Third Accuracy Assessment of Comcast's Internet Usage Meter

By Peter Sevcik
May 2014

Comcast provides a data usage meter system to measure and report how much Internet data traffic a subscriber consumes and generates. Comcast has engaged NetForecast to independently and continuously audit the accuracy of its data usage meter system. Our first report was published in December 2009 [1] and a second report was published in May 2010 [2]. This third report details the cumulative accuracy results from 569,376 measurements at 55 Comcast service locations for the 3-month period of January through March 2014.

NetForecast performs two types of accuracy audit approaches for Comcast: passive traffic measurements in subscriber homes and active reference tests in which a NetForecast PC and server generate the only traffic on dedicated test lines in subscriber homes. In combination, these two approaches help quickly identify any potential meter accuracy issues, and supply data needed to troubleshoot and resolve issues.

For both approaches, NetForecast performs independent traffic measurements, obtains hourly usage meter records from Comcast for the locations, and compares the NetForecast data with Comcast's records. Results from 50 passive and 5 reference locations are aggregated in this report.

Comcast established an accuracy goal for its Internet data usage meter to correctly measure traffic passing through a subscriber's cable modem within plus or minus (+/-) 1% accuracy over a month. **Our analysis validates that Comcast is successfully meeting its accuracy goal.**

This report describes how the Comcast usage meter works, the NetForecast methodology used to validate accuracy, and the analysis used to determine our findings.

NetForecast Report NFR5116

The Comcast Usage Meter

Comcast's Internet usage meter provides subscribers with information about how much traffic has traversed their residential Internet connections.

Where subscribers can find their meter report online

Meter reports are available online at Comcast's customer portal, accessible at http://customer.comcast.com. After signing in and selecting "My Services," a new page shows various account management tools along with an "Equipment" box on the right. In the Equipment box, the data usage accumulated during the month to the current date is displayed as shown in Figure 1.

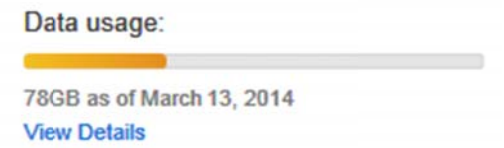


Figure 1 – Sample Thumbnail View of Comcast's Data Usage

Selecting "View Details" brings up a comprehensive page of information as shown in Figure 2. Subscribers can view usage for the current month as well as a rolling three-month history.

Selecting "For more information and frequently asked questions" brings up detailed information about the Comcast data usage plan. The page shown in Figure 2 also has a link to a Data Calculator which subscribers can use to estimate their data usage based on the number of Internet-connected devices and daily Internet activity.

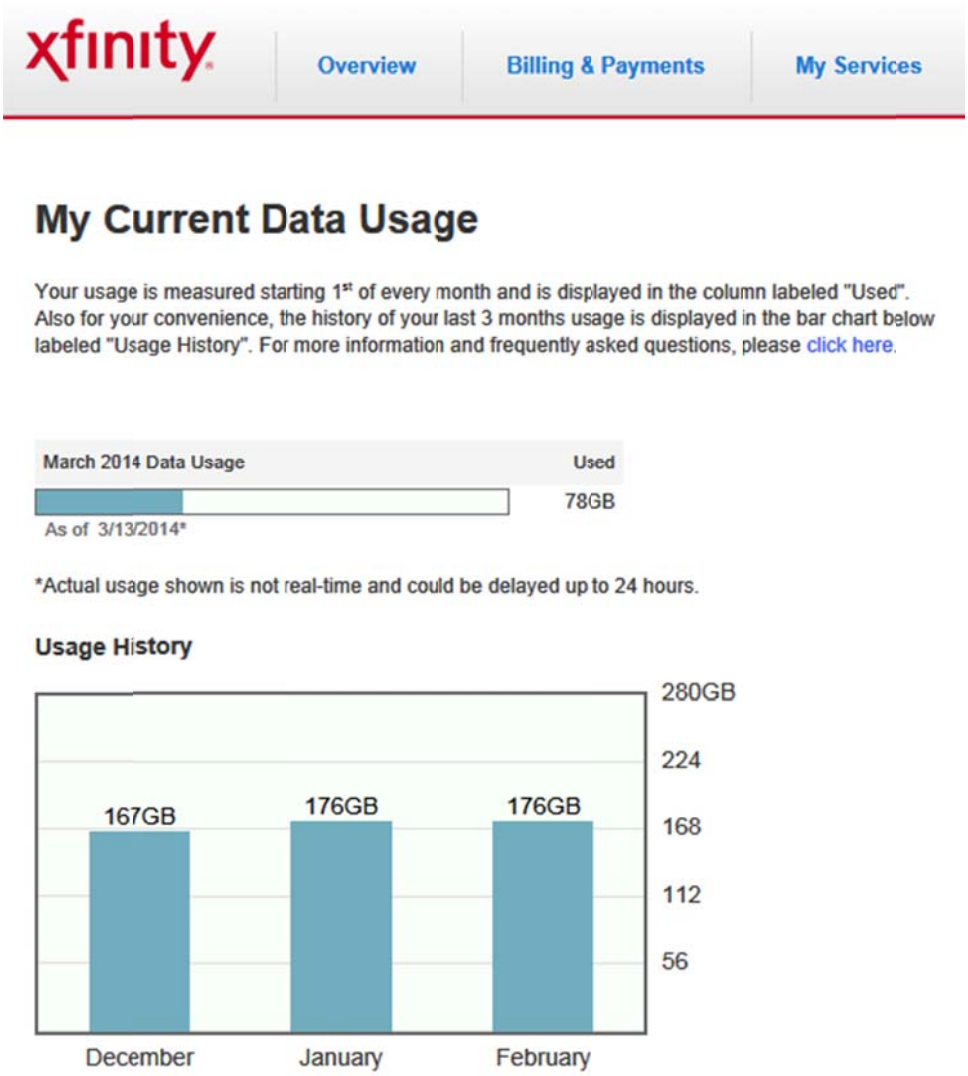


Figure 2 – Sample Subscriber View of Comcast's Data Usage Report

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

How the meter works

Comcast subscribers connect to the Comcast network through a cable modem and from there traffic travels over a local coaxial and hybrid fiber-coaxial (HFC) cable system to a Cable Modem Termination System (CMTS). The traffic then continues through Comcast's network and into the Internet.

The CMTS counts downstream and upstream traffic for each subscriber cable modem it serves. Downstream traffic flows from the Internet to the subscriber, and upstream traffic flows from the subscriber to the Internet. A CMTS periodically reports the downstream

and upstream counts in an Internet Protocol Detail Record (IPDR) as shown in Figure 3 (typically every 15 minutes, depending on the CMTS manufacturer).

As Figure 3 shows, Comcast's IPDR system collects and aggregates data from each CMTS and transfers it to the Comcast usage management platform, which associates it with the customer account and stores it in a database. When a subscriber accesses the web portal, a web service queries the usage management platform database to retrieve the usage data and display it as shown in Figures 1and 2.

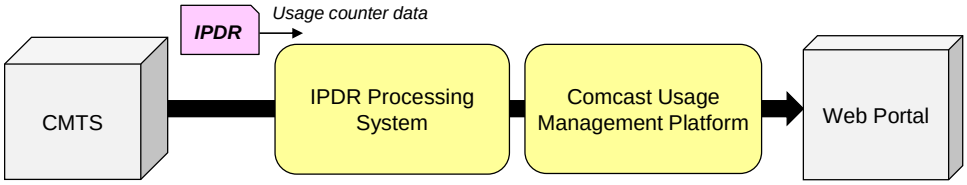


Figure 3 - How the Meter Data Is Processed

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 cable modem see a separate meter for each cable modem.

What the meter counts

All subscriber traffic destined to or from the Internet is counted, including traffic to and from Comcast's various Internet sites (www.comcast.com, www.xfinity.net, etc.). Traffic directly to or from the Internet with no Comcast involvement is counted. Traffic generated by non-Internet services delivered from servers within Comcast is typically not counted. In addition, management traffic such as SNMP polls or cable modem health checks is not counted.

The CMTS reports traffic in octets (8 bits). An octet is a telecommunications term for a byte. 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. DOCSIS is a mechanism for transporting subscriber Ethernet frames as they arrive at the cable modem across the HFC network to the CMTS. The same Ethernet transport mechanism works in reverse on the downstream direction. 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 the IP protocol and traffic generated by all other protocols above IP must be transferred, and thus are counted. Comcast does not, however, count any of the DOCSIS protocol.

> **NetForecast Validation Result:** NetForecast validated that the Comcast meter system counted all traffic as specified above.

What the meter shows

There are several steps between the time a subscriber's 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 IPDR processing system aggregates the traffic and summarizes it by hour. The usage management platform database receives the updates and prepares the results to be shown on an hourly basis. This processing introduces an overall time lag that causes the meter to update about 3 hours after the

subscriber traffic passed over the network. The Comcast goal is to have the meter update no later than 24 hours after the traffic event. These meter update time lags were not a subject of this evaluation, and so were not verified by NetForecast.

The usage management platform accumulates the up and down traffic bytes over the month, converts the total to gigabytes (GB), and rounds it down 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 to 10, etc. The result displayed in the portal is, therefore, the cumulative whole GB sum of all traffic from the beginning of the month (up + down). At the start of each month the meter resets to zero but displays "<1GB."

The meter operates on Coordinated Universal Time (UTC, also known as GMT). The "new month" meter reset occurs at midnight UTC on the last day of each month, so the change appears during the evening of the last day of the month across the US. For example, Eastern Time is typically 5 hours behind UTC, so the meter reset appears at about 7PM Eastern Time.

NetForecast Internet Usage Meter Accuracy Validation Methodology

All measurements were performed using the NetForecast UMap service delivery platform shown in Figure 4. The UMap system is enabled by measurement and reporting software embedded into customized, fully-featured, wireless home routers supplied and supported by NetForecast. Once installed, the routers register with the NetForecast UMap management and reporting system.

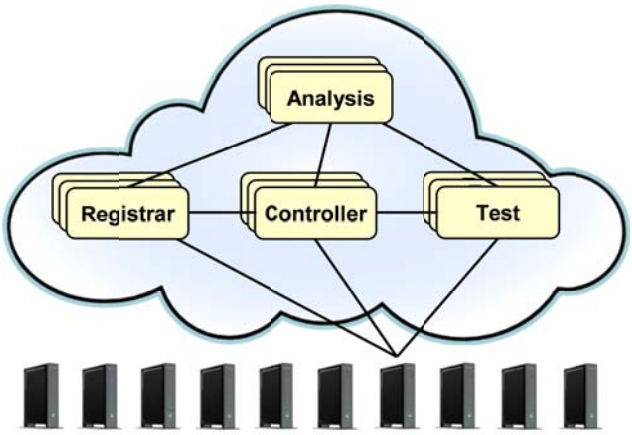


Figure 4 – NetForecast UMap Service Delivery Platform

UMap is a parallel data usage meter system that is deployed in subscribers' homes. As shown in Figure 5, the measurement software counts real user traffic traversing the Internet connection, generates synthetic traffic tests (also counted), and reports results to the UMap management system, which aggregates the data and generates reports.

The UMap platform has a high degree of security and resiliency, and strong safeguards to protect the privacy of participants.

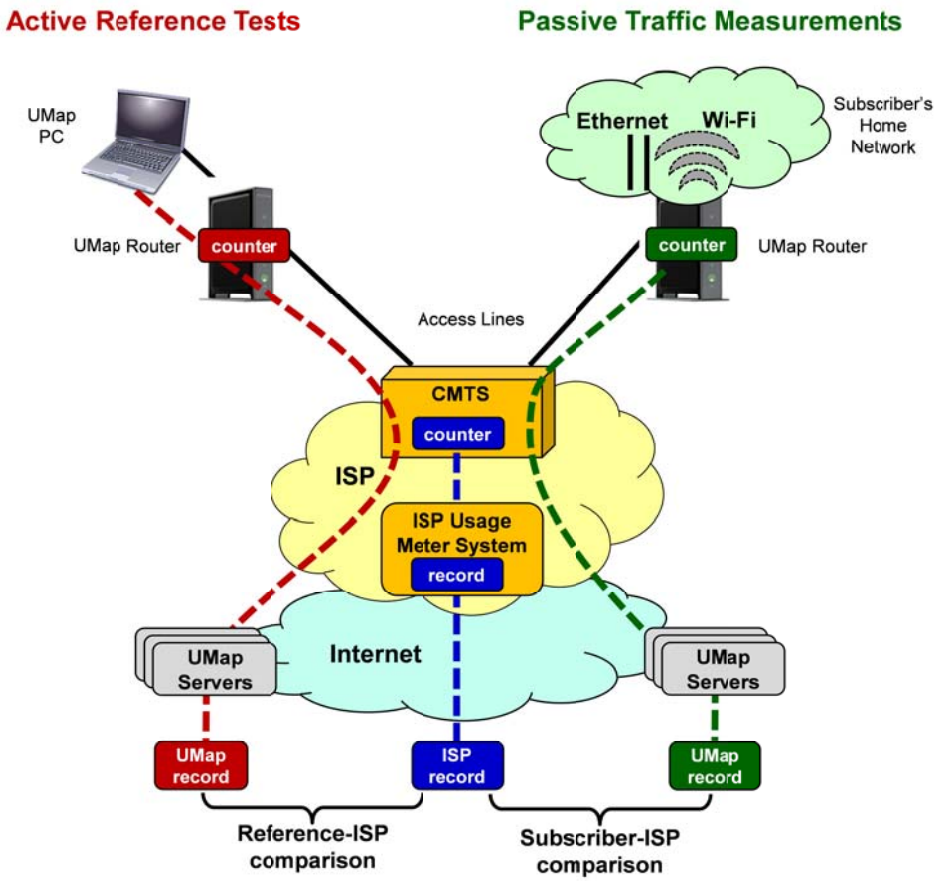


Figure 5 – The NetForecast Methodology

The UMap system operating on the Comcast network performs two types of measurements: active and passive. This report is based on results from 5 active and 50 passive measurement sites. The passive measurement sites were allocated to ensure that they are geographically distributed across Comcast's U.S. network, and on their full range of CMTS makes and models.

Active reference testing

This description applies to the red dashed line of data flow in Figure 5.

For the active reference test locations, NetForecast installs a test laptop PC running NetForecast software on each Internet connection. We use FTP accounts on NetForecast servers on the Internet. NetForecast obtains subscriber accounts and performs usage meter accuracy validation testing for all active reference test locations. Both downstream and upstream testing are performed under a variety of conditions.

The NetForecast testing involves scripts performing an FTP file transfer from a test server to the laptop. The tests consist of repeatedly transferring files of varying sizes in complex patterns. Tests are also performed as uploads from the laptop to the server. The script generates a log file documenting the transfer results and capturing detailed timing information for each transfer.

Extreme care is taken to ensure that only test traffic is sent or received through the cable modem. The laptop and router are the only devices connected to the cable modem. The laptop is cleansed of all applications that could generate traffic not needed for the tests—

and applications are configured to neither request nor receive any software updates. Finally, no remote management of the laptop is scheduled to occur during testing.

For each test, NetForecast produces carefully documented records of up and down traffic. These records include:

- Detailed network traffic information as seen at the PC, and
- Upstream and downstream traffic processed by the router firmware.

Passive traffic measurements

This description applies to the green dashed line of data flow in Figure 5.

For the passive locations, the UMap system continuously measures real-user traffic traversing the home Internet connections of many actual subscribers. Passive measurements take place under real-world conditions—i.e., it relies exclusively on the subscriber's home traffic. This enables measurements to be performed at many locations, thus providing comprehensive geographical coverage. Unlike active testing, which is not continuous, passive traffic measurement provides data for every hour during the month to give additional insight into meter accuracy.

Calculating a single error sample

UMap data from either the active or passive approach is adjusted to assure that UMap measurements count the same overhead as described above. NetForecast then aligns the hourly usage records from UMap and Usage Management Platform so that the same hours are compared. Daily sums are genenerated for each site. NetForecast applies the following formula to the UMap and the Comcast Usage Management Platform daily traffic measurement records.

$$\boldsymbol{Error} = \frac{(\boldsymbol{Comcast\ Record} - \boldsymbol{UMap\ Record})}{\boldsymbol{UMap\ Record}}$$

If the error result is positive, the meter is overreporting. If the error is zero, the meter is as accurate as it can be, and if the error result is negative, the meter is underreporting. Results are shown as a percentage. Each site-day error result is an error sample.

Analysis and Findings

Comcast has defined a goal for the accuracy of its Internet usage meter. NetForecast independently validated the meter relative to Comcast's stated goal. This section analyzes the error samples and presents NetForecast's findings regarding Comcast's actual performance relative to its accuracy performance goal.

As the usage meter system auditor, NetForecast has no stake in the design or implementation of the Comcast usage meter. The objective of our measurement system is to assess whether or not the Comcast meter system complies with the meter goal.

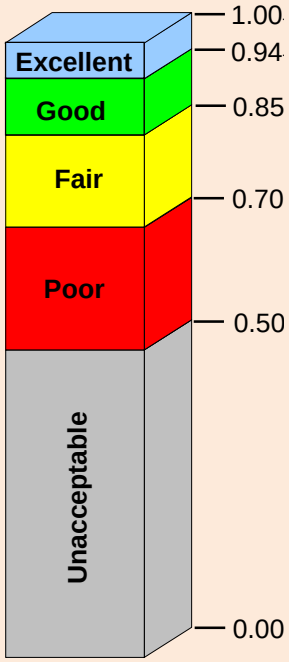
Comcast meter accuracy goal

Comcast established an accuracy goal for its Internet data usage meter to correctly measure traffic passing through a subscriber's cable modem within plus or minus (+/-) 1% accuracy over a month.

Apdex was first defined to report on many application performance measurements (response times in seconds). The methodology was formally adopted as an open standard by the Apdex Alliance in 2005. Apdex has been implemented by several vendors and is used by many enterprises. Today the Alliance has more than 1,000 members and 175 LinkedIn members. See apdex.org to learn more.

This is the first application of the Apdex methodology to usage meter error analysis. It is true to the fundamental principles of Apdex: Convert many values into a simple meaningful number that properly reflects the user's perspective of performance achieved relative to a performance target.

It is very difficult to achieve an Apdex score of 1.0 in real world situations. The Alliance adopted a scale shown below for the scores indicating good and poor performance scores.



To meet this goal, Comcast must assure that each of the two meter views presented to subscribers is accurate within +/- 1%. The two views shown in Figure 2 are defined as follows:

> Cumulative Daily Usage: This is the sum of usage by day from month start to month end. The usage bar in the top of Figure 2 is analogous to an automobile gas gauge. It shows how much "fuel" (Internet capacity) has been used. Subscribers can use this gauge to manage data usage over the month. In a typical month, 30 samples per site can be used for error evaluation.
>
> Month End Usage: This is the total usage for the complete month as shown at the bottom of Figure 2. This value is the same as the last value in the cumulative daily usage view. Subscribers see the total volume of consumption, which provides a historic view of usage. The month-end view provides only one sample per site for error evaluation.

NetForecast compares its independently measured traffic counts of synthetic and/or real user traffic with the counts generated by Comcast's meter system for the same traffic, to determine whether the comparisons fall within Comcast's target accuracy range.

Given differences in the nature and sample size of the two Comcast views, NetForecast performs different error analyses for each view as described below. NetForecast's analysis objective is to evaluate the accuracy of the two views from the perspective of *what matters to the subscriber*.

Cumulative Daily Error Analysis

Background: Analyzing and reporting on the error of a system is complex. One could take a simple approach of averaging the error of all samples. If the system has a significant bias, the simple average would show that bias (e.g., most samples are 10% low). However, if a system is fundamentally accurate, the mean (average) or median will be essentially zero. The median of the Comcast cumulative daily error samples is -0.1%, which is well inside the Comcast specification of +/-1%. But that result tells us that half of the samples are higher and half are lower. Many samples may be far from the median; therefore, usage meter accuracy assessment should not use averages.

Although under most circumstances subscribers receive accurate meter information, it is not uncommon for a meter system to occasionally provide inaccurate meter information affecting only a few subscribers. These infrequent events are typically called the "long tail of a statistical distribution." This may seem unimportant, but if the long tail encompasses a large sample size, the meter cannot be considered accurate.

Assessing Accuracy from the User's Perspective with Apdex: To provide clear insight into accuracy, NetForecast applies the Accuracy Performance Index (Apdex)—an open standard that defines a method for reporting the meaning of many measurement samples from the user's perspective. Apdex provides a uniform way to analyze and report on the degree to which measured accuracy meets a specific goal. Error samples are placed into the following categories each month.

Compliant: Number of samples within the +/-1% meter specification. These samples clearly meet the goal.

Marginal: Number of samples between -1% and -50% (underreporting). Underreporting is outside the specification; however, some modest underreporting can be tolerated since the subscriber is not harmed by some traffic not being counted.

Incorrect: Number of samples below -50% (excessive underreporting). Excessive underreporting (errors more than -50%) are not tolerated by the user. Excessive underreporting makes the meter useless (think of it in terms

of a gas gauge underreporting by 50%) and undermines the subscriber's trust in the meter.

Incorrect: Number of samples above +1% (any overreporting). Any overreporting above the 1% specification is a serious problem that users will not tolerate.

The Apdex methodology converts many measurements into one number on a uniform scale of 0 to 1 (0 = completely inaccurate; 1 = perfect accuracy). The resulting Apdex score is a numerical measure of accuracy performance. The Apdex formula is the number of compliant results, plus half of the number of marginal results, plus none of the incorrect results, divided by the total number of samples:

$$\text{Apdex [T]} = \frac{\text{Compliant} + \frac{\text{Marginal}}{2}}{\text{Total samples}} = \text{Score (0 to 1) [T]}$$

Incorrect samples have zero value

Target T or specification is always shown with the score

The value "T" is the target accuracy range of (+/-) 1% (the Comcast goal) and is always shown with an Apdex score.

Figure 6 shows the distribution of January through March error samples that illustrate the categories described above to apply the Apdex methodology to meter accuracy.

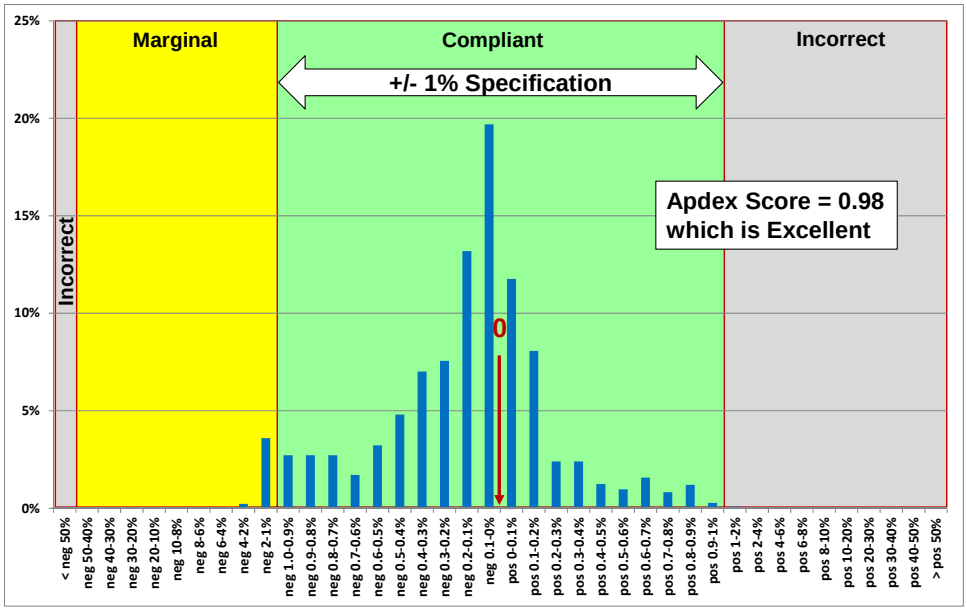


Figure 6 – Comcast's Cumulative Daily Apdex Results

NetForecast's Cumulative Daily View Accuracy Finding: The Apdex score for the cumulative daily measurements for the 3-month period is 0.98 [+/-1%], which is in the "Excellent" range.

Month-End Error Analysis

Background: There are only 55 month-end values per month (one per site). As with cumulative daily results, averaging is not informative. It is important to examine the two edges of accuracy—how far was negative error from zero and how far was positive error from zero?

Simply stating the absolute greatest deviation from zero can lead to extreme results that are misleading outliers. For example, if all but one of the samples are within +/-1% and the one outlier is at +5%, then the maximum of +5% does not accurately represent the general edge of positive error. Trying to apply standard deviation analysis is difficult because there are too few samples and their distribution is not normal.

Error Range Percentiles: The month-end error samples are ranked from most positive to most negative as shown in Figure 7. The top 5% and the bottom 5% of the samples by rank are identified. These two samples define the edges of the majority of samples. The error values of these samples are the worst underreporting and overreporting of 90% of the dataset. This is the error range within which 90% of the month-end site errors fall.

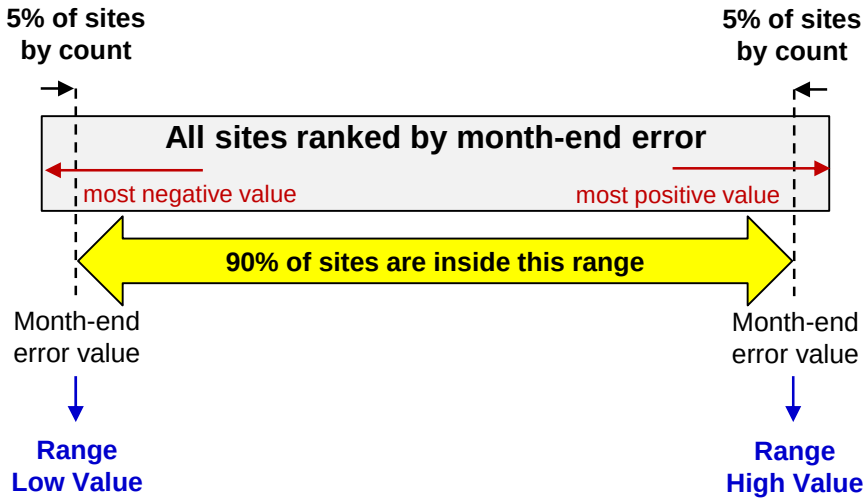


Figure 7 – Calculating the Error Range

NetForecast's Month-End View Accuracy Finding: The month-end error range for the 3-month period is -0.75% to +0.36%, which is well within the goal of +/-1% with a bias towards underreporting.

Useful Information if You Want to Do Your Own Testing

If you wish to perform your own Internet Usage meter validation testing, it is important to be aware of factors that may cause your measurements to vary from what the meter states. One such factor is where you measure. If you measure from the network, you will see protocol overhead that you will not see from a computer. Another factor affecting measurements is packet loss, and another is the presence of "unexpected" traffic.

Avoiding binary vs. decimal confusion

The Comcast meter reports in gigabyte increments. One gigabyte is a binary number not to be confused with one billion bytes. The following table illustrates the danger of applying decimal notation to byte counts. This is important to understand because counting traffic by billions of bytes will result in an apparent (but not actual) 6.9% error in 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

Decimal and binary values are often mixed. For example, 10MB means 10 times 1,048,576 bytes. A display showing 1,000MB is actually stating 1,048,576,000 bytes, which is still not a GB as shown in the table above. It requires 1,024MB to reach the value of a GB.

Where you measure matters

You can gather your own usage information either from a computer or from the network on your premises. A computer can track what is downloaded to/uploaded from it, but it does not report network protocol overhead because such data is hidden within the PC operating system (you need special instrumentation software to see all the protocol traffic).

For example, if one looks at the size of a file on a PC, that value does not include any protocol overhead which may lead the user to erroneously conclude that the Comcast meter is overreporting.

If you measure traffic in the network, you will see the payload traffic plus overhead from protocols like TCP/IP and Ethernet, which generally add up to about 6% to 9% overhead to the payload traffic for large packets and a larger percentage for small packet traffic like VoIP. The meter system counts the traffic as seen on the wire, which includes the payload plus protocol overhead, so it should closely match the network view.

Packet loss affects the meter

Any reliable end-to-end protocol (like TCP) has a mechanism to retransmit packets lost in transit. Packet loss will add to the traffic seen by the meter in some situations and subtract in other situations.

"Unexpected" traffic

If you look closely at your Internet usage, chances are you will see unexpected bytes register on your meter. Most Internet users don't know how much traffic their household produces. Here is a sampling of traffic sources that may surprise you.

Peer-to-peer (P2P) file sharing applications like BitTorrent, often used to exchange music, movie, and other types of large files, are a common source of unexpected traffic. Typically P2P agents operate automatically in the background on your network exchanging large amounts of data over your Internet connection. Users are often unaware of the total traffic consumed by P2P over a month.

One likely surprise is the number of traffic-generating devices and users in your home. Most subscribers have a wireless (Wi-Fi) router that provides access not just to PCs, but to everything from smart phones, tablets, digital video recorders and printers, to game consoles and cameras. Many of these non-PC devices "phone home" to a manufacturer or support service, and, for convenience, these automated connections are invisible to the user, so you may be 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 your home. Also, a neighbor or "wardriver" may use your connection without your permission or knowledge.

The most likely source of unexpected traffic, however, is PC software. Most popular software has automated update features which download and install updates. This transparent automation is for your convenience and protection, but the traffic it generates may come as a surprise.

Although each program update download may be small, when you multiply a modest download by the number of programs calling for updates and the number of PCs in the house, such traffic can be substantial. Furthermore, in some cases vendor default settings are aggressive, checking each hour and downloading every possible option, even if they aren't needed. For example, a software program may automatically load its interface in a dozen languages for a monolingual household.

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

In addition, many news and information services preload content onto a subscriber's PC or tablet. The content often arrives overnight for convenient early morning viewing. Of course, users don't read all the content every day, but they probably do enjoy the speed with which content appears on the screen. Fresh content may also be sent overnight 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 that it never consumes more than 1GB because it overwrites the old content with fresh content each night. As modest as this may seem from a device storage point of view, that 1GB did consume 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. For example, traffic may be going to digital video recorders such as TiVo or streaming boxes like Roku. A user in the home may have rented a movie using a subscription to a steaming service—and you may get more bits than you pay for because 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, the traffic is generated for your convenience but without your knowledge.

Most of these traffic sources are low, but some can be unexpectedly high if they aggressively load content. You should check your software settings and align update size

and frequency to your needs, bearing in mind the amount of traffic generated. Additional information about hidden traffic is available at reference [3].

Tracking down rogue traffic

If you cannot account for a high traffic volume on the meter and suspect some rogue consumption, we recommend performing a controlled test. Plan for a solid period of time when the home can become "digitally silent" (e.g., overnight or on a weekend when traveling). At the start of the silent period, record the data usage displayed on the Comcast meter, then 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 computer and log back into the Comcast usage meter portal, or check from another network. If true digital silence was achieved, the meter should not have incremented at all during any day. Because Comcast rounds down, values may increment to the next GB if the previous value was close to the rounding threshold before the silent period began. Subsequently, however, it should remain at that value for the rest of the digital silence period. If there is significant traffic shown, then there is certainly some other traffic consumer connected through the router.

Conclusions

Comcast established an accuracy goal for its Internet data usage meter to correctly measure traffic passing through a subscriber's cable modem within +/-1% over a month. Based on 569,376 measurements from January through March 2014, NetForecast validates that the Comcast Internet data usage meter cumulative daily view achieves an Apdex accuracy score of 0.98, which is a score of "Excellent." The month-end view is accurate within a range of -0.75% to +0.36% which is well within the +/-1% goal.

Both outcomes validate Comcast's +/-1% stated goal. Based on our measurement results, subscribers should be able to rely on the meter's accuracy. Comcast has implemented continuous accuracy testing aimed at measuring and maintaining meter accuracy.

References

1. Sevcik, *Comcast Usage Meter Accuracy*
   NetForecast Report 5101, December 2009.
2. Sevcik, *Comcast Usage Meter Accuracy (Updated)*
   NetForecast Report 5101, May 2010.
3. Sevcik, *Empowering Internet Users to Manage Broadband Consumption*,
   NetForecast Report 5109, presented at The Future of Internet Economics,
   Technology Policy Institute, June 15, 2012.

About the Author

Peter Sevcik is President of NetForecast and is a leading authority on Internet traffic and performance. Peter has contributed to the design of more than 100 networks, including the Internet, and is the co-inventor of three patents on application response-time prediction and congestion management. He works extensively with the SamKnows system in support of the FCC Measuring Broadband America project, analyzing operational integrity and performing deep data analysis. He can be reached at peter@netforecast.com.