Performance Evaluation of TCP Vegas over Optical Burst Switched Networks

Basem Shihada and Pin-Han Ho
University of Waterloo

Qiong Zhang
Arizona State University

The 6th International Workshop on Optical Burst/Packet Switching (WOBS), Oct. 2nd 2006
Outline

- Introduction
- TCP Vegas over OBS
- TCP Vegas Performance Modeling
  - over Barebone OBS
  - over OBS with Burst Retransmission
- Simulation and analytical result comparison
- Conclusion
Introduction

• TCP
  – Many Internet applications depend on TCP for reliable data transfer
  – Loss-based TCP (e.g. Reno, Sack) and delay-based TCP (e.g. Vegas)
  – TCP Vegas improves throughput 37% to 71% compared to TCP Reno

• OBS
  – Burst contention occurs even at low traffic loads
    • Bufferless
    • One-way signaling
  – Contention resolution schemes
    • Burst deflection
    • Burst retransmission
    • Reduce burst loss probability
    • Introduce additional delay
TCP Vegas Congestion Avoidance

1. Compute the \textit{Expected} throughput

\[
\text{Expected} = \frac{cwnd}{\text{BaseRTT}}
\]

(\textit{BaseRTT} is the minimum measured RTT)

2. Compute the current \textit{Actual} throughput

\[
\text{Actual} = \frac{cwnd}{RTT}
\]

3. Compute the \textit{Diff}=\textit{Expected} – \textit{Actual}

\[
\text{Diff} = cwnd(1 - \frac{\text{BaseRTT}}{RTT})
\]

4. Adjust the next \textit{cwnd} size as follows,

\[
cwnd = \begin{cases} 
  cwnd + 1 & \text{diff} < \alpha \\
  cwnd & \alpha \leq \text{diff} \leq \beta. \\
  cwnd - 1 & \text{diff} > \beta
\end{cases}
\]
TCP Vegas over OBS

- Vegas can NOT detect congestion in a barebone OBS network
  - *RTT* in the barebone OBS network varies little!!!

**What happens for Vegas over an OBS network with burst retransmission or burst deflection?**

- Case 1: low traffic loads
  - Burst retransmission and burst deflection introduce lower burst loss, but higher burst delay
  - Vegas detects RTT increase due to retransmission or deflection
  - Results in unnecessary reduction of the *cwnd* size

- Case 2: high traffic loads
  - Higher congestion in OBS results in even higher burst delay
  - A proper delay-based TCP congestion avoidance mechanism can take advantage of this observation
Performance Modeling

Goal: to understand the behavior of delay-based TCP over OBS networks, so as to design a proper TCP congestion avoidance mechanism for OBS networks

Assumption:

- High access bandwidth: TCP packets in a single cwnd are assembled to a single burst (i.e. TCP fast flow)
- IP access networks are not congested
- Burst loss probability and burst contention probability in OBS networks are given
TCP Vegas over an Barebone OBS

Slow-start-to-slow-start period (SS2SS):
- A to B (slow start), starts at 2 and doubles every other round
- B to C (transition period), increases by 1 every RTT due to $\text{Diff} < \alpha$
- C to D (loss free period), remains unchanged due to $\alpha \leq \text{Diff} \leq \beta$
- TO period due to a burst loss

Average RTT is very close to $\text{BaseRTT}$
TCP Vegas over OBS with Burst Retransmission

SS2SS period:
- A to B (slow start)
- B to C (transition period)
- C to G (loss free period): the size of $cwnd$ is modeled as Markov Chain
- TO period

The $cwnd$ that is retransmitted in the OBS layer has longer delay than BaseRTT
TCP Vegas over OBS with Burst Retransmission (cont)

- The probability of a successful round in which a burst experiences contention but successfully retransmitted

\[ p_{sr} = \frac{p_c - p}{1 - p} \]

- The probability of a successful round that does not experience burst contention

\[ p_{nc} = \frac{1 - p_c}{1 - p} \]
Vegas Throughput over Barebone OBS

![Graph showing Vegas Throughput over Barebone OBS](image-url)
Vegas Throughput over OBS with Burst Retransmission
Conclusion

• Modeled delay-based TCP Vegas over
  – A barebone OBS network
  – An OBS network with burst retransmission

• Future Work
  To design a delay-based TCP congestion avoidance mechanism suitable for OBS networks
TCP Vegas Congestion Control

- **Slow Start**
  - *The cwnd size exponentially increases every other round*
  - Vegas exits the slow start if the *diff* exceeds a threshold $\gamma$

- **Packet Retransmission**
  - Vegas retransmits the packet when duplicate *ACKs are received* or when a sending packet timeouts.
<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p$</td>
<td>Burst dropping probability</td>
</tr>
<tr>
<td>$p_c$</td>
<td>Burst contention probability</td>
</tr>
<tr>
<td>$p_{nc}$</td>
<td>Probability of no burst contention</td>
</tr>
<tr>
<td>$p_{sr}$</td>
<td>Probability of a burst contended but successfully retransmitted through burst retransmission</td>
</tr>
<tr>
<td>$\alpha$, $\beta$</td>
<td>Vegas throughput thresholds in packets</td>
</tr>
<tr>
<td>$\text{BaseRTT}$</td>
<td>Minimum measured RTT</td>
</tr>
<tr>
<td>$R$</td>
<td>Expected round trip time without burst retransmission</td>
</tr>
<tr>
<td>$\text{RTT}_r$</td>
<td>Expected round trip time with burst retransmission</td>
</tr>
<tr>
<td>$B$</td>
<td>Vegas throughput</td>
</tr>
<tr>
<td>$H$</td>
<td>Expected number of packets submitted during timeout (TO)</td>
</tr>
<tr>
<td>$\text{RTO}$</td>
<td>TCP retransmission timeout</td>
</tr>
<tr>
<td>$\text{TOP}$</td>
<td>Duration of a sequence of TOs</td>
</tr>
<tr>
<td>$X$</td>
<td>Number of consecutive successful rounds</td>
</tr>
<tr>
<td>$A$</td>
<td>Duration of a sequence of consecutive successful rounds</td>
</tr>
<tr>
<td>$Y$</td>
<td>Number of packets sent before a timeout expiration</td>
</tr>
<tr>
<td>$W_{\text{max}}$</td>
<td>Maximum congestion window size in packets</td>
</tr>
<tr>
<td>$W_0$</td>
<td>Cwnd size in Vegas stable state</td>
</tr>
</tbody>
</table>
$W_0$ and $W_0'$

- $W_0$ - the cwnd size when $Diff$ reaches $\alpha$ from a smaller value

$$\alpha \leq Diff = W_0 \left(1 - \frac{\text{BaseRTT}}{R}\right) \leq \beta$$

$$\frac{\alpha R}{R - \text{BaseRTT}} \leq W_0 \leq \frac{\beta R}{R - \text{BaseRTT}}$$

$$W_0 = \min \left( \frac{\alpha + \beta}{2} \times \frac{R}{R - \text{BaseRTT}}, W_{\text{max}} \right)$$

- $W_0'$ - the cwnd size when $Diff$ reaches $\beta$ from a larger value

$$W_0'(1 - \frac{\text{BaseRTT}}{\text{RTT}_r}) \leq \beta$$

$$W_0' = \left[ \frac{\beta \text{RTT}_r}{\text{RTT}_r - \text{BaseRTT}} \right]$$
Analysis for TCP Vegas over Barebone OBS

- **Number of packets sent from A to B**
  \[ Y_{AB} = 2 \sum_{i=0}^{\log_4 W_0} 2^i = 2^{\log W_0} - 4 \]

- **The duration from A to B**
  \[ A_{AB} = 2\left(\frac{\log W_0}{4}\right)R = 2(\log W_0 - 2)R \]

- **Number of packets sent from B to C**
  \[ Y_{BC} = \sum_{i=W_0/2}^{W_0-1} i = \frac{3W_0^2}{8} - \frac{W_0}{4} \]

- **The duration from B to C**
  \[ A_{BC} = \left(\frac{W_0}{2}\right)R \]
Analysis for TCP Vegas over Barebone OBS

- **Number of rounds in C to D**
  \[
  S_{lossfree} = \begin{cases} 
  0, & \text{if} \ (1 - \frac{p}{p} < 2(\log W_0 - 2) - \frac{W_0}{2}), \\
  \frac{1 - p}{p} - 2(\log W_0 - 2) - \frac{W_0}{2}, & \text{otherwise}.
  \end{cases}
  \]

- **Duration from C to D is**
  \[A_{CD} = RS_{lossfree}\]

- **Number of packets sent from C to D is**
  \[Y_{CD} = W_0 S_{lossfree}\]

- **Vegas TOP is similar to Reno**

- **Vegas throughput over barebone OBS is**
  \[
  B_{barebone} = \frac{Y_{AB} + Y_{BC} + Y_{CD} + E[H]}{A_{AB} + A_{BC} + A_{CD} + TOP} \\
  E[H] = \frac{p}{1 - p}, \quad E[TOP] = RTO \frac{f(p)}{1 - p}
  \]
Analysis for TCP Vegas over OBS with Burst Retransmission

- **Average cwnd size from C to G**
  \[
  E[W_{CG}] = W_0\pi_0 + (W_0 - 1)\pi_1 + \ldots + W_0\pi_{(W_0 - W_0')}
  \]
  \[
  = \sum_{i=0}^{W_0-W_0'} [(W_0 - i)\pi_i]
  \]

- **Duration from C to D is**
  \[
  A_{CG} = (p_{nc}R + p_{sr}RTT_r)S_{\text{lossfree}}
  \]

- **Number of packets sent from C to D is**
  \[
  Y_{CG} = E[W_{CG}]S_{\text{lossfree}}
  \]

- **Vegas throughput over OBS with burst retransmission is**
  \[
  B_{ret} = \frac{Y_{AB} + Y_{BC} + Y_{CG} + E[H]}{A_{AB} + A_{BC} + A_{CG} + TOP}
  \]