Potential Use Cases of Internet Congestion Control Research

draft-mcdysan-iccrg-usecases-00

Dave McDysan

Outline

- Motivation and Background
  - Maastricht technical plenary focusing on “congestion pricing”
  - Multiple valid viewpoints of what congestion means (e.g., network provider, economic)
- History
  - Presented in conex wg in Beijing and Prague (Usage/Volume tier)
  - Some interest, but determined to be out of scope of current conex wg charter
- Proposed Use Cases
  - Usage Tier/ Volume Feedback
  - Feedback on Time of Day, Day of Week Charging
  - Recharging for Implementing Congestion Pricing
  - Inequity of Heavy versus Light Users
    - May be implementable using conex mechanism
- Conclusions & Recommendations

Motivation and Background

- Value proposition centers on incentives (i.e., congestion pricing) and cost of providing marginal capacity
- Timescales of congestion pricing and example responses
  - Short (ms to sec): ECN
  - Medium (min to hrs/days): Traffic Engineering
  - Long (mon to yrs): provisioning marginal capacity
- Challenges identified in Maastricht Technical Plenary not (completely) addressed by current conex wg charter
  - 20% of the users generate 80% of the traffic and create unfairness
  - Volume-based pricing makes it difficult for users to manage costs incurred
  - Customers will pay a premium for unmetered use
  - A form of congestion pricing is "recharging" (e.g., "free shipping") where someone other than the end user pays for incurred congestion.
  - Some form of adaption, such as time-shifting, route-shifting, or moderating demand is required to bottlenecks in service provider networks
- 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
- Multiple valid viewpoints exist for congestion, some not completely addressed in conex include: [Bauer 09]
  - Network Operator and Economic

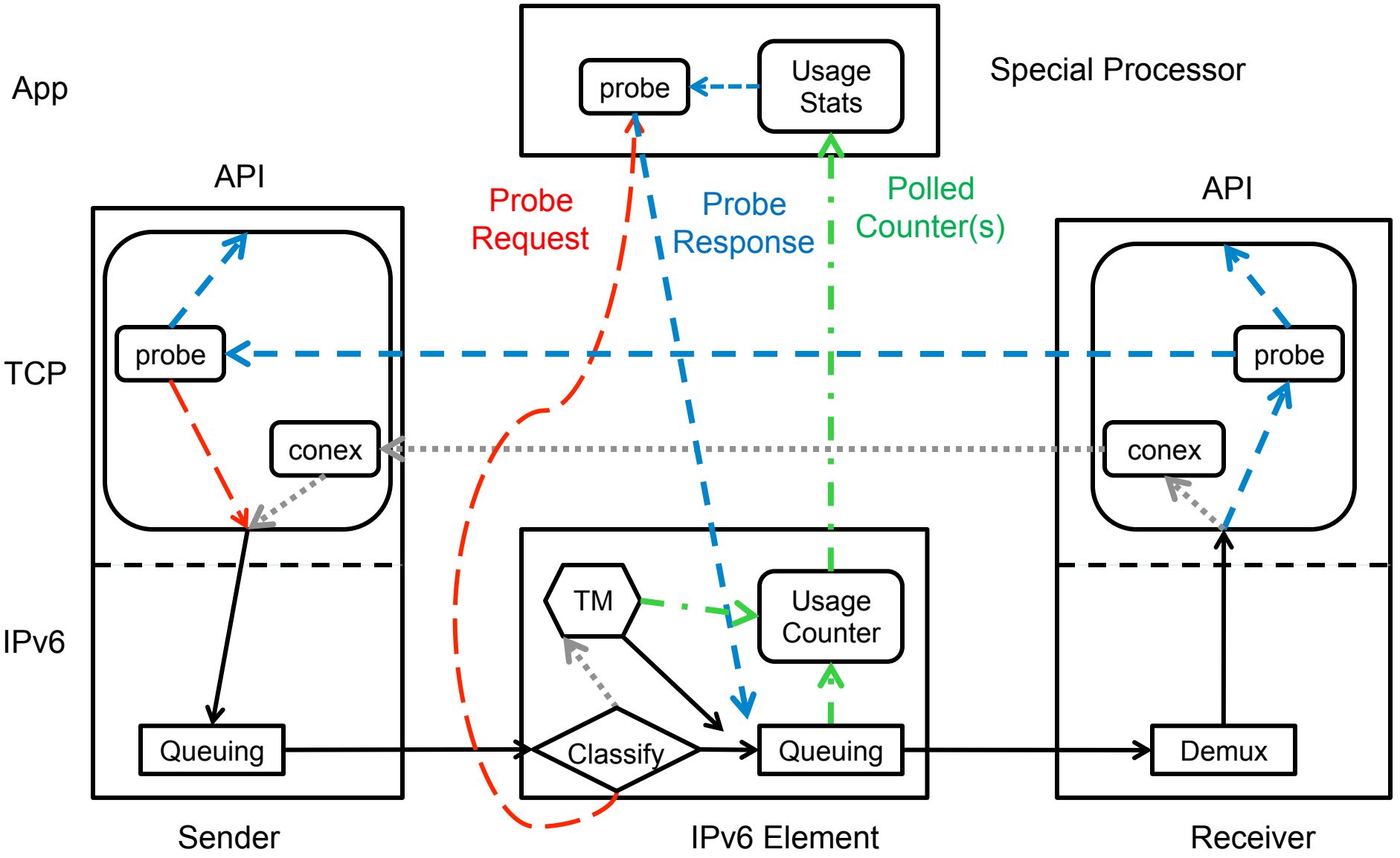
Usage Tier/ Volume Feedback

- Problem Statement
  - Complex for users to track/manage volume usage
  - Volume counting doesn't discriminate between heavy usage when congestion occurs or doesn't
  - Need better incentive for LEDBAT style and/or lower effort transport
- Objectives
  - Inform receiver and sender of cumulative volume and tier crossing trend
  - Inform receiver and sender whether congestion counting is occurring
  - Standardize on means to indicate to receiver and sender sets of packets not subject to congestion counting
  - Enable a means for recharging

Additional Mechanisms

- Usage/volume counter similar to a forwarding queue in conex, but operates over much longer timescale
- Since timescale is large, no need to feed forward information in each packet as in conex
  - Most benefit occurs for long-lived, heavy volume flows
    - e.g., video streaming or large file transfer
- Could use experimental TCP extensions and IPv6 hop-by-hop options header to implement feed forward “probe” packets from sender to receiver
  - Requires cooperation between TCP sender and receiver similar to that assumed in Conex
  - Needs to be part of TCP flow (e.g., possible experimental use of urgent pointer?)
  - “Probe” packets at IPv6 nodes don’t require fast path processing
    - these packets could be handled by a “special processor”
    - Could possibly be done using the OpenFlow protocol

Block Diagram of Conex & Additional Mechanism



Probe Request Packet

- Periodically transmitted by sender
- Intercepted by IPv6 element supporting experimental codepoints and forwarded to Special Processor
- Probe Request Contents
  - Request information on the receiving users usage/ volume tier
  - Request statistics on usage
  - Request threshold trend report
  - Request not counting this flow since it is lower effort

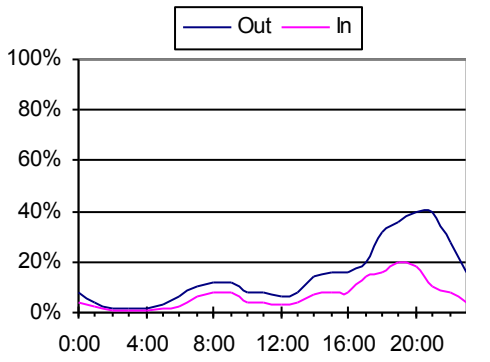
Probe Response Packet

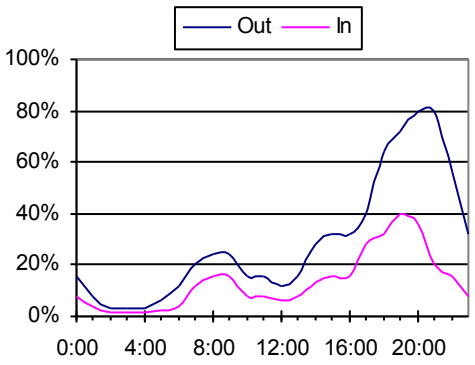
- Generated by Special Processor from Polled Usage Counters and IPv6 Element config
- Delivered to receiver (and API) and relayed back to sender (and API)
- Example Contents
  - Duration and cap for the volume measurement tier
  - Packets and octets received/sent
    - Total, conex marked, dropped, lower effort
  - Fraction of the usage tier already used
  - Tier crossing alert if current trend persists
  - A pointer (e.g., URL) and identification of authentication method that for queries
    - alternative charging methods (e.g., recharging)
    - secure method for accessing counters, configuration data
  - Other congestion measures (e.g., Shapley value)

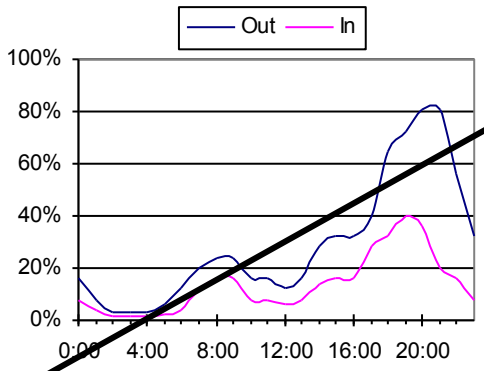
Feedback on Time of Day, Day of Week Charging

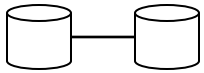
- Problem Summary
  - Congestion occurs when offered load approaches provisioned capacity, which occurs shortly before need to provision additional capacity.
    - Productive use of restoration capacity results in congestion occurring during peak periods AND failures,
    - Reserved restoration capacity produces congestion during all peak periods
  - Without Conex, peak utilization of 70-80% occurs typically without loss occurs at aggregate network bottlenecks
  - Assuming short term Conex achieves 90% utilization during peak periods, a gain of 10-20% appears feasible
  - If traffic increases ~75% per year then short term Conex defers marginal capacity provisioning by a small number of months
  - Looking over an entire day, typically 100-1000% unused capacity exists at network bottlenecks.
- Proposed Conex Use Case
  - Congestion exposure supporting incentives (e.g, pricing) motivating users/ content providers to time shift traffic to off-peak periods can defer need to provision marginal capacity by potentially years
    - Authenticated feed forward information could increment different counters
  - Use of only historical traffic patterns insufficient since exceptional events do occur, and longer term congestion exposure useful to handle these cases.

Traffic Growth and Capacity Provisioning Example

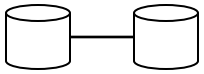




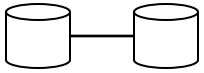




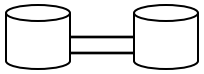
1Qx0



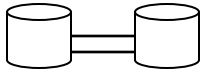
1Qx1



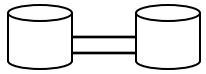
3Qx1



4Qx1

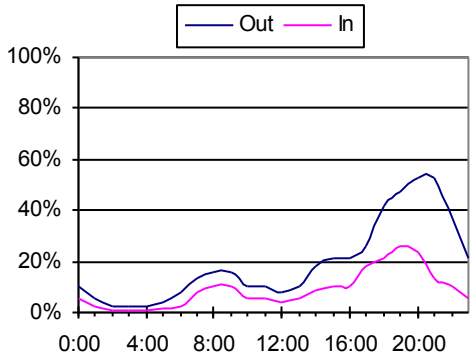


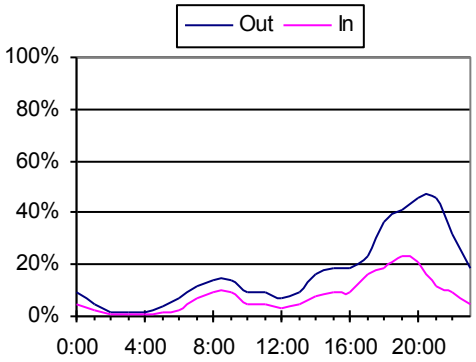
1Qx2

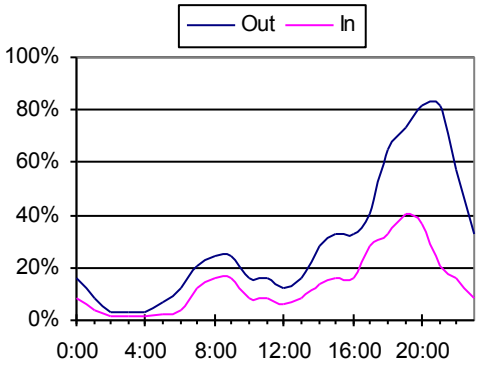


4Qx2

Short-term conex mechanism could defer marginal capacity provisioning approximately one Quarter (1Q) at these points

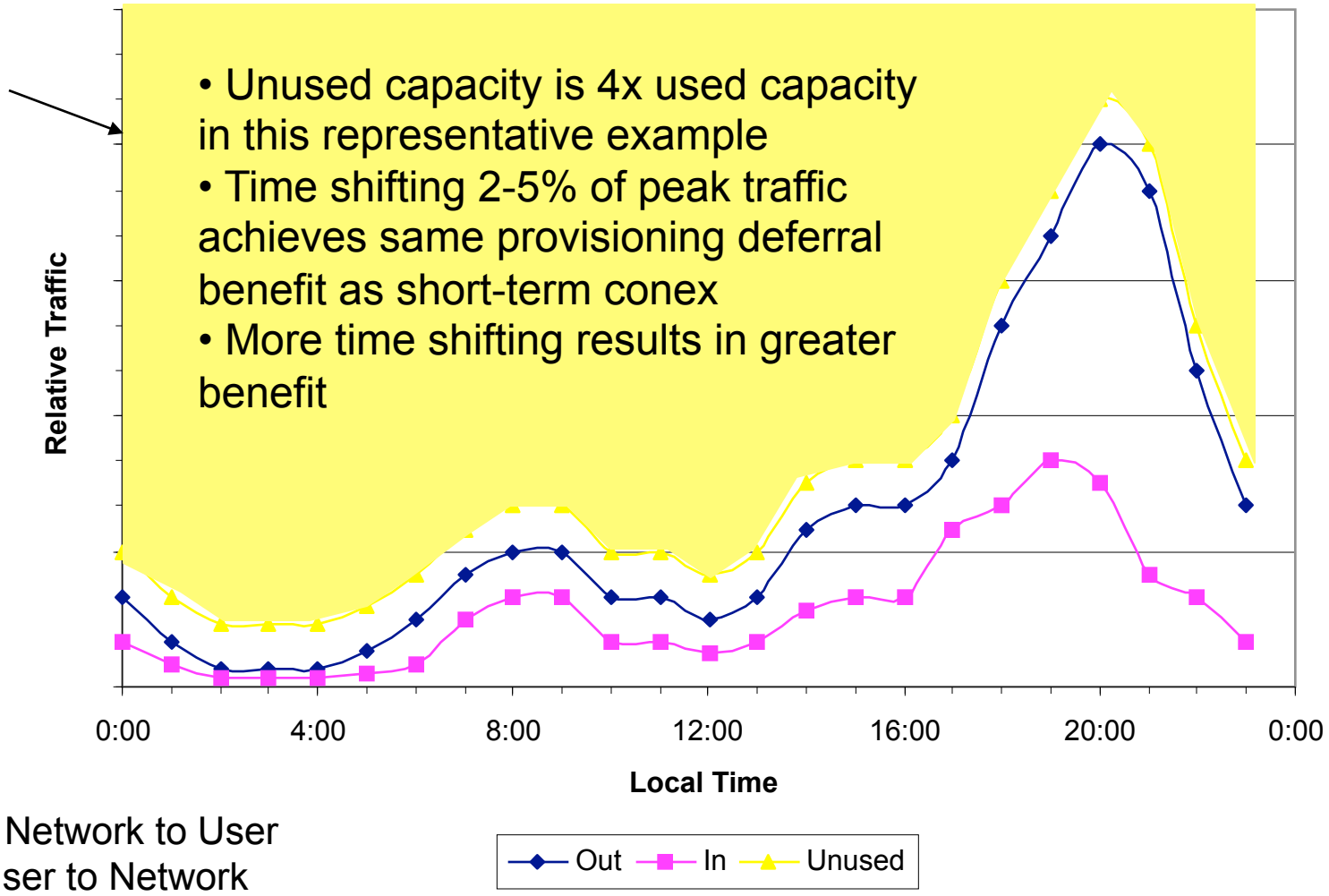






Example of Time Shifting Potential Reduction of Provisioned Capacity

80% link utilization



Out is Network to User
In is User to Network

Recharging for Implementing Congestion Pricing

- Problem Summary
  - Recharging (i.e., someone other than receiver pays) for usage causing congestion is an important incentive not currently covered in conex
  - Congestion can be for a shared, aggregate queue per current conex use case draft, and/or other congestible resources (e.g., burstiness measure, usage tier, time of day)
- Use Case
  - Augment TCP (and/or higher layer protocols) to feedback one or more congestion measures, e.g., short term but also with traffic profile, usage tier, and/or Time of Day (or a pointer to this information)
  - Include information in forward direction so that IP devices react appropriately (e.g., increment different counters)
  - Authentication method needed to valid third party charging, prevent spoofing

Inequity of Heavy versus Light Users

- Problem Summary
  - In some networks 20% of users are Heavy and generate 80% of the traffic
  - In bandwidth tiered network, remaining 80% of users are light but charged same as heavy users in same tier
    - Bandwidth tiers often implemented using a hierarchical scheduler, with the outermost scheduler being a non-work conserving shaper (or policer)
      - See DSL Forum/ BBF TR-059 for an example
  - Access network engineered for peak period and when near capacity provisioning upgrade event, congestion can occur
    - During these time heavy users create much more congestion volume (e.g., 16x) as compared with light users
- Proposed Conex-Based Use Case
  - Integrate (i.e., average) conex short term measurement over a longer time period
    - See draft for method proposed by Toby Moncaster
  - Could be used as means to invoke different forms of policing/ shaping, input to traffic engineering, and/or alternative method for incentives

Conclusions & Recommendations

- Short term congestion of a shared queue serving aggregate traffic is not the only congestible resource in some service provider networks
- Better methods to allow users, service and network providers to address congestion are needed
- Longer-term congestion feed forward/back mechanisms easier to implement experimentally (i.e., software) for research as compared with per-packet short term conex (i.e., hardware)
- Define potential use cases of research interest, define experimental code points, write some code and do some experimentation!