Status of this Memo

This Internet-Draft is submitted to IETF in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF), its areas, and its working groups. Note that other groups may also distribute working documents as Internet-Drafts.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

The list of current Internet-Drafts can be accessed at http://www.ietf.org/ietf/1id-abstracts.txt

The list of Internet-Draft Shadow Directories can be accessed at http://www.ietf.org/shadow.html

This Internet-Draft will expire on April 17, 2011.

Copyright Notice

Copyright (c) 2012 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust’s Legal Provisions Relating to IETF Documents (http://trustee.ietf.org/license-info) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.
Abstract

This draft proposes some use cases for consideration by the iccrg. These use cases are in addition to and/or complement those described by the conex working group[UseCases]. The conex working group had determined that these use cases were out of scope of the current charter, and/or could potentially be built out of the conex abstract mechanism. The focus of the use cases in this draft is on forms of congestion exposure that involve resources other than queues and timeframes other than real-time. Background and motivation for these use cases is first discussed. Expanded material on the usage tier/volume use case is also provided.

Table of Contents

1. Introduction...................................................2
2. Conventions used in this document..............................3
   2.1. Acronyms..................................................3
   2.2. Terminology............................................3
3. Motivation and Background......................................4
4. Potential Use Cases for Experimental Research...............5
   4.1. Inequity of Heavy versus Light Users.................5
   4.2. Feedback on Time of Day, Day of Week Charging........6
   4.3. Recharging for Implementing Congestion Pricing........7
   4.4. Usage Tier/ Volume Feedback..........................7
      4.4.1. Objectives for Addressing this Issue............7
      4.4.2. Potential Support Using Abstract Mechanism.....8
      4.4.3. Additional Support Using other Measures and Mechanisms9
5. Security Considerations....................................11
6. IANA Considerations........................................11
7. Acknowledgements............................................11
8. References..................................................11
   8.1. Normative References...................................11
   8.2. Informative References................................11
9. Acknowledgments.............................................12

1. Introduction

This draft proposes some use cases for inclusion in the conex Working group charter’s deliverable for an informational RFC covering use case description. These use cases are in addition to and/or complement those described in [UseCases], and focus on forms of congestion exposure that involve resources other than queues and timeframes other than real-time.

Section 3 provides some motivational background and a statement of problems involved with congestion pricing, with references to the presentations by experts in this area at the IETF 78 Technical Plenary in Maastricht.
Section 4 provides text for each of the above use cases in a mechanism independent manner.

As requested in the Beijing meeting, this is an individual draft that expands on the usage tier/volume use case from [McDysan]. The feedback recorded in the Beijing conex meeting minutes was that a number of people were interested in this idea, but that it was out of scope of the current charter, and/or could potentially be built out of the conex abstract mechanism.

Section 3 provides some motivational background and a statement of relevant problems involved with congestion pricing.

Section 4 provides text for the usage/volume tier feedback use case. It contains a section that covers a problem statement, objectives for resolving this issue, potential approaches for implementing this use case employing currently defined conex mechanisms, and a description of additional measures and mechanisms that could solve the stated problem and issues.

2. Conventions used in this document

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in RFC-2119 [RFC2119].

2.1. Acronyms

conex = congestion exposure

2.2. Terminology

The following is a quote from the CONEX working group charter:

"... develop a mechanism by which senders inform the network about the congestion encountered by previous packets on the same flow ... at the IP layer, such that the total level of congestion is visible to all IP devices along the path"

Despite its central role in network control and management, congestion is a remarkably hard concept to define. The discussions in [Bauer09] provide a good academic background. [RFC6077] defines it as "as a state or condition that occurs when network resources are overloaded, resulting in impairments for network users as objectively measured by the probability of loss and/or delay." An economist might define it as the condition where the utility of a given user decreases due to an increase in network load. The economic implications are obviously different for an end user versus a service provider.
3. Motivation and Background

This section provides references to relevant presentations given by experts on congestion pricing from the IETF 78 Technical Plenary in Maastricht.

Successful adoption of experimental congestion pricing/exposure/information sharing protocol(s) must address use cases that provide significant value to users, content providers and service providers. Central themes to this value proposition are incentives (i.e., congestion pricing) and the cost of providing marginal capacity [Varian, Johari].

There are three time scales over which congestion pricing can operate [Johari]: short (milliseconds to seconds), medium (minutes to hours to days) and long (months to years). Currently, the short term congestion signal is lost packets or a specific indication of congestion of a particular resource (e.g., a ECN indication for queue congestion), as stated in the conex charter. Setting congestion price as the marginal value of capacity is useful for medium timescales via traffic engineering and longer time scales via provisioning [Johari].

Some congestion exposure problems are challenging to address and are not completely addressed in conex [UseCases]:

- 20% of the users generate 80% of the traffic and create unfairness with certain resource sharing [Varian]
- Volume-based pricing makes it difficult for users to manage costs incurred, [Varian]. Note that a significant number of ISPs implement some form of usage/volume caps in at least some parts of the world [NewAmerica].
- Customers will pay a premium for unmetered use [Varian]
- A form of congestion pricing is "recharging" (e.g., "free shipping") [Varian], where someone other than the end user pays for incurred congestion.
- In general a content or service provider has hard capacity constraints at certain bottlenecks in their infrastructure (e.g., server capacity, router interface/queue service rate). Some form of adaption, such as time-shifting, route-shifting, or moderating the demand is required to adapt to these constraints [Kelly].

If conex exposes congestion without damage (e.g., loss) then many forms of adaption are feasible, as long as incentives are aligned with the signaled congestion [Kelly]. The use cases in the next section focus on forms of adaption that enable certain sets of incentives that are not completely covered in [UseCases].
4. Potential Use Cases for Experimental Research

The following use cases have been deemed to be out of scope for the current conex wg charter. They may be candidates for specific areas of research.

4.1. Inequity of Heavy versus Light Users

In many networks, 20% are heavy users generating 80% of the traffic [Varian]. This means that a heavy user generates 16 times the traffic as a light user as a medium term (e.g., monthly) average. But, in a bandwidth-tiered flat priced network, heavy and light users often pay nearly the same price since pricing is based upon the short-term (milliseconds) bandwidth measure of a shaper and/or a policer.

During non-peak periods, the resources of a service or content provider are underutilized and the marginal cost of capacity is small.

The access network is provisioned and traffic engineered for peak capacity of all users, and when congested, heavy users create 16 times the congestion of small users.

However, during peak use periods, a heavy user may send at near the bandwidth tier while light users may send intermittently. There is a need for a means for service providers to equitably assign costs to heavy versus light users. For example, the light users may pay less if they were charged by volume, as described in another use case.

A congestion measure of burstiness (e.g., ratio of peak rate to average rate over a longer interval than the tiered bandwidth shaper or policer) could be helpful in this use case. In general, a bursty packet flow is light (e.g., web surfing) and a non-bursty packet flow is heavy (e.g., viewing a lengthy HD video). The destination could perform this measure and feedback it back to the sender. It may only be necessary to feedback this measure for heavy users, since the absence of such feedback could be inferred as an indication of a light user. The sender could insert some processed version of this feedback measure this at the IP layer so that all IP devices could be aware of whether this is a heavy or light user.

The following use case is an example of how the conex indication of congestion may be useful to handle the heavy versus light user case. The following was based upon discussions with Toby Moncaster.

Over a time period related to the statistical multiplexing or economic congestion interval (e.g., many seconds to minutes to hours) total up the number of bytes that have been congestion marked and the total number of bytes sent per end-user. Compute the ratio of congested bytes to total bytes. This measures the average rate per user.
Quantizing users into classes using one threshold on total and another threshold on ratio results in a grid that identifies four classes of user:

<table>
<thead>
<tr>
<th></th>
<th>Volume</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ratio</td>
<td>Large</td>
</tr>
<tr>
<td>High</td>
<td>Heavy User</td>
</tr>
<tr>
<td>Low</td>
<td>LEDBAT User</td>
</tr>
</tbody>
</table>

(Where "LEDBAT User" includes other Less-than-Best-Effort algorithms.)

Figure 1: Four Classes of User

Note that Bursty and Heavy Users contribute more to congestion marking, but a Bursty user contributes less overall congestion marking and may be creating shorter periods of queue filling as compared with heavy users. LEDBAT and light users create less to congestion marking, with LEDBAT users able to transfer more volume as compared with light users since LEDBAT users back off before congestion marking occurs. An operator might reasonably take this into account in their shaping algorithms.

4.2. Feedback on Time of Day, Day of Week Charging

Congestion occurs when the offered load approaches that of the provisioned capacity, which often does not occur until shortly before there would be a need to provision additional capacity. Depending upon how restoration capacity is allocated by a service/content provider, congestion may only occur during peak periods when a failure is present.

Without Conex, utilization averaged over several minutes can be as high as 70 to 80% in typical network bottlenecks without loss that would reduce TCP effective throughput (i.e., goodput). A short term congestion control (sub-second to seconds) method that could increase utilization to above 90% and still avoid loss would only increase effective capacity by 10-20%.

If traffic increases at 50-75% per year, then a 10-20% increase in effective capacity means that the provisioning interval is only shifted by a few months. This handling of short term congestion use case alone may not be sufficient motivation for a service provider to deploy congestion handling measures.

However, in many points of a network, the majority of usage occurs during peak periods (e.g., a few busy hours) while much spare capacity exists off peak. The product of the spare capacity (bits/second) and
the non-peak interval (seconds) that could carry traffic ranges from 2 to 10 times of the traffic carried during the peak period.

Congestion exposure and congestion pricing that enables users and content providers to time shift traffic to off-peak periods that would have otherwise been sent during peak periods can reduce provisioned capacity cost by as much as several hundred percent.

Historical time of day usage patterns could be employed to time shift traffic, but often maintenance actions are performed during the off-peak periods, making the prediction of congestion using these methods less reliable. An automatic method for detection of congestion during off-peak periods is highly desirable.

4.3. Recharging for Implementing Congestion Pricing

There should be a means to recharge (i.e., someone other than the receiving user pays) for usage that causes congestion during peak demand period versus that which does not [Varian].

If TCP were augmented with information related to the form of congestion, including not only short term as covered in [UseCases], but also including usage tier, Time of Day, or burstiness then sufficient information to implement sender pays (e.g., Content provider) versus receiver pays (e.g., end user) could be implemented. When the sender includes this information in the IP layer, then usage tier counting and TOD counting could be accounted for differently in IP devices in the path between sender and destination. Such an indication of recharging would need be to be authenticated in some way.

4.4. Usage Tier/ Volume Feedback

Usage/volume caps may be arranged into multiple tiers with different pricing based upon monthly volume. This results in some problems as described in the next section. Next, high-level objectives to address these issues are then proposed. Then potential ways that already defined means [Mechanisms] may be employed are described. Finally, to address some of these issues, other measures and mechanisms that could possibly better meet the objectives are described.

4.4.1. Objectives for Addressing this Issue

Provide a way to inform a receiver of the usage/volume incurred to a moment in time. Ideally, this would also include the usage/volume time period (e.g., a month). This should be accomplished without having the user logging in to a portal to retrieve usage information, for example, as described in [Feamster11], but should be provided as an interface usable (potentially as an API) to any authorized application.
Provide a way to inform a receiver of a trend that if usage continues at the same rate then a specific usage/volume tier will be crossed.

Indicate to a receiver whether usage/volume counting is occurring in a different way when congestion measure of a particular form (e.g., loss, ECN marking) is occurring.

Standardize a way to mark packets in a way (e.g., [Lower Effort]) in conjunction with some form of conex signaling that indicates usage/volume counting will not occur (or are counted separately) under the condition that these packets do not create congestion. A means to ensure that these marked packets do not create congestion and do not impact best (and better) effort marked packets is also required.

Enable a means for recharging to occur, where usage/volume counting does not occur for the receiving user since some other party has agreed to incur the cost of usage/volume for that flow.

4.4.2. Potential Support Using Abstract Mechanism

The conex abstract mechanism [Mechanism] defines implicit signaling of loss and explicit signaling of ECN marking. It also defines re-echoed signal for loss and ECN marking based upon feedback carried by TCP from a receiver back to the sender (in an RFC to be developed by the conex wg as defined in the charter). If this feedback mechanism is designed to be extensible, then a variety of forms of feedback could be developed for use in experiments.

Counting usage/volume differently for congested packets (or congested intervals) based upon [Mechanism] re-echoed congestion experienced signals seems straightforward. This could be a local matter for the IP node which implements usage/volume tier counting.

Also, counting packets marked as Lower Effort differently is a local matter. How to ensure that these packets do not interfere with best effort could be implemented by Diffserv methods locally and at other potential bottlenecks.

Since packets subject to counting in a usage/volume cap may not occur during congestion intervals, reinsertion of such counting information using the re-echoed signals that indicate congestion does not seem possible since the same bits cannot represent usage counting and congestion experienced.

What is missing from the current conex mechanism is a feed forward path operating over a longer timescale that contains sufficient information which can be provided to an individual application to meet the stated objectives.
4.4.3. Additional Support Using other Measures and Mechanisms

Usage/volume counting has some aspects similar to that of a congestible queue, but on a much longer timescale, as follows:

- Instead of a queue which is typically sized for O(10 ms) at the sending rate, usage/volume counting occurs on a timescale O(month).
- A usage/volume tier is a threshold on a long term usage counter, similar to the way ECN marking can be a threshold on in a queue.
- Queue loss is similar to a usage/volume counter crossing from one tier into the next.
- A usage/volume tier trend warning is similar to a rate estimate for ECN marking based upon queue fill rate, as is described in PCN.

Therefore, in an abstract way a usage/volume counter can be viewed as a congestible resource, but in some ways not the same as a congestible queue. If this information is to be fed forward in a way observable at the IP layer and fed back at the transport layer (e.g., TCP), then additional packet and transport fields and/or mechanisms may be better suited to this purpose.

Furthermore, instead of feeding forward information in each IPv6 packet as in [Mechanism], usage/volume congestible resource information can occur much less frequently (e.g., many minutes to hours). The following is an outline of such measures and mechanisms.

The basic idea is based on the fact that the sender and receiver need to be cooperating using the same experimental extensions to TCP, and that if TCP can carry some of the additional information, then the scarcity of IPv6 header bits is avoided. Furthermore, as described previously, "fast path" processing is not required for this use case which has resulted in conex specifying the destination options field [conexdestopt]. Instead, the hop-by-options field of the IPv6 header could be used [RFC2460], [IPv6Format] which would only require "slow path" processing. A mechanism to allow the experimental sender to send a "probe" in the IPv6 packet (e.g., using an experimental IPv6 protocol type) could be used by an intermediate IP node(s) to forward the "probe" packet to a special processor (which may be separate from the routers' processor). This special processor could use a polled version of usage/volume count information per user and could also be configured with subscription information (e.g., usage/volume cap tier, cap duration), and threshold settings.

The handling of this "probe" IPv6 packet and associated TCP segment needs to be done within the TCP flow. It could use an Out Of Band mechanism similar to the urgent data capability in TCP. (For example, an experimental usage of the Urgent bit could possibly be employed.) The special processor could insert additional measures and implement
some of the proposed mechanisms and then modify/augment the "probe"
TCP (urgent-like) segment with the requested information and forwards
this modified "probe" packet toward the receiver via the intermediate
IP node. Packets from the receiver back to the sender could be sent
directly, or could be directed through the special processor at
intermediate node(s), depending upon the specifics of the use case
involved.

A consequence of the above extension of measures and mechanisms is
that the sender and receiver now have much more information which
could be used to solve the stated problems and meet the objectives.

The information carried in a "probe" TCP segment could include:

- The service being requested, for example:
  - Request information on the users' usage/volume tier
  - Request statistics on usage
  - Request threshold trend report
  - Request not counting this flow since it is lower effort
  - Request recharging

- Information that could be provided by the "special processor"
  includes:
  - Duration and cap for the usage volume measurement tier (e.g., a
    month)
  - The absolute count of packets and octets received/sent, and/or
    fraction of the usage tier already used
  - Count of packets and octets received/sent which experienced
    congestion
  - Count of packet and octets received/sent that were marked as
    Lower Effort
  - Estimate of whether the user will exceed the usage tier if the
    historical usage rate to the reporting instant continues
  - A pointer (e.g., URL) and identification of the authentication
    method that would enable other queries, and/or implement
    alternative charging methods (e.g., recharging)
  - Other measures related to the "congestion" of a usage/volume
    tier use case (or possibly other use cases as well).
One example of a different type of measure is described in [Stanojevic]. In this paper, the Shapley value [Shapley] is used instead of a 95-th percentile measure of hourly usage measurements across a month. The Shapley value has the following desirable intuitive properties [Shapley]: individual fairness, efficiency, symmetry and additivity. Although the 95-th percentile measure is not directly related to the usage/volume tier use case, the authors state that this is a case they plan to address in future research. A mapping, such as an approximation to the Shapley value described in this paper, could be a way to compress the usage/volume tier feed forward/ feedback information into a smaller number of bits that represents the incentives described in the objectives section.

5. Security Considerations

In the proposed mechanisms there are indications that could be spoofed and/or used to game counting and congestion feedback mechanisms, and therefore an authentication mechanism may be needed when this information is handled at the TCP/IPv6 layer in the sender to destination direction or at the TCP layer in the destination to sender direction.

6. IANA Considerations

None

7. Acknowledgements

The idea of how to use the experimental mechanism being developed by the conex working group to address heavy versus light users was suggested by Toby Moncaster. The idea of not counting lower effort traffic against a usage/volume cap was suggested my Mikael Abrahamsson on the conex mailing list. Lars Eggert provided the reference to [Stanojevic] on the conex mailing list as an example of a different form of congestion measure.

8. References

8.1. Normative References


8.2. Informative References


[Johari] Ramesh Johari, Stanford University, "The information in congestion prices: milliseconds to years," IETF 78 Technical Plenary, 29 July 2010


9. Acknowledgments

This document was prepared using 2-Word-v2.0.template.dot.

Copyright (c) 2012 IETF Trust and the persons identified as authors of the code. All rights reserved.
THIS SOFTWARE IS PROVIDED BY THE COPYRIGHT HOLDERS AND CONTRIBUTORS "AS IS" AND ANY EXPRESS OR IMPLIED WARRANTIES, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE ARE DISCLAIMED. IN NO EVENT SHALL THE COPYRIGHT OWNER OR CONTRIBUTORS BE LIABLE FOR ANY DIRECT, INDIRECT, INCIDENTAL, SPECIAL, EXEMPLARY, OR CONSEQUENTIAL DAMAGES (INCLUDING, BUT NOT LIMITED TO, PROCUREMENT OF SUBSTITUTE GOODS OR SERVICES; LOSS OF USE, DATA, OR PROFITS; OR BUSINESS INTERRUPTION) HOWEVER CAUSED AND ON ANY THEORY OF LIABILITY, WHETHER IN CONTRACT, STRICT LIABILITY, OR TORT (INCLUDING NEGLIGENCE OR OTHERWISE) ARISING IN ANY WAY OUT OF THE USE OF THIS SOFTWARE, EVEN IF ADVISED OF THE POSSIBILITY OF SUCH DAMAGE.

This code was derived from IETF RFC [insert RFC number]. Please reproduce this note if possible.

Authors’ Addresses

    Dave McDysan
    Verizon
    Email: dave.mcdysan@verizon.com