Breathing New Lifespan into HFC: Tools, Techniques, and Optimizations

Dr. Robert L. Howald, Fellow Technical Staff, ARRIS

ARRIS

Contents

Introduction

Cable operators have seen downstream bandwidth grow at 50% per year (CAGR) for many consecutive years. The trend, often referred to as Nielsen's Law, has held firm for the 20+ years and will be assumed to be a relevant guideline for assessing the future, along with variants we shall discuss. There are reasonable arguments for long-term limits of media consumption [2,7] that we will consider, although predicting applications has been difficult, and services not yet foreseen may keep the trend alive beyond media consumption.

Cable operators manage this persistent growth under the spectrum constraints of their current legacy service offerings, mostly video, which consume the vast majority of the total available spectrum. Tools for improved bandwidth efficiency are used to balance the growth of legacy services such as HD and VOD as data traffic is increased. Tools and strategies are outlined in [4,10].

Recently, the industry initiated the DOCSIS 3.1 effort, which has an objective to achieve at least 10 Gbps of downstream and 1 Gbps of upstream. This is another major tool for enabling this continued growth, and places cable on par with PON targets, while network migration steps can deliver similar average user capacity.

In this paper, we will take a look at the service growth challenge with an analysis tool concept designed to quantify the problem, introduce and describe in detail the architecture and technology evolutions in play to handle projected requirements, and then revisit our analysis to assess what these can accomplish against this growth.

The Capacity Management Timeline

A sample analysis representative of the issue facing MSOs can be charted on a Capacity Management Timeline as shown in Figure 1.

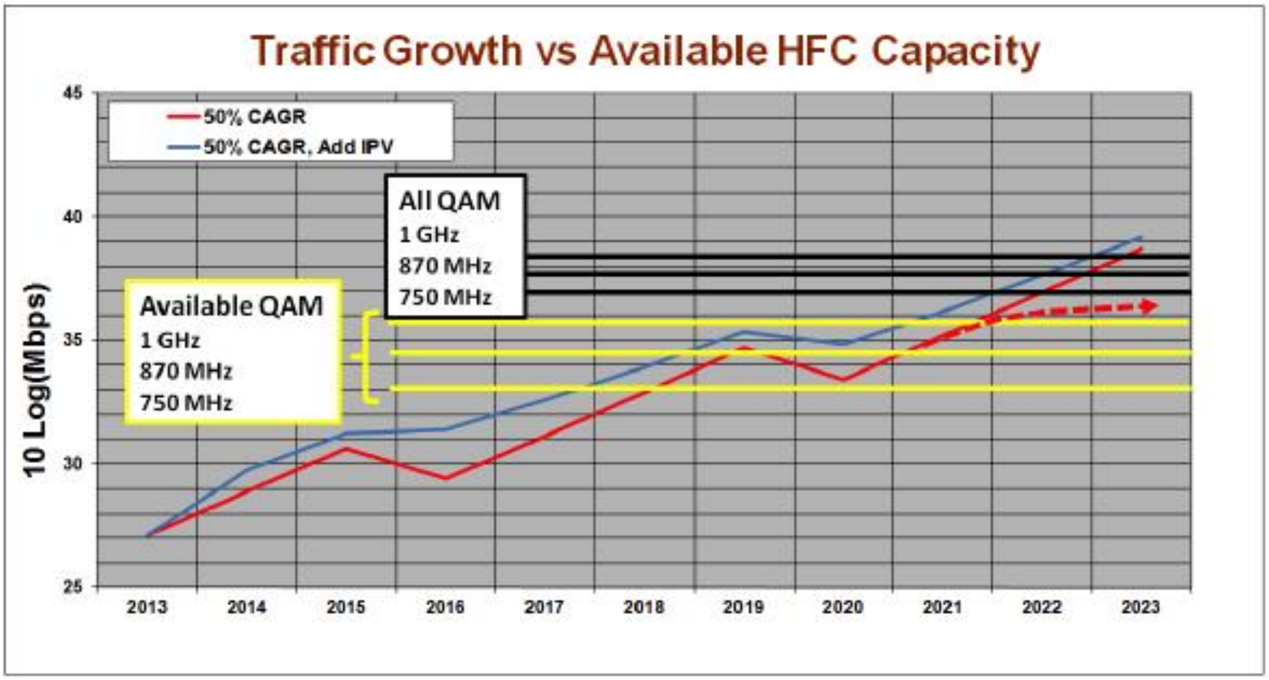


Figure 1: A Capacity Management Timeline Guides Service and Architecture Evolution

Figure 1 shows various threshold lines drawn that represent the point at which capacity of that particular configuration quantified by the threshold line is exhausted. The purpose of this paper is to look at the technology and techniques available that move such thresholds higher to allow more growth, and consider elements that are favorable from a lifespan point of view that affect the trajectories themselves.

So that we can fully appreciate the information in Figure 1 for later use, we will briefly detail the concept of the Capacity Management Timeline. This visual analysis approach allows operators to understand the timing implications of Compound Annual Growth Rate (CAGR) and service evolution. Understanding what it portrays is necessary to make a comparison of the before-and-after of the topics discussed throughout the paper.

The Intersection of Traffic, Services and Architecture

The growth of IP data (DOCSIS) is shown by the red and blue trajectories trending upward with a slope that represents the 50% CAGR. These upward bound trajectories are broken at particular years that represent service group splits (node splits). The blue trajectory has an underlying 50% CAGR, but also includes the introduction of new DOCSIS channels specifically set aside for IP Video (IPV).

Various thresholds are drawn horizontally representing capacity limitations set by the entire forward band using 256-QAM (in black), and the same spectrum examples but offset by channel slots “not available” for growth (yellow, “Available”). In this case, we limited this to 69 slots that were unavailable today. This was based on 60 analog carriers and 9 additional to account for an 85 MHz mid-split for an assumed upstream expansion that takes place over the ten-year window. Legacy digital services obviously coexist, but the idea here was to capture the offset from the all-digital cases (“black”) and the power of the analog reclamation step for comparison among capacity management tools.

Of course, any combination of legacy services that add up to 69 channels consumed would yield the same answer. For example, an all-digital downstream broadcasting 200 SD and 100 HD channels would consume about 60 slots. This is just one example – any combination of services can be analyzed and many have been, such as in [4,10]. Many specific customer examples have also been analyzed in this fashion, and contrary to what the yellow thresholds might indicate, operators generally do not have *any* room for DOCSIS growth. Actual thresholds are right on top of the current state of DOCSIS consumption. However, this discussion is about new capacity methods more so than bandwidth management [4,10]. We will focus this discussion more on how far “North” we can move capacity thresholds on Figure 1.

Let’s take a snapshot of the “today” state from Figure 1 so we can assess the gains we make with the various next generation tools. The aggressive growth of traffic versus time when evaluating against the spectrum constraints looks threatening for HFC’s sustainability. The vertical axis is a logarithmic representation to effectively capture compounding growth. Thus, 30 dB represents 1 Gbps and 40 dB represents 10 Gbps. Whenever a trajectory crosses a threshold, that threshold has run out of capacity. For the two cases shown here, the best case scenario with two splits (timed differently than shown for some spectrum cases) manages through 6-8.5 years of IP data growth, without deploying other tools to manage spectrum.

Analog reclamation, Switched Digital Video (SDV), more efficient video encoding, and IP video are all potential tools to help manage the available capacity for growth. The customized use of the Capacity Management Timeline is precisely for this purpose – based on an individual operator’s legacy services, growth expectations (data and video), and architecture variables, it is possible to chart out a migration path that allows operators to project their investment needs and timing.

IP Video Transition

There are two growth trajectories on the curve, and these represent a couple of ways to think about and quantify the transition to IP Video. First, note that IP Video will initially be a simulcast, and remain so for many years. Legacy services will co-exist as the video line-up transitions to availability over the IP network. This creates the so-called "simulcast bandwidth bubble," whereby the end state of bandwidth consumption may have an excellent outlook, but the path to getting there is limited by effectively redundant programming.

The two trajectories represent these two views:

(1) IP growth at a CAGR of 50% continues to occur, and then on top of that we must add more DOCSIS channels for the IP Video service

(2) IP growth at 50% CAGR has been driven by streaming video services like Netflix for the past several years (not conjecture). 50% CAGR continues because content that used to be elsewhere now joins the IP realm. In this view 50% captures the IP video transition already – this is just the new CAGR growth "engine."

After enough years, as can be seen, the difference becomes very small because the spectrum size needed for IPV is fixed and eventually is overwhelmed by persistent, aggressive CAGR.

The number of IP Video channels required can be determined by analyzing the serving group size, programming line-up, and encoded video bit rates, and understanding the use dynamics of primary screen, secondary screen, and VOD viewing. Also key is a statistical understanding of viewership learned from years of IPTV and SDV deployments. An analysis tool has been developed that does this calculation, and which is publicly available at www.motorola.com/multicast-unicast-calculator/.

A sample case was run with a large SD and HD programming line-up and high penetration of DOCSIS service (70%). The output is shown in Table 1 below. After two splits, about 20 DOCSIS channels are required to meet the IP Video needs (or at one split and 50% penetration early). This is what is added to the 50% CAGR for the blue trajectory, and it was added as 4+8+8 channels over a period of 7 years.

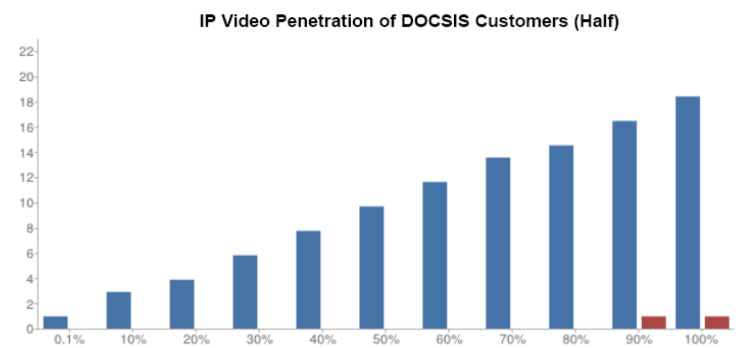
Growth Contraction?

There is one other line of thought regarding the 50% CAGR growth; which recognizes that growth is being driven by streaming video. This line of thinking is that video quality only increases to a point at which there is no value to improving it [2,7] from a human perception standpoint. It is not completely settled science when that is, but pretty settled that it is finite. The notion that an asymptote exists out in the future associated with video/data consumption (only) is shown by the dashed red line beginning in the year 2021 in Figure 1.

Total DOCSIS Capacity Required to Deliver Managed IP Video Services

Total Bandwidth Requirements - Case 2

Service Group Size HHP:	125	Peak IP Video Devices / Home:	2.5 Active
Max Video Penetration:	70%	At Peak Penetration:	87.5 Homes
Percentage of Video Devices Active at Peak:	70%	At Peak Viewership:	153 Streams



IP Video Penetration of DOCSIS Customers Charts

Total Downstream Channels @ 100% Unicast

Potential Channel Savings with Multicast

Table 1: Calculating How Many DOCSIS 3.0 Video Channels

There are three principles to this perspective:

(1) Assuming media consumption driven bandwidth, we can quantify maximum video quality bit rates that have service value.

(2) Recognition that humans have a limited ability to multi-task, in particular with video. While simultaneous secondary screens during a primary viewing may be common, humans have limited ability to focus on multiple things at once with comprehension.

(3) Use of IP devices/home and tied to residential demographics which are generally available statistics.

We can also reason that the CAGR engine has been steady for 20 years simply to keep up with increasingly higher levels of human media experiences:

(1) Alphanumeric characters

(2) Voice

(3) Images (pictures)

(4) Music

(5) Low speed video

(6) SD Video

(7) HD 1.0

The suggestion here is that perhaps the speeds supportive of the best video quality likely to be practical represent a logical tapering point of CAGR for media consumption as we can fathom it today. There are obviously long-term benefits to HFC networks and migration planning if this does come to pass, as can be concluded by evaluating the implications of the red arrow in Figure 1. We will revisit the implications of this traffic growth philosophy after evaluating our lifespan growth possibilities enabled by new capacity.

Having set the stage for the evaluation of lifespan objectives, let's now look at the component parts designed to paint a prettier picture for that objective, how they do so, and how much they offer.

Capacity Optimization

Theoretical capacity is based on two variables – bandwidth (spectrum allocated) and the Signal-to-Noise Ratio (SNR). Shannon Capacity is the well-known limit, and represents the maximum error-free rate that can be achieved in additive white Gaussian noise (AWGN). It is given very simply as

$C = [B] \text{ Log}_2 \, [1+\text{SNR (dB)}]$ (1)

This can actually be even further simplified for cable networks, in particular for the downstream, relying on high SNR assumptions. If the SNR is high it can be shown that capacity is essentially directly proportional to bandwidth, B and SNR *expressed in decibels* (dB):

$C \approx [B]\,[SNR\ (dB)] / 3$ (2)

This simplification of Shannon Capacity is accurate asymptotically within less than 0.5% with increasing SNR above 15 dB.

Clearly according to (2), more capacity is available with higher SNR, but with logarithmic proportionality. For example, 50% more spectrum yields 50% more capacity, but so does 50% more SNR. However, turning a 30 dB SNR into a 45 dB SNR is a significant network performance leap. Nonetheless, it is certainly the case that more SNR means more capacity, and architectures that create higher SNR – deeper fiber, digital optics, home gateways – create potential capacity.

Shannon Capacity is a theoretical concept, and Shannon does not describe either waveform types or codes to use in his famous treatise. For real systems, of course, we deal in signal waveforms and modulation formats to exploit the spectrum. Through this, SNR has two key practical components:

(1) Improving the link SNR itself, which translates to modulation formats. The link has many contributing noise dependencies – architectural, technology in the optical and RF network, and equipment fidelity and CPE technology itself. The relationship of evolution variables to net SNR impact is a comprehensive accounting of these pieces.

(2) Forward Error Correction's (FEC) role in capacity is played out through "SNR" in equations (1) and (2). The best codes enable a given M-QAM format and level of bandwidth efficiency at a lower SNR. Or, alternatively, for a given SNR, the best codes enable the highest order M-QAM formats of the most bandwidth efficiency.

The next section takes a look at the foundational elements of maximizing capacity – optimally exploiting the channel using modern physical layer technology tools.

Adding to the Physical Layer Toolkit

M-QAM Formats

Today's cable systems implement a maximum M-QAM format of 256-QAM (8 bps/Hz) downstream and 64-QAM (6 bps/Hz) upstream. These represent upgrades in efficiency from prior use of 64-QAM for digital TV downstream and 16-QAM upstream. Through architecture evolutions (deeper fiber) and technology improvements (optical & RF fidelity, DFB return lasers) cable has already gone through at least one major round of improving bandwidth efficiency, and most of it many years ago. Plenty of years have passed since a major technology refresh can pay dividends.

Figure 2 shows the current modulation profiles and a couple more that are anticipated as certainties in next generation systems – 1024-QAM and 4096-QAM. "In-between" profiles (512-QAM and 2048-QAM, not shown) are assumed eligible candidates as well. In the figure, all of the M-QAM formats are shown for an *equivalent uncoded BER* of 1e-8. Since they are 6 dB apart for each step up in density, the SNRs are therefore 28 dB, 34 dB, 40 dB, and 46 dB, for 64-QAM, 256-QAM, 1024-QAM, and 4096-QAM, respectively. At the very least, the latter (46 dB) should give pause to the thought of supporting that M-QAM capability over HFC.

Higher order formats can be constructed and, as we shall see, may be worth considering, but are not shown. They do not exist in simulation tools at this point!

A common end-of-line HFC cascade performance requirement for digital channels is a 42 dB SNR with digital channels typically set 6 dB below analog channels. Given that 256-QAM requires 34 dB (1e-8) without coding, and up to 4 dB less than this by DOCSIS specification with a J.83 Annex B error mitigation subsystem included, it is apparent why today's networks are very successful with 256-QAM. In fact, some are likely able to support 1024-QAM robustly using similar "J.83B" tools. Some lab evaluations have indicated this is likely to be the case [9].

However, even just considering HFC SNRs, the 1e-8 SNRs required of 2048-QAM (43 dB) and 4096-QAM (46 dB) clearly indicate extra "help" is necessary to achieve these with robustness. It can come in the form of FEC, architecture modifications, technology improvements, or all of the above, as long as we can find the dBs necessary to close the link.

A View from the Field

*F*igure 3 shows some extremely valuable pioneering work done by a major North American MSO – a first of its kind that indicates with a large statistical sample what Cable Modems are telling us their channel SNR looks like [15]. Other MSOs are now gathering such statistics as well to help the industry engage in proper technology choices based on real data.

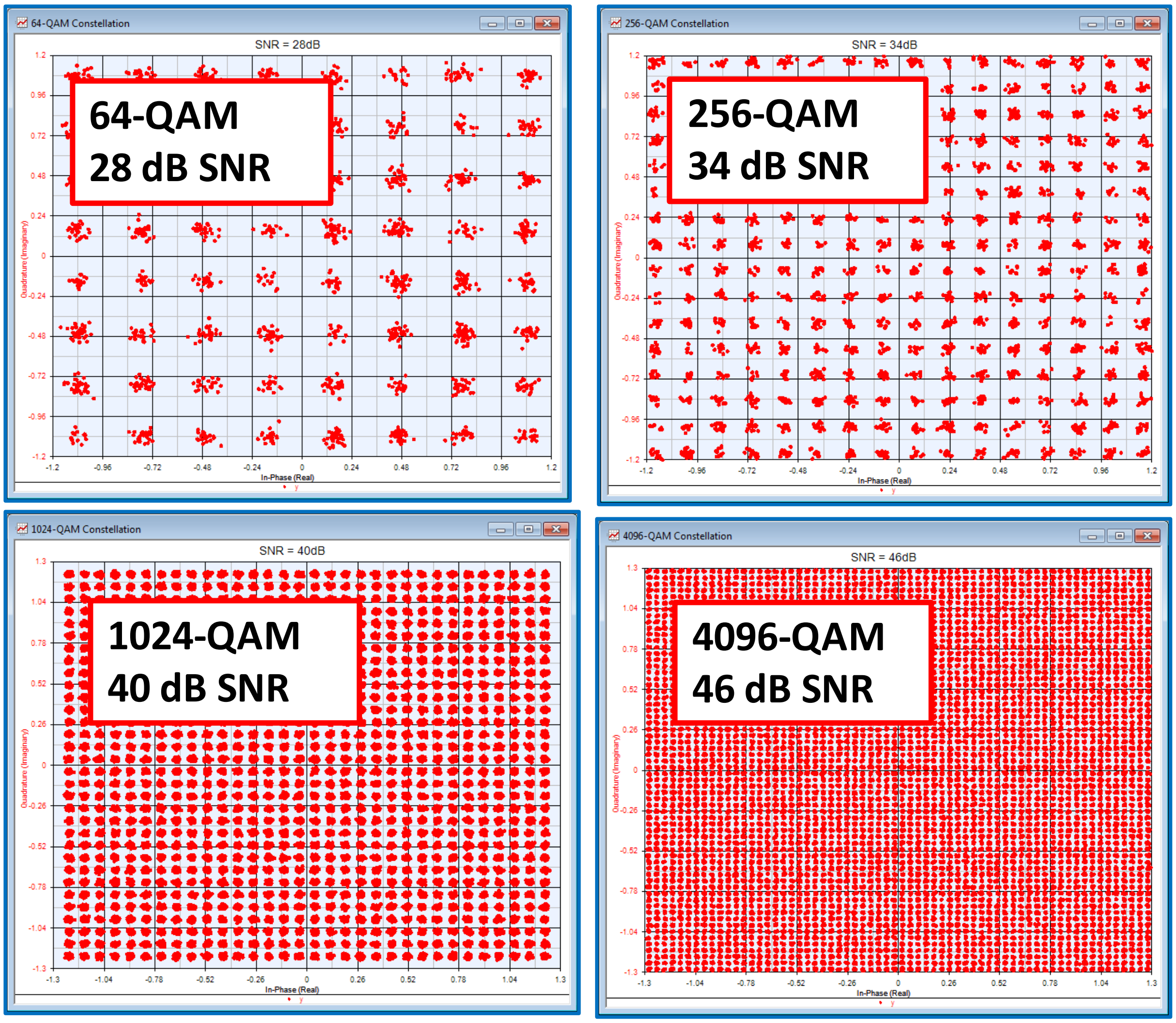


Figure 2: Increasingly Spectrally Efficient M-QAM Format

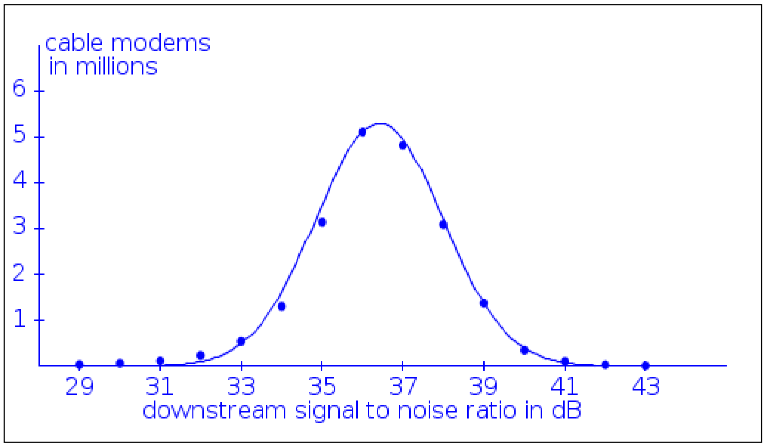


Figure 3: Major MSO Cable Modem SNR Distribution [15]

There are important differences between CM reported SNR and HFC delivered SNR, as we can easily determine by the delta between the HFC delivered 42 dB number (or better) and the SNR scale in Figure 3. The most important ones are:

(1) The CM actually measures and reports MER, which includes all impairments on the channel, all the way to the CM demodulator. Thus, it includes the CM contribution itself.

(2) The CM's contribution is strongly dependent on the location of the CPE in the home. It is a dominant noise contributor at low CM input levels.

(3) The CM was implemented for high performance of 256-QAM, which is 12 dB less sensitive than 4096-QAM.

(4) The maximum measurement fidelity itself of MER is likely in the low-to-mid-40s.

Figure 3 will prove valuable in defining QAM formats and techniques to optimize their use. While the absolute SNR numbers may be biased towards lower values relative to a new generation of technology and architecture evolution, the spread of the distribution is illustrative of the variation across the network that can be better exploited for capacity management.

The Magic of FEC

Advances in FEC have straightforward PHY design effects – better FEC reduces the SNR required to achieve a particular QAM format, increasing bandwidth efficiency and throughput for a given link performance. Today's go-to code family is Low Density Parity Check Codes (LDPC). LDPC codes have been mathematically around for many years. However, as has been the case with other codes (e.g. Reed-Solomon), they have came into vogue as the speed of computation has become sufficient to enable real-time operation of these extremely resource-intensive large block size codes. The first standard to define an LDPC code was DVB-S2 in the early 2000s, but since that time codes from the LDPC family have become part of G.hn, MoCA™, WiMax, Wi-Fi, and DVB-C2, among others. The reason is simple – they get closest to the Shannon bound, maximizing capacity, and efficient ways to implement them cost-effectively are now available.

In Figure 4, we show the DVB-C2 family of LDPC codes [14] and the M-QAM potential available, including 64-QAM through 4096-QAM. Observe the SNR requirements enabled by LDPC under the "Highest Code Rates" label in Figure 4 (90%). These are the nearest apples-to-apples comparisons to the error correction scheme used by J.83B downstream today.

The true power of LDPC can be seen in the SNRs required to deliver vanishingly low error rates in Figure 4 and Table 2. Table 2 summarizes the SNR gains available for the QAM profiles compared to the uncoded case [6]. The FEC, of course, comes with a 10% efficiency penalty (for the 90% code rate). However, 10% efficiency hit for *9-11 dB* of SNR gain is a powerful trade-off – *one-tenth* the SNR tolerated for this small loss of efficiency. The 46 dB of uncoded 4096-QAM SNR previously mentioned, for example, reduces to 35 dB as shown in Figure 4 – pretty impressive! The 9-11 dB range of SNR advantage in Table 2 is a testament to the power of LDPC codes. We will compare this to today's downstream FEC in the next section.

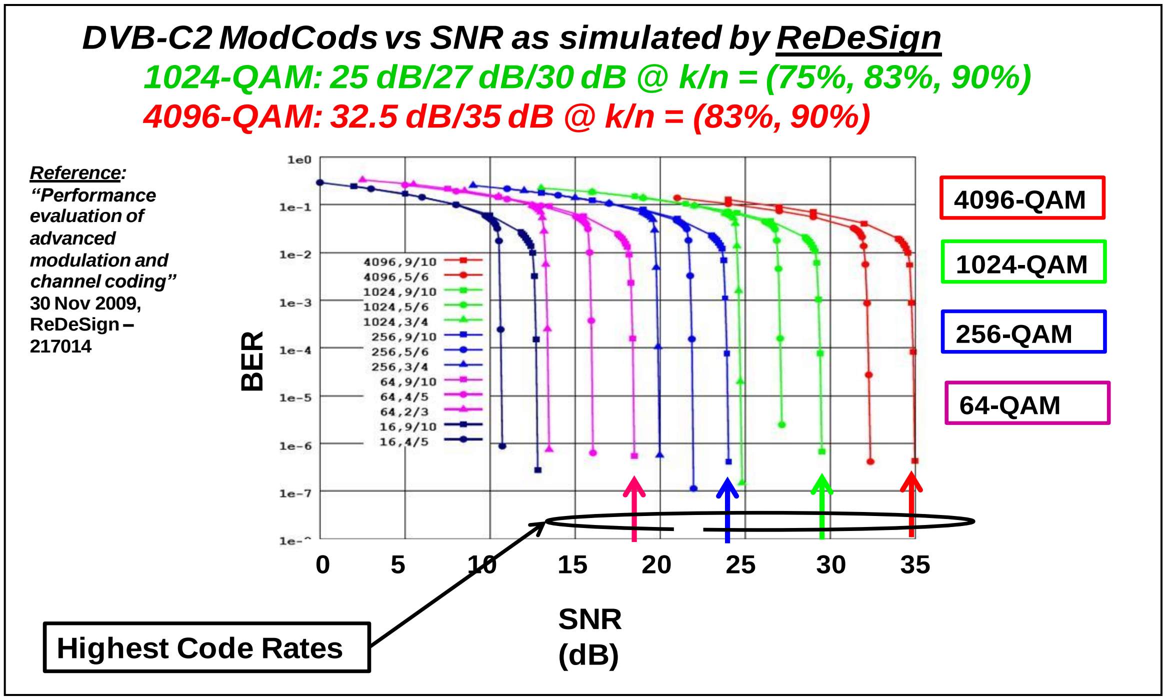


Figure 4: Bandwidth Efficient M-QAM Enabled by LDP

As impressive as Table 2 may look, M-QAM constellation pictures truly put the role of FEC into perspective. To emphasize this "magic," we show the constellations of 1024-QAM and 4096-QAM in Figure 5. The SNRs shown are 3-4 dB *higher* than the SNR threshold for low error rate (error free) performance in Figure 4. Figure 5 is the "picture is worth a thousand words" version of Table 2, illustrating the power of FEC to clean up what is quite an incoming mess.

	Uncoded ~1e-8	**LDPC** DVB-C2 @ 90%	**SNR** "Gain" (dB)
64-QAM	28	19	9
256-QAM	34	24	10
1024-QAM	40	30	10
4096-QAM	46	35	11

Table 2: Coding Gain of LDPC FEC

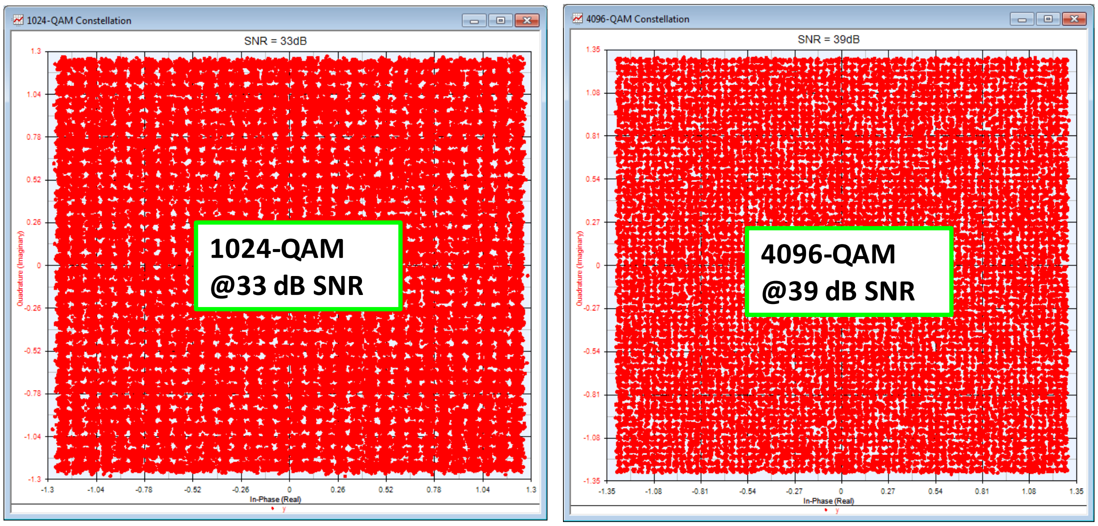


Figure 5: Amazing Error Free: The Power of LDPC Forward Error Correction

FEC II – How Does it Do That?

We can precisely identify the dB of FEC advantage of LDPC versus today's ITU J.83 Annex B downstream (as well as for the upstream). Refer to Figure 6 [1].

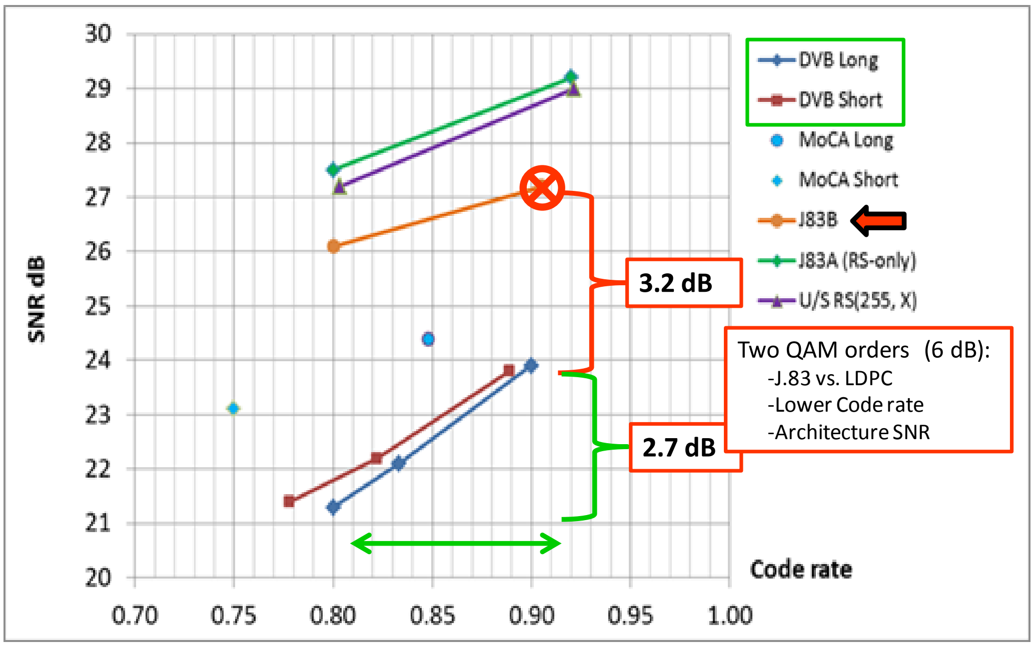
In Figure 6 (simulations by Intel), we can compare SNR vs. Code Rate for the old and new FEC choices. For the downstream, J.83B (orange) can be compared against the DVB-C2 short (red) and long (blue) codeword. The plot is based on 256-QAM, with the expectation that similar relationships will hold for other M-QAM formats for a well-designed code. Note that J.83 Annex B does not actually have variable code rate, but varying the Reed-Solomon code rate enables a relevant and straightforward simulation while allowing apples-to-apples code rate comparisons.

Figure 6 identifies how, with LDPC alone, we could actually manage a two-order increase in modulation profile – a 6 dB theoretical SNR gap – using a combination of the code family and code rate, if this were desirable, as follows:

- Labeled by the orange crosshair and bracket, LDPC at the same code rate provides about 3.2 dB of SNR gain (red bracket) compared to J.83B. A 3 dB change is roughly the equivalent of one half-step modulation order, such as 256-QAM to 512-QAM.

- At the cost of efficiency, by reducing the code rate by about 10% (to 80%), another 2.7 dB can be gained, for a total of 5.9 dB, or nearly 6 dB (green bracket and horizontal arrow).

Thus, a little more than 3 dB comes from the change in code family, and the rest comes from a 10% drop in the code rate. Since the code rate is an efficiency reduction, some or the rest of the difference to get to a 6 dB difference, such as 256-QAM to 1024-QAM, might instead be made up, for example, with architecture or technology evolution in the HFC network.



Reference – Mission is Possible: An Evolutionary Approach to Gigabit-Class DOCSIS, 2012 Cable Show Spring Technical Forum

Figure 6: LDPC vs. J.83 Annex B Comparison (Downstream) [1]

We can perform the same analysis for the upstream, as shown in Figure 7 [1]. Today's upstream *does* have a selectable code rate. The cases for t=10 and t=16 symbol-correcting are shown in the simulation (courtesy of Intel). We show two MoCA™ codes and compare to the MoCA™ short code. The availability of shorter codeword sizes is essential to match the upstream packet size distribution.

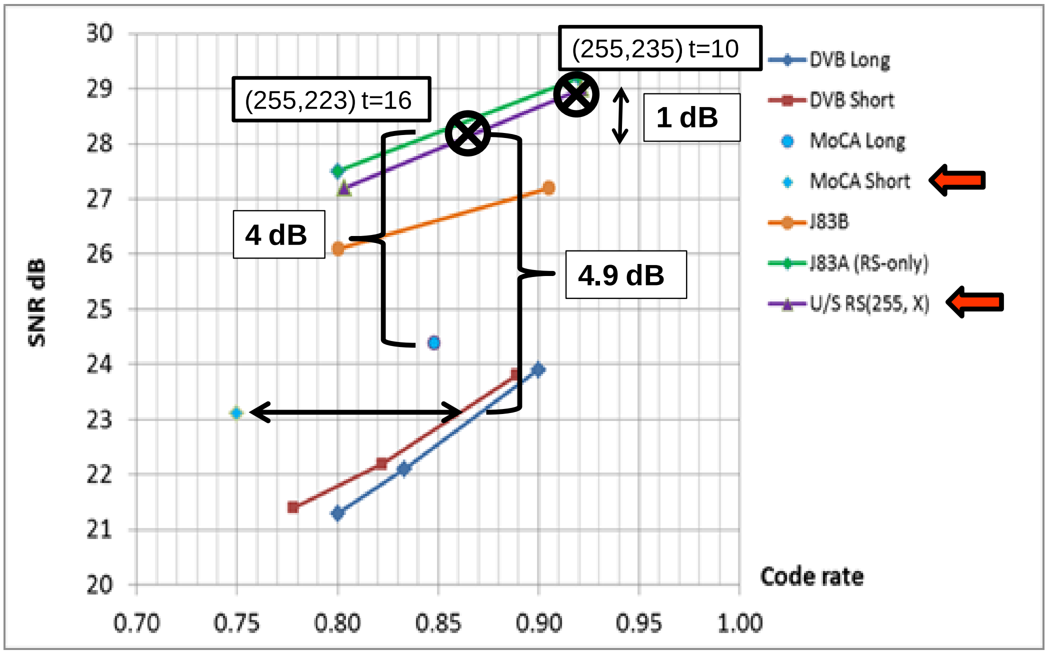


Figure 7: LDPC vs. Reed-Solomon Upstream Comparison (Upstream) [1]

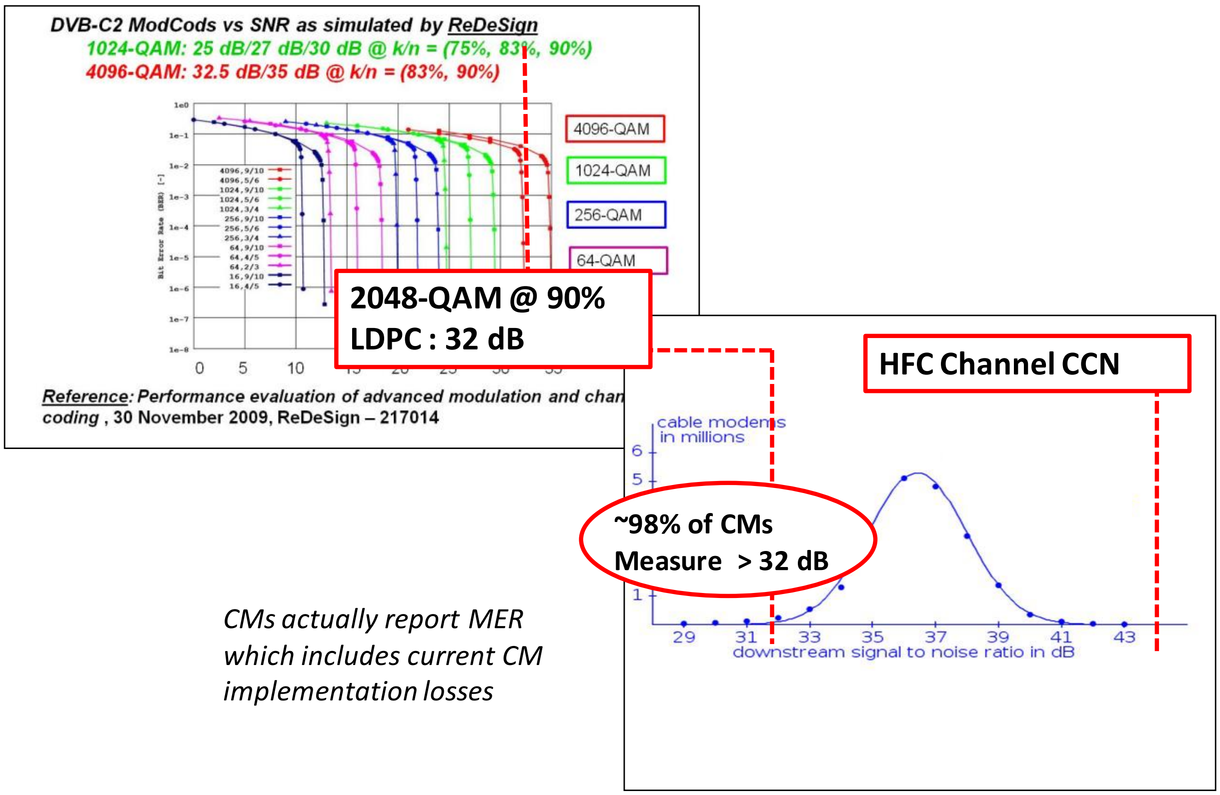
As Figure 7 shows, we can again work out the potential for a two modulation order improvement. Using the MoCA™ short code (blue diamond), we note that the SNR requirement is (4.9 + 1) = 5.9 dB lower than the t=10 error correcting. This comes at the cost of lower code rate (by 17% - significant) and thus lost efficiency. The efficiency loss when comparing the MoCA long code to the t=16 case is much less (2%), but we do not achieve 6 dB, only 4 dB. However, we might consider upstream technology or architecture improvement that offers 2 dB of additional SNR link budget to close the gap.

Since the upstream optical technology tends to be the dominant factor in the upstream SNR, the ability to directly affect the upstream bandwidth efficiency is more straightforward than the downstream. Head-end de-combining is another area where instantly accessible dB can affect the upstream bandwidth efficiency potential.

M-QAM, FEC, SNR: Connecting the Dots

With knowledge of both lower M-QAM thresholds enabled by LDPC FEC, and a well-quantified awareness of the SNR on the receiving end by fielded cable modems, we can connect the dots between the two to examine the potential for new downstream capacity. Figure 8 shows the two together to begin this comparison [16].

The Figure 3 distribution on the lower right – a classic Gaussian bell curve – shows an average of about 36.5 dB and a 2σ variation of about 3 dB. This puts over 95% of the measured modems from this large sample between 33.5 dB and 39.5 dB (±2σ).



Reference: http://www.ieee802.org/3/bn/public/mar13/howald_3bn_01_0313.pdf

Figure 8: M-QAM Potential Based on Today's Measured MER Characteristics [16]

The Figure 4 QAM-FEC simulations repeated in Figure 8 do not include the mid-step constellations. However, they are easily estimated, and in this case the estimate for 2048-QAM for the 90% code rate would be that it is 3 dB lower than the 4096-QAM SNR requirement of 35 dB, or 32 dB. On the CM distribution curve, this represents a performance achieved by about *98%* of the modems. This shows, not surprisingly, that using only 256-QAM leaves potential capacity on the table. Note that 256-QAM @ 90% DVB-C2 LDPC requires a 24 dB SNR, which only re-emphasizes the point.

Of course, this does not account for added operator margin required for robustness. A substantial margin is used by field technicians to guarantee a robust 256-QAM downstream today. Figure 6 shows the 27.5 dB of SNR required for 256-QAM in the J.83B downstream. Typically, operators will look to obtain about 35 dB (operator dependent) to "certify" an install at a customer's home [18]. We will address the margin topic more specifically in a subsequent discussion about downstream optimization, as we anticipate that this paradigm will change. However, for now, we can recognize that using a 2σ spread's lower SNR edge of about 33.5 dB in Figure 8, and subtracting the equivalent 7.5 dB margin we are left with 26 dB as an SNR. Based on Figure 4, this would support 1024-QAM with a code rate close to 80%.

Lastly, note the "HFC Channel CCN" label and red line on the lower right of Figure 8. CCN stands for Composite Carrier to Noise, accounting for both AWGN and digital distortion build-up which looks like AWGN from a noise floor perspective. It is the HFC plant equivalent of SNR. This line describes what the plant can deliver at end of line (EOL). Minimum performance of 42 dB has previously been mentioned, while typical performance is higher, such as that shown here. The point here is that the HFC channel, if properly implemented, is *not* limiting capacity from an SNR (CCN) perspective.

In summary, it should be obvious that 256-QAM is not the best-case bandwidth efficiency possible in the downstream. More bps/Hz are available if we desire to chase after them. Moreover, some of the most important capabilities to obtain these bits is already in place, in particular around the HFC channel quality, as is understood in terms of minimum EOL today, and even as reported by the CM SNR data in Figure 3, which accounts for a broader set of variables which will only improve with architecture and technology evolution. Therefore, if we need more bits, they are not far from reach. And, as Figure 1 implies, the question "do we need them" has already been answered.

The Role of OFDM

An element hidden by the capacity equations in (1) and (2) is the accuracy of a constant, static, and spectrally flat assumption of SNR. In many systems today – particularly wireless – the SNR can be quite dynamic when moving throughout a cell, for example For other channels, such as cable, it is not particularly dynamic, but does vary across the area it serves both geographically and with respect to frequency of operation.

Also, the frequency response of the channel has large implications on the receiver design and its ability to perform close to the spectral efficiency that the channel SNR suggests it should achieve. For wireless, moving across a cell in a metro area creates a difficult multi-path environment. In cable channels, a wide range of ripple and slope may exist due to static channel multi-path (micro-reflections in cable-speak) conditions as well as due to the nature of having a multi-octave RF distribution network and serving uncontrolled home coaxial architectures.

Variable and unpredictable channel conditions are specifically where multi-carrier systems (e.g. OFDM) come into play. The fundamental OFDM concept is shown in Figures 9-11.

The fundamentally different characteristic of OFDM is replacing classic single-carrier QAM, such as the 6 MHz and 8 MHz channels used today for nearly all QAM signals on the cable, with many narrower, subcarriers, and sending these subcarriers in parallel. This is depicted in Figure 9. Narrow means kilohertz-type of narrow. As a practical example, 10 kHz subcarriers would mean there are 600 of them inside a 6 MHz "normal" channel slot in North America. As in single carrier technology, the subcarriers themselves carry QAM, which is why we study QAM modulation formats in detail regardless of RF waveform type. In the ideal AWGN environment, the two techniques perform equivalently.

The other uniquely interesting OFDM characteristic is that the narrow subcarriers overlap by design, as shown in Figure 9. They get away with this (clearly, classic frequency division multiplexing, or FDM, could not) by maintaining a relationship among subcarriers that connects their spacing to the symbol rate so that they remain orthogonal. Ideally, orthogonality ensures that, by the nature of the waveform integration during demodulation, subcarriers do not interfere with one another.

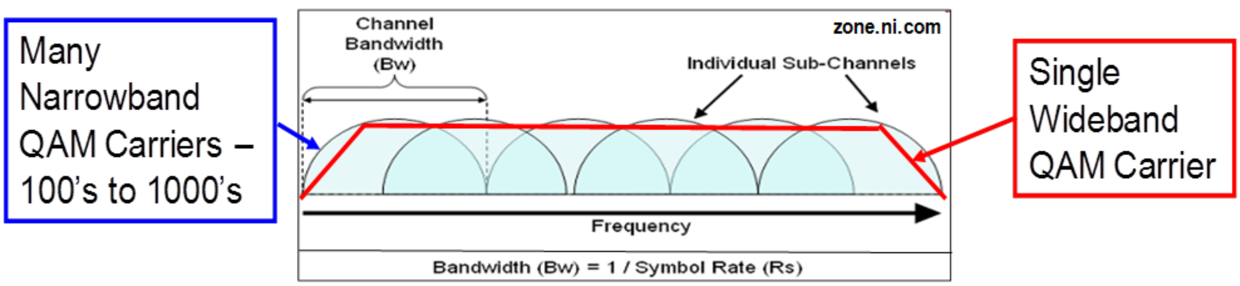


Figure 9: The Multicarrier (OFDM) Concept: Frequency Domain [20]

In the time domain, this "zero interference" quality is shown in Figure 10 whereby integrating (detection) over the period shown for one of the subcarriers has the others summing to zero. Figure 11 shows the frequency and time aspects together. All subcarriers are sent in parallel during a symbol transmission, and the process is repeated at the next OFDM "symbol" transmission.

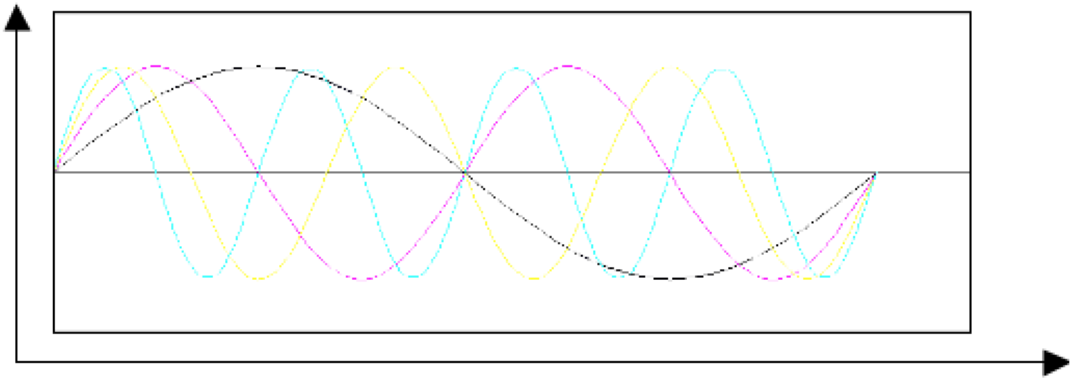


Figure 10: The Multicarrier (OFDM) Concept: Orthogonality in the Time Domain [13]

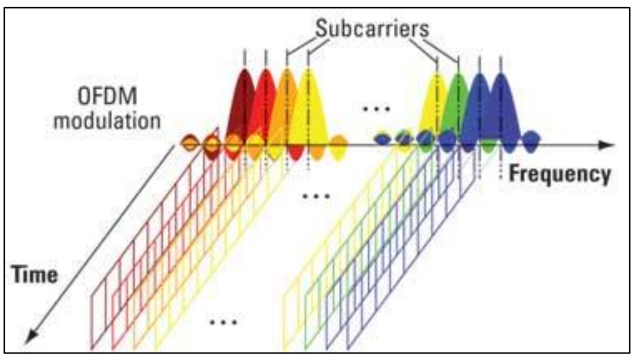


Figure 11: OFDM - Frequency and Time Domain [19]

The next transmission does not immediately follow the first (at least in terms of payload transmission) – this is one of the fine details of OFDM system design we will not get into here, but which deals with how OFDM effectively performs the function of "equalization." OFDM uses what is called a "cyclic prefix" (CP) to delay a new data transmission beyond the multi-path window.

The whole OFDM idea sounds unnecessarily complex, and indeed this was once the case. Like FEC, the multi-carrier concept was invented by brilliant engineers who noted many of the potential benefits reaped from this approach to accessing a channel many years before the implementation became practical. We will not get into implementation details, but OFDM was largely made practical, and actually quite simple, with advances in real-time computing power than enabled wideband, high-speed, high resolution FFTs that could be processed in real-time.

Shannonizing with OFDM

A good way to interpret the OFDM approach in terms of its capacity-maximizing effect is to write the expression for capacity in (2) in "long" form:

$C \approx (1/3)\sum_{\Delta f}[\Delta f]\ [P(\Delta f)\ H(\Delta f)\ /\ N(\Delta f)]_{dB}$ (3)

Here, instead of bandwidth, we have used a summation of spectrum chunks using a set of small frequency increments, Δf. The sum of all Δf increments is the bandwidth available, B. Instead of SNR, we have broken it down into its components: signal power (P), noise power (N), and channel response (H) – each also over small Δf increments. In practice each Δf represents the width of one OFDM subcarrier.

The total capacity above is then simply the summation of the individual capacities of chunks of spectrum. The purpose of the form used in (3) is to recognize that channels may not have a fixed SNR characteristic, such as due to expected non-flat frequency response variations and uncharacterized spectrum above today's 1 GHz forward band. In this case, the capacity of a not-flat SNR region can be calculated by looking at it in small chunks that, because of their narrow width, themselves approximate flat channels. A similar argument applies when there is, for example, interference. The affected OFDM subchannels will have a lower SNR (in this case S/(N+I). This flexibility is a key advantage of multi-carrier modulation such as OFDM – very narrow channels, each of which can be individually optimized.

For a single carrier transmission, it becomes increasingly difficult for wider and wider channels to achieve the same effect without complex, and sometimes impractical equalization techniques and interference mitigation mechanisms. Or, in the case of DOCSIS, it becomes impractical to channel-bond more and more single-carrier channels without incurring excessive complexity and inefficiency.

The long-form capacity equation above demonstrates why OFDM is often better suited to achieving the best throughput possible, as compared to single-carrier techniques in channels with poor or unknown frequency response, and in particular, when that response is time-varying.

The HFC downstream is typically very high SNR and generally well-behaved. However, it can be subject to large broadband frequency response variations when signal reflections are high. The downstream is also increasingly susceptible to 4G interference as these deployments increase, as well as interference sources that have existed for years. Outside the current downstream – above 1 GHz – plants are likely to vary widely as there are no requirements to be met or equipment specifications that can be used to help define the spectrum, though the coaxial cable medium clearly can be exploited beyond 1 GHz.

In the upstream, the channel is much less predictable than in the downstream, particularly at the low end of the band, and burst noise events are more prevalent than in the downstream. Furthermore, the upstream is as likely if not more so than the downstream to see a bandwidth extension into new territory, such as 85 MHz and even to 200 MHz. However, because of its "funneling" architecture, interference that may be localized and insignificant in the downstream today may impact the channel for all in the upstream when the diplex is adjusted for more upstream spectrum. The FM radio band is the most obvious candidate, should the upstream extend beyond 85 MHz. The interference-protection properties of OFDM will be valuable in this case, as it is in the troubled part of the return band today.

Note that in the upstream, the likely multi-carrier candidate is actually OFDMA, or Orthogonal Frequency Division Multiple Access. The principles of the signal waveform are the same, but in the case of OFDMA, different sub-channels can be allocated to different users simultaneously, an attribute important to efficient use of the upstream. The difference between OFDM and OFDMA is shown in Figure 12. We will generally use just "OFDM" to refer to the technology in both upstream and downstream.

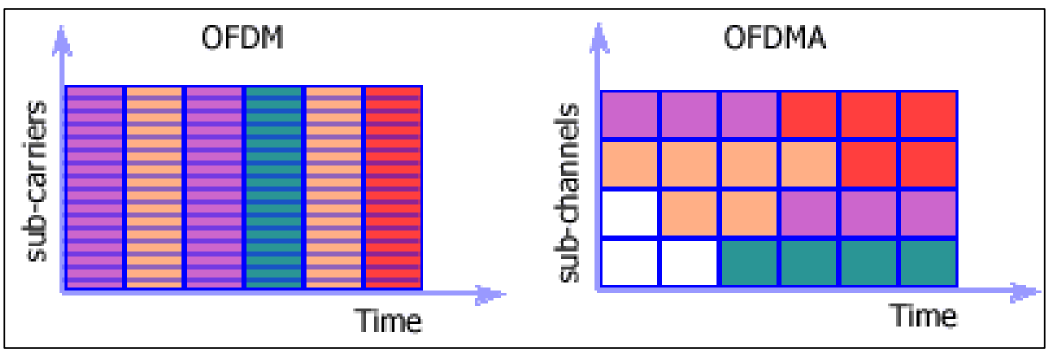


Figure 12: OFDM vs. OFDMA [12]

As discussed previously, supporting more bandwidth-efficient M-QAM profiles over HFC has little to do with whether we are discussing single carrier QAM or OFDM-QAM. When it comes to SNR (AWGN), system performance is identical. OFDM's most valuable HFC role is to overcome frequency response characteristics and unknown channel quality and manage interference conditions to yield the best probability of maximum SNR exploitation for capacity. Very wideband (high-speed) operation is also a major plus. Historically, OFDM applications have been linked by this common thread – unknown or poor RF channels – and the benefits it provides in those cases are being brought to the cable environment. In the downstream, the most questionable spectrum would be the band above 1 GHz, and in the upstream the entire channel is more suspect, but especially so 5-20 MHz.

Relative to bandwidth above 1 GHz, Figure 13 [1], shows the range of insertion loss characteristics of various models of a *single tap* type above 1 GHz for "1 GHz" specified taps. It is clear that any given tap, much less a cascade of taps, will be highly unpredictable from system to system, and even from RF leg to RF leg in the same system.

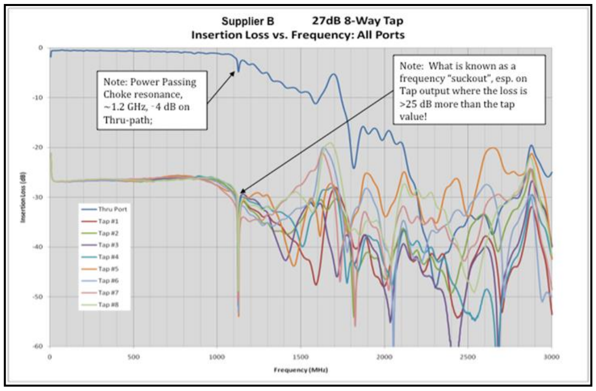


Figure 13: Unpredictable Frequency Above 1 GHz [1]

There are other important OFDM benefits not associated with system performance. Some of these are listed in Table 3.

The second point in Table 3 is perhaps the next strongest argument for OFDM for HFC, albeit it a more practical "operations" one. With so much spectrum and service evolution anticipated over the next decade, the granular spectrum management enabled by OFDM through flexible subcarrier allocation (using some but not all subcarriers) is a valuable tool when working around a full band of legacy spectrum.

Why OFDM?
Optimizes Channel Capacity, in particular for unknown, uncharacterized, and hostile interference channels
Granular spectrum allocation beneficial during band plan and service transitions
Multiple sources of supply and likely cable investment
Consistency with **other standards** and cable network extensions (wireless, EPoC)
OFDM + LDPC to Layer 1 as IP is to Layer 3 – likely final RF step (little more capacity worth exploiting)
Implementation complexity favors OFDM over TDM afor wideband channels with linearity distortions
More **Spectrally Efficient** Wideband Channel than NxFDM, 2-D Multiple Access (OFDMA)

Table 3: Why Cable OFDM?

Other points in Table 3 worth mentioning include the increasing ability to do computationally complex operations in real time. OFDM implementation – once the major obstacle – has become a strength through IFFT/FFT functionality that forms the core of transmit and receive operations.

This implementation advantage leads to one of the final strong, business-oriented arguments for OFDM. As an ecosystem, the number of suppliers of OFDM technology and the range of industries engaged in it enlarges the pool of technology resources and leverages tremendous economies of scale. The wireless industry and Home LAN products in particular both represent very high volume applications.

Impairments: Single Carrier and OFDM

OFDM puts a different signal type on the wire, and because of that it responds differently to some of the common impairments of cable – unique (CTB/CSO) or otherwise (additive interference, phase noise). We mention these two important ones here, but for a fuller treatment refer to [6]. An understanding of the differences will be critical to properly specifying and operating OFDM on the cable channel, and analysis of these effects is ongoing.

CW Interference

Single carrier techniques combat narrowband interference through adaptive filtering and equalization mechanisms. OFDM, on the other hand, deals with narrowband interference by avoidance. Also, what may be "narrow" for a single carrier QAM signal may not be narrow relative to an OFDM subcarrier. Figure 14 shows OFDM impinged upon by two interferer types – a CW carrier and a modulated waveform of some unspecified bandwidth that is similar to OFDM subcarrier spacing.

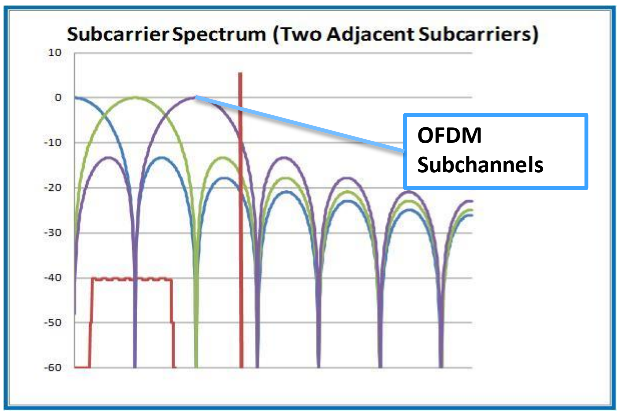


Figure 14: Interference as Seen by OFDM

Subcarriers imposed upon by an interferer can be nulled or modulated with a more robust modulation profile. The effect is a capacity loss, but generally a modest one because only a limited number are affected. Compared to SC-QAM, OFDM offers graceful degradation via lost capacity, as opposed to a thresholding effect at some intolerable level of interference. This could be viewed as both pro and a con. SC-QAM, for example, may find low levels of interference essentially invisible from a detection perspective, a scenario well represented by analog CSO/CTB distortion beats in the forward path.

CTB and CSO, when analog video is present, also have more of a deterministic quality – always preferred – in location, level, and whether they will even be present or not. Figure 15 compares 6 MHz SC-QAM and OFDM-QAM with respect to CTB/CSO interferers.

Two key characteristics stand out:

(1) Distortion beats are no longer necessarily narrow relative to the subcarrier bandwidth, on average. The distortion bandwidth and amplitude vary slowly, however, and these peaking effects can have well-documented implications for QAM performance and interleaver depth.

(2) Beat amplitude is much higher relative to SC-QAM since each subchannel is a small fraction of the total signal power in, for example, 6 MHz. For the 600 subcarriers per 6 MHz example, this is 27 dB. So, CTB/CSO of 53 dBc is now 25 dBc! And that is just the average, not including its amplitude modulation characteristics. Clearly, for OFDM, the FEC will be required to deliver error-free bandwidth efficiency.

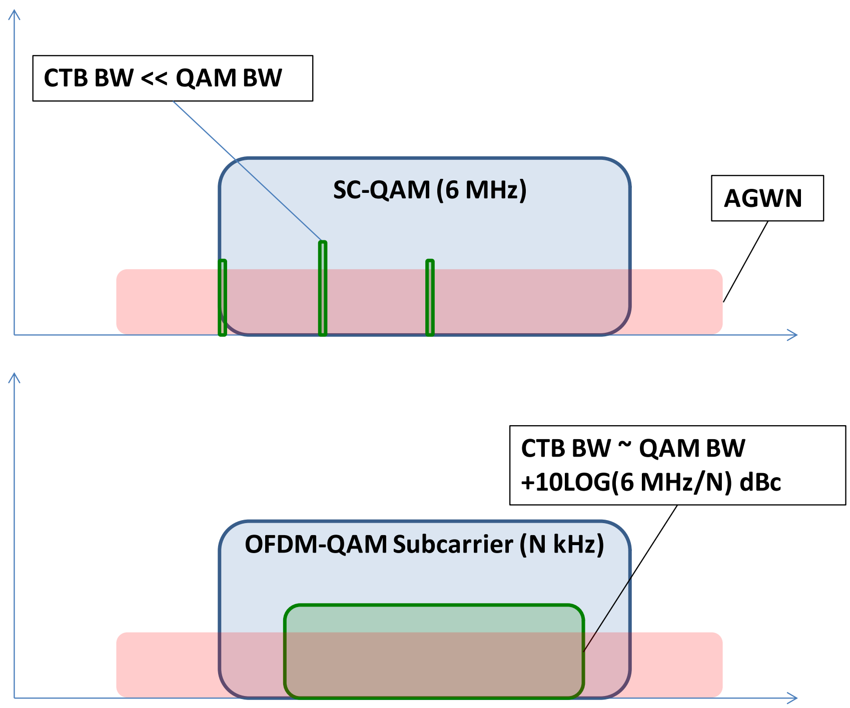


Figure 15: CTB/CSO Interference – SC-QAM vs. OFDM-QAM

OFDM system design and choice of parameters for the error mitigation subsystem are used to overcome interference in the channel whether the mechanism is distortion beats or additive interference. The latter is being observed in some cable systems in LTE bands.

Phase Noise

OFDM creates an interesting scenario with respect to phase noise degradation. A typical assumption for SC-QAM is "slow" phase noise. The exact spectral mask is less important – only the untracked rms phase noise matters. For OFDM, with many narrow subcarriers, the phase noise mask will typically extend *beyond* the subchannel bandwidth. Figure 16 shows a characteristic low-pass shape of untracked phase noise (two cases of different "bandwidth") against an OFDM sub-channel spectrum.

Phase noise thus includes two degradation mechanisms for OFDM. There is an error common to all subcarriers related to the "in-band" effects and known as "common phase error" or (unfortunately) CPE. This component is often largely tracked out and therefore of little consequence, and has the classic "rotation" effect on each subcarrier (thus "common"). The other typically more impactful component is that associated with Interchannel Interference (ICI) as the mask cross into other sub-channel bands and these effects are summed. This phase noise effect is additive, uncorrelated noise, which is better than rotational CPE, but the ICI effect from adjacent subchannels has the potential to be quite high.

Both CPE and ICI effects must be accounted for in system design, and the techniques for doing so are well understood. There are just significant differences in how to approach the solution compared to the traditional single carrier design, and we will be attempting to do so with much more sensitive, higher bandwidth efficiency, M-QAM formats.

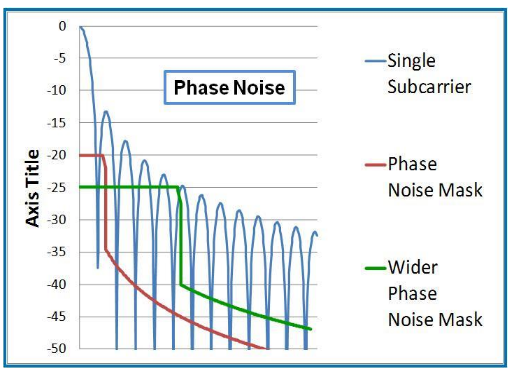


Figure 16: The Shape of the Phase Noise Mask is Critical for OFDM

Towards a Layer 1 Standard

OFDM offers a robust way to exploit spectrum above 1 GHz, which will be necessary to achieve the objective of 10 Gbps for DOCSIS 3.1. It also provides advantages in the downstream and upstream as interference sources arise going forward, and provides robustness at the low end of the return that can only be managed with S-CDMA today. In addition to its capacity optimizing capability, because of its granularity of spectrum allocation, OFDM provides bandwidth efficient flexibility for systems undergoing service and spectrum evolution, which could prove very valuable. With the transition to IP, continual enhancements of HD and on-demand services, managing a full downstream and nearly full upstream, and the expectation of a new diplex crossover sometime in the future, this is a valuable benefit.

In the long-term, OFDM-QAM plus LDPC FEC, because of its capacity optimizing capability in any channel and implementation simplicity, can be viewed for OSI Layer 1 (PHY) what IP has become for OSI Layer 3 – a de-facto go-to standard. This enables potentially significant simplification of network evolution over time. Virtually all modern RF systems across multiple industry segments implement some form of OFDM – 4G Wireless, Wi-Fi, MoCA™, G.hn, HomePlug AV, 802.11n, and VDSL. This end-state scenario would be very similar to what is currently happening in the all-IP transition, where in that case, we are simplifying at the network layer. The standardization would extend to the lower layers of the stack and include some components of a software-defined architecture.

♪ Are You Ready for Some 4k? ♫

The performance of HFC networks in the downstream is very well understood from decades of achieving fidelity acceptable for analog video. Some typical performance numbers are shown in Table 4 for the case of 60 analog carriers on a 750 MHz system over a range of cascade depths for two different return spectrum scenarios.

These first four columns are referenced to analog levels, so for digital they must be lowered 6 dB. This is listed in the far right column as "QAM CCN." Again, CCN captures all noise floor components – AWGN and digital distortions – and is for all intents and purposes HFC's SNR. Digital distortion contributors are many and largely independent, so a Gaussian assumption is reasonable.

An important and expected result from the table is the improvement in the CNR and QAM CCN as the cascade shortens and service group size gets smaller. As fiber penetrates deeper, average bandwidth per subscriber is increased, but also the channel quality improves. An RF cascade has the effect of cascading degradation at every amplification stage in the downstream – both noise and distortion. In the upstream this is also the case, but to a lesser degree of importance, while the shrinking service group size has more significant benefits to channel quality, associated primarily with interference funneling.

		750 MHz				
		60 Analog				QAM
		CNR	CSO	CTB	CCN	CCN
Return	N+0	54	64	67	53	47
5-42 MHz	N+3	51	60	63	50	44
	N+6	50	58	60	48	42
	N+0	54	64	67	53	47
5-85 MHz	N+3	51	60	63	50	44
	N+6	50	58	60	49	43

Table 4: Downstream Performance vs. Cascade

Figure 17 illustrates the fiber deep concept from a cable serving area footprint perspective. Note that service group splitting may also be achieved simply though a segmentable node, with no effect on the cascade depth. In this case, it is the upstream channel primarily that benefits.

Network Nirvana

A somewhat natural architectural end-state vision for HFC is business-as-usual node splitting culminating ultimately in an N+0 system – a passive coaxial last mile with no RF actives after the node. The benefit, in terms of channel performance, can be observed from Table 4, where now everyone is the "N+0" column.

Besides the channel quality improvement afforded by N+0, a very important advantage of the "fiber deepest" architecture is the ease with which new capacity can be exploited without the existence of actives in the path. Actives involve diplexers, which add obstacles to adjusting spectrum allocations, and their ability to supply quality "excess" bandwidth to 1.2 GHz the way taps may is probably more questionable, being active circuits.

It is assumed that HFC link performance can be maintained as the spectrum shifts to higher spectrum in the case of 1.2 GHz. The loading effect of *increased* total spectrum, such as 108 MHz-1200 MHz, can also be calculated.

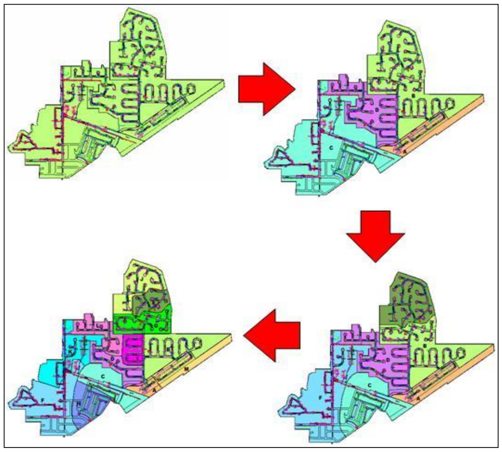


Figure 17: Fiber Deep Segments a Serving Area

We will recognize the N+0 benefits in subsequent quantization through its effect on aggregate capacity in a serving group, the SNR it enables in fully evolved FTLA architectures, and the new spectrum it tees up for exploitation of new capacity.

Downstream M-QAM Readiness

Using QAM requirements (Figure 4, Table 2) and Table 4 performance for 750 MHz networks with analog loading, we can derive what M-QAM bandwidth efficiency can be delivered to CMs over a range of HFC and home architecture variables prevalent in a typical network deployment. This is shown in Figure 18. Shown are the above calculated CCN values of 42 dB (N+6), 44 dB (N+3) and 47 dB (N+0), labeled via the pink vertical lines. Any other relevant spectrum/loading scenario can, of course, be evaluated.

Black horizontal threshold lines represent variations of the drop/home architecture, in this case assuming a fixed tap port level of 15 dBmV (analog reference). Different drop lengths have different loss, and the amount of splitting at the home ahead of the CM also varies the level into the receiver. An assumption for architecture evolution is a Point-of-Entry (POE) gateway-type approach, and we therefore limit the splitting to 4-way maximum. This is a *major* assumption for the evolved home architecture with important implications associated with the CM SNR contribution. For an assumed CM Noise Figure of 10 dB and a Tx/Rx MER (effectively all non-channel implementation losses) of 43 dB, we can observe the range of M-QAM formats that can be supported across the variables shown.

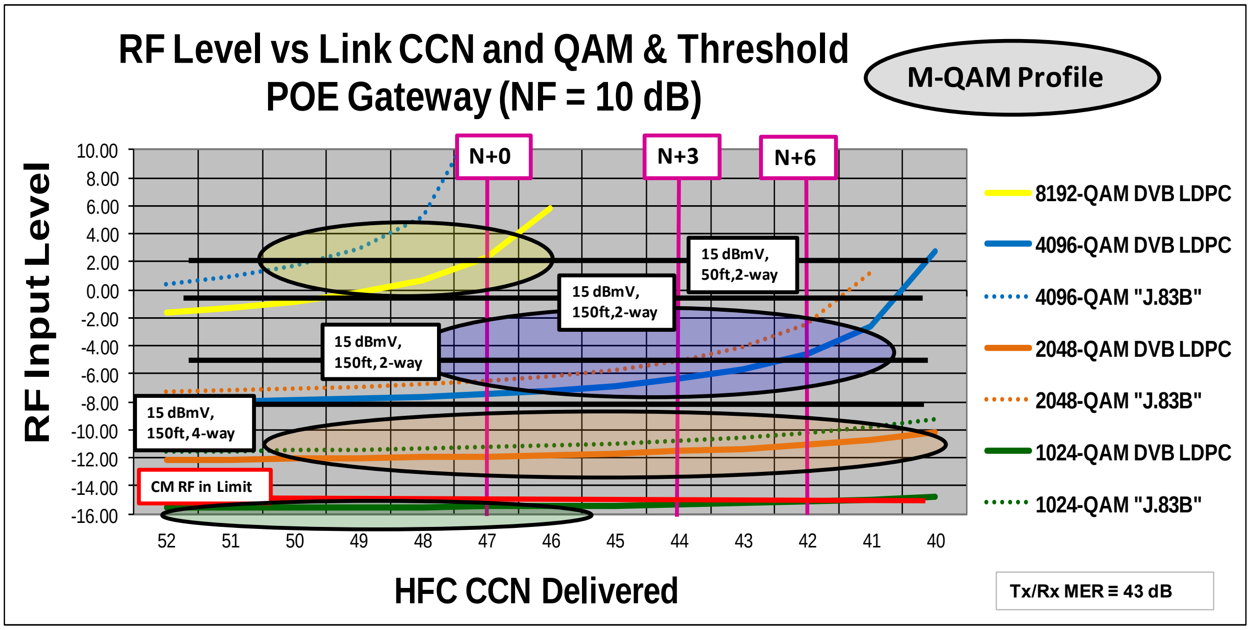


Figure 18: HFC Geography and Home Plant Architectures Means a Range of SNR

The QAM + FEC profiles that suit this set of conditions are identified by the colored ovals, which circle the region of operation of the combined variables. In this case, depending on where in the plant a subscriber was located and what drop/home architecture exists, four different QAM formats might be obtainable. Again, we have not discussed margin, but for a simple "margin" philosophy, consider that each CM reports an SNR, and the CMTS selects the next lower (more robust) profile as margin. In such an example, the same number of formats would exist, but instead of ranging from 1024-QAM to 8192-QAM, they would range from 256-QAM to 2048-QAM.

Key items that Figure 18 reveal under these specific assumptions and a relatively limited range of drop/home variables are listed below. Note also that tap port levels – fixed in this analysis – are difficult to keep aligned to a very small range over a series of taps in a string.

(1) The plant capability, at least, puts 8192-QAM in play as a possibility. It is not the limiter as architecture evolution continues either to N+0 or to a remote PHY approach.

(2) There is only minor sensitivity to the range of HFC performance for 4096-QAM out to at least N+6. There is almost no sensitivity for 2048-QAM. The drop/home architecture is the more significant factor for the NF assumed.

(3) The sensitivities to HFC CCN would increase with lower NF CPE, but the average modulation profile possible for a given drop/home architecture would also be more efficient as a result of the lower NF. An 8 dB NF may be reasonable without excessive CM cost burden.

(4) 1024-QAM with a J.83B flavor of FEC would actually be achievable today.

The evolution from 750 MHz architectures to 870 MHz architectures to 1 GHz architectures has by and large, been about expanding the bandwidth with new optical and RF technology while achieving equivalent EOL worst-case noise and distortion performance. Thus, Figure 18 is relevant, though not exact in the better performance cases, as the forward band extends to 870 MHz and even to 1 GHz. The conclusion – a range of SNR performance – remains, with the range similar but slightly compressed (on the order of 1 dB) due to the minimum performance being the same, and the better performing scenarios – the shorter cascades – degraded from 750 MHz performance by the heavier loading.

A similar statement can be made for RF loading as digital loading replaces analog loading. Clearly discrete distortions such as CSO and CTB reduce considerably in exchange for more digital distortion components, and their contribution would be reflected in CCN or MER degradation that must be managed. Again, variations are small and mostly would be reflected in the better performing cases since end-of-line targets are typically non-negotiable minimums.

Intriguing about Figure 18 is the range – there is obviously capacity left on the table if a 4096-QAM set of users is only receiving at 1024-QAM or 256-QAM. Is there a better approach? We discuss this in the next section.

Something New: Switched Broadcast

Figure 3 and Figure 18 tell us very similar things from different perspectives. In Figure 18, we have the HFC plant telling us the channel quality is there to significantly improve bandwidth efficiency if we can get sufficient level to the home. Left unsaid by Figure 18 is that, assuming we are looking to improve capacity through higher order M-QAM profiles, it would be entirely reasonable to expect CM sensitivity to improve over what is achieved today. The same can be said of fidelity requirements of the equipment on both ends; but again, Figure 18 does not tell us anything about that. It just says –"I, your HFC plant, can deliver your SNR requirement if you care to exploit it."

Figure 3, however, does include fidelity component as well as the plant component. It just is based on "legacy" fidelity components. And yet, still it is reporting to us "I have an SNR reserve for a lot of my modems that can do better than 256-QAM."

Recall, in Figure 3, we identified a 2σ range (> 95%) of CMs of about 6 dB. Therefore without even attempting to accommodate the other ~5%, we can identify three QAM thresholds in 2σ – 6dB being one square order apart, and the half-step QAM format between. For example the range 1024-2048-4096 QAM is a 6 dB range. The 4096-QAM users are losing out on some possible capacity if they are only running at the 1024-QAM level.

The above is the basis of a switched broadcast approach to the downstream, and Figure 19 captures this with the use of Figure 3. This is also referred to as Multiple Modulation Profiles [15,18]. As described, Figure 3 is somewhat the CM equivalent of Figure 18, but with the equipment limitations built into the SNR reported. This fact is good in that it is a very practical representation of reality. At the same time, because it is an MER measurement, it may capture equipment effects that specifically are insufficient for needs beyond 256-QAM. And the home variations that may not exist in a POE deployment are also represented. These differences would skew positively the Figure 3 distribution.

In any event, because we *do* care about the remaining 5% of CMs not in the 2σ range – both above and below, we can, for example, split Figure 3 up as shown in Figure 19. This example creates five regions. More regions can be created to cover the distribution with more granularity, as shown in the black "dashed" lines creating two intervals between the colored ones shown. Such granularity is not easily available with modulation formats without going outside of the rectangular M-QAM family, but can be achieved through the use of different code rates of the LDPC FEC. Again referring to Figure 3, the relationship of lower code rates leading to lower SNR requirements is possible to determine if this granularity is desired.

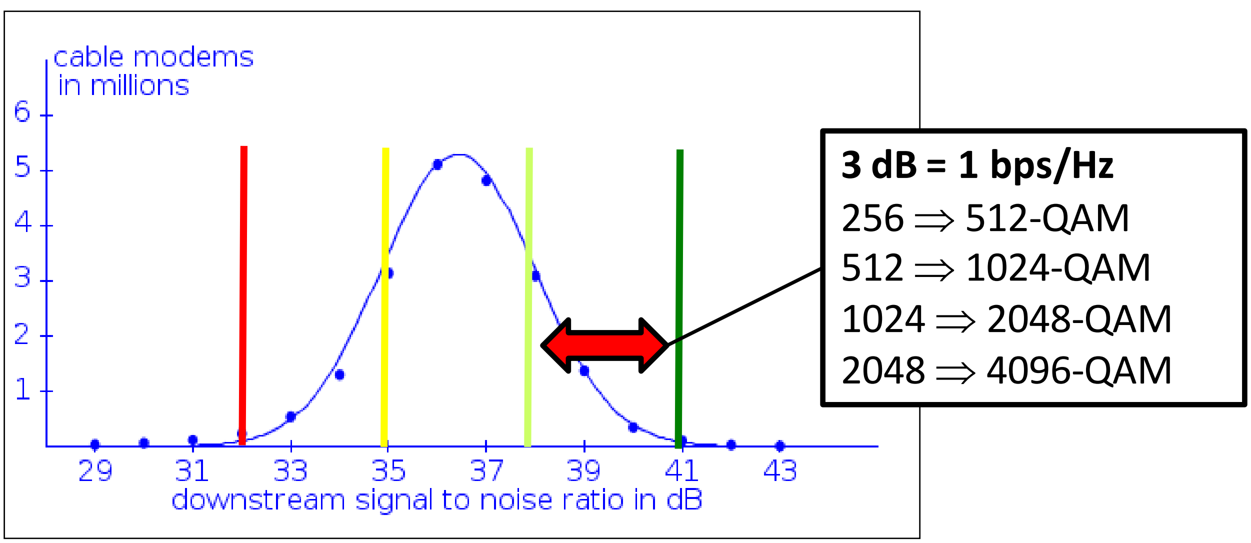


Figure 19: Multiple Modulation Selections Exploit the Range of CM SNRs

If a CM has a choice of M-QAM profiles, then it can select the one that optimizes its capacity. Every CM can do this, and select among the QAM-FEC "buckets" that suit its estimated channel performance, optimizing the average capacity usage. Each bucket represents a group of modems subscribed to the same "broadcast" profile in what amounts to a switched broadcast downstream.

It is straightforward to make a simple estimate of the relative gain above a single 256-QAM broadcast selection using the above distribution and choosing the average M-QAM profile as 2048-QAM. Based on Figure 3 and Figure 8, the threshold for 2048-QAM is about 32 dB SNR. If we use 35 dB for margin purposes (one QAM profile lower than what a modem reports gets selected), then this threshold is about 1σ lower than the statistical mean of the distribution in Figure 3. The calculation easily follows, with 256-QAM being a vanishingly statistically small percentage but which must be upheld as a fall-back QAM profile, leading to the following:

- ~16% @ 1024-QAM
- 68 % @ 2048-QAM
- 16% @ 4096-QAM

This, of course averages 11 bits/symbol, or 11/8 = 37.5% capacity gain. Using a margin of 6 dB, similar, for example, with what is used today for single-profile 256-QAM, reduces this efficiency gain to about 25%. Note that the maximum efficiency gain is 50% at 4096-QAM – 12 bits/symbol of 4096-QAM to 8 bits/symbol of 256-QAM. Per Figure 18, it appears with N+0 – and probably some enhancements to receiver performance – 8192-QAM is within reach, and up to 62.5% spectral efficiency maximum gain.

The calculation with increased margin identifies one of the other significant advantages of a switched broadcast approach as opposed to traditional broadcast of 256-QAM only. The 256-QAM only (or any single profile selected) must be able to be received by all robustly to be an effective solution. It therefore enforces a "lowest common denominator." Whatever the least capable CM can achieve is what everyone receives. Since there are often outliers and corner cases, these lowest-performing devices drive the total channel capacity as well as the margin allotted to ensure robustness. With no other option to handle a connection problem other than a truck roll, operators tend to ensure a very conservative field margin when a CM is deployed [18] – again, about 35 dB as a typical number for a receiver slicer threshold of 27.5 dB and a DOCSIS minimum BER/SNR requirement set at 30 dB SNR.

By contrast, a switched broadcast enables a CM that is experiencing connection problems – such as counting excessive codeword errors – to switch down to a more robust profile and (likely) continue to have service. This can work in the other direct as well – as plant evolutions occur or home architecture re-engineered, modems can move to more capable profiles for more throughput. As such, because plant variations of performance tend to be very slow under normal conditions, periodic updating of the QAM buckets is possible, and interrupt mechanisms potentially permissible when problems ensue, it is foreseen that the need to run the downstream operation with 6-10 dB of margin for robustness will no longer be necessary. A reasonable recommendation, for example, again might indeed be to choose a profile that is one half-step more robust than the profile that a CM reports that it can support.

Figure 20 diagrams how a switched broadcast approach may operate in practice. The approach does not come for free. There is an obvious increase in complexity in the MAC scheduling function, which now must schedule groups of modems instead of blasting out traffic with little knowledge of the receiving aggregate.

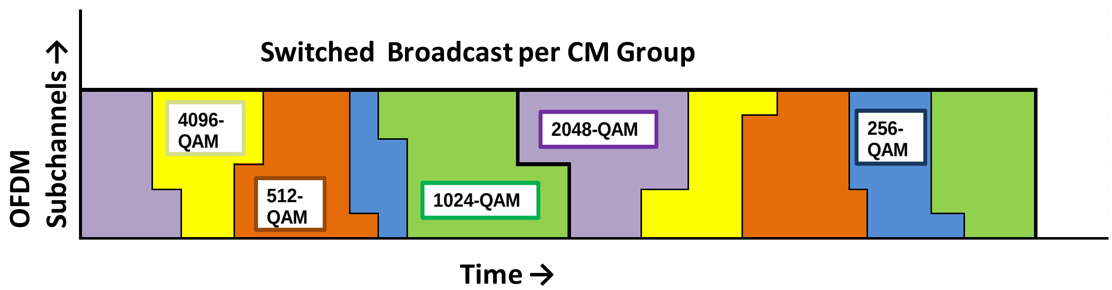


Figure 20: "Switched" Broadcast Modulation Profiles Exploit the Range of CM SNRs

Additionally, a major element of how efficiently an OFDM channel can be used is the choice of cyclic prefix (CP) and its relationship to the symbol time that was identified previously. In short, the CP manages the reflection energy by basically waiting out multi-path. CP is selected to outlast the echo, but in so doing removes time of payload transmission from the channel, costing efficiency.

To manage the complexity of switched broadcast, the same CP for each profile segment should be used. Similar to the SNR least common denominator that drives the idea in the first place, we are now subject to a CP lowest common denominator. The CP does not actually have to be chosen to completely outlast all of the echo content – residual intersymbol interference (ISI) is acceptable so long as it does not significantly affect the total SNR. However, the SNR degradation due to residual ISI of CP must be tolerable for the *highest* QAM profile. Lower QAM profiles may have been capable of a shorter CP and a more efficient usage because they could have tolerated more SNR degradation due to residual ISI. This causes a quantifiable loss of efficiency.

Upstream 85 MHz: Ready and Able

A key component of the evolution of HFC is enhancing the upstream. For many years, it has been recognized that to DFB return optics is required to take advantage of DOCSIS 2.0 and DOCSIS 3.0 capabilities, in particular around 64-QAM at 5.12 Msps. With DFBs assumed coming into place virtually everywhere a high-capacity upstream is desired, and DFB technology having advanced considerably since earlier generation lasers, we can now earnestly look at the ability of today's return optics to support beyond 64-QAM modulation and beyond 42 MHz of spectrum. The limitations of only 37 MHz of upstream, especially with a significant portion of it heavily polluted, demands that a wider band upstream be available in the future. DOCSIS 3.1 sets 1 Gbps of upstream as an objective.

In Figure 21, typical performance of an upstream DFBs at nominal link length over the 85 MHz mid-split bandwidth is shown. Also included is the RF noise contribution of a deep cascade (N+6) combined four ways (dashed blue). CMTS receiver sensitivity for high-sensitivity DOCSIS 3.0 upstream receivers is also included to arrive at a net channel response (solid blue).

New PHY performance thresholds using LDPC FEC assumptions (MoCA™ Short) of Figure 7 are shown on Figure 21 with 6 dB of margin to allow for burst receiver implementation complexities and operating margin for the more dynamic upstream channel environment. As in the downstream, system performance suggests we can be much more capable than we are running in most upstreams today.

Of course, the upstream picture is a little more variable than the downstream. There are still major performance limitations where FP lasers still exist and where Head-end combining to limit port counts effectively combines noise and halves the available SNR per each combine. Gradually, however, we expect these situations to melt away and be left with DFB links (or digital return, roughly the same performance), uncombined, and over smaller serving groups that also begin to take a bite out of the upstream additive interference problem. Figure 21 allows us to see where this potentially takes us.

It shows that with new LDPC-based FEC, 1024-QAM is possible for high performance DFB optical links available today over the full 85 MHz. There is about 13 dB of dynamic range (DR) above the threshold – not as much range as today's 64-QAM over 42 MHz but above the 10 dB DR standard typically used to define sufficient robustness for the link itself. Clearly, for the given link performance, the FEC is making an important difference for 1024-QAM support.

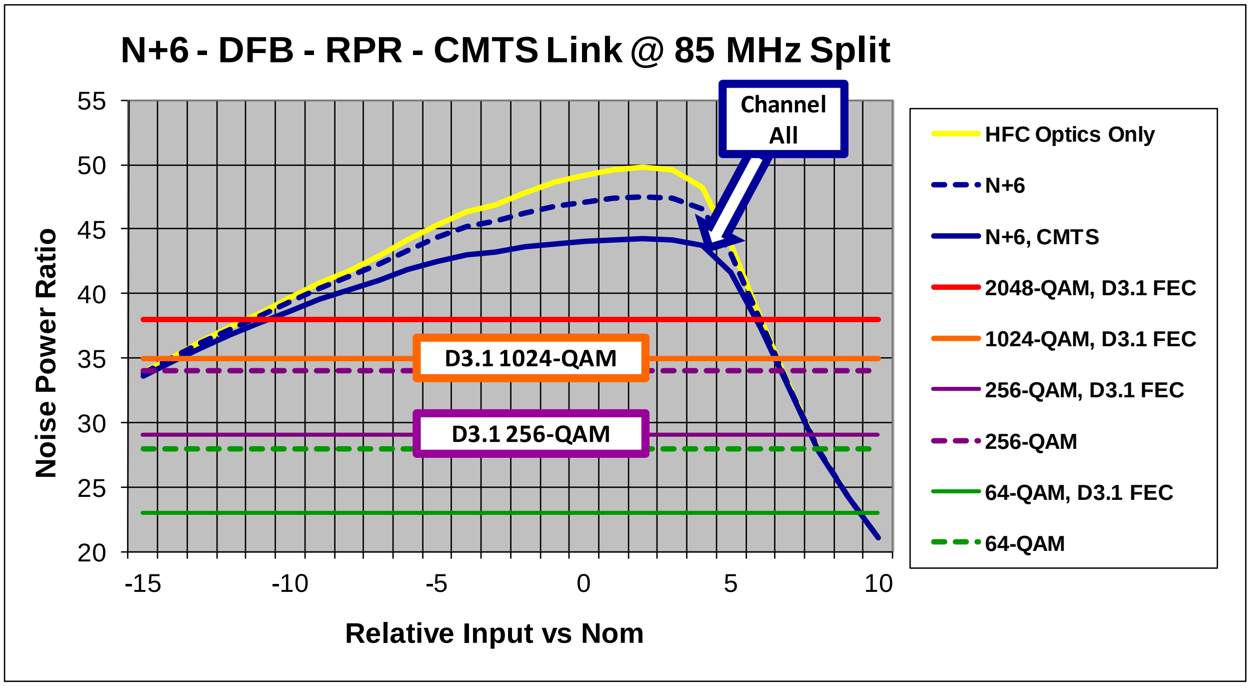


Figure 21: Modern DFBs, Improved CMTS Sensitivity, 85 MHz of Spectrum, and New FEC Create New Upstream Capacity

As Figure 21 also shows, 2048-QAM has precisely 10 dB of dynamic range, so it is actually borderline sufficient. Extended link lengths, minimum guaranteed performance, or older DFB (1 mw) lasers might yield insufficient dynamic range for typical robustness. Also, not all CMTS receivers are created equal, and without a high sensitivity receiver, net end-to-end DR will be degraded. Lastly, the complexity of 2048-QAM itself probably demands an allowance for additional implementation loss and/or dynamic range threshold. It would be very premature to state that the upstream is capable of 2048-QAM until further analysis can be done and bust receiver artifacts better understood. But that the link is in the ballpark of "good enough" is encouraging.

Figure 21 indicates why all of the M-QAM formats in Figure 2 are worth considering for the upstream as well as the downstream. The technology and architecture variables are falling into place to make these possible from a plant perspective, shifting the performance burden to the complex task of burst receivers.

Up, Up, Up and Away

High-split describes essentially anything that extends beyond the DOCSIS-defined 85 MHz mid-split, but has for many years implied an upstream spectrum of 200 MHz. In Figure 22, we extend the prior analysis to this case of a 200 MHz split. All laser characteristics are assumed the same, so the calculation is based only on signal loading loss associated with the sharing of a fixed power into the laser over a wider bandwidth.

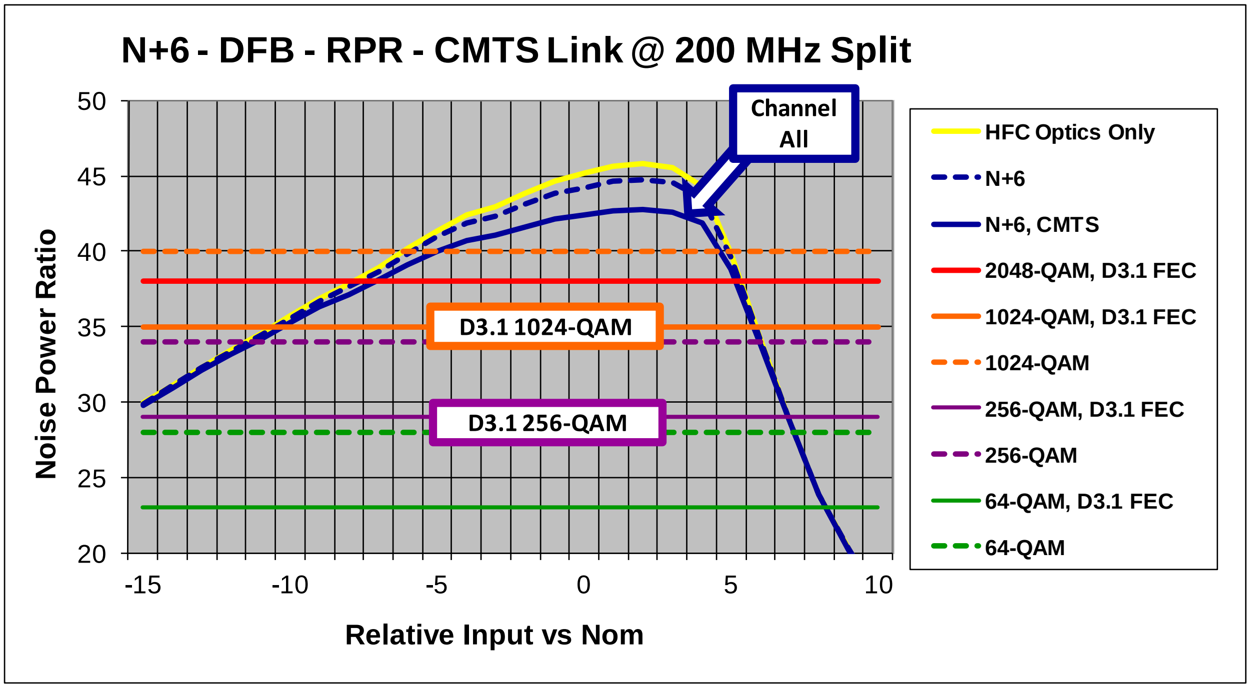


Figure 22: Extending the Upstream to 200 MHz

Encouragingly, without any improvements assumed, aside from the laser itself performing identically over the wider bandwidth, the 256-QAM mode is supported robustly despite the loading loss with plenty of dynamic range. This should not be too surprising since 256-QAM has been proven in the upstream using today's technology [5,6].

The 1024-QAM case still has sufficient DR, but it is now a borderline case at exactly 10 dB under typical conditions using today's DFB technology at nominal length. It would likely not scale in every situation as adequately robust. We can see that 2048-QAM now has insufficient dynamic range as well as a low operating margin of about 3 dB.

On the bright side, there is nothing to suggest we are out of hope expecting to get 1024-QAM (or higher) across a 200 MHz linear return link. We are within single digit dB ranges of achieving key robustness objectives – the kind of dB differences that technology evolution usually overcomes with time and development.

Based on the evolution of technology (high power DFBs), modulation profile (256-QAM), and spectrum (85 MHz proven and 200 MHz projecting well), the upstream is well along the way down the path of achieving a 1 Gbps objective and covering an expanded spectrum range with high bandwidth efficiency. Furthermore, 1024-QAM upstream appears immediately within reach with new FEC and modern DFBs, consistent with the fact that 256-QAM can be achieved today with "old" FEC.

Lastly, robust 2048-QAM from an HFC link performance does not seem like a stretch already, and is, in fact borderline acceptable from the plant SNR perspective for the 85 MHz case.

New Capacity = New Spectrum

Equations 1 and 2 were pretty clear about the role spectrum plays in finding new capacity. We know already that the best we can expect from spectral efficiency is about 50% in the downstream with 62.5% perhaps attainable eventually. In the upstream, these numbers are 67% (1024-QAM) or up to 83% (2048-QAM).

Most MSO's concern is currently around downstream because of the persistent aggressive growth. We observed this in Figure 1, and recognized how threatening this could become. Meanwhile, upstream CAGR has stagnated, putting little pressure on the urgency of relieving the inherent spectrum bottleneck of 42 MHz. As segmentation is occurring, driven by the downstream, the benefits of average upstream bandwidth per home are available to the upstream as well, assuming it is simultaneously segmented as is usually the case. However, the benefit does not actually always accrue, because this is often handled by a combining function in the HE until the traffic demands a new upstream port.

The percent new bits per second of capacity and the lifespan they represent can easily be converted to lifespan metrics through the concept of Traffic Doubling Periods (TDPs). Some simplified relationships are shown in Table 5 below.

TDP Years	**CAGR** %	**Simple** %
1.7	50	50
2	41	40
3	26	25
4	19	20
5	15	15

Table 5: Traffic Doubling Period Relationships

Table 5 is very useful for back of the envelope calculations in the range of CAGRs meaningful to cable. Obviously, if the downstream is growing at 50% CAGR and we add 50% more capacity immediately tomorrow, that step is worth about 12 months of lifespan (wow, is that all?). However, if it settles to about 40% CAGR and we add 62.5% capacity, then we add about 17 months (still, is that all?). Indeed, while these do not sound like much in isolation, this is the nature of trying to deal with the exponential (growth) with the linear (bandwidth efficiency enhancement). This is why a set of tools and techniques must be considered. For example, the picture is less ugly as we saw in Figure 1 when segmentation is included in the equation. Each segmentation is equivalent to one TDP.

And the situation is less ugly in the upstream. Even at 25% (high), a couple of node splits means 3 + 3 + 3 = 9 years of lifespan without any improvement in spectral efficiency or new spectrum. And, there is actually more spectral efficiency gain available percentage wise simply because the QAM profiles begin lower.

We will delve back into lifespan in the next section. One thing quite clear, however, is that spectral efficiency is but one part of the lifespan extension equation, and a relatively modest one at that in some cases. Node splits are incremental business-as-usual methods that deliver more average capacity as well. However, even in this case there are often diminishing returns trying to split serving groups evenly. And, it is well understood that we have hardly used the entire spectrum that can be made available on the coaxial medium. Thus, there is significant interest in finding ways to exploit new spectrum.

Based on this recognition that spectrum is critical to adding capacity in the HFC network, Figure 23 is an example of a likely long-term spectrum evolution [1] over time. A possible "final state" for bandwidth allocation on the coaxial cable is also shown, albeit with some ambiguity around the return-forward crossover band. The industry is beginning to settle around a "high-split range in the region of 200-300 MHz.

Note that by using the downstream above 1 GHz, we are extending a relatively well-behaved channel into an uncharacterized area where it will suffer more attenuation, at a minimum. We first saw this in Figure 13. However, the downstream bandwidth may need to increase, if only to offset the loss due to growth in the upstream band should it extend to 200 MHz band or greater. This downstream extension is shown in Figure 23 by the block labeled "New NG Forward." The use of OFDM has unique value in addressing this uncharacterized band, as discussed.

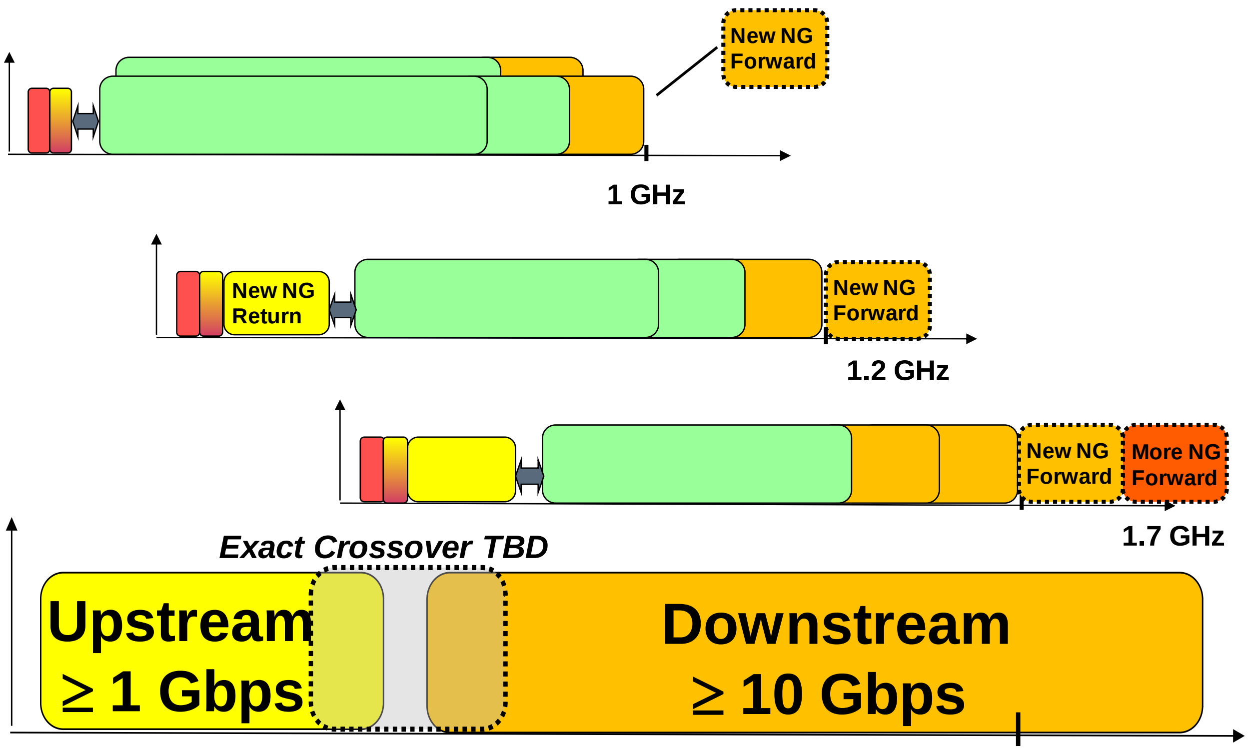


Figure 23: Possible Long-Term HFC Spectrum Evolution

In the upstream, we are instead extending a partially troubled channel into an area where we expect, in general, a better environment. The upstream today gradually becomes well-behaved with increasing frequency above about 15 MHz in North America. As we extend the spectrum above 42 MHz, cleaner bandwidth will become available, enabling more bandwidth-efficient use. The FM band is, of course, an area where characteristics may be less friendly for upstream due to funneling if we extend to 200 MHz. The implications of use of this band must be determined.

Referring to the stages shown in Figure 23, note that the upstream evolution takes place as an extension to mid-split, and subsequently an extension beyond this labeled "New NG Return." The idea is that the 85 MHz mid-split is available in current DOCSIS 3.0 and HFC technology today, and offers a very long window of upstream lifespan and service rate growth to the 100 Mbps threshold. At some point in the architecture migration, the new phase of upstream to achieve 1 Gbps can be introduced. At this point in time, perhaps due to service evolution such as IP Video and legacy removal, the downstream may be prepared to accommodate a loss in spectrum to upstream use. Otherwise, this may be the point in time to extend downstream. This appears to be the more likely scenario, due to the likely slow withdrawal of legacy services and the need therefore to simulcast, burdening downstream spectrum. Initially, the extension may simply be excess bandwidth above 1 GHz such as 1.2 GHz before evolving to a broader chunk of bandwidth exploitation above 1 GHz if necessary.

An example of the "excess bandwidth" of a single 1 GHz tap such as those very commonly deployed today is shown in Figure 24 [3].

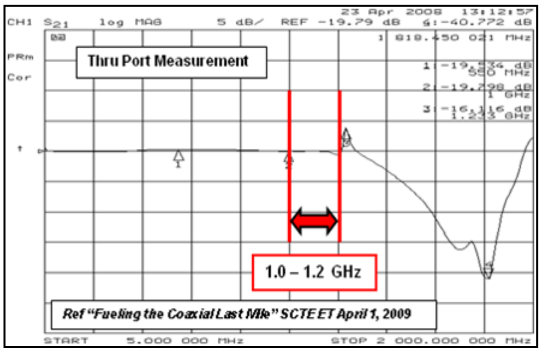


Figure 24: 1 GHz Tap "Excess" Bandwidth

Note that Figure 24 captures one single tap. In an actual RF leg, there will be multiple taps, and in an HFC cascade, there will be amplifiers and taps following a fiber optic node, and these will all cascade to create an aggregate frequency response. This is important to understand, since it is *not* the case that most actives in the field are 1 GHz. See [3] for further discussion.

Going beyond 1.2 GHz will most likely be necessary to achieve 10 Gbps of useable capacity, in particular with an extended upstream band to 200 MHz. Figure 24 makes it clear that there is not much hope for this on 1 GHz taps, especially in a cascade. Taps with wider bandwidth capability are certainly possible. However, removing old taps and replacing them with new ones is time consuming, costly, and intrusive. With a faceplate change option, the ability to convert a 1 GHz tap to a 1.7 GHz tap can occur with minimal downtime, decreasing the expense to the operator. Figure 25 shows an example frequency sweep of tap faceplate technology.

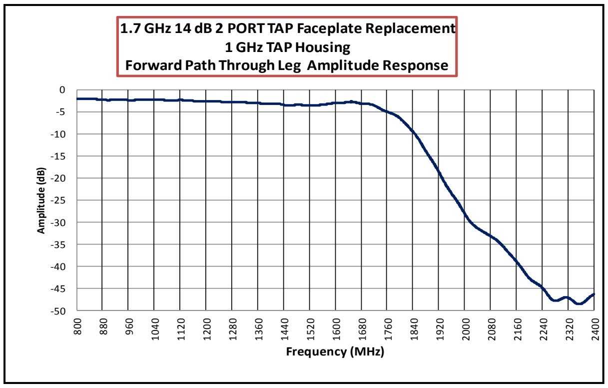


Figure 25: Creating New Forward Bandwidth with Tap Modification

How long will it take to migrate the HFCs spectrum? The period over which the spectrum evolution in Figure 23 occurs is probably on the order of 15 years. However, there is probably an equally strong likelihood that it is either 7-8 years or never. The play out of service trends and growth will have a major impact on the outcome because of the very large long-term implications of, for example CAGR remaining at 50% or settling to, say 30% and how it is trending at the time. It's the difference between 30 Gbps and 7 Gbps of aggregate capacity after ten years. It's the difference of whether spectral efficiency and service group segmentation is sufficient, or new spectrum is clearly necessary. We will discuss further using the Capacity Management Timeline in the next section.

No Free Launch

While new RF spectrum to exploit is exciting from a capacity standpoint, it has some challenges in practice, since the band where we are adding spectrum is of higher RF attenuation. There is an optical loading component as well, but since transmitter loading is flat, this added spectrum costs only about 0.8 dB relative to 1 GHz performance. Of course, 0.8 dB can be meaningful. As Figure 18 or Figure 19 indicates, for example, the shift has an impact on the percentage of users that fall into a particular M-QAM profile bucket, with the shift skewing modem distributions down and lowering the net capacity increase. It is less than 0.8 dB if we assume that the upstream extends to 85 MHz for example, but the difference is small.

However, the RF spectrum placed on the coax is uptilted for frequency dependent attenuation. Adding new RF load to the high end increases the total power disproportionately. This is easily quantified. Figures 26-28 demonstrates the issue. In Figure 26, a 550 MHz downstream of analog carriers with digital loading to 1 GHz at 6 dB digital de-rate and typical uptilt is shown. The total RF power out of the active with this loading is 70 dBmV.

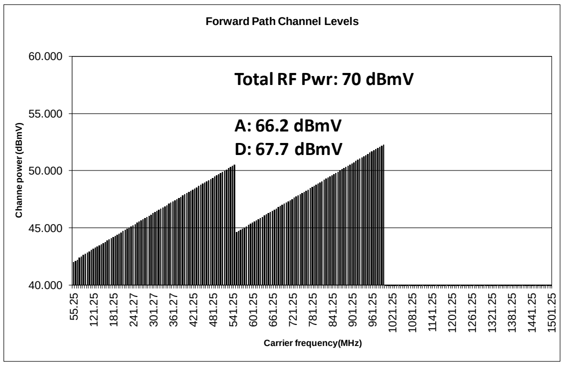


Figure 26: Typical Downstream Analog + Digital Loading

In Figure 27, the extra 200 MHz of bandwidth is loaded, and just this 200 MHz increases the total RF power load by almost double, at 2.7 dB. Fortunately, GaAs RF technology has given way to GaN technology, which is capable of higher output at equivalent distortion performance compared to typical fielded GaAs amplifiers. This has generally been engineered as higher in-band RF levels across the band by 2-3 dB or used as bandwidth extensions of older plants that maintain active spacings, but it is expected that this technology, now on its second generation, will enable 1.2 GHz at equivalent 1 GHz performance.

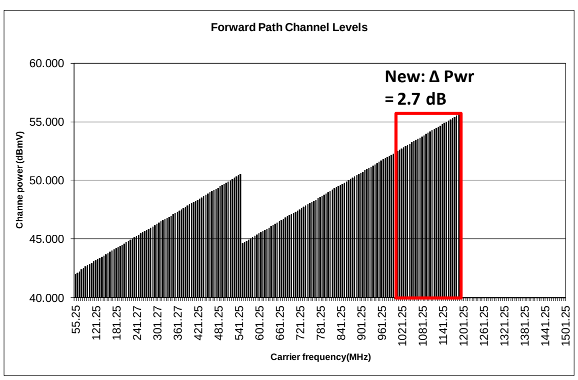


Figure 27: Tilted RF Outputs Increase Total Power Disproportionately

Now consider Figure 28, where the bandwidth is extended to 1.6 GHz. In this case, implementation of the band by simply loading the spectrum at the identical tilt and at the equivalent relative PSD is not sustainable, requiring 7.2 dB more RF power, over a 5x increase. The 5x interpretation is useful for considering the impact to the DC power budget needed to drive the additional RF power! Note that this 7.2 dB is the case even under the low end loading assumption shown in Figure 28, which is a high-split return band with a forward path that begins at around 250 MHz. Also, the optical loading for this case, should laser transmitters achieve equivalent performance parameters, is now 2.1 dB – much more significant.

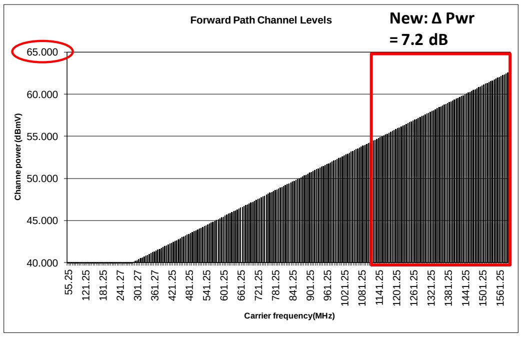


Figure 28: Downstream Bandwidth Extension to 1.6 GHz Creates an RF Power Dilemma

In summary, the addition of 1.2 GHz of bandwidth looks to be a manageable extension using current techniques for optical loading (or remote PHY) and RF distribution, including tilted RF outputs used as they are today. This is not the case at this point for enabling up to 1.6 GHz. For this case, it appears use of the band may best be served by the linear optical band extending, for example, to 1.2 GHz as part of the spectrum evolution shown in Figure 23.

Then, at some point in the future should the additional bandwidth be necessary, with the tap capability seeded into the field and the architecture migrated to N+0 so that RF actives are not an obstacle, an overlay solution that implements this additional bandwidth can be deployed, perhaps as part of a remote PHY strategy. This region is labeled "More NG Forward" in Figure 23, and as a separate overlay would also benefit from the fact that it represents a stand-alone narrowband solution.

Adding it All Up

Taking inventory of the itemized list of evolutionary techniques that offer the opportunity to breathe new lifespan into the HFC network, we have covered the following:

- Figure 1 – Analog spectrum and eventually other legacy must be assumed removed.
- Figure 2 – Employing advanced M-QAM formats will offer more spectra efficiency and more capacity over a given spectrum.
- Figure 4 – New LDPC FEC allows these more bandwidth-efficient M-QAM formats to be possible at SNRs that exist over today's HFC network performance as shown in Figure 8.
- Figure 9, Table 3 – OFDM does not itself provide "new" capacity, but its channel optimizing nature enables the full capacity of M-QAM and FEC to be obtained in all

environments and its use allows bandwidth-efficient use over an expanded RF spectrum range.

- Table 4 – HFC fiber deep evolution provides increasingly better channel SNR performance for the aforementioned M-QAM and FEC to exploit.

- Figure 17 – Fiber deepest (N+0) maximizes the channel SNR delivered over HFC *and* opens up the opportunity to expand spectrum above 1 GHz for fresh, new capacity. It service group shrinking benefits produce over 90% aggregated BW savings.

- Figure 18 – Analysis of fiber deep architectures combined with home architecture evolution via a POE gateway approach shows 1024-QAM through 8192-QAM within reach, with 8192-QAM as a potential maximum bandwidth efficient profile – a 62.5% increase over 256-QAM in spectral efficiency.

- Figure 19-20 – Switched broadcast allows CMs to be classified into buckets based on their channel quality. This increases average capacity, maximizes a user's experience, and allows the operator to eliminate the waste of dB dedicated to margin associated with a lowest common denominator single broadcast profile approach.

- Figure 21-22 – Modern DFBs, high sensitivity DOCSIS 3.0 receivers, shorter cascades, elimination of HE combining, advanced QAM profiles (Figure 2) and new FEC (Figure 4) do for the upstream what similar M-QAM, FEC, and architecture evolution do for the downstream. 64-QAM to 256/1024/2048-QAM (33-83%) is realistic.

- Figure 23-25 – Spectrum evolution frees up room for both downstream and upstream growth. Downstream 1.2 GHz or 1.6 GHz from 950 MHz or 500 MHz total is a 20%-90% increase. The upstream spectrum increase is over 100% (85 MHz) to over 300% (200 MHz).

We can lay these evolutions out on a downstream Capacity Management Timeline as in Figure 1. We can perform a similar Capacity Management analysis for the upstream. We will actually start with the upstream, since a key conclusion from that translates into details of the downstream Capacity Management Timeline calculations. And, as indicated, it is the less urgent of the two at this point for most operators and is generally managed effectively in concert with downstream segmentation.

Lifespan Management – Upstream

Similar to the calculations in the downstream, we can project how long our upstream will last under various growth and service scenarios. We base the projections on 80 Mbps deployed today, such as two 64-QAMs and one 16-QAM at 6.4 MHz.

Unlike the downstream, CAGR for the upstream tends to be less predictable – spiking when Napster and YouTube were introduced for example, and lagging for several years during other periods of time, including the last several years. Because of this and because of the hard limit of upstream spectrum with fewer tools to manage it than the downstream, we can create a form of the analysis that is simpler to interpret. This is shown in Figure 29. With a typically more variable CAGR range, it also handily provides a sensitivity analysis perspective around incorrect CAGR projections.

Figure 29 reveals why both an expansion beyond 42 MHz is not as urgent as it might seem it should be. The recent CAGR range is identified, and it has been relatively modest. Referring to Table 5, even the 25% CAGR has a three year doubling period, so with two node splits along the way in concert with the downstream, six years of latent growth lifespan is available even if fully congested today. This becomes 8 years for the 20% CAGR. This is readily apparent from the nearly full 100 Mbps ATDMA solid yellow trajectory, and the 9-year lifespan after two splits that exists for it. The timing of course, becomes the key to those splits. Figure 29 results are best synchronized with Figure 1 for the purpose of the timing of the node segmentations.

Using SCDMA and two splits and a 20% CAGR, more than a 10-yr window is available, with the first split lasting out to 7 years.

With an 85 MHz mid-split, DOCSIS 3.0 technology with a 20% CAGR provides 7 years of life, and with two splits this extends to 15 years! It is for this reason, and that this bandwidth can support a service rate of 100 Mbps – not practical over 5-42 MHz return – that an 85 MHz is potentially a long-term upstream bandwidth phase, and not an “interim” one.

Lastly, with a 1024-QAM capable upstream (red), even a 25% CAGR has 15 years of aggregate capacity life with an 85 MHz mid-split. There is almost a full decade using an aggressive 40% CAGR upstream. Again, this is pretty simple to get a feel for from Table 5: 40% is approximately a 2-yr TDP and 85 MHz will be on the order of 550 Mbps of capacity. 80 Mbps will not quite be three TDPs of 2 years – or slightly less than 6 years. Two splits are two TDPs or 4 years, so slightly less than 6 + 4 or slightly less than 10 years total.

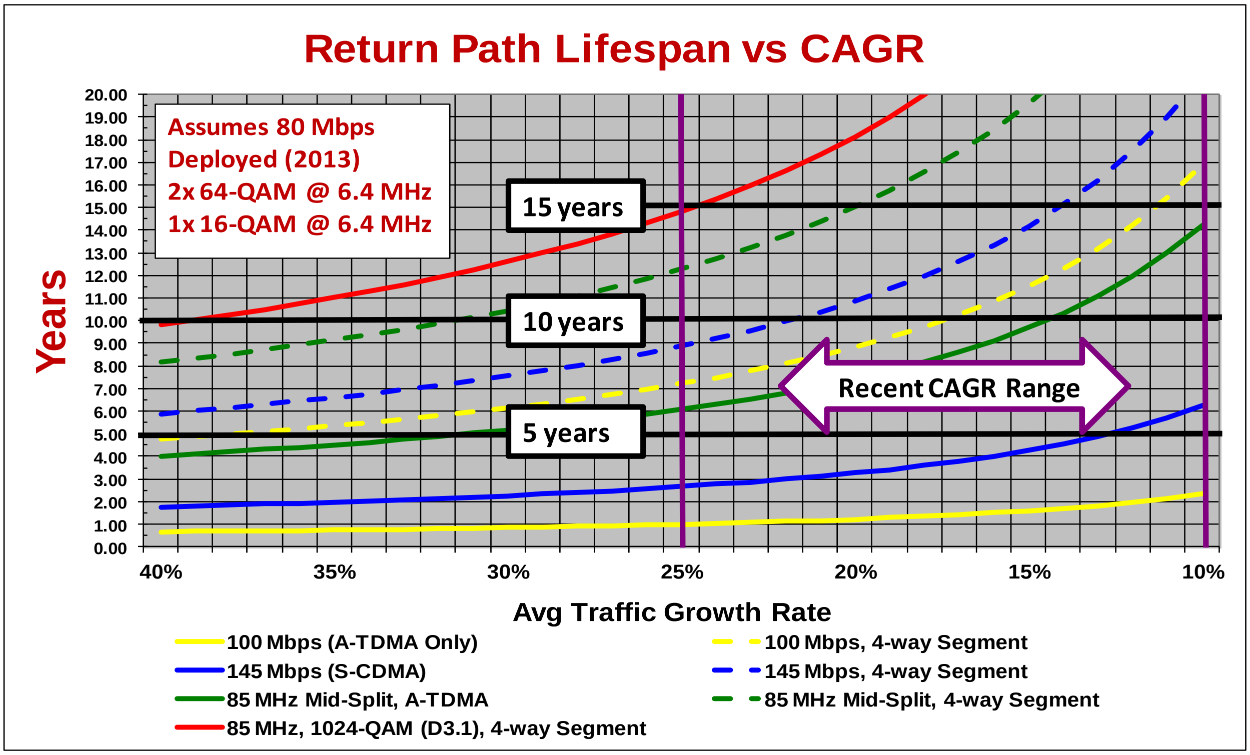


Figure 29: Upstream Capacity Management

At nominal CAGRs of today, should that continue for many years, a capacity-enhanced 85 MHz upstream offers so much room for new growth in terms of lifespan that it does not make very much sense to plan what the next step should be after it. It does of course, make sense to keep a watchful eye on nominal CAGRs. Unfortunately, the 85 MHz mid-split cannot achieve the 1 Gbps DOCSIS 3.1 objective. However, because of legacy constraints around STB signaling, and the squeeze that extended upstream spectrum places on the downstream, it may be unwise to consider a broader extension for some time in any case.

To put 1 Gbps in a CAGR perspective, it is four TDPs to exceed the threshold from 80 Mbps. At 4 years per TDP at today's CAGRs, this means 16 years to the need date from an average consumption standpoint.

A path that first evolves to 85 MHz, covering well over a decade of new capacity, delivering 100 Mbps speeds, and maintaining legacy out-of-band signaling is the most prudent phased approach. In so doing, technology to later enable a very simple switch to a wider band upstream should be deployed to minimize a subsequent costly plant touch, albeit the need for this capacity in the upstream for residential services looks well into the next decade, and many of the actives that include such technology may be removed by fiber deep evolution by the time of need.

Downstream Lifespan: Worth the Price?

An updated version of the Figure 1 downstream Capacity Management Timeline is shown in Figure 30. Because of what we have learned examining the upstream, the spectrum used for forward band allocation is an 85 MHz mid-split, and therefore a 108 MHz forward band starting point. Let's examine what we have achieved between Figure 1 and Figure 30. Note that most of the Figure 1 data is shown on Figure 30, but we have simply extended the timeline for another decade in recognition of, and to evaluate, the additional lifespan we have enabled. To reduce clutter, we only show one of the "All-QAM" thresholds of 256 QAM, and chose the 870 MHz middle case.

Today's downstream spectrum is basically full. Just ask any MSO. Technically, there is no "room for growth." Managing from this completely full state through the balancing act of old services and technology to new services and technology is the challenge addressed in detail in [4,10]. As mentioned in the first section, the objective of this paper was to see how far north we could shift capacity thresholds on the chart given the list of the architecture and technology variables that we are throwing at the problem, and what that buys us.

Using 870 MHz as an example we can make the following assessment based on the blue "IP Video" trajectory (little difference from red long term) and the green "Max Consumption" trajectories, comparing the lifespan ranges. This is shown in Table 6.

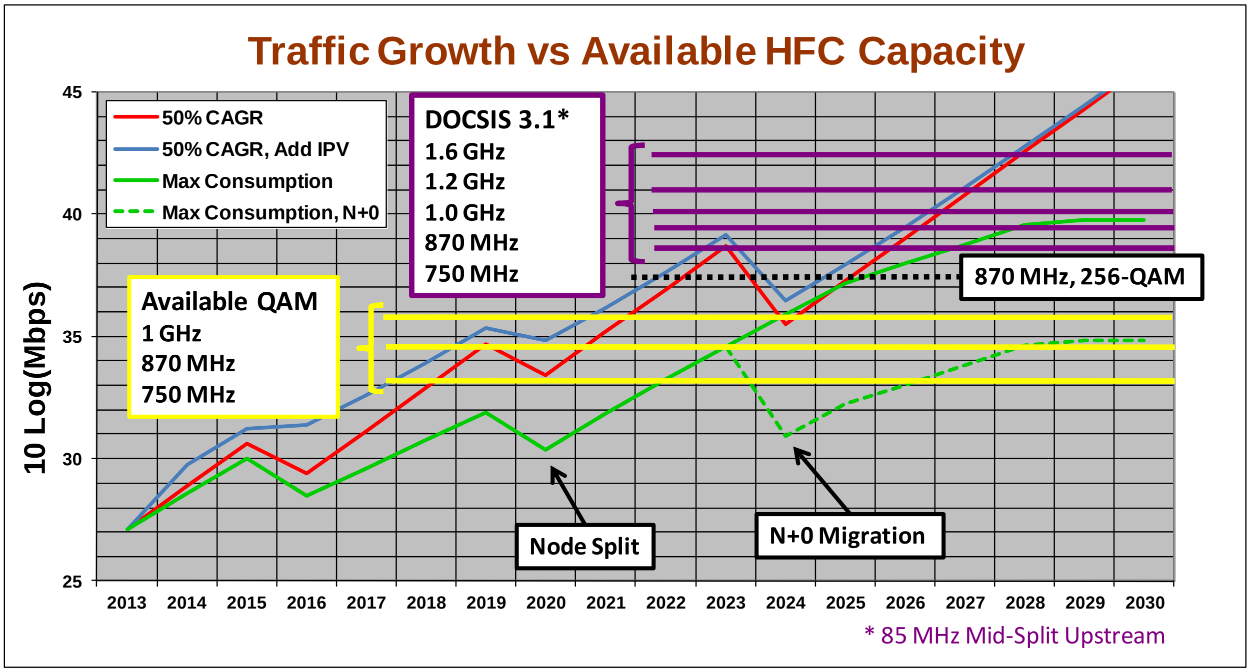


Figure 30: Capacity Management Timeline: Breathing New Lifespan into the Network

	"Available" 870MHz	"All QAM" 870MHz	"New QAM" 870MHz	"New QAM" 1200MHz	"New QAM" 1600MHz
50% CAGR w/IPV					
Two Splits	7	9	10	11.5	13.5
N+0	10	12	13	14	15
Settled CAGR					
Two Splits	10	12.5	14	Sufficient	Sufficient
N+0	Sufficient	Sufficient	Sufficient	Sufficient	Sufficient

Table 6: Capacity Management Timeline: Years of HFC Lifespan Starting with 870 MHz

The lifespan increases shown in Table 6 are significant, though not necessarily entirely comforting. Focusing on the top row of Table 6, from the "today" case of "Available QAM," the increases are as small as 3 years with a maximum of 8 years result for the 50% CAGR case. There are a pretty significant number of possible expensive steps invested in to achieve those 8 years. It does not feel like an enormous bang for the buck. The differences are even smaller when compared to the Gbps available in a full spectrum of 256-QAM – "All QAM" – at 4.5 years maximum.

This really just comes down to the implications of the relentless nature of the long term aggressive CAGR assumption, if such assumptions turn out to hold. Linear gains to combat exponential processes make it difficult to keep pace. We will consider the alternative in the next section. However, if we convert the absolute years of Table 6 to lifespan percentage gained, at least the perspective looks, perhaps, more in line with expectations.

In Table 7, we tabulate this percent lifespan gain for the three cases of advanced QAM efficiencies for the three spectrum cases relative to "Available QAM" and relative to the capacity of an 870 MHz forward band full of 256-QAM carriers ("All QAM").

	"New QAM" 870MHz	**"New QAM" 1200MHz**	**"New QAM" 1600MHz**
"Available" QAM			
Two Splits	43%	64%	93%
N+0	30%	40%	50%
All QAM			
Two Splits	11%	28%	50%
N+0	8%	17%	25%

Table 7: Capacity Management: Percent Lifespan Extension, Persistent CAGR

For the persistent CAGR case, the combination of tools provides percent lifespan gain that in some of the cases is compelling – from 30% but to as high as over 90% – using what is defined as "Available QAM" as a starting point. This term basically boils down to executing on all the following bandwidth efficiency tools:

(1) Analog reclamation

(2) New Spectrum

(3) New Modulation Profiles

(4) Service Group Segmentation

The "All QAM" percent lifespan extension is much more modest. This basically represents the above steps, but not including the analog reclamation as part of the advantage gained. This should not be surprising based on the prior discussion on TDPs and CAGRs. If we can manage 50% new bandwidth efficiency gain, but we are unable to add any actual new bandwidth, than a 50% growth in traffic in a year will consume that gain immediately. This is basically what the one year gained in the columns for "All QAM" 870 MHz and "New QAM" 870 MHz tell us. It reemphasize that new capacity for new lifespan relies on *a set of evolutions* for new technology to combat exponentially growing traffic. Part of this is absolutely the critical removal of existing less efficient legacy technologies, and in particular much less efficient analog services.

In absolute terms, we can see that the most "Northbound" thresholds of 1.2 GHz and 1.6 GHz carry us into the middle of the next decade with two node splits, and late into the next decade for N+0. While the years added in Table 6 do not seem particularly large in most cases, this is an extremely long window of time from a technology standpoint – too far out to do much in the way of strategic planning for capacity management. It is actually a very beneficial observation window for trends in service growth that impact subsequent migration decisions.

Table 6 and Figure 30 prove the obvious – "CAGR forever" always wins the bandwidth battle – it is just a matter of the time scale chosen to find defeat. This is, of course a key reason why attention to service growth trends during the window of the next decade or so will be critically important to pondering future architectural evolution. In the case of continued persistence, the only step left likely to be of sufficient ROI is FTTH.

Asymptotic Growth

Perhaps the more intriguing Capacity Management Timeline result and conclusions they imply are the cases of "Max Consumption." These trajectories, drawn from analysis in [7], are based on the premise that streaming video as the driver of CAGR arrives at a quantifiable maximum video bit rate and demographic-based concurrency per-household peak. If this is so, then the CAGR associated with streaming video will taper the traffic growth towards this maximum over the course of time that these video evolutions take place. The green trajectories demonstrate this hypothesis, which effectively suggests that 50% CAGR forever is not practical to sustain, with other reasoning for this settling also described in the first section.

The conclusions for these possibilities are intriguing. In the case of N+0, there will be sufficient bandwidth as long as the network is extended to the 1 GHz point – a standard HFC solution available today. And this is the case with simply 256-QAM available!

If instead only two splits are used, and no extension to N+0, then we coincidentally arrive at the same answer of 1 GHz minimum of spectrum, but in this case we must take advantage of increased bandwidth efficient modulation formats discussed herein. Importantly, though, the extension to either 1.2 GHz or 1.6 GHz is not required under this assumption of asymptotic CAGR, at least not for the end state – but perhaps for managing the "simulcast bubble." These cases of sufficient bandwidth are labeled as such in Table 6.

Moreover, when the settled CAGR cases (about 35% average is used) *do* represent enough growth to breach a threshold of capacity, they are nonetheless extending the lifespan by a meaningful, if not huge, 3-4 years, and this is with just a relatively modest CAGR setting. This is evidence that the power of compounding, while potentially very threatening, can also be quite forgiving if its most aggressive tendencies settle.

Conclusion

The persistence of compounding data growth demands that the industry respond in order that its service delivery capabilities are not compromised. Several avenues of evolution are available, and all may need to be deployed in order for the network to be a sustainable service delivery platform for the long-term. In composite, they put cable on a path to match or exceed PON targets capacity and bandwidth per user. We have observed that only business-as-usual plant segmentation, part of any logical plan, may buy time but appears insufficient to ensure a comforting lifespan for MSO business planners based on growth trends today. We introduced the Capacity Management Timeline as a way to visualize and plan service growth and architecture investment.

We have looked at each of the key avenues available for capacity growth in detail. To increase capacity, the answers are simple as Shannon has shown in (1). We need more bandwidth and higher SNR.

With higher SNR, we can use more spectrally-efficient modulation profiles. To most aggressively exploit these profiles, we need Forward Error Correction that maximizes the profile we can use for a given SNR. To maximize SNR, we need architectural evolutions that deliver the highest performance from the HFC channel, which means taking the fiber as deep as possible – and such that we advantageously also share it over fewer subscribers.

Also, we need more bandwidth to operate these high-efficiency M-QAM profiles over. There are two components to more bandwidth – removing the legacy services and finding new spectrum.

In the paper, we described how all of these technologies and evolutions mesh with the cable network and quantifiably described the possibilities they create for new capacity.

Armed with these tools, we put together a longer-term picture of the Capacity Management Timeline. We determined that the upstream appears currently less threatening, and quantified the lifespan possibilities for the fully evolved downstream. We observed the extension of lifespan we might expect under a persistently aggressive CAGR with our new capacity possibilities, and learned the obvious – relentless CAGR always wins, only the time of defeat changes.

However, we also learned under specific assumptions that we can add up to 8 years on the lifespan – a long observation window to see just how persistent this CAGR beast is. And, perhaps most interestingly, we learned that with some modest CAGR settling, the HFC lifespan outlook clears up nicely, and offered reasoning why this could be the true future of cable service delivery.

Furthermore, we learned that with settled CAGR and N+0 evolution, "HFC forever" emerges. And even the two split case does not necessarily require spectrum expansion in the long term. An end state that survives indefinitely appears achievable, but we simply cannot be certain until we have recorded more years of service evolution and CAGR history. If so, then perhaps the greatest challenge will have turned out to be the challenge we face today – surviving the simulcast bubble by managing service and traffic growth on top of old services that will prove difficult and complex to retire.

Acknowledgements

The author would like to thank the following individuals for their contributions to the content of this paper, providing invaluable technology and industry insights:

Rob Thompson

Jack Moran

Chris Jeblonski

Phil Miguelez

John Ulm

Fred Slowik

Mike Emmendorfer

John Chapman (Cisco)

Dr. Tom Kolze (Broadcom)

David Urban (Comcast)

John Chapman (Cisco)

Matt Schmidt (CableLabs)

References

(1) Chapman, John, Mike Emmendorfer, Dr. Robert Howald, Shaul Shulman, *Mission Is Possible: An Evolutionary Approach to Gigabit-Class DOCSIS*, 2012 Cable Show, Boston, MA, May 23-25.

(2) Howald, Dr. Robert L, *Boundaries of Consumption for the Infinite Content World*, SCTE Cable-Tec Expo, New Orleans, LA, October 20-22, 2010.

(3) Howald, Dr. Robert L, *Fueling the Coaxial Last Mile*, SCTE Conference on Emerging Technologies, Washington DC, April 2, 2009.

(4) Howald, Dr. Robert L, *Looking to the Future: Service Growth, HFC Capacity, and Network Migration*, 2011 Cable-Tec Expo Capacity Management Seminar, sponsored by the Society for Cable Telecommunications Engineers (SCTE), Atlanta, Ga, November 14, 2011.

(5) Howald, Dr. Robert, *The Grown-Up Potential of a Teenage PHY*, 2012 Cable Show Spring Technical Forum, Boston, MA, May 21-23.

(6) Howald, Dr. Robert L, *Time on Our Side: Using the Past to Pave the Way Forward*, The Cable Show Spring Technical Forum, May 20-22, Boston, MA.

(7) Howald, Dr. Robert L, Dr. Sean McCarthy, *Accounting for Techies: Taking it to the Ultra*, The Cable Show Spring Technical Forum, June 10-12, Washington, DC.

(8) Howald, Dr. Robert L, Dr. Sean McCarthy, *Bits, Big Screens and Biology*, The Cable Show Spring Technical Forum, May 20-22, Boston, MA.

(9) Howald, Dr. Robert**,** Michael Aviles, and Amarildo Vieira, *New Megabits, Same* Megahertz: Plant Evolution Dividends, 2009 Cable Show, Washington, DC, March 30-April 1.

(10) Howald, Dr. Robert L. and John Ulm, Delivering Media Mania: HFC Evolution Planning, 2012 SCTE Canadian Summitt, March 27-28, Toronto, ON, Canada.

(11) Ulm, John and Gerry White, *Architecture & Migration Strategies for Multi-screen IP Video Delivery***,** 2012 SCTE Canadian Summit, March 27-28, Toronto, CA.

(12) http://maninair.in/orthogonal-frequency-division-multiple-access-ofdma/

(13) http://www.dipol.ie/newsletter2/inf_dipo_2009_32.html

(14) http://www.ict-redesign.eu/fileadmin/documents/ReDeSign-D08_Report_on_Cable_Channel_Model.pdf

(15) http://www.ieee802.org/3/bn/public/adhoc_mmp/adhoc_mmp_urban_01.pdf

(16) http://www.ieee802.org/3/bn/public/mar13/howald_3bn_01_0313.pdf

(17) http://www.ieee802.org/3/bn/public/mar13/howald_3bn_02_0313.pdf

(18) http://www.ieee802.org/3/bn/public/nov12/schmitt_01a_1112.pdf

(19) http://www.lightwaveonline.com/articles/print/volume-26/issue-9/industry/ofdm-to-power-high-bit-rates-in-next-gen-optical-access-networks.html

(20) http://www.ni.com/cms/images/devzone/tut/a/bdc8dcae1315.gif

Appendix –Updated HFC Channel Model [16]

The IEEE 802.3bn Working Group has adopted an HFC Baseline Channel Model for purposes of developing a system design for Ethernet PON over Coax (EPoC) technology. This channel model is included here for reference. While cable channel models have been developed in the past, a primary focus for this baseline was to arrive at "Typical" values for channel parameters, rather than exclusively limit or worst-case parameters. While the limit cases are also part of the effort, it was deemed very important for the task at hand – how much capacity can be squeezed out of the HFC network – to better understand the network's ability to "scale up" in capacity.

The table below is to allow simulation teams to develop and optimize system designs suited to maximizing link capacity. Other scenarios are shown in [17].

Notes	
1	If not defined otherwise, assume typically behaving link but where the behavior is the worst (freq, location)
2	Frequency dependence of coax for broadband calculations: Loss B (dB) = Loss A (dB) x SQRT(B/A)
3	Reference virtual port level for 6 MHz signal at 1 GHz; 15 dBmV Tap port level, 100 ft drop, 2-way splitter
4	(Max Freq - OFDM BW) spectrum range used for drop loss
5	Small drop slope effect on calculation
6	SCN includes HFC geography impact (location in cascade depth)
7	50 kHz Subchannel Reference, Live Video, fully contained within subchannel
	Subcarriers with Interference (50 kHz): Every 70 subcarriers, a cluster of three interfers: I_0, I_{0+}25 kHz, I_0-25 kHz
8	Typ = CTB/CSO Worst Case Freq; Good CTB/CSO in low-distortion band, Analog contiguous at low end of band
	NCTA measurement method (avg); Error rate simulation should account for PAR and peak durations
9	Worst spectrum regions for CTB and CSO are not the same
10	D/S Burst Characterization in process; BW based on percentage of errored carriers in 8-Channel wide DOCSIS CM
	Duration based on large scale CM sweep of UCER with known interleaver settings; Levels per ReDesign channel model
11	Laser Clipping PSD captured in SCN for out-of-band EPoC Signals
12	Typical tilt, first tap, not equalized, 50 ft drop assumed (Minimum drop impact)
13	Echo mask range for a Single Dominant echo - Does not imply an assumptions about multiple echoes.
14	Meas@700-800 MHz, representive of 99% of modems

ARRIS

System Description				
HFC D/S Spectrum		1.0 GHz		
Cascade Depth		N+3		
Channel Loading		48 Analog (32 removed for D3.1) + 75 Digital		
Optical Architecture		Linear Optics 1310 nm (nominal link length) Legacy		
Home Architecture		Up to max drop length & 2-way splitter		
	#	**Parameters**	**Baseline Channel**	**Notes/Dependency**
Spectrum	1	**Frequency range**	54 MHz - 1 GHz	Note 1
	2	**OFDM Bandwidth**	192 MHz	
RF Level	3	**OFDM Power at CPE Input (dBmV)**		Notes 2-4
		6 MHz BW	-2	
		24 MHz BW	4	
		96 MHz BW	10	
		192 MHz BW	14	Note 5
SNR	4	**SCN Ratio (Signal to Composite Noise Ratio)**	44	Note 6
		Variation over 6 MHz BW (dB)	N/A	Reference Basis 6 MHz
		Variation over 24 MHz BW (dB)	1.5	
		Variation over 96 MHz BW (dB)	2.5	
		Variation over 192 MHz BW (dB)	3.0	
Interference				
Narrowband	5	**CTB Interference (20 kHz BW)**		Notes 7, 8
		# of interfered subcarriers @ 35-40 dBc	0%	
		40-45	1%	
		>45	0%	
	6	**CSO Interference (20 kHz BW)**		Note 9
		# of interfered subcarriers @ 35-40 dBc	0%	
		40-45	0%	
		45-50	2%	
		>50	0%	
	7	**Narrowband Interference (Other)**		
		Bandwidth (MHz)	N/A	
		Level, dBc (PSD)	N/A	
Wideband	8	**Burst Interference**		Note 10
		Bandwidth (MHz)	30	
		Level, dBc (PSD)	-20	
		Duration (usec)	16	
		Period (Hz)	Infrequent	
	9	**Impulse (white) Noise**		Note 11
		Level, dBc (PSD)	N/A	
		Duration (nsec)	N/A	
		Period (kHz)	N/A	
Freq Response				
Amplitude	10	**Amplitude Slope**		Note 12
		dB/MHz	0.01	
	11	**Amplitude Variation**		SCTE Definition, Echo not included
		(dB pk-pk/6 MHz)	1	
		(dB pk-pk/24 MHz)	3	
		(dB pk-pk/192 MHz)	5	
		(dB pk-pk/Total DS BW)	9	
Phase	12	**Group Delay Variation, nsec**		
		Over 24 MHz		
		Mid Band	25	
		Band Edge (24 MHz)	145	
		Over 192 MHz		
		Mid Band	200	
		Band Edge (24 MHz)	320	
Echo	13	**Echo Profile, dBc**		Notes 13, 14
		.5 usec	-20	
		1 usec	-25	
		1.5 usec	-30	
		2 usec	-35	
		3 usec	-40	
		4.5 usec	-45	
		5 usec	-50	
Spurious Modulation	14	**AM/Carrier hum modulation %**	3%	

Upstream

Note this Table, by choice of priorities of interest, represents a remote demodulation architecture (no linear optical return).

System Description				
HFC U/S Spectrum	300 MHz			
Node Architecture	N+3			
Channel Loading	Remote Tx/Rx			
HE Architecture	N/A - EPON Return			
Premise Architecture	Two Way Combining			
	#	**Parameters**	**Baseline Channel Conditions**	**Notes/Dependencies**
Spectrum	**1**	**OFDM Bandwidth**	192 MHz	
	2	**Frequency range**	100-292 MHz	
Path Loss	**3**	**Path Loss (dB)**	44	Max loss to first active
		Variation Freq, 24 MHz BW	1	Note 1
		Variation Freq, 96 MHz BW	2.5	
		Variation Freq, 192 MHz BW	5	
Added Noise	**4**	**Input Noise PSD**	- 115 dBmV/Hz	Contibutions of
Interference	**5**	**FM Band Interference**		
Narrowband		Bandwidth	8	Spectrum Overlap
		Level, dBc (PSD)	-40	Note 2
	6	**Common Path Distortion**		
		dBc	N/A	
		% effected subcarriers	N/A	
	7	**Other Bands**	TBD	New Upstream spectrum
		dBc	-50	Note 3
		% effected subcarriers	1	50 kHz subcarriers
Wideband	**8**	**Burst Interference**		Note 4
		Bandwidth (MHz)	TBD	Non-white characteristics (Note 5)
		Level, dBc (PSD)	0	
		Duration (usec)	1	
		Period (Hz)	1000	
Freq Response				
	9	**Amplitude Slope**	N/A	Captured in Path Loss Range
Amplitude	**10**	**Amplitude Variation**		not included
		(dB pk-pk/24 MHz)	1.5	
		(dB pk-pk/96 MHz)	2.5	
		(dB pk-pk/192 MHz)	3	
Phase	**11**	**Group Delay Variation (nsec)**		
		Over 24 MHz		
		Mid Band	25	
		Band Edge (24 MHz)	280	
		Over 48 MHz		
		Mid Band	50	
		Band Edge (24 MHz)	305	
		Over 192 MHz	575	
Echo	**12**	**Echo Profile, dBc**		Note 6-7
		.5 usec	-16	
		1 usec	-22	
		1.5 usec	-29	
		2 usec	-35	
		3 usec	-42	
		4.5 usec	-51	
		5 usec		
Spurious Modulation	**13**	**AM/Carrier hum modulation**	5%	
	Notes			
	1	Path Loss adopted for consistency although return path, although RF actives include upstream gain		
	2	Measured samples in MSO location of high field strength environment		
	3	Projected (for 50 kHz) from acceptable D/S interference level for analog video band (now		
	4	U/S burst characterization in process; Ref CableLabs 1997 Report "Characterization of Upstream Transient		
	5	No linear optical return - no U/S Laser Clipping (white)		
	6	Measured Upstream CM (97% criteria) extrapolated to band (30 MHz measured to 100 MHz)		
	7	Echo mask range for a Single Dominant echo - does not imply an assumptions about multiple echoes		