Workgroup: opsawg
Internet-Draft:
draft-jilongwang-opsawg-cybersmap-07
Published: 27 November 2022
Intended Status: Informational
Expires: 31 May 2023
Authors: WJL. Wang, Ed.
Tsinghua University
MCC. Miao, Ed.
Tsinghua University
ACQ. An, Ed.
Tsinghua University
ZSY. Zhuang, Ed.
Tsinghua University

Design of the native Cyberspace Map

Abstract

This memo discusses the design of the native cyberspace map which is stable and flexible to describe cyberspace. Although we have accepted the cyberspace as a parallel new world, we even have not defined its basic coordinate system, which means cyberspace have no its basic space dimension till now. The objective of this draft is to illustrate the basic design methodology of the native coordinate system of cyberspace, and show how to design cyberspace map on this basis.

Status of This Memo

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at https://datatracker.ietf.org/drafts/current/.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

This Internet-Draft will expire on 31 May 2023.

Copyright Notice

Table of Contents

1. Introduction

There is a new space created by Internet, together with computer networks, telecommunication networks, termed as cyberspace. It is an interactive domain that includes users, softwares, processes, information in storage or communication, applications, services .etc. Unfortunately, we even have not defined its basic coordinate system and even the native map.

Traditional well known coordinate systems seem feasible to visualize and represent cyberspace. However, both coordinate systems have some drawbacks. Although geographic coordinate system(GCS) vividly shows geographic information of cyberspace in geographic map, it only visualizes a tip of iceberg of cyberspace and hardly describes the characteristics of cyberspace (e.g. host, service) all at the once from cyberspace point of view. Network coordinate system (NCS) focuses on visualizing network topology with node representing host (or IP address) and edge representing network distance between two

hosts. NCS tries to represent and visualize cyberspace from network perspective. It is easy to hierarchically represent different parts of cyberspace in network topology map. However, NCS is a frequent change network due to distance changes and host connection status and it is difficult to visualize the whole cyberspace.

This demo discusses and defines a native cyberspace coordination model based on AS number and IP address following the principle of robustness, orthogonality and effectiveness. It can present cyberspace in a concise and intuitive manner and user can easily filter out the specific details of interest. Based on our cyberspace coordination model, we also propose a prototype system of native cyberspace map which can be used as the basic tool for network management, network security and network resources search .etc. The firstly proposed overall design methodology can help to establish the native cyberspace map as a unified backplane for visualization in the future.

1.1. Requirements Language

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in RFC 2119 [RFC2119].

2. Terminology

This document does not describe standard requirements. Therefore, key words from RFC 2119 [RFC2119] are not used in the document.

Manager:An entity that acts in a manager role, either a user or an application. The counterpart to an agent. A 'management client' in NETCONF terminology.

IANA:Internet Assigned Numbers Authority, an organization that oversees global IP address allocation, autonomous system number allocation, media types, and other IP-related code point allocations.

Different granularities of cyberspace: representing the degree of visual cyberspace such as AS, Metropolitan area network, Local area network, IP blocks .etc.

Network resources: including physical resources such as traditional network facilities and access devices, as well as virtual resources such as application services and information resources, which can be detected using software or hardware tools based on certain methods, techniques and standards

3. Use cases

Our cyberspace map CAN provide a unified drawing backplane, and express the cyberspace in a multi-scale, multi-dimensional and multi-view way. Drawing the measured network data on the unified backplane CAN be skillfully applied to the expression of network resources, the monitoring and management RFC 1052 [RFC1052] of network elements and the prevention of cyberspace security, etc. The following sections highlight some of the most common framework for native cyberspace map use case scenarios and are in no way exhaustive.

3.1. Network Management

Network resources management: The main concern of network managers is to have a direct and macroscopical visualization of network resources, so that they could manage network resources efficiently. In other words, based on the different sizes of network they manage, network managers have the demands to visualize network resources at different granularity. For example, network carriers mainly focus on the AS-level network and consider the resources with IP blocks, while the campus network administrators take care of the local area network and manage the resources at the specific IP addresses. Fortunately, our following Cyberspace map provides the ability to show the different granularities of cyberspace by setting the order n of Hilbert curve mapping algorithm.

Network traffic monitoring:Network traffic contains the information of IP addresses RFC 791 [RFC791] and port. Therefore, the representing of network traffic in our cyberspace map is helpful for network managers to monitor the current network traffic status and realize network anomaly detection concisely and intuitively. At the large network level, monitoring traffic exchange between ISP networks is helpful to understand network traffic status, to realize quality of service analysis and congestion prediction, and to achieve reasonable bandwidth allocation between large networks. At the LAN level, regional traffic analysis is helpful to extract user network behavior characteristics. For example, monitoring TCP135 port traffic activity of target IP and discovering potential infection mode of Blaster worm CAN prompt closing abnormal host port to repair vulnerabilities for security management.

3.2. Network Security

At present, network security problems are emerging one after another RFC 3631 [RFC3631], how to detect and visualize these phenomena has always been the focus and difficulty of the network security and management field. Instead of physically attacking the physical host of geospatial, the security attacks usually involve virus infection

against IP addresses and the vulnerabilities of corresponding hosts or perform DDoS attacks on specific IPs. Therefore, the traditional geographic coordinate system is difficult to reveal the original attack form of network.

Our cyberspace map based on IP addresses CAN reveal security issues from a higher level. In detail, it CAN intuitively express the distribution of DDoS attackers and attacked IP addresses, and further express the spread of infected IP addresses. To Assist security analysts to better understand and prevent attacks, effectively cut off the infection transmission path, and implement attack shielding and prevention. In addition, by telescopically displaying more specific information such as the AS, Network, and Organization to which the attacker IP belongs, it CAN help the corresponding network security administrators carry out effective vulnerability repair.

4. Selection on Basic Coordinate Vectors

It is still suffering a big challenge to construct a native coordinate system, given the large amounts of network data and the ability to represent sufficient level of detail of interest to the different level of administrators. To tackle these problems, we look for the stable numbering system (coordination) in cyberspace as the basic coordinate vectors to construct the cyberspace coordinate system. With deep understanding of cyberspace, we observes a number of alternative choices such as IP address space, Autonomous System (AS) number space RFC 4983 [RFC4983] , MAC address space, Domain name space RFC 1034 [RFC1034] and port number space RFC 6056 [RFC6056]. These coordinates are stable and widely adopted that almost all objects in cyberspace possess them as identifiers so that they are able to project the cyberspace in its own space. We are discussing each coordination in the following:

4.1. IP address

An IP address is a unique fingerprint assigned to each host when connecting to network. It serves two primary functions. It is used as a network interface identification of host and it also provides the location of that host in cyberspace, similar to a physical address(longitude and latitude) in geographic space. An IP address is a unique address that makes it very suitable as a base vector in cyberspace. It locates host and allows host to send and receive information and communicate with a specific host in cyberspace. An IP address is composed of a fixed bit number, the total number of IP address is constant. Since the total number of IP address doesn't change with network status, it is a robust vector in cyberspace, defined as Address Space.

4.2. Port

An port number is composed of a 16-bit binary number with the fixed total number. An port number is often come up with an IP address when establishing a connection and is orthogonal to IP address. An IP address is the network address of a host in address space, while port number is the logic address of a specific service in that host. For instance, an address may be "IP address:216.38.1.15,port number: 80", written as 216.38.1.15:80 which represents a web service on a specific host. An port number combining with an IP address locates relevant information in cyberspace at a finer granularity. While the total number of port also doesn't change with network status and it is orthogonal to address space, it is a suitable and robust vector for representing and visualizing cyberspace, defined as Logic Space.

4.3. AS number

ASN, defined for routing policy on the internet, is a collection of connected IP under the control of network operators. The AS number is composed of a 16-bit binary number with the fixed total number and the AS number is also a stable numbering system. Each AS contains a set of IP addresses and the relationship between IP address and AS are operated by RIRs. Therefore, AS is also regarded as the location of aggregated objects in cyberspace. Projecting the cyberspace into AS space provide the aggregated characteristics of IP address space. It is also an effective way to demonstrate cyberspace if the viewer want to visualize the AS level information of cyberspace such as the AS topology.

4.4. MAC Address

MAC address, defined as Media Access Control Address, is a unique identifier of network interfaces through a physical network segment. In other words, it's an identifier of hardware that uses Ethernet, which can also be referred as physical address or hardware address. Since the MAC address is the stable numbering system that is composed of 12 characters, so it could be used for the coordination of cyberspace. Furthermore, the cyberspace is created by the physical network resource with MAC address, so that we can project the cyberspace into MAC address space which is traced into each physical host.

4.5. Domain Name

Domain name is alphabetic which is easier to remember. For example, the domain name has a formed name e.g. www.apple.com, which is the identification of Apple company. Domain name is a stable numbering system which is not change with network status, however, it is impossible to enumerate because the length of domain name can be

variable. Projecting the cyberspace into domain name space only provide the detailed web information of cyberspace.

4.6. Conclusion

We discuss some alternatives that can be used as network space coordinates. Each coordinate is a candidate for constructing a cyberspace coordinate system. Obviously, projecting network space to MAC address space and domain name space is not very effective, which may lead to poor visualization of cyberspace. The former may lead to sparse visualization, because most MAC addresses are not connected to the Internet, while the latter only provides detailed network information considered as a small part of the cyberspace. As for IP address space, port space and AS space which can be regarded as the location of object in cyberspace, they can be selected as the basic coordinate vectors to demonstrate cyberspace.

5. Construction of native Cyberspace Map

After determining the basic coordinate vectors, i.e. IP address, port and AS, the specifications for the design of cyberspace maps based on these coordinates will be described in detail. Similar to ground military systems with 2-D horizontal coordinates or 3-D Cartesian coordinates, we define three types of map suitable for different scenarios.

5.1. IP Map

Effectively presenting the IP address in our IP map is an extremely challenging problem for decades. One of the primary causes of this problem is that the total unique IP addresses is about 4 billion (IPv4), each of which needs to be visualized in the map. We have to make creative use of various techniques, and it is also significant to visualize IP addresses with meaningful aggregations where possible. The one-dimensional IP map expresses the network elements in the form of lines and points discretely and unintuitively. Therefore, we introduce the space filling curves to design a unified drawing backplane, and realize the association mapping between one-dimensional IP address space and two-dimensional IP address space. That is, the network is gathered to two-dimensional space plane with length and width are both the n-th power of 2, where n represents two-dimensional space order. The space filling curves mainly include Z curve, C curve, Gray curve, Hilbert curve.

Hilbert space algorithm is optimal for the continuity and regional of space filling. It can shows a two-dimensional visualization of an IP block of 10.0.0.0.0/24, where the IP sub-blocks of 10.0.0.0/26,10.0.0.64/26,10.0.0.128/26 and 10.0.0.192/26 are adjacent. The Hilbert curves CAN provide people the ability to view

cyberspace elements in aggregated or non-aggregated mode. For non-aggregated mode, the IPv4 address space REQUIRED the order n equals 32, which is preferable when detailed IP addresses need to be examined. While for aggregation mode, the order n needs changing for visualizing different granularities of cyberspace elements, which is beneficial when viewing data from an AS or a network backbone. For example, prefix 10.0.0.0/16 CAN be aggregated to a grid with setting the order equal to 8. Based on the Hilbert curve, the IP address could be extrapolated from one dimension into two dimension to generate the 2-D IP Map with coordinate(X,Y).

It CAN be used in various security-related applications, such as network resources management, Internet interruption and secret scanning of Botnet coordination. compared to the geographic coordinate system ,it CAN realize the search, positioning and description of managed elements at different network levels (AS, Network, Organization, IP address) instead of continuously zooming in geographic locations without a clear network hierarchy. It CAN represent multi-aspect information of cyberspace all at the once. In additional, benefit from the regionality and aggregation of our coordinate system, the administrator CAN perform unified management and configuration and operates on IP address blocks of key resources such as links and backbone networks.

5.2. IP-Port Map

In order to represent the detail information for cyberspace, it can extent the basic two-dimensional spatial plane drawn by the Hilbert curve mapping algorithm into the three-dimensional map by adding the logical port orthogonal to the IP address. Although the basic coordinate system constructed by the IP address can better locate the cyberspace elements to the corresponding hosts and visualize the IP attribute of the them, it would be difficult to describe cyberspace from different cognitive perspectives such as services, which are of great interest to people. Therefore, aside from the IP address, the logical port is RECOMMENDED to be used effectively to visualize cyberspace by constructing the 3-D IP-Port map.

Specifically, the port numbers from 1 to 65536 CAN be represented on the z-axis and the height of each item CAN be used to visualize the traffic data of this port. In this three-dimensional IP-Port map, the traffic volume data that people concern about can be easily represented to perform diagnosis of flow anomaly. In addition, the different network aggregation of traffic data can be simply realized by zooming in/out. It CAN reflect the cyberspace elements more accurately and comprehensively compared to the two-dimensional IP map. It also CAN be used for application layer management, such as abnormal application monitoring and application layer traffic monitoring.

5.3. AS Map

The above IP map and IP-Port map constructed based on the IP address can better express cyberspace in most scenarios. They visualize the essential characteristics of the cyberspace (IP dimension space) compared to the geographic map, and retain the adjacent attributes between the IP addresses,express different granularities of cyberspace IP address prefixes, services, traffic .etc in aggregated or non-aggregated mode. In additional, the inherent existence of the IP address makes them more stable than the topological map. However, in some scenarios, such as representing the network traffic and attack characteristics of an AS in cyberspace, the assignment of IP address segments under an AS MAY be discontinuous, resulting in poor visualization of the IP address-based map, although continuous IP addresses remain adjacent through the Hilbert curve.

Here we define a native AS map model to represent cyberspace. Similar to the IP map, we use the Hilbert mapping algorithm to visualize the one-dimensional ASN, and construct the two-dimensional coordinate plane(2-D AS Map) to represent the AS information, which is similar to the expression of national information by latitude and longitude in the geospatial model.

Next, considering the IP address is a critical element of cyberspace, we also construct the 3-D IP-AS map model. The allocation time sequence of the IP address under the AS is RECOMMENDED to be a third-dimensional basic vector, which is orthogonal to the AS address, and its positive direction indicates the sequence is increasing, realizing the analysis and mapping of the IP address in cyberspace. Specifically, the Z-axis mapping algorithm is defined as follows:

Input : an IP address P

Output : the coordinate of Z-axis

1. Get the AS where the address P is located based on the IP database.

2. There are n IP addresses IPs=[IP1,IP2,IP3,IP4,IP5,IP6,...,IPn] under this AS, and their corresponding allocation time is T=[T1,T2,T3,T4,T5,T6,...,Tn], where the unallocated IP address allocation time is defined as MAXINT > max Allocated time[T1,T2,T3,T4,T5,T6,...,Tm], accurate to the second.

3. for i from 1 to n:

4. dict[IPs[i]]=T[i]

5. dictnew=sort(dict)

6. z= dictnew.index(P)

7. return z

z=10000 indicates that an IP address is located at the 10000th position after being sorted according to the allocation time. According to the Hilbert algorithm and the Z-axis mapping algorithm, the positioning coordinate (X, Y, Z) are used to analyze and map an IP address, and many cyberspace resource elements can be located based on the key identification IP address of communication.

Instead of representing the topological relationship using abstract points and lines, it provides the ability to describe and express in a detail and native manner compared to the map of Internet topology. At the same time, the AS backplane is fixed so that some changes in links will not affect the entire map, which also reflects the superiority of AS Map.

6. Acknowledgements

The authors would like to thank the support of Tsinghua University and National Key Research and Development Program of China under Grant No.2016YFB0801301 and 2016QY12Z2103.

7. IANA Considerations

This memo includes no request to IANA.

8. Security Considerations

This document only defines a framework for network resources categorization. This document itself does not directly introduce security issues.

9. Normative References

[RFC1034] Mockapetris, P., "DOMAIN NAMES - CONCEPTS AND FACILITIE", RFC 1034, November 1987, <https://www.rfc-editor.org/rfc/rfc1034>.

[RFC1052] Cerf, V., "IAB Recommendations for the Development of Internet Network Management Standards", RFC 1052, April 1988, <https://www.rfc-editor.org/rfc/rfc1052>.

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", RFC 2119, March 1997, <https://www.rfc-editor.org/rfc/rfc2119>.

[RFC3631] Bellovin, S., "Security Mechanisms for the Internet", RFC 3631, December 2003, <https://www.rfc-editor.org/rfc/rfc3631>.

[RFC4983] Vohra, Q., "BGP Support for Four-octet AS Number Space", RFC 4983, May 2007, <https://www.rfc-editor.org/rfc/rfc4983>.

[RFC6056] Larsen, M., "Recommendations for Transport-Protocol Port Randomization", RFC 6056, January 2011, <https://www.rfc-editor.org/rfc/rfc6056>.

[RFC791] Postel, JB., "Internet protocol", RFC 791, September 1981, <https://www.rfc-editor.org/rfc/rfc791>.

Authors' Addresses

Jilong Wang (editor)
Tsinghua University
Beijing
100084
China

Email: wjl@tsinghua.edu.cn

Congcong Miao (editor)
Tsinghua University
Beijing
100084
China

Email: 1010988944@qq.com

Changqing An (editor)
Tsinghua University
Beijing

100084
China

Email: acq@tsinghua.edu.cn

Shuying Zhuang (editor)
Tsinghua University
Beijing
100084
China

Email: 17751034616@163.com