Broadband Benchmarking: On the FCC Issues for Inquiry

Paper #0, 7 Pages

Abstract

The Federal Communications Commission (FCC) has recently taken major decisions in an attempt to regulate broadband in the US, creating disputes between itself and ISP providers. One such policy was the radical increase of broadband speed benchmark, that motivated both parties to investigate these regulations from different perspectives. This led to the Commission publically requesting comments on issues pertaining to "advanced telecommunication capabilities" to design future benchmarks based on metrics other than broadband speed.

In this work, we provide a much required input to the FCC's open questions, supported by our analysis of usage patterns, rather than just the FCCs speed benchmarks. We analyse users with high tier access links and motivate the need of multiple benchmarks based on peak usage of different types of users. We also show that beyond a certain speed, user behavior is not significantly impacted by further increases in broadband capacity. Therefore, motivating the need to define benchmarks based on metrics other than broadband speed to truly offer "advanced" capabilities.

1 Introduction

The large impact of broadband Internet in our daily lives, and its rapid increase in providing services, requires regulation of all involved parties, such as all transit and content providers. The Federal Communications Commission (FCC) holds the important responsibility of overlooking public and private sector initiatives to ensure advanced broadband deployment throughout the US. As part of this responsibility (since the amendment to the Telecommunications Act of 1996 (section 706 [4]) [1]), the FCC has been reporting results of its annual inquiry regarding the availability and timely deployment of "advanced telecommunications capability" to all Americans.

If the FCC determines that *"advanced telecommunications capability"* is not being deployed to all Americans in a reasonable and timely fashion, the Commission is required to "take immediate action to accelerate deployment of such capability by removing barriers to infrastructure investment and by promoting competition in the telecommunications market." [1]. The FCC accomplishes this role by following three major steps: (a) *defining a benchmark*[2] to evaluate "advanced telecommunications capability", (b) *auditing capabilities* of services offered by Internet Service Providers (ISPs), and (c) *spending the federal budget* to increase ISP capabilities, and accomplish their defined standards.

As independent researchers in the measurement community, it is our responsibility to ensure that the FCC defines broadband benchmarks and standards in a sensible manner. Efforts to define such benchmarks have always been based on aggregated data analysis by the FCC . The source of data are mixed across tiers and locations, and collected using different methodologies; particularly vantage points in the data were treated similarly without studying their actual usage requirements [3]. Based on such data, the FCC made a decision to aggressively increase the broadband benchmark threshold from 4 Mbps/1 Mbps to 25 Mbps/3 Mbps [1]. This prompted a sharp response from ISPs offering technologies that suddenly did not qualify as broadband (such as DSL).

We support the move to a higher broadband standard based on the increasing share of high quality entertainment traffic in home broadband consumption [5, 6], as well as the possibility of offering advanced capability (such as telemedicine). But, we strongly motivate the Commission to revisit their evaluation criteria in reaching more meaningful standards. We also realize that the FCC requests comments on the issues of broadband benchmarking [section 6] in future. To this aim, we use a dataset collected purposefully to validate and improve the FCC policy on broadband benchmark.

Our study is based on Comcasts analytic data collected from home gateways for their customers in Salt Lake City, Utah. This data consists of byte transfers collected continuously every 15 minutes from two types of users; control: users that pay and use a high capacity access link (105 Mbps), and test: users that pay for 105 Mbps but are actually offered a much higher capacity access link (250Mbps) without their knowledge. This decision was specifically anticipated to enable researchers to investigate the question: does increase in bandwidth change usage behavior?

Our contributions in this work are as follows:

- We reaffirm the importance of measuring performance during the prime-time peak hours, and urge the FCC to revisit and standardize the measurement and interpretation of **prime-time ratio**.
- We shed light on the user's and ISP's **differing perspectives of *capacity utilization***, and show that indeed, the user's behavior does change when offered a higher capac-

[1] This issue is discussed in point 1, and point 12 of the Eleventh Broadband Progress Report, FCC No. 15-10A1, respectively [3]

[2] This refers to a threshold, beyond which the access link will be considered broadband capable.

ity link even though the overall utilization stops increasing after a certain upper limit.

- We explore a user taxonomy based on **varying usage behavior within the same tier**, and suggest that the FCC adopt multiple benchmarks to better characterize broadband availability, deployment, and adoption in the US.

In this work, after offering details about our data and how it is sanitized in [section 2], we proceed by describing our methodology and evaluation criteria to study changes in broadband usage [section 3]. Next, we analyze spacio-temporal usage patterns as well as peak usage (during prime-time). Our analysis indicates that for high capacity access links, increasing capacity does not result in higher utilization, and motivates the need of multiple benchmarks based on varying usage within the same tier [section 4]. In the last section, we comment on the issues raised by FCC to investigate a better benchmarking solution for the US [section 5].

2 Data Source

Our dataset consists of network usage byte counters reported by Comcast gateways every 15 minutes from October 1, 2014 to December 29, 2014. There are two sets of broadband tiers that were used to collect this data: control set, consisting of households with a 105 Mbps access link, and the test set, consisting of households that were paying for a 105 Mbps access link, yet were receiving 250 Mbps instead. Users in the test set were selected randomly and were not told that their access bandwidth has been increased. There were more than 15000 gateway devices in the control set, with varying usage over the three months, and about 2200 gateway devices in the test set. TODO: confirm - these were reported by Comcast gateways right?

2.1 Data Description

The raw data sets provided by Comcast consisted of the test set, and 8 separate control sets consisting of more that 15k unique households, over different date ranges within the three months. Each dataset contains the following relevant fields: Device ID, sample period time, service class, service direction, IP address, and the bytes transferred in the 15 minute sample slot, as described in table 1.

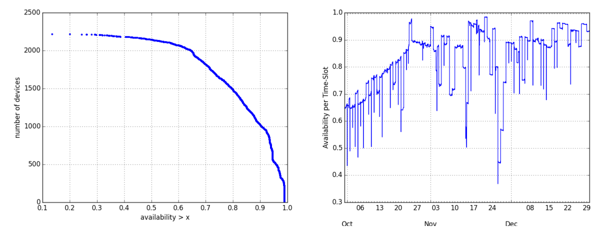
Field	Description
Device_number	Arbitrarily assigned CM device identifier
end_time	Fifteen minute sample period end time
cmts_inet	Cmts identifier (derived from ip address)
service_direction	1-downstream, 2-upstream
octets_passed	Byte count

Table 1: *Field Descriptions for Comcast Dataset by Comcast*

2.2 Data Sanitization

Our initial analysis of data transferred per time slot showed that certain gateway devices were responsive only for brief periods. We also noticed that certain time slots had a very low response rate throughout the dataset. TODO: Why? asked comcast - waiting for response .

We evaluate the fraction of responsiveness of a gateway throughout the dataset, as well as the fraction of responsiveness per time slot, and call this the **availability**. Figure 1 shows how the number of devices decreases for a higher availability requirement. Based on the common trend of this plot throughout the test and control datasets, we decided to only choose gateway devices with an availability of at least 0.8.



(a) *Availability by device* **(b)** *Availability by date*

Figure 1: *Availability, based on gateway device responsiveness. (Make common eps plot of availability – 8 control sets (half filtered) + test sets vs availability.)*

Finally, we sliced the sanitized test set based on the date range of each individual control set for comparison. We compared each of these tests individually to ensure that there are no outliers. We refer to the test and control sets in this case simply as datasets $set_1 - set_8$, where *set* is test or control . We also sliced and combined the sanitized data to give us control and test data for each month, referred to as set_{oct}, set_{nov}, set_{dec}. Finally, we combine all control sets to form a large concatenated dataset over the same date range as the complete test dataset, and we refer to this simply as set_{full}.

In the following analysis, we only present results for set_{full}, unless the behavior of an individual dataset varies significantly from the overall behavior and requires mention.

3 Methodology

In this section we describe how the dataset collected is both granular as well as unbiased, and enables us to study usage behavior in a controlled setting. We specify the evaluation criteria we use to compare usage patterns of the test and control sets.

In the following, we define capacity as the access link capacity offered by the ISP to the user. In the case of the Comcast dataset, capacity is 105 Mbps for the control set and 250 Mbps for the test set. Similarly, we define usage as the actual traffic (bytes uploaded or downloaded) by the household being studied. We use usage across 15 min time slots to derive the average data rate and compare it to the capacity in our analysis.

3.1 Relevance of the Data

Study Byte Counters: The purpose of this work is to study the usage characteristics, irrespective of the application responsible for such usage. Limiting ourselves to just byte counters makes our analysis easily extendible to any ISP, and the FCC, interested in doing a similar study at a larger scale, without the risk of leaking PII. A study of applications has already been performed extensively by Sandvine [], as well as other researchers.

Granularity of 15 minutes: Broadband usage evaluated by commercial groups [], or governmental survey bodies, usually employed by the FCC, tends to focus on aggregated usage statistics over months, long term trends, and applications. In our work we specifically focus on data transferred in 15 minutes, to avoid short term bursts that max out the capacity, but account for long term heavy flows (such as real time entertainment and voip calls) that will continuously max out the access link. This gives us a granularity fine grained enough to study major changes in usage characteristics (such as peak trends) while ignoring short term bursts of traffic (such as browsing)

Note that byte counter readings collected every 15 minutes from multiple households were synchronized for consistency in measurements.

High Tier Measurements: We limit ourselves to analyzing usage patterns in the high capacity access link tier only. The test dataset was collected by increasing the capacity from 105 Mbps to to 250 Mbps for 2200 randomly selected users, without their knowledge. This served a two-fold purpose in avoiding biases that studies on usage and capacity suffer from: (a) *Avoid behavioral change bias:* offering users with high capacity a further increase without their knowledge avoids the risk of behavioral changes that may occur when one purposefully buys a higher bandwidth connection; and (b) *Avoid frustrated user bias:* users already have a high capacity that gets upgraded, instead of opting for an upgrade because their previous capacity was insufficient for their usage. Studying datasets with these biases will always show a positive correlation between usage and capacity, and by examining a single high capacity tier, we avoid this.

Single ISP, Same Location: No bias between service plans, pricing model, and traffic treatment. Controlled setting. Paths + performance should be similar and unbiased by the ISP as data is from one city. Also avoids local behavioral biases (if any).

This gives us a highly controlled setting to study usage behaviors in an unbiased manner across a very large set of users (15k control and 1500 test households). Thus we believe that are conclusions will be representative of broadband behavior in a general American urban city. We expect the baseline behavior of all users to be similar, and in fact, interpret any differences between the control and test set behavior as aggregate changes that occurred due to the an increase in access link capacity.

3.2 Evaluation Criteria

We analyze based on three criteria:

- usage patterns
- peak utilization
- prime time ratio
- peak ratio
- traffic asymmetry

The Importance of Measuring Peak Usage:

- Capacity planning is concerned about prime-time and peak behavior, i.e., max data rate per device is more important than the average rate of that device.
- Idle hour behavior is expected to stay the same, regardless of access link
- peak utilization definition

Prime-Time Ratio definition of prime time

- Sandvine defines Network Prime-Time ratio to measure the concentration of network usage during the prime-time evening hours.
- FCC says prime-time is 7-11 PM.
- Prime-Time ratio = absolute levels of network traffic during an average peak period hour with an average off-peak hour.
- Measure PT = avg data rate during a peak hour period : off-peak period.

Peak Ratio

- define to measure imbalance in a day
- for each device,peak ratio = 90-%ile data rate per day / median data rate per day
- this metric can also be used to segregate users, based on who peaks at what time [section 5]

4 Results

TODO: This is the results section

We present a comparison between the usage characteristics of the control[3] and test[4]. We interpret the results as change in usage behavior due to increase in access link bandwidth.

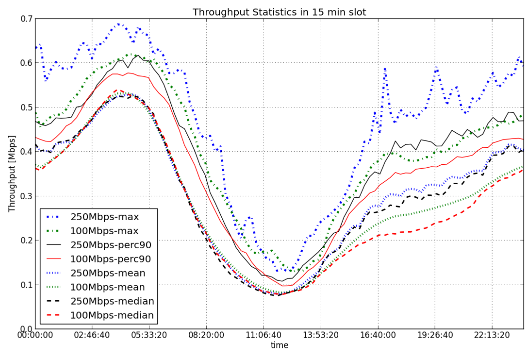
4.1 User Behavior

Issue #1.. Does aggregate user behavior differ?

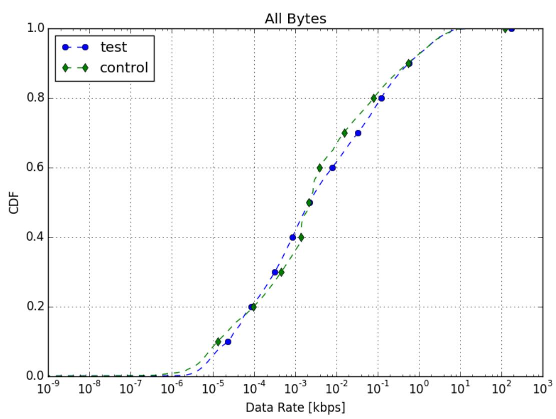
- Figure 2a
- Plot the aggregate data rate using bytes transferred per 15 min slot.

[3] gateway byte counters from users with a 105 Mbps access link
[4] gateway byte counters from users with a 105 Mbps access link

- Average across devices, then aggregate across days for each time slot.
- Peak data rates are reached between 8PM-12AM (2-6 UTC) TODO: replace with fixed time plot .
- Usual average usage patterns have a small peak in the morning (10AM), then dip, and then rise up again in the evening [5]. In contrast, there is no trough in the average usage for the patterns observed in our set.
- No troughs is consistent even on patterns observed weekly instead of daily.



(a) *agg (days) over means (devices)*



(b) *CDF of data rate per time slot for all devices (agg view of data)*

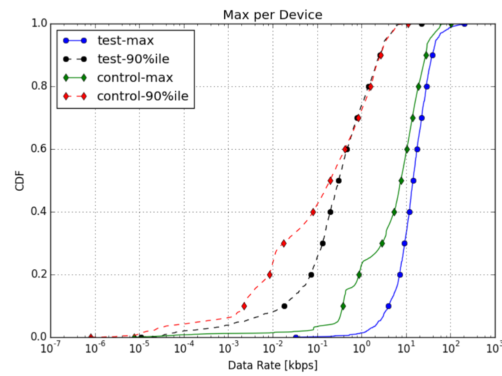
Figure 2: *User Behavior: Overall not much change due to capacity increase*

- Figure 2b
- Distribution of average data rate (kbps) per 15-min time slot
- Very similar distributions of bytes transferred
- Median data rate 2bps for 3 months x thousands of devices!

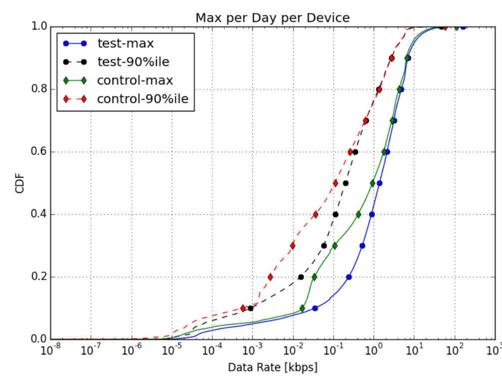
4.2 Peak Utilization

Issue #2.. Does the maximum (or 90-%ile) data transferred by the device differ?

- Figure 3a
- Calculate the maximum data rate for a device over the three months, and compare max rate (per device) for test set and max rate (per device) for control set
- test set has higher (max) average data rate below 10 kbps TODO: sanity check numbers, redo plot
- 30% of devices in the control set have a max data rate of 2 kbps while 30% of test set has a max data rate of 10 kbps.



(a) *CDF of max per device*



(b) *CDF of max per device daily*

Figure 3: *Peak Utilization: The maximum data rate varies for test and control set for low data rates, and this variation is present daily.*

- Figure 3b
- Max seen by a device per day, in 3 months
- Consistent increase for daily max data rate per device
- 15 min granularity misses information
  - short faster data bursts
  - better video quality
  - baseline is different

4.3 Prime Time Ratio

Issue #3.. Does the prime-time ratio differ?

- Figure 4
- 8p - 12a shows a higher prime-time ratio than 7p - 11p (as stated by FCC) in the Comcast data set. The curve in convex, i.e. there is only one peak time in a day.
- Distribution of prime time ratios over control and test set is very similar
- The average prime-time ratio for control sets is consistently higher than test set
- Test set has 10% lower PT than control set
- Diurnal pattern shows high ratio for 5 weekdays, and low ratio on weekends.
- Also easy to see Thanksgiving, black friday, Christmas, etc. holidays

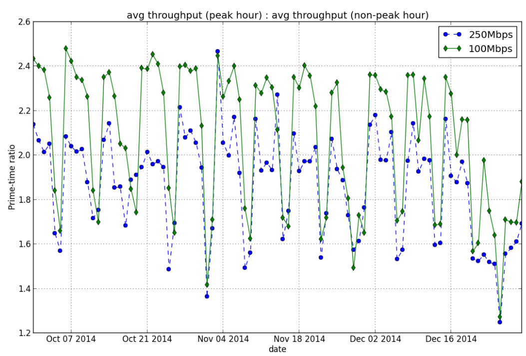


Figure 4: *Prime Time ratio showing weekly pattern + differences during holiday periods (Thanksgiving, Christmas)*

- Figure 4
- 8p - 12a shows a higher prime-time ratio than 7p - 11p (as stated by FCC) in the Comcast data set. The curve in convex, i.e. there is only one peak time in a day.
- Distribution of prime time ratios over control and test set is very similar
- The average prime-time ratio for control sets is consistently higher than test set
- Test set has 10% lower PT than control set
- Diurnal pattern shows high ratio for 5 weekdays, and low ratio on weekends.
- Also easy to see Thanksgiving, black friday, Christmas, etc. holidays

4.4 Peak Ratio

Issue #4.. How much does the traffic vary in a single day?

- Figure 5
- test-set sees higher variance in usage per device
- implies that ISPs should condition their networks to 50 times median usage per user

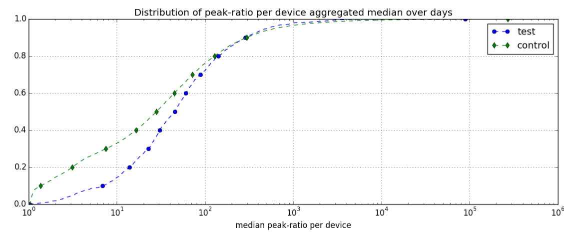


Figure 5: *Median peak ratio per device showing that test set has higher daily variance, which goes upto 50 times)*

4.5 Traffic Asymmetry

- TODO: maybe this doesn't need a separate section, just comment on asymmetry in each of the above
- Cisco vs Alcatel: upload is increasing vs there's still too much download
- Claim that its reducing due to uploads, but content is mostly download. Is the ratio still 10:1 (FCC thinks its 25:3)
- Compare with Sandvine asymmetry stats
- Talk about 3 Mbps comparison with observed uplink in control and test set

5 Discussion

- results that are contradictory
- results that data suggests that is wrong
- a better way to measure and offer broadband based on utilization?

5.1 User Taxonomy

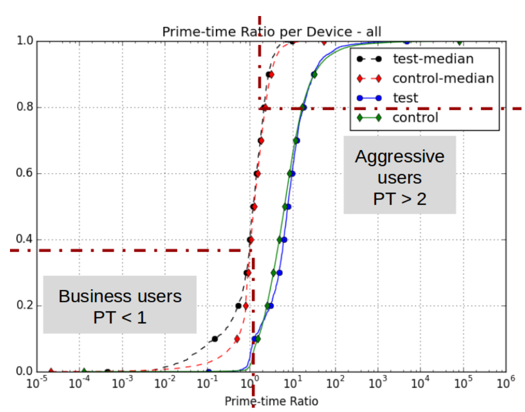


Figure 6: *(old) Prime Time ratio + usage can be used to divide users into four sets: aggressive all time + non aggressive all time, aggressive peak time, aggressive non-peak time (business hours))*

- Figure 6
- Sandvine does it by real-time entertainment usage

- cord cutters (¿ 85%ile): uses 11 times more data than the rest
- typical subscriber
- non streamers (¡ 15%ile): uses just 0.5

- the segregation splits business and casual users based on usage and motivates our discussion on multiple benchmarks based on usage.
- 30% PT < 1: possibly businesses with normal work-hours
- 20% PT > 2: aggressive prime-time streamers
- taxonomy: aggressive all time + non aggressive all time, aggressive peak time, aggressive non-peak time (business hours)

TODO: TO PLOT :

- peak ratio cdf vs no of devices
- peak ratio cdf vs time of day where peak occurred
- no of devices cdf vs time of day where peak occurred

6 Issues for Inquiry

In the FCC report No. 14-113, released on Aug 5, 2014 [2], the Commission asks some relevant questions about broadband usage, and requests comments from the community to improve its decision making process. We summarize their hypothesis and comments with regards to speed benchmarking as follows[5]

Issue #13. *Peak Time Usage.* Peak usage is defined as the average data consumed between 7:00 to 11:00 pm on weeknights. The FCC asks if evaluating peak time usage is an efficient metric, or if average usage over a day should be considered instead.

Issue #14. *Broadband Speed Benchmarks.* Recently, the FCC defined the minimum broadband speed to be 25 Mbps for downlink, and 3 Mbps for uplink, anticipating future usage. The questions asked by the Commission are as follows: (1) What is the required broadband speed for moderate usage? (2)How should the Commission forecast future household broadband uses to justify such a benchmark?

Issue #16. *Adoption Based Benchmarks.* The FCC's previous experience suggests that broadband benchmarks should be based 70% adoption rate to encourage providers to increase broadband penetration. This would motivate a benchmark of 1 Mbps, however, the uplink speed benchmark has been set to 3 Mbps. Does this aggressive limit adequately anticipate the increasing use of symmetrical services, such as two-way video calling?

Issue #19. *Popular Speed Tier Benchmarks.* The FCC is considering setting future speed benchmarks based on the fastest speed tier with a substantial customer subscription. Does a higher speed tier suggest that service of that speed is necessary to enable users to originate and receive high-quality voice, data, graphics, and video telecommunications?

Issue #22. *Other Speed Benchmarks.* Broadband requirements are not uniform throughout the nation. Some users will have significantly greater needs. The FCC is interested to know whether it should opt for multiple benchmarks depending on user scenario, usage, occupation, etc.?

[5]Note that the issue # corresponds to the point in the Eleventh Broadband Progress Report, FCC No. 15-10A1 [3]

References

[1] Joshua Romero, IEEE Spectrum. FCC Redefines Broadband: Lack of Competition Now Obvious, January 2015. (Cited on page 1.)

[2] Federal Communications Commission. Tenth Broadband Progress Report No 14-113, February 2014. (Cited on page 6.)

[3] Federal Communications Commission. Eleventh Broadband Progress Report No 15-10A1, February 2015. (Cited on pages 1 and 6.)

[4] Federal Communications Commission. Telecommunications Act of 1996, February 2015. (Cited on page 1.)

[5] Sandvine. Global Internet Phenomena Report - 1H, April 2014. (Cited on pages 1 and 4.)

[6] Sandvine. Global Internet Phenomena Report - 2H, November 2014. (Cited on page 1.)