Faculdade de Engenharia da Universidade do Porto

U.PORTO

FEUP **FACULDADE DE ENGENHARIA**
UNIVERSIDADE DO PORTO

Software Systems Architecture

Homework #02 - Design the Internet, part 2

Team T23:
Anete Pereira (up202008856)
Bárbara Carvalho (up202004695)
David Fang (up202004179)
Milena Gouveia (up202008862)
Pedro Correia (up202006199)

Introduction

The Internet is essential in today's society given the fact that it allows communication between a large number of devices dispersed around the world. However, the design behind it is frequently perceived as a complex concept and in order to provide a simple explanation of it, McCauley *et al*. wrote an article that will serve as the basis for the present essay in order to address some questions related to the internet's essential design choices that contribute to its success, the rationale behind them, the significance of the title, and finally, we will compare our own suggested design with the real internet.

The Internet's Design Choices

The design of the internet is based on a set of choices made by its creators and developers, where the essential ones are the service model, a four-layer architecture and three crucial mechanisms (routing, reliability, and resolution).

In addition to these elements, the Internet's success can also be attributed to two other crucial decisions highlighted in the paper: the adoption of a modest and general service model and the emphasis on infrastructural autonomy.

The decision to adopt a modest but general service model was crucial in enabling the Internet to accommodate a wide range of applications while leveraging the adaptability of hosts to handle the uncertainties of best-effort Internet service. Unlike traditional networks binded to specific applications, the Internet's service model made no guarantees, allowing new applications to flourish without being constrained by predefined requirements. This approach not only encouraged innovation but also facilitated a rapid increase in network speeds, as the service model remained relatively simple and adaptable. It also embraced a model that allowed for continuous evolution and adaptation along with the technological advancements.

Furthermore, the emphasis on infrastructural autonomy within the Internet's standards-driven framework has been crucial in promoting diversity and innovation in network operations. This autonomy grants entities such as autonomous systems (ASes) and naming domains the freedom to deploy their own technologies, routing protocols, and management policies without the need for extensive coordination with other entities. ASes, for instance, are only required to support IP and participate in the interdomain routing protocol BGP, allowing them to adopt L1 and L2 technologies and intradomain routing protocols without coordinating with other ASes. Similarly, naming domains can define their own name management policies and name resolution infrastructures within the framework of supporting the DNS protocol. This flexibility in deploying and managing infrastructure allows organizations to customize their networks to suit their unique needs while also allowing them to adapt them over time without being limited by rigid standardization.

Rationale Behind the Decisions

As previously stated, routing is one of the three design choices of the internet's architecture. It has an important role when directing information packets to their intended destination and can be observed at the Internet-working Layer, L3, where it is implemented by routers, or at the Network Layer, L2, where it is implemented by switches.

At L3, the routers use a forwarding table in order to determine where to transmit the received packet based on its final destination, and it can handle two scenarios. Firstly, if the router is directly connected to the packet's target network, the packet is simply forwarded to it. Second, if the previous scenario does not occur, the forwarding table takes into consideration the destination address and guides the packet to a router that brings the packet closer to its destination, i.e., the next-hop router.

It's important to note that while having a crucial role in information sharing, the routing mechanism can sometimes encounter challenges that need to be addressed, such as loops. These can occur when a packet goes through a previously visited router and consequently, it will never reach its intended destination.

There are several approaches for avoiding loops, including the **spanning-tree protocol**. It plays an important role at L2 because it alters the networking graph in order to generate a spanning tree, i.e., a cycle-free networking graph. Next, the packet is flooded to all hosts, given the fact that each switch sends it to all connections in the spanning tree, beyond the packet's final destination. This enables the address resolution protocol (ARP), that is used by the switches to memorize the link to a specific host, which has the advantage of reducing the packet flooding, considering that the path between two nodes on a spanning tree is unique and thus, when sending information to a host, we can use the same link where we originally received a packet from that host.

One other insight arises from the adoption of the Domain Name System that translates human-friendly domain names into machine-readable IP addresses. This system operates in a hierarchical structure, starting with a root level at the top, following top-level domains (TLDs), second-level domains, and subdomains. Root servers direct queries to TLD servers, which manage specific domain extensions like .com or .edu. Second-level domains represent entities (e.g., example.com), while subdomains allow for further division (e.g., blog.example.com). This hierarchical and distributed naming system provides a decentralized mechanism for mapping domain names to IP addresses while distributing workload, simplifying administration and contributing to the stability and scalability of the system.

By extracting the complexity of IP addresses to common names it also makes it easier for users to navigate across the Internet. Moreover, the decentralized nature of the DNS guarantees resilience and fault tolerance, as the system can still function even in the face of errors. Overall, the DNS facilitates global connectivity and communication in a user-friendly and efficient manner.

Analysis of the Article's Title

The title "Extracting the Essential Simplicity of the Internet" encapsulates the authors' endeavor to explain in a simple way the core principles underlying the Internet's design. As mentioned in the paper, the authors consider that "*The Internet's design is underappreciated because its beauty is buried beneath an avalanche of implementation details.*", thereby in this paper they propose a structure to simplify the understanding of the Internet by focusing on key design points, revealing its fundamental simplicity amidst apparent complexity.

This perspective aligns with the broader theme of the paper, which endeavors to refine the essential simplicity of the Internet. While the Internet's design principles may inherently embody elegance and straightforwardness, the multitude of protocols, technologies, and operational intricacies involved in its implementation can obscure this fundamental simplicity.

The necessity to simplify the understanding of the Internet also arises from its historical roots in military research. Originating from projects like ARPANET, which aimed to establish resilient communication channels among military and academic institutions, the Internet was designed to meet the demands of robustness, reliability, and ease of use, especially under challenging conditions. In this context, simplicity was paramount for ensuring effective communication and data exchange among diverse entities.

Thus, the title of the paper not only reflects the authors' goal of uncovering the essential simplicity of the Internet but also resonates with the historical imperative for simplicity rooted in its military origins.

Comparison of Own Design with Real Internet

When thinking about the design of the internet, our first approach was to think about how things should work together. This made us feel like creating an architecture capable of assuring reliability, availability and scalability was too hard of a task to accomplish. Consequently, our focus was mainly pointed to explaining how our design choices work, instead of why or what led us to make these decisions. For this reason, when analyzing the real Internet design, we noted the big role that the rationale behind the choices made took in the success of what the Internet is today. Additionally, this methodology allows an easier understanding of the requirements and expectations to accomplish.

Although the real Internet design was made with modesty and simplicity in mind, it showed a high level of detail that our architecture didn't achieve. Right from the start, the architecture introduced in the article is based on the concept of modularity and a layered structure. By breaking down complex systems into manageable components and delineating clear boundaries between layers, scalability becomes a possibility that our design could benefit from. Furthermore, the flexibility and adaptability afforded by the modularity was a key piece missing that would allow our design to support changes and growth overtime. Additionally, it would also bring other advantages like fault isolation, ensuring that the impact of failure remains localized, sparing the rest of the system from undue harm.

Despite the fact that our choice for organization and structuring was quite different from the real system, some base concepts adopted were similar, like for example the definition of network and, on top of that, the attempt to organize the networks hierarchically and in domains based on geographical areas. Network domains play a pivotal role in organizing and managing network resources, providing logical boundaries within which devices and

users can communicate and access shared resources. To facilitate the access to websites and services, our design would benefit from the implementation of an important technology named Domain Name System (DNS) used in the traditional design of the internet. It would serve as the internet's address book, translating human-readable domain names into machine-readable IP addresses, thus facilitating the routing of data across the network.

Going into the topic of routing, we created the concept of dealers that have a role similar to routers and switches in a way that both serve as the backbone of network connectivity and data transmission. However, the definition of dealers lacked some specificity about the forwarding of the information to its destination. A message would contain a sequence of bytes that included a destination ID, the source ID and its contents. Dealers offered a more customizable configuration, dispersing routing across multiple nodes to mitigate single points of failure and enhance network resilience. But in comparison to the real Internet design, this solution showed scalability issues, particularly as the size and complexity of the network increased. Consequently, routers and switches prove to be a better option since it better accommodates the increasing numbers of devices and users. By definition, routers are responsible for forwarding data between different networks based on destination IP addresses, while switches manage traffic within individual networks based on MAC addresses. This notion of addresses provides a robust addressing framework that supports efficient communication and network management.

When it comes to calculating the distance between a router and a network, we opted to use distance in terms of latitude and longitude. While latitude and longitude are commonly used to represent geographic locations on the Earth's surface, they proved to not be the best choice of metrics. The physical distance between two points using this metric may not accurately reflect the network distance or connectivity between them and it easily oversimplifies the complexity of the connections. Considering these factors, our design would improve if we explored alternative metrics such as latency, packet loss and bandwidth availability, as in the real Internet design, which would provide more accurate and actionable insights into network performance.

Conclusions

In summary, this essay elucidates the fundamental design choices behind the internet's architecture, revealing the underlying principles and rationale guiding its development. Through an exploration of service models, layered architectures, routing mechanisms, and infrastructural autonomy, we uncover the intricate framework that promotes the internet's success. Moreover, the analysis of the article's title emphasizes the ongoing effort to extract the internet's inherent complexity into accessible principles, based on its historical requirements for simplicity. By analyzing our own design considerations with the realities of the internet's infrastructure, we recognize the importance of modularity, scalability, and standardized protocols in ensuring the network's adaptability and resilience. Ultimately, this examination underscores the delicate balance between simplicity and complexity in internet design, highlighting the imperative of thoughtful architectural decisions to sustain the internet's vital role in modern society.

References

1. McCauley, James, et al. *Extracting the Essential Simplicity of the Internet*. Vol. 66, no. 2, 20 Jan. 2023, pp. 64–74, https://doi.org/10.1145/3547137.