Reaching the Goals in the National Broadband Plan

September 3, 2010

ADTRAN®

Abstract

The Federal Communications Commission (FCC) has set several goals for consumer broadband access in the National Broadband Plan, including:

- 100 million homes with access to 50 Mbps downstream and 20 Mbps upstream by the year 2015.
- 100 million homes with access to 100 Mbps downstream and 50 Mbps upstream by the year 2020.

This white paper addresses the capability of Gigabit Passive Optical Network (GPON), Active Ethernet, VDSL, and DOCSIS 3.0 to meet those goals in the last mile. It also estimates the backhaul capacities that will be required to meet the goals in access networks of different sizes.

We find that all of the last mile technologies listed above are capable of meeting the FCC goals, although DOCSIS 3.0 will require significant network engineering constraints to meet the year 2020 goal. We also find that the backhaul capacity required to meet the year 2020 goal varies over a wide range. The required capacity will depend much more on the growth in user demand between now and then than on target rates set by the FCC.

Section 1 Introduction

In its National Broadband Plan (NBP) [1], the FCC has set a goal that “100 million U.S. homes should have affordable access to actual download speeds of at least 100 Mbps and actual upload speeds of at least 50 Mbps by 2020.” An intermediate milestone related to this goal is that “by 2015, 100 million U.S. homes should have affordable access to actual download speeds of 50 Mbps and actual upload speeds of 20 Mbps.” While the NBP never specifically defines the term “actual,” it does use the term in reference to average speed,[1] median speed,[2] and “actual speeds and performance achieved with a given probability (e.g., 95 percent) over a set time period (e.g., one hour) that includes peak use times.”[3]

This white paper examines the ability of access networks to meet the NBP goals listed above. The FCC refers to the 95th percentile when addressing “technical broadband measurement standards and methodology” (in Recommendation 4.3), so that definition is used for “actual speed” throughout this paper. For last mile technologies that share capacity between multiple users, and for backhaul connections, required capacity is estimated using a high level Monte Carlo simulation. Brief descriptions of the Monte Carlo simulation and the principles behind it are provided in appendices at the end of this paper (*Appendix I or Appendix II*).

Estimating required capacity requires that demand first be estimated, which is done in *Section 2*. The demand projections derived in that section are applied to the simulations used to estimate required capacity in later sections. *Section 3* addresses the speeds that can be achieved by last mile technologies including GPON, Active Ethernet, VDSL, and DOCSIS 3.0. *Section 4* estimates the backhaul capacities required by second and middle mile networks of different sizes and derives a linear approximation that can be used to predict required capacity.

All simulation results are shown to two significant digits. Since the simulation results are based on demand values predicted up to 10 years into the future, as well as analytic, non-specific demand distributions, they need to be interpreted more as illustrating the relationships between the different network parameters, and less as a prediction of the absolute capacities that will be required out to the year 2020.

Section 2 Projected Demand

If we are to assess network requirements against the NBP goals, we need to project expected traffic demands out through the year 2020. A projection this far into the future is by definition speculative, so rather than trying to predict an exact value we will use a range of values.

2.1 Mean Busy Hour Traffic

A previous ADTRAN® white paper [2] derives mean busy hour traffic on a per-household basis, using data from the Cisco® Visual Network Index [3] and other sources. *Table 1* shows the year-over-year growth rates for busy hour traffic as derived in [2]. The overall Compound Annual Growth Rates (CAGRs) for years 2009-2014 are 31 percent and 25 percent for the downstream and upstream directions, respectively. However, the year over year numbers show a trend with the highest increases in 2010 and 2011, followed by slower growth, and the 2013-2014 increase is only 17 percent and 18 percent in the downstream and upstream directions. We’ll do projections based on the highest and lowest year-over-year values as well as the five-year CAGR in the hopes of bracketing the actual figure. The results are shown in *Table 2*.

Table 1—Year-over-year Growth for Busy Hour Traffic

Direction	2009-2010	2010-2011	2011-2012	2012-2013	2013-2014	CAGR 2009-2014
Down	45.1%	48.1%	26.9%	22.5%	17.0%	31.3%
Up	28.6%	32.7%	22.8%	20.9%	17.9%	24.5%

1, 2 National Broadband Plan, page 21.
3 National Broadband Plan, page 45.

Since *Table 1* shows a trend in which year-over-year growth rates decrease over time, it may be tempting to just extrapolate the results past 2014 using the lowest growth rates (for the 2013-2014 period in both directions). Several factors support a more conservative approach, however. First, the traffic figures projected for years further into the future are by nature less reliable than near term values, so basing long term growth on GAGR values predicted three or four years in advance is inadvisable. Second, the Cisco data may be conservative. While they have estimated global Internet traffic growth rates from 34 percent to 40 percent in recent updates to the VNI, Minnesota Internet Traffic Studies (MINTS) has estimated volume-weighted annual growth rate from 2002 to 2009 at about 67 percent [4], and Cisco has tended to revise their data upwards with new updates to the VNI.

Table 2—Mean Busy Hour Traffic Projections to Year 2020

Direction	Estimate	CAGR	Mean Busy Hour Traffic (kbps)					
			2015	2016	2017	2018	2019	2020
Down	High	48.1%	645	956	1,415	2,094	3,101	4,591
	Mid	31.3%	573	752	988	1,297	1,703	2,237
	Low	17.0%	510	597	699	818	957	1,120
Up	High	32.7%	120	159	211	280	372	494
	Mid	24.5%	113	140	175	217	270	337
	Low	17.9%	107	126	148	175	206	243

2.2 Demand vs. Traffic

The Cisco VNI values, and the ADTRAN busy hour traffic values derived from them, forecast traffic (the load carried by the network) rather than demand (the load offered to the network by users). While we ideally need demand projections in order to estimate performance, we are forced to use traffic projections as a proxy since we know of no sources forecasting user demands directly. If the offered load and the carried load are approximately equal, the substitution should not make a substantial difference in the results. We need to address the question of how close the two values typically are under the conditions of interest.

In a lightly loaded network, congestion is very infrequent and offered load and carried load are nearly equal. Conversely, in a heavily congested network the carried load may be only a small fraction of the offered load. Our simulations usually emphasize the region between these two extremes, in which congestion occurs just frequently enough to drive performance below the target rate approximately five percent of the time. When simulating these threshold conditions we find that the difference between mean offered and carried loads is less than 10 percent (frequently less than five percent). So while the two values are not equal, the difference seems acceptable given that they remain within a percentage range smaller than the range of traffic values predicted for the year 2015 in *Table 2*.

2.3 Demand Distribution

The values projected in *Table 2* do not address the distribution of the demand, only the average. The Monte Carlo simulation uses a bounded Pareto distribution for the demand. This is an analytic demand model which doesn't address specific applications.

There are two good reasons for using this model as opposed to a model that takes any specific set of applications (such as streaming video) into consideration. First, an application-specific model requires a large amount of data. There are many classes of applications (current consumer applications include streaming audio and video, peer-to-peer, VoIP, online gaming, email, Web browsing, interactive video, and ambient video), each with its own set of parameters for session initiations, flow size, clustering, etc. The parameters are not static—for instance, the average size for Web pages grew by a factor of five from 2001 to 2009 [5]. A meaningful application-dependent demand distribution would have to include all of that data, some of which is not readily available in the literature.

Second (and more important), the mix of applications in use today is not likely to be representative five or 10 years from now, nor can we predict what it might be at that time. Predicting the future of the Internet is notoriously difficult, and new "killer apps" have been known to appear almost literally overnight [6].

While we can't predict the future mix of applications, a general understanding of statistics coupled with historical data [6, 7] allows us some confidence in the Pareto distribution as an abstract model. Absent some regulating mechanism to the contrary (such as bandwidth limits or penalty pricing), it's reasonable to expect that a small minority of users will demand a disproportionately large share of resources and that the large majority of users will use less than the average for the population. The Pareto distribution provides a good fit for this model.

Finally, we note that the high end of the range shown in *Table 2* is nearly enough to support continuously streaming HD video in every household served by an access network. While we don't expect that to be the case, it shows that the range in the table covers a wide spectrum of potential applications.

All shared capacity and rate results shown in this section are generated by Monte Carlo simulations as described in Appendix II. A total of 10,000 trials are simulated in each Monte Carlo set. The result for each trial is the speed achieved by a single user attempting to receive (or send) data as fast as possible, within a population of users with random demands based on the defined distribution. The overall result for the simulation as used here is the 95th percentile speed from the set of speeds generated by the Monte Carlo trials.

3.1 GPON

GPON shares last mile capacity between the subscribers (typically 32 subscribers, but the number can be 64 or even a maximum of 128) in a Passive Optical Network (PON). Since GPON shares last mile capacity among multiple subscribers, its capacity needs to be tested against the NBP requirements. *Table 3* shows the 95th percentile rates achievable in year 2020 using current generation GPON (2.5 Gbps downstream, 1.25 Gbps upstream) shared by the maximum value of 128 users on a single PON.

Table 3—Rates for GPON with 128 Users, Year 2020 Projected Demands

CAGR	95th percentile rate (Mbps)	
	Down	UP
High	1,100	1000
Mid	1,700	1,000
Low	2,000	1,100

Even for the worst case conditions shown in *Table 3*, the rates shown are at least an order of magnitude better than the 100/50 Mbps requirements in the NBP. Given these results, we can consider the GPON last mile to be virtually transparent relative to those requirements. Put another way, the user rates achieved on GPON access networks (designed to meet NBP goals of 100 Mbps downstream and 50 Mbps upstream) will be dependent on second and middle mile design, rather than the last mile.

GPON's shared bandwidth is sufficiently high to support video services in any of a number of formats in addition to Internet service. For instance, if 1,500 Mbps is reserved for a combination of multicast and unicast video services, the remaining 1,000 Mbps is still enough to support a 95th percentile rate of 140 Mbps downstream with 128 users.

The values in *Table 3* are for current generation GPON, or ITU-T Recommendation G.984. Future generations of the technology (XGPON1 or 10GEPON) will support substantially higher rates under the same conditions.

3.2 Active Ethernet

With 1 Gbps dedicated to each user in each direction, Active Ethernet in the last mile supports the NBP goals by definition. As with GPON, the user rates achieved on Active Ethernet access networks will be dependent on second and middle mile design.

3.3 DSL

DSL in the last mile is dedicated to a single subscriber, so the actual rate that can be delivered over the last mile is equal to the peak rate. The peak rate delivered by DSL varies by loop length and by the way in which the technology is deployed. *Table 4* shows the loop reaches over 24 AWG copper loop (from Broadband Forum) at which VDSL2 can deliver the NBP target rates under the following deployment scenarios:

- VDSL2 delivered over a single pair.
- VDSL2 with vectoring. Vectoring is a technique in which Multiple Input Multiple Output (MIMO) signal processing techniques are applied to exploit the loop-to-loop transmission paths for their information value, significantly increasing the effective bandwidth on the cable as a whole.
- VDSL2 with two-pair bonding, in which two copper pairs are used to deliver services to each subscriber.
- VDSL2 with vectoring and two-pair bonding.

Loop reaches are shown both for the upstream/downstream combined goals and for rates that meet the downstream goal only (with the corresponding upstream rate shown in parentheses).

Table 4— VDSL2 Loop Reach at FCC Target Rates (24 AWG)

Target Rates	VDSL2	VDSL2 with vectoring	Bonded VDSL2 (2 pairs)	Bonded VDSL2 with vectoring
100/50 Mbps (2020 goal)	N.A.	1,350 ft	1,400 ft	2,200 ft
50/20 Mbps (2015 goal)	1,500 ft	2,500 ft	2,600 ft	2,900 ft
100 Mbps down	N.A.	1,400 ft (45 Mbps)	1,500 ft (46 Mbps)	2,700 ft (32 Mbps)
50 Mbps down	1,500 ft (23 Mbps)	2,700 ft (16 Mbps)	4,400 ft (8.4 Mbps)	5,000 ft (6 Mbps)

The loop reaches shown in *Table 4* are based on currently defined frequency band plans. Modifications to the existing plans would allow additional enhancements in loop reach optimized to the FCC goals.

As long as the dedicated last mile provides at least the required rate, the FCC goals for year 2015 and 2020 can be met by DSL access networks. As with all access networks, the second and middle mile design needs to support the required rates (see *Section 4*)—however, the last mile design can be determined independent of average demand predictions.

3.4 DOCSIS 3.0

DOCSIS 3.0 supports channel bonding, so providing the rates required by the NBP is primarily a matter of finding the right combination of the number of bonded channels and the number of subscribers sharing those channels. In the downstream direction, each channel supports a shared capacity of 38 Mbps.

Channel capacity in the upstream direction is dependent on the number of subscribers sharing the channel. As long as the subscriber population is relatively low (on the order of 200 or less), each channel supports a shared capacity of about 27 Mbps. Once the number of subscribers grows beyond that limit (which is not fixed, as it depends both on population and coaxial plant layout) the shared capacity for each channel drops to about 17.5 Mbps due to noise funneling. Networks with more than 200 subscribers are not considered in the figures below.

Figure 1 and *Figure 2* show the numbers of bonded channels required to provide service to different numbers of subscribers with performance meeting the NBP year 2015 and year 2020 goals. The figures show that the number of channels dedicated to data transmission can be traded off against the network size. Year 2015 goals can generally be met with three to seven bonded downstream channels and two to three bonded upstream channels, serving populations that range from 32 to 192 subscribers. The Year 2020 goals are much more difficult. Depending on annual growth in demand, DOCSIS 3.0 networks may need as many as 33 bonded downstream channels to serve 192 subscribers. The upstream 2020 goal is slightly easier, requiring from three to eight bonded channels.

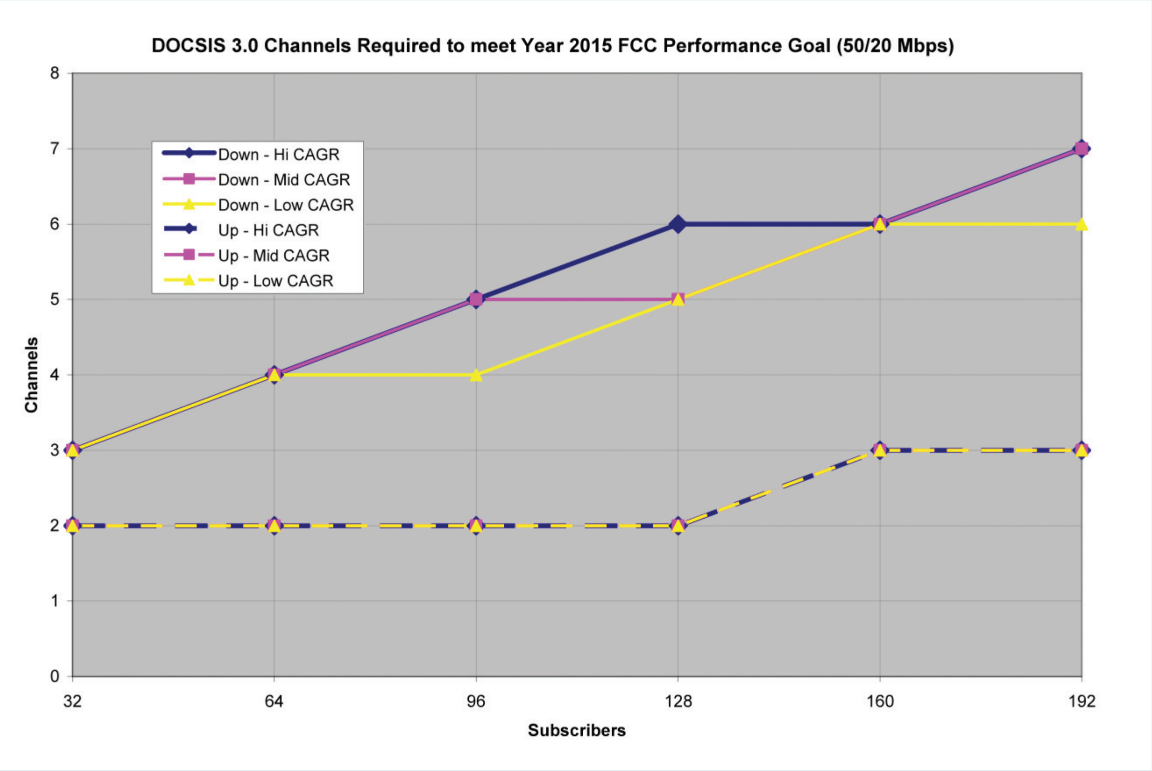

Figure 1—DOCSIS 3.0 Channels Required, Year 2015

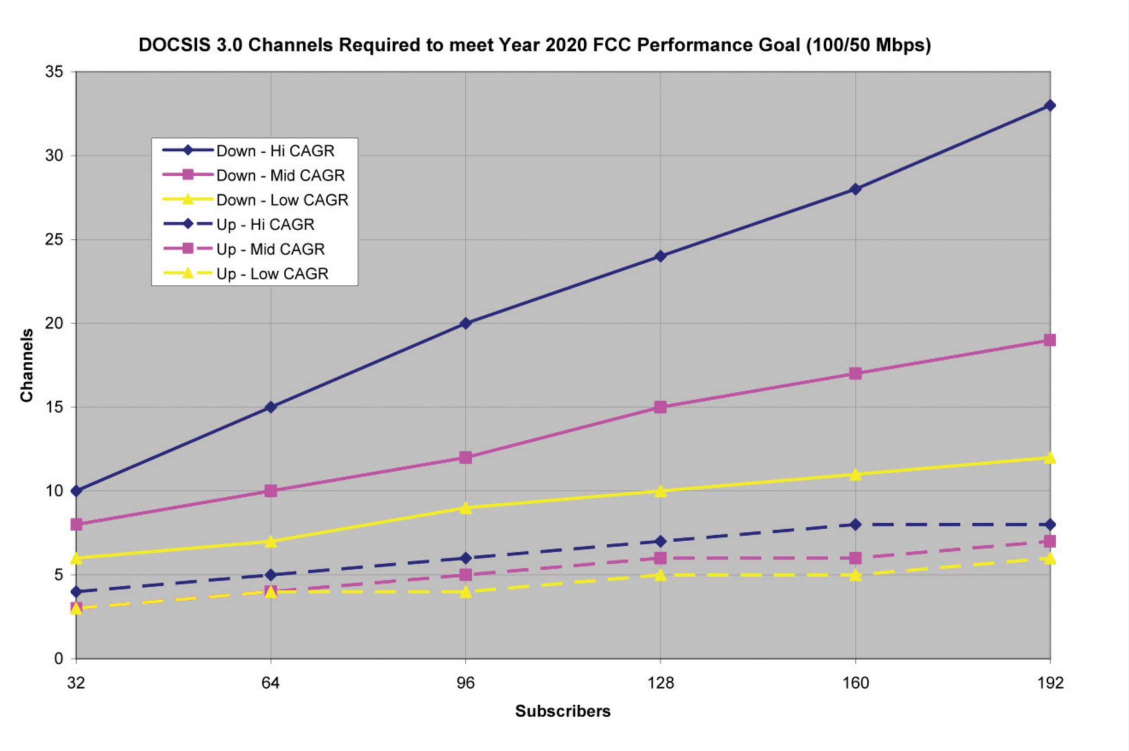

Figure 2—DOCSIS 3.0 Channels Required, Year 2020

Section 4 Second/Middle Mile Capacity

Rates in the aggregation network can be simulated largely independent of the technology in the last mile. While simulation results vary slightly for the different technologies due to different bounds on the input distributions, these variations are primarily artifacts of the distribution model. More importantly, any variation due to last mile technology is swamped by the much larger variation due to the range of projected demands, especially for year 2020 results. So, a single set of required backhaul capacities is presented that can be considered applicable independent of the last mile technology.

Backhaul capacity requirements were generated for access subnetworks with 32, 128, 512, 1024, and 2048 subscribers. The backhaul capacities required to meet the NBP goals for year 2015 are shown in *Table 5*, and the capacities required to meet year 2020 goals are shown in *Table 6*. As with the previous section, all capacity and rate results are generated by Monte Carlo simulations with 10,000 trials in each set.

Table 5—Backhaul Capacity Required to Meet 2015 Goals (50 Mbps down/20 Mbps up, 95th percentile)

CAGR	Direction	# subscribers				
		32	128	512	1,024	2,048
High	Down	105 Mbps	200 Mbps	500 Mbps	870 Mbps	1,650 Mbps*
	Up	32 Mbps	55 Mbps	115 Mbps	190 Mbps	330 Mbps
Mid	Down	100 Mbps	185 Mbps	460 Mbps	790 Mbps	1,500 Mbps*
	Up	31 Mbps	53 Mbps	110 Mbps	180 Mbps	310 Mbps
Low	Down	95 Mbps	175 Mbps	420 Mbps	720 Mbps	1,350 Mbps*
	Up	31 Mbps	51 Mbps	105 Mbps	175 Mbps	300 Mbps

Table 6—Backhaul Capacity Required to Meet 2020 Goals (100 Mbps down/50 Mbps up, 95th percentile)

CAGR	Direction	# subscribers				
		32	128	512	2,048	2,048
High	Down	380 Mbps	920 Mbps	3,000 Mbps*	5,900 Mbps*	12,000 Mbps*
	Up	94 Mbps	170 Mbps	410 Mbps	690 Mbps	1,300 Mbps
Mid	Down	270 Mbps	550 Mbps	1,550 Mbps	2,900 Mbps*	5,800 Mbps*
	Up	82 Mbps	145 Mbps	310 Mbps	510 Mbps	920 Mbps
Low	Down	200 Mbps	370 Mbps	900 Mbps	1,550 Mbps	2,900 Mbps
	Up	76 Mbps	125 Mbps	250 Mbps	400 Mbps	700 Mbps

As we start to apply higher capacity cases to the high level Monte Carlo simulation, the results start to show unrealistically high values for average utilization. As the network capacity grows significantly larger than either the average demand or the target rate, the simulated target rate becomes achievable under conditions that start to resemble network congestion. While in a real network this would not be tolerated due to indicators such as increasing packet loss; the high level simulation does not account for this. So, a check is added to make sure that average utilization will not exceed 80 percent. This is indicated in the above tables by asterisks on values which have been modified due to this check.

Note that the backhaul capacities required to meet the year 2020 goals are up to seven times as large as those required to meet the year 2015 goals, even though the target rates themselves are only twice as large. This is because between the years 2015 and 2020, average demand increases substantially—depending on the CAGR used, the increase is from twofold to sevenfold. The capacity required to serve an access network scales roughly with the product of the average demand times the number of users. The required capacity is much less dependent on the desired actual rate (which is the inverse of saying that the actual rates achieved are very sensitive to small percentage changes in network capacity).

When graphed, the values in *Table 5* and *Table 6* show a nearly linear trend for networks with more than 100 subscribers, so a linear approximation may be appropriate to generalize the results. We know that as the network gets larger and the overall average demand becomes the dominant factor in the capacity required, the Monte Carlo simulation results become unrealistically high and we must limit the minimum required capacity so that average utilization doesn't exceed a threshold (we are using 80 percent). This implies that the slope for the linear approximation vs. demand should not be less than 1/80% = 1.25. We also know that as the network gets smaller, the target rate becomes increasingly dominant in determining required capacity. Finally, we want a conservative approximation that may overestimate, but will not underestimate, the required capacity.

We achieve an approximation that satisfies the above criteria by setting the slope to 1.25 and choosing the lowest Y-intercept, normalized by the desired target rate that results in all approximated values being at least as large as all values resulting from the Monte Carlo simulations. *Equation 1* provides the approximation, where:

- C = the required capacity,
- n = the number of subscribers served by the network,
- d = the average demand per subscriber, and
- r = the desired rate at the 95th percentile.

$$C = 1.25nd + 2.5r \qquad \textbf{(Equation 1)}$$

Figure 3 and *Figure 4* show required capacity for the year 2015 and 2020 goals, respectively. The lines show capacity estimated by Equation 1 and the discrete points show the simulation results from *Table 5* and *Table 6*.

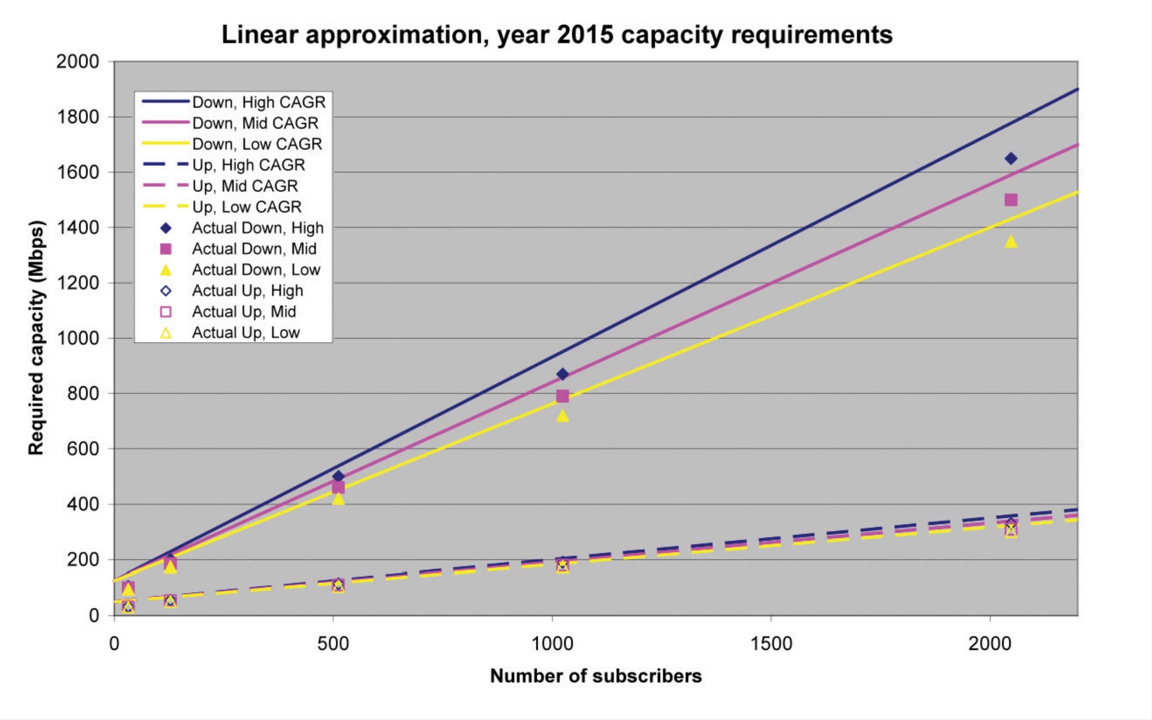

Figure 3—Estimated vs. Simulated Required Capacity, Year 2015

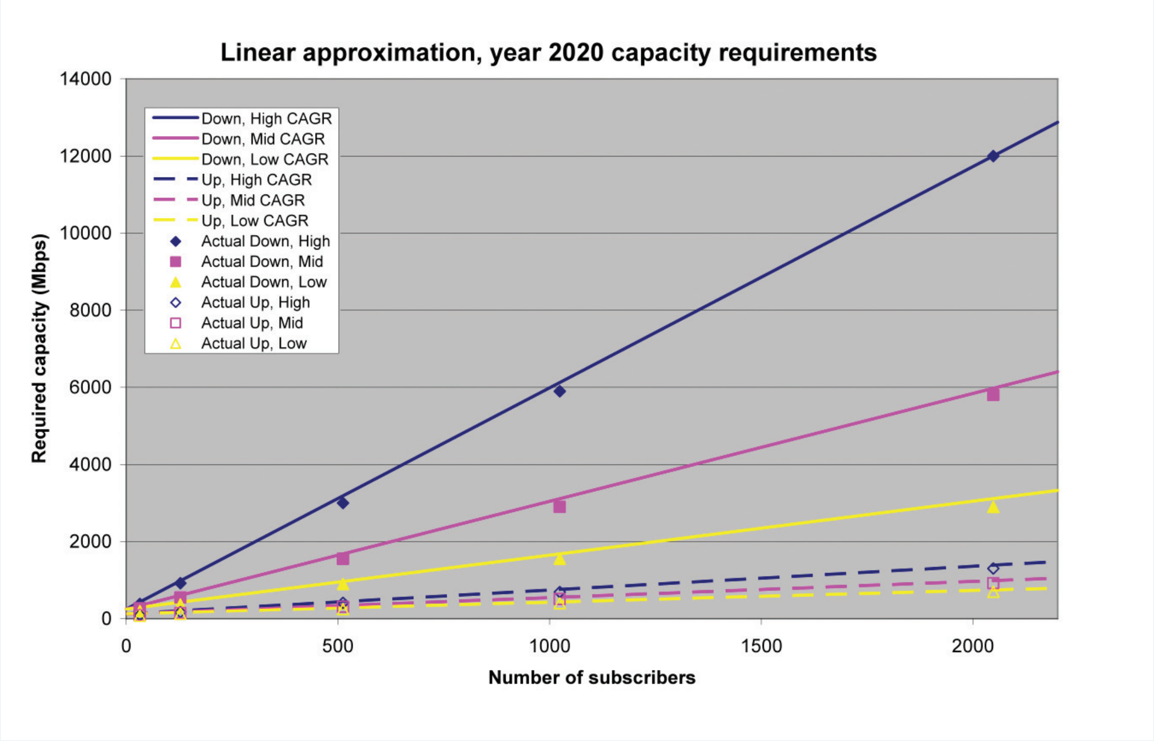

Figure 4—Estimated vs. Simulated Required Capacity, Year 2020

Section 5 Conclusions

The network capacities required to meet the goals outlined in the NBP are dependent not only on the goals themselves, but also on the parameters of the demand that users will place on access networks as the target years approach. Demands for the target years are predicted using high, mid-level, and low CAGRs derived from the most recent Cisco Visual Networking Index.

Any of the last mile technologies included in this white paper can meet the year 2015 and year 2020 goals listed in the NBP. Some of the technologies (GPON and Active Ethernet) meet the goals with an order of magnitude in margin. VDSL2 combined with vectoring and/or pair bonding meets the year 2020 goals. Depending on long term traffic growth and the size of the subscriber pool served, DOCSIS 3.0 may require as many as 33 bonded channels to meet the year 2020 goals.

Required backhaul capacity is simulated independent of the last mile technology and is found to have a linear trend for access networks serving hundreds or thousands of users. A linear approximation to estimate required capacity is derived from the simulation results.

The results show that, especially for larger networks, the required network capacity is much more dependent on the demand placed on the network by users than it is on the desired rate. The expected growth in demand between now and 2020 is such that the 100 Mbps "actual rate" goal set by the FCC will have relatively little effect on the backhaul capacities required that far in the future.

Section 6 References

[1] Federal Communications Commission, "Connecting America: The National Broadband Plan," March 16, 2010.

[2] ADTRAN, "Updated busy hour consumer Internet traffic projections," June 24, 2010.

[3] Cisco, "Cisco Visual Networking Index – Forecast and Methodology, 2009-2014," June 2 2010, http://cisco.biz/en/US/solutions/collateral/ns341/ns525/ns537/ns705/ns827/white_paper_c11-481360_ns827_Networking_Solutions_White_Paper.html last accessed on June 23 2010.

[4] http://www.dtc.umn.edu/mints/2002-2009/analysis-2002-2009.html, last accessed on June 23 2010.

[5] ADTRAN, "Defining Broadband: Network Latency and Application Performance," June 18 2009.

[6] Floyd, S. and Paxson, V., "Difficulties in Simulating the Internet," IEEE/ACM Transactions on Networking, 2001, volume 9, pp 392-403.

[7] Gerber, A., Houle, J., Nguyen, H., Roughan, M., and Sen, S., "P2P, The Gorilla in the Cable," 2003, available at http://www.research.att.com/~sen/pub/p2pCable2003.final.pdf

[8] Jha, S. and Hassan, M., "Engineering Internet QoS," Artech House, Boston, 2002.

Appendix I: Capacity, Demand, and Actual Speed

The “actual rate” experienced by a user on a broadband network is a bit of a misnomer. Even when a user is actively using a broadband connection, the activities performed most of the time do not use all of the available bandwidth on the network connection. For example:

- Non-real time activities like Web browsing and email access place intermittent, bursty demands on the network. Due to the nature of the communication protocols used for these applications (such as TCP and HTTP), factors such as end-to-end network latency frequently prevent the data transfer from taking place at the available network rate, even during the bursts of traffic.
- Streaming media from sites like YouTube requires the transfer of large files. While faster transfer minimizes the time required to fill the playout buffer and begin playout, transfer rates that are many times faster than the playout rate do nothing to enhance the user’s experience once playout has started. So, servers that source streaming media usually split it into smaller pieces and interleave pieces destined for different users, minimizing the response time for any one user but also limiting the transfer rate below what the network might otherwise support.
- Real time applications like VoIP and video calls operate at a given rate (which may be adjustable in specified increments). They do not use excess available bandwidth.
- Of course, even when a user is actively consuming data that has been delivered over the network connection (such as reading a Web page), the connection itself may be idle.

So when we measure “actual rate,” we are actually measuring the rate that a user can get when he requests as much as possible from his connection—that is, we are generating a test to measure that rate. The resulting rate varies over time due to network capacity, traffic scheduling and management techniques, and the traffic being generated by the community of users sharing the network. While the interactions between these elements are complex, in many cases the “actual rate” can be summed up by the following (admittedly oversimplified) statement:

When a user measures his actual rate, the result reflects the bandwidth available on the network—that is, the bandwidth that is not being used by everyone else.

The above statement stems from the fact that most users are not trying to maximize use of their connection at any given time, but that the user testing his performance is doing just that. While this sounds artificial (and a test is always somewhat artificial by definition), the results are still valid—they reflect the actual performance the network is capable of providing when required. The key point is that the resulting performance is as dependent on other users as on the network itself. Specifically:

- For a network with fixed capacity, as the average user demand increases, the actual rate decreases.
- The actual rate is sensitive to relatively small changes in average demand.
- The actual rate is also sensitive to relatively small changes in network capacity.

The above points are illustrated by the following example. Consider a network serving 600 users with a fixed total bandwidth of 100 Mbps. The demand placed on the network varies over time, with an average of 100 kbps per user. When the total demand from all users is summed, it averages 60 Mbps, but it varies such that the 5th and 95th percentile demands are 50 Mbps and 80 Mbps, respectively. This is illustrated in *Figure 5*.

Figure 5—Example of Capacity, Demand and Actual Rates

When a user measures his actual rate by performing a speed test, the result will correspond to the bandwidth available on the network at the time of the test. 95 percent of the time, that user will get a result (ignoring factors like peak rate caps) of 20 Mbps or better.

What happens if the average demand increases by a small factor? *Figure 6* shows the same network with average demand increased by 20 percent. Now, the 95th percentile actual rate is only 4 Mbps—a fivefold decrease.

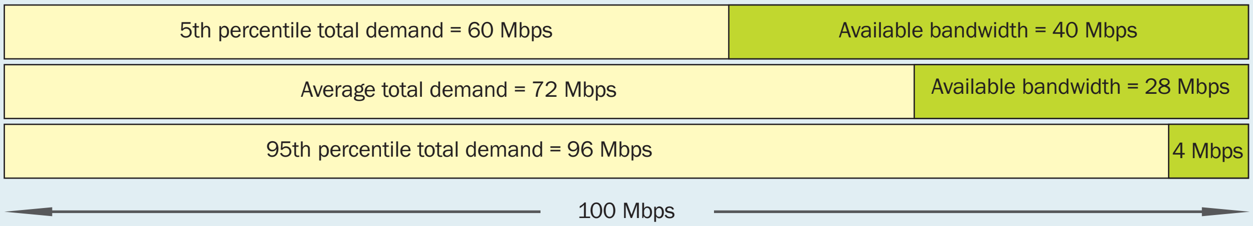

Figure 6—20 Percent Increase in Average Demand

Appendix I: Capacity, Demand and Actual Speed (continued)

Restoring the 95th percentile performance (a fivefold increase) requires an increase in capacity of only 16 percent, shown in *Figure 7*.

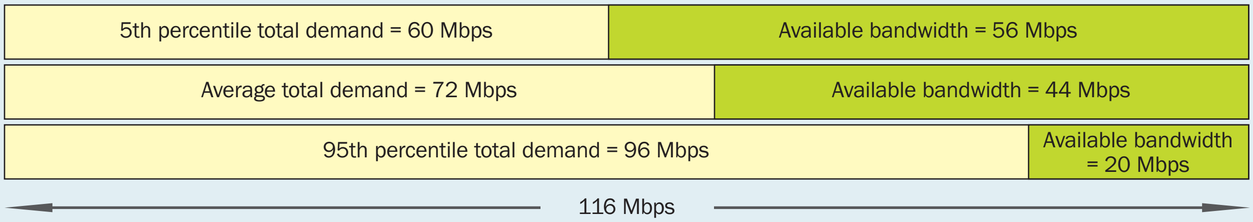

Figure 7—95th Percentile Performance Restored

Of course, performance testing on real broadband access networks is more complex than the previous statement and example, but the basic concept illustrated above is true in the general sense. Some additional factors include:

- When a user initiates a speed test, this application usually replaces any other demand he might have contributed to the total on the network.
- The actual rate results from the user's traffic being scheduled with traffic from other users, usually at multiple points in the network. There are many ways to manage traffic that can affect the outcome of a performance test.
- Even if the traffic management is ideally fair (in the sense of max-min fairness as defined in [8]), the actual rate will reflect the available bandwidth only if each other user's demand is less than that available bandwidth. Otherwise, the bandwidth will be allocated between users such that each gets a fair share.
- Performance results are also affected by service limitations such as bandwidth caps, peak rate limits, and differences between individual service contracts.

Appendix II: Simulation Methodology

The high level simulation model is designed to allow rapid estimation of expected speeds for a given set of input conditions. For each trial, the load offered by each user is generated as a single value of a random variate with the desired load distribution. This value (expressed in bits per second) represents the offered load normalized over a time which is both long enough to mask packet-level effects and short enough to be of interest. Although the time period is unspecified, it can be thought of as on the order of seconds—as in the length of time it might take a user to test the speed of his connection.

There is no assumption that each user is attempting to access the network at a constant speed during this period of interest. A better way to interpret the offered load would be as an average value resulting from each user trying to transmit data at the maximum allowed rate over some fraction of the time period of interest.

After all offered loads are generated, the load for a single user is replaced by the maximum load supported by the distribution. The carried load achieved by this user (who is now, by definition, attempting to transfer as much data as possible over the entire period of interest) is the primary outcome for each trial.

The carried load for each user is determined using max-min fair share (aka maximum fairness) [8], an idealized algorithm which is approximated by implementations such as deficit round robin and weighted fair queuing. In max-min fair share, flows (beginning with the smallest loads) receive their requested bandwidth up to a maximum beyond which total utilization would exceed 100 percent. The remaining flows each receive the maximum value. An example is shown in *Figure 8*, where the offered load (shown in blue) totals 127 Mbps but the network capacity is only 100 Mbps. In the example, flows with offered load of less than 6.7 Mbps receive the full requested rate and the remaining flows each receive 6.7 Mbps, resulting in a total carried load (shown in red) of 100 Mbps.

Max–min Fair Share Example

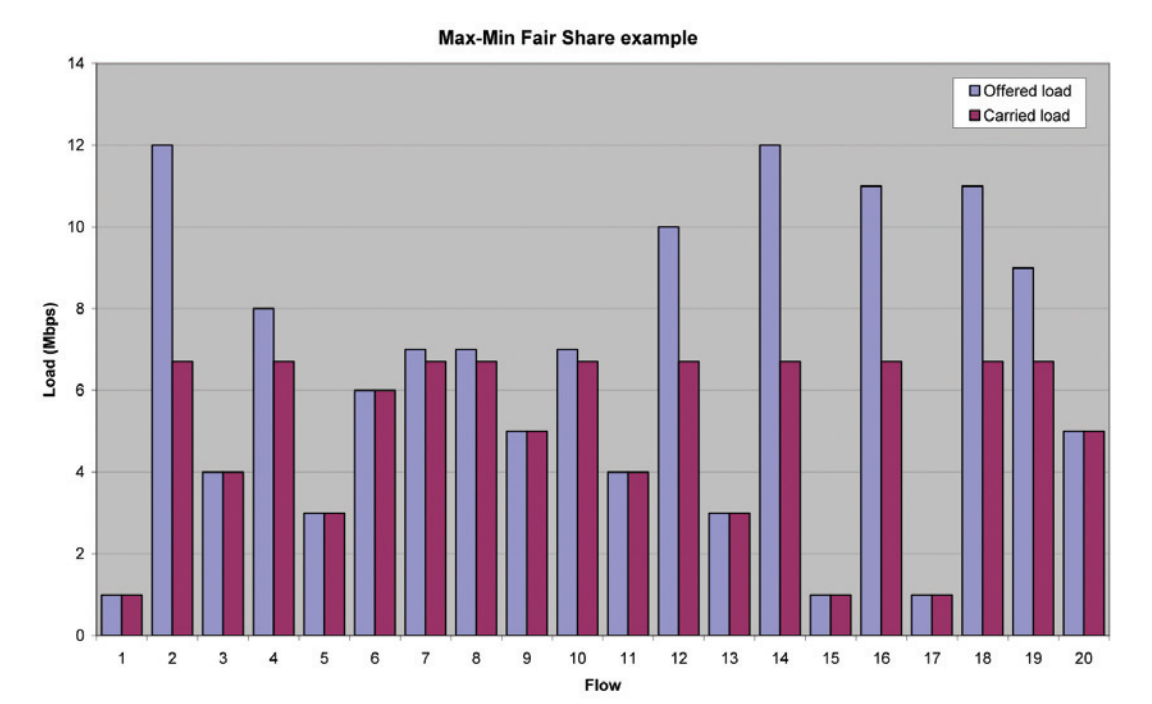

Figure 8—Max–min Fair Share

As noted above, the primary result from each trial is the carried load achieved by the test user. Additional results such as minima and maxima are also saved. The Monte Carlo simulation performs the requested number of trials and generates an output file with individual trial results and a summary of statistics for the set of trials.