Technology-Agnostic Reliability for HFC and FTTH

A technical paper prepared for presentation at SCTE TechExpo24

Jiten Patel
Director of Network Operations
Altice USA
Jiten.Patel3@alticeusa.com

Bryn Chung
Sr Director, Software Development
Altice USA
Bryn.Chung@alticeusa.com

David Wang
Principal Engineer
Altice USA
David.Wang@alticeusa.com

Joe Foster
Director of Network Engineering
Altice USA
Joe.Foster@alticeusa.com

Sean Ni
Sr Network Engineer
Altice USA
Sean.Ni@alticeusa.com

Table of Contents

Title **Page Number**

List of Figures

Title **Page Number**

List of Tables

Title **Page Number**

1. Introduction

With the growing presence of broadband providers utilizing both Fiber to the Home (FTTH) and Hybrid Fiber-Coaxial (HFC) technologies, a heightened demand for technology-agnostic data points for detecting reliability concerns arises. These data points are crucial for enabling providers to identify network anomalies across both networks. It is possible to monitor each network on its own, there are data points that are relevant to both technologies.

This paper demonstrates the utilization of Dynamic Host Configuration Protocol (DHCP) lease data, Customer Premise Equipment (CPE) uptime and CPE Flaps for detecting network and service reliability as well as frequent service interruptions experienced by customers. Through the analysis of these data points, broadband providers can pinpoint impairments within both their FTTH and HFC infrastructures. Additionally, we will explore how this data can be leveraged to identify issues common to specific CPE types, firmware versions, or network topography.

Leveraging technology-agnostic data presents a practical solution for identifying network impairments within FTTH and HFC infrastructures. These methods offer broadband providers a cost-effective and streamlined approach to uphold network performance and reliability.

2. Data Life Cycle

Before diving into the various datasets used in this paper, it is important to understand the data life cycle shown in Figure 1 below. This illustrates a typical approach to convert data into insights and more.

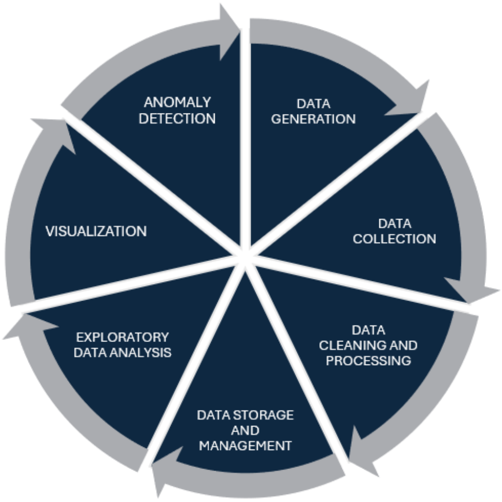

DATA GENERATION
DATA COLLECTION
DATA CLEANING AND PROCESSING
DATA STORAGE AND MANAGEMENT
EXPLORATORY DATA ANALYSIS
VISUALIZATION
ANOMALY DETECTION

Figure 1 – Data Life Cycle Stages

The process starts with **Data Generation**; without data, subsequent stages cannot commence.

Concerning the data discussed in this paper, many entities have exerted significant efforts to define protocols and standards employed in networking and telecommunications. These protocols and standards play a crucial role in ensuring interoperability and provide the data discussed in this paper. For this paper's purposes, some standards and protocols that will be discussed are Technical Report 069 (TR-069), Simple Network Management Protocol (SNMP), Secure Shell (SSH), and the Data-Over-Cable Service Interface Specifications (DOCSIS®) specification.

There is a large amount of data that is generated, but not all data can be collected or used. The **Data collection** phase is where relevant information for the project at hand is identified and captured using the appropriate methods. Once the data collection is implemented, the phase of **Data Cleaning and Processing** begins. The primary goal here is to convert data from its original, unprocessed state into something more accessible and usable. The processed data is then stored in databases or datasets for further usage. **Exploratory Data Analysis** (EDA) is a key phase where the focus is on understanding the patterns and characteristics of the data. It uses statistical and visual tools to explore the data's structure, identify patterns, trends, and potential challenges. EDA offers a comprehensive view of the data through visualizations, summary statistics, and correlation analyses. This guides practitioners towards informed decisions based on data insights. Acting as a compass, EDA directs the data analysis journey, revealing the data's intricacies and informing the development of effective insights. **Visualization** creates graphical representations of the information, making it easier to communicate the analysis. Finally, **Anomaly detection** helps identify the deviations in the data set, it can identify data objects or pattens that deviate from a dataset's normal behavior and help service providers maintain reliability.

3. DHCP

The DHCP protocol acts as a service heartbeat for CPE, detecting irregularities that may indicate poor service levels. It is one of the essential services, alongside Bootstrap Protocol (BOOTP), Trivial File Transfer Protocol (TFTP), Address Resolution Protocol (ARP), Domain Name System (DNS), Network Time Protocol NTP and Time of Day (ToD), required before customers receive high-speed internet (HSI). Beyond being a service-enabled service, DHCP has a renewal timer and serves as the first sign of life (FSOL) for various session requests. Its unique properties and extensive metadata make DHCP ideal for preliminary diagnostics.

DHCP is a client/server protocol that automatically provides an Internet Protocol (IP) host with its IP address, subnet mask, default gateway and other related configuration information. The protocol is based on RFCs 2131 and 2132 which define DHCP as an Internet Engineering Task Force (IETF) standard based on Bootstrap Protocol (BOOTP), a protocol with which DHCP shares many implementation details. DHCP allows hosts to obtain required IP configuration information from a DHCP server, removing the need for manual configurations of the TCP/IP stack. DHCP simplifies network management by automatically assigning IP addresses to devices, known as clients, on a network. It ensures that each device has a unique IP address, preventing conflicts and simplifying network administration. DHCP operates based on a client-server model, where the server manages a pool of IP addresses and leases them to clients for a specified period.

Below are some key components of DHCP:

Server: The server responsible for managing IP address allocation. It stores a pool of IP addresses and assigns them to clients upon request. The server also maintains a database of leased IP addresses and their associated clients.

Client: The device requesting an IP address from the DHCP server. This can be any network-enabled device, such as computers, smartphones, printers, and IoT devices. For this paper, it will be service provider CPE.

Relay Agent: A network device that forwards DHCP requests from clients to a DHCP server, especially useful in larger networks where clients and servers are on different subnets.

The DHCP process involves several steps- Discover, Offer, Request, and Acknowledge (DORA) for DHCPv4 and Solicit, Advertise, Request, Reply (SARR) for DHCPv6. Since the IPv6 protocol stack has no concept of a Broadcast packet the initial solicit packet is transmitted using Multicast as opposed to broadcast in IPv4 Discover packet's case.

Figure 2 below shows how a DHCP transaction flow works for HFC. The process begins with the client broadcasting a DHCP Discover message to the network. This message is sent as a broadcast because the client does not yet have an IP address and does not know the IP address of the DHCP server. The Discover message contains the client's MAC address and other identifying information. Upon receiving the Discover message, one or more DHCP servers on the network respond with a DHCP Offer message. This message includes an available IP address from the server's pool, the subnet mask, the lease duration, and other network configuration details such as the default gateway and DNS servers. The client receives the Offer message(s) and selects one. It then responds with a DHCP Request message, indicating its acceptance of the offered IP address. This message is also broadcasted to inform all DHCP servers that the client has chosen an offer, preventing other servers from reserving the same IP address. Finally, the selected DHCP server sends a DHCP Acknowledge (ACK) message to the client. This message confirms the IP address assignment and provides any additional network configuration parameters. Upon receiving the ACK message, the client configures its network interface with the assigned IP address and other settings. Clients must renew their leases periodically to continue using their assigned IP addresses, as DHCP leases are time bound. The renewal process involves the following steps according to the specifications defined in RFC2131: when half of the lease time has elapsed (T1 Timer), the client sends a DHCP Request message directly to the server that granted the lease. If the server is available and the IP address is still valid, it responds with a DHCP ACK message, extending the lease. If the client does not receive a response to its renewal request, the client will continue to send via unicast a request message at half the remaining lease time until the request is fulfilled (T2 Timer - seven-eighths of the total lease time). If the lease renewal has not been fulfilled after T2 Timer, the client broadcasts a DHCP Request message to all DHCP servers and subsequent Requests will be broadcast. This is known as the rebinding process, allowing any available server to extend the lease [1]. If the client fails to renew or rebind the lease, the IP address lease expires, and the client must initiate the DHCP process again to obtain a new IP address.

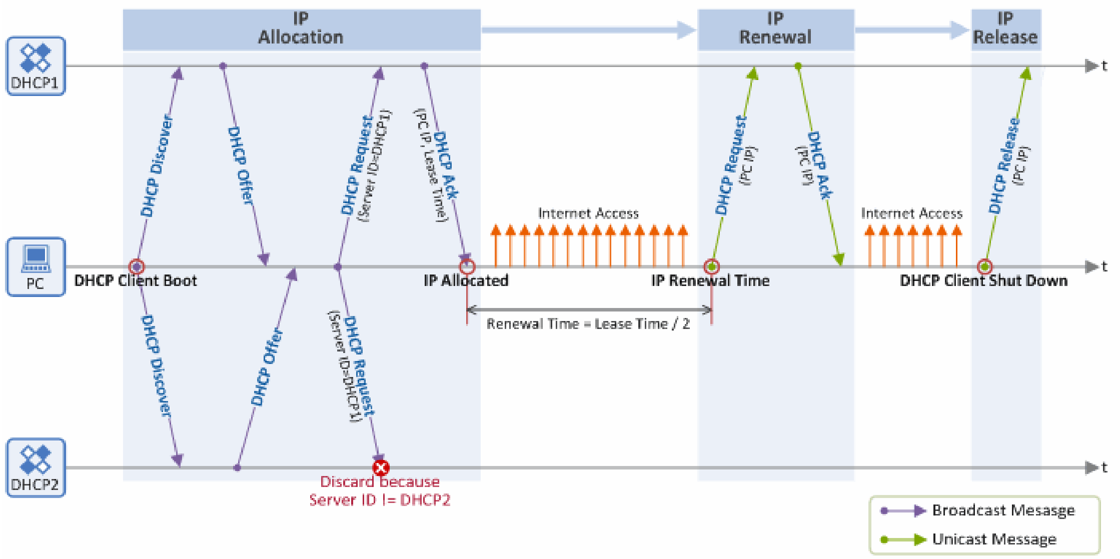


Figure 2 - DHCP protocol procedures for HFC (from [2])

Table 1 compares the HFC and FTTH technologies in terms of technical features. Although Physical Layer, Data Link Layer and Network functions are different, the use of DHCP protocol at the application layer is the same for both the technologies.

Table 1 – Functional equivalency for HFC and FTTH

	HFC	FTTH
Communication Medium	Fiber + Coaxial	Fiber Only
Communication Protocol	DOCSIS	Gigabit Passive Optical Network (GPON)/ Ethernet passive optical network (EPON)
IP Routing	Cable modem termination system (CMTS) routing	Broadband Network Gateway (BNG) routing
Command and Control	Managed via CMTS, DOCSIS protocol	Managed via Optical Line Terminal (OLT), Optical Network Unit Management Control Interface protocol (OMCI) or Operations, administration, and management (OAM) protocol
IP assignment	DHCP	DHCP

The flowchart in Figure 3 below describes the process of a DHCP request for FTTH services. The process begins with the activation of the fiber gateway, ONU and OLT are part of this activation. The Fiber Gateway (ONU/CPE) sends a broadcast DHCP Discover message, to discover available DHCP servers. The OLT adds Option 82 to the DHCP Discover packet; this option is used for the DHCP relay agent to include information on the client's point of attachment. The DHCP Discover message, now with Option 82, is broadcasted on the network. BNG acts as a proxy between the ONU and Authentication, Authorization, and Accounting (AAA) because DHCP client (RG) do not use the Remote Authentication Dial-In User Service (RADIUS) protocol. The BNG maps the values from the DHCP packet to RADIUS Attribute Value Pairs (AVPs) and sends an Access Request to the AAA RADIUS server. If the client is authenticated, the RADIUS server sends back an Access Accept message with the client's configuration information. The BNG then sends a DHCP Offer message to the client, offering IP configuration parameters. The client responds with a DHCP Request message, indicating that it accepts the parameters offered by the BNG. Finally, the BNG sends a DHCP Acknowledge message to the client. This marks the completion of the DHCP process.

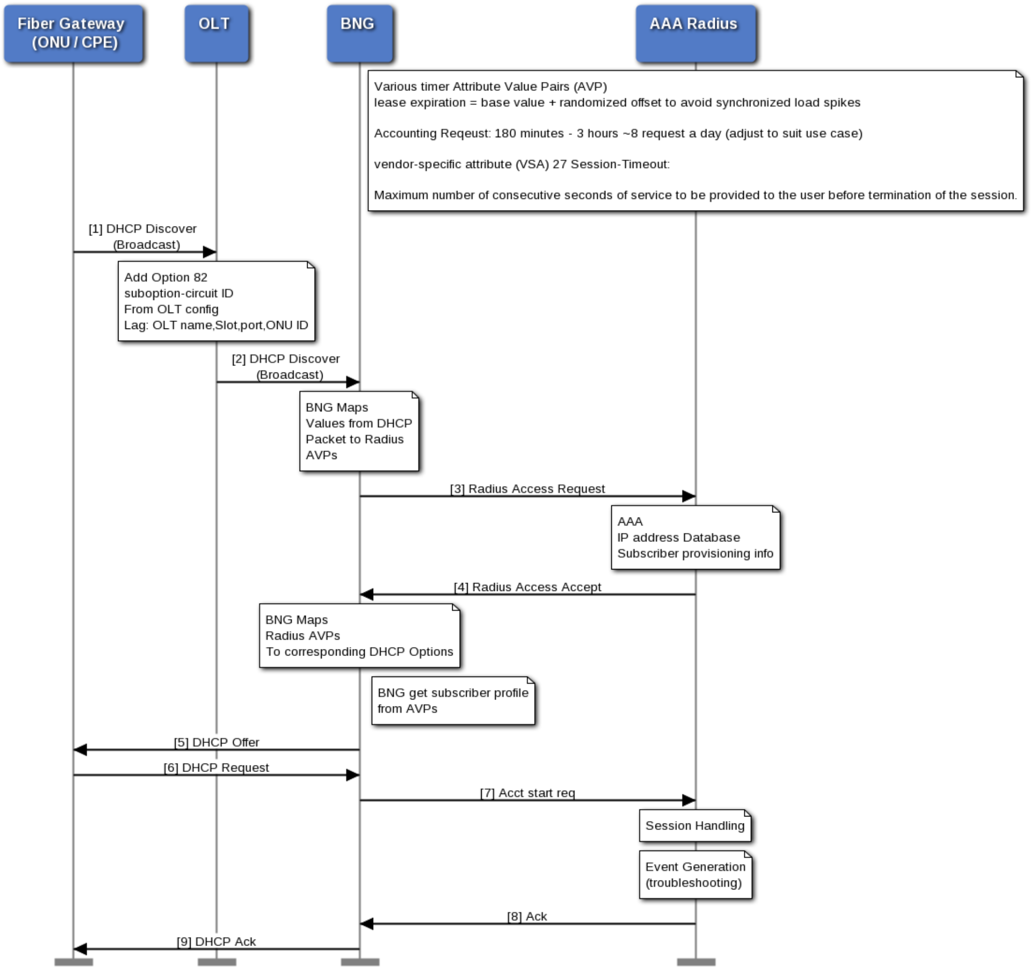

Fiber Gateway (ONU / CPE)
OLT
BNG
AAA Radius
Various timer Attribute Value Pairs (AVP)
lease expiration = base value + randomized offset to avoid synchronized load spikes
Accounting Reqeust: 180 minutes - 3 hours ~8 request a day (adjust to suit use case)
vendor-specific attribute (VSA) 27 Session-Timeout:
Maximum number of consecutive seconds of service to be provided to the user before termination of the session.
[1] DHCP Discover (Broadcast)
Add Option 82 suboption-circuit ID From OLT config Lag: OLT name,Slot,port,ONU ID
[2] DHCP Discover (Broadcast)
BNG Maps Values from DHCP Packet to Radius AVPs
[3] Radius Access Request
AAA IP address Database Subscriber provisioning info
[4] Radius Access Accept
BNG Maps Radius AVPs To corresponding DHCP Options
BNG get subscriber profile from AVPs
[5] DHCP Offer
[6] DHCP Request
[7] Acct start req
Session Handling
Event Generation (troubleshooting)
[8] Ack
[9] DHCP Ack

Figure 3 - DHCP protocol procedures for FTTH

Table 2 – DHCP Lease configurationTable 2 below shows a typical configuration for Lease duration that any service provider uses. Each service provider uses their own configuration based on the number of IPs available to them, it is key to understand how the DHCP lease interval is configured for the DHCP transaction data to be useful.

Table 2 – DHCP Lease configuration

CPE Type	Lease duration	Renew Time (T1 timer)	Lease Requests per day (based on configuration)
Cable Modem	4 days	2 days	0 or 1 request
Voice Over IP modem	2 days	1 day	1 request
Set Top Box	2 days	1 day	1 request
Router Gateway	1 days	12 hours	2 requests

The reasons why DHCP can be utilized to identify customer experience difficulties are outlined in the subsequent section. High volumes of DHCP requests from CPE can signal network connection problems. Usually, a CPE might start a DHCP requests either to renew its IP lease or to obtain a new IP upon restarting. Nevertheless, a surge in such transactions could lead to numerous potential issues. If a CPE frequently loses connection to the network, it will repeatedly initiate DHCP requests to obtain an IP address. The DHCP server itself could also be a cause for elevated lease requests, the server could become overwhelmed by too many incoming requests, causing delays or failures in assignment. Problems with the server's software or hardware could impede its ability to efficiently handle IP leases. Alternatively, the CPE might be experiencing internal issues that lead to repeated DHCP requests. Faults with the ethernet interface could make the CPE lose its IP address, prompting it to consistently request a new one. Firmware errors within the CPE might similarly trigger anomalous DHCP traffic. In some cases, relentless DHCP requests indicate malicious activities or security issues. For example, a Denial of Service (DoS) attack targets the DHCP server with an overload of requests to exhaust available IPs and prevent legitimate CPE connections. A compromised CPE could also engage in constant DHCP queries as part of an attacker's harmful actions.

With a solid grasp of DHCP operations, be it DHCPv4/DHCPv6 or HFC/FTTH technologies, the DHCP transaction logs follow a comparable structure. Each DHCP transaction is compiled and recorded, which will be examined in the following section, focusing on the methods of collection and utilization.

3.1. DHCP Data Collection

IP Address assignment is needed for gateway equipment in HFC and FTTH networks for High-Speed Internet Service (HSI). This DHCP Transaction information and useful metadata such as Option 60 and Hostname are collected by a home-grown mechanism in near real time as leases are granted to CPE. Scalability and interoperability should be considered when building such a collection system, depending on multiple factors including lease interval, number of devices, software features, geographic diversity and network diversity. Figure 4 below provides a high-level architecture for data collection. IP addresses are granted by DHCP servers on the HFC network and the BNG/RADIUS servers on the FTTH network. Syslog is a universal protocol between granting systems that can leveraged for data exports. A central destination is configured on the DHCP and BNG/RADIUS servers to send their syslog formatted log of IP address grants. This syslog data can then be captured and placed on a central database. The syslog messages are forwarded to a home-grown syslog listener. The syslog messages are then queued and processed in a first in first out method. This approach enhances the system's ability to scale and handle a high volume of request efficiently.

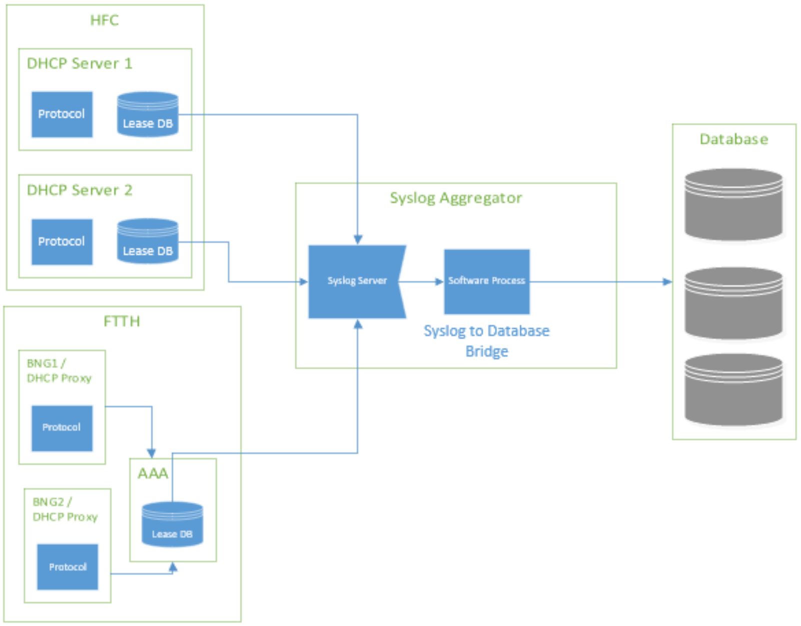


Figure 4 - DHCP Data collection flow

3.2. DHCP Data Cleaning and Exploratory Data Analysis

After gathering the raw DHCP transaction data, it can be cleansed and analyzed. It's important to establish criteria for failures to pinpoint CPEs with atypical activity patterns. Several methods exist to transform this raw data into valuable insights. Below are three viable strategies for detecting failures.

The first approach is extracting the number of lease requests per day for the CPE. Determine the number of lease requests that is typical for a CPE each day based on the T1 timer configuration (example shown in Table 2 above). If the CPE is sending the DHCP request more than two times the daily expected value, that suggests abnormal behavior. To convert the raw data into usable information, each CPE will be marked as pass or fail based on the daily DHCP request transaction log. Table 3 below shows an example of how the processed data looks. CPE4, CPE6 and CPE9 are marked as failing because they sent greater than two DHCP requests on that day.

Table 3 – Example: DHCP request per day failure criteria

CPE	# of lease Requests on a given day	Failure flag
CPE1	1	pass
CPE2	0	pass
CPE3	1	pass
CPE4	4	fail
CPE5	0	pass
CPE6	5	fail
CPE7	0	pass
CPE8	2	pass
CPE9	3	fail

The second approach is to calculate the duration between the two consecutive DHCP requests for the CPE. If the duration between the two requests is less than the renew time, that request can be marked as failing, and the CPE can be marked as failing on that day. Table 4 below shows how the processed data looks. CPE4, CPE6, CPE8 and CPE9 are marked as failing because they sent DHCP requests before their renew time.

Table 4 – Exampe: Duration between two DHCP requests

CPE	Duration between the two DHCP requests (hours)	Failure (renew time is 48 hours)
CPE1	48 hours	pass
CPE2	50 hours	pass
CPE3	52 hours	pass
CPE4	40 hours	fail
CPE5	49 hours	pass
CPE6	2 hours	fail
CPE7	51 hours	pass
CPE8	1 hour	fail
CPE9	10 hours	fail

The Third approach involves utilizing anomaly or outlier detection techniques to ascertain on which day the number of DHCP requests fall outside the expected range for a CPE. Should the volume of DHCP request be classified as an outlier, that CPE should be flagged as failing for the given day. Given that a CPE's DHCP request data typically adheres to a standard distribution, multiple models and methodologies exist to pinpoint outliers. Table 5 below shows an example of using the outlier approach. In this example, Day 7 data is compared with the previous days for the CPE. CPE4 sends four DHCP requests on Day 7, looking at the previous days the DHCP requests were consistently one per day. Hence CPE4 is marked as failing on Day 7.

Table 5 – Example: DHCP request per day anomaly detection

	# of DHCP requests on that day							
CPE	Day 1	Day 2	Day 3	Day 4	Day 5	Day 6	Day 7	Day 7 - Failure
CPE1	1	0	1	0	1	0	1	pass
CPE2	1	2	1	2	1	2	1	pass
CPE3	1	1	1	1	1	1	1	pass
CPE4	1	0	1	1	0	1	4	fail
CPE5	1	0	1	0	1	1	0	pass
CPE6	1	0	2	1	0	1	5	fail
CPE7	1	0	0	0	0	1	0	pass
CPE8	1	0	2	2	1	1	2	pass
CPE9	1	2	2	1	2	2	3	pass

Irrespective of the approach used, the goal is to convert the raw transaction logs into useful information that pinpoints the CPEs outside of the normal behavior when it comes to DHCP transactions. Once the failing and passing CPEs are distinguished, the next step is to convert this data into metrics, so the baseline can be defined and compared against.

EDA will also help flush out any data characteristics that should be excluded or drive the decision to select the failing criteria. For example, a CPE that is connected to the network, but the customer account is in non-pay status, those CPEs would send DHCP requests more often, because they won't receive the response.

Once the data is cleaned, processed and stored, the next step of the methodology is to define the metric. Metric is a standard of measurement used to quantify and evaluate performance. The metric that can be used is the "Number of CPEs with frequent DHCP Requests" and a more standardized version would be "Percent of CPEs with frequent DHCP Requests". The summarized metric data can be processed and stored in tables for quicker retrieval and visualization. To get the most benefit from this data, summarization of the data can be done at various dimensions or attributes. Some of the key attributes that can be used are network topology (Node, CMTS, Headend, Region, OLT, BNG) and CPE attributes (Model, Firmware, CPE Type).

4. CPE Uptime

Uptime refers to the amount of time a device has been continuously operating without interruptions. It is a crucial metric in the context of network devices, indicating the stability and reliability of the internet connection provided to customers. For CPE, uptime can be a measure of how long the device has been running since the last restart, reset, or power cycle. Monitoring uptime is essential for understanding and enhancing the customer experience. A high uptime value suggests that the CPE has been operating smoothly without frequent interruptions, indicating a stable and reliable internet connection. Conversely, low uptime can signify frequent disruptions, low uptime might be due to device malfunctions, network issues, firmware bugs, or external factors like a power outage. Uptime data helps in diagnosing potential problems; when a customer reports intermittent connectivity issues, examining the CPE uptime can reveal if the device has been reconnecting frequently. By regularly monitoring uptime, service providers can identify trends and preemptively address issues before they affect the customer. For instance, a pattern of decreasing uptime across multiple devices in a specific area might indicate broader network problems that need attention.

Uptime is a vital metric for evaluating the performance and reliability of CPEs in both FTTH and HFC networks. However, the implications and factors affecting uptime can vary between these technologies. FTTH delivers internet services directly to homes using fiber-optic cables. This technology is less susceptible to interference and signal degradation. Issues affecting uptime are often related to hardware (like the Optical Network Unit), external physical damage to the fiber cables, or occurrences of service provider network outages. HFC combines fiber-optic and coaxial cable technologies to deliver broadband services. Fiber is used for back-haul network, while coaxial cables are used for the last mile to the customer's premises. Uptime in HFC networks can be influenced by a range of factors including signal interference, noise, and the quality of the coaxial network. Equipment like cable modems and amplifiers also play a critical role in maintaining stable uptime. In HFC networks, maintaining high uptime is more challenging due to additional active and complex components that have potential to fail on the access network.

4.1. CPE Uptime Data Collection

This paper concentrates on three key uptime metrics: SNMP System Uptime, Primary Service Flow Activation Uptime, and TR-069 Uptime, all of which are summarized in Table 6 below. Service providers have the flexibility to choose any of these available uptime data sets for further analysis. Collecting data is a resource-demanding process since it requires harvesting information from all CPE within the network. While an hourly interval for data collection is deemed necessary, service providers may opt for a less frequent schedule to obtain more accurate data. In the below section, certain commands and sources to obtain the data are provided.

SNMP System Uptime: This is the total time that the cable modem has been powered on and operational. It resets every time the modem is restarted or loses power or T3/T4 timeouts or Interface flaps. This data can be obtained using SNMP via Object Identifier (OID) Info: 1.3.6.1.2.1.1.3.0: {iso(1) identified-organization(3) dod(6) internet(1) mgmt(2) mib-2(1) system(1) sysUpTime(3)}

SNMP Command:

```
snmpget -Of -v2c -c <community string> <cmip> 1.3.6.1.2.1.1.3.0
```

```
.iso.org.dod.internet.mgmt.mib-2.system.sysUpTime.sysUpTimeInstance = Timeticks:
(75387500) 8 days, 17:24:35.00
```

The above command for single cable modem returns the time (in hundredths of a second) since the system was last re-initialized. Service providers can use a parallel bulk SNMP polling system within the Operational Support System (OSS) or utilize a Network Management System (NMS) to regularly collect data from all cable modems.

Primary Service Flow Uptime: For the network to function properly, all CMs MUST support at least one upstream and one downstream Service Flow. These Service Flows are called the upstream and downstream Primary Service Flows. The Primary Service Flow needs to always be provisioned to allow the CM to request and to send the largest possible unconcatenated MAC frame. The CM and CMTS MUST immediately activate the Primary Service Flows at registration time. Also, if a Primary Service Flow of a CM is deactivated that CM is de-registered and MUST re-register. [3]

The primary service flow activation time is when the primary data flow is active. Primary service flow is also used to avoid ephemeral Packet Cable Multimedia (PCMM) gate set and short-lived Unsolicited Grant Service (UGS) flow.

Downstream and Upstream primary service flow uptimes can be collected using the various SNMP Object Identifiers (OID) from the CMTS.

CMTS OID examples that can be used to capture primary upstream and downstream uptimes:

'docsQos3ServiceFlowSID': '.1.3.6.1.4.1.4491.2.1.21.1.3.1.6.'

'docsQos3ServiceFlowDirection': '.1.3.6.1.4.1.4491.2.1.21.1.3.1.7.'

'docsQos3ServiceFlowPrimary': '.1.3.6.1.4.1.4491.2.1.21.1.3.1.8.'

'docsQos3ServiceFlowTimeActive': '.1.3.6.1.4.1.4491.2.1.21.1.4.1.4.'

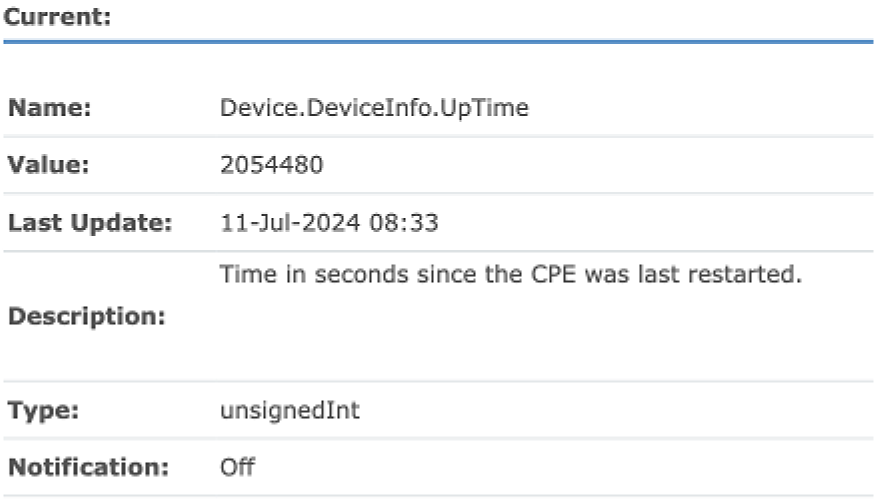
The below example provides context on the difference between SNMP System Uptime and SNMP Primary Service Flow Uptime:

When a modem is power cycled, both CM and Router Gateway (RG) will reset and have similar uptime. However, when the RF interface (F connector) is disconnected from the DOCSIS Gateway long enough the CM will reset but SNMP System Uptime might not reflect on the CM side, but the SNMP Primary service flow uptime will be reset. As seen below, the CM shows 40 days uptime, but the primary service flow shows one day, which indicates that the RF interface was disconnected, and service was interrupted.

SNMP System Uptime: "40 days, 12:26:54"
SNMP Primary Service Flow Activate Time (docsQos3ServiceFlowTimeActive): "1 day, 1:23:34"

Since this is not a standard practice and not something available off the shelf, operators can implement a custom SNMP polling system to monitor the entire cable modem population. This system can periodically collect SNMP Primary Service Flow uptime from all the modems.

TR-069 Uptime: This is similar to system uptime but specifically refers to the ACS managed gateway device. The gateway uptime is the total time that the gateway has been powered on and operational. It resets every time the gateway is restarted or loses power. This applies to both the HFC and FTTH gateways and the uptime data can be accessed through the TR-069 Data model (DeviceInfo.UpTime). This parameter provides the total time in seconds that the device has been up and running. Figure 5 below shows an example gateway uptime collected from the DeviceInfo.UpTime data model object.



Current:

Name:	Device.DeviceInfo.UpTime
Value:	2054480
Last Update:	11-Jul-2024 08:33
Description:	Time in seconds since the CPE was last restarted.
Type:	unsignedInt
Notification:	Off

Figure 5 – Example: TR-069 Uptime

For service providers without ACS management, Secure Socket Shell (SSH) command can be run on a gateway to capture this data.

SSH command:

Command 1: uptime

Output: 12:34:12 up 23 days, 18:41, load average: 2.43, 2.40, 2.41

uptime gives a one-line display of the following information. The current time, how long the system has been running, and the system load averages for the past 1, 5, and 15 minutes.

Command 2: cat /proc/uptime

Output: 2054503.21 2454449.70

This file contains two numbers (values in seconds): the uptime of the system and the amount of time spent in the idle process.

To collect this data from all the managed gateways, service providers can generate periodic reports from the ACS system, or they can create a customer SSH polling system to mass collect this data by using either of the two options mentioned above.

Table 6 – Uptime data sources

	Standalone Modem	Cable Modem Gateway	FTTH Gateway
SNMP System Uptime	SNMP OID 1.3.6.1.2.1.1.3.0	SNMP OID 1.3.6.1.2.1.1.3.0	NA
DS Primary Service Flow Uptime	CMTS OID	CMTS OID	NA
US Primary Service Flow Uptime	CMTS OID	CMTS OID	NA
TR-069 Uptime (only if under ACS management)	NA	TR-069/ACS (Device.DeviceInfo.UpTime)	TR-069/ACS (Device.DeviceInfo.UpTime)
SSH (if no TR-069 implemented)	Model specific	Uptime command	Uptime command

Uptime behaviors are firmware/model/SoC (System on Chip) specific. Validation and document CPE behaviors as part of certification process is highly recommend. This way service providers can leverage the data collected to its fullest potential.

4.2. CPE Uptime Data Cleaning and Exploratory Data Analysis

As explained in the DHCP section, once the raw data is gathered, to put it in action it should be cleansed and processed.

To derive the daily uptime percentage, the first step is to determine how many seconds in a day the modem was up and running. As the raw uptime value for each CPE is collected hourly, finding the difference between the current hour counter and previous hour counter will provide the number of seconds the modem was up for that hour. Whenever the device reboots, the counter will reset and thus provides valuable information on how long the CPE was operational. Table 7 below illustrates how the raw data can be converted to the information that is needed. At 3:00, the counter was reset, and the counter value was 2,500; indicating that the modem was up for 2,500 seconds in that hour. At 4:00, the CPE did not respond to the polling request, indicating that the CPE would have been down for that duration, at 5:00 the counter was read as 200, which indicated that the uptime for that hour would have been 200 seconds. Upon performing EDA, various scenarios will be uncovered which will shape the refinement of data cleansing and processing. One such scenario is illustrated between 8:00 and 10:00 hour, the poller was not able to retrieve the counter values for 8:00 and 9:00 polls, but 10:00 poll showed the value of 10,800, which would indicate that the modem was up for the entire duration, but data collection failed. In such case, excess values for the hour would be distributed to previous hours to account for the missed polls as shown in the “Cleaned Uptime seconds” column in the Table 7 below.

Table 7 – Example: Uptime Calculation

Hour	Counter in seconds	Uptime seconds	Cleaned Uptime seconds
0:00	54,000	3,600	3,600
1:00	57,600	3,600	3,600
2:00	61,200	3,600	3,600
3:00	2,500	2,500	2,500
4:00	Not available	0	0
5:00	200	200	200
6:00	3,800	3,600	3,600
7:00	7,400	3,600	3,600
8:00	Not available	0	3,600
9:00	Not available	0	3,600
10:00	18,200	10,800	3,600
11:00	21,800	3,600	3,600
12:00	25,400	3,600	3,600
13:00	29,000	3,600	3,600
14:00	32,600	3,600	3,600
15:00	36,200	3,600	3,600
16:00	39,800	3,600	3,600
17:00	43,400	3,600	3,600
18:00	47,000	3,600	3,600
19:00	50,600	3,600	3,600
20:00	54,200	3,600	3,600
21:00	57,800	3,600	3,600
22:00	61,400	3,600	3,600
23:00	65,000	3,600	3,600
Total Uptime			78,300
Total Time			86,400
Uptime percentage			78,300/86,400=90.6%

Within the three metrics collected, DHCP and CPE Flaps are comparable because they both reflect how often service interruptions occur, whereas the Uptime metric represents the length of time these disruptions last. The fundamental concept is to isolate CPEs that show an uptime less than X percent over the course of a day. In this paper, we will define a failing CPE as one with a daily uptime percentage below 99%. That gives wiggle room of ~15 minutes for any power fluctuations, firmware updates and occasional reboots that might be triggered by customer behavior. Identifying what the failing percentage should be is part of the EDA and the goals set by the company. Table 8 below shows an example of how the processed data can be utilized, since the goal set per CPE is 99% uptime, CPE4 and CPE6 are marked as failing on that day.

Table 8 – Example: Uptime failure criteria

CPE	Uptime Percentage	Failure (less than 99% uptime)
CPE1	99.9%	pass
CPE2	99.1%	pass
CPE3	100%	pass
CPE4	96.3%	fail
CPE5	100%	pass
CPE6	98.2%	fail
CPE7	99.4%	pass
CPE8	99.8%	pass
CPE9	99.3%	pass

The metric that can be derived from uptime dataset is the ‘Number of CMs with high downtime’, and a more standardized version would be ‘Percent of Devices with high downtime’. Like the DHCP data approach, CM uptime data can be condensed into different dimensions and archived for future retrieval and display.

5. CPE Flap

CPE flap in this paper’s context refers to frequent disconnection and reconnection of a CPE device to the network. These can be caused by various factors such as power fluctuations, hardware issues, signal interference, software bugs or network issues. When a CPE flaps, it temporarily disrupts the network connection and re-establishes the connection with the service provider's network. Monitoring the frequency and patterns of flaps is critical for assessing and improving customer experience. A stable connection with minimal flaps indicates a reliable service, which is crucial for a positive customer experience. Service providers can use CPE flap data to identify broader trends and address issues proactively. For example, if multiple CPEs in a specific area are flapping frequently, it might point to a localized network problem that needs investigation. CPE flaps are relevant across both HFC and FTTH technologies and serve as a powerful indicator for reliability.

For HFC, CM Registration / De-Registration type traps are specific types of notifications used in network management to monitor the status of cable modems. These traps are a record of registration status of the cable modems. CM Registration traps are sent by the CMTS when a cable modem successfully registers with the network. The registration process is necessary to establish and optimize communication between the CM and the CMTS, ensuring proper service level entitlements providing efficient and reliable internet access for the user. CM De-Registration traps are sent by the CMTS when a cable modem de-registers from the network, either due to a voluntary action (e.g., user disconnects the modem) or due to a network issue (e.g., loss of signal, reboot).

For FTTH, as per GPON specifications, there are alarms and performance monitoring mechanisms to detect link failure. Certain alarms that are detected at OLT would indicate flap, such as Loss of signal for ONUi (LOSi), Loss of Signal (LOS) and DGi (Received dying-gasp of ONUi). There are certain other alarms defined in the OAM functions that indicate a disruption or degraded service, those should be explored by the service providers when they implement this mechanism. [4]

In this paper, the methods utilized to detect service disruptions included registration/deregistration and alarms; however, several alternative techniques exist. Among these alternatives for CM, there is the CM event log, CMTS command line interface, and NMS/SNMP modem polling. For gateway CPE, SSH is another viable method to extract reboot counter for the gateway. Service providers can select any of these options to identify service interruptions and leverage the collected data for analysis.

5.1. CPE Flap Data Collection

For HFC, NMS tools or SNMP trap receivers can be used to receive and unpack CM Registration / De-Registration type traps from the CMTS. The receiving system should decode the information in the SNMP trap payload according to the CMTS vendor specific Management Information Base (MIB) configuration. This data can then be captured and placed on a central database. Scalability should be considered as these traps will be received for every registration state for every CM.

For FTTH, NMS Tools or proprietary systems provided by network equipment manufacturers can be used to receive and interpret alarms from ONU/ONTs. Alarms from these tools can be logged into a database for consumption for future analytics.

5.2. CPE Flap Data Cleaning and Exploratory Data Analysis

CPE Flap and DHCP data are very similar once they are transformed from their raw form into information. As was done with DHCP data, failing criteria can be defined as more than two flaps per day. As shown in the Table 9, CPE4 and CPE6 are marked as failing due to frequent flaps (>2) in a day.

Table 9 – Example: Flaps per day failure criteria

CPE	# of flaps on a given day	Failure
CPE1	1	pass
CPE2	0	pass
CPE3	1	pass
CPE4	3	fail
CPE5	0	pass
CPE6	9	fail
CPE7	0	pass
CPE8	2	pass
CPE9	1	pass

The metric that can be derived from flap dataset is the "Number of CPEs with frequent flaps" and a more standardized version would be "Percent of CPEs with frequent flaps". Like what was done for the DHCP data, the CPE Flap data can be summarized at various dimensions and stored for retrieval and visualization.

6. Vizualization and Anomaly Detection

Data visualization and anomaly detection are crucial stages in the data life cycle, enhancing data interpretation. Data visualization transforms complex data sets into graphical representations such as charts, graphs, and maps. This visual representation allows stakeholders to quickly comprehend patterns, trends, and insights, making data more accessible and digestible. Effective visualization highlights key data points and relationships, facilitating better understanding and communication of findings. It transforms raw data into actionable insights, which are crucial for performance tracking, and decision-making. By turning large volumes of data into visual summaries, visualization helps to uncover hidden patterns and correlations that might be missed in text-based data. Anomaly detection identifies data points that deviate significantly from the norm within a dataset. These anomalies can indicate errors, failures, or new, unexpected insights. It helps in the early identification of outliers that could indicate errors, allowing for timely corrective actions. Detecting anomalies is essential for spotting issues and ensuring actions can be put in place. Together, data visualization and anomaly detection provide a robust framework for exploring, understanding, and maintaining data, ensuring that insights derived are both accurate and meaningful. Below are some examples of visualization that help understanding the data and their uses.

Figure 6 below shows a line graph that represents 'Percent of Devices with high downtime' metric over time. The y-axis is labeled as 'Percent of Devices with high downtime' which ranges from 0% to 3.5%. The black line that fluctuates around the average represents the actual data points. The lines labeled '+/- 3 Standard Deviations' represent the range within which approximately 99.7% of the data points should fall if the data follows a normal distribution. The areas marked as 'Data Anomaly' are significant deviations from the average. These are instances where the percentage of devices with high downtime deviated significantly from its normal range, which could be cause for investigation or concern. In summary, this chart appears to be monitoring performance consistency over time, using statistical process control methods. The 'Data Anomaly' labels highlight periods where the failure rate was unusually high.

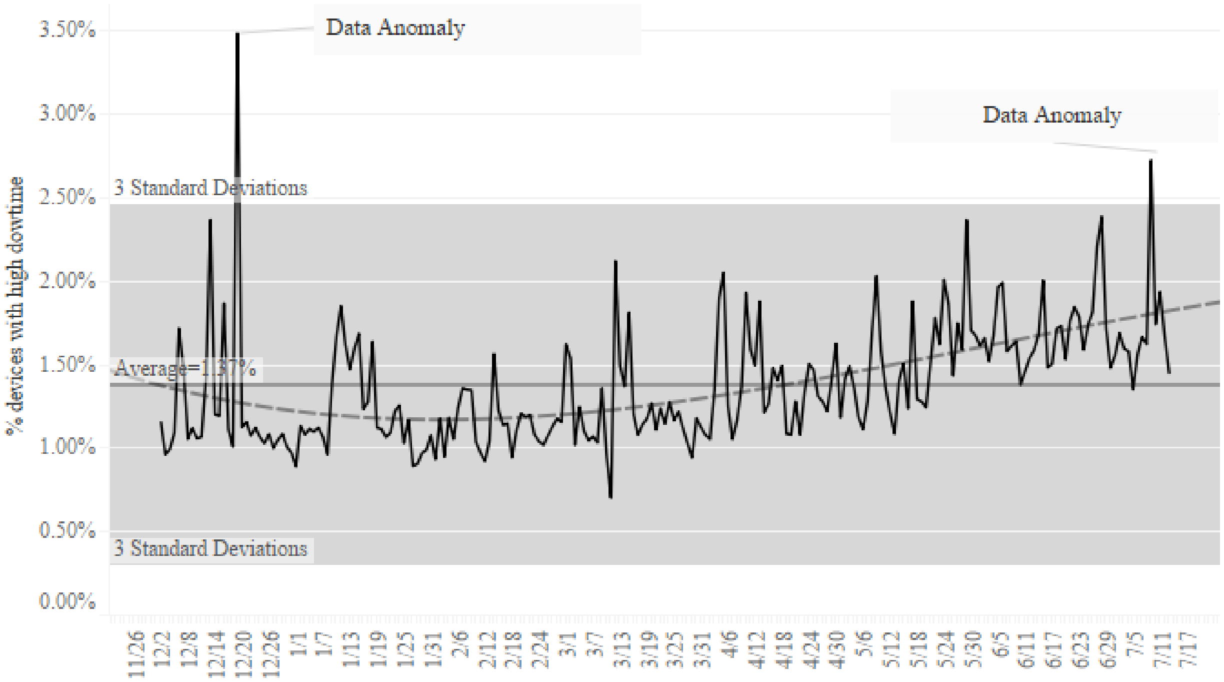


Figure 6 – Uptime KPI Trend

Figure 7 below show two scatter plot graphs. The scatter plot on the left shows each facility, the scatter plot on the right shows each CPE type**.** Both scatter plots, plot 'Number of devices' on the vertical axis, against 'Percent of devices with high downtime" on the horizontal axis for different facilities and CPE types. The points are scattered primarily in the left region of the plot, which indicates that most of the facilities/CPE models show similar behavior. However, there are some facilities/CPE models that have a higher failure rate and would need further investigation. These anomalies, when detected and actioned upon promptly, will result in improved network reliability. These graphs are useful for visualizing the distribution of device failure rates across different facilities and CPE types. They highlight the typical failure rates as well as anomalies and outliers, which could be critical for business analysis or decision-making processes.

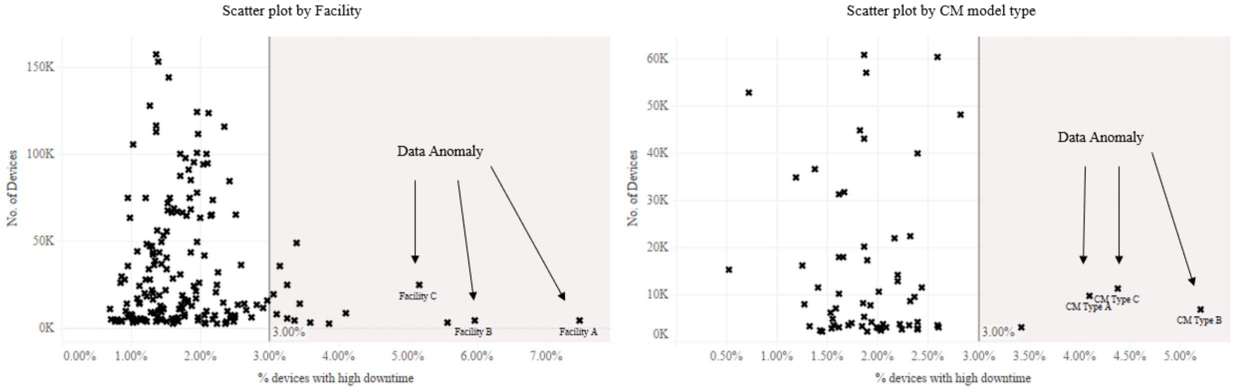


Figure 7 – Uptime Scatter Plot

Figure 8 below illustrates the monitoring of continuous improvement using the ‘Frequent Flap’ metric. The chart documents the lifecycle of a certain CPE model. Broadband providers incorporate new CPEs featuring the latest technology and innovation as part of their strategy to remain competitive and meet consumer expectations. Initially, when a CPE is introduced, the percentage of devices frequently flapping ranges from 2% to 3%. Since this metric is under constant scrutiny and any anomalies are flagged in comparison to existing CPE model types, engineering teams are prompted to investigate the root causes. This leads to the discovery of bugs and defects, which are subsequently rectified in new firmware releases, yielding an observable improvement over time. As more firmware updates are rolled out, additional issues are addressed, contributing to ongoing enhancement that can be graphically tracked. In essence, the chart presents a visual account of how frequent flapping among devices fluctuates over time, especially after different firmware upgrades, offering a means to assess the influence of these updates on CM functionality and the user experience.

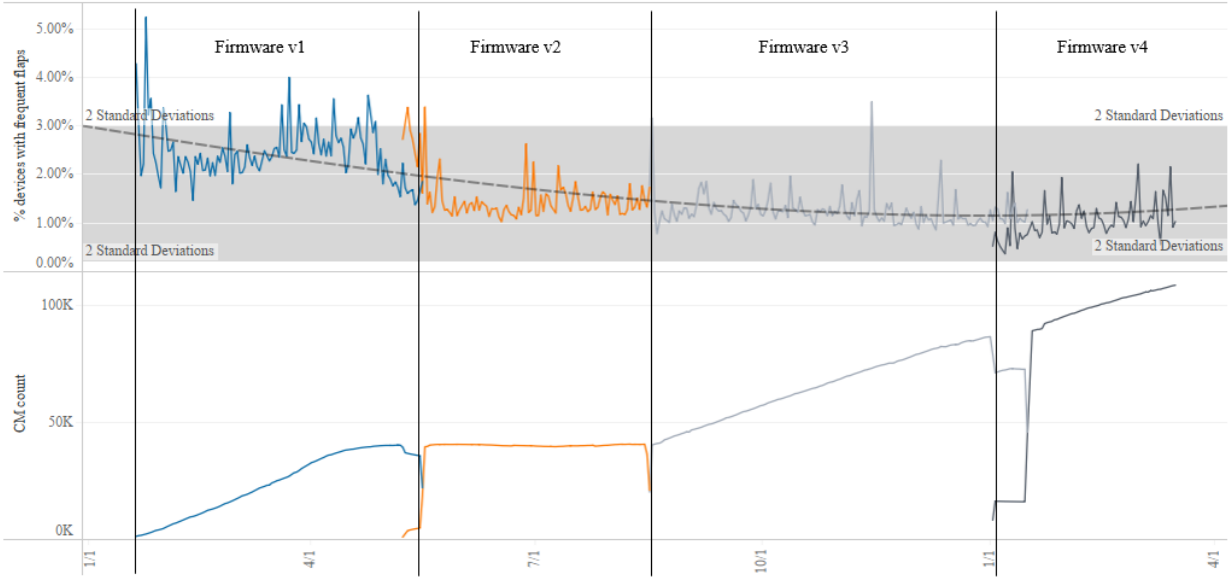


Figure 8 – CPE Flap Failure Trend for CM model by Firmware version

7. What lies ahead

Incorporating Artificial Intelligence (AI) and Machine Learning (ML) will significantly enhance data analysis by automating complex processes, uncovering deeper insights, and enabling predictive capabilities. These advancements are particularly beneficial for analyzing datasets related to DHCP, uptime, and CPE flaps, helping ensure network reliability and availability. AI and ML can automate the analysis of large datasets involving DHCP logs, uptime records, and CPE Flap events. Traditional methods that have been used in the past for diagnosing issues within these datasets can be time-consuming and prone to human error. As seen throughout the paper, there are certain thresholds used to indicate failure or issues, but these were based on human interpretation of the observed data and understanding of the process. Automation with AI/ML will not only accelerate these processes but also enhance accuracy and consistency. Machine learning models excel at identifying complex patterns and correlations within large datasets. By analyzing DHCP logs, uptime records, and CPE flap events ML can detect the trends more accurately than the approach used in this paper. Furthermore, AI/ML can correlate failure data points with outages and power disruption events, thereby filtering out unnecessary noise from the data. With AI/ML, these datasets can be correlated to each other, to provide predictive analytics capacities. AI/ML's predictive analytics can forecast potential network problems by analyzing historical data from DHCP, uptime, and CPE flap logs. Predictive models can anticipate DHCP failures, identify signs of potential downtime, and detect indicators of imminent reboots. This foresight allows for proactive maintenance, minimizing network disruptions. Prescriptive analytics takes this a step further by offering actionable recommendations based on these predictions, guiding network operations teams on preventative measures to optimize network performance. AI/ML systems continuously learn from new data, enhancing their accuracy and effectiveness over time. This self-improving capability ensures that AI/ML-driven analysis adapts to changing network conditions and emerging issues. As an example, as new patterns of DHCP failures or CPE Flaps events are detected, the models refine their predictions and recommendations, leading to better issue resolution and network stability.

8. Conclusion

The paper highlights that both HFC and FTTH technologies share similar data points. These data points sometimes come from identical sources or systems, like for DHCP and uptime, and can occasionally be collected from various sources yet translated into the same type of information as with CPE flap. Additionally, these metrics have a conceptual similarity and are complementary, allowing them to support each other's data. As service providers begin to adopt a blend of HFC and FTTH technologies, it becomes essential to employ technology-agnostic data points to monitor network performance and identify reliability issues. Utilizing these data points enables service providers to locate and mitigate network irregularities effectively, thus maintaining consistent service for customers. The entire data life cycle, from collection to anomaly detection, plays a pivotal role in turning raw data into practical insights that improve customer experience. Real-world applications of these approaches offer guidelines for service providers to enhance their network operations and maintain high service reliability levels. Furthermore, with the growth of AI and ML technologies, these data sets could be further leveraged to automate intricate operations, refine pattern recognition, and deliver predictive and prescriptive analytics. AI and ML can advance the analytical processes of DHCP, uptime, and CPE flap data sets, contributing to more robust and consistently operational networks.

Abbreviations

AAA	authentication, authorization, and accounting
ACK	dhcp acknowledge (ack)
ACS	auto configuration server
AI	artificial intelligence
ARP	address resolution protocol
AVPs	attribute value pairs
BNG	broadband network gateway
BOOTP	bootstrap protocol
CM	cable modem
CMTS	cable modem termination system
CPE	customer premise equipment
DHCP	dynamic host configuration protocol
DHCPv4	dynamic host configuration protocol version 4
DHCPv6	dynamic host configuration protocol version 6
DNS	domain name system
DOCSIS	data-over-cable service interface specifications
DORA	discover, offer, request, acknowledge
EDA	exploratory data analysis
EPON	ethernet passive optical network
FSOL	first sign of life
FTTH	fiber to the home
GPON	gigabit passive optical network
HFC	hybrid fiber-coaxial
HSI	high-speed internet
IETF	internet engineering task force
IP	internet protocol
IPv4	internet protocol version 4
IPv6	internet protocol version 6
MAC	media access control
MIB	management information base
ML	machine learning
NMS	network management system
NTP	network time protocol
OAM	operations, administration, and management
OID	object identifier
OLT	optical line terminal
OMCI	optical network unit management control interface protocol
OSS	operational support system
PCMM	packet cable multimedia
PMIP	proxy mobile ip
RADIUS	remote authentication dial-in user service
RFC	request for comments
RG	router gateway
SARR	solicit, advertise, request, relay
SNMP	simple network management protocol
SSH	secure socket shell

TCP	transmission control protocol
TFTP	trivial file transfer protocol
ToD	time of day
TR-069	technical report 069
UGS	unsolicited grant service

Bibliography & References

[1] R. Droms, "Dynamic Host Configuration Protocol," *IETF RFC 2131,* 1997.

[2] Netmanias, "Understanding the Basic Operations of DHCP," 23 October 2013. [Online]. Available: https://www.netmanias.com/en/?m=view&id=techdocs&no=5998&xtag=dhcp-network-protocol&xref=understanding-the-basic-operations-of-dhcp.

[3] CableLabs, "Data-Over-Cable Service Interface Specifications DOCSIS 3.0," *MAC and Upper Layer Protocols Interface,* Vols. CM-SP-MULPIv3.0-C01-171207, 2017.

[4] International Telecommunication Union, "Gigabit-capable passive optical networks (G-PON): Transmission convergence layer specification," *SERIES G: TRANSMISSION SYSTEMS AND MEDIA, DIGITAL SYSTEMS AND NETWORKS,* no. G.984.3, 2014.

[5] T. Stobierski, "8 STEPS IN THE DATA LIFE CYCLE," Harvard Business School, 2 Feb 2021. [Online]. Available: https://online.hbs.edu/blog/post/data-life-cycle. [Accessed 1 7 2024].

[6] J. Case, M. Fedor, M. Schoffstall and J. Davin, "A Simple Network Management Protocol (SNMP)," *IETF RFC 1157,* 1990.

[7] K. McCloghrie and M. Rose, " Management Information Base for Network Management of TCP/IP-based internets: MIB-II," *IETF RFC 1158,* 1991.

[8] e. a. K. McCloghrie, " Structure of Management Information Version 2 (SMIv2)," *IETF RFC 2578,* 1999.

[9] Broadband Forum, "TR-069 CPE WAN Management Protocol," *TECHNICAL REPORT,* no. Amendment 6 Corrigendum 1, 2020.