Delay and Delay Variation Simulation Results for Multi-hop Conventional Ethernet Cases with Bursting/Bunching

Geoffrey M. Garner
gmgarner@comcast.net

SAMSUNG Electronics

IEEE 802.3 ResE SG
2005.08.19

Outline

- Introduction
- Simulation models and assumptions
- Case 1
  - ***Add additional scenarios as needed***
  - Results
- ***Add additional cases as needed***
- Conclusions

Introduction

- Reference [1] presented initial simulation results for transport of time-sensitive traffic over conventional Ethernet
  - Considered simple, one-hop, two-switch network
    - one switch-to-switch link with end-devices attached to each switch
- Results showed that unfiltered peak-to-peak delay variation for competing Constant Bit Rate (CBR) traffic streams whose rates differ slightly from nominal can be appreciable compared to the requirements for digital audio and video [2]
  - For 3 CBR streams, 50% link utilization, and 256 byte packets, unfiltered delay variation was almost 50 μs in one case
    - For sufficiently small frequency offsets (e.g., 1 ppm or less), phase-locked loop (PLL) filtering at the egress did not reduce the delay variation appreciably
  - This exceeds the requirements for uncompressed digital video and digital audio, and is close to the limit for compressed digital video (50 μs) [2]
  - For 6 CBR streams, 50 % link utilization, and 256 byte packets, unfiltered delay variation exceeded 100 μs in one case, and was reduced to just over 80 μs by filtering
    - This exceeds the limits for digital audio and video
  - Also considered adding a best-effort stream with maximum size packets, though this did not appreciably change the time-sensitive stream results

Introduction (Cont.)

- Discussion during the presentation of [1] indicated it would be of interest to show a worse case, with multiple hops
- Further discussion in a subsequent ResE SG conference call indicated it would be of interest for the multiple hop case to resemble the bursting/bunching scenarios described in [3] (see Annex F of [3] for details)
  - Combine N traffic sources from N locally-attached end devices at a switch, and transport over a link to a downstream switch
  - Replicate this configuration N times, so that the downstream switch has N incoming links
  - Drop the traffic from $N-1$ of the sources from each incoming link at the downstream switch to locally-attached end devices (the number of locally attached end devices is therefore $N(N-1)$)
  - Transport the remaining N streams (one from each incoming link) over an outgoing link to a downstream switch
  - Repeat the above scenario at the downstream switch, i.e., replicate the above configuration N times
  - For k stages, the total number of sources at the ingress grows like N^k

Introduction (Cont.)

- The analysis in [3] is mainly qualitative, i.e., is carried out by graphically representing packets at various times
  - The analysis assumes worst-case arrival patterns (i.e., packets from competing time-sensitive streams always arrive simultaneously)
- The analysis in [3] considers both queueing at the input of each switch and queueing at the output of each switch (these are separate cases)
- It was felt it would be desirable to simulate this scenario as a case that is possibly worse than those considered in [1]
- In addition, it is of interest to consider total delay for multi-hop cases
  - While total delay was not explicitly discussed in [1], end-to-end delays for a path consisting of the ingress link, single switch-to-switch link, and egress link were on the order of at most 300 μs, and were this large only for the case that included a single best-effort stream with maximum size packets
    - The longest path through the network was between 100 and 200 m, and the propagation delay was 1.755×10^8 m/s (the default minimum propagation speed in Opnet, which assumes dispersion representative of the medium and configuration
    - With this assumption, propagation delay is of the order a few μs, and is therefore negligible (and would be negligible with no dispersion)

Simulation Models and Assumptions

- As in [1], OPNET simulation tool was used to simulate packet delays
  - OPNET contains models for full-duplex Ethernet MAC and for Ethernet bridges
  - Models were modified (as indicated in [1]) to include priority classes
    - Priority queueing is non-preemptive
- Considered basic topology as described in [3] and summarized in Introduction
  - At each stage, combine N previous stages
    - Each previous stage supplies N traffic streams to this stage
    - Drop the traffic from $N - 1$ of the traffic streams from each previous stage
      - Therefore, need $N(N - 1)$ end devices at this stage
    - Carry one traffic stream from each incoming link (from each previous stage) over an egress link to the next stage
      - Therefore, N traffic streams are carried to the next stage
- Assume all packets are maximum size
  - 1500 bytes (Opnet adds Ethernet overhead)
- Assume 100 Mbit/s Ethernet links
- Assume the maximum path length through the network is 100 – 200 m, and the propagation speed is as in [1], i.e., 1.755×10^8 m/s

Simulation Models and Assumptions (Cont.)

- Assume all time sensitive traffic streams have the same nominal rate, with a small frequency offset
  - Frequency offset is different for each competing stream, and is chosen on input
    - This captures the fact that Time-sensitive video and audio clients have specified nominal rates, but are allowed to differ from those nominal rates by specified frequency tolerances
  - Nominal rate is chosen based on the number of streams per switch at the network ingress and desired link utilization
    - Input rate (and time between packets) is constant
- OPNET model assumptions (same as in [1])
  - Two priority classes
    - Time-sensitive traffic gets high priority
    - Best-effort traffic gets low priority
    - Priority queueing is non-preemptive
    - Queueing is first-come, first-served (FCFS) within each priority class
  - OPNET model for full-duplex Ethernet MAC is used (with priorities added)
  - OPNET contains spanning tree and rapid spanning tree algorithms
    - Use rapid spanning tree algorithm here

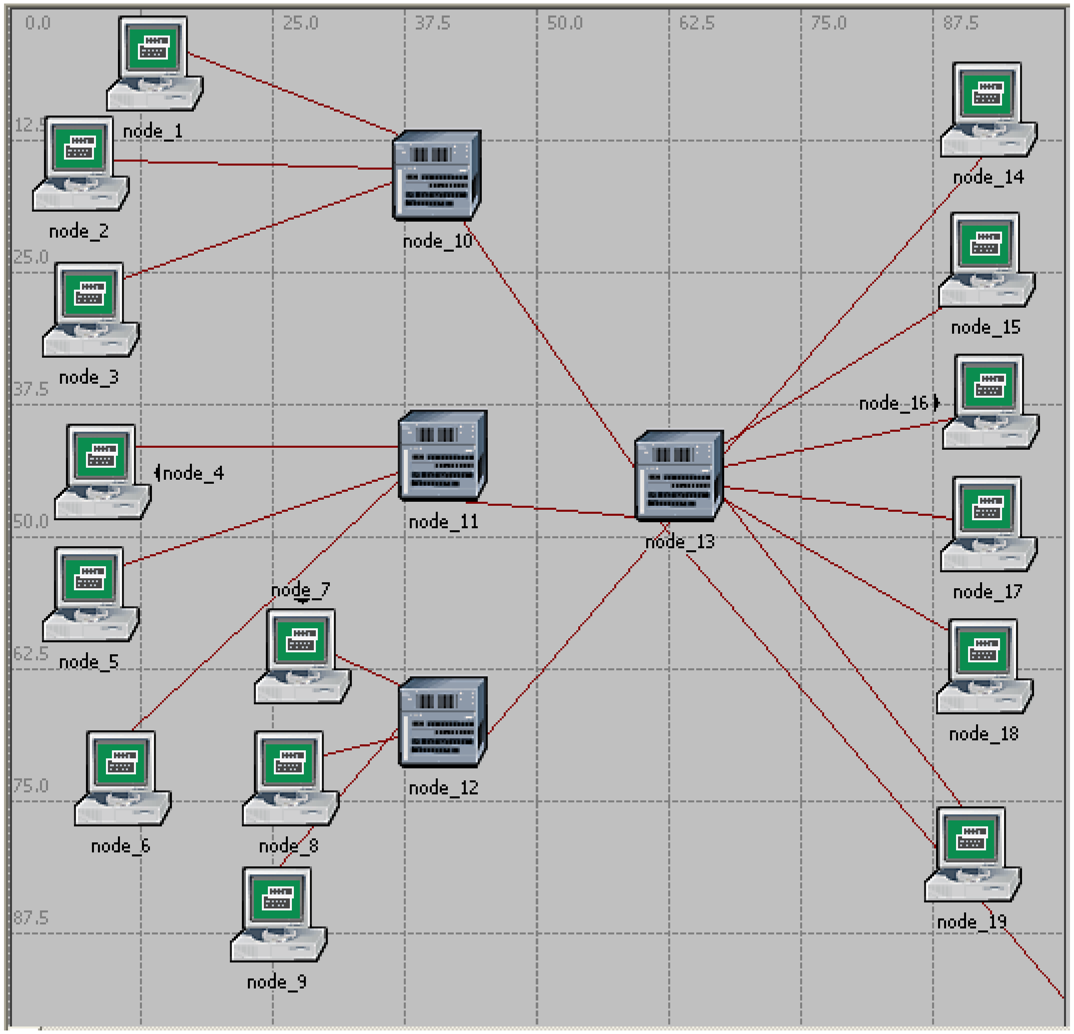
Simulation Case 1

- Three sources at the ingress of each switch at the initial stage ($N = 3$)
- 4 stages ($k = 4$), i.e., a traffic stream that is not dropped at an intermediate stage traverses 4 switch-to-switch links (plus one ingress and one egress link)
  - Total of 81 traffic sources
    - We needed to restrict the numbers of hops and/or sources/switch to keep the total number of sources manageable (e.g., $N = 3$ and $k = 7$ would have produced 2187 traffic sources; $N = 6$ and $k = 7$ would have produced 279936 traffic sources.)
- All traffic is time-sensitive
  - Packet size is as given above (1500 bytes plus Ethernet overhead added by Opnet)
  - Nominal packet arrival rate for each stream is 1333.33 packets/s
    - Nominal time between packets is 0.00075 s
    - Resulting switch-to-switch link utilization is approximately 50% (results from 3 traffic streams)
      - Utilization per stream assuming nominal arrival rate and excluding Ethernet overhead is 16%
- Network topology is shown on the next 3 slides
  - It was convenient to use the subnet capability of Opnet, due to the large number of traffic sources and hierarchical structure of the network

Simulation Case 1 (Cont.)

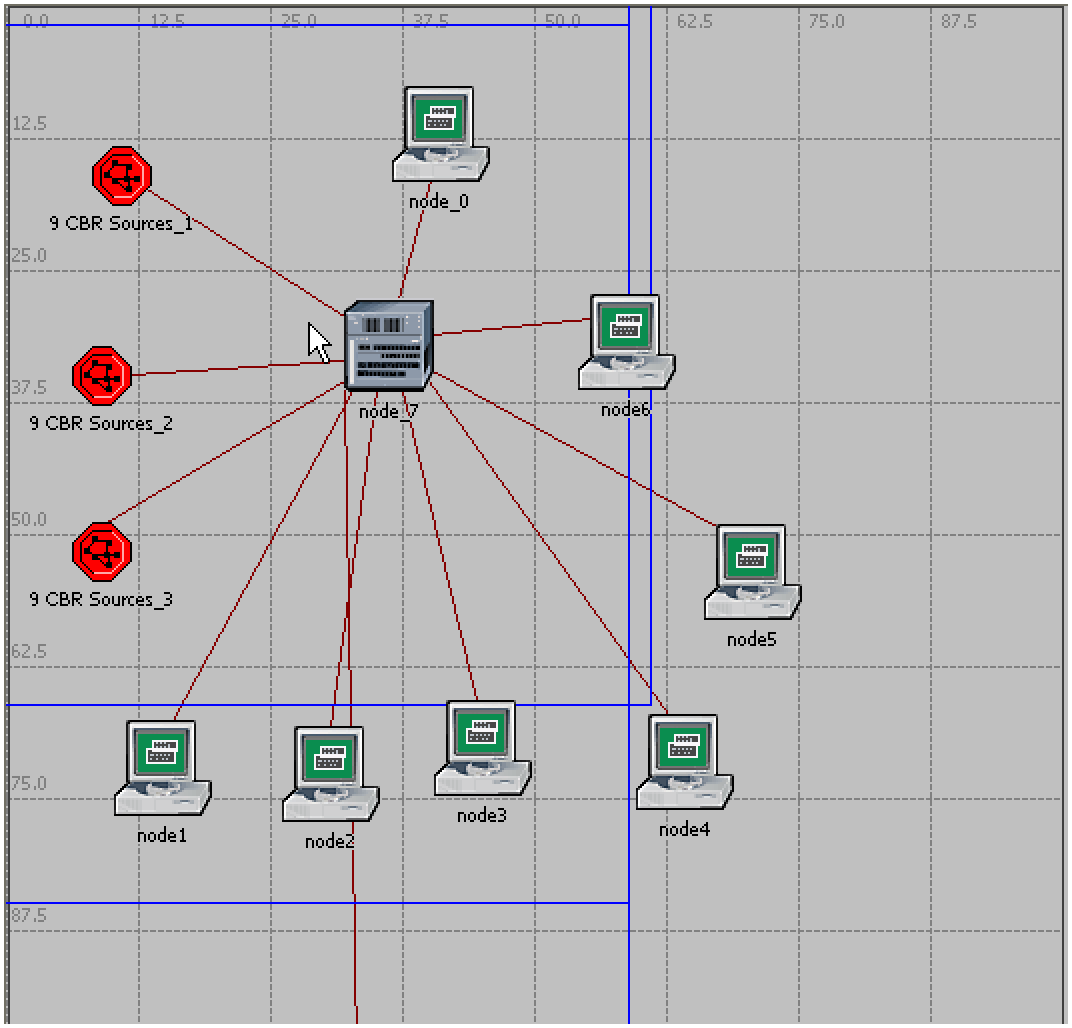
- Simulate for 255 s, with traffic turned on at 5 s
  - Needed to add small amount of best-effort traffic in reverse direction to ensure each destination node would be in the forwarding database of each switch (otherwise get flooding and link utilizations that exceed 100%)
    - Node_0 sink in stages 3 and 4 is used for some of this reverse traffic

Simulation Case 1 – Stages 1 and 2



Switch to Switch link
utilization = 50%

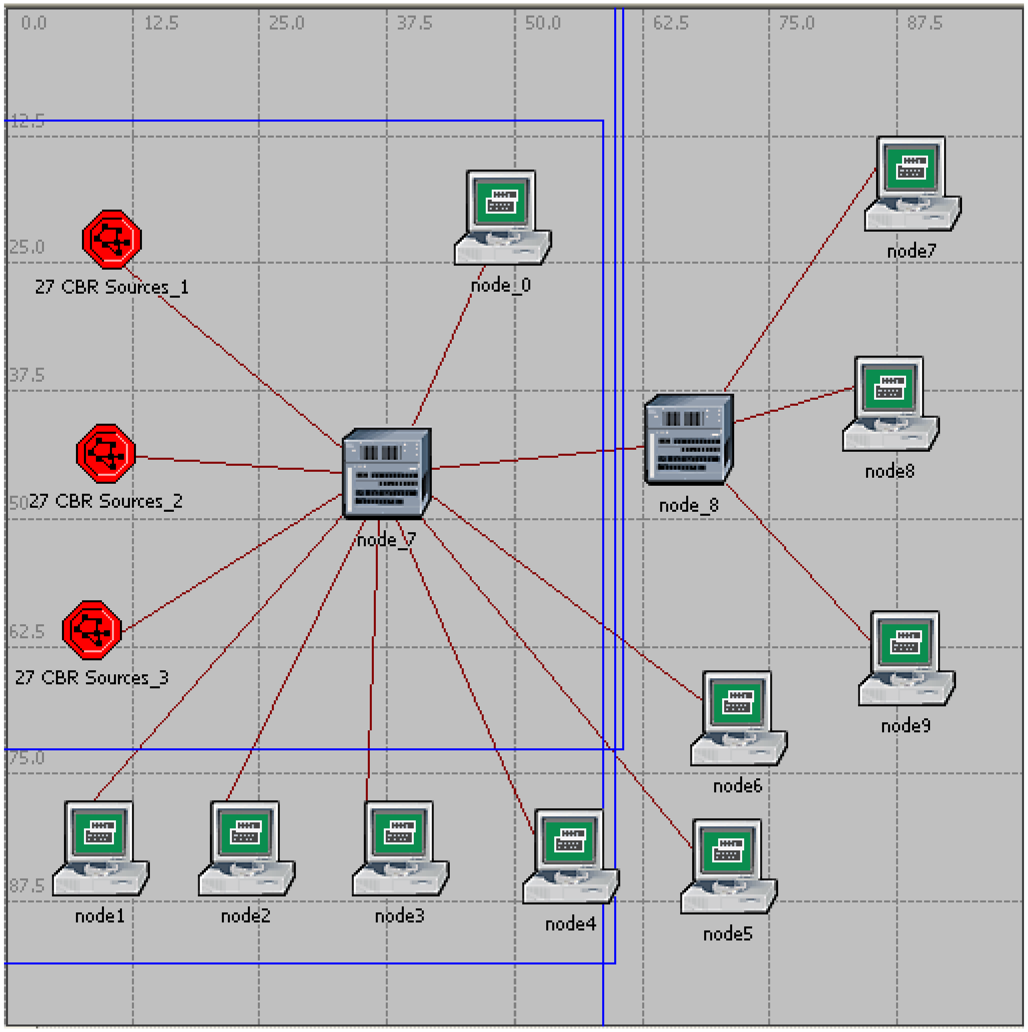
Simulation Case 1 - Stage 3



Switch to Switch link utilization = 50%

Each subnet (the red octagonal icons labeled 9 CBR Sources_1, _2, and _3 represent a stage 1 and 2 network as shown on the previous slide

Simulation Case 1 - Stage 4



Switch to Switch link utilization = 50%

Each subnet (the red octagonal icons labeled 27 CBR Sources_1, _2, and _3 represent a stage 3 network as shown on the previous slide

Simulation Case 1 Traffic Streams – Stages 1 and 2

- Node 2 to node 14, rate offset by -100 ppm
- Node 3 to node 15, rate offset by +100 ppm
- Node 5 to node 16, rate offset by -50 ppm
- Node 6 to node 17, rate offset by +50 ppm
- Node 8 to node 18, rate offset by -75 ppm
- Node 9 to node 19, rate offset by +75 ppm
- Streams from nodes 1, 4, and 7 are transported on link to stage 3 (see following slides for details)

Simulation Case 1 Traffic Streams - Stage 3

- 27 CBR Sources_1 subnet
  - 9 CBR Sources_1 subnet, node 4 to node 1, rate offset by –10 ppm
  - 9 CBR Sources_1 subnet, node 7 to node 2, rate offset by +10 ppm
  - 9 CBR Sources_2 subnet, node 4 to node 3, nominal rate
  - 9 CBR Sources_2 subnet, node 7 to node 4, rate offset by +10 ppm
  - 9 CBR Sources_3 subnet, node 4 to node 5, rate offset by –10 ppm
  - 9 CBR Sources_3 subnet, node 7 to node 6, nominal rate
- 27 CBR Sources_2 subnet
  - 9 CBR Sources_1 subnet, node 4 to node 1, rate offset by –10 ppm
  - 9 CBR Sources_1 subnet, node 7 to node 2, rate offset by +10 ppm
  - 9 CBR Sources_2 subnet, node 4 to node 3, nominal rate
  - 9 CBR Sources_2 subnet, node 7 to node 4, rate offset by +10 ppm
  - 9 CBR Sources_3 subnet, node 4 to node 5, rate offset by –10 ppm
  - 9 CBR Sources_3 subnet, node 7 to node 6, nominal rate

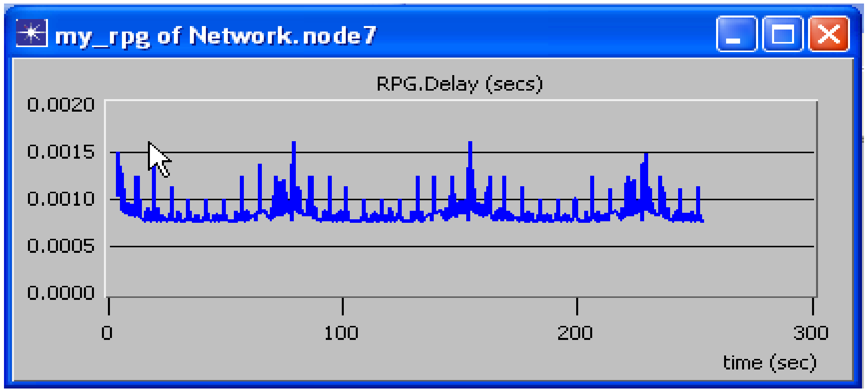
Simulation Case 1 Traffic Streams – Stage 3 (Cont.)

- 27 CBR Sources_3 subnet
  - 9 CBR Sources_1 subnet, node 4 to node 1, rate offset by –10 ppm
  - 9 CBR Sources_1 subnet, node 7 to node 2, rate offset by +10 ppm
  - 9 CBR Sources_2 subnet, node 4 to node 3, nominal rate
  - 9 CBR Sources_2 subnet, node 7 to node 4, nominal rate
  - 9 CBR Sources_3 subnet, node 4 to node 5, rate offset by –10 ppm
  - 9 CBR Sources_3 subnet, node 7 to node 6, nominal rate
- Note that the 3rd and 4th streams are different in all three 27 CBR subnets

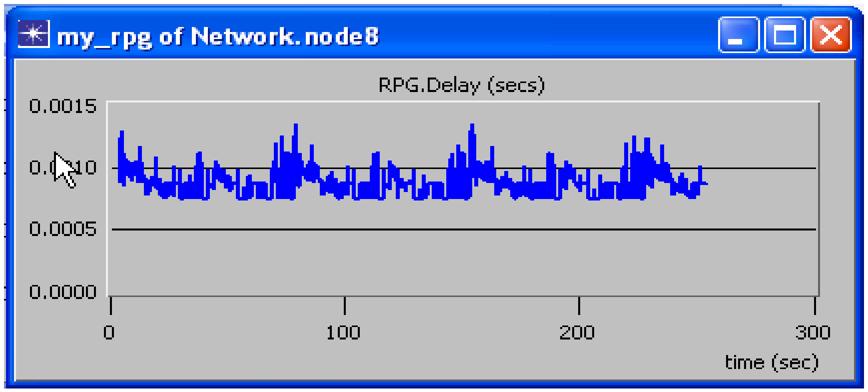
Simulation Case 1 Traffic Streams – Stage 4

- 27 CBR Sources_1, 9 CBR Sources_1 subnet, node 1 to node 7, nominal rate
- 27 CBR Sources_1, 9 CBR Sources_2 subnet, node 1 to node 1, rate offset by –10 ppm
- 27 CBR Sources_1, 9 CBR Sources_3 subnet, node 1 to node 2, rate offset by +10 ppm
- 27 CBR Sources_2, 9 CBR Sources_1 subnet, node 1 to node 8, rate offset by –10 ppm
- 27 CBR Sources_2, 9 CBR Sources_2 subnet, node 1 to node 3, rate offset by +10 ppm
- 27 CBR Sources_2, 9 CBR Sources_3 subnet, node 1 to node 4, nominal rate
- 27 CBR Sources_3, 9 CBR Sources_1 subnet, node 1 to node 9, rate offset by +10 ppm
- 27 CBR Sources_3, 9 CBR Sources_2 subnet, node 1 to node 5, rate offset by –10 ppm
- 27 CBR Sources_3, 9 CBR Sources_3 subnet, node 1 to node 6, nominal rate

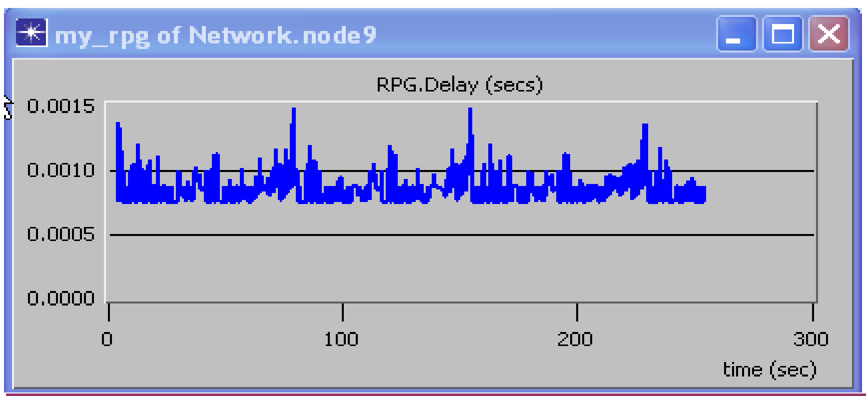
Simulation Case 1 Stage 4 Results for 4-Hop Streams



- 27 CBR Sources_1, 9 CBR Sources_1 subnet, node 1 to node 7, nominal rate

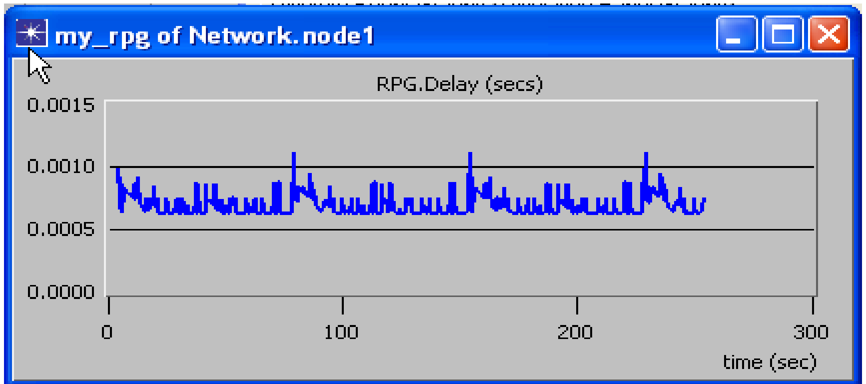


- 27 CBR Sources_2, 9 CBR Sources_1 subnet, node 1 to node 8, rate offset by –10 ppm

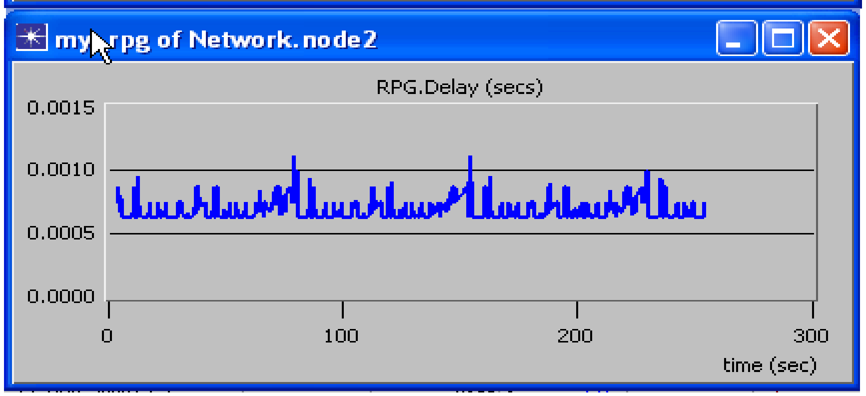


- 27 CBR Sources_3, 9 CBR Sources_1 subnet, node 1 to node 9, rate offset by +10 ppm

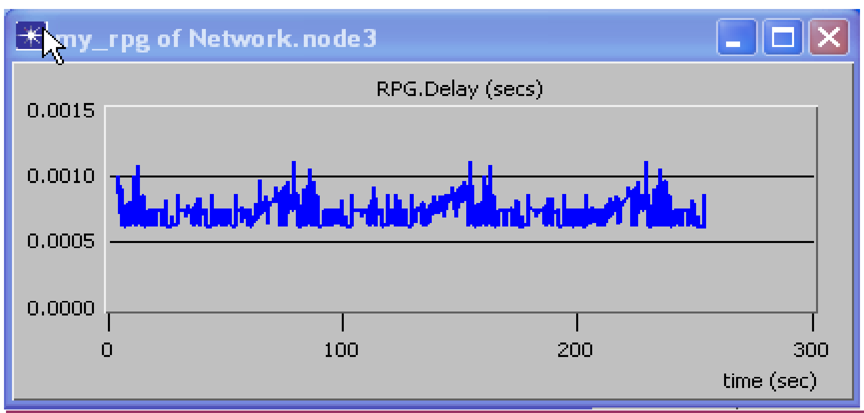
Simulation Case 1 Stage 3 Results for 3-Hop Streams



- 27 CBR Sources_1, 9 CBR Sources_2 subnet, node 1 to node 1, rate offset by –10 ppm

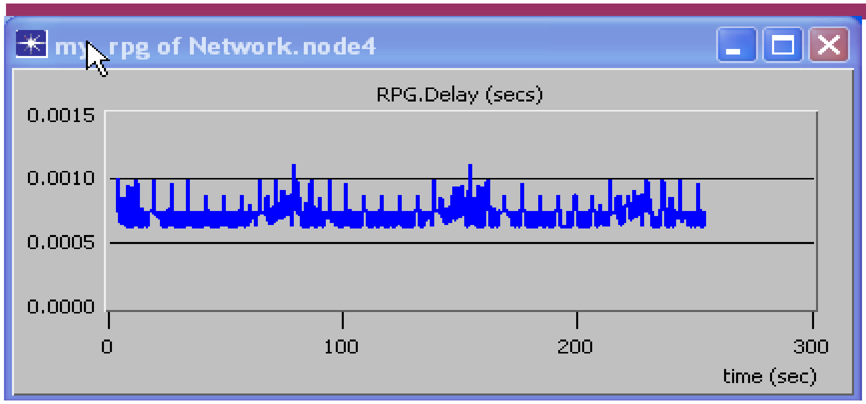


- 27 CBR Sources_1, 9 CBR Sources_3 subnet, node 1 to node 2, rate offset by +10 ppm

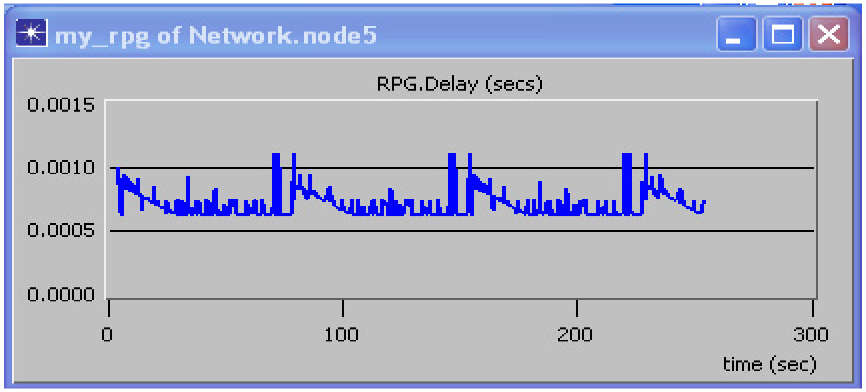


- 27 CBR Sources_2, 9 CBR Sources_2 subnet, node 1 to node 3, rate offset by +10 ppm

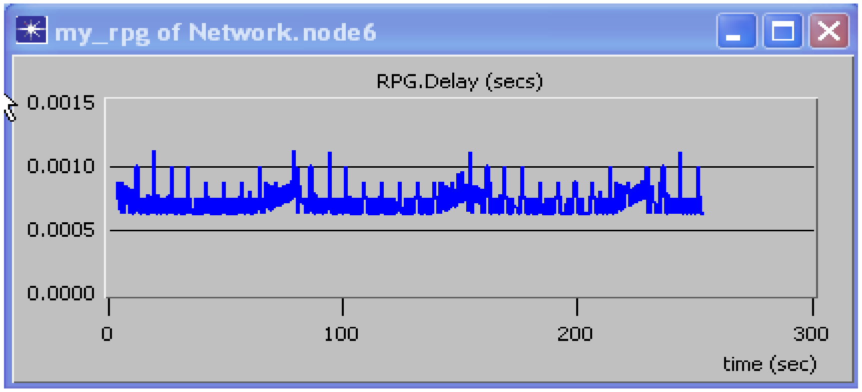
Simulation Case 1 Stage 3 Results for 3-Hop Streams



- 27 CBR Sources_2, 9 CBR Sources_3 subnet, node 1 to node 4, nominal rate

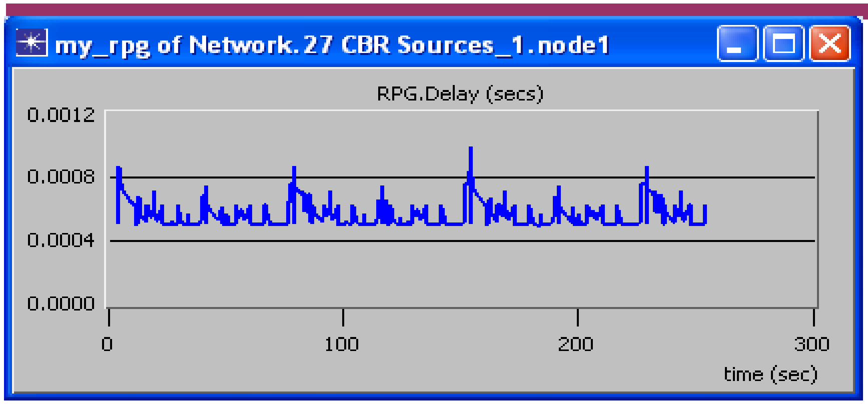


- 27 CBR Sources_3, 9 CBR Sources_2 subnet, node 1 to node 5, rate offset by –10 ppm

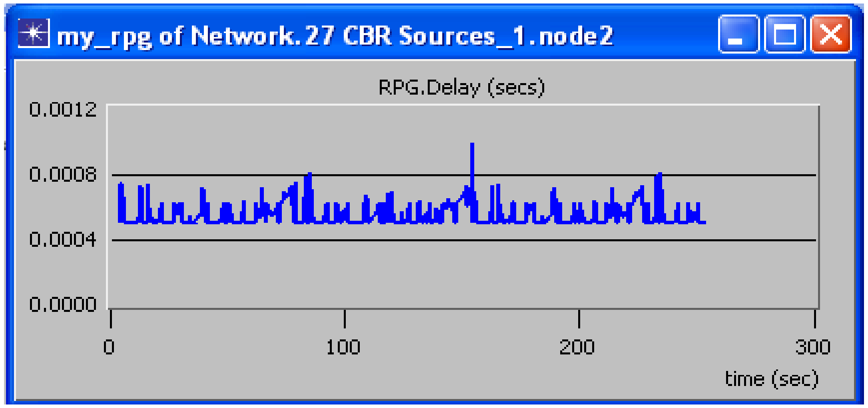


- 27 CBR Sources_3, 9 CBR Sources_3 subnet, node 1 to node 6, nominal rate

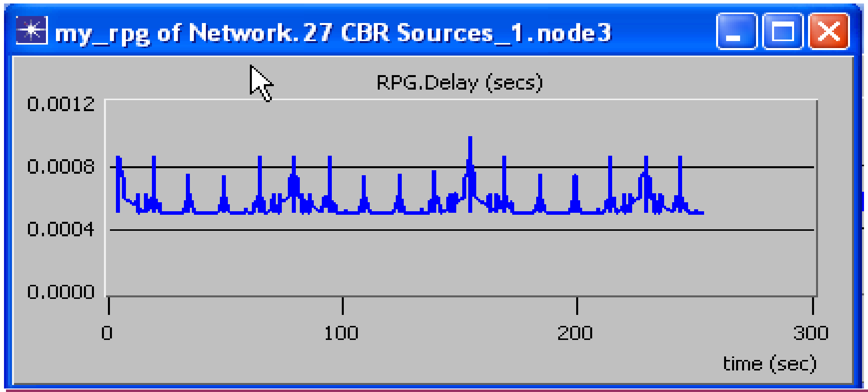
Simulation Case 1 Stage 2 Results for 2-Hop Streams



❑27 CBR Sources_1, 9 CBR Sources_1 subnet, node 4 to node 1, rate offset by –10 ppm

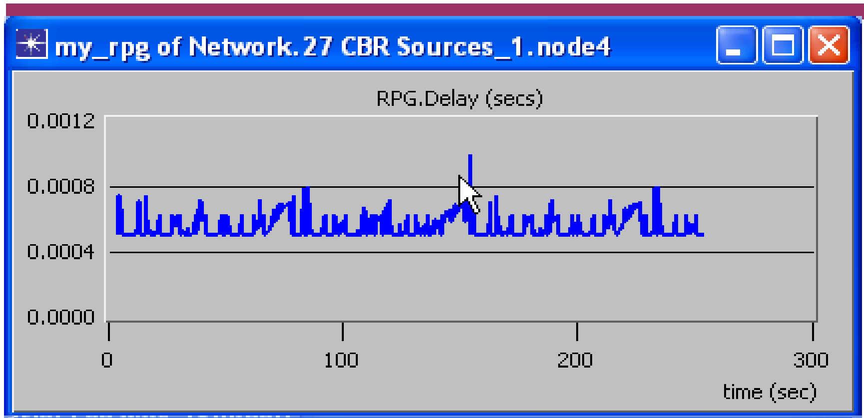


❑27 CBR Sources_1, 9 CBR Sources_1 subnet, node 7 to node 2, rate offset by +10 ppm

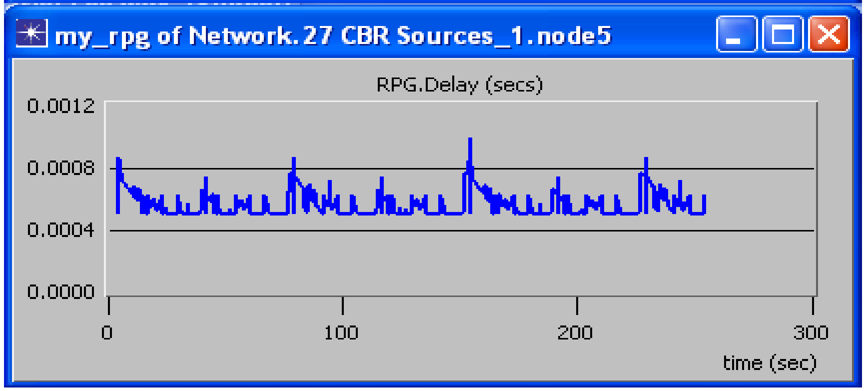


❑27 CBR Sources_1, 9 CBR Sources_2 subnet, node 4 to node 3, nominal rate

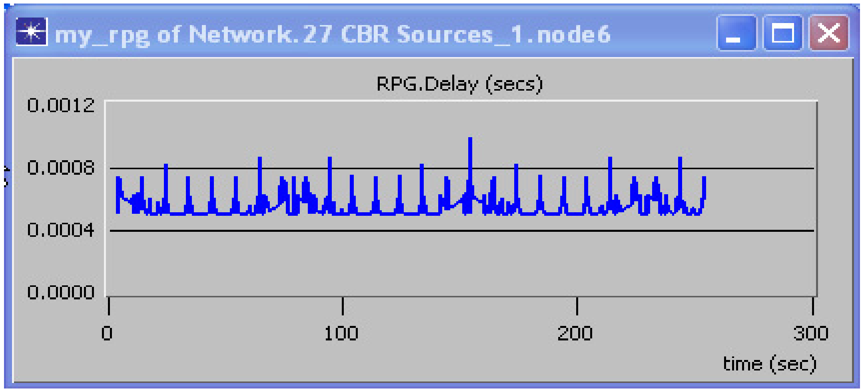
Simulation Case 1 Stage 2 Results for 2-Hop Streams



- 27 CBR Sources_1, 9 CBR Sources_2 subnet, node 7 to node 4, rate offset by +10 ppm

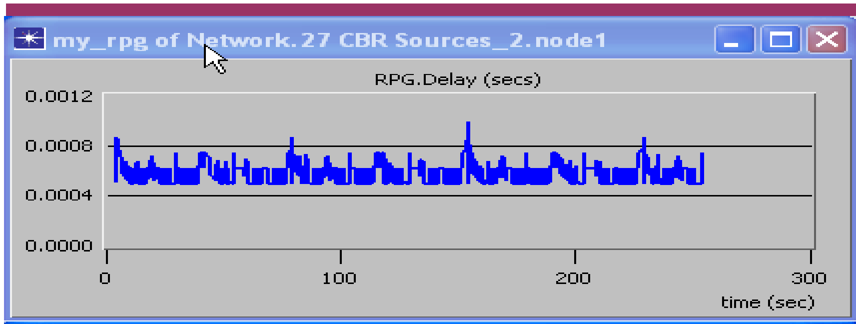


- 27 CBR Sources_1, 9 CBR Sources_3 subnet, node 4 to node 5, rate offset by –10 ppm

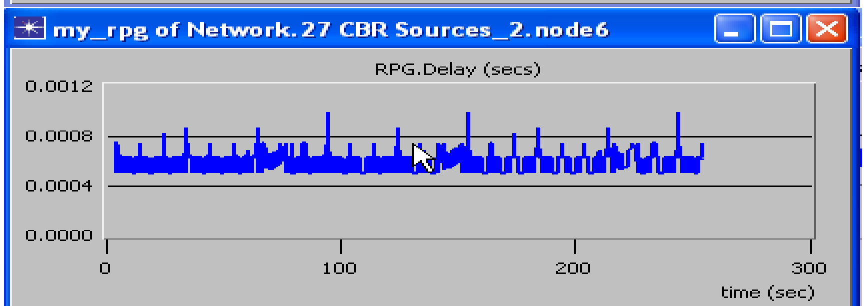


- 27 CBR Sources_1, 9 CBR Sources_3 subnet, node 7 to node 6, nominal rate

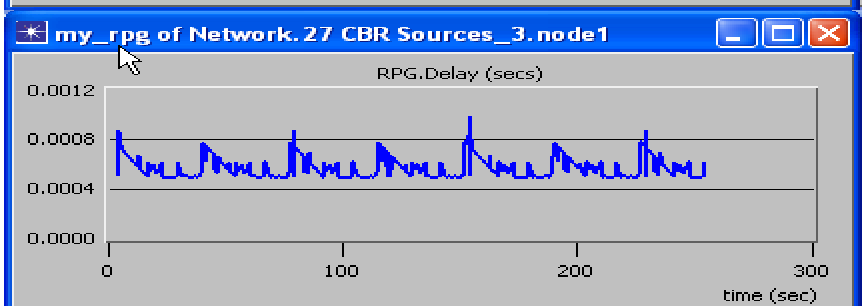
Simulation Case 1 Stage 2 Results for 2-Hop Streams



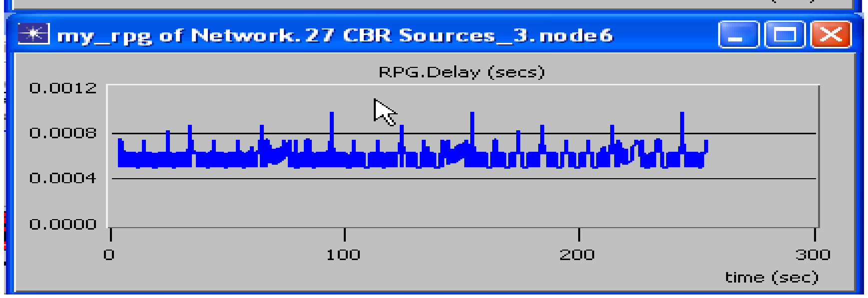
❑27 CBR Sources_2, 9 CBR Sources_1 subnet, node 4 to node 1, rate offset by –10 ppm



❑27 CBR Sources_2, 9 CBR Sources_3 subnet, node 7 to node 6, nominal rate

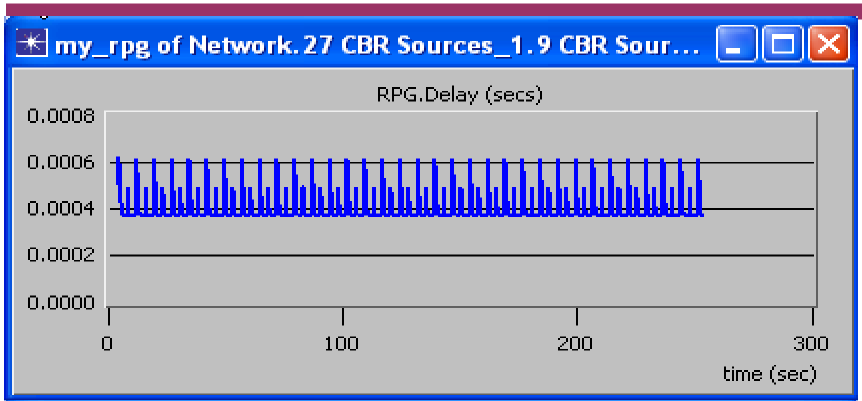


❑27 CBR Sources_3, 9 CBR Sources_1 subnet, node 4 to node 1, rate offset by –10 ppm

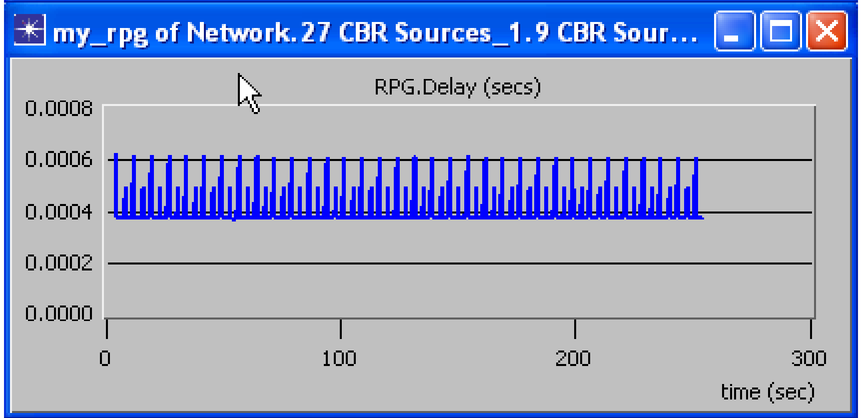


❑27 CBR Sources_3, 9 CBR Sources_3 subnet, node 7 to node 6, nominal rate

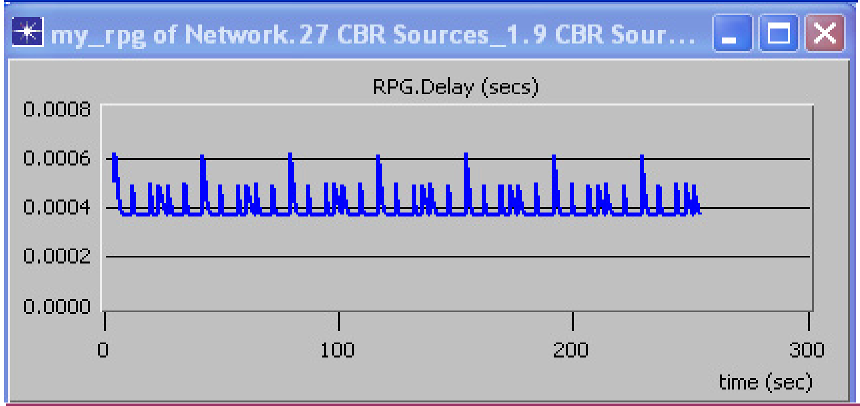
Simulation Case 1 Stage 1 Results for 1-Hop Streams



- 27 CBR Sources_1, 9 CBR Sources_1 subnet, node 2 to node 14, rate offset by –100 ppm

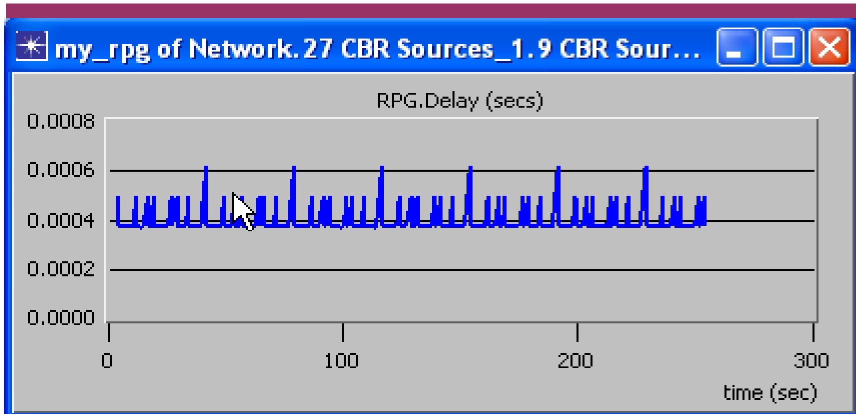


- 27 CBR Sources_1, 9 CBR Sources_1 subnet, node 3 to node 15, rate offset by +100 ppm

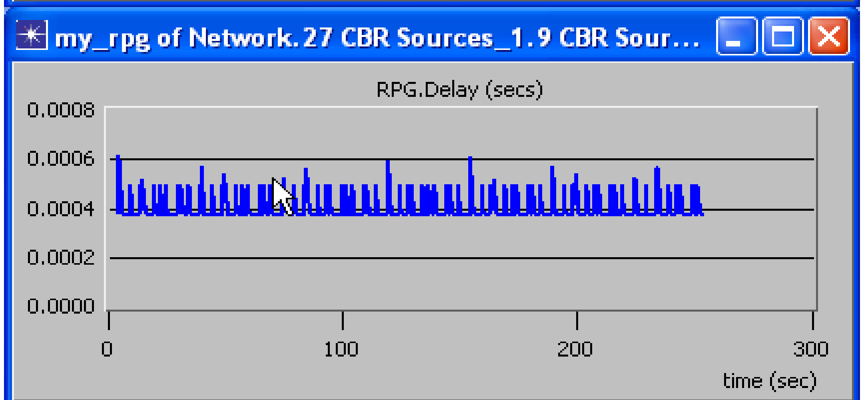


- 27 CBR Sources_1, 9 CBR Sources_1 subnet, node 5 to node 16, rate offset by –50 ppm

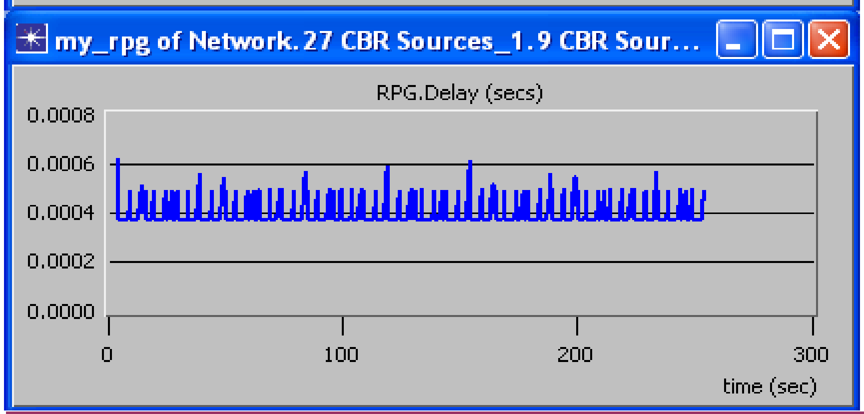
Simulation Case 1 Stage 1 Results for 1-Hop Streams



- ❑27 CBR Sources_1, 9 CBR Sources_1 subnet, node 6 to node 17, rate offset by +50 ppm

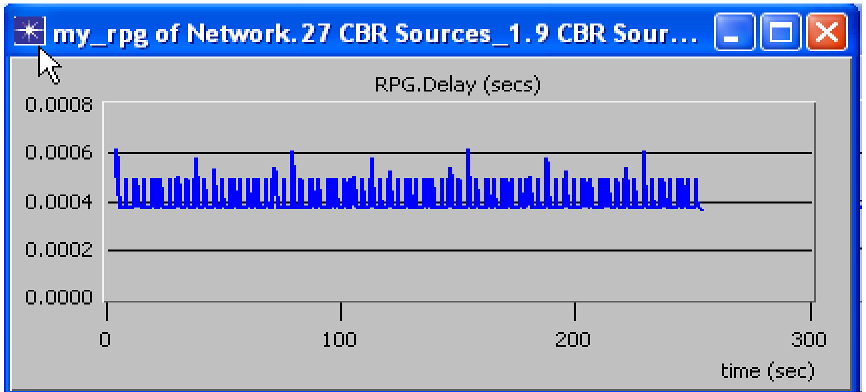


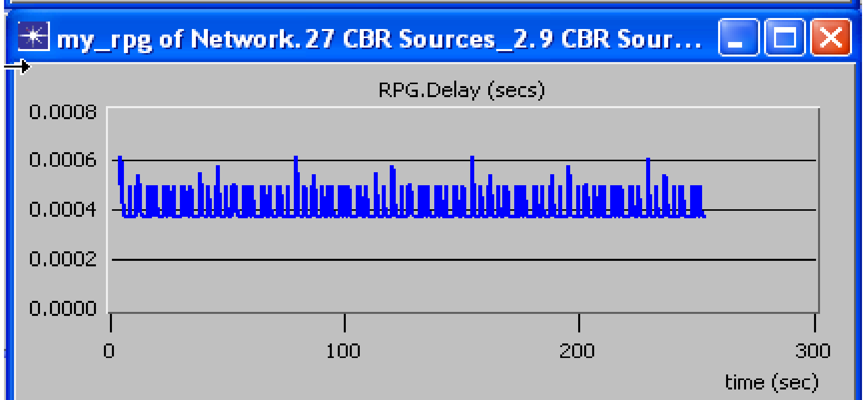
- ❑27 CBR Sources_1, 9 CBR Sources_1 subnet, node 8 to node 18, rate offset by –75 ppm

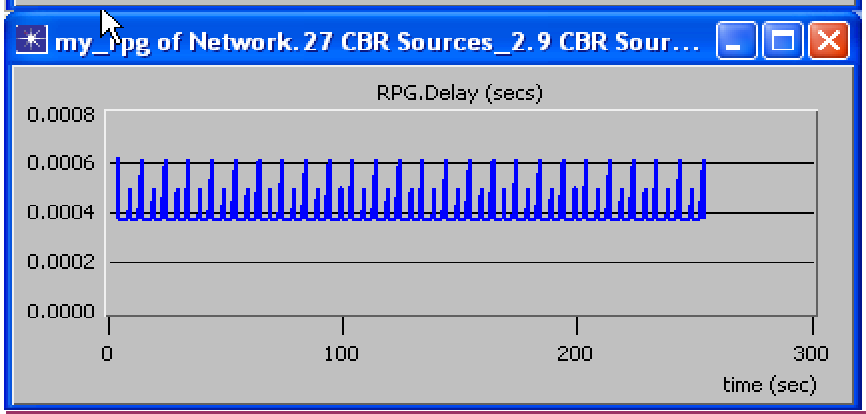


- ❑27 CBR Sources_1, 9 CBR Sources_1 subnet, node 9 to node 19, rate offset by +75 ppm

Simulation Case 1 Stage 1 Results for 1-Hop Streams

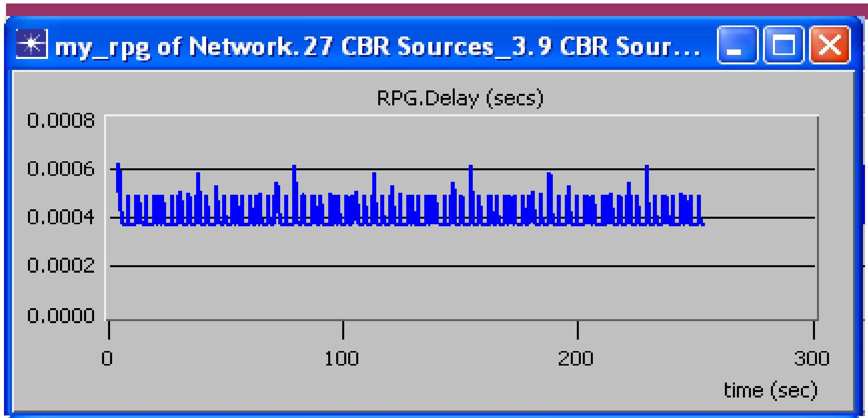




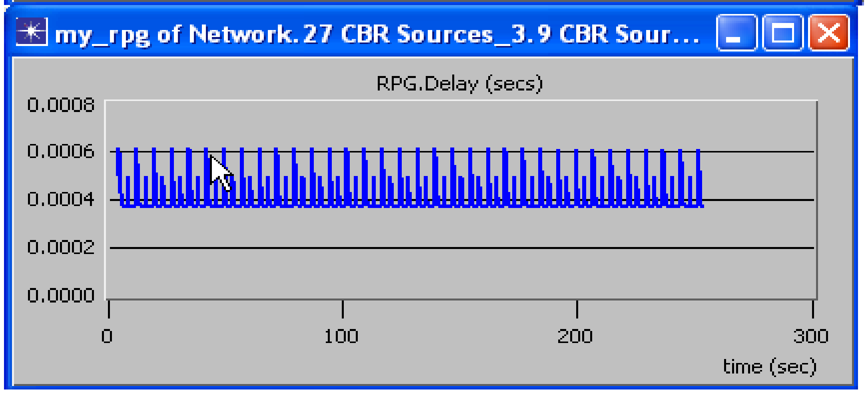


- 27 CBR Sources_1, 9 CBR Sources_2 subnet, node 2 to node 14, rate offset by –100 ppm
- 27 CBR Sources_2, 9 CBR Sources_2 subnet, node 2 to node 14, rate offset by –100 ppm
- 27 CBR Sources_2, 9 CBR Sources_2 subnet, node 9 to node 19, rate offset by +75 ppm

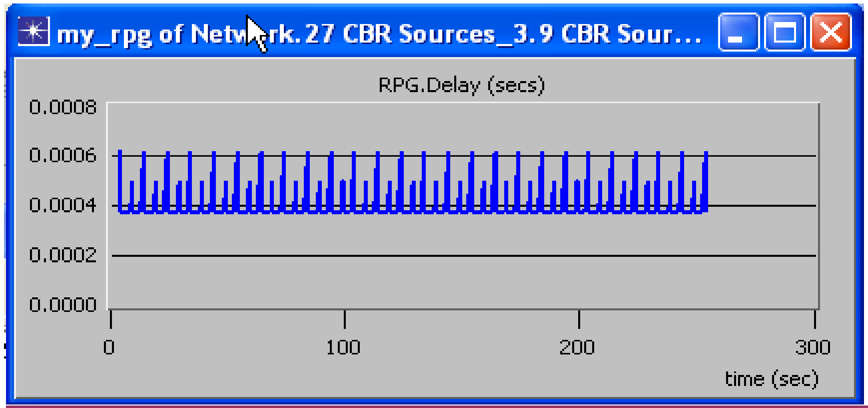
Simulation Case 1 Stage 1 Results for 1-Hop Streams



- 27 CBR Sources_3, 9 CBR Sources_2 subnet, node 2 to node 14, rate offset by –100 ppm



- 27 CBR Sources_3, 9 CBR Sources_2 subnet, node 9 to node 19, rate offset by +75 ppm



- 27 CBR Sources_3, 9 CBR Sources_3 subnet, node 2 to node 14, rate offset by –100 ppm

Summary of Peak Delay and Peak-to-Peak Delay Variation

Number of Hops	Peak Delay (ms)	Peak-to-Peak Delay Variation (μs)
1	0.62	250
2	1.0	500
3	1.5	510
4	1.6	850

Conclusions

- Peak delay and peak-to-peak delay variation increase with number of hops, as expected
  - Peak delay reaches 1.6 ms after 4 hops
  - Peak-to-peak delay variation reaches 850 μs after 4 hops
- Peak delay is slightly below worst-case that would be obtained for this 4-hop case
  - Worst-case for 3 contending CBR streams at a switch occurs when 2 frames are queued when a frame arrives (for link utilization < 100%)
  - Then, for contention occurring at 4 switches (in a 4-hop path), the delay due to contention is (4)(2)(frame transmission delay) = 8(frame transmission delay)
  - Also have transmission delay for the frame itself on the 4 switch-to-switch links plus the two access links
  - Then total delay due to transmission and queueing, in worst case, is 14(frame transmission delay)
  - Then worst-case delay (neglecting propagation delay since it is much smaller) is
    - $14(12000+8(38)\text{ bits})/10^{8}\text{ bits/s}) = 1.72 \times 10^{-3}\text{ s} = 1.72\text{ ms}$

Conclusions (Cont.)

- If number of contending traffic streams is increased from 3 to 6 (still for a 4-hop case), would expect worst-case delay due to contention to increase by (4)(3)(frame transmission delay) = 12(12000+8(38) bits)/10^8 bits/s) = 1.48 ms
  - Worst-case total delay in this case would be 1.72+1.48 ms = 3.2 ms
- If number of hops is increased from 4 to 7 (still for 3 contending traffic streams), would expect worst-case delay to increase to [7(2)+9](frame transmission delay] (i.e., 2 contending frames at each of 7 switches plus 9 total transmission delays (switch-to-switch plus access links))
  - Worst case total delay in this case would be 23(12000+8(38) bits)/10^8 bits/s) = 2.83×10^{-3} s = 2.83 ms

Conclusions (Cont.)

- Peak-to-peak delay variation is slightly below worst-case that would be obtained for this 4-hop case
  - Worst-case peak-to-peak delay variation is equal to the worst-case delay due to contention at the switches, as this is the component of delay that is not always present
  - Then worst-case peak-to-peak delay variation is
    - $8(12000+8(38)\text{ bits})/10^8\text{ bits/s}) = 9.8 \times 10^{-6}\text{ s} = 980\ \mu\text{s}$
- Note that the amount by which the worst case peak delay exceeds the actual peak delay (1.72 ms – 1.6 ms = 0.12 ms) and the amount by which the worst case peak-to-peak delay variation exceeds the actual peak-to-peak delay variation (980 μs – 850 μs = 130 μs = 0.13 ms) are approximately equal, as expected
- Results obtained for 1, 2, and 3 hops are consistent with similar worst-case analyses for these cases (with 3-hop results below worst-case results by approximately 0.1 ms)
- While peak delay does not exceed 2 ms for 3 contending traffic streams and 4 hops, the results indicate that it will exceed 2 ms for 6 contending traffic streams with 4 hops, and 3 contending traffic streams with 7 hops

Conclusions (Cont.)

- Worst-case peak-to-peak delay variation is just below 1 ms (i.e., 980 μs) for 3 contending traffic streams with 4 hops
  - For 6 contending streams with 4 hops, this increases to 5(4)(frame transmission delay) = 2.46 ms
  - For 4 contending streams with 7 hops, this increases to 3(7)(frame transmission delay) = 2.58 ms
- Therefore, while peak-to-peak delay does not exceed 2 ms for 3 contending traffic streams and 4 hops, the results indicate that it will exceed 2 ms for 6 contending traffic streams with 4 hops, and 3 contending traffic streams with 7 hops

Conclusions (Cont.)

- The results indicate the following rules of thumb may be used to estimate worst-case delay and worst-case peak-to-peak delay variation for an arbitrary N hop path through a network
  - Worst case end-to-end delay = $[(N + 2) + \Sigma_{j = 1 \text{ to number of switches}} (\{\text{number of incoming links at switch } j\} - 1)][\text{frame transmission delay}]$
    - Assumes propagation delay is negligible (must be added if it is not negligible)
  - Worst case peak-to-peak delay variation = $[\Sigma_{j = 1 \text{ to number of switches}} (\{\text{number of incoming links at switch } j\} - 1)][\text{frame transmission delay}]$

References

1. Geoffrey M. Garner and Felix Feng, *Delay Variation Simulation Results for Transport of Time-Sensitive Traffic over Conventional Ethernet*, Samsung presentation at July, 2005 IEEE 802.3 ResE SG meeting, San Francisco, CA, July 18, 2005.
2. Geoffrey M. Garner, *End-to-End Jitter and Wander Requirements for ResE Applications*, Samsung presentation at May, 2005 IEEE 802.3 ResE SG meeting, Austin, TX, May 16, 2005.
3. *Residential Ethernet (RE) (a working paper)*, Draft 0.136, maintained by David V. James and based on work by him and other contributors, August 10, 2005, available via http://www.ieee802.org/3/re_study/public/index.html