Access Network Node Augmentation Optimization through an Operational Digital Twin Model of Fiber Node Plant Topology

A technical paper prepared for presentation at SCTE TechExpo24

Mehul Patel
Distinguished Architect
Comcast
mehul_patel@comcast.com

Sinan Onder
Vice President, Software Architecture
Comcast
sinan_onder@comcast.com

Emma Romero
Student Co-op
Comcast
emma_romero@comcast.com

Maher Harb
Distinguished Engineer, Data Science
Comcast
maher_harb@comact.com

Table of Contents

Title **Page Number**

List of Figures

Title **Page Number**

List of Tables

Title **Page Number**

1. Introduction

Comcast operates and manages its network to be optimized for delivery of bandwidth for its subscribers. Within this context, management of broadband capacity is one of the biggest challenges. The hybrid fiber coaxial (HFC) plant, which combines optical fiber and coaxial cable, is commonly deployed across the network footprint. The ability to scale speeds and bandwidth is dependent on the number of subscribers, subscriber traffic patterns, and how spectrum is configured on a fiber termination point or node. At the point of saturation, or when a certain usage threshold is crossed after optimizing spectrum or device capability mix, the node goes into an alarm state. Once an alarm is detected, our problem statement is to determine the best way of constructing and designing new plant augmentations to resolve the problem. Specifically, we need to provide optimum construction designs either by creating new fiber termination points deeper in the network (action known as node-split) or by balancing the radio frequency (RF) ports on an existing node housing (action known as node-rebalance). With the HFC network consistently evolving in technology and hardware and with increasing supply and demand, it becomes a necessity to optimize network construction designs and to optimally augment the network for current and future capacity needs, while efficiently keeping cost at a minimum.

The current software and systems used to solve RF capacity supply-demand problems are mainly designed to help make human-driven processes more efficient, rather than to help drive automated data-driven solutions. There are many human-driven processes to manage the current optimization of the network capacity. Furthermore, each group/team managing an aspect of the network capacity problem collects vast amounts of data and tends to develop their own interpretation of the data, leading to highly proprietary and fragmented solutions. Due to the many human-driven and complex processes and solutions, the ability to create data-driven solutions becomes challenging. Accordingly, we tend to settle for augmentation actions that are reactive and that are less impactful and their ability to address the underlying issue. With the growing demand and consumption of bandwidth, cable network operators must quickly evolve in providing automated solutions for optimization of their RF capacity so that actions shift from being an “art” that is highly dependent on the knowledge and expertise of the designer, to being a “science” that is rooted in data, evidence, and analytics.

Another aspect of the problem is the inability of the current processes and systems that manage RF capacity to keep up with the evolution of the network. Cable network operators undergoing a transformation defined by the 10G roadmap, many new variables are introduced. These include highly dynamic orthogonal frequency division multiplexing (OFDM) & orthogonal frequency division multi access (OFDMA) channel capacity defined by the Profile Management Application (PMA), multiple OFDM/A channels on the same spectrum, and full duplex (FDX) channels supporting symmetrical Gbps speeds. The current processes that are human driven are not optimized to make use of the new technologies being introduced. For example, the current methods used to provide RF capacity solutions are mainly tailored to split a node or rebalance the load across a node housing’s RF ports. These methods may not be as efficient or cost effective due to the constraints of the current algorithms and rules; they also tend to overlook low-cost software-based mitigations actions (e.g. shifting spectrum from D3.0 to OFDM). Construction costs are also increasing, so when we predict that construction is needed as a last resort to solve capacity problems, we must be confident that the action is optimized such that the node will not trigger an alarm once again within a couple of years following the augmentation action. In a nutshell, we need to be thorough and surgical in how we resolve RF capacity issues.

In this paper, we propose digitization of the network and use of the operational digital twin concept as a framework to provide optimal solutions for RF capacity supply-demand problems. We introduce data-driven algorithms to recommend physical or logical changes in the access network and traffic reallocation

© 2024 Comcast. SCTE TechExpo24 logo used with permission from SCTE

models utilizing the digital twin plane. This paper also shows how we can use simulations in the digital twin plane to provide efficient and optimized data-driven outcomes and options. Hence, the true benefit of these simulations is to enable a more optimized and predictive approach to managing RF capacity within the HFC network.

2. Operational Digital Twin Concept

2.1. Network Digital Transformation or Digitization

Digital transformation is a journey from manual and human-driven processes to a digital paradigm, enabling more end-to-end automation. Digitization is distinct from virtualization: while virtualization involves creating a virtual instance or representation of a physical resource to replicate its function, digitization extends further. Beyond the creation of digital objects representing physical resources, digitization also necessitates the development of digital capabilities that interact with and optimize these physical resources, leveraging the digital objects. Figure 1below illustrates the concept of virtualization in the context of Comcast's virtualized cable modem termination system (vCMTS).

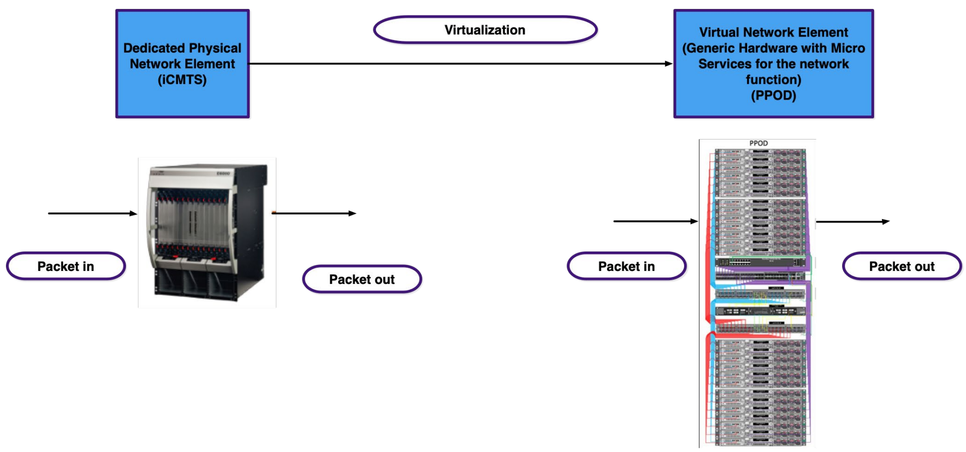

Figure 1 - Schematic illustrating the virtualization concept when applied to Comcast's CMTS platform.

Our Comcast principle of digitization relies on three core components: digital entity, digital capability, and digital mobility as described below:

- **Digital Entity:** The domain objects important to the business (such as remote physical layer (PHY) device (RPD), node segment, amplifier, service, truck roll) within the digital plane, where there exists a single source of truth.
- **Digital Capability:** The digital actions performed on or towards the digital entity, such as node segment rebalancing or node segment splitting. For example, pressing a button would trigger a truck roll.
- **Digital Mobility:** The ability to transport the source digital entities to different environments for various use cases. It involves data replication rather than the data itself.

The digitization of the network and the creation of digital entities begins with data derived from the network. This data includes information that is well-known and collected from the network, such as the HFC node, spectrum, associated node segments, and the RF plant that connects the customer premise equipment (CPE). Figure 2 below illustrates this concept. By properly digitizing the network and creating the appropriate digital domain objects, we can reap significant benefits.

© 2024 Comcast. SCTE TechExpo24 logo used with permission from SCTE

Figure 2 - Illustration of benefits of digitization of the *DAA* platform in the middle and revolving around the operational digital twin are the different capabilities to manage a network. The orange boxes represent the consistent optimizations for different operations groups.

2.2. Choice to Use Operational Digital Twin

The operational digital twin involves digitizing physical objects and their interactions within the digital plane (the Digital Entity). Our objective in creating the operationalized digital twin was not to develop yet another specialized digital twin solely for addressing RF capacity issues. Instead, the goal was to establish a virtual twin, populated with data that can be leveraged to solve a wide array of problems and use cases. By separating the digital representation from a specific solution or capability and establishing it as a standalone, first-class concept, it can be utilized by multiple capabilities.

Traditionally, digital twins are used for simulating physical world concepts to improve the efficiency of designing and building solutions by testing approaches in a simulated environment first. In Comcast's journey of network digital transformation, the primary use of the “digital twin” differs from this traditional approach. Digital twins are created to represent physical entities within the operational digital plane to enhance the consistency, accuracy, and quality of digital operations. For this reason, they are referred to as “operational digital twins”. The operational context drives the modeling of these twins, such as operational digital twins for “Intended Node,” “Design Node,” and “Operational Node.” Figure 3 below represents the traditional use of a digital twin, while Figure 4 illustrates our intended use of the operational digital twin.

© 2024 Comcast. SCTE TechExpo24 logo used with permission from SCTE

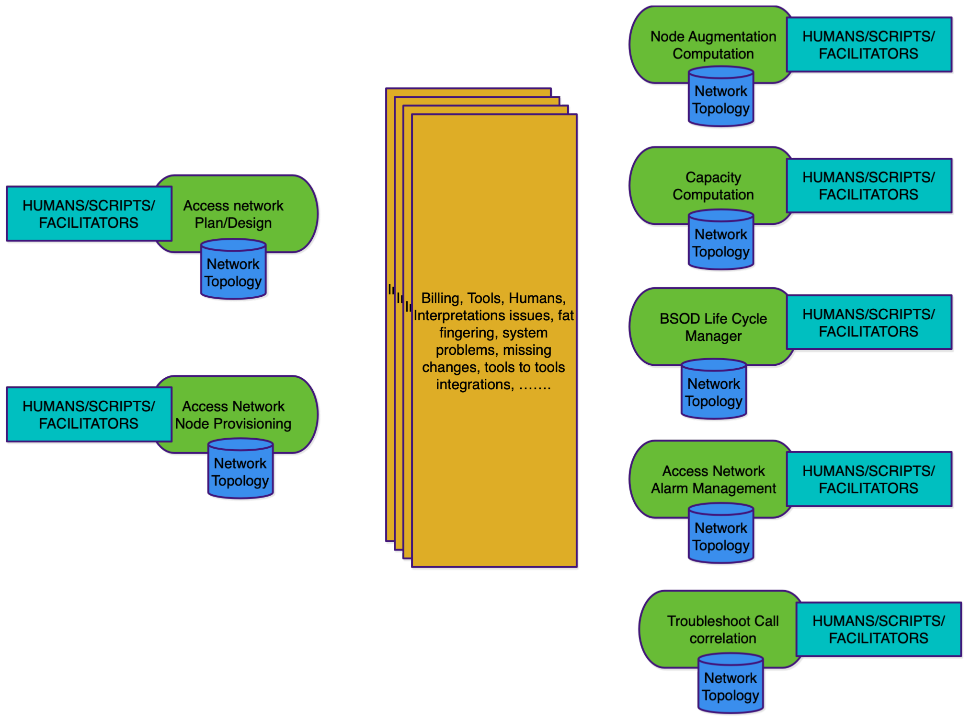

Figure 3 - This figure captures how traditional use cases create singletons of digital twin for their data.

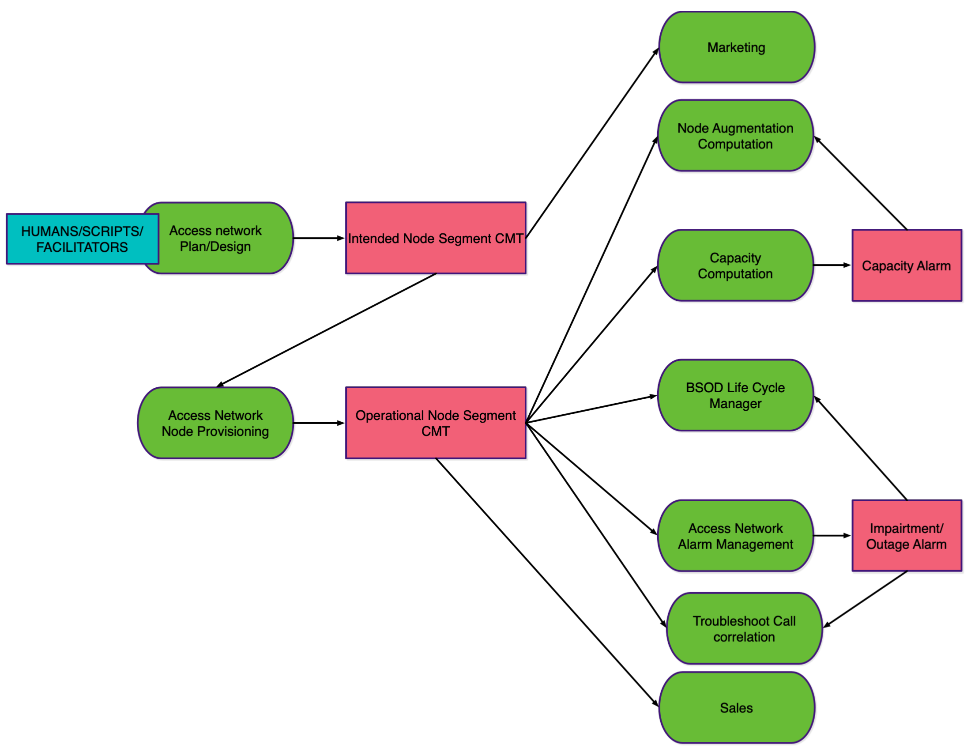

Figure 4 - Schematic illustration of how a operational digital twin can be created for many use cases and problems.

As the network evolves, the operational digital twin will adapt, incorporating data that is vital for managing and maintaining the network. This approach enables data-driven outcomes and options to address a variety of issues. Operationalizing the digital twin using operational data models like provisioning and network telemetry helps ensure the digital twin remains synchronized with the physical elements. This approach eliminates the need to maintain and manage independent data models, creating a digital twin intended for public use across multiple scenarios. Our teams benefit by not having to collect data and create additional interpreted data models, as they can utilize the digital twin generated from operationally discovered data.

Additionally, this approach allows our analytics team—focused on solving bandwidth capacity problems—to leverage real data processed by the network. This enables the team to concentrate on creating and optimizing algorithms for the problem without the need to build an environment for the algorithms. The advantages of having an operationalized digital twin environment for numerous problems and use cases increase the efficiency of our creators and optimizers, allowing them to dedicate more time to refining algorithms.

2.3. Making Use of the Operational Digital Twin

In this paper, both aspects of digital twin concepts will be utilized: the operational aspect for real-time monitoring and management, and the traditional aspect for simulations and analysis. The operational digital twin is constructed using operational node and RPD data obtained from the network back office. The plant data is derived from construction design data, which is used to map the physical HFC node topology, including nodes, amplifiers, taps, and drops. Usage and capacity data are collected from the operational poller and persist in a centralized data bucket. These data points collectively form the operational digital twin of the network model.

The initial use for this operational digital twin will be to create a data-driven solution to provide an automated and efficient HFC node topology augmentation plan. Node topology augmentation plans are created to help upgrade or expand the current capacity of an existing HFC node topology to help improve the network traffic and performance. The plans consist of either physical construction or software-based intents or changes to help manage and improve the distribution of the network traffic.

This operational digital twin serves as the foundation for the proposed approach. Subsequently, this digital twin of the HFC node topology, based on the same data models, will be employed to simulate the efficiency of a given augmentation plan. This involves recreating alternative plans using the digital twin approach, which aligns with the traditional use of digital twins for simulations and analysis. Further details of this approach will be provided in Sections 3-5.

2.4. Avoiding the Pitfalls of Previous Solutions

The proposed solution enables our algorithms to harness operational data generated by the network. Our objective is to offer data-driven outcomes and options to enhance the efficiency of our construction split processes, which can be costly when making changes to the HFC network. Previous solutions relied heavily on proprietary, rule-based outcomes and options. These solutions were region-based and focused on managing environments and collecting data to generate outcomes, relying heavily on human experience.

In contrast, our algorithms are transparent and publicly available, evolving in tandem with network changes. This approach eliminates the need to start from scratch when new technology is introduced to enhance bandwidth and capacity. The idiom “this is more art than science” is frequently used when there is not enough data for an optimum solution and decisions are made based on “expertise”. The goal of this approach is to make this process “more science than art.” Figure 5 represents traditional data and solutions to help people facilitate complicated outcomes, while Figure 6 shows how our approach is to make more data-driven and less human-involved decision-making to automated solutions and outcomes.

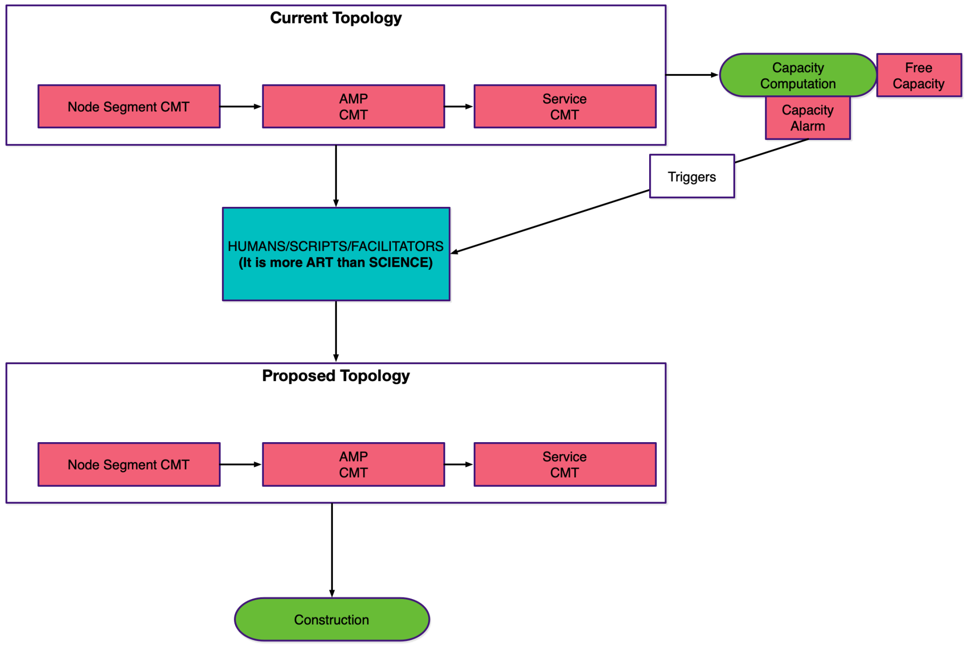

Figure 5 - Representation of traditional capacity solutions generating outcomes to drive human processes.

© 2024 Comcast. SCTE TechExpo24 logo used with permission from SCTE

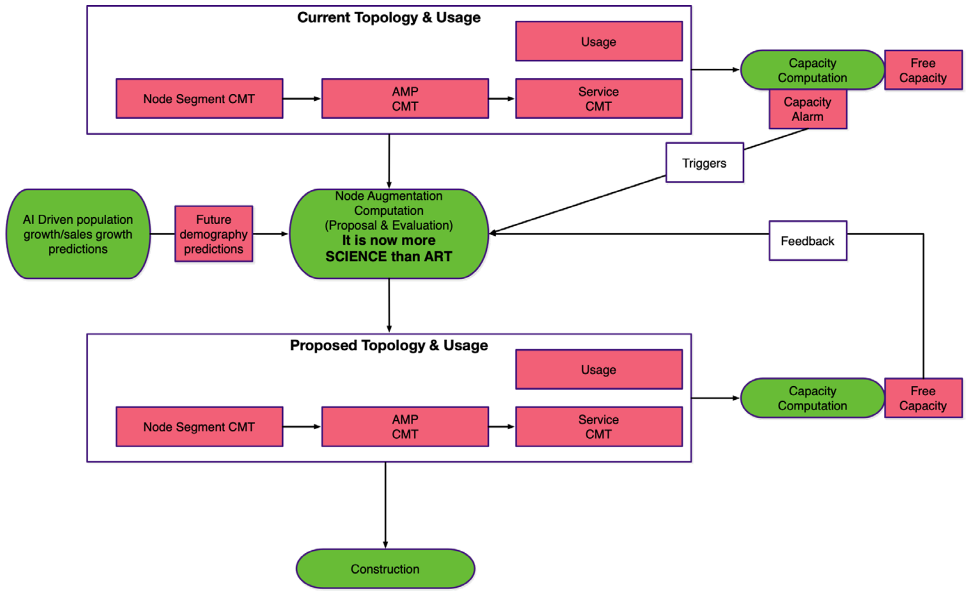

Figure 6 - Illustration of moving to a more automated and data driven solution to solve capacity issues.

3. Access Network Node Data Model

The HFC access network is represented by a graph data model that was described in an earlier SCTE technical paper [1]. For modeling RF capacity supply-demand problems, the unit of interest is the fiber node. We note, however, that the term node is loosely used by the cable tech industry to refer to an RF aggregation point. But this term is somewhat misused and misleading. Specific to the Comcast RPD platform, devices (cable modems) share a single downstream port and up to two upstream ports (also referred to as node segments). Thus, devices connected to the same RPD (node segment) share the same RF spectrum, and consequently, the downstream (upstream) traffic patterns of those devices along with the spectrum configuration (data channels, modulation order, forward error correction (FEC) configuration, etc.) dictate the availability of bandwidth at a given point in time. In this respect, a node is both an RPD (in the downstream traffic direction) and a node segment (in the upstream traffic direction). The discussions & examples below focus on manipulating the ‘node segment’ topology to address upstream capacity constraints. The same concepts apply to manipulating the RPD topology to address downstream capacity constraints. One last terminology convention to clarify is addressing the term ‘node’ as a graph node in the computer science sense of the meaning. To avoid confusion with the ‘node’ as an access network element, a graph ‘node’ will be referred to henceforth as a “vertex.” Figure 7 is a graph visualization of a node segment representing the Comcast distributed access architecture (DAA) platform. In the node segment graph, vertices may represent physical elements such as amplifiers, splitters, taps, drops, etc., or logical elements such as cable modem service groups, billing addresses, customer account numbers, etc. Edges on the other hand represent coaxial cables connecting each pair of vertices that are physical elements. The graph has a tree structure: there exist no path cycles, the root vertex is the node segment, each vertex (except for the root) has exactly one parent vertex, and the leaf vertices are the devices. Note that the data model is rich beyond what is shown in this example node. Elements including

© 2024 Comcast. SCTE TechExpo24 logo used with permission from SCTE

locations, accounts, drops, etc. were suppressed on purpose to avoid crowding the visual. Moreover, each element (whether vertex or edge) has attributes associated with it. Attributes may be common across all classes of elements or unique to a specific class. For example, physical elements have spatial (latitude-longitude) coordinates, edges representing cables have lengths, and devices have data-over-cable service interface specifications (DOCSIS®) version, device model, device firmware version, etc. On the software side, neo4j is the database platform used for managing and querying the graph objects [2]. Using a native graph database for storing, updating, and querying allows efficient execution of graph operations both in terms of the computational cost and the conciseness and ease of constructing the search queries. For more on the graph databases and associated query languages refer to neo4j graph algorithms white paper [3].

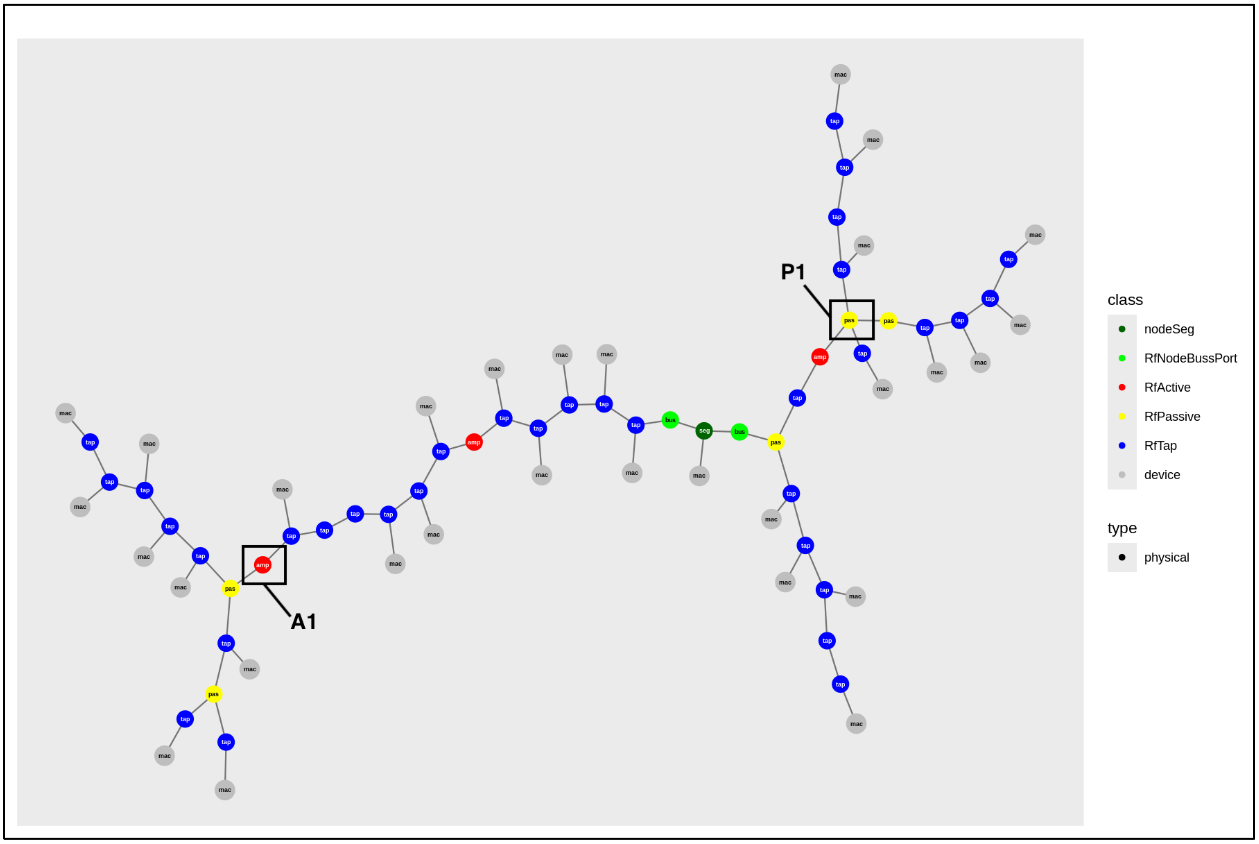

Figure 7 - Example graph visualization for a node segment showing various elements of the HFC plant. The graph has a tree structure with the node segment (dark green circle) as the root and the devices (light gray circles) as the leaves. Note, this visualization is strictly for highlighting the relationships between entities and is not informed by the actual spatial coordinates or map of the plant. Amplifier A1 and passive P1 are annotated on the figure to highlight the transformation shown in the next figure.

3.1. Node Transformations

What has been described so far can be thought of as a digital representation of the HFC plant. The graph data model supports many applications outside the topic of this paper. Specifically, it is an asset for day-to-day network operations as it holds the lineage for every connected device on the network. On the space-time dimension, we tend to think of the graph data model as the spatial dimension (topology) at the present (t_0). Transformations in space and time can unlock interesting uses for this data. This section

© 2024 Comcast. SCTE TechExpo24 logo used with permission from SCTE

focuses on the spatial dimension. To explore node augmentation scenarios, first we need to define and implement a host of transformations that shift the graph representation of the node in the digital world while keeping the time dimension fixed. For example, under a given transformation T map graph G to graph G'. Those transformations should reflect actions that happen in the real world. Below, we describe the key set of transformations that may be invoked in node augmentations. All of these were implemented under a single Python class.

3.1.1. Vertex Move

A simple "cut & insert" operation in which the tree branch under the "move from" vertex is cut from the tree and inserted under the "move to" vertex. While this is a generic tree operation, for a fiber node, the move must obey certain constraints that are imposed by the network architecture. The following are examples of such constraints: a device must be inserted under a drop, the only element that can be inserted under a node segment is the busleg (RF port on the node housing), certain classes of vertices can be daisy-chained (e.g., passives, amplifiers) while others cannot (e.g., nodes, buslegs, devices, drops). In addition to these "hard" constraints, which reflect the physical & architectural limitations, there are also "soft" constraints that are imposed by policy. For example, in principal amplifiers and passives may be cascaded/daisy-chained in any form or fashion. However, we may wish to limit the length of the amplifier cascade to less than a certain number to ensure that RF performance is within the design specifications or to support certain types of planned future plant upgrades. Both hard and soft constraints are implemented in the vertex move operation to confirm that a requested move is allowed. Figure 8 shows the same example node in Figure 7 but with the amplifier marked with "A1" moved under the passive marked with "P1".

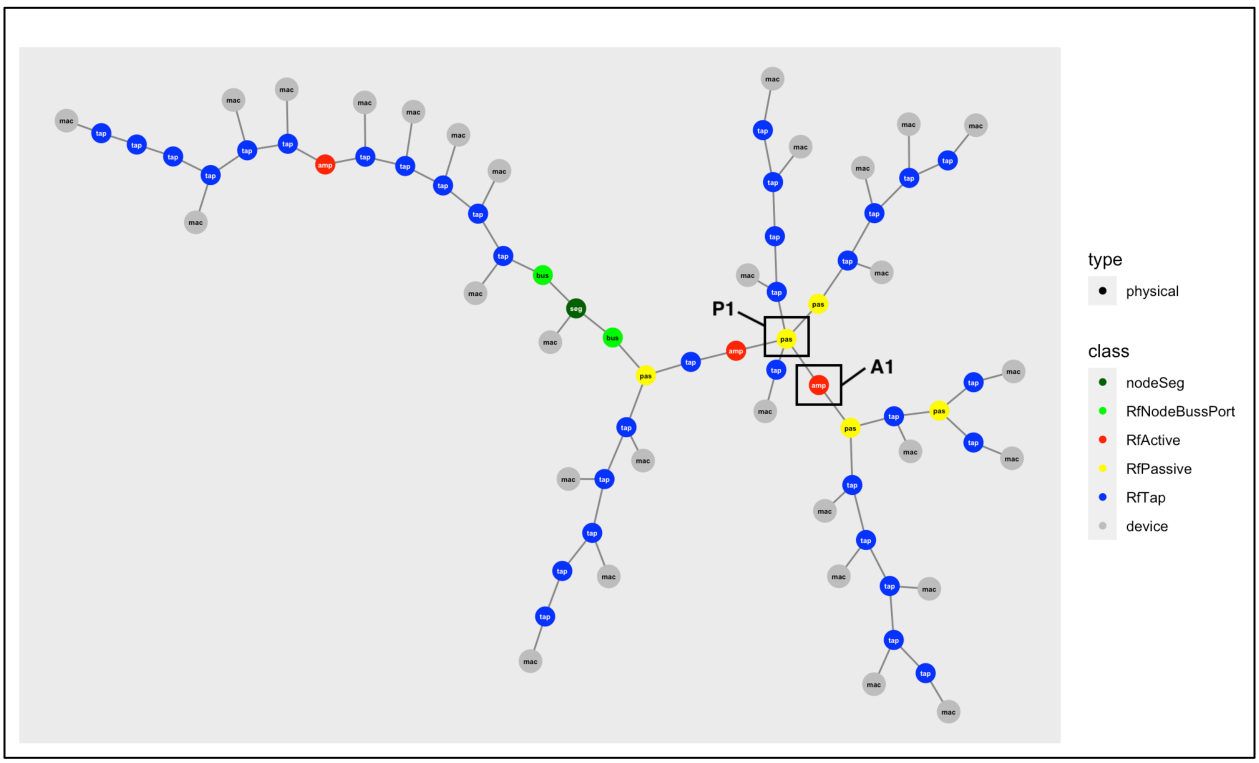

Figure 8 - The same node segment shown in the Figure 7 but with a vertex move operation applied.

© 2024 Comcast. SCTE TechExpo24 logo used with permission from SCTE

3.1.2. Vertex Creation

Certain augmentation actions require the creation of new vertices. The most notable examples are the creation of a new node segment and/or a new busleg. Once again, the creation of vertices must obey physical and architectural design. For example, a node housing has four RF ports. Therefore, creating a new busleg is allowed only if the number of existing buslegs in use is less than four. Similarly, a node housing can accommodate a maximum of 2 RPDs, each in turn accommodating a maximum of two node segments. These constraints are enforced by the vertex creation operation.

Contained within a node segment, moving and creating vertices are not useful augmentation actions to relieve RF capacity issues. Instead, they are combined to distribute devices across two node segments (or across two RPDs), thereby reducing traffic on the congested RF spectrum. Figure 9 shows an example “node split” action, in which a new node segment (N1) was created, and a branch (A1) from the node segment in Figure 7 was moved under the newly created node segment. Thus, both types of basic operations were used to fulfill this augmentation action.

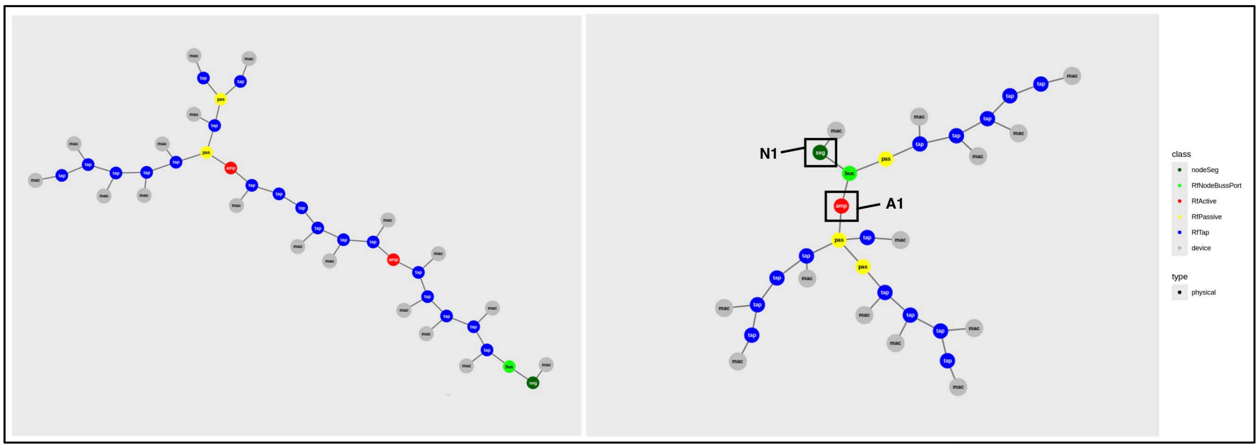

Figure 9 - The same node segment shown in Figure 7, with Amplifier A1 split from the original node segment and inserted under a newly created node segment (N2). The resulting topology are the two children node segments shown in the left (original node minus A1 branch) and right (new node) panels.

3.1.3. Amplifier-Node Swap

A more complex but often employed augmentation action is moving an existing fiber node to the location of an existing amplifier and replacing the amplifier by the fiber node. This is usually a viable option since amplifiers are the only class of elements that are powered aside from nodes. Therefore, the cost of the augmentation is relatively low compared to moving the node to a location with no access to power. Note that this action cannot be represented as a combination of create/move/delete operations since it may involve reversing some of the tree paths (i.e., the child-parent relationships). Figure 10 shows an example of such an action, in which node (N1) was moved to the location of amplifier (A1). A1 in turn was removed from the graph.

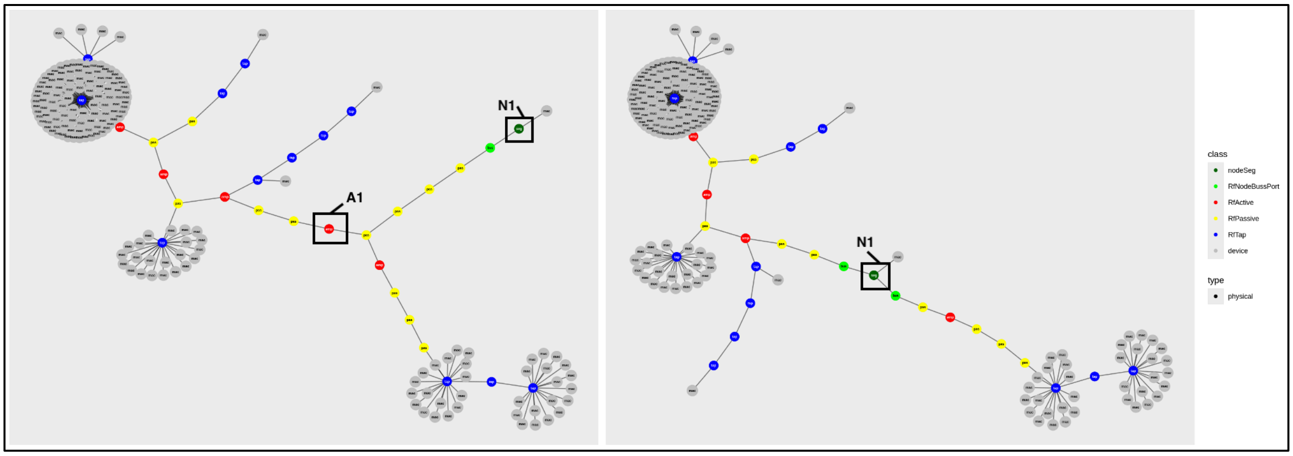

Figure 10 - Example of an amplifier-node swap augmentations action. The current node is shown in the left panel. The right panel shows the transformed node in which node location was moved to the location of the amplifier marked as A1.

3.2. Augmentation Actions of Interest

Now that the building block operations are defined, we can transform a given node that exists today to a node that exists in the digital twin plane. Such transformation will also include enrichment data that captures the main cost elements associated with the transformation. These may include the amount of additional fiber or coaxial cable that needs to be deployed, the maximum amplifier cascade length in the transformed node, whether any amplifiers are no longer needed because of the reduced node-to-device distances, etc. These cost metrics are very useful for future exercises in which the different augmentation scenarios are compared according to the costs and benefits associated with each. Table 1 is a summary of the various augmentation actions used for RF capacity planning along with the basic operations involved with each.

Table 1 - Summary of the different augmentation actions that may be invoked to relieve high spectrum utilization in the upstream or downstream traffic directions.

Current Node Topology	Transformed Node Topology	Node Housing Location Moved	Action Category	Basic Operations Involved	Fixes Traffic Issues in
1 RPD, 1 Node Segment	1 RPD, 2 Node Segments	No	Node Split	Vertex Creation, Vertex Move	Upstream
1 RPD, 2 Node Segment	1 RPD, 2 Node Segment	No	Node Rebalance	Vertex Move	Upstream
1 RPD, 2 Node Segment	1 RPD, 2 Node Segment	Yes	Node Rebalance	Amplifier-Node Swap	Upstream
1 RPD	2 RPDs	No	RPD Split	Vertex Creation, Vertex Move	Downstream & Upstream
2 RPDs	2 RPDs	No	RPD Rebalance	Vertex Move	Downstream & Upstream
2 RPDs	2 RPDs	Yes	RPD Rebalance	Amplifier-Node Swap	Downstream & Upstream

© 2024 Comcast. SCTE TechExpo24 logo used with permission from SCTE

4. Reconstructing Traffic on Digital Twin Nodes

What was achieved in the previous section is the creation of node topologies that exist only in the digital twin plane. These digital twin nodes serve as potential candidates for augmentation actions. The digital twin nodes, however, do not generate any telemetry. Understanding the impact of the candidate topologies on traffic requires that we somehow generate telemetry for these nodes within a simulated environment. The idea for creating a simulated telemetry environment informed by existing traffic characteristics is based on the following principles. The current DAA telemetry platform is extremely rich, reporting on traffic at a 5-minute frequency from the vCMTS. Furthermore, both RPD and cable modem telemetry are available at a DOCSIS channel granularity (e.g., single carrier quadrature amplitude modulation (SC-QAM) or OFDM/A). As an outcome of transforming a node in the digital twin plane, the migration of cable modems between node segments (or RPDs) is known. Therefore, traffic can be reconstructed accordingly at a node level from the individual cable modems under the node. Through this reconstruction, we would have created a simulated traffic for the digital twin nodes. Note that the definition of "traffic" may be different in the downstream vs. the upstream direction. For the results discussed below, the focus is on the upstream and the count of DOCSIS codewords within each 5-minute window is adopted as the metric for measuring traffic volume.

One additional modeling exercise is needed. The main capacity performance metric is channel utilization. This is a vCMTS-reported raw metric reflecting mini-slot occupancy in the upstream. Utilization is the key metric that triggers capacity alerts and, depending on the situation leads to augmentation plans. However, utilization cannot be allocated to individual devices. To circumvent this problem, we developed a linear regression model to calculate channel utilization based on the count of upstream codewords. Even though codewords vary in length, Figure 9 shows that the relationship between total codewords over each 5 minute interval and the corresponding channel utilization is nearly linear. A linear regression model was fit to this data (solid blue line in Figure 11). The R-squared value of the model was 0.61, indicating that codewords explain much of the variability in utilization . Such a model is specific to a given node as it considers the unique behavior and characteristics of subscribers on the node. With the channel utilization reconstructed from the aggregation of per-subscriber traffic on the digital twin nodes, we now have simulated telemetry that is normalized to mimic telemetry on the 'real' production nodes. The power of this approach is having digital twin nodes and production nodes 'live' side by side in the data model. Both are fed to the same capacity platform that crunches traffic data to generate high-level capacity reporting metrics. Thus, digital twin nodes require no special data pipeline to assess the impact of the proposed augmentation actions. Instead, as far as data pipelines & algorithms are concerned, they are treated as if they do exist as real nodes (with the understanding that their id is used to trace back to the augmentation scenario each is supposed to represent).

© 2024 Comcast. SCTE TechExpo24 logo used with permission from SCTE

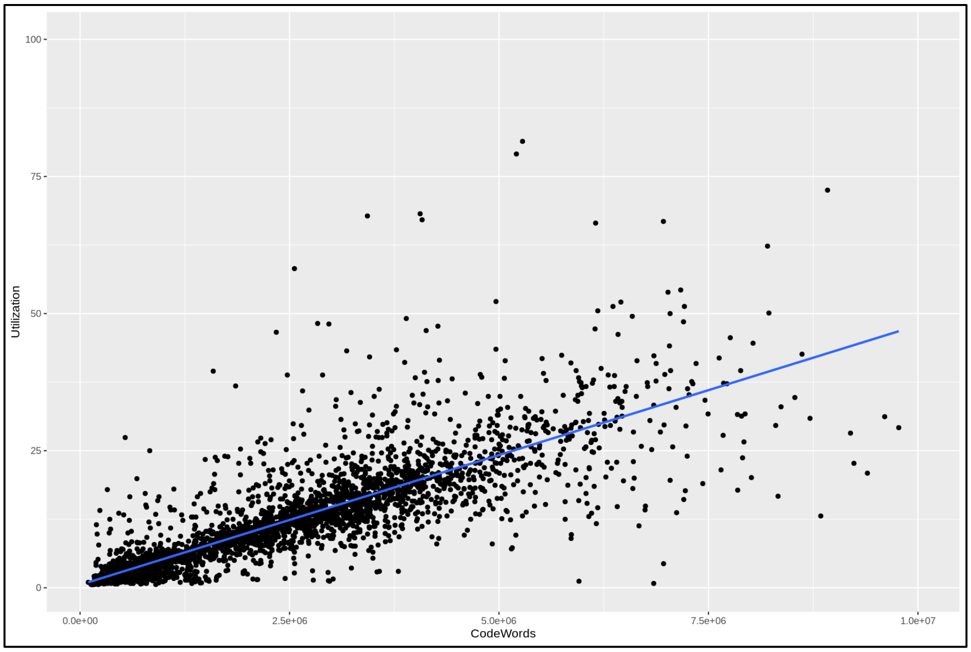

Figure 11 - Scatter plot of upstream channel utilization (y-axis) vs. total code words (x-axis). The solid blue line represents a linear fit of the relationship.

To conclude this section, we comment on another layer of modeling that can be introduced as part of reconstructing traffic on the digital twin nodes. So far, the time dimension has been treated as fixed in the present. However, augment actions can take a few months to a year to implement as they involve securing permits, construction, supply chain issues, availability of labor, etc. Adding to that, many nodes exist in locations that experience seasonal traffic (e.g. vacation spots, university campuses). It would be prudent to overlay a time series forecasting element to the traffic to project the traffic in time for a better assessment of the situation. Once again, no special effort is required to forecast traffic on the digital twin nodes. The same time series forecasting models that are employed today to project traffic on production nodes are invoked on the digital twin nodes. By forecasting traffic, we now have shifted the digital twin topology both in space (augmentation actions that change topology) and time (predicting future traffic patterns on the augmented digital twin nodes).

5. Example Simulation

In this section, we consider an example to highlight the use of the digital twin model to inform the optimal augmentation action. The starting node segment is the one shown in the left panel of Figure 12. Note that to reduce crowding in the visual, every group of cable modems connected under the same tap was grouped into a single gray circle labeled with the number of cable modems in the group. As seen in the figure, many more cable modems are connected to one of the node segments (dark green circles) than the other. As such, the upstream utilization is now higher on the larger node segment and, if not addressed, may trigger an alarm in the future. Such an imbalance is not necessarily a flaw in the design of the plant. It could develop organically by subscriber growth along the larger branch of the tree.

Once the node has been flagged as overloaded, it is fed into the operational digital twin model to determine which alternative node topology provides the “best” fix to the problem. The model runs through all possible transformations in the form of vertex moves, creations, and amplifier-node swaps and calculates the potential impact based on the reallocation of cable modems and traffic for each scenario. In this example, we restricted the augmentation action space to rebalancing the node segments by moving buslegs (light green circles) across the two node segments. We also limited the optimization objective to balancing the number of cable modems under each node segment. The optimal topology is shown in the right panel of Figure 12. We also show in Table 2 the different considered scenarios along with the objective metric for the optimization.

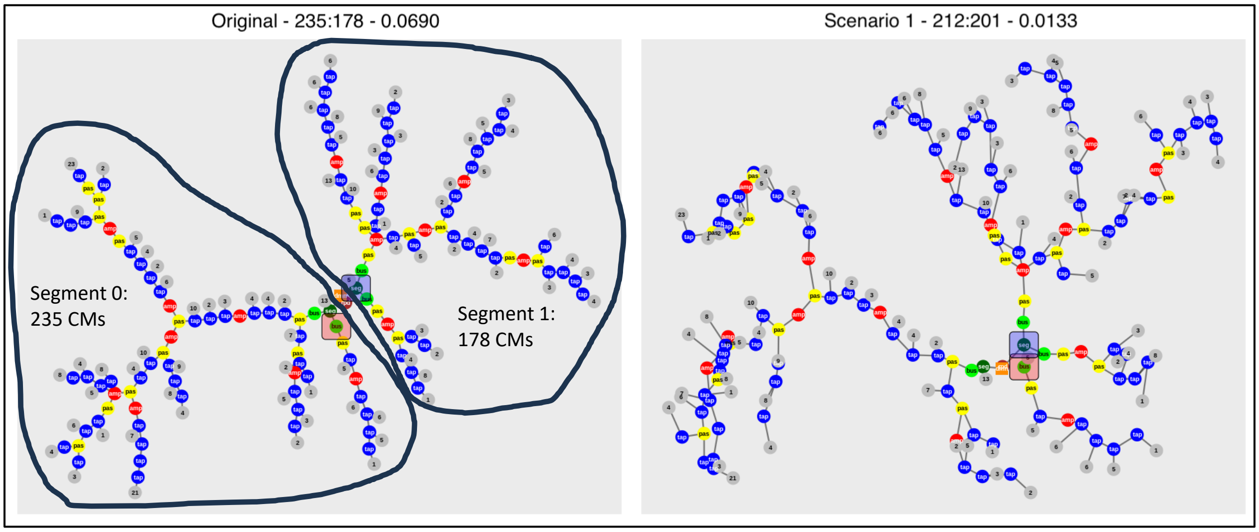

Figure 12 - The original node segment (left) and the most optimal topology (right) as suggested by the operational digital twin model. The transformation for this example was vertex move as highlighted in the figure.

Table 2 - Different augmentation scenarios for the node in Figure 12 as determined by the digital twin model. Balance ratio is defined as 1 minus the ratio of number of subscribers of the smaller node segment to the number of subscribers of the larger node (hence, a balance ratio of 0 represents a perfect split)

Scenario Id	Transformation Action	Devices on Node Segment 0	Devices on Node Segment 1	Balance ratio
0	Original configuration	235	178	0.0690
1	Node rebalance	212	201	0.01331
2	Node rebalance	253	160	0.1126
3	Node rebalance	36	377	0.4128
4	Node rebalance	390	23	0.4443

6. Conclusion

This paper proposes a data-driven approach that unifies the direct/run-time operational capabilities of the running network (e.g., network management, alarm detection, etc.) and more offline/indirect operational

© 2024 Comcast. SCTE TechExpo24 logo used with permission from SCTE

capabilities (e.g., capacity management, node augmentation management) while maximizing reuse of the operational data models for topology/usage modeling and usage computation/alarm detection in production. The operational digital twins are used in modeling/designing the traditional digital twin in the network simulation to create a closed loop to ensure that maximum effectiveness will be achieved when the node augmentation plan is executed. This is achieved by starting with what is in production, computing the best augmentation in a simulated environment, and then creating the augmentation plan to change what is in production.

The focus of the paper is on optimization of access network node augmentation, but the approach outlined here can be applied to any network optimization concept (optimization of backbone network augmentation, route optimization, traffic management, etc.) if the following functions can be implemented:

1. Operational Digital Twin for the resource to be optimized (the underlying entity, its topology and the usage modeling)
2. Cable network operators' usage computation capability
3. Ability to compute alternative resource configuration (alternative topology/configuration and/or resources)
4. Ability to project usage to the alternative topology/configuration
5. Ability to compute usage for the alternative topology (ideally this is just an instance of 3)

With careful use of "operational digital twin" and "traditional/simulated digital twin" the 1, 2, and 5 can be just existing production data and capabilities and only 3 and 4 need to be developed for new use cases. This approach not only ensures that software capabilities in production are reused to the maximum and new development is minimized, but it will also increase the consistency and quality of the outcome.

We believe that this approach will minimize surprises after the execution of the proposed augmentation by increasing the outcome's predictability and efficiency and achieving the goal of making node augmentation "more science than art."

Abbreviations

BSOD	Business Service Over DOCSIS
CMT	Comcast Mid-Tier (Data distribution)
CPE	customer premise equipment
DAA	distributed access architecture
DOCSIS	data-over-cable service interface specifications
FDX	full duplex
FEC	forward error correction
HFC	hybrid fiber-coaxial
OFDM	orthogonal frequency division multiplexing
OFDMA	orthogonal frequency division multiple access
PHY	physical layer
PMA	profile management application
RF	radio frequency
RPD	remote PHY device
SC-QAM	single carrier quadrature amplitude modulation
SCTE	Society of Cable Telecommunications Engineers
vCMTS	virtualized cable modem termination system

Bibliography & References

Include an annotated bibliography of key resources providing additional background information on your topic.

1. "How Network Topology Impacts Rf Performance: A Study Powered By Graph Representation Of The Access Network", M. Harb, K. Subramanya, R. Narayanaswamy, S. Walavalkar, and D. Rice, SCTE/NCTA technical paper (2021).
2. https://neo4j.com/
3. https://neo4j.com/graph-algorithms-white-paper/