TECHNICAL PUBLICATION

SVTA5010: Geo-Data for IPv6

Created and Approved by the Streaming Video Alliance

October 14, 2019

WORKING GROUP:

Geo (Sub-group of Networking & Transport)

GROUP CHAIR(S):

Jason Lee (Digital Envoy)
Ben Jones (Neustar)

PROJECT LEAD(S):

Jason Lee (Digital Envoy)
Ben Jones (Neustar)

1. CONTRIBUTORS

The following people contributed to this document:

- Glenn Deen (Comcast)
- Jason Lee (Digital Envoy)
- Yves Boudreau (Ericsson)
- Dan Alexander (Comcast)
- John Leddy (Comcast)
- Ben Jones (Neustar)

Notice:

This document has been created by the Streaming Video Technology Alliance. It is offered to the Alliance Membership solely as a basis for agreement and is not a binding proposal on the companies listed as resources above. The Alliance reserves the rights to at any time add, amend or withdraw statements contained herein. Nothing in this document is in any way binding to the Alliance or any of its members. The user's attention is called to the possibility that implementation of the Alliance agreement contained herein may require the use of inventions covered by the patent rights held by third parties. By publication of this Alliance document, the Alliance makes no representation or warranty whatsoever, whether expressed or implied, that implementation of the specification will not infringe any third party rights, nor does the Alliance make any representation or warranty whatsoever, whether expressed or implied, with respect to any claim that has been or may be asserted by any third party, the validity of any patent rights related to any such claim, or the extent to which a license to use any such rights may or may not be available or the terms hereof.

2. ABSTRACT

There are many different approaches to associating attributes to an IP address and many different attributes that can apply. There are also several different approaches to delivering that data. This document will focus on attributes that fall into three categories: Identity, Service, and Location. Identity can include a user's name, the upstream service provider, an enterprise network, university, or department. Services can include attributes such as a W-Fi network, cable, wireless, infrastructure, and enterprise. Finally, location attributes can identify city, state, zip codes, country, region, or geocodes. A collection of attributes would form objects with their associated values, and these objects can be related to an individual IP address or a range of addresses to create an IP addressing object. These IP address objects can also be tied to an IP address object in a parent child relationship to provide as much detail as desired by the source. This document provides a JSON object model and schema to represent how those attributes can be tied to an IP address in a common format for controlling the access to streaming video using geo-location data for IPv4 and IPv6 addressing.

Note: access controls, privacy, and the management of privileged information (PII) will not be defined here and will be left to other efforts.

2.1. Versioning

Version	Version 1.0	**Date**	October 14, 2019
• Approved initial version of the document			
Version	Version 1.1	**Date**	March 10, 2020
• Document rendered into new template. • Additional minor edits made. Added Objectives and Scope section. • Added Conclusion.			
Version	Version 1.2	**Date**	August 21, 2023
• Added new SVTA document number • Updated headings to reflect new numbers			

Table 1: Versioning

2.2. Requirements

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 RFC2119.

TABLE OF CONTENTS

3. OBJECTIVES AND SCOPE

The objectives of this document are:

- To propose a specification for Geo Data representation in JSON format that addresses both IPv4 and IPv6.

3.1. Scope

The scope of this document covers:

- A proposed JSON schema to represent Geo Data for IPv4 and IPv6 for streaming video

The scope of this document DOES NOT cover:

- Implementation methods for the JSON schema

4. INTRODUCTION

With the growth of video content distribution via streaming, it is important to understand how the delivery of this content can be restricted and controlled. Gaining a better understanding of whom, what, or where content is being requested from is critical to the user experience as well as to meeting rights holder requirements.

Location services, though, are nothing new. There are many algorithms that leverage routing, mobile data, and many other details to create a profile of whom, or what is using an IP address, and from where it is being used. These location services however are not perfect and can sometimes result in erroneous content blocking or other events that adversely affect the viewer experience. In fact, customer service issues still occur with professional sports blackouts, customer validation checks, and other issues resulting from problematic location data.

Network operators and video content distributors can benefit from a consistent way in which to publish IP addressing information that can be consumed by content delivery networks. This data can also be leveraged internally within an organization to improve network management. But this consistent approach must also take into consideration IPv6 addressing. Many network operators are migrating to this addressing approach as the availability of IPv4 addresses dries up which is especially problematic as internet-connected devices requiring a networking connection continue to skyrocket. As network operators deploy IPv6 addressing both within their own networks and to their subscriber devices, it becomes paramount to have a location-data solution that supports both IPv4 and IPv6.

A consistent format for IPv6 geolocation data interchange will help facilitate seamless transactions if the network operators, content delivery networks, third-party service providers and all others involved in the ecosystem can agree on a standard for data exchange.

The proposed JSON format would be one standardized way to interchange IP address geographic and network information. The format would be mostly intended for transfer and ingestion of IPv6 data and thus would typically be used as the response to a corresponding API request. It is not recommended as a storage format. Databases or Comma Separated Value (CSV) files will be used for IP management and storage.

This document details the three main attribute groups that will represent the IP address or prefixes key information, each containing a handful of data fields to specifically

describe. It is important to note that none of the data fields associated to IP address ranges will uniquely identify an individual or place of residence.

Although the focus on this document is for an IPv6-first specification, the proposed schema works for IPv4 as well. It's been over eight years since World IPv6 Launch day on June 6th, 2011, and the internet is still far from its completion to an all IPv6 Internet, but standards like the one proposed can help drive worldwide adoption and, more specifically, reduce friction for service providers that need a clear mechanism for delivering geo-location data within a streaming video experience given the potential for complex combinations.

5. ACRONYMS AND DEFINITIONS

This document may employ the following acronyms and other terms:

Abbreviation	Description
CIDR	Classless Inter-Domain Routing
GIS	Geographic Information System
HTTP	Hypertext Transfer Protocol
IP	Internet Protocol
JSON	JavaScript Object Notation
OTT	Over the Top
PII	Personally Identifiable information
RDAP	Registration Data Access Protocol
RIR	Regional Internet Registry
SP	Service Provider
STB	Set Top Box
VOD	Video on Demand
API	Application Program Interface
CDN	Content Delivery Network

Table 2: Acronyms and Definitions

The following are some definitions relevant to understanding the application of this document:

- **IP Address Range**—a range of contiguous IP addresses that can be identified as an individual IP address, by a start and end IP address, or by CIDR notation.
- **Object**—defined in RFC7159 as it related to JSON, an object is an unordered collection of zero or more name/value pairs, where a name is a string and a value is a string, number, Boolean, null, object, or array.

- **Entitlements (aka, Rights or Usage Rules)**—the permissions and restrictions associated with a content item or access to a service, the enforcement of which prevents unauthorized access and/or distribution.
- **Type**—the decimal value of the protocol version as defined by IANA. Given this is an IPv6 specification the number should equal 41 if there are no tunnel extension headers.

6. BUSINESS DRIVES AND APPLICATIONS

Some of the business models and business drivers in the television and streaming video industry add additional complexities in the OTT environment. These complexities can be overcome if they are considered with technologies applied to solve these challenges while maintaining (and even evolving) current and future business arrangements between distributors and content rights holders. This document addresses the following list of business drivers and considerations which must present the ability to authorize consumption of content based on the licensee's content rights:

- Geo-fencing and Geo-blocking
- Targeted and mobile advertising
- Content Delivery Network targeting

Note: the above list only represents three of the most common use cases. Others exist which are not covered in this document.

6.1. Geo-fencing and Geo-blocking

Content (i.e., television programs, movies etc.) is typically licensed by specific geography and therefore has "rights" to licensees that must be enforced by distributors and OTT providers to meet those licensing terms. In the past, television services weren't as mobile as they are today, and service was traditionally delivered to a fixed location with a fixed address. Enforcement of the licensing geography was much easier to administer to fixed locations with the Pay-TV provider having access to and enforcing first party data to geo fence the location of content distribution.

As content consumption became mobile via the internet and through such devices as laptops, smartphones, and tablets, the need to enforce these distribution rights and entitlements has become more challenging. More content is being licensed for "digital" (internet) distribution and is using IPv4 as the underlying protocol. Some companies, such as Quova (Neustar)[1], Maxmind[2], and Digital Element (Netacuity)[3] sprouted up to attempt to build databases that mapped IPv4 CIDR blocks to specific geographies. These

[1] https://www.home.neustar/security-intelligence

[2] https://www.maxmind.com/en/home

[3] https://www.digitalelement.com/

proprietary databases are used heavily to block and restrict content consumption to the proper geographies so that distributors can enforce their content licensing obligation.

These solutions are not perfect by any means but are the standard used today and are mature for IPv4. Pay TV operators have also added additional technologies to their platforms that augment these IPv4 databases to prove (and restrict) content consumption to the licensed geographies by proving that a customer is behind operator-controlled customer premise equipment such as DSL or a cable mode. GPS has also been a great tool (when available) to provide anonymous latitude and longitude information which further adds to the toolkit for proper determination of geography to content rights mapping. IPv6 however, in its adoption stage, doesn't provide the accuracy yet to be used by most TV OTT providers to enforce geography-based content licensing enforcements.

6.2. Targeted and Mobile Advertising

The current IP geo location databases are also used by targeted advertising technology companies to manage their advertising campaigns. The database is used to determine a user's general location in order to provider relevant local ads to that user instead of irrelevant ads for products or services that aren't available in their area. This targeted advertising capability has emerged in the last 10 years as a much more effective way for some advertisers to reach their core audience and provide more timely information on products and services available in their area, typically as granular as city and zip. These databases continue to become more specific and more granular as they use additional information and machine learning to determine location of a user at any point in time.

GPS information mapped to IP information can be a powerful tool to build a further, refined view of a user's location and draw a historical graph of user mobility that far exceeds the primitive use of address or latitude/longitude as indicators of current location. IPv6, heavily used in mobile networks today, still does not provide an acceptable accuracy of location but is improving, especially for devices that do not have GPS capabilities.

6.3. Content Delivery Network Targeting

Content Delivery Networks (CDNs) distribute and cache content to accelerate the delivery of resources to client devices or APIs. The objective of a CDN is to deliver that requested content from a more geographically closer location to the end user than the

origin server maintaining an authoritative copy of the data. This requires the use of IP data to direct users to the closest edge node.

As more companies employ multiple CDNs to help deliver such content (like a streaming video), they must make delivery decisions about which CDN to use based on performance- and other-deterministic data (i.e., reliability, responsiveness, scalability, etc.). In the scenario of multi-CDN use, IP data is employed by the CDN to redirect users to their closest geographical or topological (i.e., number of router hops) location. Some authoritative DNS providers offer responses based on the source of the request, using specifications like EDNS Client Subnet, which ultimately have implemented some IP geo data.

7. DATA OBJECTS

There are several attributes which can provide information about whom, what, and where an IP address is being used. These attributes can be placed into three categories which can be applied in any number of frequency or combination. These three categories include Identity, Location, and Service. The three categories are then grouped together into an IP address object that associates the details to an IP address, or range of addresses. An example JSON document which encapsulates this structure and data relationship is provided below[4]:

```
{
  "IP_Address": {
    "title": "IP Address",
    "IP": {
      "Type": "type": "[IANA Assigned Internet
Protocol Numbers]",
      "IP_Addressrange": "value",
      "Start": "value",
      "End": "value",

      "Identity": {
        "values": {
          "country": "value",
          "region": "value",
          "city": "value",
          "postal_code": "value",
        }
      },

      "Location": {
        "values": {
          "postal_code": "value",
        }
      },

      "services": {
        "values": {
          "connection_type": "value"
        }
      }
    }
  }
}
```

[4] Note: a JSON schema is provided at the end of this document

7.1. IP Address

An IP address object is the collection of Identity, Location, and Service attributes or objects related to an individual IP address, range, or prefix. The IP address object must specify four key parameters that are *prefix*, *start*, *end*, and *version*.

Prefix	Definition
IPaddressrange	CIDR notation of the individual IP address, or prefix. In the event of a range of addresses that span CIDR boundaries it would be represented as the first and last address separated by "-". (e.g. 198.51.100.3-198.51.100.27 or 2001:DB8::1-2001:DB8::3a)
Start	An integer value of the first IP address. This would be the all zero host bits of a prefix, the integer value of the first IP address in a range, or the integer value of an individual IP address
End	An integer value of the last IP address. This would be the host bits of all ones for a prefix, the integer value of the last IP address in a range, or the integer value of an individual IP address.
Type	Flag using 4 for an IPv4 address and 41 for an IPv6 address

Table 3: IP address object values

An example of the data elements in the JSON example provided previously is reflected below[5]:

IPv4

```
"IP": {
      "Type": "4",
            "IPaddressrange": "198.51.100.3-
            198.51.100.27",
            "Start": "3325256707",
            "End": "3325256731",
```

IPv6

```
"IP": {
      "Type": "41",
            "IPaddressrange": "2600:6c56:7008:201::",
```

[5] Note: 41 is the decimal number classification for IPv6 per IANA Assigned Internet Protocol Numbers.

```
        "Start":
"50512861188357975747210762182144819200",
        "End": null,
```

7.2. Identity

Much work has already been done in the field of identity management. This is not an attempt to redefine these efforts, but rather leverage them in a consolidated approach. Identity can include a user's name, the upstream service provider, an automobile, an enterprise network, university, or department. The identity attributes should not be limited to a customer name or ISP but could include whatever the provider needs to describe whom or what is attributed to an IP address.

This format can also be leveraged internally within an organization to assist with security or network operations. Knowing which department, or group that is responsible for a range of address space, can greatly reduce the response times required to address an issue with a part of the network.

Discussed later in this document, these identity attributes must have access controls. The details of these controls would not be included in the output data but would be needed to control what could access certain identity information. This could be something as simple as a bit identifying access to a defined level of detail. It may need to include a more detailed framework like that defined in RFC4745.

These access controls will be needed on the identity information to protect any PII data that may be available. The identity details associated to an IP address may be different if the query is made by a general access request that is open to the public, versus a query made by a requestor who has permissions for greater detail.

Prefix	Definition
country	ISO 3166-2 two-character or three-character code representing the country of the IP address
region	ISO 3166-2 compliant code representing the country region of the IP address
city	The full name of the city of the IP address
postal_code	The five-character or six-character postal code of the IP address

Table 4: Identity object values

An example of the data elements in the JSON example provided previously is reflected below[6]:

```
"Identity": {
"values": {
      "country": "US”,
      "region": "US-WA”,
      "city": "WALLA WALLA”,
      "postal_code": "99362”,
      }
},
```

7.3. Location

Much work has also been done in the area of location management. This is not an attempt to redefine these efforts, but rather leverage them in a consolidated approach. The location details could be as simple as an attribute pair for a zip code. It could be more complicated like a JSON object representing a geo-shape. The level of detail would also vary depending upon the access permissions of the requestor.

The data should adhere to RFC7946 “The GeoJSON Format” when it involves a geographical area. This standard defines the JSON representations for position, point, multipoint, linestring, multilinestring, and polygon.

The data should leverage RFC4119 “A Presence-based GEOPRIV Location Object Format” and RFC5139 “Revised Civic Location Format for Presence Information Data Format Location Object (PIDF-LO)” when the location information involves a civic location. These standards define the references for details like address, state, or postal code.

RFC6772 “Geolocation Policy: A Document Format for Expressing Privacy Preferences for Location Information” will play a part here for the level of detail provided according to access permissions of the requestor.

This document does not provide details of the JSON elements for the location category. For reference examples, please see the IETF RFC7946 document.

[6] Note: the schema defines the postal_code as the most specific Identity attribute. No personally identifiable attributes are a part of the schema and therefore the Identity section will not specifically identify a household or individual.

7.4. Service

There are use cases where it would be beneficial to know that an IP address is from a Wi-Fi network, versus a cable or wireless network, or through a VPN. An IP address range could serve a set of wireless devices that cover an area in a neighborhood. That device traffic however may be tunneled back through a VPN concentrator to a completely different location. One attribute of the IP address range may indicate the location of the neighborhood, while another attribute may identify the location of the VPN concentrator.

When forming parent child relationships to IP address objects, it would be helpful to have a service attribute as an additional distinction to provide for better decision-making. It may be helpful to distinguish between DSL, Cable, Fiber, Wireless, Satellite, Broadband over Power lines (BPL), or VPN (or Tunnel endpoint), among others.

This could be a simple name-value pair that is an attribute of the IP address object and would apply to all other attributes and objects within the IP address object.

Prefix	Definition
?	The speed is unknown
dialup	Modem used to connect at speeds up to 56 Kbps
broadband	High-speed connection, but specific type is unknown
cable	High-speed connection over cable infrastructure
xdsl	High-speed connection over DSL infrastructure
mobile	Cellular network connection (e.g., Cellular network connection (e.g., CDMA/EDGE/EV-DO/GPRS/3G/4G), Speeds vary greatly from < 1 Mbps to > 30 Mbps
T1	High-speed connection limited to roughly 1.5 Mbps
T3	High-speed connection at speeds up to 45 Mbps
OC3	High-speed connection at speeds up to 156 Mbps
OC12	High-speed connection at speeds up to 622 Mbps
satellite	Connection is via satellite; speeds vary but generally broadband

wireless	Connection is via wide area wireless; speeds vary but generally broadband

Table 5: Service object values

An example of the data elements in the JSON example provided previously is reflected below:

```
"Service": {
      "values": {
            "connection_type":
"?|dialup|broadband|cable|xdsl|mobile|t1|t3|oc3|oc12|
satellite|wireless"
      }
}
```

8. OBJECT COMBINATION

The identity, location, and service attributes and objects can be grouped together and tied to an individual IP address, range, or prefix to form an IP address object. All the attributes and objects tied to an IP address object would apply to all IP addresses within the defined range.

Each IP address object can have an optional child attribute. This attribute can point to additional IP address objects that refer to an IP address, range, or prefix that falls within the start and end address of the parent. This allows for greater detail to be provided to an IP address while associating other information to the covering range of IP addresses.

9. INTERFACE REQUIREMENTS

There are three interface requirements:

- Access control
- Data format
- Transport

9.1. Access Control

There would need to be data feeds available for generic information (research), but more detailed information may require a formal agreement between the consumer (other network operators, content delivery networks, geo-data services, law enforcement) and provider. It will have to be defined how the data will be structured to accommodate these different levels of access.

These levels of access, and the privacy implications that go with the data, are outside the scope of this document. The focus here is on what data can be made available, and it will be up to the implementers of this format to determine the appropriate amount of details to provide.

9.2. Data Format

The response to individual data requests, or any bulk delivery will be provided using JSON. The output should follow the formatting defined in RFC7159 “The JavaScript Object Notation (JSON) Data Interchange Format.”

9.3. Transport

HTTP is intended to be the transport layer given it is a JSON standard and the infrastructure, servers and client libraries for HTTP are widely available already. Ideally this JSON specification would be implemented via a REST API but that is ultimately left up to the system architects building the service.

10. MEASUREMENT AND METRICS

As with any engineering effort, there should be a means to measure whether the model, and its data, provides any benefit. The core benefit would be better decision making by the recipients who use the data. This could be a difference in latency measurements for delivered content. It could also be a difference in the number of unknown matches when trying to establish a determination.

Latency is an obvious metric to focus on but there are multiple measurements to consider. One is the improved user experience when content is delivered from the closest topological nodes. This will be measured in milliseconds. Another performance metric, is the time it takes to make a request, receive the response data, and extract the geographic information. This will be independent and based upon each provider's specific implementation but will be an important component to improve the overall user's experience.

A dependency of these measurements would be that the content provider is measuring a metric to begin with. Only then would it be possible to determine any change in the data. A short set of metrics should be identified to accompany the new data so progress could be measured.

11. CONCLUSION

Geo data is an important resource in controlling access to streaming video content, better targeting advertising, and providing a more tailored experience. But, a standard implementation of geo data for streaming video, especially in an IPv6 address space, is not shared among ISPs or other network operators in the streaming video delivery chain. It is hoped that this proposed JSON schema and approach can be the building block for such a standardized approach.

12. SPECIFICATION AND STANDARD REFERENCES

If you make ANY reference to standards or specs from other organizations in your paper, that's great! Just put them here in the following table:

Specification	Organization	More Information
RFC 4745	Internet Engineering Task Force (IETF)	Click here
RFC 5139	Internet Engineering Task Force (IETF)	Click here
RFC 5870	Internet Engineering Task Force (IETF)	Click here
RFC 6280	Internet Engineering Task Force (IETF)	Click here
RFC 6772	Internet Engineering Task Force (IETF)	Click here
RFC 7159	Internet Engineering Task Force (IETF)	Click here
RFC 7459	Internet Engineering Task Force (IETF)	Click here
RFC 7840	Internet Engineering Task Force (IETF)	Click here
RFC 7946	Internet Engineering Task Force (IETF)	Click here
Assigned Internet Protocol Numbers	Internet Assigned Numbers Authority (IANA)	Click here

Table 6: Specification and standards references

13. SCHEMA EXAMPLE

Based on the JSON code provided in this document, the following is the inferred schema:

```
{
      "$schema": "http://json-schema.org/draft-
04/schema#",
      "type": "object",
      "properties": {
            "IP_Address": {
                  "type": "number",
                  "properties": {
                        "title": {
                        "type": "string"
                        },
                  "IP": {
                        "type": "object",
                        "properties": {
                              "Type": {
                                    "type": "string"
                              },
                  "IP_Addressrange": {
                        "type": "string"
                        },
                  "Start": {
                        "type": "string"
                        },
                  "End": {
                        "type": "string"
                        },
                  "Identity": {
                  "type": "object",
                  "properties": {
                        "values": {
                              "type": "object",
                              "properties": {
                                    "country": {
                                          "type":
                                    "string"
                                    },
                                    "region": {
                                          "type":
                                    "string"
                                    },
                                    "city": {
                                          "type":
                                    "string"
                                    },
                              "postal_code": {
```

```
                                    "type": "string"
                                    }
                              },
                        "required": [
                              "country",
                              "region",
                              "city",
                              "postal_code"
                        ]
                  }
            },
            "required": [
                  "values"
                  ]
            },
            "Location": {
                  "type": "object",
                  "properties": {
                        "values": {
                              "type": "object",
                              "properties": {
                                    "postal_code": {
                                          "type":
                                    "string"
                                    }
                              },
                              "required": [
                                    "postal_code"
                              ]
                        }
                  },
                  "required": [
                        "values"
                  ]
            },
            "services": {
            "type": "object",
            "properties": {
                  "values": {
                        "type": "object",
                        "properties": {
                              "connection_type": {
                                    "type": "string"
                              }
                        },
                        "required": [
                              "connection_type"
                        ]
                  }
            },
            "required": [
                  "values"
```

```
                        ]
                    }
            },
            "required": [
                    "Type",
                    "IP_Addressrange",
                    "Start",
                    "End",
                    "Identity",
                    "Location",
                    "services"
                    ]
            }
    },
    "required": [
            "title",
            "IP"
            ]
    }
            },
            "required": [
            "IP_Address"
            ]
    }
```

14. TABLES AND FIGURES

14.1. Tables

15. ABOUT THE STREAMING VIDEO ALLIANCE

Comprised of members from across the video ecosystem, the Streaming Video Technology Alliance is a global association that works to solve critical streaming video challenges in an effort to improve end-user experience and adoption. The organization focuses on three main activities: first is to educate the industry on challenges, technologies, and trends through informative, publicly available resources such as whitepapers, articles, and e-books; second is to foster collaboration among different video ecosystem players through working groups, quarterly meetings, and conferences; third is to define solutions for streaming video challenges by producing specifications, best practices, and other technical documentation. For more information, please visit https://www.svta.org.

Streaming Video Technology Alliance
5177 Brandin Court
Fremont, CA 94538 USA
svta.org