Closed Loop Capacity Optimization for Extended Spectrum DOCSIS

A Technical Paper prepared for SCTE•ISBE by

Dr. Thushara Hewavithana
Senior Architect, Network Platform Group
Intel Corporation
5000 West Chandler Blvd, Chandler, AZ 85226
602-245-1468
Thushara.hewavithana@intel.com

Dr. Rainer Strobel
Lilienthalstr. 16, 85579 Neubibrg
rstrobel@maxlinear.com

Nader Foroughi
Senior Network Architect, Access Architecture & Technology
Shaw Communications Inc
2728 Hopewell Place NE. Calgary, AB T1Y7J7

Table of Contents

Title **Page Number**

List of Figures

Title **Page Number**

List of Tables

Title **Page Number**

1. Introduction

Achievable data rates in the fixed access network keep increasing. While passive optical networks (PON) move to 25 or 50 Gbit/s with IEEE 802.3ca [1] and MGfast [2] targets 10 Gbit/s aggregated point-to-point rate over twisted pair, it is time to evaluate HFC technology as a successor for 10 Gbit/s capable DOCSIS, using full duplex or 1.8 GHz bandwidth [4].

Figure 1 compares the data rate trends for different access technologies. DSL, as a point-to-point technology is at lower rates, but with a higher growth rate. While DOCSIS and PON, both shared medium technologies, follow a similar trend with lower growth rate of the aggregated rate, which is compensated by reducing the number of subscribers sharing the bandwidth as an additional measure.

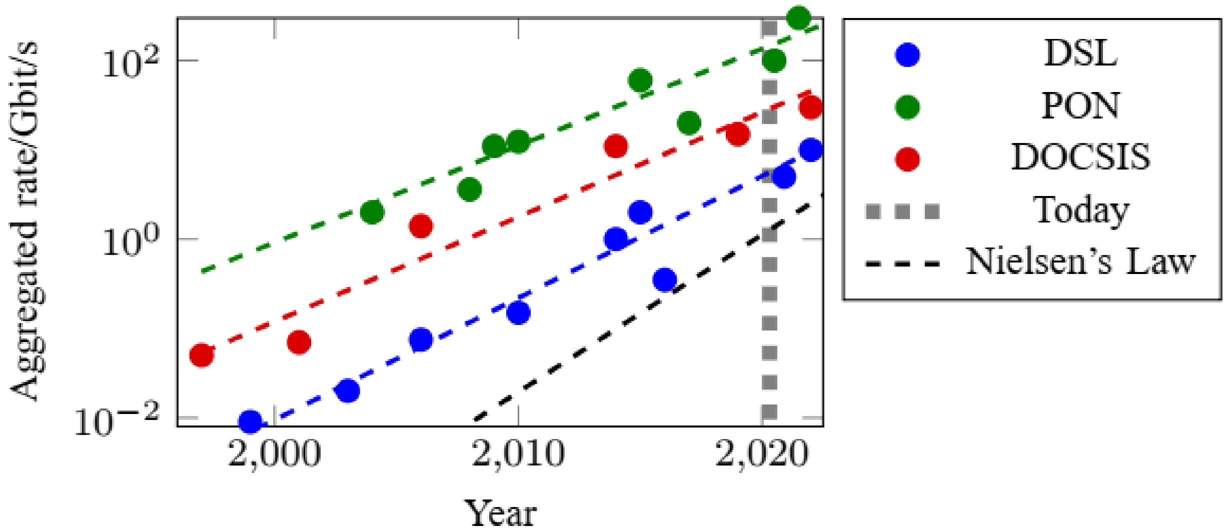


Figure 1 – Agrregate rates of different access network topologies by standard introduction year

From Figure 1, aggregated data rates around 30 Gbit/s are a competitive choice for a future DOCSIS generation, which is herein called extended spectrum DOCSIS (ESD). This will allow 20-25 Gbit/s downstream (DS) and 5-10 Gbit/s upstream (US) rates, which is comparable to a single 25G PON wavelength service that is shown in [5] to serve future access network requirements. Following the arguments of [6], this will allow for 10 Gbit/s services and cover the bandwidth growth predicted by Nielsen's law [7].

The cable industry has recognized ESD as a viable path to extend competitiveness of DOCSIS network at a fraction of cost compared to fiber deployments going forward. Under the 10G DOCSIS initiative, 1.8 GHz frequency division duplexing (FDD) DOCSIS and 1.2 GHz full duplex DOCSIS options have been included in DOCSIS 4.0 as two possible ways of getting to 10 Gbit/s node throughput, enabling low single digit Gbit/s services. We can consider 1.8 GHz FDD to be an intermediate step to get to the 3 GHz ESD and 10 Gbit/s services. In this paper, we will focus more on 1.8 GHz ESD when describing algorithms and evaluating performance results. Nonetheless we will maintain the forward compatibility of our algorithms for a future 3 GHz ESD solution.

One of the key enablers of ESD is the advancement of power amplifier (PA) technology that can support multi GHz transmit signal. However, total composite power (TCP) of these PAs does not scale with the increased spectrum beyond 1218 MHz. Therefore, the ESD communication system is limited in its transmit power. Optimal allocation of available transmit power and appropriate bit-loading (profile definition) is needed to get the maximum capacity out of the network.

In this paper, we outline a framework for closed loop optimization of the capacity of ESD systems subjected to the TCP constraint mentioned above. Cable modems provide the node with channel estimate and signal to noise ratio estimate data and the intelligent node uses this data to calculate the optimal power allocation and bit-loading for the downstream. This can be made part of the profile management application running on a virtual Cable Modem Termination System (vCMTS). We have shown that combining careful allocation of channels, closed loop optimization of transmit power, and adaptive bit-loading achieves considerable gains in data rate and reduction in TCP for network topologies currently present in MSO networks.

2. Technological and Operational Challenges with Extended Sepctrum DOCSIS

As MSO's expand their spectrums to 1.8 GHz and beyond, the expectation for the achievable modulation order at the top end of the spectrum is reduced. In traditional plant, being 750 MHz or even 1 GHz, most MSOs expect 4096 QAM to be achievable by each orthogonal frequency division multiplexing (OFDM) block deployed there. But the same is not true for 1.8 GHz and beyond. There are many reasons why that is the case but the primary one would be the current spacing that outside plant (OSP) is designed to. Due to this, any method to increase the achievable throughput in the network is highly sought after.

This section explores the technological and operational challenges of upgrading the plant to 1.8 GHz.

2.1. Plant Spacing and Drop-In Upgrades

Most OSP architectures are designed to 750 MHz and 'stretched' to 1 GHz. As a result, most of the RF power in plant actives, including nodes and amplifiers, has been utilized to overcome the existing span losses. Span loss is defined as the total insertion loss of all the elements in an HFC span, measured in dB. This includes all the plant passives such as splitters and couplers. Although the span losses today are manageable with the current amplifier gains, they will certainly become a major point of concern when the spectrum is expanded to higher frequencies. As an example, a span loss of 35 dB in a traditional 1 GHz plant equates to 49 dB in 1.8 GHz.

Although plant re-spacing is always an option, it can be extremely costly and as a result, operators will rely on the expanded power of amplifier gain chips to overcome span losses.

2.2. Total Composite Power (TCP)

It is generally understood that the 1.8 GHz amplifiers will have ~75 dBmV of total composite power (TCP) available, which is roughly the same power available in current 1.2 GHz devices. With that been said, not all of this power is available for use. As an example, Figure 2 demonstrates the trade-off between Modulation Error Ratio (MER) and TCP utilized.

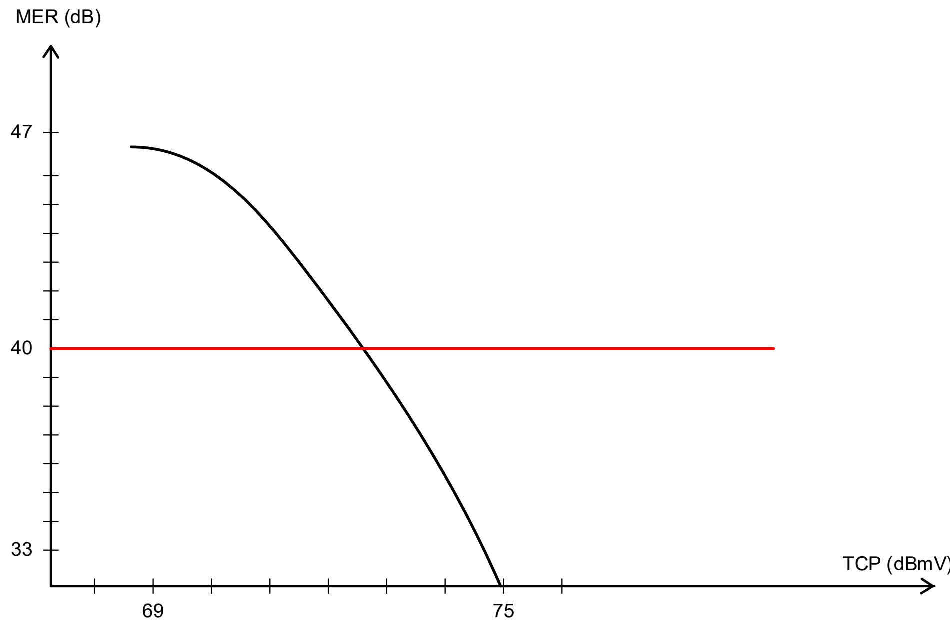


Figure 2 – Amplifier Nonlinear Distortion

As a general rule of thumb, 3 dB of back-off is needed to achieve 40 dB+ MER, which is typically what operators aim for. Along with that, the internal loss of the active device has to be accounted for, which is typically 2 dB. To summarize, there is a total of 70 dBmV of power to be utilized at the port of each active device. This can be a concern for operators given that 65-68 dBmV of TCP has already been allocated to overcome the span losses in the 'traditional' plant.

2.3. Passive Network Updates – Taps and Other Connectors

Taps and passives can be another point of concern when upgrading the OSP to 1.8 GHz. Traditionally, most operators have relied on face-plate upgrades to expand the spectrum range of plant taps and passives. This is generally accepted as a faster and more cost-effective method to upgrade the available bandwidth of taps.

Unfortunately, this might not be the case with 1.8 GHz upgrades. Based on the research and the information released by the vendor community, a face-plate upgrade of the current 1 GHz taps can potentially expand the bandwidth to ~1.6 GHz. it should also be noted that this is a best effort.

This can be a concern given the uncertainty of the maximum available bandwidth in the plant. As a result, it is generally accepted that taps and passives have to be swapped out for 1.8 GHz version. Given that the entire housing of the tap has to be swapped out as a part of this effort, most of the taps being developed will have housings that can support up to 3 GHz with future face-plate upgrades, future proofing the plant for 3 GHz upgrades.

2.4. Plant Model

In order to encompass most of the hybrid-fiber-coax (HFC) architectures deployed, the following plant models and assumptions were considered for this analysis.

Note: Trunk spans are defined as spans that are untapped. Distribution spans are tapped. Both trunk and distribution span losses include all other passive elements' insertion losses, such as splitters and couplers.

Assumptions:

- Modem:
  - Point of entry (PoE) device
- Drop:
  - Cable: RG6
  - Length: 150 feet
- Number of taps in each span:
  - 5
- Distribution span losses at 1 GHz:
  - Typical plant: 35 dB
  - Stretched plant: 37 dB
- Trunk span losses at 1 GHz:
  - Typical plant: 32 dB
  - Stretched plant: 35 dB

Figure 3 and Figure 4 summarize the parameters above.

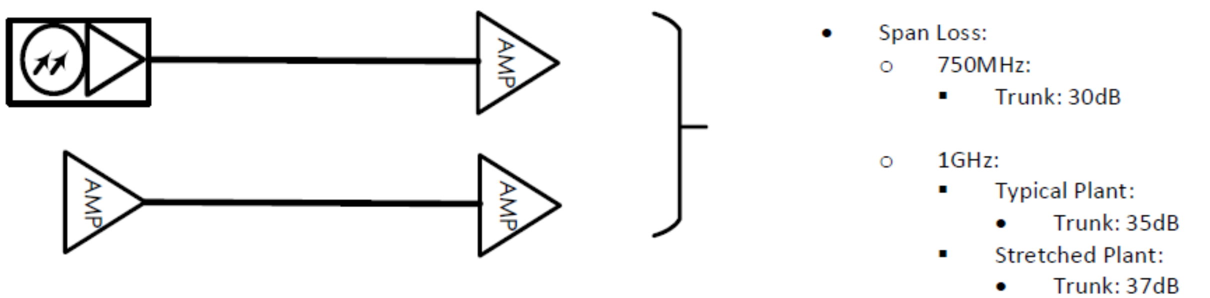


Figure 3 – Trunk Span Losses

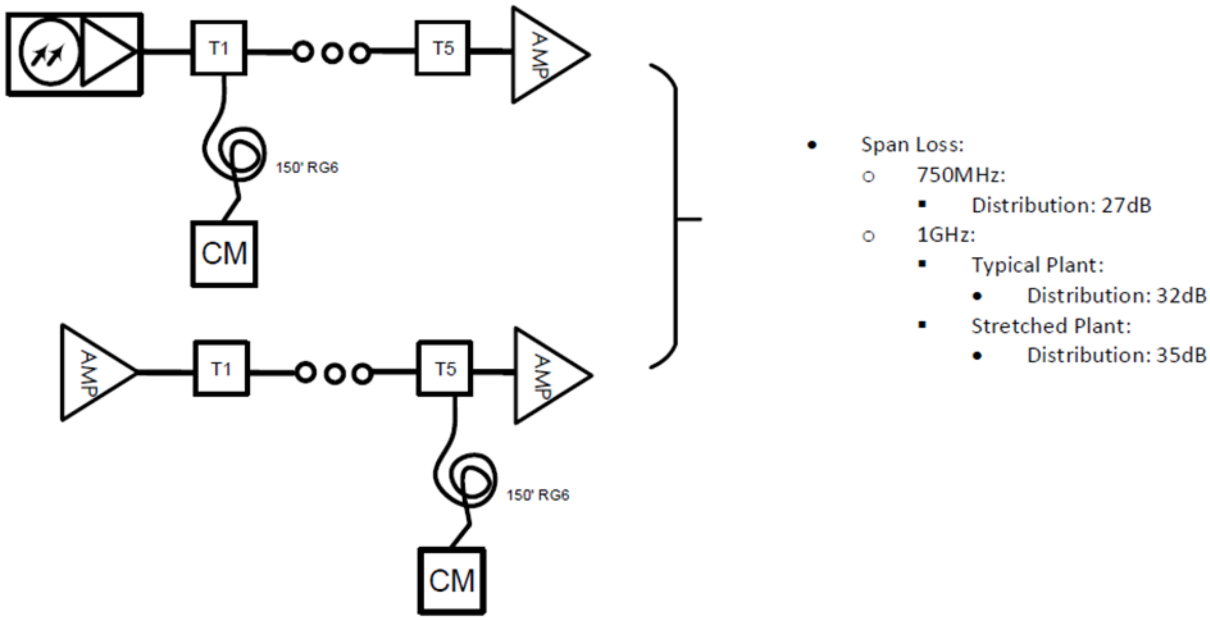


Figure 4 – Distribution Span Losses

2.4.1. N+0 vs N+X

HFC plant can be divided into two categories: passive and cascaded.

- Passive plant (N+0): where no amplifiers are used after the node
- Cascaded plant (N+X): where amplifiers are used to boost the signal multiple times to the end-of-line

When designing an N+0 plant, the main point of concern is the output power of the node. Assuming that we are operating in a distributed access architecture (DAA) plant, the primary drivers for the plant quality would be the modulation error ration (MER) of the DAA device. Since no amplifiers are used to boost the signal, no noise or distortion is added to the primary signal being generated by the DAA device.

On the contrary, when designing a cascaded plant, the following can be a concern:

- Amplifier noise contribution
- Amplifier distortion contribution

2.4.2. Noise

Designing a cascaded system for optimal carrier to noise is always a big priority for an operator. One of the biggest contributors in system design is the receive power (Rx Power) at the amplifier, given that it is one of the primary drivers for the overall system carrier to noise (C/N).

The equation below calculates the C/N of an amplifier:

$$C/_{N}\,(dB) = C_i(dBmV) + 57.4 - NF(dB) \qquad (1)$$

Where:

- C_i: input signal

- NF: Noise figure of the amplifier

Note: the number 57.4 is the noise power for QAM carriers. This value will vary marginally depending on the temperature.

Equation (1) shows the significance of the Rx power versus noise figure of the amplifier, in overall system design.

The overall system C/N can be derived from the following equation:

$$C/N_{total}\,(dB) = -10log\left\{10^{\frac{-C/N1}{10}} + 10^{\frac{-C/N2}{10}} + \cdots + 10^{\frac{-C/Nn}{10}}\right\} \tag{2}$$

Where $C_{/N_x}$ is the carrier to noise of each amplifier calculated independently.

When cascading identical amplifiers, the following approximation is typically used:

$$C/N_{total}\,(dB) = C/N_x - 10log_n \tag{3}$$

Where:

- C/N_x: the carrier to noise of a single amplifier
- n: the number of identical amplifiers in cascade.

2.4.3. Distortion

The build-up of distortions in a cascaded plant are less predictable than noise. The following equation can be used to estimate the carrier to composite triple beat (C/CTB) and carrier to composite second order distortion (C/CSO):

$$C/CTB_{total}\,(dB) = C/CTB_x - 20log_n \tag{4}$$

$$C/CSO_{total}\,(dB) = C/CSO_x - 10log_n \tag{5}$$

Where:

- C/CTB_x: the distortion of a single amplifier
- C/CSO_x: the distortion of a single amplifier
- n: number of identical amplifiers in cascade

2.4.4. Designing a Noise-Limited System

For optimal performance, operators design systems that are unity gain. This means that the loss between two amplifiers is equal to the gain of each amplifier. If the loss is less than the gain, the distortions with accumulate. Whereas if the loss is greater than the gain then the input power will be less than the desired amount, degrading the system C/N.

Due to the difficulties that come with designing a system that is both noise and distortion limited, removing one of those parameters will be optimal. Given that noise performance of amplifiers is far more straight-forward in comparison to distortion, designing a noise-limited system is a very attractive idea.

Since distortions are highly dependent on output power, designing a noise-limited system can be achieved by reducing the output power out of the node/amplifiers and making sure the signal level received at the next amplifier is high enough, based on equation 1. Figure 5 demonstrates balancing the signal level against noise and distortion in a system design.

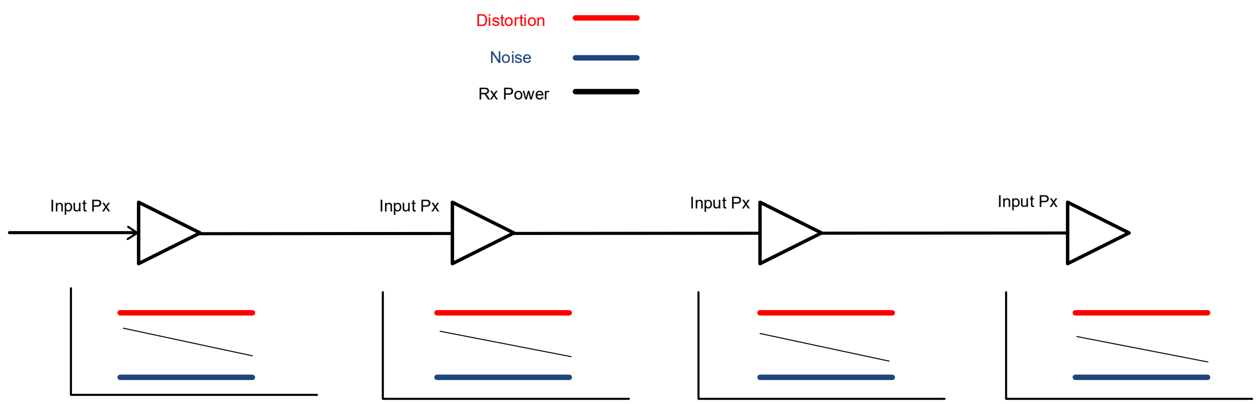


Figure 5 – Signal Level Balanced Between Noise and Distortion

2.5. Typical Plant Topologies

In this paper we focus on Node + 0 (Figure 6) and Node + 4 (Figure 7) networks as the basis for performance evaluation.

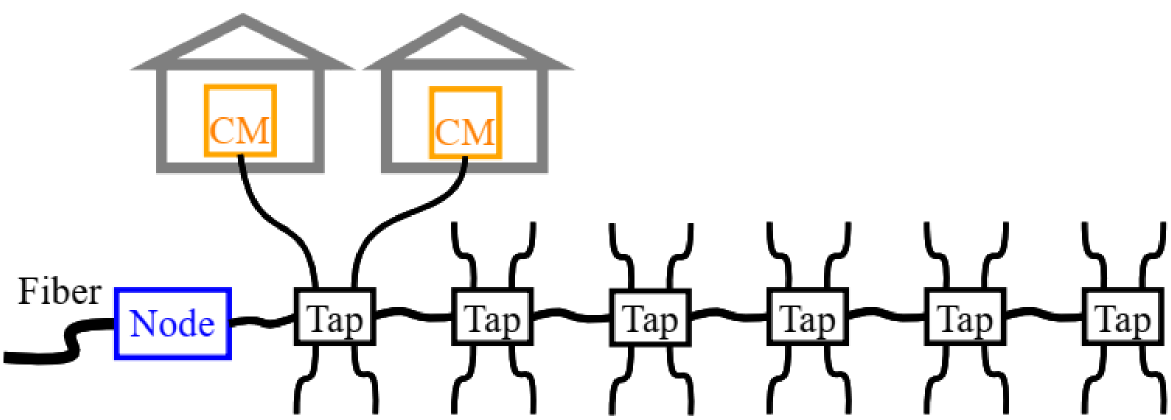


Figure 6 – Node + 0 Passive HFC Plant Topology

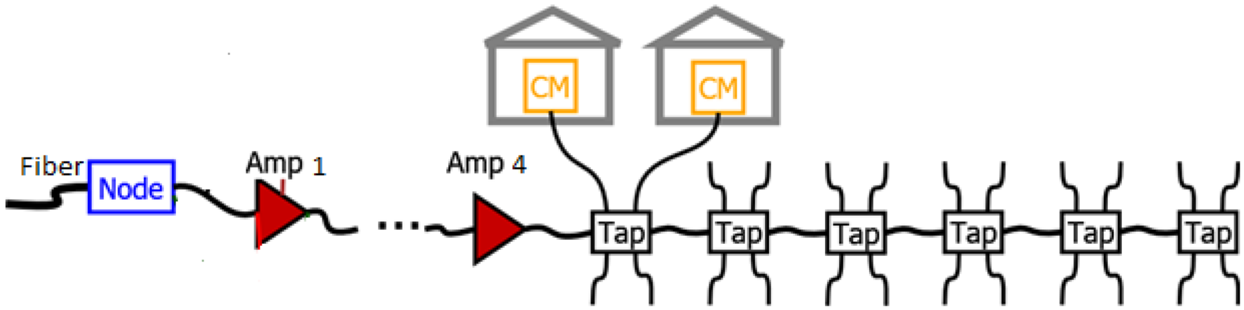


Figure 7 – Node + 4 Cascade Plant Topology

This plant topology has 5 amps total in cascade, the node amp and the 4 network amps. The network is built with a cascade of 4 trunk spans and single distribution span as described in section 2.4.

2.6. Problem Definition

As described before, the ESD communication system is limited in its transmit power. Optimal allocation of available transmit power and appropriate bit-loading (profile definition) is needed to get the maximum data rate out of the network. In this paper, we outline a framework for closed loop optimization of the throughput of ESD system subjected to TCP constraint mentioned above. We will show that combining careful allocation of channels, closed loop optimization of transmit power, and adaptive bit-loading achieves considerable gains in data rate and reduction in TCP for network topologies currently present in MSO networks.

3. Theoretical Framework for Closed Loop Throughput Optimization

In this section, we present the theoretical framework for optimizing the available transmit power for throughput of the cable network. We start off with the simpler case of Node + 0 passive network and then extend the theory to cover general case of Node + X, X>0, networks.

3.1. Node + 0 Network

The capacity evaluation for the extended spectrum HFC network requires knowledge of the channel characteristics and capacity limiting factors. The capacity limiting factor in the transmitter is amplifier distortion. At the receiver, additive white Gaussian noise and receiver distortion due to analog-to-digital conversion limits capacity. HFC transmission schemes such as DOCSIS 4.0 [4] use OFDM modulation, where the channel is partitioned into K narrowband subcarriers k = 1, . . . ,K with a subcarrier spacing Δf. Those orthogonal channels are coupled only by nonlinear distortion or a sum power constraint. The transmit power per carrier x(k) as well as the information rate per carrier b(k) can be adjusted per carrier. The data rates for a given signal-to-noise ratio SNR(k) on carrier k is given by

$$R = \eta \Delta f \sum_{k=1}^{K} \min\left(\log_2\left(1 + \frac{SNR^{(k)}}{\Gamma}\right), b_{\max}\right). \tag{6}$$

where limitations of modulation and coding are considered in terms of an SNR gap to capacity Γ [12] as well as with a limit b_{max} to the number of bits transmitted per carrier and channel use. The OFDM system requires overhead for the cyclic extension to guarantee orthogonal channels, which is considered in an efficiency factor η. Using η = 1, Γ = 1 and $b_{max} \rightarrow \infty$ gives the capacity without coding and modulation limitations.

Capacity, C, and achievable rate, R, are evaluated with respect to power constraints where the simplest case is a sum power constraint [13]. For practical systems, additional per-carrier constraints are considered, as shown in [14], Sec. 3.1.6. For this case, achievable rate, R, and capacity C for the case of Γ = 1, is the solution to

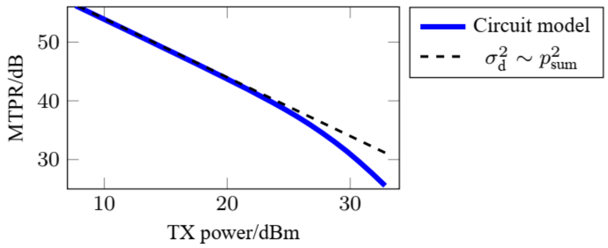
$$\begin{aligned} R_l &= \max_{x_l^{(k)}} \sum_k \log_2\left(1 + \frac{|H_l^{(k)}|^2 x^{(k)}}{\Gamma \sigma_l^{(k),2}}\right) \\ &\text{s.t. } \sum_k x^{(k)} \le p_{\max} \\ &\text{s.t. } 0 \le x^{(k)} \le p_{\text{mask}}^{(k)} \end{aligned} \tag{7}$$

where $H_l^{(k)}$ is the channel coefficient on carrier k (attenuation, phase) between node and cable modem (CM) l and $\sigma_l^{(k).2}$ is the additive white Gaussian noise variance on carrier k for CM l.

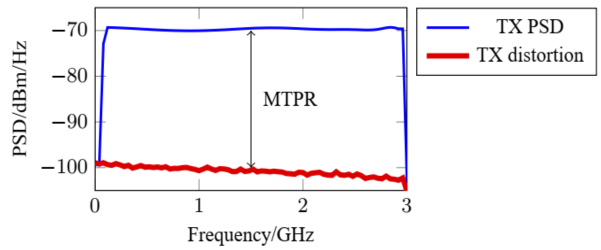
The power constraints are formulated as a sum power limit p_{max} and a spectral mask constraint $p_{mask}^{(k)}$. Limitations of the modulation alphabet size to b_{max} are incorporated into the spectral mask constraint, using $p_{mask}^{(k)} = {\Gamma(2^{b_{max}} - 1)\sigma^{(k),2}}\Big/{|H^{(k)}|^2}$. The solution to Eq. (7) is obtained by a modified water-filling algorithm as described in [14], chapter 3.1.6. Other algorithms to solve Eq. (7) have been published in [15],[17].

The sum power limit, $\sum x^{(k)} \le p_{max}$, in Eq. (7) can be seen as a simplified model for the behavior of real transmit amplifier, where the SNR and thus the data rate is limited by distortion increasing with increasing transmit power. Transmit amplifier distortion can be seen as a transmit power dependent noise source with variance σ_d^2. It is characterized in measurement and simulation by a missing tone power ratio (MTPR). MTPR is the ratio between signal power and distortion power, as shown in Figure 8 (b). It is determined as the signal level on one OFDM carrier which is transmitted with zero power while the others are transmitted at the desired level ($MTPR^{(k)} = {x^{(k)}}\big/{\sigma_d^2}$). In the following discussion, the cable modem index l is skipped without loss of generality.

Figure 8 (a) shows the increase of nonlinear distortion σ_d^2 in a 3 GHz amplifier circuit model. The MTPR decreases with increasing transmit power, $p_{sum} = \sum_{k=1}^{K} x^{(k)}$. In frequency domain, as shown in Figure 8 –(b), distortion is approximately flat.



(a)



(b)

Figure 8 – Distortion evaluation from amplifier circuit model

The dependency between distortion variance σ_d^2 is and signal power $p_{\text{sum}}(x^{(k)})$ can be described by $\sigma_d^2 = \delta\left(p_{\text{sum}}(x^{(k)})\right)^{\alpha}$. For the amplifier shown in Figure 8 (a), the constants are δ = −64 dB and α = 2. Following the argumentation of [9] a lower bound for the capacity of the nonlinear copper channel is derived.

Introducing distortion in the SNR per carrier gives the term

$$SNR^{(k)} = \frac{|H^{(k)}|^2 x^{(k)}}{\sigma^2 + \delta^{(k)}\left(p_{\text{sum}}(x^{(k)})\right)^{\alpha}} \tag{8}$$

where the distortion variance is $\sigma_d^{(k),2} = \delta^{(k)}\left(p_{\text{sum}}(x^{(k)})\right)^{\alpha}$, assuming white distortion. This gives the rate R (or capacity C with Γ = 1) according to

$$\begin{aligned} R = \max_{x^{(k)}} \sum_k \log_2\left(1 + \frac{|H^{(k)}|^2 x^{(k)}}{\Gamma\left(\sigma^2 + \delta^{(k)}\left(p_{\text{sum}}(x^{(k)})\right)^{\alpha}\right)}\right) \\ \text{s.t. } 0 \le x^{(k)} \le p_{\text{mask}}^{(k)} \end{aligned} \tag{9}$$

The derivative $\frac{\partial R}{\partial x^{(k)}}$ is given by

$$\begin{aligned} \frac{\partial R}{\partial x^{(k)}} = & \frac{\left|H^{(k)}\right|^2}{\left(\sigma^2 + \delta_k\left(p_{\text{sum}}(x^{(k)})\right)^{\alpha}\right)\left(\Gamma + SNR_k\right)} - \\ & \sum_{d=1}^{K} \frac{\left|H^{(d)}\right|^2 x^{(d)} \alpha \delta^{(d)}\left(p_{\text{sum}}(x^{(k)})\right)^{\alpha-1}}{\left(\sigma^2 + \delta^{(d)}\left(p_{\text{sum}}(x^{(k)})\right)^{\alpha}\right)^2\left(\Gamma + SNR^{(d)}\right)} \end{aligned} \tag{10}$$

and $\partial R/\partial x^{(k)} = 0$ must hold for the optimal power allocation for all carriers with $0 < x^{(k)} < p_{\text{mask}}^{(k)}$. The optimal power allocation can be found, e.g., by a projected gradient method with a step size ρ as given by

$$x_{t+1}^{(k)} = \min\left(\max\left(x_t^{(k)} + \rho\frac{\partial R}{\partial x^{(k)}},0\right),p_{\text{mask}}^{(k)}\right) \tag{11}$$

It can be shown [17] that the solution to equation (10) takes the following form,

$$\frac{1}{\mu} = \sum_{d=1}^{K}\frac{\left|H^{(d)}\right|^2 x^{(d)}\alpha\delta^{(d)}\left(p_{sum}(x^{(k)})\right)^{\alpha-1}}{\left(\sigma^2+\delta^{(d)}\left(p_{sum}(x^{(k)})\right)^{\alpha}\right)^2\left(\Gamma+SNR^{(d)}\right)} \tag{12}$$

The dependency between μ and p_{sum} is given by,

$$\frac{1}{\mu} = \frac{1}{|I_{\text{fill}}|}\left(p_{sum} + \sum_{k\epsilon I_{\text{fill}}}\frac{\Gamma\sigma_{nd}^{(k),2}}{|H^{(k)}|^2} - \sum_{k\epsilon I_{mask}} p_{mask}^{(k)}\right) \tag{13}$$

Where I_{mask} is the set of subcarriers for spectral mask constraint is active and I_{fill} is the set of subcarriers with power allocation to meet water-fill level, and $|I_{\text{fill}}|$ denotes cardinality (number of elements) of the set $|I_{\text{fill}}|$.

Hence the transmit power per subcarrier is given by,

$$x^{(k)} = \begin{cases} \frac{1}{\mu} - \frac{\Gamma\sigma_{nd}^{(k),2}}{|H^{(d)}|^2} & \text{for } k\epsilon I_{\text{fill}} \\ 0 & \text{for } k\epsilon I_0 \\ p_{mask}^{(k)} & \text{otherwise} \end{cases} \tag{14}$$

Where I_0 is the set of subcarriers where positiveness constraint given in equation (7) is active and therefore no power is allocated.

To implement the optimization scheme, the distortion parameters, δ and α must be known from an amplifier characterization. During operation, the noise conditions must be known from an SNR measurement. The algorithm performs multiple water-filling steps. As the optimization can be done by software in background during operation of the link, there is no issue with computation time.

3.2. Extention to Node + X, X > 0, Networks

The algorithm described in section 3.1 assumes a single source of transmitter distortion at the node and a single receiver noise source at each CM, as it is reflected in the SNR in Eq. (8). In case of a Node + X topology with multiple intermediate amplifiers, each intermediate amplifier represents an additional source of distortion and each amplifier input experiences additive receiver noise.

Still, it can be shown that under a certain assumption, the optimization framework of section 3.1 can be applied to the Node + X case, too. The precondition for the spectrum optimization to be applied is that the attenuation and down-tilt of the cable section, given by the transfer function $H_{span}^{(k)}$, is compensated by amplifier stage $H_{amp}^{(k)}$ at the end of the cable section, such that the power level and spectral shape of the transmit signal is the same at each intermediate amplifier $H_{span}^{(k)}\ H_{amp}^{(k)} = 1$, as shown in Figure 9.

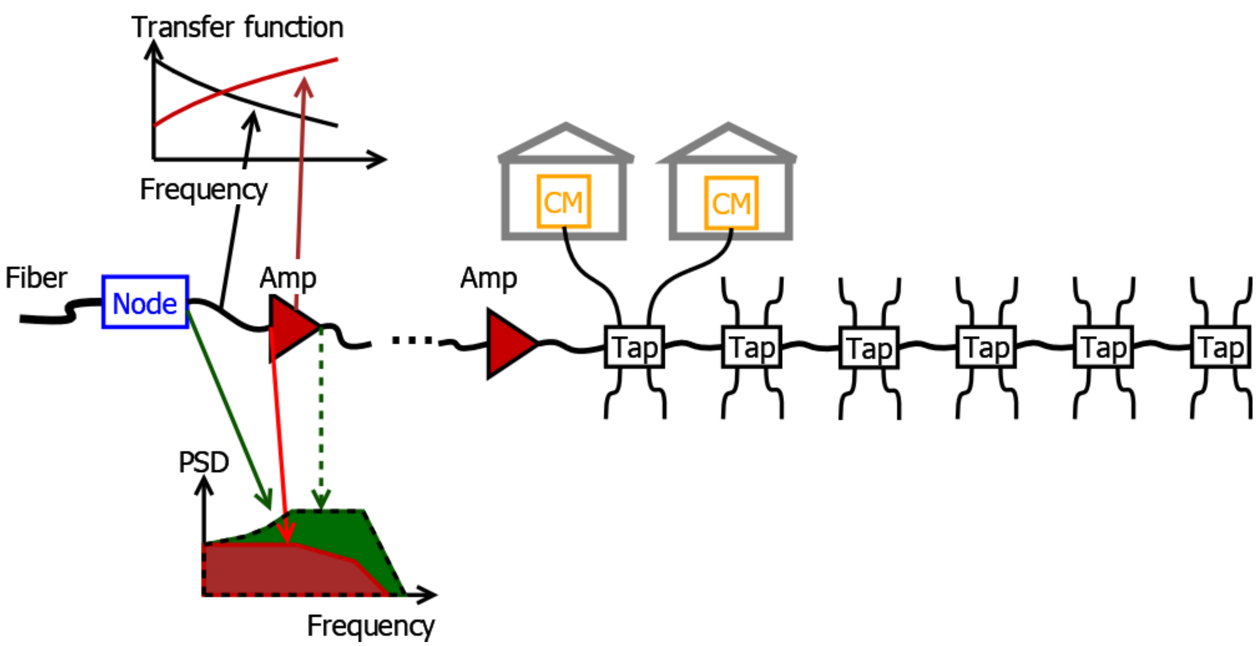


Figure 9 – Node+X topology with Power Spectrum Optimization

Mathematically, the transmit power per carrier k, $x^{(k)}$ is (approximately) the same at node output and the amplifier outputs , $x^{(k)} \approx x_{\text{amp},1}^{(k)} \approx \ldots \approx x_{\text{amp,n}}^{(k)}\ \forall\, k = 1, \ldots, K$. With this condition satisfied, all the amplifier distortion of N amplifiers can be combined into one distortion $(N+1)\ \delta^{(k)}\big(p_{sum}(x^{(k)})\big)$.

The receiver noise of the amplifiers, assuming $\sigma_{amp}^2 \ll \left|H_{span}^{(k),}\right|^2 x^{(k)}$ for all the used frequencies, can be handled as additive noise. The receiver noise of the amplifiers is summed up, referred to the CM receiver and added to the CM receiver noise σ^2 which gives the overall additive noise term to be $N\ \sigma_{amp}^2 \left|H_{amp}^{(k)}\right|^2 \left|H^{(k)}\right|^2 + \sigma^2$.

Accordingly, the algorithm of section 3.1 remains as is while performing the following substitutions

Receiver noise: σ^2 ➔ $N\ \sigma_{amp}^2 \left|H_{amp}^{(k)}\right|^2 \left|H^{(k)}\right|^2 + \sigma^2$

Transmitter distortion: $\delta^{(k)}\big(p_{sum}(x^{(k)})\big)$ ➔ $(N+1)\ \delta^{(k)}\big(p_{sum}(x^{(k)})\big)$

Compared to the Node + 0 architecture, transmitter distortion optimization is even more relevant in the Node + X architecture, as there is a higher distortion level present due to the distortion of multiple amplifiers adding up. It may also be beneficial to drive the TCP of individual amps down as much as possible to lower individual nonlinear distortion contributions from amps. We will look at strategies on achieving this objective in this in this paper.

3.3. Algorithm Description

The above framework leads to an iterative throughput optimization algorithm described below.

Initialization Step: Initialize the TCP, $P_{sum}(0)$, to the target maximum power level. Index 0 refers to initial value

Following that, the following steps are performed in an iterative loop, until the $P_{sum}(n)$ converges to a steady value. $|P_{sum}(n) - P_{sum}(n+1)| < Threshold$

Iterative Step 1: Apply water-filling described in [14] with the current P_{sum} to discover,

- I_0 = The set of subcarrier that get no power allocation
- I_{fill} = The set of subcarriers that gets power allocation to water-fill level
- I_{mask} = The set of subcarrier that gets power allocated to $p_{mask}^{(k)}$ level. For these subcarrier $p_{mask}^{(k)}$ level is hit before reaching water-fill level and therefore no additional power is wasted to reach water-filling level

Iterative Step 2: Based on subcarrier sets information from previous step and Equation (14), calculate the power allocation per subcarrier, $x^{(k)}$.

Iterative Step 3: Calculate water-filling level, $\frac{1}{\mu}$, using equation (12)

Iterative Step 4: Update the TCP, $P_{sum}(n+1)$, from equation (13)

Exit Criteria: If $|P_{sum}(n) - P_{sum}(n+1)| < Threshold$ then stop iteration. Otherwise go back to Iterative Step 1.

Pseudo code for the algorithm is given below.

Table 1 – Optimization algorithm summary

Initialization	$P_{sum}(0)$ = max allowed TCP
Iteration	Identify the sets I_0, I_{fill}, and I_{mask} (water-filling, [14]) Update $x^{(k)}$ using Eq. (14) Calculate μ from Eq. (12) with updated $x^{(k)}$ Update $P_{sum}(n+1)$ from μ, using Eq. (13)
Exit Criteria	$|P_{sum}(n) - P_{sum}(n+1)| < Threshold$

4. System Level Solution for Closed Loop Optimization

In order to apply the closed loop optimization algorithm developed above in a practical system, we first need to consider the impact of channel allocation. Channel allocation and power optimization together form our overall closed loop solution. We consider two power allocation strategies; Full spectrum allocated to each CM, and channel stacking or staggered channel allocation.

4.1. Full Spectrum allocated to each CM

Consider the mid-split US/DS partitioning with all the CMs allowed to use DS OFDM channels anywhere in the spectrum from 108 MHz to 1794 MHz. In this case, power allocation is optimized considering channel frequency response and noise of all channels for a CM connected to a particular tap. Given that each channel is potentially shared between all CMs in the node, we have to carefully select the CM that we target the optimization algorithm for (i.e. what channel frequency response and noise responses to use in the algorithm). Going for the worst-case CM (farther away from node) or the best-case CM (closer to node) may not lead to overall node throughput optimization. A CM that represents median or overage behavior, in terms of received signal quality, would lead to better results. Once the downstream power is optimized, DOCSIS has other tools, such as the profiles, to fine tune the throughput for CMs with different received signal quality.

4.2. Channel Stacking

In this case, based on the observation that the cable channel has higher losses at higher frequencies, we allocate the lower frequency channels to far away CMs and higher frequency channels to close in CMs. It will be shown later that this allows us to reduce the TCP of node and amps significantly, opening up potential other benefits in overall network architecture.

4.3. Implementation Considerations

Closed loop algorithm implementation considerations are described in this section. The aim is, as much as possible, to work within the current DOCSIS 4.0 standard provisions to implement the algorithm described in previous section. We also highlight aspects of the standard that can be improved to better facilitate the closed loop throughput optimization.

The algorithm described in section 3.3 requires per subcarrier channel frequency response estimates and total noise estimates from the CMs to derive the optimal power allocation. DOCSIS 3.1 and 4.0 provide following proactive network maintenance (PNM) features to help gather this information from the CM:

- Downstream Channel Estimate Coefficients: CMTS can command CM to send Channel Estimate coefficients to CMTS (section 9.3.4 of [3], [4]).
- Downstream Receive Modulation Error Ratio (RxMER) Per Subcarrier: CMTS can command CM to send MER estimates to CM (section 9.3.6 of [3], [4])

Noise power per subcarrier can also be derived from the above information.

Once the optimal power allocation solution is found, the CMTS needs to shape the downstream transmit spectrum accordingly. The DOCSIS 4.0 spec [4] has provisions for the node to shape the transmit RF spectrum under the following constraints:

- Apply a uniform up tilt (i.e. same gradient across all entire spectrum) to the transmit spectrum so that more power is allocated to higher frequencies.
- Introduce multiple step downs/ups to maintain total composite power bound (sum step downs < 10 dB)
- Spectrum discontinuities, step down/up, are only allowed in OFDM channel boundaries

Figure 10 shows an example power allocation instance allowed by the spec viewed at interface C – before the uniform tilt and power amplification (PA).

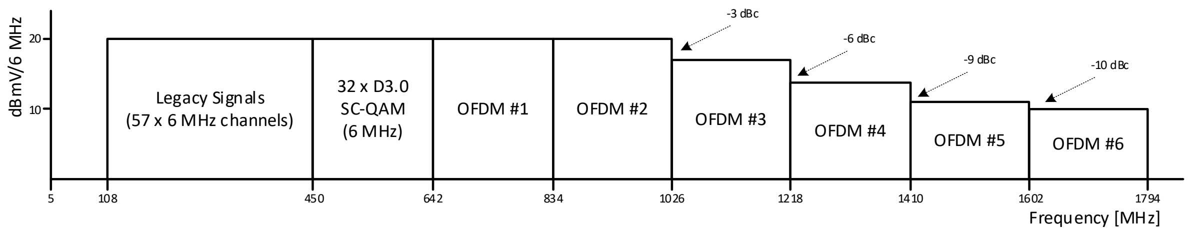


Figure 10 – Downstream Transmit Power allocation as seen from Node Interface C

At interface D (node output port), post tilt application and PA, the above spectrum appears with a uniform up tilt while keeping the same step-downs seen in Figure 10.

One key area where the spec could be amended is to allow for more flexible spectrum shapes with possible flat power spectral density (PSD) at high frequencies at interface D. This means removing the condition of uniform tilt across the entire spectrum mentioned above. This will allow for more accurate implementation of some of the transmit PSD optimization scenarios described later in this document. However, we should weigh the benefits of doing so against potential complications to the node and amp architectures as well as the operation of the network.

We can bring in a machine learning approaches to incorporate various other aspects of the system in the overall solution, such as:

- Take into account the individual CM throughput usage over time in channel allocation
- Overall node-wide throughput usage over time
- Service agreement data for individual CMs
- Prior knowledge of network topology

A profile management application (PMA) can also be used in conjunction with the items mentioned above to increase the overall performance in the distribution plant. By enabling dynamic bit-loading and profiles assigned to each service group, the overall plant throughput and stability of the plant will increase.

4.4. Create Headroom in Power Budget for Soft Flexible MAC Architecture

In this section, we briefly address an added benefit of potential TCP reduction by closed loop optimization in creating room in the overall network power budget for a soft flexible MAC architecture (FMA) solutions. Currently the node power budget is very tight already with remote PHY device (RPD) solutions. With a node power budget ranging from 160 W to 180 W for North America, the RPD and the RF power amplifiers are already using up most of this power. For example, for 85% power delivery efficiency, power left over for other potential uses are roughly 16 W and 33 W for a 2x2 node with 4 legs for overall 160 W and 180 W power budgets respectively. A large chunk of the power is taken up by the 4 PAs, which are assumed to be at 71 dBmV TCP at the node output port.

For FMA solutions, especially remote MAC device (RMD), where MAC functionality is also distributed to the node in addition to PHY, the above excess power is not sufficient for a fully software based solution using general purpose compute, let along having additional headroom for future edge computing applications. This forces a solution which is either based on ASIC/FPGA or/and a CPU with limited compute (to fit within available power budget).

If we can lower the TCP by just 2 dB, to 69 dBmV, this increases the available power for FMA and edge compute to 28 W to 45 W. As shown in section 5, with the additional power saving achievable with closed loop optimization combined with channel stacking, it is possible to reduce the TCP by as much as 6 dB for1.8 GHz ESD. Even for 1.2 GHz DOCSIS, we can explore channel stacking with power optimization to get some reduction in power consumption. This opens the possibility of a soft MAC architecture and also leaves enough headroom for other future edge compute applications.

5. Simulation Results and Discussion

Rate results for different network topologies, channel allocation strategies and power allocation strategies are given in this section. Node + 0 and Node + 4 network topologies are considered. Allowing all CMs to access the entire downstream spectrum vs channel stacking is also considered. In terms of power allocation, three different allocation strategies are compared:

- Tilted TX PSD to receive flat RX spectrum at CM.
- Flat TX PSD as implied by conventional water filling solution.
- Optimal power allocation based on the algorithm described in this paper.

We explore both the mid and high split plant scenarios:

- Mid-split: DS starts at 108 MHz
- High-split: DS starts at 258 MHz

5.1. Full Optimization for Node + 0 and Node + 4

In this section, we explore the optimization of transmit PSD for capacity without any other constraints imposed by operational considerations. This is partly an academic exercise to gain insight into the properties of the power allocation algorithm

5.1.1. Node + 0 Network, Mid-Split

Figure 11 shows the throughput for CMs connected to different taps for the three different power allocation strategies.

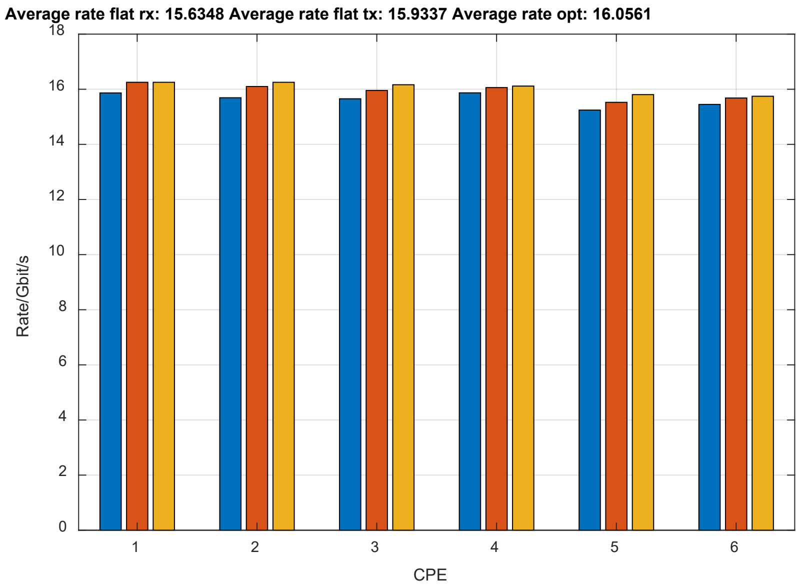


Figure 11 – Rate for flat RX (blue), flat TX (red) and optimal (yellow) power allocation

Throughput gain from using the optimal algorithm is limited in this case.

5.1.2. Node + 4 Network, Mid-Split

The same three power allocation strategies described in section 5 are used here. It is assumed that the distribution cable sections have 35 dB attenuation at 1 GHz. The Node + 4 network is constructed with four straight trunk cable sections with amplifiers followed by a cable section.

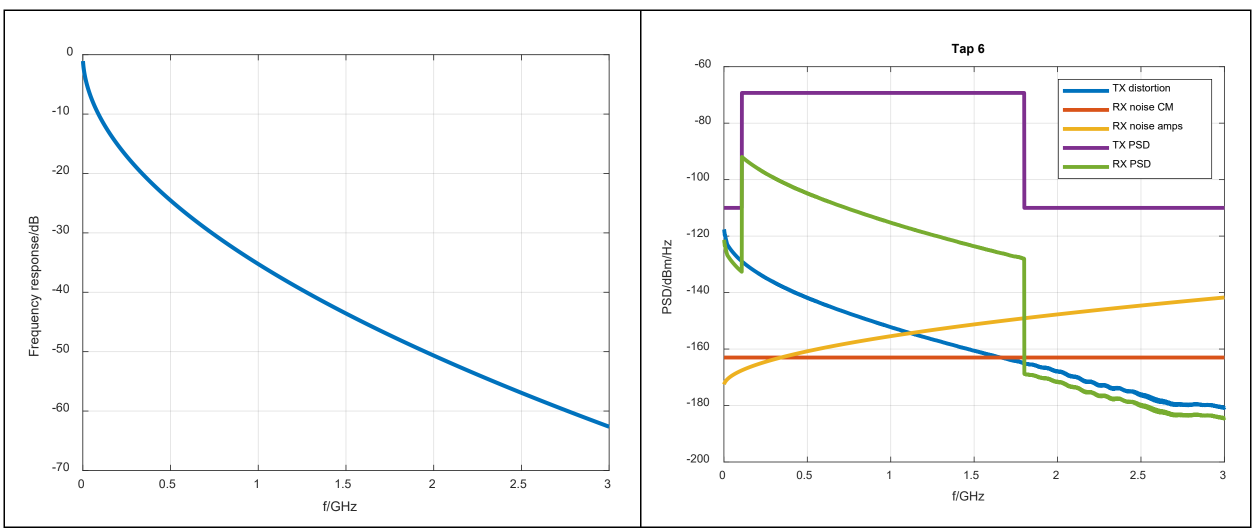


Figure 12 – Attenuation of the cable sections (left) and signal and noise PSDs at the last Tap (right)

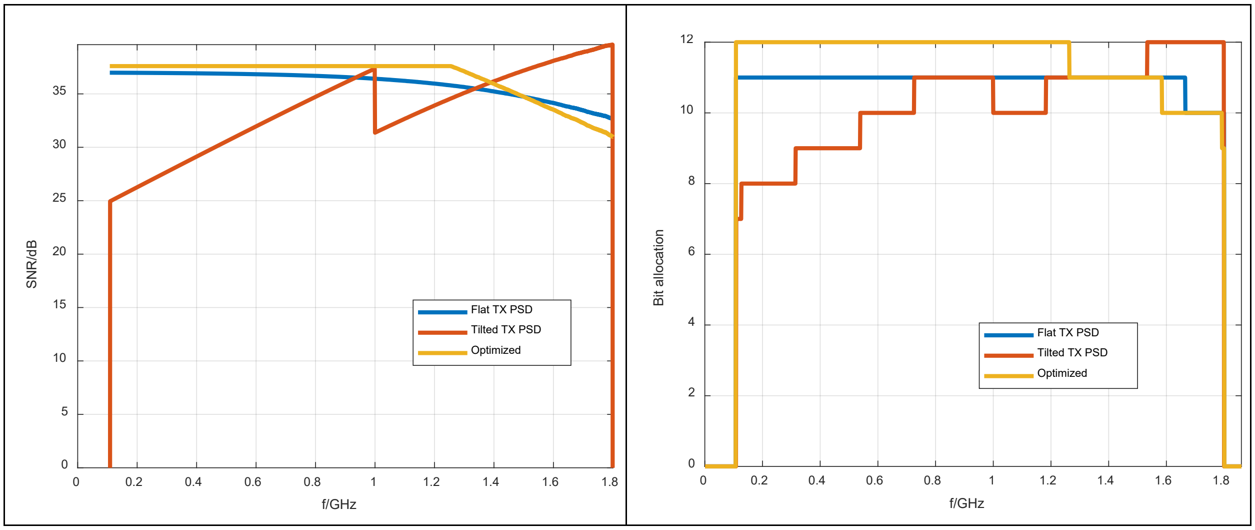


Figure 13 – Node + 4 mid split SNR at Tap 3 (left) and bit allocation at Tap 3 (right)

The distortion of all 5 transmit amplifiers (CMTS and 4 amplifiers) is the same. Each amplifier compensates the channel attenuation and tilt of preceding cable segment completely (segment unity gain) such that the transmit spectrum is the same at each stage. For the noise figure of the amplifiers, 2 amplifiers with 5 dB and 2 amplifiers with 10 dB are assumed. Additional 3 dB of losses in the receiver is assumed (account for losses in passive connectors, diplexers, etc).

Figure 14 shows the throughput for CMs connected to different taps for the 3 power allocation strategies. Up to 10% Rate improvement is achieved with the optimal power allocation.

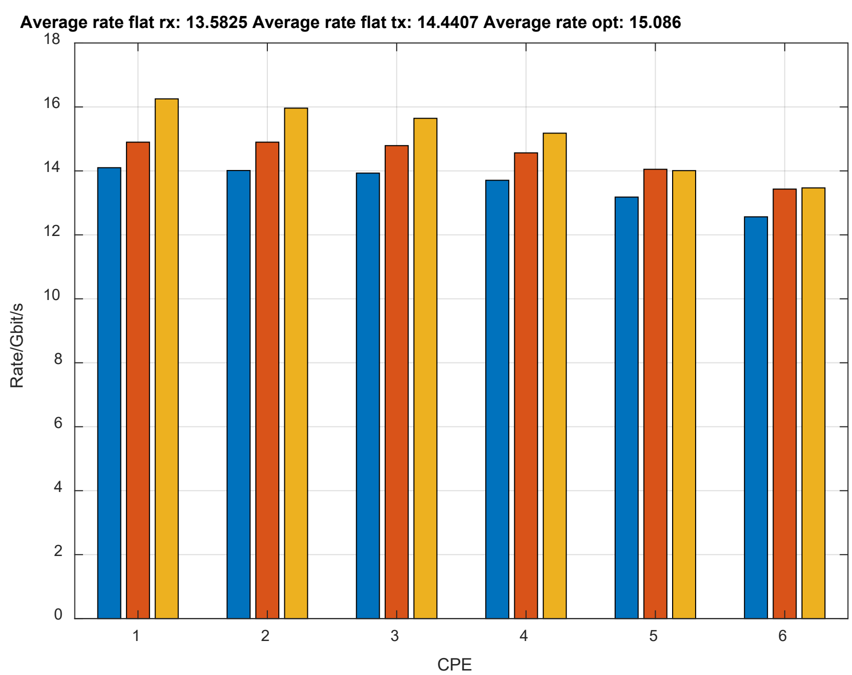


Figure 14 – Node + 4 mid-split Data rates for tilted TX- flat RX PSD (blue), flat TX PSD (red), and optimized PSD (yellow)

Figure 15 shows transmit power spectrum for the 3 power allocation schemes (left) and TCP when optimizing throughput for a certain tap (right). In case of tilted TX PSD and flat TX PSD, the transmit power spectrum and TCP don't depend on the tap. When optimizing throughput for long taps (5, 6), the optimal spectrum shape is flat, while optimizing for short taps (1) gives a tilt for all frequencies as an optimal shape. When optimizing for the medium taps (2-4), the optimal spectrum shape has a tilt at low frequencies followed by flat region at high frequencies. The tilt at low frequencies indicates that the performance at these frequencies are nonlinear distortion dominated. On the other hand, at high frequencies the dominance of thermal noise (i.e. signal is more attenuated and hence closer to thermal noise floor) makes the flat spectrum optimal (classic water-filling comes into play). TCP graphs shows the TCP at node when optimized for tap number given in x-axis.

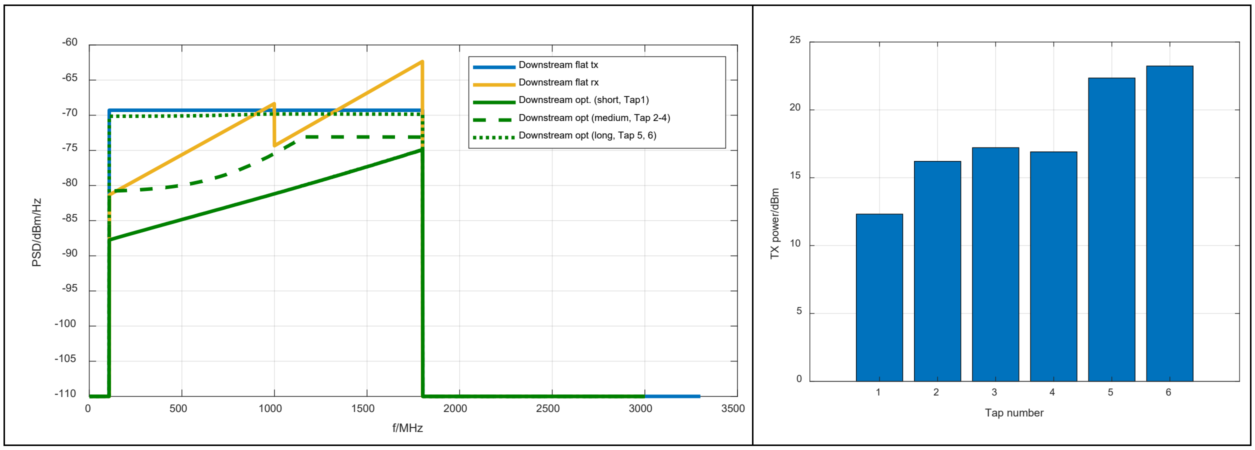


Figure 15 – Node + 4 mid-split Transmit PSDs and TCPs for different allocation schemes

TCP requirement is dominated by the far away CMs, requiring the full 71 dBmV (22.25 dBm) of power.

5.1.3. Node + 0 Network, High-Split

We've repeated the test for high split note to get the following rate and TCP results. TCP requirement is dominated by far away CMs demands.

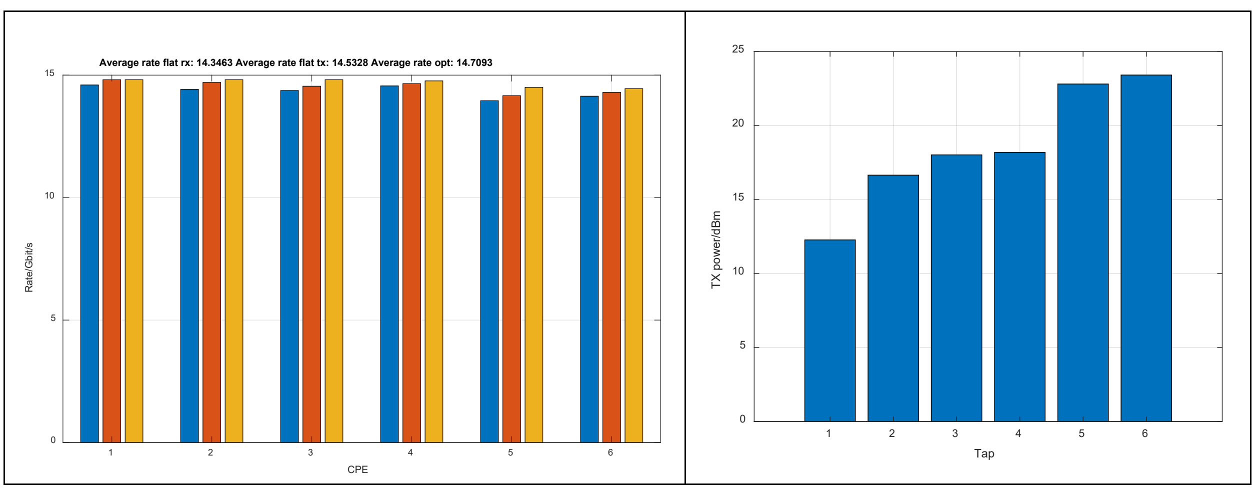


Figure 16 – Node + 0 high-split Data rates for flat RX (blue), flat TX (red), and optimized (yellow) and corresponding TCP

Optimal PSDs are shown in Figure 17. Although we can maintain a uniform tilt in spectrum for the close in and medium range CMs, long range CMs forces a flat spectrum at high frequencies.

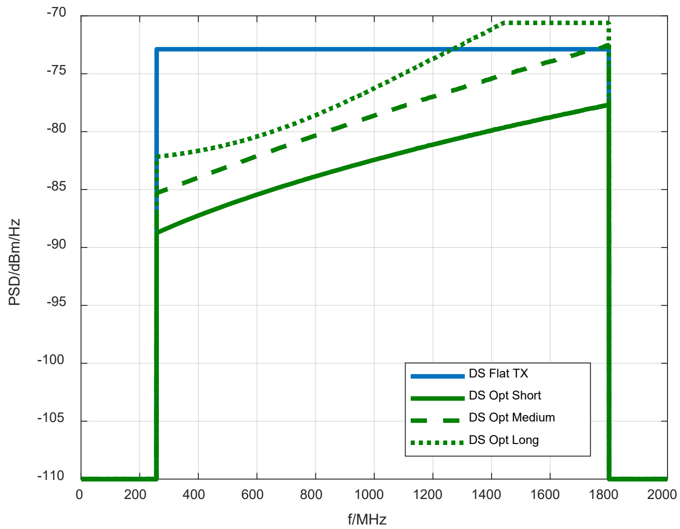


Figure 17 – Node + 0 high-split Transmit PSDs for different allocation schemes

5.1.4. Node + 4 Network High-Split

Up to 10% rate improvement achieved with optimal power allocation as show in Figure 18 (left).

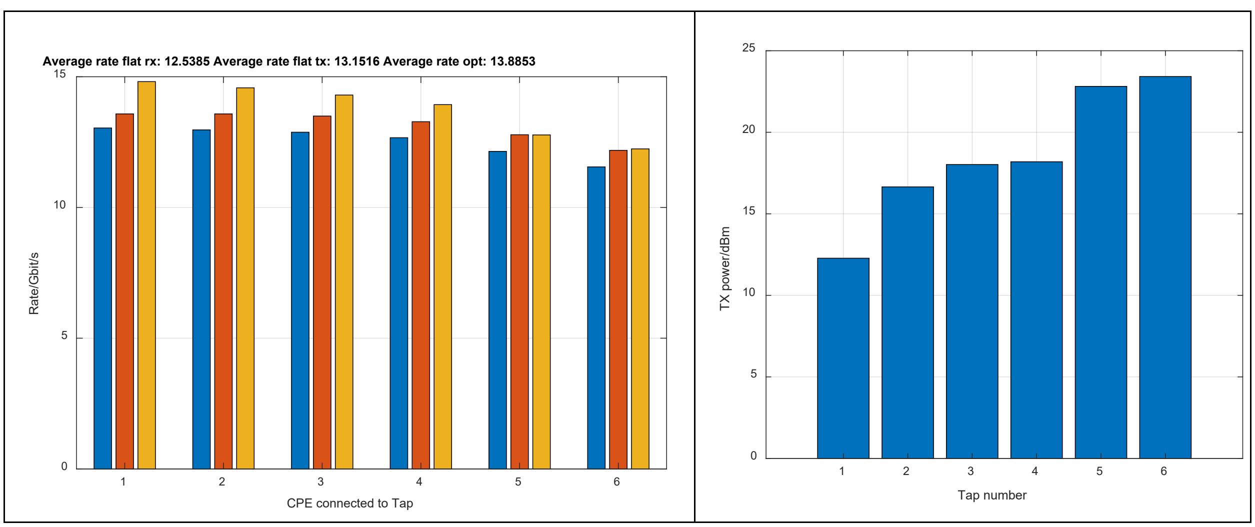


Figure 18 – Node + 4 high-split Data rates for flat RX (blue), flat TX (red) and optimized (yellow) and the corresponding TCP

Power allocation across frequency for the 3 allocation schemes is shown in Figure 19.

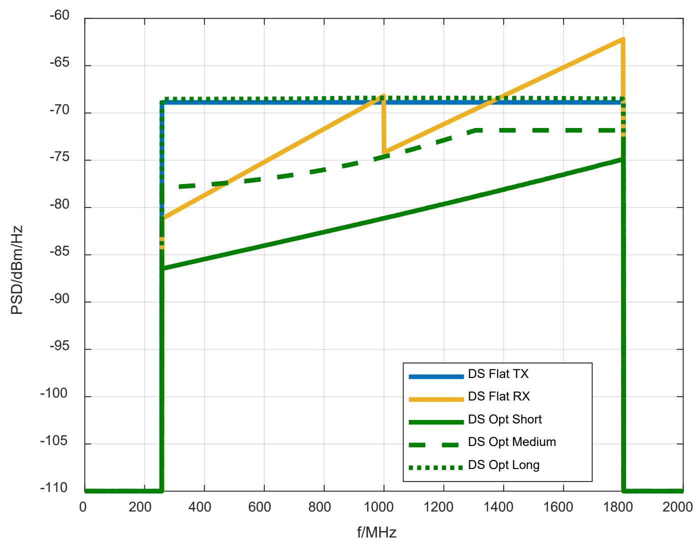


Figure 19 – Node + 4 high-split Transmit PSDs for different allocation schemes

5.2. With Staggered Channel Allocation – Channel Stacking

In these tests, we exploit the property of the cable channel that it has higher losses at higher frequencies. To minimize the impact of higher losses in higher frequencies, we allocate these higher frequency channels to close in CMs and lower frequency channels to far away CMs. Consider to hypothetical channel allocation (non-DOCSIS) shown in Figure 20. The channel estimate for each allocated channel is used in the optimization algorithm.

Modem / Channel	1-4 (Tap 1)	5-8 (Tap 2)	9-12 (Tap 3)	13-16 (Tap 4)	17-20 (Tap 5)	21-24 (Tap 6)
1						
2						
3						
4						
5						
6						
7						
8						
9						

Figure 20 – Channel to CM allocation for PSD optmization

Figure 21 shows PSD for the above channel allocation scheme. A significant reduction is TCP (up to 6 dB) is achieved through this channel allocation scheme without losing any throughput.

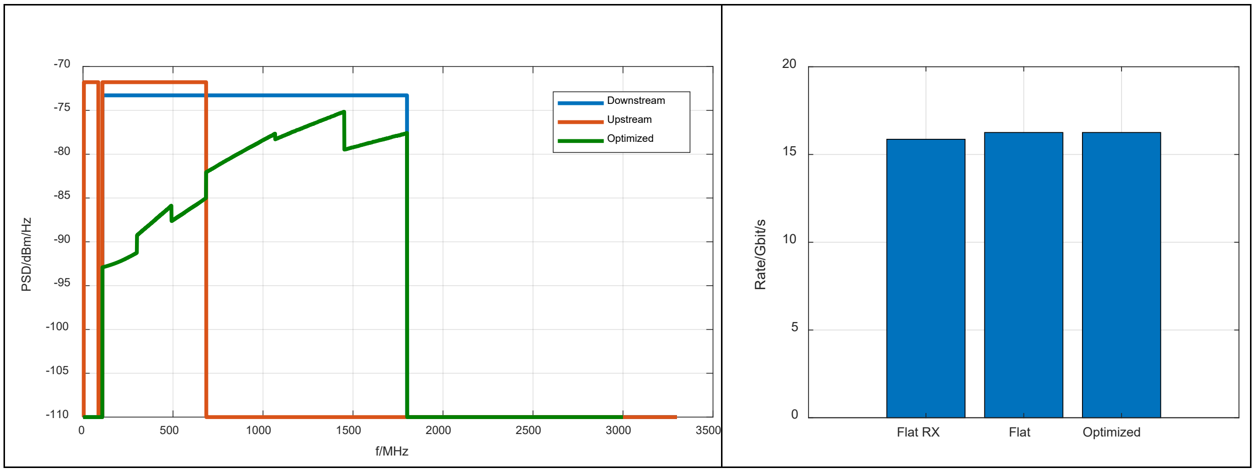


Figure 21 – Power allocation vs Data rates

However, this is not quite a practical allocation scheme. In practice we would allocate set of channels to a group of CMs and take advantage of statistical multiplexing to improve the utility of the spectrum.

Figure 22 shows a practical channel allocation scheme where channels allocated to long range CMs are indicated in green and channels allocated to short range CMs are indicate in yellow. Within each channel set, the CM corresponding to the channel estimate used in the optimization algorithm is indicated in dashed lines.

Modem Channel	1-4 (Tap 1)	5-8 (Tap 2)	9-12 (Tap 3)	13-16 (Tap 4)	17-20 (Tap 5)	21-24 (Tap 6)
1						
2						
3						
4						
5						
6						
7						
8						
9						

Figure 22 – Channel stacking for two service groups

Note that each CM is allocated 5 OFDM channels, which will allow us to maintain 10 Gbit/s for the service group (SG).

SG A:
- CMs in Taps 1 to 3.
- Allocated channels 5 to 9

SG B:
- CMs in Taps 4 to 6
- Allocated channels 1 to 5

This channel allocation enables what we call 10G DOCSIS and it will appear like a node-split in the sense that we are supporting two 10G SGs using available spectrum.

By optimizing the power allocation for each channel to match at least one of the taps in each SG, we are optimizing the average throughput of each SG (assuming each user has roughly the same probability of using each channel).

The alternative scheme shown in Figure 23 uses the same channel allocation, but the channel frequency responses used in the optimization algorithm is taken from the medium range CM within each SG. This is the default channel allocation method used in following tests unless mentioned otherwise.

Modem / Channel	1-4 (Tap 1)	5-8 (Tap 2)	9-12 (Tap 3)	13-16 (Tap 4)	17-20 (Tap 5)	21-24 (Tap 6)
1						
2						
3						
4						
5						
6						
7						
8						
9						

Figure 23 – Default channel stacking scheme for two service groups

In practice, with additional information, we could make more data driven approach to allocating channels and deciding which tap/CM channel estimate we use for optimization.

5.2.1. Node + 0 Network Mid-Split

Figure 24 shows the optimal PSD and resulting data rates for the channel allocation scheme given in Figure 23 for Node + 0 Network.

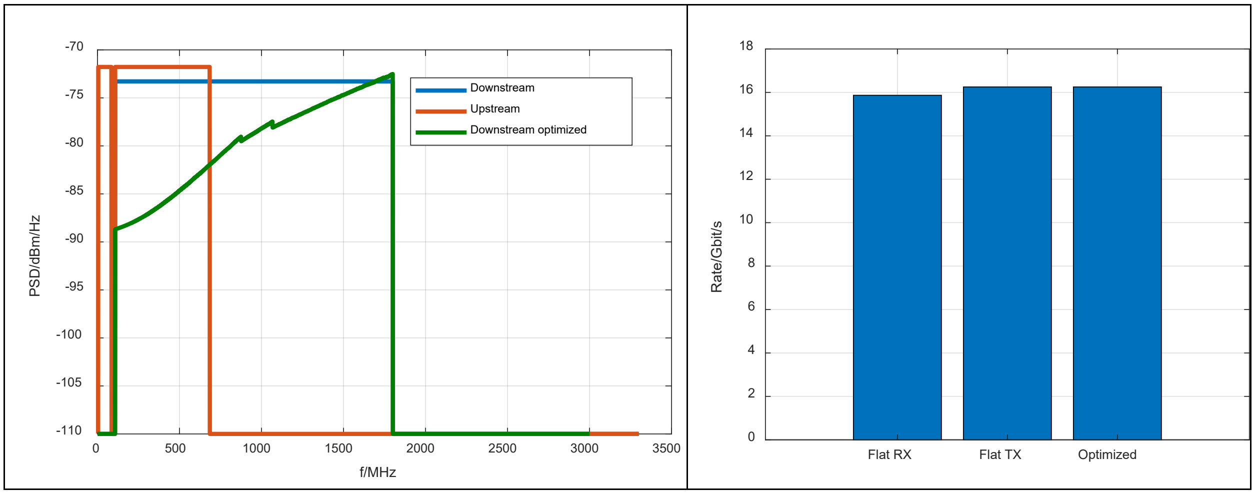


Figure 24 – Node + 0 mid-split Power allocation and Rate for channel stacking

Note that the TCP has reduced by nearly 6 dB compared to assigning all the channels to each CM. Furthermore, the overall rate is higher compared to allocating all channels to each CM.

5.2.2. Node + 4 Network Mid-Split

Channel allocation based power optimization results for Node + 4 mid-split case is shown in Figure 25. As in Node + 0 case, we are making significant savings in TCP. In addition, the optimal power allocation gives 10-15% increase in the throughput compared to non-optimal power allocation schemes.

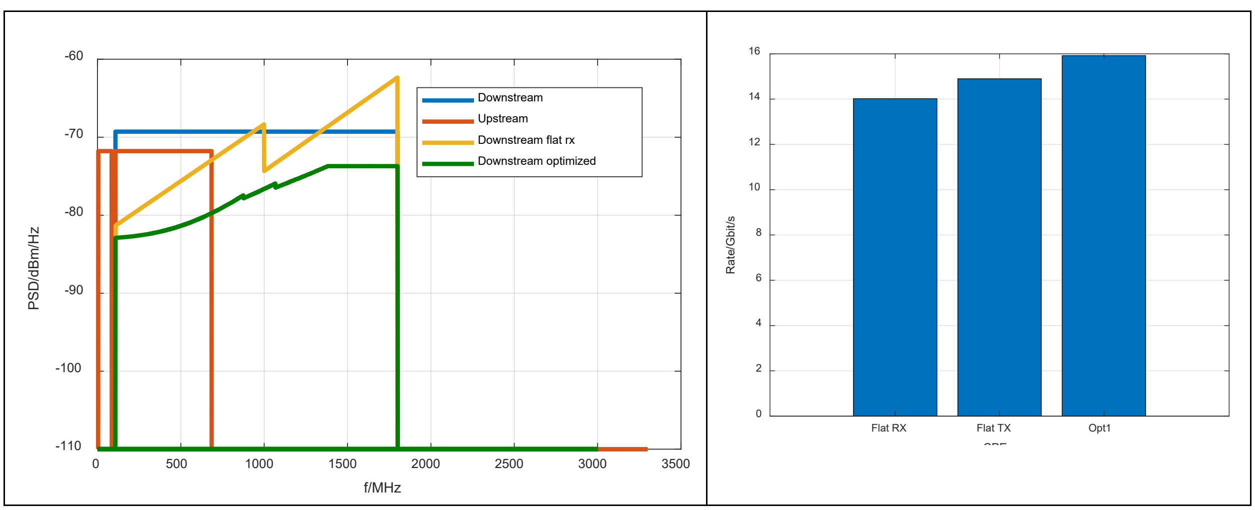


Figure 25 – Node + 4 mid-split power allocation and Rate for channel stacking

5.2.3. Node + 0 Network High-Split

Channel allocation based power optimization results for Node + 0 high-split case is shown in Figure 26. As with Node + 0 mid-split cases, we are making significant savings in TCP while not losing any throughput.

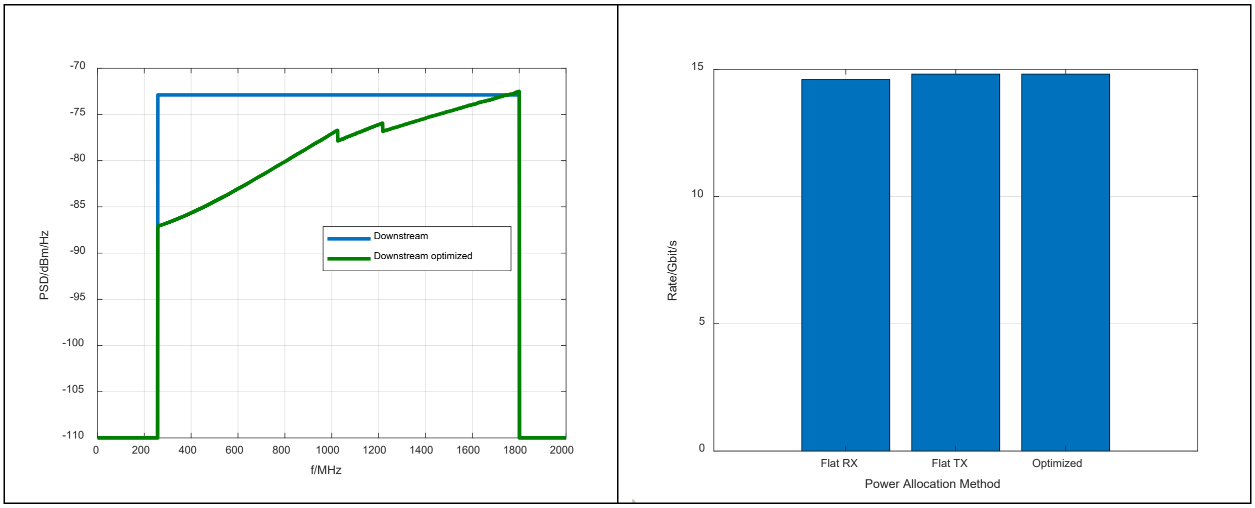


Figure 26 – Node + 0 high-split power allocation and Rate for channel stacking

The optimized TX TCP is 6 dB below the TX power used for flat TX and flat RX.

5.2.4. Node + 4 Network High-Split

Channel allocation based power optimization results for Node + 4 high-split case is shown in Figure 27. As in Node + 4 mid-split case, we are making significant savings in TCP, and at the same time improving the throughput by 10-15%.

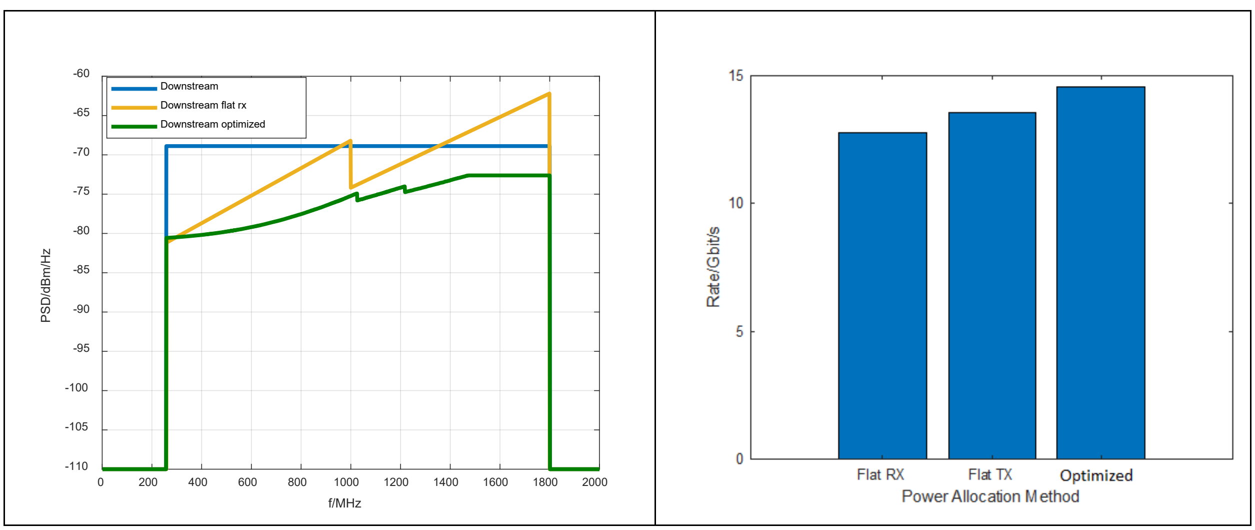


Figure 27 – Node + 4 high-split power allocation and Rate for channel stacking

6. Conclusion and Future Work

This paper shows that there are significant benefits to be gained from careful allocation of channels to the CMs and optimization of available limited transmit power in a ESD system.

For a given channel allocation, considerable throughput gain, in order of 10%, is achievable with closed-loop power optimization. Furthermore, carefully combining the channel allocation and optimal power distribution can significantly reduce the required TCP for the node and the network amps.

The additional headroom created in transmit power budget can be exercised to improve the range of the network. On the other hand, savings made to overall node power consumption budget can enable more flexible fiber deep deployment options, such as soft FMA and edge compute.

The above benefits are applicable to network topologies across the board, be it passive Node + 0, fiber deep, or conventional Node + X.

Further work is needed to quantify the effect of relaxing some of the network design principles, such as unity gain.

Abbreviations

C/CSO	carrier to composite second order distortion
C/CTB	carrier to composite triple beat
CM	cable modem
DAA	distributed access architecture
DS	downstream
ESD	extended spectrum DOCSIS
FDD	frequency division duplexing
FMA	flexible MAC architecture
HFC	hybrid fiber-coax
MER	modulation error ratio
MTPR	missing tone power ratio
NF	noise figure
OFDM	orthogonal frequency division multiplexing
PA	power amplifier or power amplification
PMA	profile management application
PNM	proactive network maintenance
PON	passive optical network
PSD	power spectral density
RMD	remote MAC device
RPD	remote PHY device
ISBE	International Society of Broadband Experts
SCTE	Society of Cable Telecommunications Engineers
TCP	total composite power
US	upstream
vCMTS	virtual Cable Modem Termination System

Bibliography & References

[1] IEEE, "IEEE P802.3ca/D1.4 Draft Standard for Ethernet Amendment: Physical Layer Specifications and Management Parameters for 25 Gb/s and 50 Gb/s Passive Optical Networks," 2018, IEEE 802.3ca/D1.4, 27 November 2018.

[2] Wang, Eric, "Draft text for G.mgfast-PHY," 209, ITU Draft Recommendation Q4/15-TD48 (190121).

[3] CM-SP-PHYv3.1, "Data-Over-Cable Service Interface Specification DOCSIS 3.1, Physical Layer Specification," CableLabs, 2013

[4] CM-SP-PHYv4.0, "Data-Over-Cable Service Interface Specification DOCSIS 4.0, Physical Layer Specification," CableLabs, 2020.

[5] Ed Harstead, Doutje van Veen, Vincent Houtsma, and Pascal Dom, "Technology Roadmap for Time-Division Multiplexed Passive Optical Networks (TDM PONs)," Journal of Lightwave Technology, vol. 37, no. 2, pp. 657–664, 2019.

[6] Ed Harstead and Randy Sharpe, "Forecasting of access network bandwidth demands for aggregated subscribers using monte carlo methods," IEEE Communications Magazine, vol. 53, no. 3, pp. 199–207, 2015.

[7] Jakob Nielsen, "Nielsen's Law of Internet Bandwidth," 2018.

[8] Cloonan, Tom and Al-Banna Ayham and O'Keeffe Frank, "Using DOCSIS to Meet the Larger BW Demand of the 2020 Decade and Beyond," 2016, Arris White Paper.

[9] Partha P Mitra and Jason B Stark, "Nonlinear limits to the information capacity of optical fibre communications," Nature, vol. 411, no. 6841, pp. 1027, 2001.

[10] NetCommWireless, "CTTdp Unit (4 ports)," July 2016, Specification Sheet NDD-4200 R1.

[11] Krapp, Steven, "Virtual Fiber - 100 Gbps over Coax," in SCTE 2017 Fall Technical Forum. IEEE, 2017.

[12] J.M. Cioffi, "A multicarrier primer," ANSI T1E1, vol. 4, pp. 91–157, 1991.

[13] Thomas M Cover and Joy A Thomas, Elements of information theory, John Wiley & Sons, 2012.

[14] Rainer Strobel, Channel Modeling and Physical Layer Optimization in Copper Line Networks, Springer, 2019.

[15] Danny Van Bruyssel, Yannick Lefevre, Vladmir Oksman, Rainer Strobel, Miguel Peeters, and Dong Wei, "G.mgfast: Working text for advanced coding scheme," January 2019, Contribution ITU-T Q4/16-C17-R1 (190121).

[16] Rob Howald, "Evaluation of In-Home Architecture variants for FDX and Rationale," April 2017, Comcast FDX f2f Meeting.

[17] Rainer Strobel and Thushara Hewavithana, "Power Spectrum Optimization for Capacity of the Extended Spectrum Hybrid Fiber Coax Network", IEEE ICASSP, 2020.

