One More Bit Is Enough

Yong Xia, Shivkumar Kalayanaraman,
Rensselaer Polytechnic Institute
Lakshminarayanan Subramanian, Ion Stoica,
University of California, Berkeley
SIGCOMM’05

Presenter: Max Podlesny
Discussion leader: Ben Wun

Forecast

- Designing and implementing a simple, low-complexity congestion control protocol
Drawbacks of Congestion Control Schemas

- End-to-end and explicit feedback based schemas
- Degree of congestion is not revealed in end-to-end schemas (TCP, TCP+AQM/ECN)
- Delay-based methods are sensitive to minor delay variations
- Hard to deploy in Internet schemes using explicit rate feedback
  - A lot of bits to encode the congestion-related information (XCP)
  - IP headers do not have those bits

Problem statement

- Designing a congestion control protocol having:
  - high utilization (comparable to XCP)
  - reasonable fairness
  - ability to be deployed in current Internet
  - low persistent queue length
  - negligible packet loss rate
Main ideas

- Decouple efficiency control and fairness control
  » Use of different modes for achieving efficiency and fairness
- Use link load factor as the congestion signal
  » Load factor = users’ demand / link capacity
  » Switching controller
  » Congestion degree indicator

How VCP works

- Checking load factor at each step
- Possible states:
  » Low load ([0;80%])
  » High load ([80%;100%])
  » Overload ([100%; ∞])
- Codes of states
  » Low load: (01)
  » High load (10)
  » Overload (11)
Load factor transition point

- Requirements for transition point:
  - Achieving high utilization
  - Multiplicative Decrease should move the system to the high-load state
  - If being in low-load state one Multiplicative Increase step should move the system to the high-load state
- $\beta > 0.95$ induces 14 RTTs to halve cwnd
- $\beta = 0.5$ induces reduction of network utilization
- $\beta = 0.875$ satisfies the requirements
- Load factor transition point is 80%

Load factor estimation

- Tradeoff between two requirements
  - Monitoring the reaction on feedback
  - Avoidance queue buildup
- 75%~90% of flows have RTTs < 200ms
- Period of estimations is 200 ms
- Calculation of load factor:

$$\rho_l = \frac{\lambda_l + \kappa_q \cdot \tilde{q}_l}{\gamma_l \cdot C_i \cdot t_\rho}$$
Congestion control adjustments

- AI: \( \text{cwnd}(t+\text{rtt}) = \text{cwnd}(t)(1+\xi) \)
- MI: \( \text{cwnd}(t+\text{rtt}) = \text{cwnd}(t) + \mathcal{H} \)
- MD: \( \text{cwnd}(t+\delta t) = \text{cwnd}(t)\beta \)
- \( \text{rtt} = t_\theta \)

Congestion control parameters

- RTTs are equal for all flows
  - AI: \( \mathcal{H} = 1 \)
  - MI: \( \mathcal{H}(\rho) = \kappa(1-\rho)/\rho \)
    - No information about exact value of \( \rho \)
    - \( \mathcal{H}(\rho) \) is minimum at \( \rho = 80\% \)
    - \( \mathcal{H}(0.8) = 0.0625 \)
    - Stability of the algorithms requires that \( \kappa = 0.25 \)
  - MD: \( \beta = 0.875 \)
- RTTs are heterogeneous
  - AI: \( \mathcal{H}_s = \mathcal{H}\text{rtt}/t_\theta \)
  - MI: \( \mathcal{H}_s = (1 + \mathcal{H})^{\text{RTT}/t_\theta} - 1 \)
  - MD: \( \beta = 0.875 \) as it is performed once per period of load factor estimation in case of congestion
  - Scaling for fairness:
    - \( \mathcal{H}_{\text{rate}} = \mathcal{H}(\text{rtt}/t_\theta)^2 \)
VCP parameter setting

<table>
<thead>
<tr>
<th>Para</th>
<th>Value</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>$t_p$</td>
<td>200 ms</td>
<td>the link load factor measurement interval</td>
</tr>
<tr>
<td>$t_q$</td>
<td>10 ms</td>
<td>the link queue sampling interval</td>
</tr>
<tr>
<td>$\gamma_t$</td>
<td>0.98</td>
<td>the link target utilization</td>
</tr>
<tr>
<td>$\kappa_q$</td>
<td>0.5</td>
<td>how fast to drain the link steady queue</td>
</tr>
<tr>
<td>$\kappa$</td>
<td>0.25</td>
<td>how fast to probe the available bw (MI)</td>
</tr>
<tr>
<td>$\alpha$</td>
<td>1.0</td>
<td>the AI parameter</td>
</tr>
<tr>
<td>$\beta$</td>
<td>0.875</td>
<td>the MD parameter</td>
</tr>
</tbody>
</table>

Analysis of VCP Model Stability

- $\kappa \leq 0.5$ makes sure asymptotic stability of user’s rates
- High utilization, fairness, zero steady-state queue length and zero packet loss rate are achieved at the equilibrium
Simulations parameters

- ns-2
- Link capacities: [100Kbps; 5Gps]
- RTTs: [1ms; 1.5s]
- Numbers of long-lived, FTP-like flows: [1; 1000]
- Arrival rates of short-lived, web-like flows: [1per s; 1000per s]
- Data packet size: 1000 bytes
- ACK packet size: 40 bytes
- Time of simulations: no less than 120s
- Utilization and throughput are averaged over 500ms
- Queue length and cwnd are sampled every 10ms

Results: one bottleneck

- Utilization is at least 93%
- Utilization gap is 7%
- Very low capacities (100Kbps) induce bottleneck average queue increase to 50% of the buffer size
Results: one bottleneck

- Utilization is 85%~94% for RTTs > 800ms
- The average queue is < 5% of the buffer size
- The average queue is up to 15% for very low RTT (1ms)

Results: one bottleneck

- High utilization
- The average queue is < 5% of the buffer size
Results: multiple bottlenecks

- Average utilization is 94%
- The average queue is < 0.2% of the buffer size
- Zero packet drops

Results: Fairness

- VCP reveals good fairness but its fairness converges significantly longer than XCP
Conclusion

- Development of VCP
  - Simplicity, low-complexity, usage for BDP networks
  - High utilization, reasonable fairness, low persistent bottle-neck queue, negligible packet loss rate
  - Two bits to encode the network congestion information, i.e. no changes of the IP header

Questions