

Copyright(C)2002 Youki Kadohaveshi. All rights reserved.



ネットワーク資源の競合と調停

- フロー制御 (flow control)
 - 速度の違いを吸収
 - 配送順序の入れ替わりからの回復
 - 重複、パケット廃棄、ビット誤りからの回復
- 輻輳制御 (congestion control)
 - 輻輳を抑えつつ帯域を共有

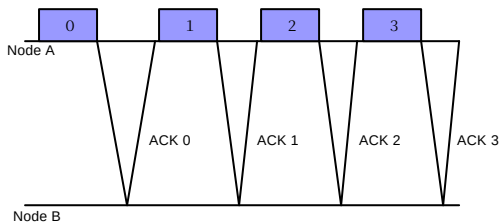
Copyright(C)2002 Yuuki Kadibayashi. All rights reserved.

フロー制御

- Stop-and-Wait
- Go Back n (a.k.a. sliding window)
- さまざまな方式が存在
 - ネットワーク側が速度等を申告するもの
 - ホスト側で速度を推定しながら動くもの
- ホスト側で自律的に動く方式はネットワーク側での状態管理を必要としないため、より大規模な運用が可能
- Autonomous → Less state management → Scalable

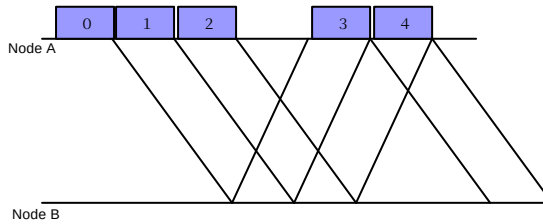
Copyright(C)2002 Yuuki Kadibayashi. All rights reserved.

Stop-and-Wait



Copyright(C)2002 Yuuki Kadibayashi. All rights reserved.

Go Back n



Copyright(C)2002 Yusaki Kadibayashi. All rights reserved.

TCP におけるフロー制御

- sliding window
- window size
- パケット番号ではなくシーケンス番号で
- End to end

Copyright(C)2002 Yusaki Kadibayashi. All rights reserved.

TCP における輻輳制御

- Fair-share model
 - End to end
- window size の増減
 - additive increase
 - multiplicative decrease
- window size 増加アルゴリズムの切り替え
- packet drop で輻輳が起きたことを知る

Copyright(C)2002 Yusaki Kadibayashi. All rights reserved.

Window size の増やし方

- slow start threshold (sssthresh) を基準としてcongestion window (cwnd) の増やし方を変える
- (アルゴリズムの概要)
- On receiving an ACK:
 - if (cwnd < sssthresh) {
 - /* slow start */
 - cwnd += 1;
 - } else {
 - /* congestion avoidance */
 - cwnd += 1 / cwnd;
- }

Copyright(C)2002 Yosuke Kadobayashi. All rights reserved.

Window size の増やし方

- Slow startの時
 - ☐ exponential increase
- Congestion avoidanceの時
 - ☐ additive increase
- (Illustrated p. 311)

Copyright(C)2002 Yosuke Kadobayashi. All rights reserved.

Window size の減らし方

- (アルゴリズムの概要)
- On detecting packet drop:
 - sssthresh = cwnd / 2;
 - if (timeout) {
 - cwnd = 1;
- }
- (Illustrated p. 315)

Copyright(C)2002 Yosuke Kadobayashi. All rights reserved.

Packet drop

- 2つの方法で検出
- Duplicate ACK
- Retransmission Time Out (RTO)
- タイムアウトはどのように判定すればよいか？
 - RTO ~ RTT Estimator

Copyright(C)2002 Yusuki Kadobayashi. All rights reserved.

RTO Calculation

- $Err = M - A$
 - $A \leftarrow A + gErr$
 - $D \leftarrow D + h(|Err| - D)$
 - $RTO = A + 4D$
 - ☐ A: smoothed RTT
 - ☐ D: smoothed mean deviation
 - ☐ g: gain for the average (1/8)
 - ☐ h: gain for the deviation (1/4)
- (Illustrated p. 304)

Copyright(C)2002 Yusuki Kadobayashi. All rights reserved.

More topics on TCP

- Fast retransmit
- Fast recovery
- Selective Acknowledgment
- Explicit Congestion Notification
- Interaction with RED
- TCP extensions for wireless links
- ...

Copyright(C)2002 Yusuki Kadobayashi. All rights reserved.

TCP enhancements

- 細かな実装の違いが存在
 - ☐ Reno
 - ☐ NewReno
 - ☐ SACK
 - ☐ ...
- ベンダ独自の工夫が裏目に出ることも

Copyright(C)2002 Yuuki Kadibayashi. All rights reserved.

まとめ

- フロー制御
 - ☐ Stop-and-Wait
 - ☐ Go back n
- TCPにおけるフロー制御
 - ☐ Sliding window
- TCPにおける輻輳制御
 - ☐ Slow start
 - ☐ Congestion avoidance

Copyright(C)2002 Yuuki Kadibayashi. All rights reserved.

課題

- tcpdump, tcpillust, あるいは tcptrace を使って以下のディレクトリにある2つのトレースを比較せよ。違いは何か。同じ点は何か。
- /usr/local/share/tcpillust/Infomation_Network_t-kadai01 (mint のみ)
- 学籍番号 % 4 == 0 の学生: AIT, AIT2
- 学籍番号 % 4 == 1 の学生: ITB, ITB2
- 学籍番号 % 4 == 2 の学生: MIT, MIT2
- 学籍番号 % 4 == 3 の学生: WIDE, WIDE2

Copyright(C)2002 Yuuki Kadibayashi. All rights reserved.

tcpdump

- `cd /usr/local/share/tcpillust/Infomation_Network_I-kadai01`
- `tcpdump -r AIT`
- 09:45:10.547843 mint111.aist-nara.ac.jp.1195 > www.cs.ait.ac.th.http: S 1442476782:1442476782(0) win 16384 <mss 1460> (DF)
- 09:45:11.116022 www.cs.ait.ac.th.http > mint111.aist-nara.ac.jp.1195: S 3854747835:3854747835(0) ack 1442476783 win 17520 <mss 1460> (DF)
- ...

Copyright(C)2002 Youki Kadibayashi. All rights reserved.

tcpillust

- `cd /usr/local/share/tcpillust/Infomation_Network_I-kadai01`
- `tcpillust AIT`

Copyright(C)2002 Youki Kadibayashi. All rights reserved.

tcptrace

- `mkdir ~/infonet1; cd ~/infonet1` 等
- `ln -s /usr/local/share/tcpillust/Infomation_Network_I-kadai01/* .`
- `tcptrace -G AIT`
- `ls *.xpl`
- `a2b_owin.xpl a2b_ssize.xpl a2b_tsg.xpl b2a_rtt.xpl b2a_tput.xpl`
- `a2b_rtt.xpl a2b_tput.xpl b2a_owin.xpl b2a_ssize.xpl b2a_tsg.xpl`
- `xplot a2b_rtt.xpl`
- グラフの読み方は <http://www.tcptrace.org/>

Copyright(C)2002 Youki Kadibayashi. All rights reserved.
