Comcast Usage Meter Accuracy

Peter Sevcik
May 2010
Update to the original December 2009 document

NetForecast Report
NFR5101

Comcast has launched a new usage meter that monitors and reports on how much traffic a subscriber consumes each month. NetForecast was engaged by Comcast to perform independent testing and analysis of the usage meter's accuracy. NetForecast had no role in the design or implementation of the Comcast usage meter. This report documents the results of our technical audit of the usage meter accuracy for Comcast subscribers served by the Cable Modem Termination System (CMTS) equipment supplied by Arris, Cisco and Motorola.

NetForecast performed an extensive battery of tests and studies of the meter system over a ten month period. NetForecast generated test traffic, performed its own independent traffic measurements and calculations, obtained meter data from Comcast for the same test traffic, and compared NetForecast's predicted results with the Comcast meter results. Comcast had no prior knowledge of the values NetForecast predicted.

Comcast's stated goal is that the usage meter correctly reflect traffic passing through a subscriber's cable modem within plus-or-minus 1.0% accuracy over the month. Our analysis validates that the accuracy of the Comcast meter for subscribers is within plus-or-minus 0.6% over the month, well within Comcast's stated goal.

The Comcast Usage Meter

Comcast's Acceptable Use Policy (http://www.comcast.net/terms/use/) includes a restriction regarding "excessive use." Comcast defines excessive use as any value above 250 Gigabytes (GB) per month per Comcast High-Speed Internet residential customer account. The new Comcast usage meter is designed to provide subscribers with information about how much of their usage allowance has been consumed over the month.

Where subscribers can find their meter report online

Meter reports are available online at the Customer Central portal. Subscribers can access the portal at http://customer.comcast.com After logging in, selecting the "Users and Settings" tab shows various account management tools along with box labeled "My devices" that displays the amount of the 250GB allowance that has been used within the current month. Clicking on "View details" brings up the view shown in Figure 1. Subscribers can view the usage for the current month as well as usage information for up to three rolling months history.

Selecting the highlighted Learn More link displays an extensive FAQ document with information about Comcast's Acceptable Use Policy.

A subscriber can access the Customer Central portal from any browser from any ISP's network, allowing subscribers to check their meter when away from home.

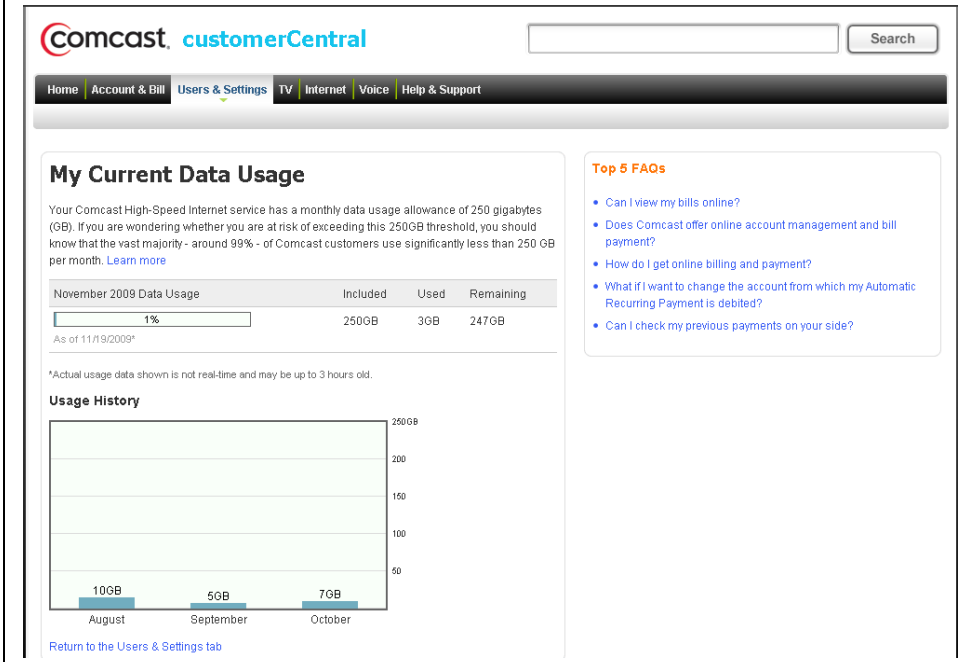

Figure 1 – Sample Subscriber View of Usage Meter

How the Meter Works

Comcast High-Speed Internet subscribers are connected to the Internet through a series of devices and technologies. The subscriber's traffic path begins at the Cable Modem (CM) and travels over a local coaxial or Hybrid Fiber-Coaxial (HFC) system that terminates at a Cable Modem Termination System (CMTS), which is typically located in a headend or hubsite. The traffic then continues through the Comcast network and onto the Internet, with the typical support and interconnection services supplied by any Internet Service Provider (ISP). The specification for how a cable modem communicates with the CMTS is defined in the Data Over Cable Service Interface Specification (DOCSIS), which is an international standard developed by CableLabs and a large group of participating companies.

Subscriber traffic is measured by the CMTS for each cable modem it serves. The CMTS keeps separate incrementing counters for traffic traveling upstream (i.e., from the subscriber to the Internet) and downstream (i.e., from the Internet to the subscriber). The status and value of the counters is periodically reported in an Internet Protocol Detail Record (IPDR). The IPDR specification is managed by the TeleManagement Forum (TM Forum), an international non-profit industry association. A DOCSIS Management Information Base (MIB) defines how traffic is counted in the IPDR.

Each CMTS periodically reports the IPDR for each CM (typically every 15 minutes depending on the CMTS manufacturer). IPDRs are gathered from the CMTSs by the Active Resource Manager (ARM) system, which has been supplied to Comcast by Active Broadband Networks. The ARM system processes the traffic data, which is then stored in the Comcast Enhanced Mediation Platform (CEMP). The subscriber

usage meter request invokes a web service to query the CEMP database by account number and MAC address to retrieve the usage for the customer meter display. The major steps in the process are shown in Figure 2.

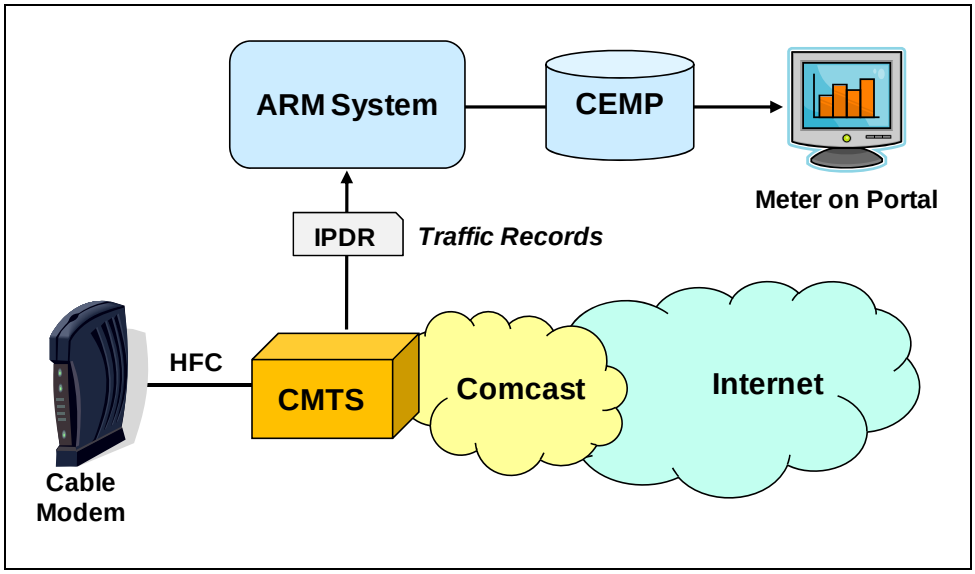

Figure 2 - How the Meter Data Is Processed

The subscriber sees results for devices that are authorized under the subscriber's account. This is typically the cable modem which is identified by its MAC address. Subscribers with more than one device see a separate meter for each device.

What is counted as subscriber traffic

The CMTS reports traffic in octets (8 bits). An octet is a telecommunications term for a byte. All traffic destined to or from the Internet is counted, including traffic to and from Comcast's various Internet sites (www.fancast.com, www.comcast.net, etc.). Traffic destined to or from Comcast's non-Internet services (e.g., Comcast Digital Voice, digital video, etc.) is not counted.

In addition to subscriber traffic, a very small amount of management traffic such as traffic from SNMP polls or cable modem health checks is also counted. After studying the background traffic associated with modem management, NetForecast concluded that it represents less than 1GB over an entire month. Although NetForecast accounts for this traffic in our validation analysis, we do not believe it represents enough volume to move the meter one unit. We will explain this conclusion in more detail later.

The traffic counted includes all of the octets that must be transported over the HFC network between the cable modem and the CMTS in each direction. The DOCSIS specification defines the mechanisms for transferring subscriber traffic across the HFC network. Both ends of the network are terminated by an Ethernet interface. In essence DOCSIS is a mechanism for transporting subscriber Ethernet frames as they arrive to the cable modem across the HFC network to the CMTS. Of course the same Ethernet transport mechanism works in reverse on the downstream side of the connection.

Since the fundamental traffic load from the subscriber is an Ethernet frame in either direction, all other protocols placed into an Ethernet frame are counted as subscriber traffic. This means that traffic generated by IP protocol overhead as well as traffic generated by all other protocols above IP must be transferred and therefore are counted.

What the meter shows

There are several process steps between when a subscriber packet moves through the cable modem and when the meter results appear on the Comcast.net portal. Each of these steps takes time. We have already described the time lag associated with CMTS traffic reporting—generally every 15 minutes. The ARM system aggregates the traffic and summarizes it by hour. The CEMP database receives the updates and is prepared to show the results on an hourly basis. This processing introduces a time lag that causes the meter to update between 1 to 3 hours after the traffic was sent on the network. Typically it updates within 2 hours after a traffic event.

The CEMP accumulates the traffic UP and DN in bytes over the month. It then converts the total to Gigabytes and *rounds down* the result to whole Gigabytes. Whole number rounding down means that a unit value is not shown until a full decimal value is accumulated, e.g., 9.9 is rounded down to 9, and 10.1 is rounded down 10, etc. Therefore, the result displayed in the portal is the cumulative whole GB sum of all traffic from the beginning of the month (UP+DN).

At the start of each month the meter is reset to zero but shows "<1GB." The background traffic described above consumes less than 1GB, so a subscriber with a powered up cable modem and router but no devices connected will see <1GB on the meter for the entire month.

The meter operates on universal time (UT or GMT). This means that the "new month" which will show a meter reset to zero appears in the evening of the last day of the month across the US. For example, Eastern Time is typically 5 hours behind UT or GMT.

How Much Is a Number?

The Comcast meter reports Gigabytes, which is a binary number not to be confused with the similar decimal number. There is an easy numbers trap that appears to make the two systems the same. A thousand is often referred to as the metric kilo, followed by a million that starts with the same "M" as mega. But in fact these are very different values. The following table illustrates the difference.

Counting traffic by billions of bytes will result in a -6.9% error relative to the meter which uses binary numbers. A negative error indicates that the value is low relative to the standard value. Counting in decimal generates higher values relative to the meter.

Binary		
KB	Kilobyte	1,024
MB	Megabyte	1,048,576
GB	Gigabyte	1,073,741,824
TB	Terabyte	1,099,511,627,776

Decimal	
Thousand (Kilo)	1,000
Million	1,000,000
Billion	1,000,000,000
Trillion	1,000,000,000,000

NetForecast Meter Accuracy Validation Methodology

NetForecast acquired regular Comcast High-Speed Internet accounts in various locations on the Comcast network. We instrumented the accounts with a test laptop PC running Windows XP and a Linux-enabled Linksys router on which we installed Tomato router firmware. In addition we used FTP accounts on various NetForecast servers on the Internet. Figure 3 describes the NetForecast instrumentation.

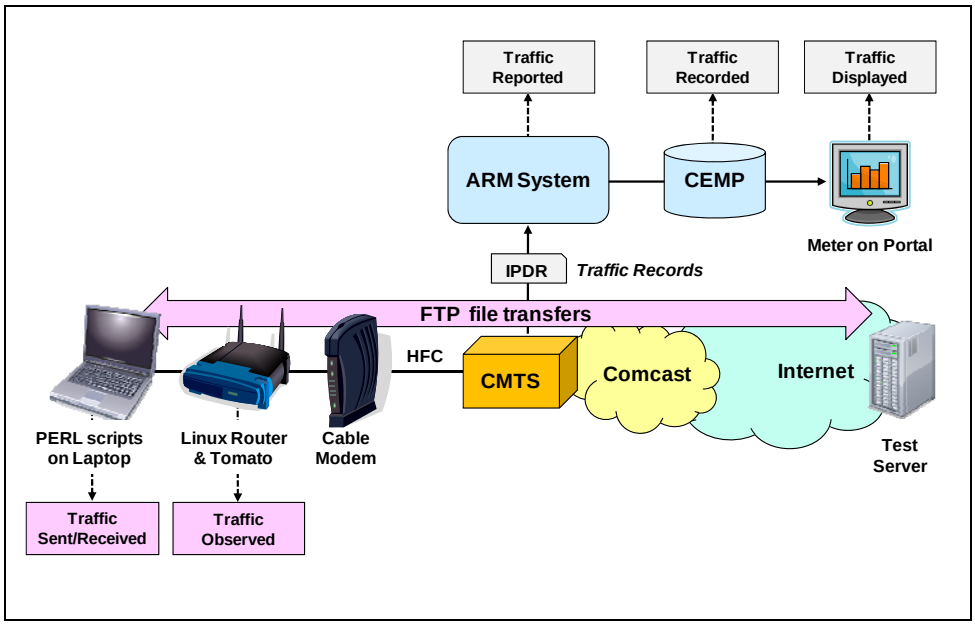

Figure 3 - The NetForecast Instrumentation

The basic NetForecast test involved a Perl script performing an FTP file transfer from one of our test servers to the laptop. The test consisted of repeatedly transferring files of several file sizes in complex patterns. These tests were also performed as uploads from the laptop to the server. The Perl script generated a log file documenting the transfer results and detailed timing information for each transfer.

Extreme care was taken to ensure that no traffic other than the test traffic was sent or received through the cable modem. First the laptop and router were the only devices connected to the cable modem. Second the laptop was cleansed of all applications that could generate traffic not needed for the tests. Furthermore Windows XP and other applications were configured not to ask for or receive any software updates. Finally, remote management of the laptop was carefully scheduled not to occur during testing.

For each test, we produced our own carefully documented records of traffic sent up and down as shown in the lower left of Figure 3.

- The first record was the basic file size as recorded on a typical PC. That value was adjusted to account for FTP/TCP/IP and Ethernet overhead both upstream and downstream. This became a calculated estimate of the actual traffic sent and received on the wire.

- The second record came from the Tomato firmware which also recorded the traffic it processed both upstream and downstream to the cable modem.

We continuously received meter data from three key places in the Comcast meter system during testing phase. These are depicted in the upper right of Figure 3.

- First we received IPDR records from the ARM collecting IPDR data from the CMTS involved in testing. The calculations were validated under controlled conditions at the NetForecast test lab. This was an early indicator of the CMTS traffic measurement accuracy.

- Second, we received hourly traffic records from the CEMP. This provided a detailed preview of the result presented on the meter portal.

- Finally we logged into each of the Comcast accounts to observe the meter as displayed on the meter portal (data shown to the subscriber). This completed the end-to-end view of meter accuracy.

Meter accuracy validation is the comparison of the three traffic reports from Comcast against the real traffic observations from the NetForecast tests. The NetForecast observations were seen by NetForecast only and were not shared with Comcast. The Comcast data was supplied by Comcast to NetForecast without Comcast's knowledge of which tests were performed and when they were performed.

NetForecast performed these tests from May through November 2009 under a variety of conditions. It is important to note that NetForecast only saw IPDR data and meter records for cable modems that were on NetForecast test accounts.

Results of NetForecast's Independent Meter Validation Study

Comcast's goal for the usage meter is to show traffic consumed by a suscriber's cable modem to within plus-or-minus 1.0% accuracy over the month.

Based on the results of our extensive testing and analsysis, NetForecast validates the accuracy of the Comcast meter for subscribers connected to CMTS equipement supplied by Arris, Cisco and Motorola to be accurate within plus-or-minus 0.6% over the month. Furthermore, the reporting system maintains that accuracy throughout all the elements of the meter system—up to and including the final view as seen on the customer portal. This statement applies when comparing the meter with the actual traffic sent to and from the cable modem, which includes some protocol overhead as explained above.

If there is a failure in the measurment, recording, processing, or storage of the traffic data, then the overwelming bias of such failure is to show less traffic than the susbcriber actaully sent. The system is designed to err towards showing less traffic than showing more traffic than was sent by the subscriber.

Reconciling Multiple Views of Traffic

To successfully reconcile what the meter shows relative to your own traffic consumption, it is important to understand several views of traffic. To illustrate this, let us assume the following hypothetical case. A subscriber downloads a 1GB file using FTP to his/her PC each hour for 20 hours and then stops and does nothing during the next 4 hours (the experiment takes 24 hours). There is no other PC or device connected to the Comcast service during the 24 hours.

Although there are many possible views of traffic in this case, we will focus on the three shown in Figure 4.

Computer: The computer downloaded 20GB and therefore shows 20GB on the computer's disk. If the subscriber kept a cumulative count of the traffic, he would note that the traffic as seen on his disk drive incremented by 1GB each hour from hour 1 through hour 20.

Network: Each 1GB file was transferred over FTP/TCP/IP protocols and then placed into Ethernet frames that were sent via the cable modem over the CMTS link. In addition, the server sent TCP acknowledgments (window updates) to the client. Finally, some client-server protocol set-up and handshaking occurred. If FTP/TCP/IP/Ethernet protocols were used in this case, they would add about 6.2% overhead to the traffic seen on the wire.

Meter: The meter system counts the traffic as seen on the wire (file size plus the 6.2% protocol overhead). However, the meter typically updates the customer portal 2 hours after the traffic actually traversed the wire, therefore, the meter showed a low cumulative count relative to the computer or network until it caught up.

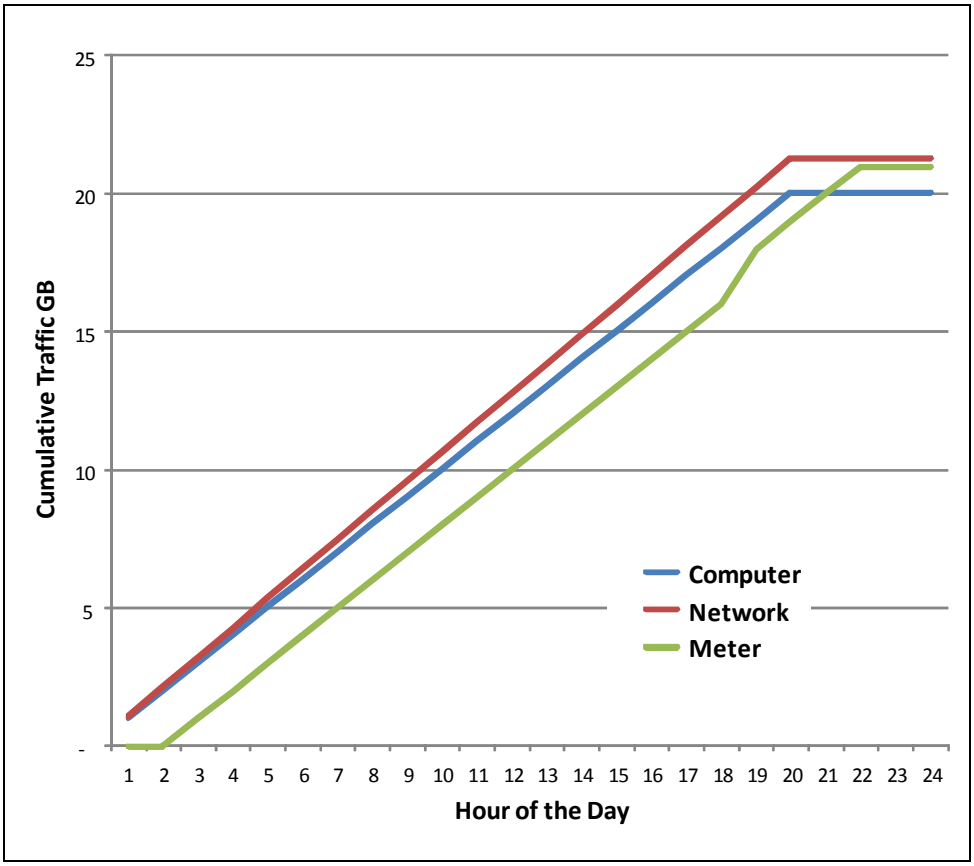

Figure 4 - Traffic Over the Day

Notice the increasing difference between the computer sum and the network sum. This is the 6.2% protocol overhead. Also notice the jogs in the meter line. This is the effect of the meter rounding down, which forces the meter to jump suddenly when a full GB count is reached.

Note that during the 2 hours after the test ends all of the lines flatten and no changes occur. Also notice that the meter value nearly catches up to the network value. The reason that it does not completely catch up is again due to rounding down.

Figure 5 shows another view of the experiment, tracking the percent difference between the results from the computer traffic view and the network traffic view relative to the meter view.

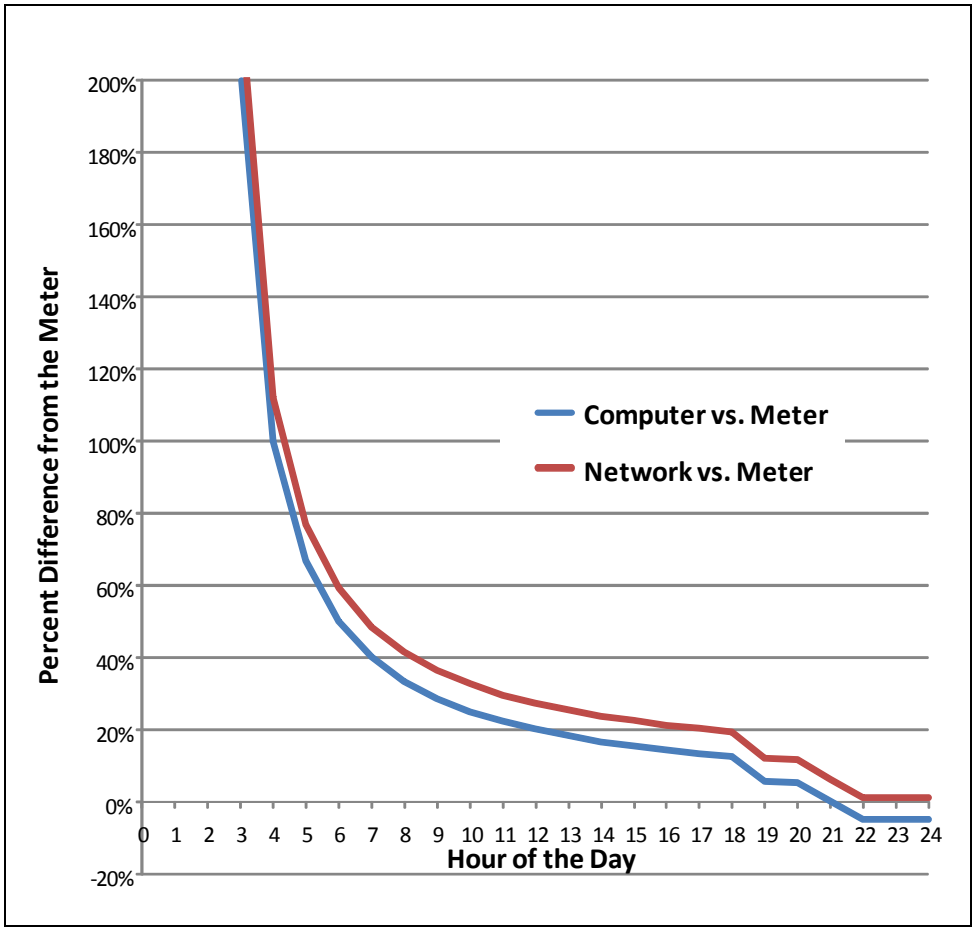

Figure 5 - Apparent Errors Converge Over Time

Although at first glance this divergence, especially in the early hours of measurement, may seem to be due to errors in the meter, it is not. Both the computer and the network show that much more traffic was sent relative to what the meter reports. This is due to the 2-hour lag in the meter count. In the beginning the meter has still not changed despite the fact that 2GB has moved into the computer.

The differences become progressively lower as each hour adds a smaller amount of traffic relative to the total that has accumulated during the test. After hour 20 when traffic stops, the network line almost matches the meter. It does not match exactly due to the rounding down effect. As time progresses the computer registers that it received less than the meter shows. This is the 6.2% overhead attributable to the protocol effect described earlier.

The plus-or-minus 0.5% accuracy NetForecast validated reflects how closely the network (red) line comes to zero (matching the meter). *NetForecast compared the error calculations based on the exact meter value without rounding down*.

As the month progresses the difference induced by the meter rounding becomes smaller. For example, if the network traffic (true traffic processed) is 10.9GB, the meter will read 10GB, an apparent underreporting of -8.3%. If later in the month the total usage were to climb to 100.9GB, the meter will read 100GB. At this level the difference is only -0.9%. At the critical 250GB per month traffic limit, the subscriber can send 250.99GB and the meter will still show 250GB, which is a -0.39% difference in the subscriber's favor.

Protocol overhead varies

Traffic due to protocol overhead will vary by protocol and many other conditions. The FTP example is one of the lowest overhead methods of moving content. Subscribers should be aware that there are many higher overhead protocols scenarios. For example:

- Web-based applications have much higher overhead and the overhead varies based on the web site design.

- Many advanced web pages are actually complex Java applications operating on top of the browser. The Java application must load and then execute to render the page, and it often makes additional requests for data during execution. The final output seen by the user on the screen may be a small fraction of what was transmitted to the browser.

- Security in the form of SSL or an enterprise VPN adds overhead.

- If content is delivered from a combination of source server and a CDN, it will generate some additional traffic. In general, the more servers involved in delivering the same content, the more traffic will be generated to redirect the browser, open new connections, etc.

Retransmission adds to the meter

Any reliable end-to-end protocol (like TCP) has a mechanism to retransmit packets lost in transit. For this reason packet loss will add to the traffic seen by the meter in two situations:

Downstream Traffic: If the loss occurs between the CMTS and the cable modem, then the packet will be sent again by the server and counted as an additional packet. Packet loss that occurs on the Internet or other parts of the Comcast network will not arrive to the CMTS-CM link and will not be seen the first time. Under these conditions, the second packet is the only one seen and counted.

Upstream Traffic: Packets that are sent from the subscriber to the server (e.g., backups) and are lost will again be retransmitted. If the loss occurs between the CMTS and the server (anywhere on the Internet), the first packet was already counted and the second packet will also be counted. However, if the loss occurs on the CMTS-CM link then the CMTS never saw the first packet and only the second one will be counted.

Sources of "Unexpected" Traffic on the Meter

A subscriber is likely to see traffic on the meter that he or she is not expecting. Many Internet users are unaware of the amount of traffic produced to support the wide range of things they do on the Internet. Here is a brief sampling of traffic sources that might surprise subscribers.

The first traffic source that may surprise subscribers is the number of devices and users in a subscriber's home. The meter counts all traffic sent and received by the cable modem, and most subscribers have a wireless (Wi-Fi) router that provides access to many devices. Traffic can be generated by more than just PCs. Any device that has access to the wireless router is a potential Internet traffic generator—including smart phones, game consoles, digital video recorders, printers, cameras, etc. Many non-PC devices "phone home" to a manufacturer or supporting service. These automated connections are transparent to the user as a convenience so the user is unaware of the traffic generated.

In rare cases a PC could be hijacked and generate traffic that has nothing to do with any user in the home. Also, the subscriber's Wi-Fi network may have been "hacked" by a neighbor who is adding traffic without the homeowner's knowledge.

The most likely source of unexpected traffic, however, is from software running on PCs throughout the home. The Windows operating system and most popular software have automated update programs. These updates often download and are installed automatically without the need for user intervention. The automation is generally designed for the convenience and protection of the consumer, but the traffic it generates may come as a surprise.

Each program update download may be modest in size, however, when you multiply a modest download by the number of programs calling for updates and the number of PCs in the house, the traffic attributable to updates can be substantial. Furthermore, in some cases the vendor default update settings are very aggressive, with some default settings checking each hour and downloading every possible option even though they are not all needed. For example, a software program may load its interface in a dozen languages even though all household members only know how to read English.

Another possible "surprise" upstream traffic source is online file backup, uploading to photo sharing sites, etc. Again, the backup software or service settings may be more aggressive than needed.

In addition, many news and information services preload content onto their subscriber's PC or smart phone over the Wi-Fi home network. The content often arrives overnight for convenient viewing in the morning. Of course the user does not read all of the content each day, but likes the speed with which content appears on the screen. The fresh content may also be sent to a smart phone or MP3 player to be viewed or listened to during the morning commute.

Assume each night's upload is only 1GB, which takes up a modest 1GB on the device's storage, and assume too that it never consumes more than 1GB because it overwrites the old content with fresh content each night. As modest as that may seem from a device storage point of view, that 1GB did consume Internet bandwidth each night adding up to 30GB over a month on the meter (plus protocol overhead).

Finally there may be unexpected traffic to non-PCs. A large volume of traffic may be going to digital video recorders such as TiVo. A user in the home may have rented a

movie from Amazon, Netflix. Blockbuster, etc. Renting the movie will be a known traffic-generating event, however, many services also preload the start of other movies as well as trailers to make them instantly available should they be called for. As in other situations described above, traffic is consumed for the consumer's convenience but without his or her knowledge.

Most of these traffic sources are low, but some can be unexpectedly high if they aggressively load content. Subscribers should check their software settings and align update size and frequency to their needs, bearing in mind the amount of traffic that generates.

Tracking down rogue traffic

If a subscriber cannot account for a high traffic volume on the meter and suspects some rogue consumption, then we recommend performing a controlled test. Plan for a solid period of time when the home can become "digitally silent" (overnight or on a weekend when traveling). At the start of the silent period log into the Comcast meter portal and note the consumption value. Then immediately turn off all devices that can access the Internet. Make sure, however, to keep the router and cable modem operating.

At the end of the digital silence turn on one PC and log back into the Comcast meter portal, or you can check from an Internet cafe or other means while you are away. If true digital silence was achieved, the meter should not have incremented by more than 1GB. If there is more than 1GB use over even several days, then there is certainly some other traffic consumer connected through the router.

If the digital silence experiment worked, then carefully add devices back to the home network while watching the meter. Note that the meter only increments once per hour, so it may take some time to find a rogue traffic source. On the other hand, the home may simply be a highly connected place that is leveraging many aspects of the Internet, and the traffic may be entirely due to legitimate use.

Conclusions

The Comcast usage meter is highly accurate to within plus-or-minus 0.6% over the month. Based on our test results, subscribers should be able to rely on the meter's accuracy to better understand how much Internet traffic they are consuming, and how to manage that traffic if they wish.

The meter will shine a new light on a previously unknown and misunderstood aspect of the digital age. NetForecast believes that this information will allow consumers to become better informed, and better informed consumers will help positively shape the Internet's future.

About the Author

Peter Sevcik is a Principal of NetForecast and is a leading authority on Internet traffic, performance and technology. Peter has contributed to the design of more than 100 networks, holds the patent on application response-time prediction, and pioneered many performance management techniques including Apdex. At BBN Peter played a pivotal role in the development of the Internet. He can be reached at peter@netforecast.com.

R3

NetForecast helps enterprises and vendors understand and improve the performance of networked applications.

Additional information is available at: www.netforecast.com