Outline

• MaxNet as an alternative to TCP
• Linux implementation of MaxNet
• Demonstration of fairness, quick convergence, etc.
• Incorporating ADPM with MaxNet
Shortcomings of TCP

Sensitivity to losses

Round-trip time (RTT) unfairness

\[
\frac{rate_2}{rate_1} = \left( \frac{RTT_1}{RTT_2} \right)^2
\]

Slow start-up times for small flows
MaxNet Overview

Transmission rate (congestion window)

Price

Echo price

Sender

Router

Receiver

Adjust window
MaxNet Benefits from Explicit Feedback

• Avoid queueing, loss instead of reacting to it
  ▫ Avoid hindering flows
• Flows achieve max-min fairness
  ▫ Maximize the minimum transfer rate
  ▫ Without affecting minimum, maximize second minimum rate, repeat.

Max-Min Fairness
Other Protocols using Explicit Feedback

- **XCP** (Katabi, Handley, Rohrs 2002)
  - At capacity, new flow bandwidth allocation too slow
- **RCP** (Dukkipatti, Kobayashi, Zhan-Shen, Mckeown 2005)
  - Signals RTT for more responsiveness
- **REM, JetMax, RED/ECN, many others**
TC on Routers to Implement MaxNet

Packet Arrival

Packet Parsed with MaxNet field

TC Hook

TC calls MaxNet code

Packet now has price data

Price Update Calculation

• Update based on rate of traffic since last update

Packet Price Update

• Only if router price > packet price
Choice of Host Kernel

- Linux kernel version 2.6.23.1
  - Benefits from improvements in SACK (selective acknowledgements)
- Modular Framework
  - Consolidates much of MaxNet code
Linux Modular Framework

Hooks from Modular Framework

- **Cwnd_event** – Extracts MaxNet Option
- **Cong_avoid** – Changes congestion window per algorithm
- **Ssthresh** – Sets threshold for slowstart
- Many Others...

Linux Kernel Modular Networking Framework
TCP Header

- TCP Header
- SACK Header
- SACK 1
- SACK 2
- MaxNet Option

MaxNet Option Format

MaxNet Option (42)
1 byte

Option Size (6)
1 byte

Echo Price (3 bytes)

Price (3 bytes)
Sender Targets for Equilibrium

- Equilibrium Rate (log scale for constant relative precision)
  \[ x_{\text{max}} e^{-\text{price}/T} \]
- Variable \( \xi \) for stability
- Host keeps track of variables: \( \xi \), price
  - Values in fixed point format
Sender Acknowledgement Arrival

ACK Arrival

Packet Parsed with MaxNet option

cong_avoid hook to MaxNet code

Price Updated
- Take price from ACK

Congestion window Update
- Recalculate window based on new price

Next Packet Transmitted
Receiver Data Arrival

Packet Arrival

Packet Parsed with MaxNet option

Echo price in ACK Packet

ACK Transmitted

Solution: Have MaxNet treat delayed ACKs as increase in RTT, not congestion
Experiment Topology
Experiment Topology

Legend:
- Hardware Router
- Software Router
- Server
Experiment Topology

Legend:
- Hardware Router
- Software Router
- Server
Experiment Topology

Legend:

- □ Hardware Router
- □ Software Router
- ○ Server
Example Data - Fairness and Quick Convergence
Bursty Data Flows

- Data transmitted/received in bursts
- Router price and CPU load increase temporarily
- Results in slower transfer rates
- TCP Pacing as a solution
Future enhancements

• “Incompatibility” with TCP
  ▫ Greedy TCP will hog bandwidth
  ▫ MaxNet backs off and stops sending

• Encrypted packets
  ▫ MaxNet option in TCP header – routers cannot use
Advanced Deterministic Packet Marking (ADPM)

- Single bit (ECN) in IP header marks packets
- Determine bit value with IPid field hash
- More packets passed yield closer estimate to actual price
ADPM

• Benefits
  ▫ Enables MaxNet to work with encrypted packets
  ▫ Routers should not look inside TCP layer

• Drawbacks
  ▫ Decreases accuracy of congestion level estimation
Summary

• MaxNet: Router assisted scheme
  ▫ Avoids some pitfalls of TCP
• MaxNet ported to 2.6.23.1
  ▫ Modular framework
  ▫ Improved SACK handling
• Address packet encryption with ADPM