Boundaries of Consumption for the Infinite Content World

SCTE Cable-Tec Expo®
New Orleans, LA
October 20-22, 2010

Dr. Robert L. Howald
CTO Office
Motorola Home & Mobile

Abstract

Cable operators have continued to exploit the flexibility of the HFC architecture as services have expanded in breadth and capacity usage. The diverse tools available have enabled the architecture to evolve to capably support multiplicative year-on-year traffic growth. We have shown in prior published analysis that, using available tools and unexploited capacity, there is still significant headroom left in HFC to deliver yet more downstream growth, extending the life of the network for many more years. In addition to this positive outlook, the state of fiber deep migration has put operators in the enviable position of having multiple evolutionary options moving forward as services continue to grow. In particular, considering traffic growth timelines, technology trends, and cost implications, it is reasonable to now consider whether architecture evolution means the continuation of fiber deep extensions, augmenting the network with new access technologies, or replacing HFC altogether with an FTTP technology. Each option has important legacy support, cost, and growth projection implications.

This analysis will focus on the service and traffic growth component of the equation, the trends of which are a key element used by MSOs to draw conclusions about the above-mentioned migration alternatives. However, we will instead look at residential media consumption-based services from the top down. That is, rather than using historical growth curves to predict the future, we calculate a potential consumption endpoint based on how much and what kind of media might accumulate in a world limited by what can reasonably be consumed, and at a level at which it could most aggressively be implemented in a differentiated fashion. This estimate of a "full consumption household" includes simple components, such as family sizes and screen counts, and more complex physiological, and behavioral components. These more scientific components include the resolution of the human visual system and its impacts on quantization of pixels, encoding, and refresh rate. It also includes the growth and capability to meaningfully multi-task and multi-consume. Finally, it considers the scientific/behavioral element of screen sizes, large and small. From this top-down perspective, we can project what evolutionary time span aligns with these limitations. We can re-assess network migration alternatives from an ability-to-deliver standpoint, and understand critical architecture thresholds. And, we can ponder what non-media consumption-based services might drive new growth.

I. Introduction

Decades of broadband growth and an ever-increasing range of video services has given operators a sound historical basis upon which to base future growth trends, which is critical for business planning. Service growth and subscriber satisfaction with the portfolio of media brought to their doorstep provides new revenue opportunities. However, understanding what is required to enable new growth and those personalized, sophisticated experiences determines the success of the overall business venture. MSOs must constantly be evaluating the capital expenditures (CAPEX) necessary, and the return on investment (ROI) they provide. In addition to CAPEX, operator choices have an effect on operating expenditures (OPEX) and efficiencies. Key decisions must be made for upgrading hubs, homes and access networks, each of which has different implications to equipment and maintenance spending, and planning the operation. The prevailing MSO approach has been a very successful pay-as-you grow – investing just-in-time to capitalize on technologies as they mature to cost effectiveness and reliability, and as the consumer demand requires. This has worked extremely well because of the latent capacity in the initial HFC builds, which incrementally was mined as necessary by extending fiber, adding RF spectrum, adding WDM optical technologies, and delivering digital and switched services.

It is natural to wonder whether the upward trajectory of broadband bits-per-second to the home continues unabated. Historically, applications have evolved to fill newly created bandwidth or accelerate new bandwidth requirements forward, often without much warning or prediction of the application itself. The trend of a ~50% per year compound annual growth rate (CAGR) seem relentless. Can it continue forever? Will we actually be providing Terabit per second (Tbps) services in the 2030s? Can we predict what might require such bandwidth, considering we have not been able to do so very well in the past? Is there, in fact, an asymptotic endpoint to this growth? Well, answers to such questions may be little more than guesses. However, what if we instead ask – can we at least draw some boundaries around what we understand that *could* limit consumption? In the case of media consumption, there are some mathematically quantifiable limits to consider. Basically, our consumption of video media is limited by the fact that we are human beings with all of the physical and behavioral constraints that involves. The creation of high-quality video and processing thereof is encroaching on a point where "better" video may not be possible in terms of subscribers recognizing or perceiving a difference in the experience. Combined with demographic and behavior factors, this is what will be explored in the paper.

With broadband growth trajectories regularly thrashed about in the technical community and blogs, and regular reporting of statistics of Internet and consumer traffic, why care about quantifiable limits? The simple answer is to see what networks could be in store for in terms of storage and bandwidth and compare the orders of magnitude. The years out make it unreasonable to use the word "plan," but it would seem helpful to not have a continuous upward trajectory with a question mark at the end if we can do better. As the cable access architecture continues to see deeper and deeper fiber deployment, as telco providers build out xPON architectures, and as FTTP has entered the MSO space in the forms of commercial services and RFoG technology, it is now reasonable to take this long - perhaps ultimate - view look.

Note that we explicitly state that we are calculating *media consumption* limits. This is because it is very likely that the subsequent era "Moore/Nielsen Law" growth will be something other than residential media consumption, the video component of which is driving the current phase of growth, on the heels of the peer-to-peer phase. Interestingly, according to "The Law of Accelerating Returns" [7], Moore's law of computing power is in fact the *fifth* exponential growth paradigm in the history of computing – we just happen to be in the growth window of time that it becomes widely observable by the masses. Other non-media consumption applications, some of which already are underway, include machine-to-machine communications, healthcare-related services, and more sophisticated artificial and integrated intelligence technologies [7].

II. Projected Traffic Growth

A. A Wrinkle in Time

The broadband growth trajectory has kept up its relentless pace in the current era because of video. Pictures begat low resolution video, begat YouTube, begat Hulu. Nowadays, IPTV delivery over the cable network is part of virtually every MSO's plan, with the only unresolved questions being "when" and, operationally "how." It all has happened quite quickly, though there is much hard work in the trenches that goes unobserved that leads to these breakthroughs.

Figure 1 highlights the critical inflection point behind the growth. It shows the convergence of the two key trends – data rate provided and video rate required – that created the impetus to begin rethinking the delivery architecture for video. Some would say it marks the inevitability of video distribution, much like voice, becoming an IP-based service over the access network. IP distribution in the regional architectures, Headends, and home networks is quite typical already. In the case of video, IP delivery on the HFC access network would necessarily be over the DOCSIS® infrastructure today. Future technologies that exploit more of the coaxial spectrum may deviate from this simply to absorb the continuous growth projections. In any case, it is the continued advance of DOCSIS data rates, shown in Figure 1, crossing paths with delivery rates for standard definition video, coincident with the emergence of MPEG-4 part 10 (H.264 AVC, or simply H.264) encoding that increased the momentum for IPTV over cable.

Note that this crossover occurred roughly five years ago. At that time, of course, H.264 was less mature, data services had been engineered primarily for static webpage delivery and "near" real-time peer-to-peer file transfers, and DOCSIS 3.0 was incomplete. In addition, newer adaptive rate technologies aimed at enhancing the streaming video experience were yet to be developed. As such, we could mostly treat ourselves online to very poor quality, stop-and-go, limited action content sources.

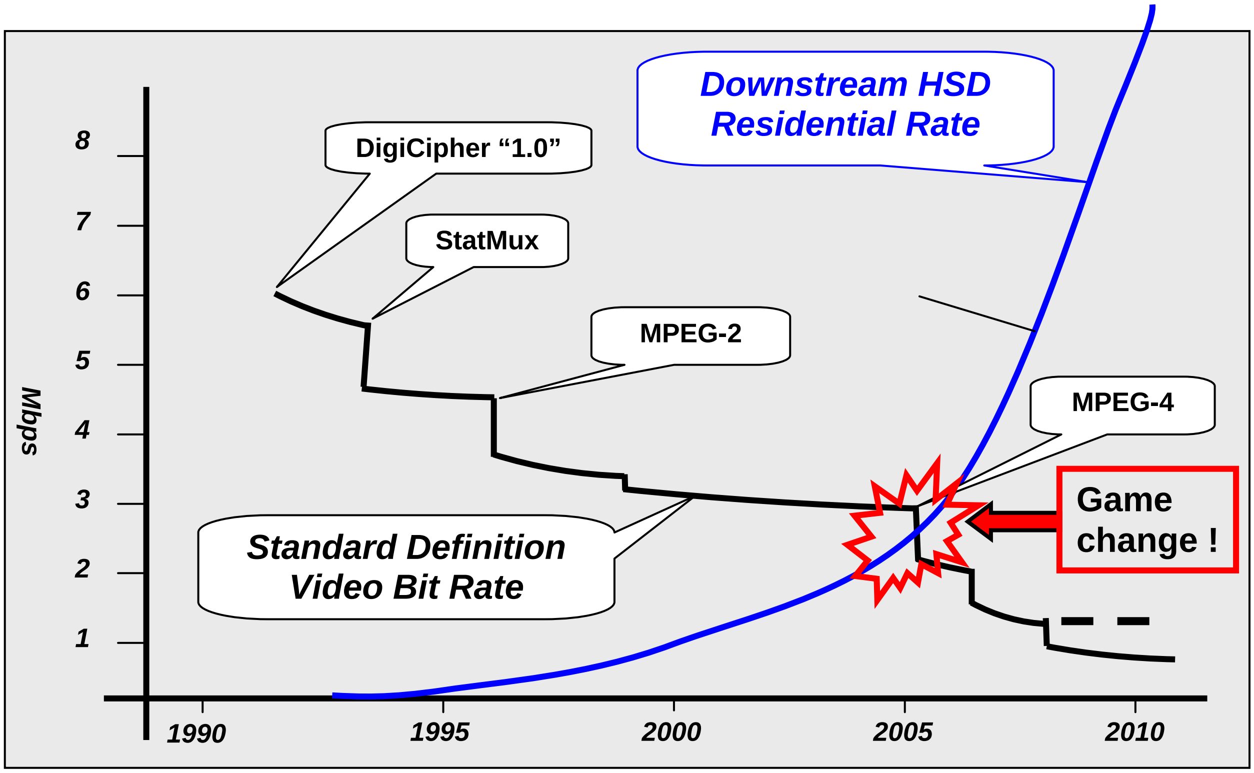

Figure 1 – Downstream Internet Speeds vs Digital Video Requirement

B. Historical Trends

Figure 2 shows a traditional trajectory of broadband growth, beginning with a 2009 spectral allocation of 400 Mbps, or 10 QAMs of unicast. This estimate is based on a mix of VOD and HSD (DOCSIS) downstream channels, such as 6 VOD and 4 DOCSIS, or 8 and 2. While this represents the aggregate bandwidth as one spectrum line-up, there is some devil in the details. It is common for the VOD QAMs to extend across more than a single serving group of data today. For example, if the DOCSIS carriers serve a single node, then the VOD carriers might serve two nodes. As such, the growth forward from ~400 Mbps has room for a portion of the bandwidth seeing a VOD service group split opportunity, acting as a course correction of the average growth pace, much like a node split. In this sense, it is conservative. On the flip side, however, slower trajectories that may result from efficiencies associated with a transition to IP video delivery would not necessarily apply to the VOD (CBR) unicast.

As DOCSIS begins to consume more of the HFC bandwidth, and to smaller and smaller serving groups, a true IP-centric unicast growth trajectory becomes a more representative model. Network DVR (n-DVR) would also contribute and possibly accelerate this trend, reversing any statistical benefits to network transport requirements associated with the overall trend of increased viewing of recorded content on home DVRs.

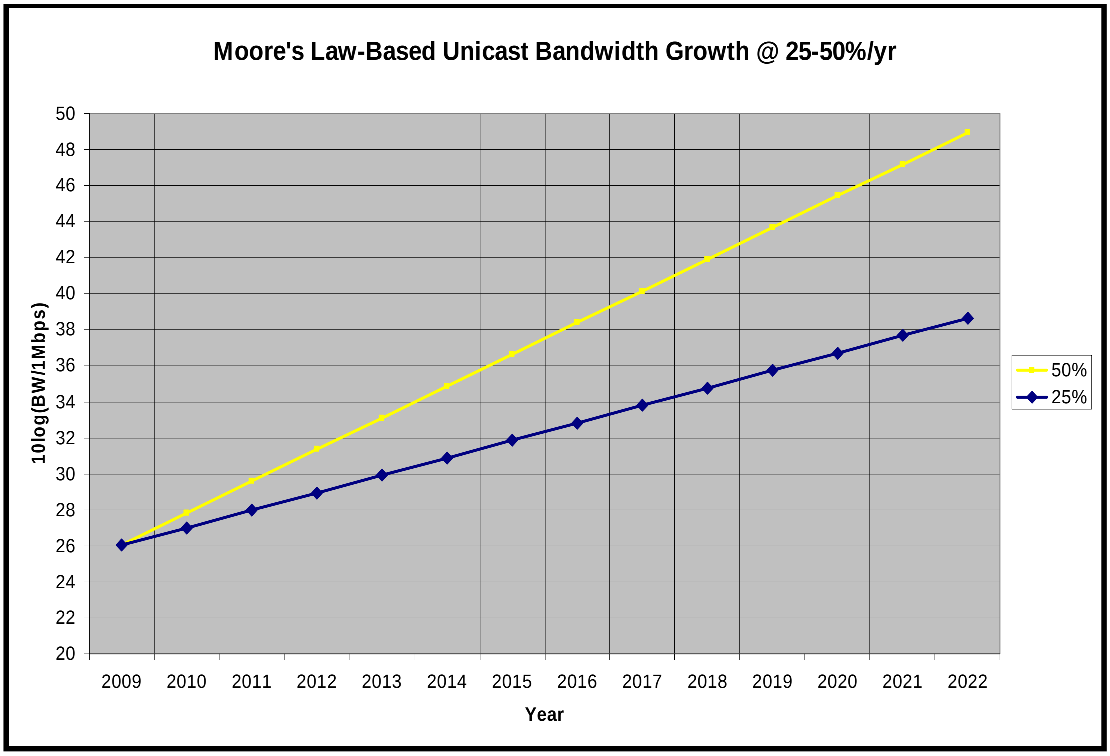

Figure 2 – Common Broadband Growth Trajectories - Nielsen's Law

Referring once again to Figure 2, the linear, in dB, trajectory takes us to a 10 Gbps of spectrum allocation (40 dB) between 2017 and 2023, for growth within the 25%-50% CAGR range. Unfortunately, 10 Gbps, though possible over the access network, is not readily achieved over HFC as we use it today [5].

<u>C. The Next Decade of Growth</u>

While a 10 Gbps requirement creates challenges, there are factors that intervene favorably as we consider traffic growth trajectories moving forward. For example, it is certainly reasonable to consider our growth trajectories in the context of the larger picture of network segmentation relief, which increases average bandwidth per subscriber. Refer to Figure 3. This segmentation relief would be as shown in the blue and purple trajectories in Figure 3 below for 25% and 50% growth, respectively. In this case, segmentation amounts to a service group size reduction by a factor of two, which creates (ideally) 3 dB on the plot.

In addition to service group reduction, there are other means to exploit technology advances in favor of extending the life of the network. Two such means are shown here. First, consider the dotted blue and purple trajectories of growth in Figure 3, labeled "w Video Eff." This stands for "with video efficiencies," which means we account for improved video bandwidth usage via H.264, used exclusively for IPTV in this example. We also include a factor to scale the bandwidth further through the use of variable bit rate (VBR) delivery over DOCSIS 3.0 with channel bonding. This is in contrast to the CBR technique used over single MPEG-TS QAMs in today's digital video delivery.

Finally, while 10 QAMs may be allocated today for unicast, use of analog reclamation and switched digital video can increase the available pool of QAMs for new narrowcast services. Different combinations of analog and digital broadcast (including SDV for modeling purposes) offer different available size QAM pools, up to 96 QAMs of narrowcast for 870 MHz systems under specific broadcast assumptions – the top 40 programs in SD & HD. Thresholds of available QAMs are given by the horizontal dashed lines of different colors. Also, note the black dotted horizontal reference for 10 Gbps. This represents, basically, the limit of the RF coaxial last mile today, under limitations imposed by 1 GHz field equipment. By making further changes in the plant, this number can increase further [5][6].

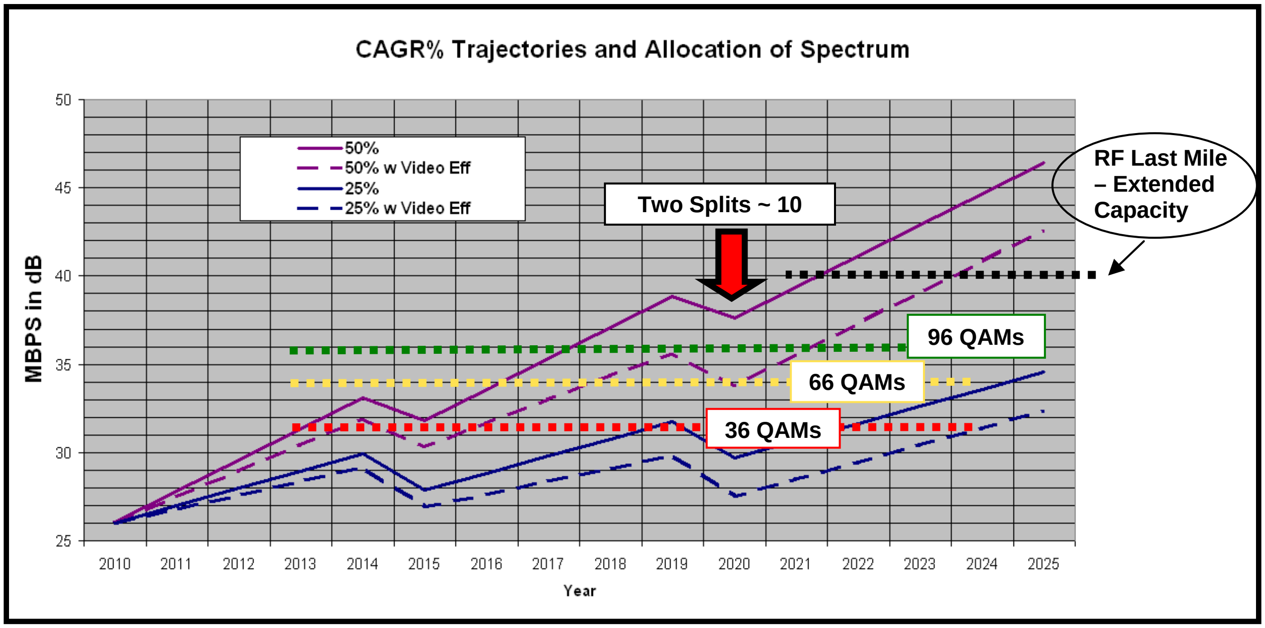

Figure 3 – Broadband Growth, Corrections, and Thresholds

D. Endpoints in Sight?

Figure 3 brings us back to the question posed at the introduction. Do the upward bound trajectories simply continue on and on? Extrapolating Figure 3, it would suggest that traffic in the Tbps (60 dB) range enters the conversation in the 2030s and could demand attention (for those who are not retired before that!). Or, under some imposed criteria – such as considering only media consumption – can we begin to draw the box around what worst case possibilities exist, and what new generations of video might entail? Referring to Figure 4 – are there media consumption endpoints that may come into play as we architect the long-term network? What are some of the things that might limit the amount of media that one home consumes? We refer to this household-based media consumption boundary as a *Full Consumption Household (FCHH).*

The components of interest are relatively straightforward, and include scientific, demographic, and behavioral elements, such as:

- What are physiological and perceptual limitations of the human visual system in space and time?
- How do display, setting, and field of view (FOV) factor in?
- Where are the evolution of new formats and better encoding headed?
- How many media-consuming subscribers per home?
- What about multi-tasking behaviors and devices?

These pieces of information, which themselves may have underlying components, can be used to create a simple model of media consumption boundary conditions. Maximizing the delivery to a single household by summing of all media rates calculated is useful and insightful, especially for the home LAN industry. However, it also has a worst case, over-engineered quality to it, in particular relative to a network aggregate. When combined with demographics, behavioral assumptions, and traffic engineering factors, we can understand the values in the aggregate – whatever form aggregation points might take down the road.

Note the calculation is not considered as early requirements for future equipment. It is instead upper boundaries, around which inspired conversations regarding model inputs and assumptions, and "begin with the end in mind" planning can take place.

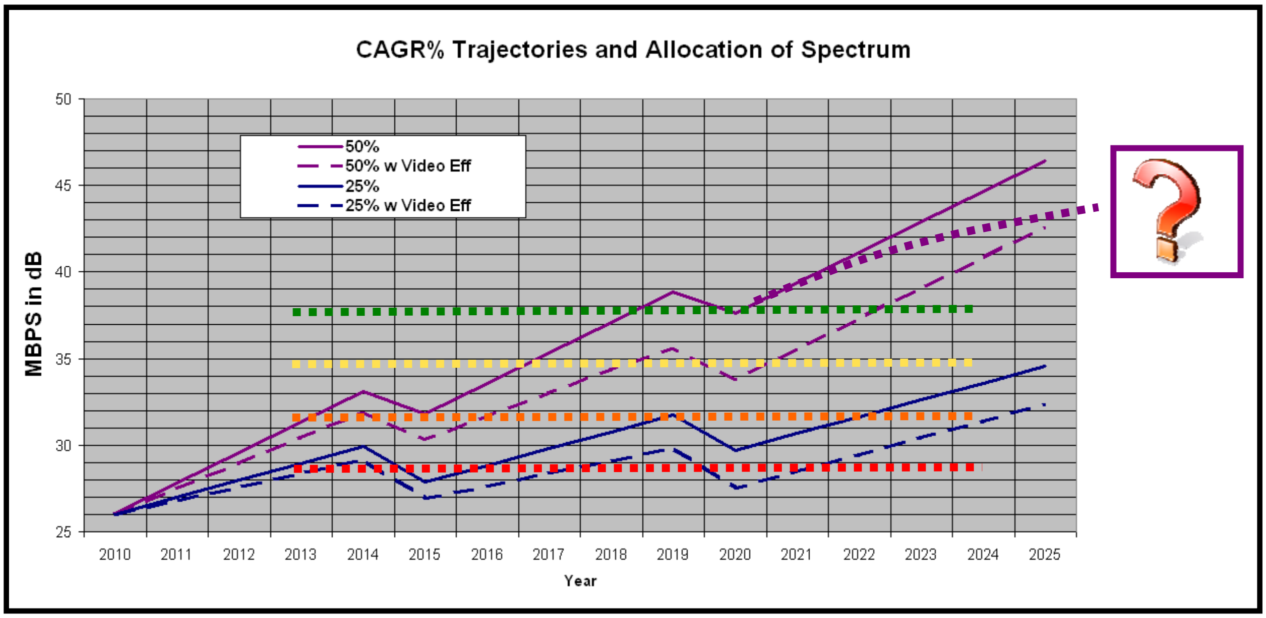

Figure 4 – Household Media Consumption: Asymptotic Endpoints?

II. Human Perception and Limitation

A. Spatial Resolution

Size Resolution

Visual acuity is the term used to describe how well our eyes can resolve distinct objects. The regular prescription glasses visit to the eye doctor is based on measuring visual acuity so that it can be readjusted via corrective lenses. The test themselves and the field of study in general uses concepts and science developed more than a hundred years ago, and much still is applied today. In particular, many references point out that a person with normal (20/20) vision can resolve elements in a pattern that are $1/60^{th}$ of a degree, or one arc-minute, separated. “20/20” refers to evaluations of ability for people with “normal” vision to identify such predefined patterns at 20 ft away (or 6 m, thus 6/6 in “metric” vision). As originally referenced by Snellen, it is meant to convey what is minimally acceptable as “normal” vision. It does not convey the best visual acuity that can be achieved, as it is often used casually. Also important to understand is that the acuity test is done using extremely sharp lines (letters or literally lines) at very high contrast ratios in static configuration, none of which applies to typical viewing conditions for video. Many good examples of misleading conclusions because of the interdependency of these variables are described in [16]. In general, the one arc-minute reference as the visual limit of resolution, while not wildly inaccurate, is nonetheless an oversimplification of the ability of the eye based on the above aged reference designed for visual acuity testing. And, it is not very representative of a typical video viewing environment.

Other acuity references – it is surprising how many and varied they are – use 30 seconds (half arc-minute) or less, such as 28 seconds [2], or 18 seconds [20]. Physiology-based calculations, based on the receptors lining the retina or the aperture itself, calculate 20-30 seconds of arc as the limit. Other references call out such resolution as theoretical capability, but also that it is further limited in the subsequent processing that leads to actual perception of images. Please see the references section at the end of the paper for a more comprehensive look at literature. “Normal” 20/20 acuity – one arc-second – is behind some of the guidelines that derived HDTV resolution and viewing distance, as well as newer HDTV generations in development, such as 7680 x 4320 (16x HD) Ultra-High Definition TV (UHDTV) to be discussed.

In video display terms, translated to linear distance, one arc-minute pixels from the 20 ft visual acuity test point means 1080p resolution HD needs a 134 inch display. Of course, common viewing distances are closer than 20 ft, resulting in display sizes reducing accordingly into the ranges that are more familiar for the same spatial resolution qualities. From another perspective, viewing 1080p resolution on an HDTV set means a horizontal viewing angle of about 33° degrees in order to absorb all of the resolution detail offered by the high definition programming. Display sizes and recommended viewing distances will be discussed in Section II-B – Displays & Formats.

Spatial resolution is inseparable from display size in defining the video experience. This is simply because a fixed number of pixels and varying display sizes mean varying pixel sizes. Varying pixel sizes means a different viewing point is “ideal” for resolving the detail. Too far removed, and they eye is not sharp enough to notice. Too close and the digitization may be viewable, the viewing angles may be uncomfortable, or both. Viewing angle, however, also relates to the feeling of presence to the viewer. A wider field of view (FOV) creates a stronger feeling of presence, in general. We discuss FOV below and elaborate on displays and formats in the next section.

Field of View

The eye's field of view is about 140° in the horizontal and 90° in the vertical, or roughly 30% of a 4π steradian sphere [4]. This FOV range is not uniform across references (can you tell that human vision is analog yet?), but those mathematically used herein are all are within about 12% of another, and each about one-third of the steradians of a full sphere. This 140° by 90° field is about 30x more in terms of the FOV provided by the guidelines we will see below for the recommended HD viewing experience according to the Society for Motion Picture and Television Engineers (SMPTE). THX Cinema Service recommends similar criteria, on average, for a theatre experience. From a visual FOV perspective in setting upper boundary video rates, it can be supposed that to guarantee full resolution in any direction, we should consider 30x the pixels, display size, and position that ensures 140° H and 90° V. It would never make any sense, physically, to have a larger FOV.

Combining acuity and field of view (FOV), estimates of the total number of "megapixels" for human vision range from 200-600 Megapixels. Specific examples in the references used herein include 221 Megapixels [2] and 576 Megapixels [20]. Note the difference! It is easily mathematically attributable to a slight difference in the FOV used, but to a larger difference in the acuity limitation (18 seconds versus 28 seconds). The author in [2] has subsequently written that the 28 seconds value is considered conservative and to be updated.

While this is indeed the field of view through the range of peripheral vision, the resolution of the eye drops off away from center of focus very strongly. The actual number of pixels we can perceive at one time, including the weighting of this resolution drop-off, has been estimated at about 15 million pixels – or 5% using 300 megapixels of vision. Nonetheless, since it is not realistic to assume where a viewer is focusing at any given time, or assume all viewers present are focusing in the same area, there is little that can be done to accommodate this, although in [2] this is a key driver for doing the analysis therein. In the gaming environment in particular, different assumptions about the focus of the gamer might be able to better exploit this limitation. We will not account for intelligently exploiting the pixel space of perception, but refer the reader to [2] for a discussion of the concept of variable resolution pixels.

In addition to pixel perception and FOV, wide viewing angles can result in an unpleasant viewing experience due to symptoms similar to motion sickness, an effect strongly dependent on the viewer's visual acuity. Newer research has been done here and will be discussed [10].

⇒ We will consider both spatial resolution limitations and the full visual field of view, in the spirit of an upper bound of potential consumption

⇒ We will account for "FOV" viewing in two ways

- Scaling of the Steradian ratio *and* pixel size using 20 arc-second for visual acuity

- Scaling of UHDTV by the solid angle ratio: (FOV in Sr / UHDTV in Sr). UHDTV will be described in more detail below. Its FOV is about 2 Sr, with spatial resolution based on one arc-minute of resolution.

The former represents boundary conditions for video delivery, the latter more in line with an HDTV 3.0 scenario – where UHDTV represents HDTV 2.0.

Practically speaking, an extended FOV can mean a shorter viewing distance (or a larger screen). The smaller distance may be limited by comfort, as previously described. Larger distance in addition may be limited by physical wall space. In either case, being closer or larger challenges pixel size as it relates to visual acuity – reason to increase the pixel count based on a better resolution reference of 20 seconds for physiological boundary rate purposes.

Color

Human vision's tri-chromatic image receptors ("cones") are exploited in color reproduction, although indirectly applied. Video processing includes correction for the nonlinear nature of how the color spectrum is received (gamma correction), as well as a mapping of the three primary colors into brightness (luminance) and two color (chrominance) components. The remaining paragraphs may be old hat material for experienced video engineers. But, because we manipulate these results for bandwidth purposes, we will describe with some level of detail color space processing.

It has been estimated that humans perceive about 10 million or less colors at once. Thus, three colors already corrected for wavelength response, and allocating eight bits per shade to red, green, blue (RGB), provides 2^(8x3) = 2^24, or >16 million colors, known as "True Color." The color space range of True Color, therefore, exceeds human perception – whereas even one less bit and certainly two could under represent the color space. Note that higher "bit depth" of color encodings exist. However, these are typically used in recognition that it provides bit depth "margin" for further post-processing that may negatively affect the color space. In this way, and in A/D converter parlance, the post-processed bit depth still yields sufficient resulting "effective" bits of color, hopefully keeping undesirable artifacts imperceptible. As observed above, a bit depth of eight does not leave much room for loss of color space resolution below the threshold for human perception.

In implementation, each pixel is defined by three bytes – one luminance and two chrominance. While not three independent colors, it is nonetheless three independent variables defining the color space, just in a different way. It is done so because it is well understood that the eye is most sensitive to luminance, and less so to chrominance. Thus, for chrominance components, shortcuts can be taken with less impact. Since bandwidth and storage requirements were key considerations, chrominance sub-sampling is applied in a way that is deemed sufficient for imperceptibility.

The shortcut is to have a fraction of the number of chrominance samples, spreading them across several pixels, compared to luminance, saving bandwidth and storage in the process. This is what is being described when "4:2:2" or "4:2:0" sub-sampling is performed. These mean half as many chrominance samples, and one-fourth as many, respectively, translating to savings of one-third or one-half in bandwidth and storage. No sub-sampling would be described as "4:4:4." 4:2:0 is the mechanism supported through MPEG-2. For MPEG-4, however, new "high" profiles are supported including 4:2:2, and even no sub-sampling. Other profiles in MPEG-4 go even further, allowing 10-bit samples (instead of 8) and in the most extreme case using no sub-sampling, up to 14 bits. Several reasons exist for the more extreme modes, such as matching source content format and handling, lossless transformation, content distribution and contribution, and editing and post-processing [21].

⇒ We will model color sampling from today's implementation of 8-bit @ 4:2:0, through no sub-sampling and up to a bit depth of 10 bits per channel – some margin of degradation at any process along the way up to the display

B. Displays & Formats

Resolution & Screen Size

With the advent in particular of HDTV, development in QUAD HD (2x in each dimension) and UHDTV, and the human vision concepts above, the video and CE industries have a strong understanding of the relationship among resolution required, screen size, and viewing distance. We elaborate further on these practical conclusions here. Similar to the color space description, video engineers and home theatre aficionados may skip this section, but we detail somewhat in order to provide context for extending the recommendations.

Just as visual acuity is measured and referenced to object sizes at defined distances, the display size and placement relative to the viewer is a key piece of the resolution requirement equation. Figure 5 is a straightforward way to see how these factors interact [18] based on recommendations provided by multiple professional organizations, home theatre experts, and retail manufacturers. Generally, for a fixed resolution (linear trajectories on the plot), a larger screen size is best viewed further away. For a fixed screen size, higher resolutions are best viewed by sitting closer to allow for the full benefit of the increased detail on the display. Finally, for a fixed distance from the display and the higher the format resolution, the larger the screen size should be. As a simple example of the value of such guidance, a simple 50 inch screen, if viewed more than 20 ft away or greater, will begin to lose the benefit of HD at 720p, and provide an experience more akin to Standard Definition 480p. Sitting too close, such as 5 ft away on a 100" 1080p screen, threatens quality due to distinguishing of pixels. This chart thus also explains the increased pixel count of UHDTV based on a 100" display recommendation and wider viewing angle (closer).

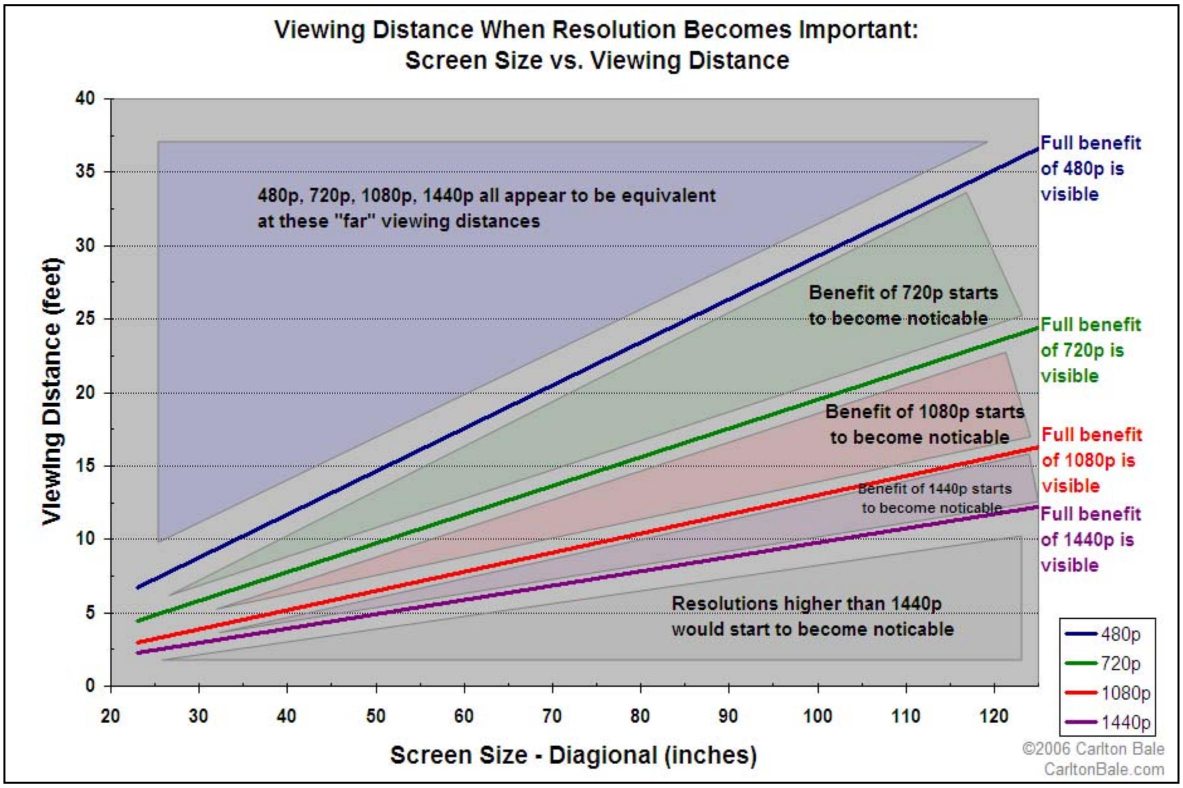

Figure 5 – Resolution – Seeing Everything That You Paid For

The screen size – viewer distance guidelines come from different organizations and retailers. While they tend to cluster around similar recommendations, they are not in complete agreement, usually due to the inherently varied perspectives of the organizations to the issue, such as, for example, what sells more TVs. The range of recommendations varies from about 1.5x-2.5x of display size for viewing HD content, with the lower end corresponding to 1080 resolution. There is no shortage of websites willing to give opinions on the optimum range [12][14][18][22].

The recommendations are correlated to an assumption about visual acuity as it relates to the distance to the display and the ability to resolve the image detail. For example, the recommended optimum fields of view are given as about 30 deg (SMPTE) or 40 deg (THX) in the horizontal plane. Table 1 shows the consistency between the SMPTE recommend viewing field horizontal resolution and the maximum recommended distance beyond which the detail provided by the content cannot be appreciated by the viewer.

In the vertical plane, simpler guidelines say do not do anything to induce neck strain, and maintain at least a 15 deg vertical field of view, with a maximum of 35 degree viewing angle (neck strain).

Table 1 – Viewing Angle and Resolution

	SMPTE 30 deg FOV	Fully Resolution Maximum
Display Size	Distance in Feet	
60	8.1	7.8
80	10.8	10.5
100	13.6	13.1
120	16.3	15.7
10	1.4	1.3

UHDTV / Super Hi-Vision

While there are well-established professional guidelines on the best viewing conditions for HD, there is still reasonable debate about what is “optimum.” In particular, the use of a horizontal field of view of 30° (SMPTE) and 40° (THX) are part of the debate. HD, at least, was the first system that began considering the limitations of the human visual system.

Studies show that a viewer perceives a more immersive experience when the angle is widened up to about an 80° viewing field [12]. The HD value of 30° came about through research noting that this feeling of presence strengthens considerably as the viewing angle exceeds 20°. However, the research indicates that it continues to strengthen past 30°. Japan Broadcasting Corp (NHK) determined through their research during the development of UHDTV that a 100° angle is the recommended viewing field [3]. Recent studies by NHK [10] also determined that viewing comfort, measured in terms of the symptoms of motion sickness, leveled off at just above the 80° “immersive” viewing angle.

From the same NHK study, visual acuity was examined, but as it applies to spatial resolution of actual images, as opposed to test patterns. It included tests both for determining the ability to discriminate different resolutions, but also to perceive subjective “realness” of pictures compared to actual objects. “Realness” increased sharply to about a spatial resolution of 0.6 arc-minute pixels, increasing further beyond but much more slowly.

Now, if we use this 0.6 arc-minute value and a vision FOV of 140° x 90°, we calculate the full scene as 125 Megapixels. However, UHDTV, while a much larger total pixel count than HDTV, uses the same one arc-minute resolution basis, resulting in a scene of the same FOV being about 45 Megapixels. In addition, the recommended viewing position for UHDTV around a 100” screen (100° x 67°) is only about one-half of the above FOV. The total pixel count then becomes about 23 Megapixels. This (32 Mpixels > 23 Mpixels) represents how a pixel count of 8k x 4k, or 32 Megapixels, is deemed sufficient for UHDTV. It remains consistent with visual acuity requirements at the recommended viewing distance, yet is still backwards compatible with HD formats at 4x in each plane. Note also that (coincidentally in this case) the total pixel count of approximately 32M exceeds by a factor of two the estimated total number of pixels that can be simultaneously “perceived” given prior discussions regarding 15 million perceivable pixels.

UHDTV specifies only a 60 Hz progressive frame rate, and a 10-bit color depth. The 60 Hz frame rate is curious – perhaps the focus (pardon the pun) on spatial resolution distracted NHK on in-depth study of dynamic resolution, i.e., resolution with motion? We will discuss dynamic resolution further below. In [10], it is indeed suggested that it is time to re-evaluate frame rate because dynamic resolution is also impacted by the increasing screen sizes.

Finally, NHK predicts that, based on a typical 30 year broadcast technology lifecycle, the middle part of the next decade (2020s) is the strike point for this format to become commercial.

UHDTV has a role in the calculations ahead as follows:

⇒ We will consider the use UHDTV format as an end-user service – in addition to using its guidelines as a reference for a visual "field of view" display

⇒ More generally, but stated here for completeness, we will NOT consider any multiple screen environments of different views – such as screens in front and to the sides for the purpose of being more fully immersive

3DTV

The screen size – distance relationship runs into trouble with 3DTV because the 3D experience is, by design, trying to add depth. Virtual extreme points are purposefully created such that the content on the screen may look closer or farther. The "optimal" 2D location cannot be optimal for all depths. There is much work to be done here, but some simple guidelines are [22]:

1) 3D viewing is more of an individual preference situation for viewing distance than the carefully crafted scientific basis of 2D viewing. Much of this has to do with some fundamental challenges that the brain must resolve to manage its way through the 3D experience, which vary individual to individual.

2) Scientifically, the challenge posed to the brain involves a 3D "comfort zone" whereby, outside of this zone, viewing may be less enjoyable and more disorienting. The result is a much narrower range than 2D viewing for an enjoyable experience.

3) A currently recommended approach is to have the physical distance be guided by the average of the closest and furthest virtual extreme points of the 3D display. The result is that, generally, closer than the 2D recommended distance will be better for 3D. In the example given in [22] based on one viewership study of 3D, the distance was decreased by about 20% compared to the 2D (visual acuity-based) distance. A conclusion of this would be that when purchasing a 3DTV, consider the next size higher (Best Buy will recommend the same!).

We will consider 3D formats in developing the numerical model. While the optimization of 3DTV is a multipronged work in progress, what we care most about here is its impact on the bandwidth to deliver to the home, which can at least be ballpark estimated. Our estimates range from 1.3x - 1.7x scaling factor, and likely higher in early versions, depending on approach.

⇒ Given the abundance of unknowns, we will use 1.5x scaling factor (50% more bandwidth) to represent the average increase in pixels to store or transport 3D content

C. Dynamic Resolution

The term "dynamic resolution" refers to the ability to resolve spatial detail of objects in motion. The 30 Hz (interlaced), 50 Hz, and 60 Hz frame rates have origins in AC line rates, and thus are not scientifically developed rates tied to video observation and testing. They simply exceeded what was known at the time about 40 Hz rates causing undesirable flicker. Most early analysis on frame rate was to ensure that motion appeared realistic (seamless), as opposed to a sequence of still shots. There was less focus on eliminating other artifacts of motion, such as smearing effects. The above frame rates have since become embedded in tools and equipment of the production, post-processing, and display industries. As such, HDTV standards today are based on the 60 Hz interlaced or progressive frame rate. It has been suggested [1,10] that with larger and brighter displays of higher resolution, the frame rates in place based on practical implementation limitations of the 1930s era ought to be reconsidered, while recognizing a need to maintain backward compatibility. Further research on the phenomenon of human perception of video content as a function of frame rate was inspired by a study in 1980 that showed an improvement of up to two points on the CCIR *5-point* scale when the frame rate on a scene that included fast moving action was increased to 80 Hz [1]. On the 5-point scale, "poor" becomes "good," which is not insignificant.

As displays become larger and of higher resolution and contrast, the challenges to effectively displaying motion increases, because the edges to which movement is ascribed are now sharper as well as longer. Yet, as spatial resolution has improved, temporal resolution has not. This is in part because there is room to grow from the interlaced to progressive scan as a natural first step. Interlaced video itself is a nod to the imbalance of motion representation – exchanging spatial resolution for a higher rate of image repetition to better represent motion than a progressive scanning system of the same bandwidth. With the 60p improvement being exercised, the same question can be asked as with respect to spatial resolution – what is good enough? When can humans not perceive a difference? As you can likely guess, there is not a firm answer to this question, either. There are also not as many opinions as in the case of spatial resolution, which brings in the photography industry. By contrast, the "motion representation" arena typically includes both video engineers and the gaming community.

So, what is the bottom line of these studies? Humans stream video continuously in a physiological sense, so the question is around the processing engine in the brain. Various papers and web sites make estimates, some based on informal testing or loose observation, others more along the lines of the scientific method [1][11][17]. These sources describe tests where frame rates of 100-300 fps show perceived improvements compared to 60 fps. The difficulty of performing this type of testing – content and equipment basics – limits how much has been learned.

⇒ For our model, based on the above mentioned studies, we will allow for the possibility of a 2x and 5x increase in frame rate

There are potentially positive encoding implications to these higher frame rates, which we will consider below.

D. Encoding & Perception

The cable digital video world is dominated by MPEG-2 encoded content today, distributed via MPEG-2 Transport Stream (TS). The newest commercial evolution of the MPEG encoding standard is H.264. On average, H.264 will provide on the order of a 50% reduction in bandwidth of video delivery. The 50% is a gross oversimplification, as the gain is very content dependent, with higher action content being more difficult to pull down by 50%, while simpler content is much more efficiently encoded. Nonetheless, for our simple model and aggregate assumptions, we will make use of this average 50% average reduction compared to MPEG-2. Scaling the end result by a different factor is a straightforward modification.

After H.264 we will find, along a typical encoding lifecycle, H.265, on which early work has already begun. Note that it can be anticipated that higher frame rates are likely to be more efficiently encoded – require more horsepower, certainly, but provide better gains. The simple reason is that frame-to-frame variation and temporal aliasing should be able to me more aggressively exploited [1]. Because of this, we will also assume that, in the world of H.265, the 50% efficiency objective is again fully achieved. Early estimates of 20% (mathematical gain) and up to 50% (math plus perception) might otherwise suggest a more conservative scaling for H.265. However, we will assume higher frame rates more readily enable scaling efficiency goals.

⇒ For H.265, we will assume the objective efficiency gains can be achieved, or 50%

For simplification, we will apply 50% across the board, rather than differentiate 60 Hz encoding gain from 120 Hz and 300 Hz. Again, different relative scaling is straightforwardly applied.

One final assumption to be made is the efficiency with which we can assume channelized MPEG-2 TS-based QAM transport, which uses CBR traffic engineering, is improved in IP-based transport using bonded DOCSIS 3.0 channels and VBR delivery.

⇒ We use a 30% efficiency improvement for VBR average bandwidth gain

Note this is DOCSIS 3.0 channel bonding and VBR efficiency contributions only – we have already accounted for the conversion to H.264 with the baseline HD assumption.

Finally, we will use the following constant bit rate value as the reference point for subsequent bit rate calculations:

⇒ Baseline reference: MPEG-4 HD 720p or 1080i CBR Average = 8 Mbps

From this reference point, we will calculate a variety of configuration-dependent rates through scaling factors described in all prior discussions.

III. Demographics

A. Households

The household composition of the United States is very well quantified by the census bureau [19], including projections about what is to come. Outside of the US, the same mathematics would apply of course, but based on a different set of statistical distributions. Figure 6 shows a bar chart of US household size through 2007 [13]. This data can be used to find what a single household maximum consumption might be for, say, 90% of the households of four people or less. Using the distribution itself can also be used for the traffic engineering of the "average home" peak, and for the aggregate. In calculations to come, we say >90% and add the five person household to the home and neighborhood aggregates.

In addition to the "90%"maximum household threshold, the average household size number is also useful, in particular for these larger neighborhood aggregates. In the US, the average size of all households is 2.6, and the average size of family households is 3.2, both of these being 2009 projections. And, in both cases the number is gradually declining and has been for many years. Interestingly, the age distribution of children 18 yrs of age and under is virtually uniform, whether viewed in blocks (<5 yrs old, 5-9, 9-14, or 15-19), or in individual years. This is relevant, as age range guides the likelihood of a user consuming his or her own content, and also whether they are likely to be multi-tasking – i.e. surfing the web while watching TV.

B. Behaviors

With respect to multi-tasking, a recent study [8] found that three out of four Americans multi-task – watching TV while surfing the Internet – and that the trend is growing year-on-year. Small laptops, more powerful WiFi-enabled smart phones, and devices like the iPad are sure to keep the trend alive. For slightly more than half, the Internet is actually the main focus while multi-tasking. The content on the TV is rarely (7%) the subject of their browsing. Instead, browsing is generally the larger websites or a social networking activity. Thus, even when primary screen viewing today, that may be only part of what is going on by a media consumer.

- ⇒ We will consider average and "maximum" family sizes as defined by >90%
- ⇒ We will consider that multi-tasking is taking place for some users in the home
- ⇒ We will apply family distributions and cable penetration to consider the aggregate from a traffic engineering perspective
- ⇒ We will not combine multi-tasking with 3D viewing

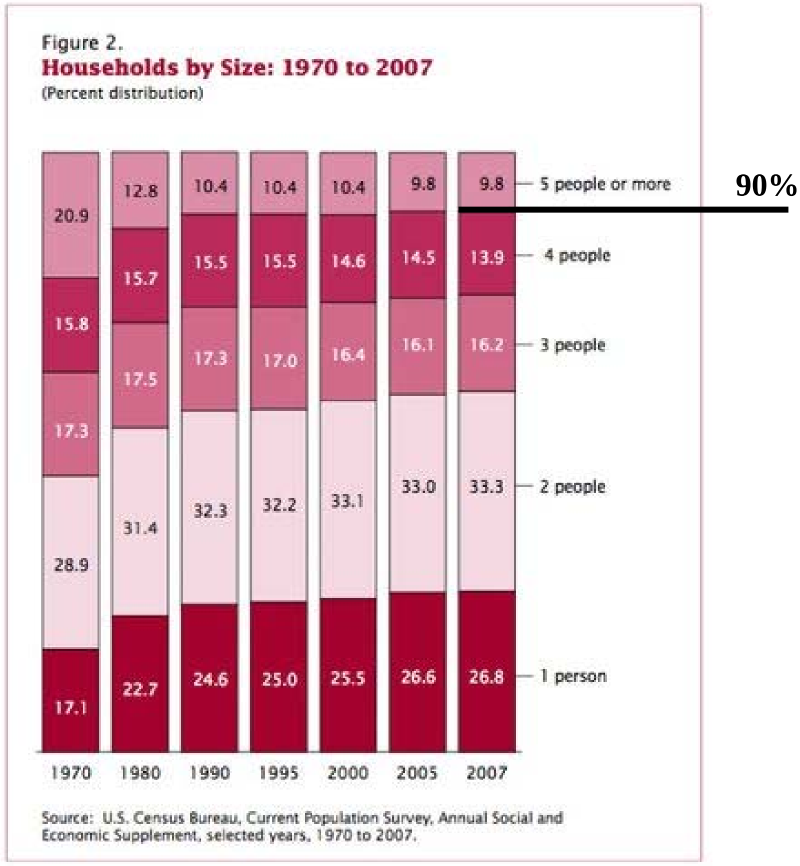

Figure 6 – US Household Distribution and Trend

IV. Full Consumption Household

A. Summarizing Factors

The list of contributing factors as we discussed them above is summarized below.

⇒ We will consider both spatial resolution limitations and the full visual field of view, in the spirit of an upper bound of potential consumption

⇒ We will account for "FOV" viewing in two ways

- Scaling of the Steradian ratio *and* pixel size using 20 arc-second for visual acuity
- Scaling of UHDTV by the solid angle ratio: (FOV in Sr / UHDTV in Sr). UHDTV will be described in more detail below. Its FOV is about 2 Sr, with spatial resolution based on one arc-minute of resolution.

⇒ We will model color sampling from today's implementation of 8-bit @ 4:2:0, through no sub-sampling and up to a bit depth of 10 bits per channel – some margin of degradation at any process along the way up to the display

⇒ We will consider the use of UHDTV format as an end-user service – in addition to using its guidelines as a reference for a visual "field of view" display

⇒ More generally, but stated here for completeness, we will NOT consider any multiple screen environments of different views – such as screens in front and to the sides for the purpose of being more fully immersive

⇒ Given the abundance of unknowns, we will use 1.5x scaling factor (50% more bandwidth) to represent the average increase in pixels to store or transport 3D content

⇒ For H.265, we will assume the objective efficiency gains can be achieved, or 50%

⇒ We use a 30% efficiency improvement for VBR average bandwidth gain

⇒ Baseline reference: MPEG-4 HD 720p or 1080i CBR Average = 8 Mbps

⇒ We will consider average and "maximum" family sizes as defined by >90%

⇒ We will consider that multi-tasking is taking place for some users in the home

⇒ We will apply family distributions and cable penetration to consider the aggregate from a traffic engineering perspective

⇒ We will not combine multi-tasking with 3D viewing

The maximum video stream rate and FCHH screen rate is a key objective. The nature of the scale of the maximum theoretical rates suggests that they dominate all other aspects of multi-tasking of media. As such, for the boundary FCHH calculation, we simply determine its value, estimate a peak for the >90% of homes case, and further estimate implication by weighting it by the US household distribution and traffic engineering variables. We do not consider secondary multi-tasking behaviors as relatively meaningful to bandwidth. There is so much in the unknown in the build up of worst case boundary conditions, that it did not make sense to add further conjecture about secondary behaviors.

However, in addition to the extreme boundary case, we consider cases that we refer to as "HDTV 2.0" and "HDTV 3.0." In these cases, we back-off some of the worst case boundary variables in a way that may make evolutionary sense over the course of time. Generally, they include combination of UHDTV and HDTV services, and enhanced versions thereof. In these cases, we consider primary screens, secondary screens, age distributions, and user behaviors. The assumptions in these cases are:

1) One FOV primary screen (home theatre)

2) UHDTV screen available to all other users
3) Smaller multi-tasking screen supporting 1080p HD

B. The Big Equation

Ultimate Boundary Conditions

The boundary condition for video rates are shown in Table 2, where the damage is quite clear for 3D viewing, at very high frame rate, and removing chrominance sub-sampling – all told 15x worth of increased storage and or bandwidth alone. While the highest number in Table 2 is nearly 57 Gbps, it seems quite reasonable to assume H.265, or perhaps something beyond, is available to help combat very long-term bit growth.

Another major scaling factor is the frame rate assumption. As we have discussed, as display have gotten larger, brighter, and higher resolution, temporal resolution has recently come to the front burner as a topic of interest. Studies indicate advantages in testing done up to 300 Hz. This is somewhat speculative due to the small amount of testing and variations in viewing conditions, but we consider that such findings may be paving the way for increases in frame rate as well. Under this assumption and H.265, we estimate enabling storage and bandwidth based on rates of **19.0 Gbps** and **28.4 Gbps** would cover any foreseeable situation, beyond which no additional advantage would be gained.

Yet, it is very much unclear the direction of higher frame rates. Under the assumption that doubling of frame rate delivers sufficient dynamic resolution, our calculations estimate that 2D and 3D services would be ensured to come in under **7.6 Gbps** and **11.4 Gbps**, respectively, beyond which there is no advantage to be gained.

Table 2 – Ultimate Video Service Boundary Rates (Gbps)

2D	3D	H.265 - 2D	H.265 - 3D	Scan Type	Quantization	Subsampling	Frame Rate
10.1	15.2	5.1	7.6	Progressive	Bit Depth 10	Chroma 4:2:2	120 Hz
15.2	22.7	7.6	11.4	Progressive	Bit Depth 10	Chroma 4:4:4	120 Hz
25.3	37.9	12.6	19.0	Progressive	Bit Depth 10	Chroma 4:2:2	300 Hz
37.9	56.9	19.0	28.4	Progressive	Bit Depth 10	Chroma 4:4:4	300 Hz

We combine these boundary numbers with demographics to arrive at Table 3 below. The worst case peak boundary for the FCHH based on *The Ultimate Boundary Condition* in Section IV-B is simply the >90% household size of five multiplied by the video rate, while the weighted average considers both the distribution of household sizes and the age distribution. It suggests that a weighted average insurmountable boundary for >90% of the homes passed is **48 Gbps** under the maximum frame rate condition of 300 Hz. At 120 Hz, this limit becomes about **19 Gbps**.

Table 3 – Ultimate Video Service Boundary Rates (Gbps) with Demographics

		Max Peak Rate	Weighted Avg	Fr Rate
Theoretical Worst Case		(Gbps)	(Gbps)	(Hz)
H.265	**1 FOV Screen/user**	57	19	120
H.265	**1 FOV Screen/user**	142	48	300

We apply traffic engineering of the access network in the next section, alongside similar calculations for HDTV 2.0 and HDTV 3.0.

The Matrix

Now that we have calculated worst case boundaries, we can examine all of the aforementioned variables for intervening cases along the way that describe HDTV 2.0 and HDTV 3.0. The result is a massive matrix of permutations, and a correspondingly large number (and range) of FCHH bandwidth estimates for the discussed formats and scaling variables. A subset of the permutation matrix is shown in Table 4.

The scale factors associated with the columns in Table 4 are as discussed in Section IV – Summarizing Factors, including the reference baseline of an H.264 HD CBR average rate of 8 Mbps. Note that further incremental encoding gains of MPEG-4 HD are themselves very likely. Just under 8 Mbps gets 5 HDs into a 6 MHz QAM, however, so this represents that expectation. And, again, resulting outputs simply scale as [Assumed Rate/8 Mbps].

We repeat some of the scaling factors once more explicitly for convenience:

- Channel Bonded VBR Efficiency Factor: 0.7x
- 3D Multiplier: 1.5x
- H.265 Efficiency vs. H.264: 50%
- Quantization: 1.25x (8-bit to 10-bit)
- Chrominance Sub-sampling: 1.33x, 2.0x for 4:2:2 or 4:4:4 relative to 4:2:0
- Frame Rate: 2x, 5x
- HD to QUAD HD / UHDTV / FOV: 4x / 16x / 30x

Each of the first three numbers above is open to debate about the range of values to assume. VBR efficiency in some studies is as high as 50% or more. 3DTV is likely to begin at a greater than 50% premium, but as more advanced use of multi-view coding (MVC) and 2D+Depth techniques evolve over left/right-eye, a rate increase over 2D of about 30% have been estimated. Given the immaturity of 3DTV, and until more is known about MVC capability and what is considered a satisfactory 3D experience, it seems reasonable to be more conservative. Finally, H.265 has an objective of being twice as efficient as H.264. H.264 has indeed achieved this, on average, over MPEG-2, and we have discussed the favorable conditions relative to temporal resolution in Section II-C – Dynamic Resolution.

The Table 4 subset is again a single video stream calculation, and only for the 300 Hz frame rate – but includes each of the other intervening variables to calculate rates at this highest frame rate. The frame rate element, which has a significant role in the calculations, is also very straightforward to extrapolate downward from this table: for 120 Hz the matrix is multiplied by 40%, while for 60 Hz it is multiplied by 20%. Note that the grayed out cells represent accurate calculations, but parameter settings not associated with the UHDTV as it is drawn up today.

The maximum calculated video rates is about **6.3 Gbps for H.264 3D** and **3.2 Gbps for H.265 3D** as video rate boundaries – highlighted in Table 4. To provide some perspective on a couple of the major scaling factors associated with complex interaction of video perception and video format, these numbers become instead about **1.7 Gbps and 850 Mbps** if the frame rate tops out at 120 Hz, and 4:2:2 chrominance sub-sampling is allowed. For 2D viewing under those same conditions, they become abut **1.1 Gbps and 560 Mbps**.

Table 4 – Subset of Calculation Permutations
Primary Screen Bit Rate (Mbps)

2D	3D	H.265 - 2D	H.265 - 3D	Scan Type	Quantization	Subsampling	Format	Frame Rate
112.00	168.00	56.00	84.00	Interlaced	Bit Depth 8	Chroma 4:2:0	QUAD HD	300 Hz
224.00	336.00	112.00	168.00	Progressive	Bit Depth 8	Chroma 4:2:0	QUAD HD	300 Hz
140.00	210.00	70.00	105.00	Interlaced	Bit Depth 10	Chroma 4:2:0	QUAD HD	300 Hz
280.00	420.00	140.00	210.00	Progressive	Bit Depth 10	Chroma 4:2:0	QUAD HD	300 Hz
149.33	224.00	74.67	112.00	Interlaced	Bit Depth 8	Chroma 4:2:2	QUAD HD	300 Hz
298.67	448.00	149.33	224.00	Progressive	Bit Depth 8	Chroma 4:2:2	QUAD HD	300 Hz
186.67	280.00	93.33	140.00	Interlaced	Bit Depth 10	Chroma 4:2:2	QUAD HD	300 Hz
373.33	560.00	186.67	280.00	Progressive	Bit Depth 10	Chroma 4:2:2	QUAD HD	300 Hz
224.00	336.00	112.00	168.00	Interlaced	Bit Depth 8	Chroma 4:4:4	QUAD HD	300 Hz
448.00	672.00	224.00	336.00	Progressive	Bit Depth 8	Chroma 4:4:4	QUAD HD	300 Hz
280.00	420.00	140.00	210.00	Interlaced	Bit Depth 10	Chroma 4:4:4	QUAD HD	300 Hz
560.00	840.00	280.00	420.00	Progressive	Bit Depth 10	Chroma 4:4:4	QUAD HD	300 Hz
448.00	672.00	224.00	336.00	Interlaced	Bit Depth 8	Chroma 4:2:0	UHDTV	300 Hz
896.00	1344.00	448.00	672.00	Progressive	Bit Depth 8	Chroma 4:2:0	UHDTV	300 Hz
560.00	840.00	280.00	420.00	Interlaced	Bit Depth 10	Chroma 4:2:0	UHDTV	300 Hz
1120.00	1680.00	560.00	840.00	Progressive	Bit Depth 10	Chroma 4:2:0	UHDTV	300 Hz
597.33	896.00	298.67	448.00	Interlaced	Bit Depth 8	Chroma 4:2:2	UHDTV	300 Hz
1194.67	1792.00	597.33	896.00	Progressive	Bit Depth 8	Chroma 4:2:2	UHDTV	300 Hz
746.67	1120.00	373.33	560.00	Interlaced	Bit Depth 10	Chroma 4:2:2	UHDTV	300 Hz
1493.33	2240.00	746.67	1120.00	Progressive	Bit Depth 10	Chroma 4:2:2	UHDTV	300 Hz
896.00	1344.00	448.00	672.00	Interlaced	Bit Depth 8	Chroma 4:4:4	UHDTV	300 Hz
1792.00	2688.00	896.00	1344.00	Progressive	Bit Depth 8	Chroma 4:4:4	UHDTV	300 Hz
1120.00	1680.00	560.00	840.00	Interlaced	Bit Depth 10	Chroma 4:4:4	UHDTV	300 Hz
2240.00	3360.00	1120.00	1680.00	Progressive	Bit Depth 10	Chroma 4:4:4	UHDTV	300 Hz
842.24	1263.36	421.12	631.68	Interlaced	Bit Depth 8	Chroma 4:2:0	FOV	300 Hz
1684.48	2526.72	842.24	1263.36	Progressive	Bit Depth 8	Chroma 4:2:0	FOV	300 Hz
1052.80	1579.20	526.40	789.60	Interlaced	Bit Depth 10	Chroma 4:2:0	FOV	300 Hz
2105.60	3158.40	1052.80	1579.20	Progressive	Bit Depth 10	Chroma 4:2:0	FOV	300 Hz
1122.99	1684.48	561.49	842.24	Interlaced	Bit Depth 8	Chroma 4:2:2	FOV	300 Hz
2245.97	3368.96	1122.99	1684.48	Progressive	Bit Depth 8	Chroma 4:2:2	FOV	300 Hz
1403.73	2105.60	701.87	1052.80	Interlaced	Bit Depth 10	Chroma 4:2:2	FOV	300 Hz
2807.47	4211.20	1403.73	2105.60	Progressive	Bit Depth 10	Chroma 4:2:2	FOV	300 Hz
1684.48	2526.72	842.24	1263.36	Interlaced	Bit Depth 8	Chroma 4:4:4	FOV	300 Hz
3368.96	5053.44	1684.48	2526.72	Progressive	Bit Depth 8	Chroma 4:4:4	FOV	300 Hz
2105.60	3158.40	1052.80	1579.20	Interlaced	Bit Depth 10	Chroma 4:4:4	FOV	300 Hz
4211.20	6316.80	2105.60	3158.40	Progressive	Bit Depth 10	Chroma 4:4:4	FOV	300 Hz

B. Household Scaling Factors

The Figure 6 demographics are used to estimate a FCHH rate for >90% coverage of the subscriber base, and based on the rate calculations and household size distribution. The household size distribution is also valuable for building an aggregate for traffic engineering purposes. In addition to the household size demographic, use of the age distribution for households with children assumed that 50% were of age to be multi-tasking capable for media consumption. As previously discussed, the age distribution of children in the US is nearly uniform in single years and in blocks of years. As such, something slightly less than half of children in a household likely fall into this category. It is approximated at 50% here, recognizing that there are a sizable percentage of single adult households. This acts to increase the probability of a multi-tasking age user under the age of 18 for a given household size.

From the full set of rate permutations that include Table 4 and two other matrices just like it using slower frame rates, three cases have been selected for further analysis of the FCHH, and subsequently for aggregate calculations. These three cases, based on the screens in the house as outlined above in Section I-A – Summarizing Factors, are:

Case 1) H.264

First Service, Home Theatre: Highest rate stream – 3D viewing, FOV screen, 300 fps, 10-bit quantization without chrominance sub-sampling (highlighted in Table 2)

Second Service: 3D UHDTV (10-bit, 4:2:2, 60 fps (per standard))

Each Additional: 2D UHDTV + HD (multi-tasker, data service video clips)

Case 2) H.265 with same services as above

Case 3) H.265, UHDTV Only @ 120 fps (no FOV screen theatre)

First Service, Second Service: 3D UHDTV @ 120 fps

Each Additional: 2D UHDTV @ 120 fps + HD (multi-tasker, data service video clips)

Accounting statistically for multi-tasking behaviorally is accommodated by the traffic engineering variable for concurrent use. Whereas a 50% peak viewing hour concurrency is used for the video services, the secondary service consisting of HD viewing such as social networking video clips is oversubscribed by 10:1, or concurrent use is assumed engineered to 10%. Note that this is significantly higher than today's HSD-type engineered numbers on the order of 1% or less, and is meant to reflect that the trend of video sharing of high-quality video in real time will necessarily drive concurrency up. A competing data point would argue that as broadband speeds increase, concurrency can be decreased because the bits get there faster and, therefore, need access to the wire less. However, this is a historically web browsing-centric paradigm.

Alternatively, 10% @ 8 Mbps HD is equivalent from an aggregate traffic modeling perspective of 80 Mbps @ 1%, or an 800 Mbps service at 0.1%. In other words, a data service model using an HSD peak service rate and concurrent use number could be selected as is traditionally done for comprehensive traffic engineering modeling. However, these would be projections for very long-term data service needs. And, in fact, alternatives to such projections driven by historical growth patterns represent a major reason this paper was written in the first place. The net difference in the household result for different approaches to modeling the "data" service, in particular when combined with the primary video rate, is relatively small. Instead, we have chosen to use *demand*-based HD video service rates at ten-fold usage to each multi-tasking user as sufficient to describe the requirements of the secondary, personal screen – in addition to a regular screen service.

C. Traffic Engineering

In addition to the demographics of the US Household distribution of Figure 6 and the service mix possibilities of a multi-user household as described above, other traffic engineering principles are used to complete Table 5. It represents a further subset of the rate permutations, driving towards taking the HDTV 2.0, HDTV 3.0, and *Ultimate Boundary* (labeled Theoretical Worst Case in Table 5) numbers, and demographics into the area of network implications. The assumptions on formats and behaviors are repeated below for reference.

Table 5 considers each of the components above:

1) Video rate matrix samples identified in Cases 1, 2 and 3 above
2) Ultimate Boundary Conditions in Section IV-B from Tables 2 & 3
3) US Household Distribution (peak single HH for 90% coverage and a weighted average) and underage demographic (applied only to weighted average)
4) Traffic engineering for penetration (60% video and data), concurrency of peak viewing hour (50%)
5) Concurrency of multi-tasking service, HDTV 2.0 & 3.0 only (HD-rate data @ 10%)

Table 5 – Single User and Aggregate User Rates

		Single HH Composite (Mbps)		Access Network - Traffic Eng (Gbps)				Video Profiles Used
		Max Peak Rate	Weighted Avg	Aggregated HHP				
HDTV 2.0 & 3.0		> 90% US HH	US HH Sizes	250	125	50	25	
H.265	**UHDTV Only, 120 fps**	2730	811	100	50	20	10	3,4
H.265	**Max FOV Primary**	3872	3198	253	127	51	25	1,2,3
H.264	**Max FOV Primary**	7745	6396	507	253	101	51	1,2,3
Theoretical Worst Case								
H.265	**1 FOV Screen/user**	56851	19329	1791	895	358	179	5
H.265	**1 FOV Screen/user**	142128	48324	4477	2239	895	448	6

Profile #	**Video Profile***				
1	Chroma 4:4:4	FOV	300 Hz	3D	60" Pixels
2	Chroma 4:2:2	UHDTV	60 Hz	3D & 2D	60" Pixels
3	Chroma 4:4:4	HD	60 Hz	2D	60" Pixels
4	Chroma 4:4:4	UHDTV	120 Hz	3D & 2D	60" Pixels
5	Chroma 4:4:4	FOV	120 Hz	3D	20" Pixels
6	Chroma 4:4:4	FOV	300 Hz	3D	20" Pixels

*All Progressive Scan, Bit Depth = 10

HH Viewing Behavior Assumptions - HDTV 2.0 & 3.0 Only
Single FOV Primary Screen, 3D Content
UHDTV Other Screens, 3D content on one
No multi-tasking with 3D
Multitasking - HD rate to 2nd Mobile Device (i.e. IPad), High concurrency HSD
(HDTV 2.0 and 3.0 Only)

Several interesting items to note looking Table 5:

- Weighted average peaks per home of 6.4 Gbps/3.2 Gbps for H.264/H.265, 300 Hz frame rate, and full field viewing under traditional acuity guidelines

- Limiting evolution to UHDTV, the peak rate under H.265 drops to 2.7 Gbps, while the weighted average drops below 1 Gbps to 811 Mbps. This large step from 2.7 Gbps down to 811 Mbps, relative to the difference between peak and weighted average in the above cases, shows the dominance of the primary "home theatre" video maximum rate

- The 7.7 Gbps per home threshold from the table calculates out to around 2022 – thus the importance of considering H.264 throughout.

- Under Case 3 – UHDTV only @ 120 fps (standardized at 60 fps today) and H.265, the hp group sizes of 25 hp and 50 hp (30 and 15 homes, respectively, with penetration factored in), the aggregate rate could actually be supported with an evolved HFC architecture, in particular using FTLA techniques that exploit the coaxial last mile [5].

- About a half a Terabit to 4.5 terabits serves as the access network aggregate boundary condition based on uncompromised video quality weighted by demographics, traffic, and size of the aggregation group.

V. Conclusion

The calculations above, and based on what we know and in some cases are still learning, provides some insight into media consumption boundaries we might consider in long-term planning for system architectures. Again, none of the calculations is meant to presume that such a scaling of the video service will necessarily come to pass in terms of residential equipment. Rather, there merely comes a point at which engineering of the video to higher and higher quality and presence no longer provides noticeable difference, or value, to a video consumer. And, we know enough about human physiological and perceptual capabilities, as well as demographics to generate upper bonds to media consumption for home and network storage and bandwidth. These factors are worth considering in the context of an otherwise ultimately infinite projected bandwidth demand, and the build up here can inspire conversation about assumptions and trends used as inputs

A worst case bound of about 28 Gbps for a single video stream has been calculated, beyond which no viewing benefit would be gained. We estimated that this would mean about 48 Gbps peak of consumption across >90% of homes, and nearly half a Terabit to 4.5 Terabits of access capacity to enable this depending on aggregation levels.

We found that backing off of pure worst case conditions (HDTV 2.0 or HDTV 3.0), rates decrease considerably, although still challenging relative to how we think of bit rates over HFC today. Examples include an average per-home peak of 6.4 Gbps, or 3.2 Gbps under an assumption of one more generation of MPEG encoding. Under the assumption of planning for a generation of consumer home video based on UHDTV, these numbers are reduced to 2.7 Gbps and 811 Mbps Gbps for 120 fps. With UHDTV built around 60 Hz today, these numbers would be halved, and in each case further reduced under stronger 2D viewing dynamics.

Total bandwidth that consider penetration, concurrent use of video and data service, ranges from about 10 Gbps to 100 Gbps for ≤ 50 hp aggregation points under the cases described above. Such sizes are highlighted because they could represent, for example, FTLA nodes in an HFC architecture.

It is important to recognize that the idea that there is a media consumption boundary does not necessarily mean that the trajectory of bandwidth delivered to the home must asymptotically level out. Further exploitation of an increasingly capable broadband pipe may include alternative applications. Examples we can wrap our minds around today include more sophisticated machine-to-machine communications running a host of background services. The nearest example today is probably healthcare services. Non-media video – security being the most common current example – flooding networks could make round the clock video sources out of virtually every connected subscriber and lead to more downstream activity associated with the processing of this information in a cloud service. In general, processing of any data as part of a cloud service could be foreseen as part of the data exchange that takes place between home, devices, subscribers outside the home, and the information services cloud in the future. Such processes could begin to take advantage of advances in artificial intelligence. The idea that artificial intelligence will at some point augment human intelligence [7], which begins to have the look and feel of many apocalyptic movies, has also been considered an inevitable part of the accelerating technology breakthrough lifecycle. All of this falls into the “anybody’s guess” category when trying to place bets on applications that feed the bandwidth monster. The point is that it is not safe to assume that bounding the media consumption number answers the eternal question: How much bandwidth is enough? Instead, the point of the paper is merely that, as far as consumption of media by humans goes, we know enough about the important parts of the equation to effectively wrestle with it and try to quantify some outcomes.

References

Authors

[1] Armstrong, M and D Flynn, M Hammond, S Jolly R Salmon, "High Frame Rate Television," BBC Research Whitepaper WHP 169, September 2008.
[2] Deering, Michael F, "The Limits of Human Vision," Sun Microsystems, *2nd International Immersive Projection Technology Workshop*, 1998.
[3] "NHK Bets on Super Hi-Vision as Future TV", Yoshika Hara, EE Times, Sept 17, 2007
[4] Hoffner, Randy, "HDTV and the Resolving Power of the Eye," www.tvtechnology.com
[5] Howald, Robert, "Fueling the Coaxial Last Mile," SCTE Conference on Emerging Technologies, Washington DC, April 2, 2009
[6] Howard, Daniel, "Lifetime Expectancy of Coax Plant," Internal Motorola Document
[7] Kurzweil, Ray, "The Law of Accelerating Returns," KurzweilAI.net
[8] Nielsen Company, Study for Yahoo May 2010, "The American Media Multi-Tasker"
[9] Sullivan, Gary J, Pankaj Topiwala, and Ajay Luthra "The H.264/AVC Advanced Video Coding Standard: Overview and Introduction to the Fidelity Range Extensions"
Presented at the SPIE Conference on Applications of Digital Image Processing XXVII
[10] Sugawara, Masayuki, and Kenichiro Masaoka, Masaki Emoto, Yasutaka Matsuo, Yuji Nojiri, "Research on Human Factors in Ultrahigh-Definition Television (UHDTV) to Determine Its Specifications"

Websites

[11] http://amo.net
[12] http://en.wikipedia.org – Optimum HDTV Viewing Distance
[13] http://krusekroncile.typepad.com
[14] http://myhometheatre.homestead.com
[15] http://psychedelic-information-theory.com "Limits of Human Perception" (Chapter 5)
[16] http:// schubincafe.com
[17] www.100fps.com
[18] www.carbonbale.com
[19] www.census.gov
[20] www.clarkvision.com
[21] www.poynton.com
[22] http://www.practical-home-theater-guide.com

Acknowledgements

Many thanks are due Ajay Luthra for his consistent willingness to share his in-depth knowledge of video and encoding technology.