The Design Principles of the Internet

Qian Liu   2014011216
E-mail: liuqian14@mails.tsinghua.edu.cn
The Department of Electronic Engineering, Tsinghua University

Abstract

The Internet, which has been an essential component of society, is designed based on quite a few principles, mainly including end-to-end arguments, layering principle and so on. This article is firstly intended to explain what these principles are and try to interpret why they are necessary to the Internet. Meanwhile, it has been a hot topic of general interest whether these traditional principles will be still valid for future Internet. In terms of this issue, the article points out that the traditional principles should not be abandoned totally; instead, they are supposed be adapted to the changing Internet environments by some adjustments. In this article, some reasonable suggestions are put forward for the design of the future Internet, as well as some personal comments on these principles.

I.   INTRODUCTION

The Internet is a communication facility designed to connect users together so that they can exchange information. After several decades' evolution, the Internet has been the most indispensable communication means and an essential component of our society. Behind the success of the Internet, however, it is always overlooked by people that the design principles of Internet play an important and crucial role. Without exaggeration, the extraordinary growth and innovation of the internet depends crucially on these principles [1]. We have to say that they are to the Internet what the bones are to a person.

There are so many principles that they cannot be thoroughly discussed in a short article, thus I select two important and influential principles, including end-to-end arguments and layering principle, which will be discussed at length in Section II.

These days, the principles have been greatly challenged. More and more applications designed in different ways are springing up, which are particularly sensitive to the traditional design principles. And there has been a phenomenon that the governments and service providers tend to interfere in the communications on Internet. Meanwhile, the emergence of firewalls, NAT (Network Address Translator) also conflict with some of the original principles. With the appearance of quite a few new requirements for Internet, the design principles are supposed to be reconsidered. Section II will specif-

1

ically talk about how we should deal with these principles respectively.

II.   CURRENT DESIGN PRINCIPLES

This section mainly focuses on the original principles of the Internet, including their basic description, advantages as well as some personal opinions. These analyses will help a lot establish a basic-level understanding of the principles, as well as make good preparations for the anatomy of the new challenges and put forward some suggestions accordingly in Section III.

i. End-to-end Arguments

A)   Basic Description

The end-to-end arguments are principles that suggest how to design an Internet. And it is "arguments" rather than singular form because the end-to-end arguments are a set of principles, mainly including transparency, decentralism and simplicity of deployment.

This set of principles were first proposed and articulated by Saltzer, Reed and Clark in 1984 [2]. Generally speaking, the end-to-end arguments are mainly about how the requirements of the applications should be met in a network system. More specifically, the most crucial question is how the applications and related devices should be designed and implemented into a system. The arguments suggest that the functions for specific applications should not be built into the core of the Internet, which means application-level functions should be built at the end points[1].

---

[1]The end points, also called end systems, is an abstract structure for the ends of the network.

The end-to-end arguments consists of a set of forward-looking opinions. After accepted and recognized, these arguments served as architectural models of the Internet and have had great influence on the development of the network technology.

B)   Why End-to-End?

It is undeniable that no matter how carefully a network is designed, it will still suffer from transmission failure with a certain probability[3]. At least for now, the fact is that there is no perfect way to deal with the inherent problems. Under such circumstances, the best way to cope with the problems is to accept them. Based on this thought, [2] chose to transfer the responsibility to the end points rather than the network itself, which means the end-to-end functions can only be performed effectively by the end points.

C)   Advantages

1)   Simplicity and Generality

The end-to-end arguments emphasizes no discrimination among applications, which indicates the network provides available resources that are not specified to any single application. The simplest application of the end-to-end arguments is totally transparent, with packets going in and coming out, and that is all that happens in the network.

Therefore, the structure of the core network is designed as simple and general as possible. Only the basic function should be preserved inside the network, which brings much convenience to future updates. What's more, the complexity as well as the cost of the core network are both reduced.

2)   Flexibility and Compatibility

2

The end-to-end arguments demand that the functions should be implemented at the end points, thus the network is suitable and operational for all applications. This compatible mechanism help increase the probability that a new application can be inserted or an old application can be modified and replaced without changing the core structure of the network. In the meantime, a variety of applications are allowed to connect and use the network. This flexibility really saves a lot of problems for application developers, because once a new application is invented, it can come into service immediately after installing the code for it on their end points [4].

3) Reliability and Security

Another significant advantage of the end-to-end arguments is the reliability. The end points are much more trustworthy than the network to perform functions of the applications. For instance, TCP (Transmission Control Protocol) is implemented at the end points, thus only the specific end points can get to know the TCP information. Meanwhile, TCP is set to deliver packets in order. And when some packets are delivered unsuccessfully, TCP will try to retransmit them until the packets eventually get through. As for the security, there are certain parts of the packets that are only concerned with end points. These information or data can be encrypted before forwarded, which is normally called the end-to-end encryption [5], so the end points can take the responsibility for ensuring the security and privacy for the information. In a word, these settings enhance the reliability and security of the network.

Conclusion

With all the advantages listed above, it can be explained why the end-to-end arguments should be used. The conclusion can be drawn that all sorts of applications with various functions are supposed to be completely implemented at the end points, and it is rarely necessary to build these applications into the core of network.

D) Personal Comments

When I learned about the end-to-end arguments for the first time, I felt kind of disappointed and skeptical. Actually, I could not help asking: can the foundation of the Internet be so simple? The end-to-end arguments are so simple and superficial that I don't believe it can support the whole architecture and be the fundamental principle of such complicated Internet.

However, after more literature research and deeper reading, I have to admit that I was wrong. The advantages above are sure enough to explain why the Internet should be designed based on this principle. Since the end-to-end arguments were proposed, the Internet has evolved dramatically, but the functions implemented inside the core network still remained rather simple and general. All the specific applications, such as e-mail, the World Wide Web, multi-player games and so on, are always implemented at the end points, which are represented by the computers and smart-phones.

In my opinion, the most profound and lasting influence of the end-to-end arguments is neither how they simplified the Internet nor how reliable the Internet become; instead, I think it is the indirect effects of these arguments that carry more weight. As we can see, the end-to-end arguments have motivated and encouraged many people to innovate and develop multifunctional applications. If the core of the network had been designed and intended for one specific application, the deployment of others would have been inhibited and restrained, and there wouldn't be

3

so many powerful applications today. There is no doubt that the end-to-end arguments have indirectly contributed to the development of modern science and technology. And without any exaggeration, these applications have imperceptibly changed our lifestyle, behavior and even our way of thinking and looking at the world.

ii. The Layering Principle

A)   Basic Description

The layering principle, also known as modularization principle, is another crucial principle in the design of Internet. Actually, not only the Internet, but almost every communication system is designed based on the layering principle. No matter it is TCP/IP model or OSI reference model, the whole Internet is divided into several layers.

Generally speaking, the layering principle is intended to decompose the functions into different layers, each of which offers different services to the higher layer or module [6].

B)   Advantages

1)   The Reduction of Complexity
The layers of the Internet have been designed and set up with highly abstract interfaces. That means one single layer doesn't need to know how other layers operates inside. The only thing it has to be informed is what the functions the higher layer is of and how to use the interfaces provided by lower-level layer. Therefore, the complexity of the Internet structure is greatly reduced.

2)   The Independence of Functionality
The functions of the Internet have been divided and distributed into different layers, which

means the functionality is isolated. For example, IP is designed to compute routes and deliver packets in order to enable the end-to-end communication. And TCP provides reliable tranmission by using congestion control at the end points. And the transport layer offers message service to diverse applications. Therefore, the Internet is actually the combination and the layers with isolated functions.

3)   The Reusability of Layers
We can conclude the reusability from an example: due to the separation of IP layer from other network technologies, IP is allowed to operates in various networks. That means the module is reusable and portable. In fact, the reusability has been a common characteristic of the Internet layers or modules, and it help a lot achieve rapid adaptation and quick fix of the applications.

C)   Personal Comments

I am amazed that the layering principle is so brilliant that even I don't understand the mechanism of the layers at all, I can still make good use of them and utilize them well. And I believe another merit of this principle in engineering application is the simplicity of troubleshooting or rapid repairing. Once some functions of the Internet break down, we can easily find where the faults are resulted from and repair them right inside the specified layer. Therefore, the stability of the Internet is improved by layering principle.

In the meantime, I'd like to reiterate the social influence. To my knowledge, the layering principle has been a common methodology in all kinds of modern systems and works well in various fields. I hold the strong belief that no matter how far it looks like among different disciplines and fields, there are bound to be some generality or similarity. The layering principle is a good ex-

4

ample. All fields should learn from each other to obtain some inspiration, which will push forward the development of human society.

III. THE PRINCIPLES FOR FUTURE INTERNET

Based on Section II, this section, mainly focusing on the usability of current design principles, is organized by questions below[7]:

• *What are the challenges that the current design principles are faced with? And are the current principles out of date for future Internet?*

• *How should the design principles meet the new requirements for Internet in the future?*

i. The End-to-end Arguments

A) Background

Over the last few years, a large number of new requirements have emerged for the Internet. These various requirements can be met through the addition of some new mechanism in the core of the network. Inevitably, the requirements has brought challenges to current design principles of Internet[8]. Will the existing principles be totally abandoned? How can we preserve the benefits of the original design principles as much as possible? The questions will be answered later. Before that, it should be first pointed out what exactly the new challenges are.

B) What Are the Challenges?

1) The Advent of Third-party Intervention

Originally, the communication or packet delivery is limited between two end points. Nevertheless, for some reason, some third parties demand that they have the authority to impose themselves between these communicating end points regardless of the desire of the end points[3]. This definitely violates the end-to-end arguments. As we can see, the third parties with different purposes include some organizations and the governments focusing on public safety. In such a situation, the end-to-end arguments are no longer applicable, and the security and reliability will be seriously destroyed.

2) The Rapid Increment of Applications

As is known to all, the Internet makes no guarantee about the packet delivery. What it can guarantee is to spare no effort to do that. When the Internet was first put into practice, the applications were so tractable that the end-to-end principles can basically meet the requirements of communications. However, lots of new applications are emerging, most of which deal with audios and videos. Under such circumstances, a much more stable and wisdom Internet service that can cope with those applications is highly demanded.

3) The Reduction of Trust and Reliability

Traditionally, the end points are willing to cooperate with each other to accomplish packet delivery or data transmission. But there have been an exponential growth in the population of Internet users, and the components of the Internet users become complicated. It is not hard to find that some people have the tendency to misuse the Internet to benefit themselves[9]. As a consequence of that, there has been less reason for normal users to trust the other end points we want to communicate with, especially the anonymous ones.

Conclusion

These days, the end-to-end arguments have been confronted with a serious situation, and it still remains to be determined whether they should be preserved or abandoned. But what we have known is that if there are chances that the

end-to-end arguments be preserved, they definitely need to be augmented or extended[10].

C) How to Deal with the Challenges

Here are some strategies or ideas to cope with the given challenges respectively.

a) The issue raised by third-party intervention is quite tricky and intractable. If only we can reject the unreasonable demands of the third party! However, if we cannot reject it, the only way might be finding another design approach which can preserve the advantages of the end-to-end arguments as much as possible[11].

b) As for the dramatic increment of applications, it is highly demanded that the packet delivery should be divided into several stages. For intermediate server in each stage, the delivery should be treated traditionally in order to assure the data can be transmitted successfully.

c) With regard to the reduction of reliability, more mechanism should be implemented inside the network[12]. Although it may complicate the structure of the core network to some extent, but the end-to-end arguments are still valid if we regard the added mechanism as a basic function inside the network.

To summarize, the end-to-end arguments are still valid and powerful, but they need to be properly extended and augmented to adapt to the new situation.

D) Personal Comments

The Internet is becoming more and more complicated, and the new requirements for the Internet are much tougher to meet. Though the end-to-end arguments are confronted with unprecedented challenges, it doesn't mean that they should be abandoned. From the interpretation above, they are still valid for the future Internet with some extensions and adjustments. Whatever, the valuable nature of the Internet deriving from the end-to-end arguments, such as encouraging innovation, should be always preserved.

ii. The Layering Principle

A) What Are the Challenges?

Based on the layering principle, the functions of each layer or module will be carried out completely before the process goes to the next layer. This indicates if some optimization needs to be done, it can only performed inside the layer separately[11]. Such constraints seems to be quite unwieldy if we want to implement the functions in an as efficient and optimal as possible way.

Besides, new things are springing up vigorously, such as firewall, NAT and so on. Unfortunately, they have violated the layer principle because they can inspect and even change the data inside the packets, which should have been out of their reach. Actually, the violation right comes from the fact that the layers separate networking functionality.

B) How to Adjust Accordingly

The layering principle used to be overly strict, which seems inappropriate today. Based on the new situation, the layering principle should be loosened[13]. For example, in order to perform some functions more efficiently, the functions of one single original layer can be repeated across multiple layers if necessary. And the relationship between different layers should be emphasized because the layers highly depend on the existence of other layers today. By introducing a new mechanism[14], we should be confident that

the challenges will be tackled properly in the future Internet.

IV. SUMMARY

This article focuses on two of the main principles of the Internet and points out how they contributed to the success of the Internet. Meanwhile, the article also analyzes the new challenges for the principles and try to give some reasonable suggestions. From my perspective, I think these principles should always evolve along with the changes of society and times. The change is continuous in Internet forever, and the design principles of the Internet also inevitably change. To quote from B. Carpenter[1]:

Principles that seemed inviolable a few years ago are deprecated today. Principles that seemed sacred today will be deprecated tomorrow. The principle of constant change is perhaps the only principle of the Internet that should survive indefinitely.

What should be emphasized is that no matter what the Internet will become in the future, all the design principles are meant to meet the requirements of society and improve the quality of peoples life. After all, science and technology are destined to provide people with better life, and the Internet is no exception.

References

[1] Carpenter, B. "Architectural principles of the Internet." *Lecture Notes in Computer Science* 11.3(1996):1867.

[2] Saltzer, J. H., D. P. Reed, and D. D. Clark. "End-to-end arguments in system design." *Acm Transactions on Computer Systems* 2.4(1984):277-288.

[3] Blumenthal, Marjory S., and D. D. Clark. "Rethinking the design of the Internet: the end-to-end arguments vs. the brave new world." *Acm Transactions on Internet Technology* 1.1(2001):70-109.

[4] Brwolff, Matthias. "End-to-End Arguments in the Internet: Principles, Practices, and Theory".*CreateSpace*, 2010.

[5] Zeidler, Howard M. "End-to-end encryption system and method of operation." US, US4578530. 1986.

[6] Clark, D. "The design philosophy of the DARPA internet protocols." *Acm Sigcomm Computer Communication Review* 18.4(1999):106-114.

[7] Shenker, S. "Fundamental design issues for the future Internet." *Selected Areas in Communications IEEE Journal* on 13.7(1995):1176-1188.

[8] Clark, D. D., et al. "Tussle in cyberspace: defining tomorrow's internet." *IEEE/ACM Transactions on Networking* 13.3(2005):462-475.

[9] Moors, T. "A critical review of 'End-to-end arguments in system design'." *IEEE International Conference on Communications, Icc* 2000:1214 - 1219.

[10] Lemley, Mark A., and L. Lessig. "The End of End-to-End: Preserving the Architecture of the Internet in the Broadband Era." *Ssrn Electronic Journal*48.4(2000):925-972.

[11] Mller, and P. All. "Future Internet Design Principles." *European Commission - Information Society and Media.*

[12] Clark, David D., and M. S. Blumenthal. "The End-to-End Argument and Application Design: The Role of Trust." *Social Science Electronic Publishing* 21.2(2007):357-390.

[13] Clark, David D., et al. "Addressing reality: an architectural response to real-world demands on the evolving Internet. " *Acm Sigcomm Computer Communication Review* 33.4(2003):247-257.

[14] Ford, A., P. Eardley, and B. V. Schewick. "New Design Principles for the Internet." *IEEE International Conference on Communications Workshops*, 2009. ICC Workshops 2009:1 - 5.