THE IMPACT OF GOING FIBER SPEEDS OVER COAX FEEDS

Alon Bernstein, Cisco Systems, Inc.
Joe Godas, Cablevision

Abstract

Data rates over cable networks are constantly increasing due to higher bandwidth demands from consumers as well as competitive pressure. With DOCSIS 3.0 the cable plant can offer speeds that rival those offered by PON (Passive Optical Networks) technologies and meet consumer demands.

Building an end-to-end solution that will support these data rates is not only about adding faster network components and updating the user service agreements. This paper will explore the specific impact of supporting data rates in the order of 100 Mbps and above over the cable network.

OVERVIEW

With DOCSIS 3.0, the typical traffic profile for a cable subscriber is going to increase from the 10mbps-30mbps range to the 50mbps-100mbps range. These traffic rates compete well against end-to-end optical technologies.

In theory it takes a DOCSIS 3.0 system and a change in the modem configuration file to get to the above rates. In practice, the added throughput has implications on the upper layer application behavior, network design and equipment design. This paper will list what these implications are and how they can be addressed in order to unleash the full potential of DOCSIS 3.0.

Most of the issues described here are not unique to cable. Any service provider network that will serve 100mbps+ speeds will have to deal with them. However, DOCSIS 3.0 is the technology that would allow, for the first time, residential costumers to have access at these rates and so it will be the first time cable service providers manage such a high speed network.

IMPACT OF HIGHER SERVICE CONTRACTS ON AGGREGATE TRAFFIC

Even before DOCSIS 3.0, in order to compete with telcos, cable Multi-Service Operators (MSOs) have been steadily increasing their traffic contracts from 2.5mbps down/500k up to 30mbps down/5 mbps up. Looking at the lessons learned from past increases will help to infer what will happen with future increases.

While one might expect that the aggregate trend would linearly track the service contract increase, or in other words, that a doubling of a service contract will result in a doubling of the aggregate traffic, it is evident this is not the case.

The following graphs represent aggregate traffic as measured at the internet border in the Cablevision network at different periods over the years.

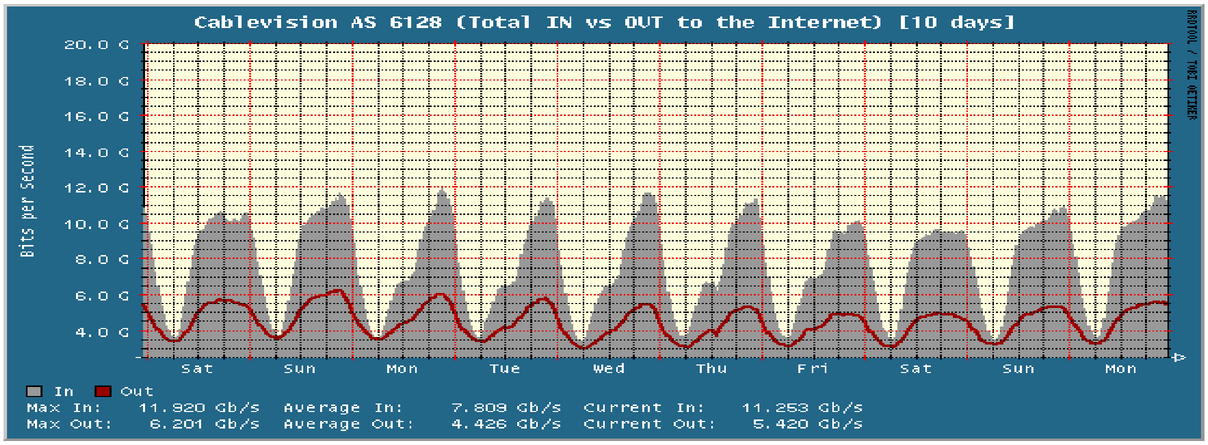

Figure 1: 1Q2004

Figure 1 depicts the aggregate traffic during 2004Q1 when an individual contract was 10mbps down, 1mbps up. As can be seen, the traffic peaks are around 12Gbps and the lows are in the 4Gbps range. The gap between the highs and lows is roughly 8Gbps.

18 months later we see traffic demands ramp up quickly as subscribers are added to the system. The modem contracts haven't changed but downstream consumption has almost doubled due to a 45% increase in subscriber acquisition (see Figure 5). Now the peaks are in the 22Gbps range, and the lows at 7Gbps. The spread has increased to 15Gbps, much more then the linear increase in subscriber bandwidth.

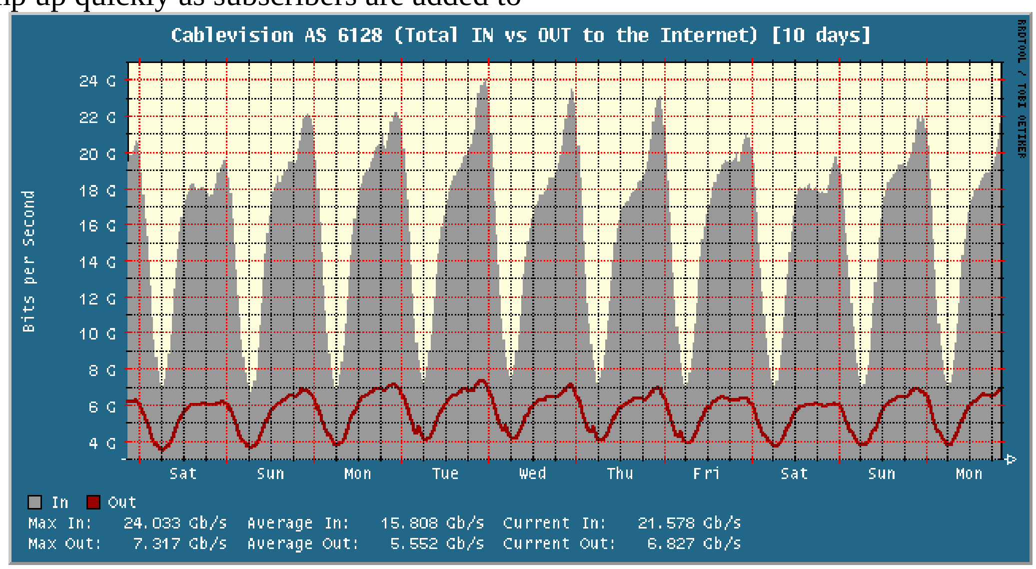

Figure 2: 3Q2005

In 2Q2006 (Figure 3) -The traffic contracts for the entire subscriber base were increased to 15mbps down and 2mbps up. The peaks rose to about 40Gbps, and the dips to 12Gbps. The spread is 38Gbps, and keeps growing much faster then expected. In this time period, a 19% increase in subscriber acquisition has produced a whopping 100% increase in traffic consumption. This surge in demand is attributed to a

50% increase in the modem contract burst capability (going from a 10Mbps to 15Mbps downstream contract).

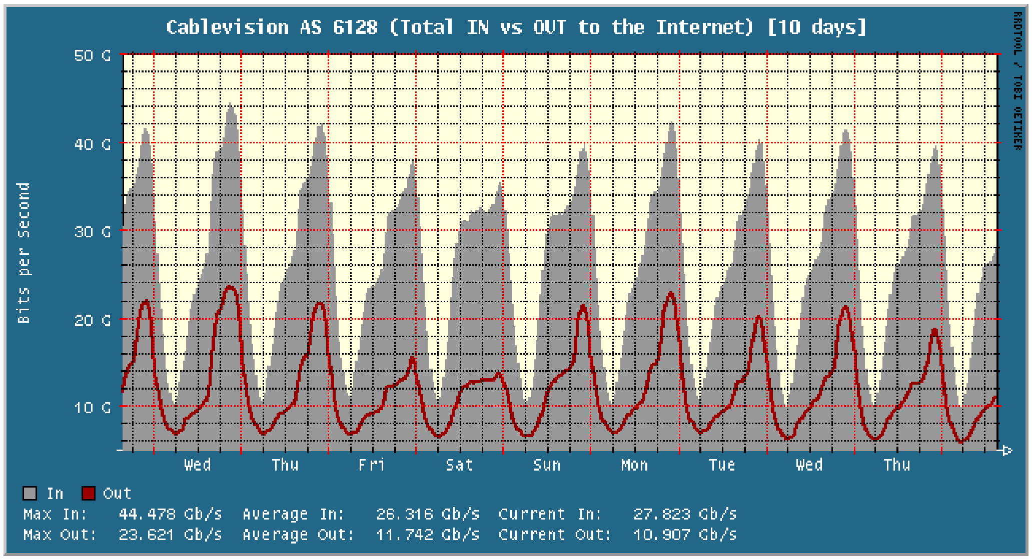

Figure 3: 2Q2006

1Q2007 is especially interesting. The first thing to note is that it is barely 6 months since 2Q2006 and corresponds to a subscriber increase of roughly 9%. Curiously enough, while the peaks have increased by 10Gbps, the lows have hardly moved from their 2Q2006 range (about 12Gbps).

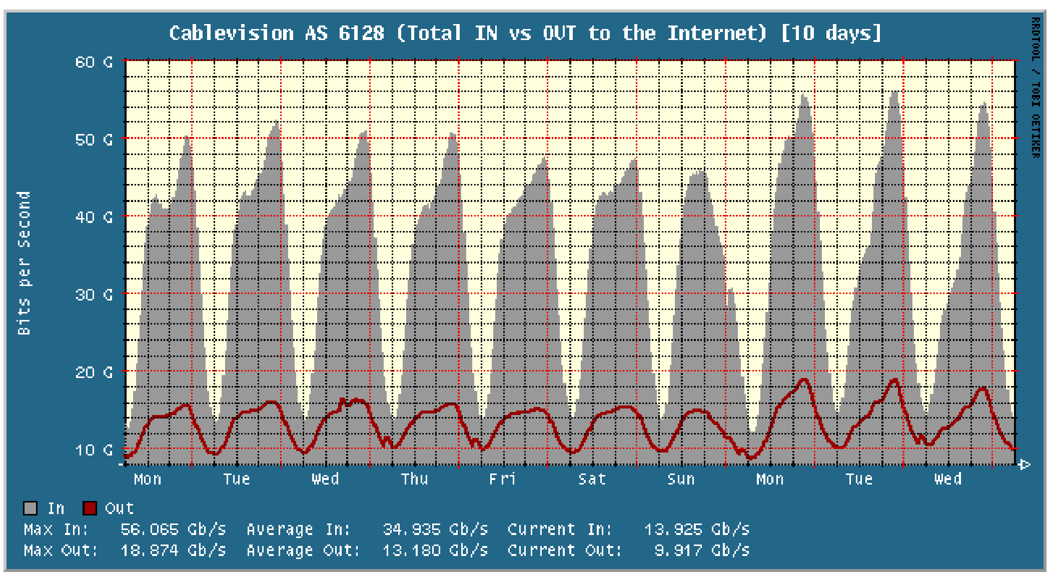

Figure 4: 1Q2007

Several observations can be drawn from these set of traffic trends:

1. Traffic rates are increasing rapidly, and non-linearly;

2. Across all graphs, the peaks and the lows occur at about the same time, so they most likely correspond to interactive use of the internet around evening time;

3. If the majority of traffic was peer-to-peer (P2P) then the lows would have been much more pronounced as people tend to leave their P2P clients on through the night. Either people shut down their P2P when not active, or some highly interactive applications (file downloading or real-time streaming) are the dominant behavior during peak time. Higher burstable contracts appear to have the effect of creating higher peak bursts, but not higher sustained use (as noted by stagnating dwell/gap growth).

Another important take away from this graph is that network capacity planning techniques will need to change. The rules previously used for lower bandwidth contracts cannot be extended to higher modem contract rates because of the increased burst capability (thus higher aggregate peak rate) of the traffic flows.

		SUBS	Delta -Subs	% Sub incr	Aggregate BPS	Delta - BPS	% BPS incr	(Kbps) BPS/Sub
2004	Jan	1,082,662	-	-	12,000,000,000	-	-	11.08
2005	August	1,569,475	486,813	45%	22,000,000,000	10,000,000,000	83.33%	14.02
2006	June	1,866,716	297,241	19%	44,000,000,000	22,000,000,000	100.00%	23.57
2007	January	2,039,259	172,543	9%	56,000,000,000	12,000,000,000	27.27%	27.46

Figure 5: Usage/Growth stats

TCP/IP PERFORMACE IN HIGH SPEED NETWORKS

DOCSIS 2.0 Transmission Control Protocol (TCP) downstream throughput is gated by the number of Acknowledgment (ACK) packets that can be sent on the upstream. This is because TCP tends to self-clock and the rate of the segments sent on the downstream depends on the rate of ACKs sent on the upstream. In other words, if a cable modem can send up to 200 ACKs per second, the downstream rate can not exceed “ACK per second X segment size” Bytes per second, or in a typical case 200*1460*8 = 2.33Mbps. DOCSIS Upstream Concatenation can help, but it does not solve the basic problem of having a limited “ACK rate”. See RFC3449 (TCP Performance Implications of Network Path Asymmetry) for more detail on TCP performance limits due to asymmetrical links.

DOCSIS 3.0 allows multiple outstanding requests and so the bottleneck imposed by ACK rates is greatly reduced. The TCP receive window size becomes the new bottleneck. The following text explains the TCP receive window issue at a high level:

TCP defines a “window” as the amount of bytes that can be sent unacknowledged to a network. They don't need to be acknowledged (yet) because they count as in-transit, meaning that due to the various delays in the network they either did not have a chance to be received by the end point, or the ACK did not have enough time to travel back. In fact, TCP attempts to keep this window at a maximum because that way it "fills the pipe" between the sender and receiver so that there are always bytes in-transmit.

The original TCP definition sets the maximum number of bytes that can be in-transit through the network to 64Kbyte because of the 16bit encoding of the window size in the TCP header. If we assume a very simplistic model where a network introduces a 100ms roundtrip delay (possible on the Internet) and that a single ACK clears a whole 64Kbyte window, then the maximal best case throughput of a TCP session can not exceed 64Kbyte * 100ms * 8 = 52Mbps, which is less then the capability of a DOCSIS 3.0 modem – and is a theoretical upper bound. In a real system it's likely to be even slower.

To make things worse, the default window size in Windows XP is set to 16K, which will limit the maximal throughput in the example about to 52/4 = 13Mbps (even below DOCSIS 2.0 capabilities).

This issue with TCP has been recognized a long time ago, and in 1992 the IETF published RFC1323 that defines TCP/IP extensions to allow larger windows sizes.

To demonstrate that this in not a cable specific problem, even on the web pages of FiOS (Fiber to the home service from Verizon), it is recommended to run an ActiveX component (Figure 6) that sets the proper entries in the windows registry in order to take advantage of the RFC1323 support in Windows.
Note that they are other ways to increase TCP throughput and any "TCP speedup" utility would try setting the same parameters that are set by Verizon.

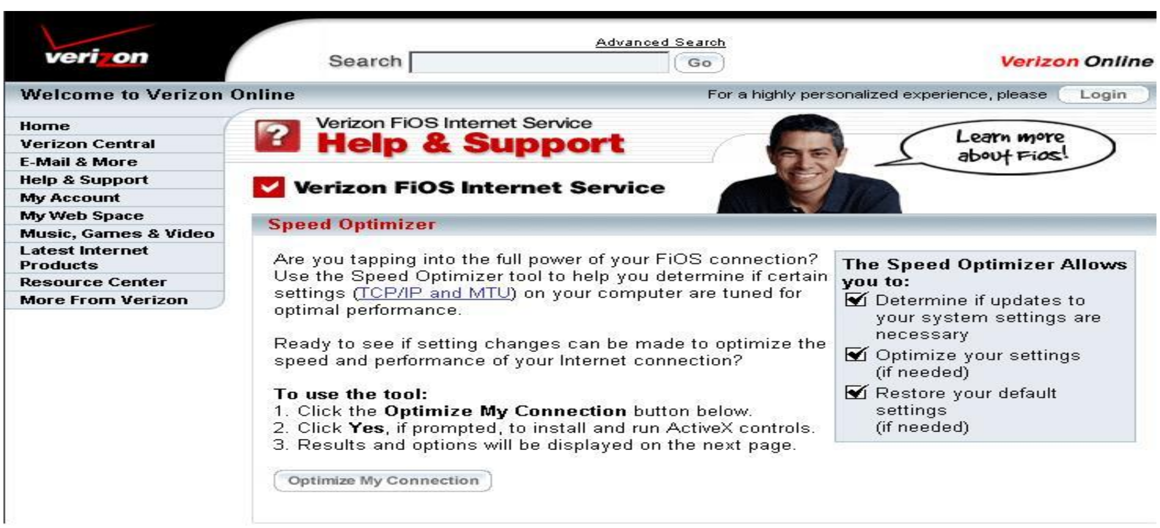

Figure 6: Verizon TCP speed up page

The following parameters are set:

1. TCP 1323 Extensions - This parameter enables enhancements to the TCP/IP protocol that provide improved performance over high speed connections;

2. TCP Receive Window - This parameter specifies the number of bytes a sender (the source you are downloading from) may transmit without receiving an acknowledgment. Modifying it determines the maximum size offered by the system (appears to be about 400K);

3. MTU (Maximum Transmission Units) - The MTU defines the largest single unit of data that can be transmitted over your connection. The FiOS network requires an MTU of 1492 bytes.

All utilities are available on the web under the general categories of "TCP optimizer", "TCP speed" etc, change these parameters. In a cable environment, and especially with DOCSIS 3.0, a user should update the operating system TCP defaults to work with TCP 1323 extensions, so that larger window sizes can be define.

IMPACT ON THE HOME NETWORK

The bottleneck link at the home used to be the internet connection. It was fair to assume that whatever the internet delivered, the home network can consume. With DOCSIS 3.0 this might no longer be the case. For example, if a user has a contract for a burst rate of 100Mbps, but the home wireless router is limited to 54Mbps, it is possible to congest the home network with data.

To allow for proper QoS handling in the event of a home network congestion DOCSIS 3.0 has the option of embedding priority bits in the 1-Byte, 3-Byte or 5-Byte DOCSIS headers. The priority bits are taken into consideration when queuing packets after the packets have been processed on the RF side and are on the way to the home network. Once they are queued to the home network, than in the event of congestion the lower priority packets would be dropped first.

The priority bits are arranged based on the Ethernet 801.D priority bits definition as depicted in Table 1:

		Number of CM output queues						
		2	**3**	**4**	**5**	**6**	**7**	**8**
Traffic Priority	**0 (Default)**	0	0	0	0	0	0	0
	1	0	0	0	0	0	0	1
	2	0	0	1	1	1	1	2
	3	0	0	1	1	2	2	3
	4	1	1	2	2	3	3	4
	5	1	1	2	3	4	4	5
	6	1	2	3	4	5	5	6
	7	1	2	3	4	5	6	7

Table 1 : DOCSIS 3.0 priority encoding

This table allows vendors to build devices with varying degrees of complexity in terms of how many output queues they have on their home network interfaces. The table defines which priority encoding goes to each queue.

Natually, if there are 8 queues then each gets a dedicated priority queue.

This feature can work in tandem with uPNP and Ethernet type flow control to preserve end-to-end QoS all the way to the home device.

HIGH SPEED AND POWER CONSUMPTION

Another side impact of higher traffic rates is the increased power consumption by networking equipment. As with other issues raised in this article, this is not a cable specific issue. It's not even a communication industry specific issue. It is a result of the current CMOS technologies, used in most of today's ASICs, reaching their physical limitations.

For example, modern oxides (the insulation layer in CMOS) are at 10-12 angstrom, that's 5 atoms thick! In addition to the fact its hard to go much lower then 5 atoms, there is already a scaling problem, since reducing the density by half would require placing a 2.5 atom thick oxide – clearly not an option…There is also a limit to the minimal voltage that a CMOS circuit can operate in. As a result, there a limit to how much power can be reduced – a limit that impacts CPUs, mobile devices, data centers and telecom.

In the past routers/switches had to look only at the packet headers (Ethernet/VLAN for switches, IP for routers) in order to forward a packet to its proper destination. As data rates went higher, the faster the forwarding rates had to go and the more power was needed to perform it. In addition to this basic forwarding the amount of "heavy lifting" per packet/byte is increasing. A sample of application that require per byte operation are:

1. Deep Packet Inspection: For reasons that relate to filtering certain network traffic type, and to protecting the network against Denial of Service Attacks, the network has to look beyond the L2/L3 portion of a packet in order to figure out what a packet is, not just where it goes to.
2. Video Streams manipulation: High touch video processing, such as transcoding, trans-rating and ad-insertion require extensive byte manipulation.
3. Encryption: While not new to cable networks, encryption is a byte-by-byte operation on a whole packet. The faster the data rates are, the faster the encryption chips have to run and the more power they need.

Hopefully, technology innovations will help reduce the power requirment as much as possible. In addition,new network architectures, such as M-CMTS, allows for distribution of function (L1/L2/L3) which is also a distribution of power. For example, if the M-CMTS packet shelf is located outside the hub, it reduces the power demands on the hub.

DUAL TOKEN BUCKET

A single token bucket definition (as the one used in DOCSIS 1.1/2.0) places restriction on the peak rate and burst size of a flow. The problem with a single token bucket definition is that for the duration of a burst the flow is not limited to the peak rate, instead, the burst rate can be as high as the total capacity of the link. As long as the burst size is minimal (in the range of a couple of Ethernet frames) this is not a significant issue, but for those customers who define large burst sizes the burst size could become an issue with assuring fairness across a large number of flows. It can further be exacerbated because of the data rates supported in DOCSIS 3.0.

The way to address extended bursts with the DOCSIS 3.0 toolkit is the newly defined “maximum sustained traffic rate” and “maximum downstream traffic rate”. These two parameters refer to how fast the traffic can flow during the traffic vs. how fast it can flow once the burst is exhausted. For example, a traffic contract can be:

- Maximum sustained traffic rate is 10mbps
- Maximum downstream traffic rate is 30mbps
- Burst size is 2Mbytes

For this contract, the user can send up to 2Mbytes of data at 30mbps, but if the user sends more then 2Mbytes the CMTS would reduce the rate to 10mbps.

TRAFFIC PATTERNS IN A HIGH THROUGHPUT NETWORK

With the bandwidth that DOCSIS 3.0 provides it is very likely that a new crop of applications will show up and put to use the added capacity. These new applications will have new traffic patterns, and therefore the network planning can not be derived from the existing traffic patterns. The first section of this paper already demonstrates how unexpected the aggregate traffic can be when individual contracts are increased. This section will discuss some of the applications that might use the bandwidth provided by DOCSIS 3.0 and their impact on the network. In one sentence, the one common theme is that video is the new “killer app”that will drive bandwidth demands.

Today, a majority of internet traffic is related to file sharing. In fact, one may argue that a good percentage of file sharing traffic is video content and therefore is a form of video on demand distribution. In that sense we are already experiencing a bandwidth explosion that is video driven. The next step to file sharing is to stream the video content and play it while it’s being downloaded, and this is already being implemented by joost (www.joost.com). The “joost” type of video streaming has a couple of fundamental impacts on network traffic patterns.

1. Since every home is a media source, as well as media consumer, the upstream bandwidth is driven higher.

2. Joost, just like a file sharing application, opens up multiple parallel TCP sessions. That means that it will try to squeeze as much bandwidth from the existing traffic contract without trying to be “fair”.

While joost and the multitude of joost-like applications that are sure to follow, represent “over the top” video (meaning video sources that are not originated by the cable MSO), a cable MSO IPTV delivery will have its own impact of bandwidth usage and traffic patterns. Although its not clear at the moment if IPTV will be a significant bandwidth drive in the cable MSO world, its worth exploring the way it is different then over the top delivery.

MSO generated IPTV streams are likely to be carried over RTP (as opposed to TCP). While TCP is a closed loop protocol, RTP is open loop, meaning that the RTP source will not slow down or try to be network friendly if the network is congested. In that case, one might ask why use RTP instead of TCP? The reason is that RTP has a much better story when it comes to delivering real-time content and that is one way to differentiate MSO content from over-the-top content. But because open-loop delivery can not recover as gracefully as TCP, the network has to be over provisioned by much more then with TCP flows. In fact, if IPTV catches on, there could be a major shift from a world where most traffic is TCP to a world where most traffic is RTP: instead of flows that regulate themselves (TCP) a network that has to be over-provisioned to accommodate non-flexible traffic flows (RTP). Also take into account the fact that a video flow is fairly long lived, and the end result is a network that future networks can't be oversubscribed by the amount used today. Note that this conclusion may be somewhat counter-intuitive since higher bandwidth, in principal, should lead to better statistical multiplexing and higher oversubscription ratios. If the internet traffic remined as it is today, this may have been the case, but it is less likely when considering future applications.

-Alon Bernstein, alonb@cisco.com
-Joe Godas, joe@cv.net