information for each of the top-level domains. Since the authoritative root nameservers operate at the top of the hierarchy, resolvers find them by referring to IP addresses pre-stored at local computers throughout the Internet.

Over the past several years, some groups have established alternate root nameservers on the public Internet that distribute different information than the information distributed by the authoritative root nameservers. These groups then seek to persuade ISPs and Internet users to replace the pre-stored IP addresses of the authoritative root nameservers with those of their alternate servers. For a variety of reasons, these alternate roots have not to date achieved a significant level of usage on the public Internet.

Fortunately, the rare usage of alternate roots has thus far limited their practical effect on the Internet. If these alternate roots were to become prevalent, however, they would have the potential for seriously disrupting the reliable functioning of the DNS. Some of the consequences include:

- *Providing the Wrong Location:* The presence of alternate public DNS roots can result in different answers being given to the same DNS query issued from different computers on the Internet, depending on whether the inquiring computer is programmed to access the authoritative root or a particular one of the alternate roots (or more precisely a domain-name resolver associated with one or the other of these). The fundamental DNS design goal of providing consistent answers to DNS queries is therefore frustrated.<sup>[1]</sup>
- *Reaching the Wrong Computer:* The main consequence of such inconsistent data is that the same domain name can identify different computers depending on where the name is used. Put another way, *Uniform Resource Locators* (URLs) are no longer uniform. Thus, typing in a Web site address at two different computers configured to reference different roots can result in reaching different Web sites—a particularly disturbing possibility if, for example, money is to change hands or privacy or security concerns are violated. Similarly, the same piece of e-mail sent to the same address from the two computers can be directed to different recipients. The return of inconsistent DNS data defeats the globally consistent resolution of domain names that is vital to the Internet achieving its promise as a universal communications and applications medium for commerce, research, education, cultural exchange, expressive activities, and other uses.
- *Consequences Unpredictable to Most Users:* The set of DNS answers that will be received (from the authoritative root or one of the several alternate roots) is not predictable by most end users. Most users on the Internet employ a local DNS resolver that is configured by another person. Few users are likely to appreciate the significance of the resolver's DNS configuration; even fewer are likely to have detailed knowledge of that configuration. As the number of users on the Internet has grown, the proportion of users knowledgeable about technical concepts such as DNS resolvers and root servers has diminished. Yet these non-technical users are precisely those for whom the Internet in general—and the DNS in particular—hold the greatest potential benefits.

- *Intermediate Hosts Add to Confusion:* Moreover, some Internet services depend on the actions of DNS resolvers employed by intermediate hosts. Alternate roots introduce the possibility that the DNS answer obtained by the intermediate host alters the character of the service in an unexpected way. A similar phenomenon can occur where one user sends another a reference to a URL, such as an e-mail reply address or a link on a Web site. If the recipient of an e-mail or the visitor to the Web site is using a computer that employs a different DNS root than intended by the sender of the e-mail or the designer of the Web site, unexpected results are likely to occur. For example, the e-mail could end up with the wrong person.
- *Cache Poisoning:* Alternate roots also introduce the possibility of misdirected Internet activities due to the phenomenon known as cache poisoning. For performance reasons, the DNS design calls for resource records to be passed around among the nameservers on the Internet, so that a resolver can obtain quicker access to a local copy of the resource record. Because the DNS assumes a single-root system, resource records are not marked to distinguish them according to the root from which they emanate. Thus, the presence of alternate roots introduces the possibility that Internet activities by those intending to use the authoritative root could be misdirected by a stray resource record emanating from an alternate root. Indeed, some malicious hacking attacks have been based on this principle, prompting the *Internet Engineering Task Force* (IETF) to propose a series of not-yet-fully-implemented improvements known as *DNS-Security* or *DNSSEC*.

(It should be noted that the original design of the DNS provided a way to operate alternate roots in a way that does not imperil stability. See “Experimentation” below for details.)

These potentially destructive effects of alternate roots have long been accepted by the vast majority of Internet engineers. Despite this broad-based recognition, some have sought to justify the alternate roots by downplaying these effects. In response, and to document what it referred to as “some of the problems inherent in a family of recurring technically naive proposals,” in May 2000 the *Internet Architecture Board* (IAB)<sup>[14]</sup> issued RFC 2826, entitled “IAB Technical Comment on the Unique DNS Root.” The IAB summarized its comments (in relevant part) as follows:

“Summary: To remain a global network, the Internet requires the existence of a globally unique public name space. The DNS name space is a hierarchical name space derived from a single, globally unique root. This is a technical constraint inherent in the design of the DNS. Therefore it is not technically feasible for there to be more than one root in the public DNS. That one root must be supported by a set of coordinated root servers administered by a unique naming authority.

“Put simply, deploying multiple public DNS roots would raise a very strong possibility that users of different ISPs who click on the same link on a Web page could end up at different destinations, against the will of the Web page designers.”

For some concrete examples of potential failures and instabilities that would likely result from alternate roots prevalently used on the public Internet, see the draft “Alt-Roots, Alt-TLDs.”<sup>[17]</sup>

In the face of the destabilizing consequences of alternate roots, as articulated by the IAB and others, ICANN’s prime directive of preserving the stability of the Internet and DNS requires an unwavering commitment to promote the continued prevalence of a single authoritative root for the public DNS. Any other course of action by ICANN would be irresponsible.

### **The Public Trust in Coordinated Assignment Functions**

The Internet’s proper operation requires assignment of unique values to various identifiers for different computers or services on the Internet. To be effective, these assigned values must be made broadly available and their significance must be respected by the many people responsible for the Internet’s operation. For example, every computer on the public Internet is assigned a unique IP address; this address is made known to routers throughout the Internet to cause TCP/IP packets with that destination address to be routed to the intended computer. Without common agreement to respect the assignment, the Internet would not reliably route communications to their intended destinations.

### **Beginnings to 1998: Central Coordination as a Public Trust**

From the very beginnings of the Internet, the technical community has recognized the need for central coordination of the unique assignment of the values of identifiers. The IANA, now operated by ICANN was created to fill this need; it now makes assignments of unique values for approximately 120 different identifier types. This responsibility has always been understood to be a public trust, and the IANA long ago adopted the motto: “Dedicated to preserving the central coordinating functions of the global Internet for the public good.”

The most commonly known of the Internet’s uniquely assigned identifiers, of course, are domain names. From the time the DNS was deployed, the Internet community made the IANA “responsible for the overall coordination and management of the Domain Name System (DNS), and especially the delegation of portions of the name space called top-level domains.”<sup>[18]</sup> As in its other assignment responsibilities, the IANA’s role is to act in the public interest, neutrally, and without proprietary motives.

### **Competition as a Value Guiding the Internet's Technical Management**

In the Internet's early years, with limited exceptions day-to-day registration activities for domain names were done by a single company (first SRI International and later Network Solutions) under the IANA's guidance.

By the mid-1990s, however, the growth and increasing commercialization of the Internet led the U.S. Government's Green<sup>[2]</sup> and White<sup>[3]</sup> Papers to note the emergence of "widespread dissatisfaction about the absence of competition in domain name registration." This dissatisfaction prompted the Green and White Papers to include the promotion of competition in registration services as one of the four values (stability; competition; private, bottom-up coordination; and representation) that should guide the Internet's technical management. Both documents made clear that, of these four values, preservation of stability was to be paramount.

Building on the IANA model of a non-profit entity carrying the public trust to perform the vital central coordination functions, the U.S. Government reconciled the need to ensure Internet stability with the desire to introduce competitive domain-name registration services as follows:

"In keeping with these principles, we divide the name and number functions into two groups, those that can be moved to a competitive system and those that should be coordinated. We then suggest the creation of a representative, not-for-profit corporation to manage the coordinated functions according to widely accepted objective criteria. We then suggest the steps necessary to move to competitive markets in those areas that can be market driven."<sup>[4]</sup>

This dichotomy recognizes that the Internet is, after all, a network (albeit a network of networks), and networks require coordination among their participants to operate in a stable and efficient manner. It also reflects the phenomenal success of the Internet's tradition of cooperatively developed open and non-proprietary standards. Those standards have provided an environment of highly interoperable systems that has allowed competition and innovation to flourish.

### **ICANN Assumes the Public Trust**

After public comment on the Green Paper, the United States Government issued the White Paper, which laid out the basic charter on which ICANN was founded and continues to operate. The White Paper re-emphasized the prime directive of stability and, to that end, the need to avoid creation of alternate roots:

"The introduction of a new management system should not disrupt current operations or create competing root systems. During the transition and thereafter, the stability of the Internet should be the first priority of any DNS management system."<sup>[5]</sup>

The United States Government then invited the Internet community to form a not-for-profit corporation to perform the “coordinated functions” that should be handled as a matter of public trust, rather than according to a competitive regime that would not be conducive to stability. Among the “coordinated functions” were management of the root-server system and decisions to introduce new TLDs:

“Similarly, coordination of the root server network is necessary if the whole system is to work smoothly. While day-to-day operational tasks, such as the actual operation and maintenance of the Internet root servers, can be dispersed, *overall policy guidance and control of the TLDs and the Internet root server system should be vested in a single organization* that is representative of Internet users around the globe.

“Further, changes made in the administration or the number of gTLDs contained in the authoritative root system will have considerable impact on Internet users throughout the world. In order to promote continuity and reasonable predictability in functions related to the root zone, the *development of policies for the addition, allocation, and management of gTLDs and the establishment of domain name registries and domain name registrars to host gTLDs should be coordinated.*”<sup>[6]</sup>

In response to this invitation for the formation of a non-profit, Internet-community-based organization, ICANN was established in 1998. ICANN was subsequently selected by the United States Government from among several proposals submitted precisely because it was open, consensus-based, and rooted in the Internet community. The establishment of ICANN had followed extensive dialogs among different constituencies of the Internet community to ensure that ICANN could be responsive to the needs of these various constituencies.

ICANN, among its other responsibilities, now acts as the coordinator for operation of the authoritative root-server system and the policy forum for decisions about the policies governing what TLDs are to be included in the authoritative DNS root.<sup>[7]</sup>

In linking the formation of ICANN to the global Internet community, the White Paper established a public trust that required that the DNS be administered in the public interest as the unique-rooted,<sup>[8]</sup> authoritative database for domain names that provides a stable addressing system for use by the global Internet community. This commitment to a unique and authoritative root is a key part of the broader public trust—to carry out the Internet’s central coordination functions for the public good—that is ICANN’s reason for existence.

### The Public Trust and the Introduction of New TLDs

It is essential that the centrally coordinated functions be performed in the public interest, not out of proprietary or otherwise self-interested motives. For this reason, ICANN was founded as a not-for-profit public-benefit organization, accountable to the Internet community. Longstanding Internet principles also require that the policies guiding the coordinated functions be established openly based on community deliberation and input. For these reasons ICANN's structure is representative of the geographic and functional diversity of the Internet, and relies to the extent possible on private-sector, bottom-up methods.

As the White Paper emphasized, the decisions about the introduction of new TLDs are appropriately done within this open, non-proprietary, and broadly representative framework, rather than by individuals or entities not accountable to the community and that ordinarily act for their own proprietary motives:

“As Internet names increasingly have commercial value, the decision to add new top-level domains cannot be made on an ad hoc basis by entities or individuals that are not formally accountable to the Internet community.”<sup>[9]</sup>

Within the framework of its commitment to a unique root system and to the stability of the Internet, last year ICANN launched a process for carefully introducing several new generic TLDs to the DNS. This introduction was fashioned as a proof of concept of the technical and business feasibility of introducing more TLDs into the DNS. Proceeding with an initial proof of concept was in response to the advice of ICANN's *Protocol Supporting Organization* (PSO) and its *Domain Name Supporting Organization* (DNSO) to proceed cautiously and in an orderly fashion. The PSO and the DNSO represent the consensus views of the technical and the user/business/other institutional communities, respectively. Generic TLDs had not been introduced for many years, and there were and still are serious questions as to what the effect of introducing new TLDs will be on the stability and reliability of the DNS; and many questions about what should be the appropriate contractual and business context.

In response to an issued RFP, forty-seven institutions and groups submitted proposals for the establishment of new TLDs. They chose to work within the community-based ICANN process, even though they knew that only a “limited number” of TLDs would be selected—at least in the first round. In fact, seven were selected, and, following a methodology which allowed for considerable community input, contracts have or will shortly be signed with these initial seven. ICANN looks forward to the successful introduction of these new TLDs and will work with the community to monitor their performance so that a community decision can be made on moving forward with the introduction of more TLDs, should this be the conclusion of the proof of concept.

### **Outside the Process**

Some private organizations have established DNS roots as alternates to the authoritative root. Some uses of these alternate roots do not jeopardize the stability of the DNS. For example, many are purely private roots operating inside institutions and are carefully insulated from the DNS. Others are purely experimental in the best traditions of the Internet and are carefully managed so as not to interfere with the operation of the DNS. These both operate within community-established norms.

Frequently, however, these alternate roots have been established to support top-level or pseudo-top-level domain name registries that are operated for profit. Yet other alternate roots have been established by certain individuals to protest the policies developed by the broader community processes for management of the authoritative root, or to express their disinterest in participating in those processes. These alternate roots have not been launched through any ICANN consensus processes, so they have not been entered into the authoritative root managed by the IANA or ICANN.

These alternate roots typically substitute insular concerns in place of the community-based processes that govern the management of the authoritative root. Their operators decide to include particular top-level domains in these alternate roots that have not been subjected to the tests of community support and conformance with consensus processes—coordinated by ICANN—that would allow their inclusion in the authoritative root. These decisions of the alternate root operators have been made with no apparent regard for the fundamental public-interest concern of Internet stability. The widespread introduction of active domain names into these alternate roots could in fact impair the uniqueness of the authoritative name resolution mechanism and hence the stability of the DNS.

In fact, some of the operators of these alternate roots state that stability is not an important attribute for the DNS. This thesis, for reasons already stated, is at fundamental variance with ICANN policy as embodied in its founding documents. Some of these operators and their supporters assert that their very presence in the marketplace gives them preferential right to TLDs to be authorized in the future by ICANN. They work under the philosophy that if they get there first with something that looks like a TLD and invite many registrants to participate, then ICANN will be required by their very presence and force of numbers to recognize in perpetuity these pseudo TLDs, inhibiting new TLDs with the same top-level name from being launched through the community's processes.

No current policy allows ICANN to grant such preferential rights. To do so would effectively yield ICANN's mandate to introduce new TLDs in an orderly manner in the public interest to those who would simply grab all the TLD names that seem to have any marketplace value, thus

circumventing the community-based processes that ICANN is required to follow. For ICANN to yield its mandate would be a violation of the public trust under which ICANN was created and under which it must operate. Were it to grant such preferential rights, ICANN would abandon this public trust, rooted in the community, to those who only act for their own benefit. Indeed, granting preferential rights could jeopardize the stability of the DNS, violating ICANN's fundamental mandate.

Alternate roots inherently endanger DNS stability—that is, they create the real risk of name resolvers being unable to determine to which numeric address a given name should point. This violates the fundamental design of the DNS and impairs the Internet's utility as a ubiquitous global communications medium. Some of these alternate systems also employ special technologies that—ingenious as they may be—may conflict with future generations of community-established Internet standards. Indeed, can there be any guarantee that these proprietary technologies can or will be adapted to future changes in Internet standards?

### **Experimentation**

Experimentation has always been an essential component of the Internet's vitality. Working within the system does not preclude experimentation, including experimentation with alternate DNS roots. But these activities must be done responsibly, in a manner that does not disrupt the ongoing activities of others and that is managed according to experimental protocols.

DNS experiments should be encouraged. Experiments, however, almost by definition have certain characteristics to avoid harm: (a) they are clearly labeled as experiments, (b) it is well understood that these experiments may end without establishing any prior claims on future directions, (c) they are appropriately coordinated within a community-based framework (such as the IETF), and (d) the experimenters commit to adapt to consensus-based standards when they emerge through the ICANN and other community-based processes. This is very different from launching commercial enterprises that lull users into a sense of permanence without any sense of the foregoing obligations or contingencies.

Moreover, it is essential that experimental operations involving alternate DNS roots be conducted in a controlled manner, so that they do not adversely affect those who have not consented to participate in them. Given the design of the DNS, and particularly the intermediate-host and cache poisoning issues described earlier, special care must be taken to insulate the DNS from the alternate roots' effects. For example, alternate roots are commonly operated by large organizations within their private networks without harmful effects, since care is taken to prevent the flow of the alternate resource records onto the public Internet.

It should be noted that the original design of the DNS provides a facility for future extensions that accommodates the possibility of safely deploying multiple roots on the public Internet for experimental and other purposes. As noted in RFC 1034, the DNS includes a “class” tag on each resource record, which allows resource records of different classes to be distinguished even though they are commingled on the public Internet. For resource records within the authoritative root-server system, this class tag is set to “IN”; other values have been standardized for particular uses, including 255 possible values designated for “private use” that are particularly suited to experimentation.<sup>[10]</sup>

As described in a recent proposal within the IETF,<sup>[11]</sup> this “class” facility allows an alternate DNS namespace to be operated from different root servers in a manner that does not interfere with the stable operation of the existing authoritative root-server system. To take advantage of this facility, it should be noted, requires the use of client or applications software developed for the alternate namespace (presumably deployed after responsible testing), rather than the existing software that has been developed to interoperate with the authoritative root. Those who operate alternate roots for global commercial purposes, however, have not followed this course.

In an ever-evolving Internet, ultimately there may be better architectures for getting the job done where the need for a single, authoritative root will not be an issue. But that is not the case today. And the transition to such an architecture, should it emerge, would require community-based approaches. In the interim, responsible experimentation should be encouraged, but it should not be done in a manner that affects those who do not consent after being informed of the character of the experiment.

### **Conclusion**

The success of the Internet and the guarantee of Internet stability rest on the cooperative activities of thousands, even millions, of people and institutions collaborating worldwide towards a common end. This extraordinary—even unprecedented—community effort has served to impel the incredible growth of the Internet. Many of these people and institutions compete intensely among themselves yet agree to do so within a common framework for the overall public good. Their collective efforts provide a policy framework for technical and entrepreneurial innovation, and the advancement of economic, social, and educational goals.

Most members of the global community and most institutions with which they are associated recognize that it is in their best long-term interests to work within these community-based processes, even if that means foregoing short-term advantages to particular individuals or groups. The over-arching principles outlined in this document override exclusive and narrowly focused self-interest.

Community-based policy development is not perfect. It may proceed slower than some would wish. The introduction of new TLDs has proceeded at deliberate speeds. Impatience in the context of Internet timescales is perfectly understandable. The outcome of orderly processes based on the wishes of the community, however, is assurance that the Internet will continue to function in a stable and holistic manner that benefits the global community, and not become captured by the self-interests of the few. That, in the minds of most, is a price worth paying.

ICANN—in deference to its public trust—will continue to collaborate with these citizens of the Internet community to advance the notions of a unique root system as a prerequisite to Internet stability, and to ensure that community-based policies take precedence. ICANN encourages responsible experimentation designed to further advance the Internet as a useful, stable, and accessible medium for the public good.

### References

- [1] Ironically, to avoid name conflicts in a multi-root system, a single-root system would need to be created—adding a higher level to the hierarchy.
  - [2] “Improvement of Technical Management of Internet Names and Addresses,” (Green Paper), 63 *Federal Register* 8825, 8827 (20 February, 1998).
  - [3] “Management of Internet Names and Addresses,” (White Paper), 63 *Federal Register* 31741, 31742 (10 June, 1998).
  - [4] Green Paper, 63 *Federal Register* at 8827.
  - [5] White Paper, 63 *Federal Register* at 31749. The Green and White Papers both made additional references to the need for a single authoritative root system. For example, in response to comments received from the Green Paper, the White Paper notes:

“In the absence of an authoritative root system, the potential for name collisions among competing sources for the same domain name could undermine the smooth functioning and stability of the Internet.”
  - [6] White Paper, 63 *Federal Register* at 31749 (emphasis added).
  - [7] ICANN’s corporate charter emphasizes its role in overseeing operation of the unique DNS root:

“... the Corporation shall ... pursue the charitable and public purposes ... of promoting the global public interest in the operational stability of the Internet by ... (iv) overseeing operation of the authoritative Internet DNS root server system ...”

ICANN Articles of Incorporation, para. 3. The phrase “the authoritative Internet DNS root server system” is decidedly in the *singular*.
- See: <http://www.icann.org/general/articles.htm>

[8] The Memorandum of Understanding between the United States Government and ICANN that governs the transfer of responsibilities from the U.S. Department of Commerce to ICANN also makes reference to the authoritative root in the singular, not in the plural:

“In the DNS Project, the parties will jointly design, develop, and test the mechanisms, methods, and procedures to carry out the following DNS management functions: ...

“b. Oversight of the operation of the authoritative root server system;

“c. Oversight of the policy for determining the circumstances under which new top level domains would be added to the root system ... ”

See also: [www.icann.org/general/icann-mou-25nov98.htm](http://www.icann.org/general/icann-mou-25nov98.htm)

[9] White Paper, 63 *Federal Register* at 31742.

[10] Eastlake, D., Brunner-Williams, E., Manning, B., “Domain Name System (DNS) IANA Considerations,” section 3.2, RFC 2929, September, 2000.

[11] Klensin, J., “Internationalizing the DNS—A New Class,” Internet Draft, work in progress, December, 2000.

[12] Internet Assigned Numbers Authority (IANA). See [www.iana.org](http://www.iana.org)

[13] Postel, J., “Domain Name System Implementation Schedule—Revised,” RFC 921, October 1984.

[14] Internet Architecture Board (IAB). See <http://www.iab.org>

[15] Eastlake, D., “Domain Name System Security Extensions,” RFC 2535, March 1999.

[16] Mockapetris, P., “Domain Names—Concepts and Facilities,” RFC 1034, November 1987.

[17] <http://www.icann.org/stockholm/draft-crispin-alt-roots-tlds-00.txt>

[18] Postel, J., “Domain Name System Structure and Delegation,” RFC 1591, March 1994.

[19] Postel, J., and Reynolds, J., “Domain Requirements,” RFC 920, October 1984.

Dr. M. STUART LYNN is President & CEO of The Internet Corporation for Assigned Names and Numbers (ICANN). Dr. Lynn has had a distinguished career in computing and information technology that dates back almost four decades. His most recent position until his retirement in 1999 was as Associate Vice President for Information Resources and Communications for the University of California Office of the President where he served as chief information officer for the combined University of California system. Dr. Lynn also served as President and Chairman of the Board of the Corporation for Education Network Initiatives in California (CENIC). Dr. Lynn has also held positions at Cornell University, UC Berkeley, Rice University, Baylor College of Medicine, IBM and Chevron. Over the course of his career, he has been active in several professional organizations including the Association for Computing Machinery (ACM) and the American Federation of Information Processing Societies. In 1994, he was elected a Fellow of the ACM. In addition, he has served on numerous boards of directors, advisory committees and as a consultant to academia, government and industry. Dr. Lynn holds a M.A. and Ph.D. in Mathematics from the University of California at Los Angeles and a B.A. and M.A. in Mathematics from Oxford University.

E-mail: [lynn@icann.org](mailto:lynn@icann.org)

# Book Review

**Web Protocols and Practice** *Web Protocols and Practice: HTTP/1.1, Networking Protocols, Caching, and Traffic Measurement*, by Balachander Krishnamurty and Jennifer Rexford, ISBN 0-201-71088-9, Addison-Wesley, 2001.

If you want to know something about the underlying workings of the Web, you can find it somewhere out there on the Web itself. But, as we all know, it is not always easy to find the page you want, and particularly not if you are in a hurry and don't want to have to wade through documentation hierarchies or download PDF files. In these cases a real book is unbeatable, if one is available. Sadly, for information about the lower reaches of Web protocols there has been no single useful printed reference source available.

## Organisation

This book fills that gap. It provides a detailed look at all the low level protocol issues as well as many other things; the book's subtitle sums it up admirably. The first section provides a brief history of the Web and its development which introduces all the important terminology and, most importantly, also says what the book is *not* about: nothing on XML (hurrah!), HTML, scripting languages, administration of Web servers, or specific products.

Section two moves on to more technical matters looking at Web clients, proxies and servers. The client chapter has a particularly useful section on spiders with an excellent table showing the names and calling hosts of the commonest spider programs. The information about proxies and servers is also of high quality and provide a solid grounding in how they interact with each other and the potential problems that can arise.

The third section looks at the protocols involved when using the Web. Starting with a concise run through TCP and the use of the DNS, the authors then glance at FTP, SMTP and NNTP, before going to a detailed examination of HTTP/1.1. In my personal experience, information on HTTP/1.1 has always been particularly inaccessible, both from the point of view of discoverability and readability, and this chapter explained several things that I had been puzzled about, especially about cache control which is rather a black art. (Also featured is a comprehensive table of HTTP return codes to which I shall turn quite often.) To finish this section of the book, there is a chapter on how HTTP interacts with TCP—a whole area that I had never really thought about before and which is much more complex that I would have thought it to be.

Next is a short section devoted to measuring and characterizing Web traffic. This a hugely contentious area and the discussion is well balanced and sensible. Following this the authors look in more detail at caching and at multimedia streaming, and manage to cover the latter topic without going into much unnecessary details about the actual bits that get sent whilst still giving a good coverage of the important material.

To round off the book, there are three chapters devoted to research topics, looking again at caching, measurement and protocol issues. Much of the material here is not directly of relevance to someone who is dealing with Web protocols on a daily basis, but there is still much here that will be of interest as the authors draw attention to places where improvements can be expected and how these might be realised.

#### **Excellent Book**

As you might expect, there is also a comprehensive bibliography and index. All in all an excellent book that is well researched, well written, and clearly set out without the excess of white space that is so common in computing books today. The price is perhaps rather high (I certainly could not recommend this as a textbook to my students—they simply could not afford it), but for people working in the industry it would be a worthwhile purchase and I think that they would soon find it an indispensable source of reference.

—Lindsay Marshall, *University of Newcastle upon Tyne*

**Lindsay.Marshall@ncl.ac.uk**

#### **Summary of Acronyms**

DNS: *Domain Name System*

FTP: *File Transfer Protocol*

HTTP: *HyperText Transfer Protocol*

NNTP: *Network News Transfer Protocol*

PDF: *Portable Document Format*

SMTP: *Simple Mail Transfer Protocol*

TCP: *Transmission Control Protocol*

XML: *Extensible Markup Language*

---

#### **Would You Like to Review a Book for IPJ?**

We receive numerous books on computer networking from all the major publishers. If you've got a specific book you are interested in reviewing, please contact us and we will make sure a copy is mailed to you. The book is yours to keep if you send us a review. We accept reviews of new titles, as well as some of the "networking classics." Contact us at [ipj@cisco.com](mailto:ipj@cisco.com) for more information.

## Call for Papers

*The Internet Protocol Journal* (IPJ) is published quarterly by Cisco Systems. The journal is not intended to promote any specific products or services, but rather is intended to serve as an informational and educational resource for engineering professionals involved in the design, development, and operation of public and private internets and intranets. The journal carries tutorial articles (“What is...?”), as well as implementation/operation articles (“How to...”). It provides readers with technology and standardization updates for all levels of the protocol stack and serves as a forum for discussion of all aspects of internetworking.

Topics include, but are not limited to:

- Access and infrastructure technologies such as: ISDN, Gigabit Ethernet, SONET, ATM, xDSL, cable, fiber optics, satellite, wireless, and dial systems
- Transport and interconnection functions such as: switching, routing, tunneling, protocol transition, multicast, and performance
- Network management, administration, and security issues, including: authentication, privacy, encryption, monitoring, firewalls, trouble-shooting, and mapping
- Value-added systems and services such as: Virtual Private Networks, resource location, caching, client/server systems, distributed systems, network computing, and Quality of Service
- Application and end-user issues such as: e-mail, Web authoring, server technologies and systems, electronic commerce, and application management
- Legal, policy, and regulatory topics such as: copyright, content control, content liability, settlement charges, “modem tax,” and trademark disputes in the context of internetworking

In addition to feature-length articles, IPJ will contain standardization updates, overviews of leading and bleeding-edge technologies, book reviews, announcements, opinion columns, and letters to the Editor.

Cisco will pay a stipend of US\$1000 for published, feature-length articles. Author guidelines are available from Ole Jacobsen, the Editor and Publisher of IPJ, reachable via e-mail at [ole@cisco.com](mailto:ole@cisco.com)

### Next ICANN Meeting, Marina del Rey, November 13–15, 2001

Many members of the *Internet Corporation for Assigned Names and Numbers* (ICANN) community wrote in response to a call for input as to whether the events of September 11 would affect their plans to travel to Los Angeles in November to attend the scheduled ICANN meetings. Almost without exception the respondents emphatically encouraged ICANN to hold its meetings and stated unequivocally that they planned to attend unless the international situation deteriorated to where travel was not practical.

Given this response and given the need to address emerging priorities, ICANN is planning to proceed with its November meeting, subject to any further serious change in the international situation that would affect travel conditions. However, as discussed below, the format of the meeting will differ significantly from what had previously been announced.

The events of September 11 have caused institutions worldwide to rethink their priorities and plans. As an international institution, ICANN is not immune. Although those events raise logistical and other concerns for holding meetings, they also underscore the need to address Internet stability issues, and security as a key component of stability. ICANN is not responsible for the overall security of the Internet. However, given ICANN's global responsibilities for the stability of the Internet's naming and addressing systems and under the new circumstances facing the international community, it would be irresponsible for ICANN not to conduct an in depth assessment of the robustness and security of these systems, and to take steps, if necessary, to strengthen the Internet in these regards. These are urgent matters and of worldwide importance.

The Internet is global in reach, as are the threats of terrorism. The events of September 11 offered a stark and tragic reminder of the incalculable importance of a reliable and secure naming and addressing system to support emergency response, personal and other communications, and information sharing. E-mail, instant messaging, and the Web, for example, all played essential roles.

Accordingly, the November ICANN meetings will focus on stability and security of the Internet's naming and addressing systems and of their operational implementation globally. This will be the overriding imperative for the meeting. As such, this will be a very different kind of meeting than previous ICANN meetings and will not follow the usual format.

At this meeting, ICANN will be seeking to promote discussion throughout the community on how to reassess areas of potential threats that could affect services within the scope of ICANN's responsibilities, how to improve readiness to meet these threats, and what additional policies or other actions should be considered and implemented to facilitate such improvements.

Clearly not all these questions will be answered in one meeting, but ICANN must now devote its energies as members of the global Internet community towards obtaining answers. Every constituency and supporting organization will be asked to report on its efforts to ensure the stability of the Internet's naming and addressing systems and what additional steps it proposes to take to improve that stability and security among its member organizations. Agenda items will be assessed for inclusion by what they contribute to the overall focus of the meeting.

Although a precise schedule has not yet been mapped out, these meetings will last three days from November 13 through 15, inclusive. Constituencies and supporting organizations will be asked to meet during this time to focus on the topic of the meeting. There will be a Board meeting at the end of the meeting to address essential business. The Board agenda will concentrate on topics where time is of the essence.

The focus of the meetings may well delay progress on some of the worthy and important initiatives that are currently underway. The effects of such delays have to be measured against the importance of ensuring the stability and security of the Internet itself. This will require patience on the part of those who may experience delays in matters of importance to them so that the ICANN community can bear down on the issue at hand.

This is only a preliminary announcement to enable attendees to firm up their travel plans. Details of the meeting will be announced as soon as possible. Please visit the ICANN Web site (<http://www.icann.org>) for further updates.

#### **Van Jacobson Receives 2001 ACM SIGCOMM Award**

Van Jacobson, the man widely credited with saving the Internet from an otherwise inevitable congestion collapse in the late 1980s, has been named the 2001 recipient of the ACM SIGCOMM Award. Jacobson is chief scientist at networking startup Packet Design, LLC.

The award is given annually by the *Association for Computing Machinery's Special Interest Group in Data Communications* (ACM SIGCOMM) to a recipient with a long and distinguished history of contributing to the field of data communications. Jacobson began his career in data communications developing control systems for the Department of Energy in the 1970s. He is best known for redesigning the TCP/IP protocol's flow-control algorithms to better handle congestion, preventing the Internet's collapse from traffic congestion in 1988–89. He is also widely recognized for his work on network synchronization effects, scalable multimedia protocols and applications, IP operations tools (for example *traceroute* and *pathchar*) and high-performance TCP implementations.

Prior to joining Packet Design as a member of the founding team, Jacobson was chief scientist at Cisco Systems, and before that had been group leader for Lawrence Berkeley Laboratory's Network Research Group.

The SIGCOMM Award has been presented every year since 1989. Prior recipients include Paul Baran, Vinton G. Cerf, David Farber and Leonard Kleinrock. ACM SIGCOMM is the world's largest professional society devoted to data communications. For more information, see: <http://www.acm.org/sigcomm/>

#### Useful Links

The following is a list of Web addresses that we hope you will find relevant to the material typically published in *The Internet Protocol Journal*. In the near future we will make these and other links available on our Web site: <http://www.cisco.com/ipj>

If you have suggestions for other pointers to include, please drop us a line at [ipj@cisco.com](mailto:ipj@cisco.com)

- The *Internet Engineering Task Force* (IETF). The primary standards-setting body for Internet technologies. <http://www.ietf.org>
- *Internet-Drafts* are working documents of the IETF, its areas, and its working groups. Note that other groups may also distribute working documents as Internet-Drafts. Internet-Drafts are not an archival document series. These documents should not be cited or quoted in any formal document. Unrevised documents placed in the Internet-Drafts directories have a maximum life of six months. After that time, they must be updated, or they will be deleted. Some Internet-Drafts become RFCs (see below). <http://www.ietf.org/ID.html>
- The *Request For Comments* (RFC) document series. The RFCs form a series of notes, started in 1969, about the Internet (originally the ARPANET). The notes discuss many aspects of computer communication, focusing on networking protocols, procedures, programs, and concepts but also including meeting notes, opinion, and sometimes humor. The specification documents of the Internet protocol suite, as defined by IETF and its steering group the IESG, are published as RFCs. Thus, the RFC publication process plays an important role in the Internet standards process. <http://www.rfc-editor.org/>
- The *Internet Society* (ISOC) is a non-profit, non-governmental, international, professional membership organization. <http://www.isoc.org>
- The *Internet Corporation for Assigned Names and Numbers* (ICANN) "... is the non-profit corporation that was formed to assume responsibility for the IP address space allocation, protocol parameter assignment, domain name system management, and root server system management functions previously performed under U.S. Government contract by IANA and other entities." <http://www.icann.org>

- The *North American Network Operators' Group* (NANOG) “...provides a forum for the exchange of technical information, and promotes discussion of implementation issues that require community cooperation. Coordination among network service providers helps ensure the stability of overall service to network users.”  
<http://www.nanog.org>
- The *Regional Internet Registries* (RIRs) provide IP address block assignments for Internet Service Providers and others. Currently, there are three active RIRs:
  - The *Asia Pacific Network Information Centre* (APNIC):  
<http://www.apnic.net>
  - *RIPE Network Coordination Centre*—the RIR responsible for Europe and Northern Africa: <http://www.ripe.net>
  - *American Registry for Internet Numbers* (ARIN)—the RIR responsible for the Americas and Sub-Saharan Africa:  
<http://www.arin.net>

Two more RIRs are in the process of formation: *AfriNIC* for Africa and *LACNIC* for Central- and Latin America.
- The *World Wide Web Consortium* (W3C) “ ... develops interoperable technologies (specifications, guidelines, software, and tools) to lead the Web to its full potential as a forum for information, commerce, communication, and collective understanding.”  
<http://www.w3.org/>
- The *International Telecommunication Union* (ITU) “... is an international organization within which governments and the private sector coordinate global telecom networks and services.”  
<http://www.itu.int>
- The *International Organization for Standardization* (ISO) “ ... is a worldwide federation of national standards bodies from some 140 countries, one from each country. The mission of ISO is to promote the development of standardization and related activities in the world with a view to facilitating the international exchange of goods and services, and to developing cooperation in the spheres of intellectual, scientific, technological and economic activity. ISO's work results in international agreements which are published as International Standards.” <http://iso.org>

This is by no means intended to be a complete list of organizations that are related to Internet development in one way or another, but this list should give you a good starting point.

This publication is distributed on an “as-is” basis, without warranty of any kind either express or implied, including but not limited to the implied warranties of merchantability, fitness for a particular purpose, or non-infringement. This publication could contain technical inaccuracies or typographical errors. Later issues may modify or update information provided in this issue. Neither the publisher nor any contributor shall have any liability to any person for any loss or damage caused directly or indirectly by the information contained herein.

---

## The Internet Protocol Journal

Ole J. Jacobsen, Editor and Publisher

### Editorial Advisory Board

**Dr. Vint Cerf**, Sr. VP, Internet Architecture and Technology  
WorldCom, USA

**Dr. Jon Crowcroft**, Marconi Professor of Communications Systems  
University of Cambridge, England

**David Farber**  
The Alfred Fitler Moore Professor of Telecommunication Systems  
University of Pennsylvania, USA

**Peter Löthberg**, Network Architect  
Stupi AB, Sweden

**Dr. Jun Murai**, Professor, WIDE Project  
Keio University, Japan

**Dr. Deepinder Sidhu**, Professor, Computer Science &  
Electrical Engineering, University of Maryland, Baltimore County  
Director, Maryland Center for Telecommunications Research, USA

**Pindar Wong**, Chairman and President  
VeriFi Limited, Hong Kong

*The Internet Protocol Journal is published quarterly by the Chief Technology Office, Cisco Systems, Inc.  
www.cisco.com  
Tel: +1 408 526-4000  
E-mail: ipj@cisco.com*

*Cisco, Cisco Systems, and the Cisco Systems logo are registered trademarks of Cisco Systems, Inc. in the USA and certain other countries. All other trademarks mentioned in this document are the property of their respective owners.*

*Copyright © 2001 Cisco Systems Inc.  
All rights reserved. Printed in the USA.*



The Internet Protocol Journal, Cisco Systems  
170 West Tasman Drive, M/S SJ-10/5  
San Jose, CA 95134-1706  
USA

ADDRESS SERVICE REQUESTED

