The Hitchhikers Guide to the Internet

The Hitchhiker's Guide to the Internet by Stephen McLaughlin is a humorous and informative guide that navigates the vast landscape of the early internet, offering tips, insights, and wit on how to explore the digital world.

RFCs

The foundational structure and operational mechanics of the Internet are encapsulated within a series of documents known as RFCs (Request for Comments). This unique process begins when an individual aims to formalize a concept by drafting a document delineating their proposal and submitting it to Jon Postel (postel@isi.edu), who facilitates the discussion as a mediator. The proposal is then digitally circulated for commentary among all interested parties, possibly undergoing several iterations of revision. Upon reaching a consensus that the idea is beneficial, the proposal is formalized by assigning it an RFC number and incorporating it into the official collection.

RFCs are categorized into five distinct classes: required, suggested, directional, informational, and obsolete. The 'required' category encompasses fundamental protocols, such as RFC-791 (The Internet Protocol), that must be implemented by any host wishing to connect to the Internet. Suggested RFCs, like RFC-793 (Transmission Control Protocol), while not mandatory, are widely adopted by network hosts to enhance Internet usability. Directional RFCs, despite receiving approval, have not been broadly applied due to lack of necessity or competition with established methods; however, adoption is encouraged to maintain a standard approach should the need arise. Informational RFCs serve to provide essential facts regarding the Internet's functionality (e.g., RFC-990, Assigned Numbers). Over time, as the Internet evolves, certain RFCs become obsolete, yet they hold historical significance in understanding the progression and modifications of Internet protocols, with new RFCs sometimes supplanting older versions without diminishing their relevance.

This systematic method ensures the Internet remains an open, collaborative platform, with RFCs serving as the cornerstone for its development and innovation, allowing for dynamic evolution while maintaining a core structure and interoperability among diverse systems.

Operating the Internet

In the chapter titled "Operating the Internet," the process of accessing and retrieving documents from the early internet is detailed through the usage of tools like telnet and FTP (File Transfer Protocol) on a BSD (Berkeley Software Distribution) system. The narrative begins with a user deciding to locate and download a document named NETINFO:NUG.DOC, which stands for The Users Guide to the ARPAnet, using telnet to access sri-nic.arpa. This is followed by an illustration of how to conduct an anonymous FTP session to actually retrieve the document.

The FTP process is described step by step: connecting to the FTP server at sri-nic.arpa, logging in as an anonymous user by providing 'myname' as a sort of password, initiating the download of NETINFO:NUG.DOC, and successfully completing the transfer of 157675 bytes in approximately 450 seconds. This sequence showcases the early internet's user experience, characterized by text-based commands and relatively simple user-server interactions.

After obtaining NETINFO:NUG.DOC, the text suggests another document, NETINFO:WHAT-THE-NIC-DOES.TXT, as another valuable resource for new users, emphasizing the richness of information accessible through these means. Further, it addresses how users can communicate with the Network Information Center (NIC) for assistance, document requests, or to report problems through various email addresses specific to different inquiries, such as general user assistance, user registration and WHOIS updates, hostnames and domain changes, computer operations, and comments on NIC publications and services.

Lastly, the chapter acknowledges those without network access, suggesting alternative means to obtain documents, indirectly highlighting the period's limitations and the dependency on electronic mail and online connectivity for information exchange. This narrative encapsulates the nascent stages of internet operation, emphasizing the technical procedures and community support structures in place to navigate the ARPAnet, the precursor to the modern internet.

Address Allocation

In this chapter focusing on "Operating the Internet," the discussion revolves around methodologies for distributing internet messages within a campus setting and the intricacies of IP address allocation necessary for connecting a local network to the internet.

Message Distribution Methods:
1. Reflector Set-up on a Local Machine: A reflector forwards messages to a campus-wide distribution list, enabling a singular message to reach a broad audience efficiently.
2. Creation of an Alias for Notesfile Access: This allows messages to be placed in a notesfile, where campus users can access the latest information at their convenience.
3. Screening by the Campus Wide Area Network Liaison: The option to have messages screened for relevance or merit before forwarding them ensures that the information distributed is of value to the campus community.

IP Address Allocation:
- Unique IP Address Requirement: Before a network can join the internet, it must be assigned a unique IP address by the Internet Systems Consortium (ISI).
- Addressing Process: The process involves acquiring an application from ISI, completing it, and submitting it back, either electronically or via postal mail. An IP address is then assigned and communicated to the applicant.
- IP Address Format: An IP address is composed of four decimal numbers separated by periods (e.g., 192.17.5.100), representing a 32-bit value divided into octets.
- Classification of Networks: IP addresses are categorized into Classes A, B, and C to accommodate various network sizes—from large to small—based on hierarchical or flat organizational structures.
- Class A is designed for very large networks, Class B for medium-sized ones, and Class C addresses support smaller networks.
- Class D and Class E addresses are reserved for multicast and experimental uses, respectively.

Strategies for Addressing and Routing:
- Subnetting for Efficient Addressing: To manage routing effectively, campuses or sites should limit the announcement of discrete network numbers to no more than two to prevent routing table overloads.
- Subnetting as a Solution: It introduces a method to utilize a single network address announcement while dividing the network internally into subnetworks using subnet masks, allowing for efficient internal and external address management.

Challenges and Considerations"
- Compatibility with Older Systems: Some older systems might not support the intricacies of subnetting,

necessitating careful planning and implementation to ensure network compatibility.

In summary, efficient communication within an educational campus and the broader digital community requires strategic planning in message distribution and IP address allocation. The adoption of subnetting offers a solution to the challenge of maintaining an expansive and efficient network while ensuring compatibility with existing internet infrastructure.

Trust Issues

Chapter 14 of the book "Operating the Internet" discusses the evolving challenges in managing the internet's infrastructure, notably trust issues among gateways and the complexities of routing and congestion. In the early days, trust among gateways, which exchange routing information, was implicit under the unified management of DARPA. However, the proliferation of multiple wide area networks under various administrations has introduced concerns about the potential for a rogue gateway to disrupt the internet. Efforts are underway to develop solutions for managing untrustworthy gateways and enhancing the routing of data across multi-homed networks.

The chapter also touches on the issue of capacity and congestion, particularly on the ARPAnet during peak hours. The planned expansion of links aims to address these problems, dictated by the future direction set by the Internet Architect and the Internet Activities Board (IAB). The IAB consists of several committees led by experts overseeing different areas of the internet infrastructure, including autonomous networks, end-to-end services, and privacy, among others.

Routing, a critical function for directing traffic from its source to its destination, is explored through the analogy of a child navigating a restaurant. The chapter elaborates on IP gateways (routers) that facilitate traffic flow between networks based on IP header information and network state. Routing protocols can vary significantly, with some requiring complete network knowledge (the "adult algorithm") and others only a subset, often used in hierarchical networks to avoid loops.

Two types of routing protocols are highlighted: static routing, suitable for small networks or as a default route in networks with a single gateway, and RIP (Routing Information Protocol), which is adapted for IP from the Xerox Network System. While static routing is reliable for simple setups, RIP is more dynamic but best suited for networks with small diameters due to its reliance on hop-count metrics, which can be problematic in networks with links of varying speeds or congestion levels.

The chapter concludes by discussing efforts to improve RIP through documentation and refinement to make it more effective for larger networks, indicating an ongoing evolution of internet management practices to address the challenges of growth, reliability, and complexity.

Gated

In the realm of internet operation, managing the transmission delay across network links presents a complex challenge. Unlike tangible measurements, such as round-trip time, which suffer from varying conditions like congestion or disparate speeds, delay metrics necessitate a nuanced approach. To approximate the time delay, a sophisticated algorithm is essential for coordinating time synchronization among nodes, despite the intrinsic approximation. Routers, specifically Hello routers, adeptly manage this task, maintaining time synchronization across a nationwide network within a precise millisecond range.

The Exterior Gateway Protocol (EGP), delineated in RFC-904, diverges from typical routing protocols by focusing on reachability rather than the quality of connections. It functions as a communication protocol

allowing gateways to indicate network accessibility without specifying the connection's efficacy. Despite EGP incorporating a metric system, its arbitrary nature—ranging from 1 to 8 to signify link quality, with lower values indicating superior quality—lacks formal standardization, limiting its practical utility. Furthermore, EGP's design quirks, which blur distinctions between adjacent metric values, further diminish its effectiveness, relegating its utility to a mere three discernible states and an "unreachable" classification within contexts like the NSFnet.

The coexistence of varied networking protocols—RIP for regional and campus networks, Hello for the NSFnet backbone, and EGP for the Defense Data Network (DDN)—raises the question of interoperability. The early internet resorted to static routing, configured per site using Fuzzball software. However, the static routing approach, while initially sufficient, proved to be brittle in dynamic networking environments. It lacked the flexibility to adapt to changing network conditions, potentially leading to inefficiencies or communication deadlocks when networks turned unreachable. This landscape underscored the growing need for dynamic routing protocols that could adjust to the internet's evolving topology, ensuring robust and uninterrupted connectivity across its multitude of networks.

"Names"

The chapter "Names" from the book "Operating the Internet" delves into the intricacies of how devices connected to the Internet use symbolic names to communicate, navigating the complexity of Internet Protocol (IP) addresses. It begins by explaining the necessity of symbolic names, given the human difficulty with remembering numerical IP addresses. This need led to the establishment of a name register at the Network Information Center (NIC), facilitating the association of human-friendly names with IP addresses.

As the Internet expanded, particularly with the introduction of workstations and microcomputers, maintaining an up-to-date host file became increasingly labor-intensive and bandwidth-consuming. This challenge was addressed by the introduction of the Domain Name Service (DNS), described in RFC-882 among others, which offers a distributed database system to map names to addresses efficiently.

The chapter further unpacks the structure and functionality of domain names, explaining that they are hierarchical and tree-structured, with the root at the right end. For instance, "uxc.cso.uiuc.edu" represents a machine named 'uxc' within the University of Illinois at Urbana's domain, denoted by 'uiuc', under the educational institutions’ domain 'edu'. The process of resolving a domain name involves querying a series of servers, from root name servers down to the specific host, to obtain the possible IP addresses for communication.

Moreover, entities may apply for their own domain name, with the primary condition being the ability to maintain two Internet-reachable machines to serve as name servers for the domain. These servers, not required to be geographically co-located, facilitate the domain's connectivity and management.

Lastly, the chapter introduces the Berkeley Internet Name Domain (BIND) Server, which implements the internet name server for UNIX systems. BIND, integrated with 4.3BSD, supports host name lookups, replacing the traditional “/etc/hosts” file, and continues to evolve with the Internet, offering a platform for reporting operational challenges and discussing future developments.

This overview encapsulates the process by which the Internet manages and resolves names, highlighting the evolution of domain name services and the ongoing innovations to support this crucial aspect of Internet architecture.

Trailers

In the chapter titled "Operating the Internet," the process of how internet protocols operate and manage data transmission, along with specific protocol behaviors and the challenges they face, are detailed with a focus on TCP/IP mechanics, trailer usage in packets, and retransmission strategies.

The chapter begins by explaining the journey of data as it's converted into packets for transmission over the internet. Applications like FTP send data to TCP, which then segments this data, adding a TCP header to each chunk. These chunks are further encapsulated within IP packets, which add their own headers before sending the data across the network. The concept of trailers is introduced as a method aimed at optimizing this process. Trailers, unlike headers, are added at the end of packets with the intention of reducing memory moves on both sending and receiving ends, potentially improving efficiency. However, the chapter points out that this practice wasn't widely adopted or standardized, resulting in operational issues when traversing through gateways that aren't designed to recognize routing information at the end of a data block. This problem is most apparent when trying to transmit long files over networks, leading to transmission failures or hang-ups because many systems and gateways do not properly handle trailers.

Further, the chapter elaborates on TCP's mechanism to ensure data integrity through retransmissions. If an acknowledgment isn't received within a reasonable timeframe, TCP retransmits the data packets. The "reasonable" timeframe is determined by TCP's retransmission algorithm, which, ideally, minimizes unnecessary retransmissions. The chapter contrasts the retransmission strategies of BSD 4.2 and BSD 4.3, noting that the former is prone to overly aggressive retransmissions, especially in environments with high delay and limited bandwidth, leading to increased network traffic. In contrast, BSD 4.3 offers a more balanced approach, quickly retransmitting a few times on the assumption of being on a low delay network before significantly reducing the frequency of retransmissions.

The chapter concludes by providing references for further reading and a comprehensive list of major RFCs (Request for Comments) documents that are fundamental to understanding internet protocols and standards, including those for UDP (User Datagram Protocol), IP (Internet Protocol), ICMP (Internet Control Message Protocol), TCP (Transmission Control Protocol), SMTP (Simple Mail Transfer Protocol), and FTP (File Transfer Protocol), among others. These RFCs serve as crucial resources for anyone looking to delve deeper into the technical specifications and operational principles of internet communications.