On Congestion Management

Analysis, history, and viable options

Mick Seaman
mick_seaman@ieee.org

On Congestion Management

- Why congestion management ?
- Dimensional analysis
- Alternative approaches
- Why not flow control
- Two viable options
- Summary

Why congestion management

- The network problem
- What would we like to control ?
- Network power

Why – the network problem

- Dropping frames radically impacts performance
- Queuing lots of frames leads to long response times and too many jobs started
- Long queues rarely improve throughput much

NOTE– some may be under a misapprehension that TCP and similar protocols only control frame transmission by noticing drops, that's not true. Delay (can) play(s) a useful part through window signaling. This allows control to come into play before dropping occurs.

What would we like to control ?

Given some traffic arrival statistic network (and as a consequence system) throughput is a function of queue depths in bridges.

- e.g. plot average queue depth against average throughput for Poisson arrivals
  - Mean queue depth 1, 66% utilization
  - Mean queue depth 10, >90% utilization
  - 99.8% confidence depths for these ~12, ~35
- All well known queuing theory, though needs traffic stats

Network power

- Minimize delay (queue depth)
- Maximize bandwidth
- Network power

  = bandwidth (bits/s) / delay (secs)

  possibly with adjustment for higher utilization
- Averages and variances to consider/optimize

Dimensional analysis

- Applicability of dimensional analysis
- Similar problems

Dimensional analysis - applicability

- Beyond checking units !
- Establish relevance of models and comparisons
  - Identify the dimensionless parameters of the underlying physics
- Present case
  - Buffered channel with interfering traffic
  - Parameter is min number of frames end to end
  - Determines nature of control theory problem

All the same problem

- 10 Gb/s, 1500 byte frames, 252 meters
- 1 Gb/s, 2.52 kilometers
- Frame relay at T1/E1 3,000 kilometers
- TCP/IP, 512 byte frames, 10 Mb/s, 2,000 km

Store and forward buffers dominate

Over 20 years intense study of these problems

Alternate approaches

- Control the number of buffers in use
  - Send flow control messages to source
  - Send flow control messages to previous hop
  - Signal forward to destination, then controls src
    - » Congestion experienced signal in forwarded frame
- Control the delay or jitter in the delay
    - » Mark relative frame timings on transmit
      (can be very crude)
      signal back to source

Equivalent approaches

Why not flow control (again)

- Forward and return delay not great contribution to delay compared to burst buffering
- Need to insert extra frames (into potential congestion)
- Hop by hop penalizes non-congested flows
- ‘Stuck’ information in changing networks
- Forward congestion signaling already in IP
  - Proven technology, dimensionally equiv. Problem
- Delay measuring ‘shim’ an alternate for protocols with no defined response to forward congestion signal

Two viable alternatives

- Forward congestion signaling already in IP
  - Proven technology, dimensionally equiv. problem
  - Already in some switches
- Delay measuring/signaling ‘shim’
  - Can be implemented in end stations without changing the switch/bridge at all
  - For protocols with no defined response to forward congestion signal

Summary

- Yet another flow control proposal ?
- Repeat questions on simulation reality
- Very extensively studied problem
- One well known solution
- One viable bridge independent alternate
  - Known control theory problem and solutions
- If it's IP use "congestion experienced"
- Protocol and bridge independent, use end station shim