Defining Network Capacity
draft-ietf-ippm-bw-capacity-02

Status of this Memo

Copyright Notice

Abstract

Measuring capacity is a task that sounds simple, but in reality can
be quite complex. In addition, the lack of a unified nomenclature on
this subject makes it increasingly difficult to properly build, test,
and use techniques and tools built around these constructs. This
document provides definitions for the terms 'Capacity' and 'Available
Capacity' related to IP traffic traveling between a source and
destination in an IP network. By doing so, we hope to provide a

common framework for the discussion and analysis of a diverse set of
current and future estimation techniques.

Table of Contents

[1](). Introduction

Measuring the capacity of a link or network path is a task that
sounds simple, but in reality can be quite complex. Any physical
medium requires that information be encoded and, depending on the
medium, there are various schemes to convert information into a
sequence of signals that are transmitted physically from one location
to another.

While on some media, the maximum frequency of these signals can be
thought of as "capacity", on other media, the signal transmission
frequency and the information capacity of the medium (channel) may be
quite different. For example, a satellite channel may have a carrier
frequency of a few gigahertz, but an information-carrying capacity of
only a few hundred kilobits per second. Often similar or identical
terms are used to refer to these different applications of capacity,
adding to the ambiguity and confusion, and the lack of a unified
nomenclature makes it difficult to properly build, test, and use
various techniques and tools.

We are interested in information-carrying capacity, but even this is
not straightforward. Each of the layers, depending on the medium,
adds overhead to the task of carrying information. The wired
Ethernet uses Manchester coding or 4/5 coding which cuts down
considerably on the "theoretical" capacity. Similarly RF (radio
frequency) communications will often add redundancy to the coding
scheme to implement forward error correction because the physical
medium (air) is lossy. This can further decrease the information
efficiency.

In addition to coding schemes, usually the physical layer and the
link layer add framing bits for multiplexing and control purposes.
For example, on SONET there is physical layer framing and typically
also some layer 2 framing such as HDLC, PPP or even ATM.

Aside from questions of coding efficiency, there are issues of how
access to the channel is controlled which may affect the capacity
also. For example, a multiple-access medium with collision
detection, avoidance and recovery mechanisms has a varying capacity
from the point of view of the users. This varying capacity depends
upon the total number of users contending for the medium, how busy
the users are, and bounds resulting from the mechanisms themselves.
RF channels are also varying capacity, depending on range,
environmental conditions, mobility and shadowing, etc.

The important points to derive from this discussion are these: First,
capacity is only meaningful when defined relative to a given protocol
layer in the network. It is meaningless to speak of "link" capacity

without qualifying exactly what is meant. Second, capacity is not
necessarily fixed, and consequently, a single measure of capacity at
whatever layer may in fact provide a skewed picture (either
optimistic or pessimistic) of what is actually available.

2. Definitions

   In this section, we specify definitions for capacity. We begin by
   first defining a baseline capacity that is simply tied to the
   physical properties of the link.

   Nominal Physical Link Capacity:
      Or NomCap(L) is the theoretical maximum amount of data that the
      link can support. For example, an OC-3 link would be capable of
      roughly 155 Mbps. We stress that this is a measurement at the
      physical layer and not the network IP layer, which we will define
      separately. While NomCap(L) is typically constant over time,
      there are links whose characteristics may allow otherwise, such as
      the dynamic activation of additional transponders for a satellite
      link.

   The nominal physical link capacity is provided as a means to help
   distinguish between the commonly used link layer capacities and the
   remaining definitions for IP layer capacity. As a result, the value
   of NomCap(L) does not influence the other definitions presented in
   this document.

   There are many factors that can reduce the IP information carrying
   capacity of the link, some of which have already been discussed in
   the introduction. However, the goal of this document is not to
   become an exhaustive list of of such factors. Rather, we outline
   some of the major examples in the following section, thus providing
   food for thought to those implementing the algorithms or tools that
   attempt to measure capacity accurately.

   The remaining definitions are all given in terms of "IP layer bits"
   in order to distinguish these definitions from the nominal physical
   capacity of the link.

   IP layer bits:
      Eight (8) times the number of octets in all IP packets received,
      from the first octet of the IP header to the last octet of the IP
      packet payload, inclusive.

   We then define a path P of length n as a series of links (L1, L2,
   ..., Ln) connecting a sequence of nodes (N1, N2, ..., Nn+1). A
   source, S, and destination, D, reside at N1 and Nn+1 respectively.
   Furthermore, we define a link L as a special case where the path size
   is one. The concept of links are consistent with [RFC2330] in that
   they represent a link-level connection between two nodes, and each
   node is not necessarily a router. However, the concept of a path
   differs slightly as intermediate nodes N2...Nn do not need to be
   routers. They may, for example, be Ethernet switches or other kinds

of layer 2 or layer 1 devices.

IP layer bits are recorded at the destination, D, beginning at time T and ending at a time T+I. Since the definitions are based on averages, the two time parameters, T and I, must accompany any report or estimate of the following values in order for them to remain meaningful. It is not required that the interval boundary points fall between packet arrivals at D. However, boundaries that fall within a packet will invalidate the packets on which they fall. Specifically, the data from the partial packet that is contained within the interval will not be counted. This may artificially bias some of the values, depending on the length of the interval and the amount of data received during that interval. We elaborate on what constitutes correctly received data in the next section.

IP Layer Link Capacity:
   We define the IP Layer link capacity, C(L,T,I), to be the maximum number of IP layer bits that can be transmitted from the source S and correctly received by the destination D over the link L during the interval [T, T+I], divided by I.

Using this, we can then extend this notion to an entire path, such that the IP layer path capacity simply becomes that of the link with the smallest capacity along that path.

IP Layer Path Capacity:
   C(P,T,I) = min {1..n} {C(Ln,T,I)}

The previous definitions specify a link's capacity, namely the IP information bits that can be transmitted across a link or path should the resource be free of any congestion. Determining how much capacity is available for use on a congested link is potentially much more useful. However, in order to define the available capacity we must first specify how much is being used.

IP Layer Link Usage:
   The average usage of a link L, Used(L,T,I), is the actual number of IP layer bits correctly transmitted from any source over link L during the interval [T, T+I], divided by I.

An important distinction between usage and capacity is that Used(L,T,I) is not the maximum amount, but rather, the actual amount of IP bits that are sent. The information transmitted across the link can be generated by any source, including those who may not be directly attached to either side of the link. In addition, each information flow from these sources may share any number (from one to n) of links in the overall path between S and D. Next, we express usage as a fraction of the overall IP layer link capacity.

   Average IP Layer Link Utilization:
      Util(L,T,I) = (Used(L,T,I) / C(L,T,I))

   Thus, the utilization now represents the fraction of the capacity
   that is being used and is a value between zero, meaning nothing is
   used, and one, meaning the link is fully saturated. Multiplying the
   utilization by 100 yields the percent utilization of the link. By
   using the above, we can now define the capacity available over the
   link as well as the path between S and D. Note that this is
   essentially the definition in [PDM].

   IP Layer Available Link Capacity
      AvailCap(L,T,I) = C(L,T,I) * (1 - Util(L,T,I))

   IP Layer Available Path Capacity
      AvailCap(P,T,I) = min {1..n} {AvailCap(Ln,T,I)}

   Since measurements of available capacity are more volatile that that
   of capacity, it is important that both the time and interval be
   specified as their values have a great deal of influence on the
   results. In addition, a sequence of measurements may be beneficial
   in offsetting the volatility when attempting to characterize
   available capacity.

3. Discussion

3.1 Standard or Correctly Formed Packets

   The definitions in this document specify that IP packets must be
   received correctly. The IPPM framework recommends a set of criteria
   for such standard-formed packet in section 15 of [RFC2330]. However,
   it is inadequate for use with this document. Thus, we outline our
   own criteria below while pointing out any variations or similarities
   to [RFC2330].

   First, data that is in error at layers below IP and cannot be
   properly passed to the IP layer should not be counted. For example,
   wireless media often has a considerably larger error rate than wired
   media, resulting in a reduction in IP Link Capacity. In accordance
   with the framework, packets that fail validation of the IP header
   should be discarded. Specifically, the requirements in [RFC1812]
   section 5.2.2 on IP header validation should be checked, which
   includes a valid length, checksum, and version field.

   The framework specifies further restrictions, requiring that any
   transport header be checked for correctness and that any packets with
   IP options be ignored. However, the definitions in this document are
   concerned with the traversal of IP layer bits. As a result, data
   from the higher layers is not required to be valid or understood as
   they are simply regarded as part of the IP packet. The same holds
   true for IP options. Valid IP fragments should also be counted as
   they expend the resources of a link even though assembly of the full
   packet may not be possible. The framework differs in this area,
   discarding IP fragments.

   In summary, any IP packet that can be properly processed should be
   included in these calculations.

3.2 Other Potential Factors

   The base definitions make no mention of hardware duplication of
   packets. While hardware duplication has no impact on the nominal
   capacity, it can impact the IP link layer capacity. For example,
   consider a link which can normally carry a capacity of 2X on average.
   However, the link has developed a syndrome where it duplicates every
   incoming packet. The link would still technically carry a capacity
   of 2X, however the link has a effective capacity of X or lower,
   depending on framing overhead to send the duplicates, etc. Thus, a
   value for C(L,T,I) and AvailCap(L,T,I) will reflect the duplication
   with the lower value.

   IP encapsulation does not impact the definitions as all IP header and

payload bits should be counted regardless of content. However,
different sized IP packets can lead to a variation in the amount of
overhead needed at the lower layers to transmit the data, thus
altering the overall IP link layer capacity.

Should the link happen to employ a compression scheme such as ROHC
[RFC3095] or V.44 [V44], some of the original bits are not
transmitted across the link. However, the inflated (not compressed)
number of IP-layer bits should be counted.

3.3 Common Literature Terminology

Certain terms are often used to characterize specific aspects of the
presented definitions. The link with the smallest capacity is
commonly referred to as the "narrow link" of a path. Also, the value
of n that satisfies AvailCap(P,T,I), is often referred to as the
"tight link" within a path. So, while Ln may have a very large
capacity, the overall congestion level on the link makes it the
likely bottleneck of a connection. Conversely, a link that has the
smallest capacity may not be a bottleneck should it be lightly loaded
in relation to the rest of the path.

Also, common literature often overloads the term "bandwidth" to refer
to what we have described as capacity in this document. For example,
when inquiring about the bandwidth of a 802.11b link, a network
engineer will likely answer with 11 Mbps. However, an electrical
engineer may answer with 25 MHz, and an end user may tell you that
his observed bandwidth is 8 Mbps. In contrast, the term capacity is
not quite as overloaded and is an appropriate term that better
reflects what is actually being measured.

3.4 Comparison to Bulk Transfer Capacity (BTC)

Bulk Transfer Capacity (BTC) [RFC3148] provides a distinct
perspective on path capacity that differs from the definitions in
this document in several fundamental ways. First, BTC operates at
the transport layer, gauging the amount of capacity available to an
application that wishes to send data. Only unique data is measured,
meaning header and retransmitted data are not included in the
calculation. In contrast, IP layer link capacity includes the IP
header and is indifferent to the uniqueness of the data contained
within the packet payload (Hardware duplication of packets is an
anomaly addressed in the previous section). Second, BTC utilizes a
single congestion aware transport connection, such as TCP, to obtain
measurements. As a result, BTC implementations react strongly to
different path characteristics, topologies, and distances. Since
these differences can affect the control loop (propagation delays,
segment reordering, etc), the reaction is further dependent on the

algorithms being employed for the measurements. For example,
consider a single event where a link suffers a large duration of bit
errors. The event could cause IP layer packets to be discarded, and
the lost packets would reduce the IP layer link capacity. However,
the same event and subsequent losses would trigger loss recovery for
a BTC measurement resulting in the retransmission of data and a
potentially reduced sending rate. Thus, a measurement of BTC does
not correspond to any of the definitions in this document. Both
techniques are useful in exploring the characteristics of a network
path, but from different perspectives.

3.5 Type P Packets

Note that these definitions do not make mention of "Type P" packets,
while other IPPM definitions do. We could add the packet type as an
extra parameter. This would have the effect of defining a large
number of quantities, relative to the QoS policies that a given
network or concatenation of networks may have in effect in the path.
It would produce metrics such as "estimated EF IP Link/Path Capacity"
or "estimated EF IP Link Utilization".

Such metrics may indeed be useful. For example, this would yield
something like the sum of the capacities of all the QoS classes
defined along the path as the link or path capacity. The breakdown
then gives the user an analysis of how the link or path capacity (or
at least the "tight link" capacity) is allocated among classes.

These QoS-based capacities become difficult to measure on a path if
there are different capacities defined per QoS class on different
links in the path. Possibly the best way to approach this would be
to measure each link in a path individually, and then combine the
information from individual links.

3.6 Time and Sampling

We must emphasize the importance of time in the basic definitions of
these quantities. We know that traffic on the Internet is highly
variable across all time scales. This argues that the time and
length of measurements are critical variables in reporting available
capacity measurements.

The closer to "instantaneous" a metric is, the more important it is
to have a plan for sampling the metric over a time period that is
sufficiently large. By doing so, we allow valid statistical
inferences to be made from the measurements. An obvious pitfall here
is sampling in a way that causes bias. For example, a situation
where the sampling frequency is a multiple of the frequency of an
underlying condition.

[4](#). **Conclusion**

   In this document, we have defined a set of quantities related to the
   capacity of links in an IP network. In these definitions, we have
   tried to be as clear as possible and take into account various
   characteristics that links can have. The goal of these definitions
   is to enable researchers who propose capacity metrics to relate those
   metrics to these definitions and to evaluate those metrics with
   respect to how well they approximate these quantities.

   In addition, we have pointed out some key auxiliary parameters and
   opened a discussion of issues related to valid inferences from
   available capacity metrics.

[5](). IANA Considerations

This document makes no request of IANA.

Note to RFC Editor: this section may be removed on publication as an RFC.

6. **Security Considerations**

   This document specifies definitions regarding IP traffic traveling
   between a source and destination in an IP network. These definitions
   do not raise any security issues and do not have a direct impact on
   the networking protocol suite.

7. Acknowledgments

   The authors would like to acknowledge Mark Allman, Patrik Arlos, Matt
   Mathis, Al Morton, and Stanislav Shalunov for their suggestions,
   comments, and reviews. We also thank members of the IETF IPPM
   Mailing List for their discussions and feedback on this document.

8. References

8.1 Normative References

8.2 Informative References

[PDM]   Dovrolis, C., Ramanathan, P., and D. Moore, "Packet
             Dispersion Techniques and a Capacity Estimation
             Methodology", IEEE/ACM Transactions on Networking 12(6):
             963-977, December 2004.

[RFC1812]   Baker, F., "Requirements for IP Version 4 Routers",
             RFC 1812, June 1995.

[RFC2119]   Bradner, S., "Key words for use in RFCs to Indicate
             Requirement Levels", BCP 14, RFC 2119, March 1997.

[RFC2330]   Paxson, V., Almes, G., Mahdavi, J., and M. Mathis,
             "Framework for IP Performance Metrics", RFC 2330,
             May 1998.

[RFC3095]   Bormann, C., Burmeister, C., Degermark, M., Fukushima, H.,
             Hannu, H., Jonsson, L-E., Hakenberg, R., Koren, T., Le,
             K., Liu, Z., Martensson, A., Miyazaki, A., Svanbro, K.,
             Wiebke, T., Yoshimura, T., and H. Zheng, "RObust Header
             Compression (ROHC): Framework and four profiles: RTP, UDP,
             ESP, and uncompressed", RFC 3095, July 2001.

[RFC3148]   Mathis, M. and M. Allman, "A Framework for Defining
             Empirical Bulk Transfer Capacity Metrics", RFC 3148,
             July 2001.

[V44]   ITU Telecommunication Standardization Sector (ITU-T)
             Recommendation V.44, "Data Compression Procedures",
             November 2000.

Authors' Addresses

    Phil Chimento
    JHU Applied Physics Lab
    11100 Johns Hopkins Road
    Laurel, Maryland 20723-6099
    USA

    Phone: +1-240-228-1743
    Fax: +1-240-228-0789
    Email: Philip.Chimento@jhuapl.edu

    Joseph Ishac
    NASA Glenn Research Center
    21000 Brookpark Road
    Cleveland, Ohio 44135
    USA

    Phone: +1-216-433-6587
    Fax: +1-216-433-8705
    Email: jishac@grc.nasa.gov

Intellectual Property Statement

Disclaimer of Validity

Copyright Statement

Acknowledgment