

Explicit PCN Marking

draft-babiarz-pcn-explicit-marking-00

Detailed Simulation Results:

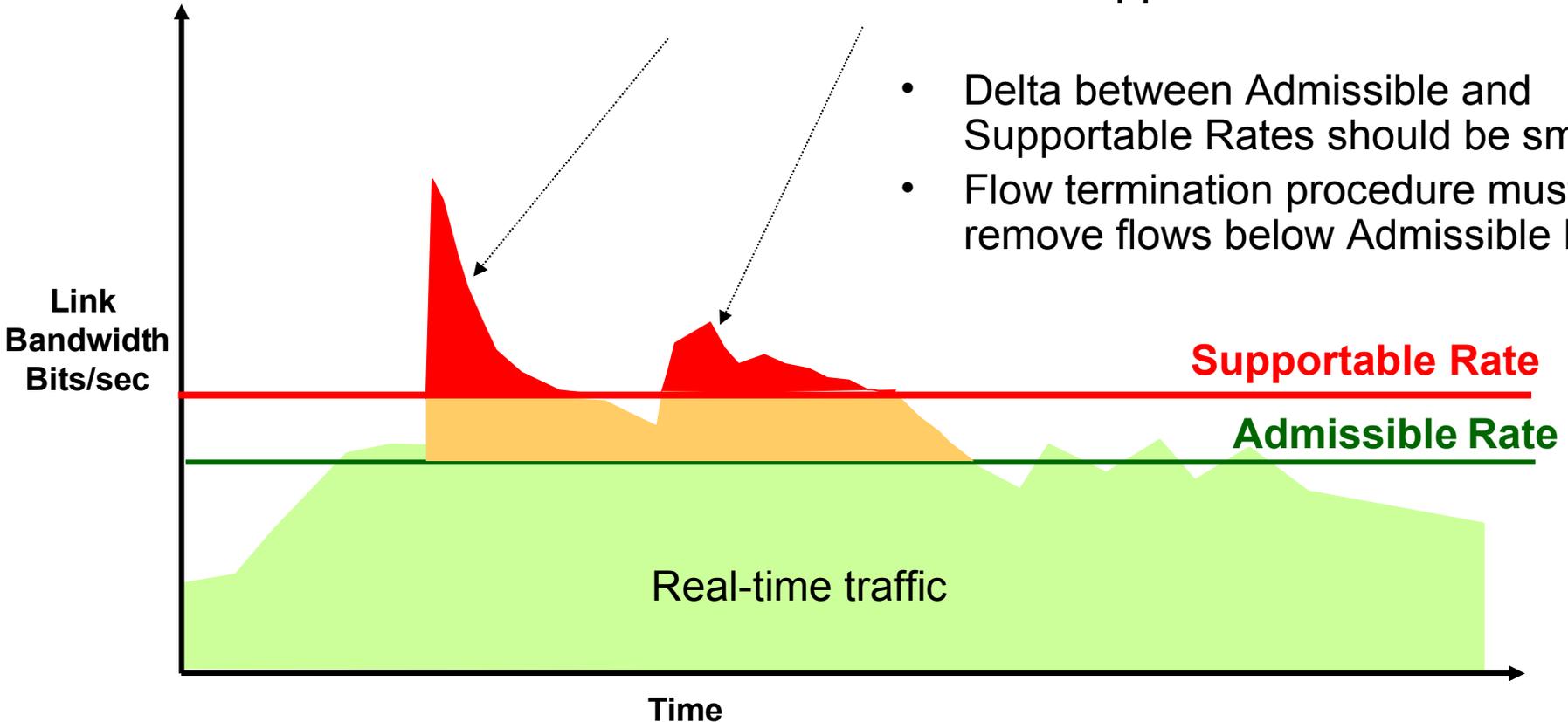
http://standards.nortel.com/pcn/Simulation_EPCN.pdf

Jozef Babiarz
Xiao-Gao Liu
Kwok Ho Chan

68th IETF, March 19, 2007

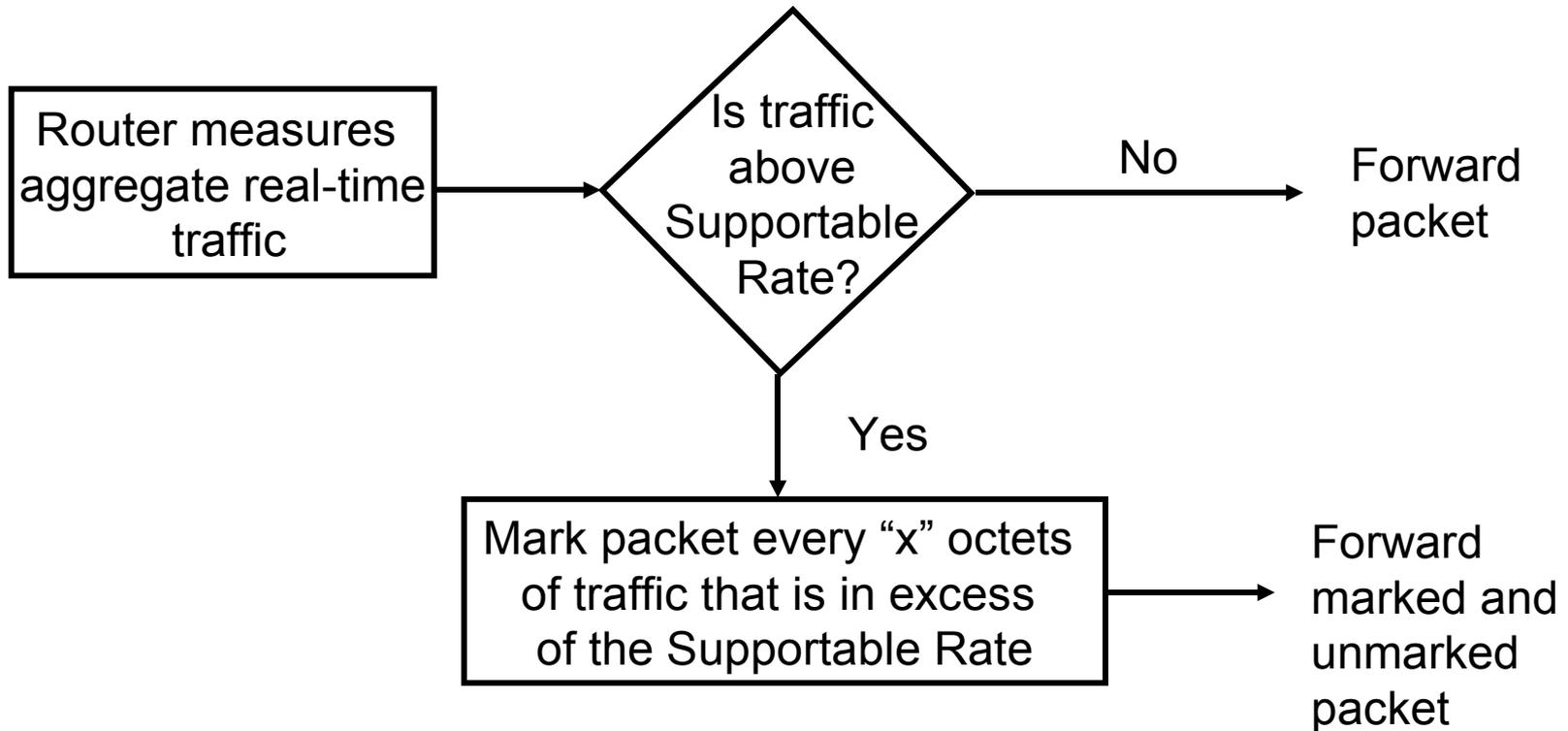
Principles and Objectives

- Focus of this draft is on excess marking and flow termination when traffic is above Supportable Rate

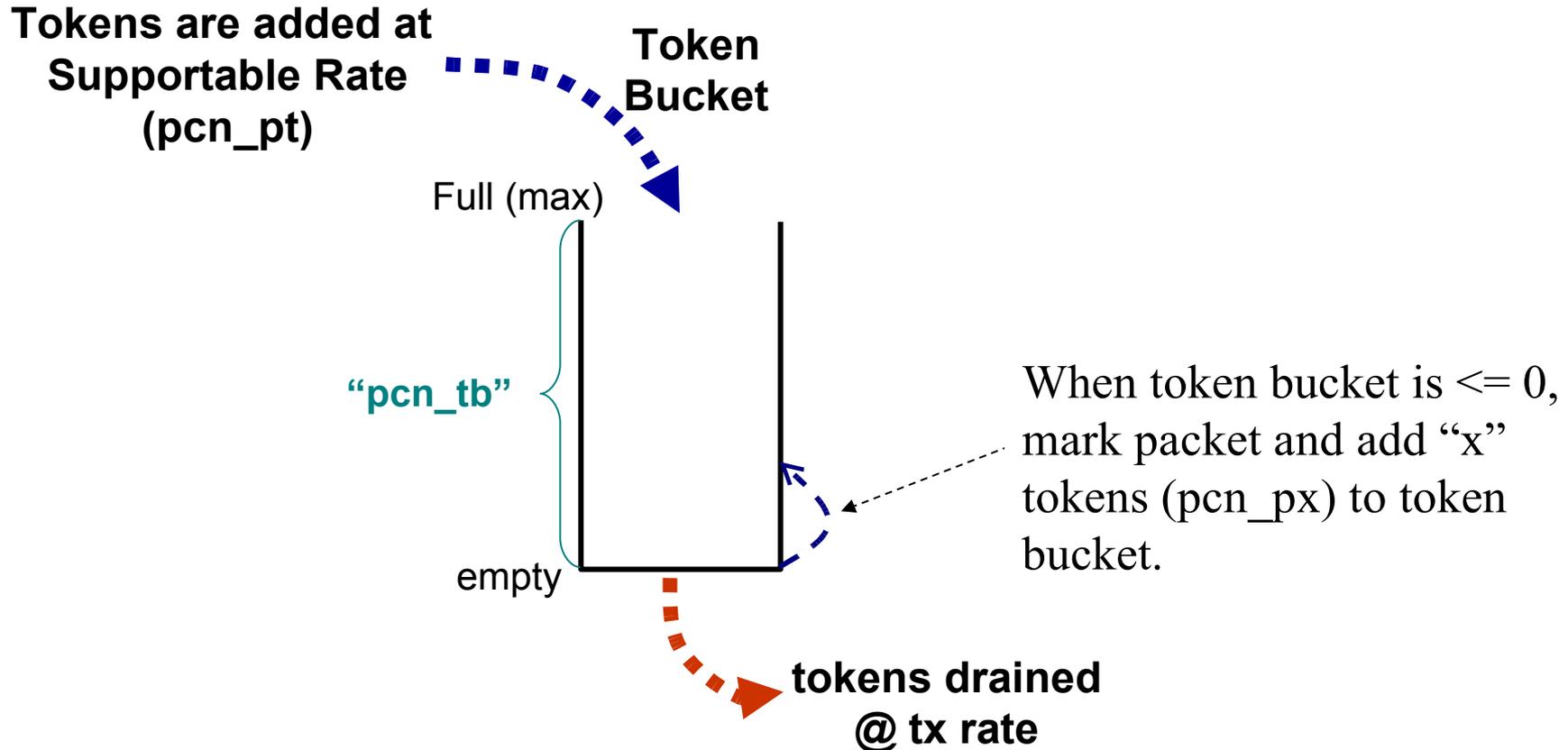


Note: Flow admission is not discussed

Explicit Marking Behavior



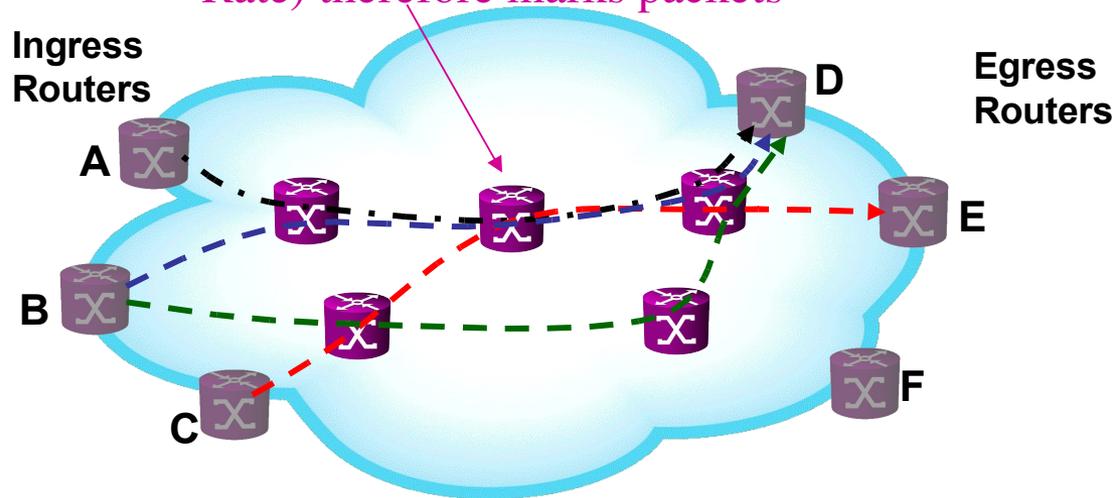
Metering/Marking Representation using Token Bucket



- When Tx rate exceeds Supportable Rate (pcn_pt) tokens are drain at faster rate than added

Operation of Explicit PCN Marking

An interior router/link experiences congestion (traffic exceeds Supportable Rate) therefore marks packets



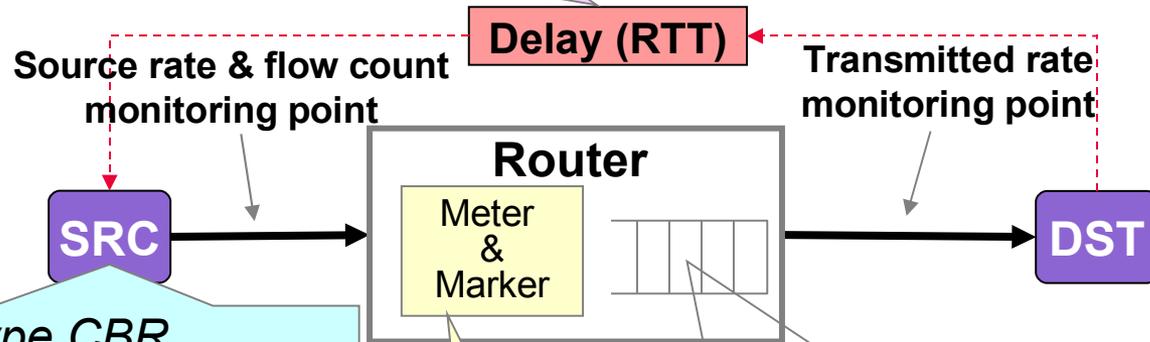
- Egress routers monitor packets, on seeing a congestion marked packet, signal to ingress router what flow(s) is/are marked, meaning what flow(s) is/are going through congestion point in the network.
- As well, egress router can also signal to ingress the amount of congestion if ingress/egress aggregate information is available.
- Ingress router blocks packets belonging to congestion marked flows and may signal to source to stop sending packets.

== == == Aggregated flows experiencing congestion

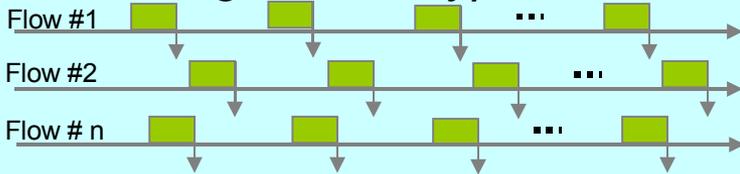
- - - Aggregated flows not experiencing congestion

Simulation Setup for Voice

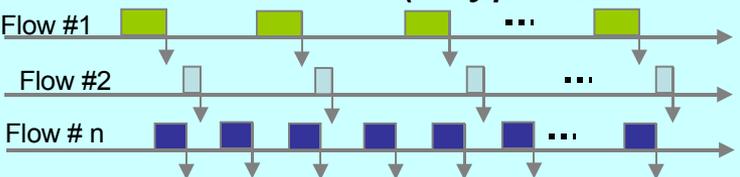
“PM” Flag with Delay = RTT (2, 10, 50, 100, 200, 800 ms)



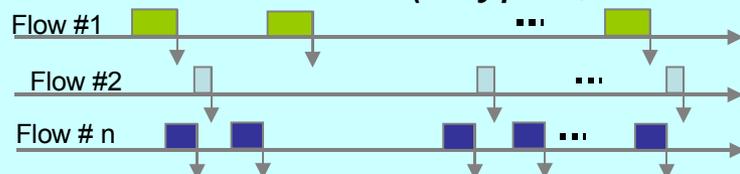
Single codec type CBR



Mixed codec CBR (3 types, 1/3 each)



Mixed codec VBR (3 types, 1/3 each)



Forwarding Rate and Queue Size

Rate: 1G, or 1.8, 1.5, 1.2 times preemption rate
Queue: large (unlimited), or programmable

Preemption Thresholds

BW in Mbps: (0.8, 8, 40)

Token bucket = pcn_tb

Marking interval = pcn_px

Note: Preemption Threshold = Supportable Rate

Simulation Results (40 Mbps)

- Large number of flows
 - Preemption level = 40 Mbps
 - No packet loss
- Preemption times: 50ms, 200ms and 800ms
 - Preemption time = RTT + processing time in endpoints/GWs
- Excess rate marking every “n” bytes
 - pcn_px = 2,064, 4,064 and 8,064 bytes
- Token bucket size
 - pcn_tb = 52K – 58K bytes
- Graph results
 - Violet trace shows number of flows
 - Orange trace shows transmitted rate (router egress)
 - Blue trace shows source rate (router ingress)

Note: Blue trace only visible if there is packet loss

Token Bucket Parameter Setting

pcn_tb = 52K bytes

pcn_px = 2,064 bytes

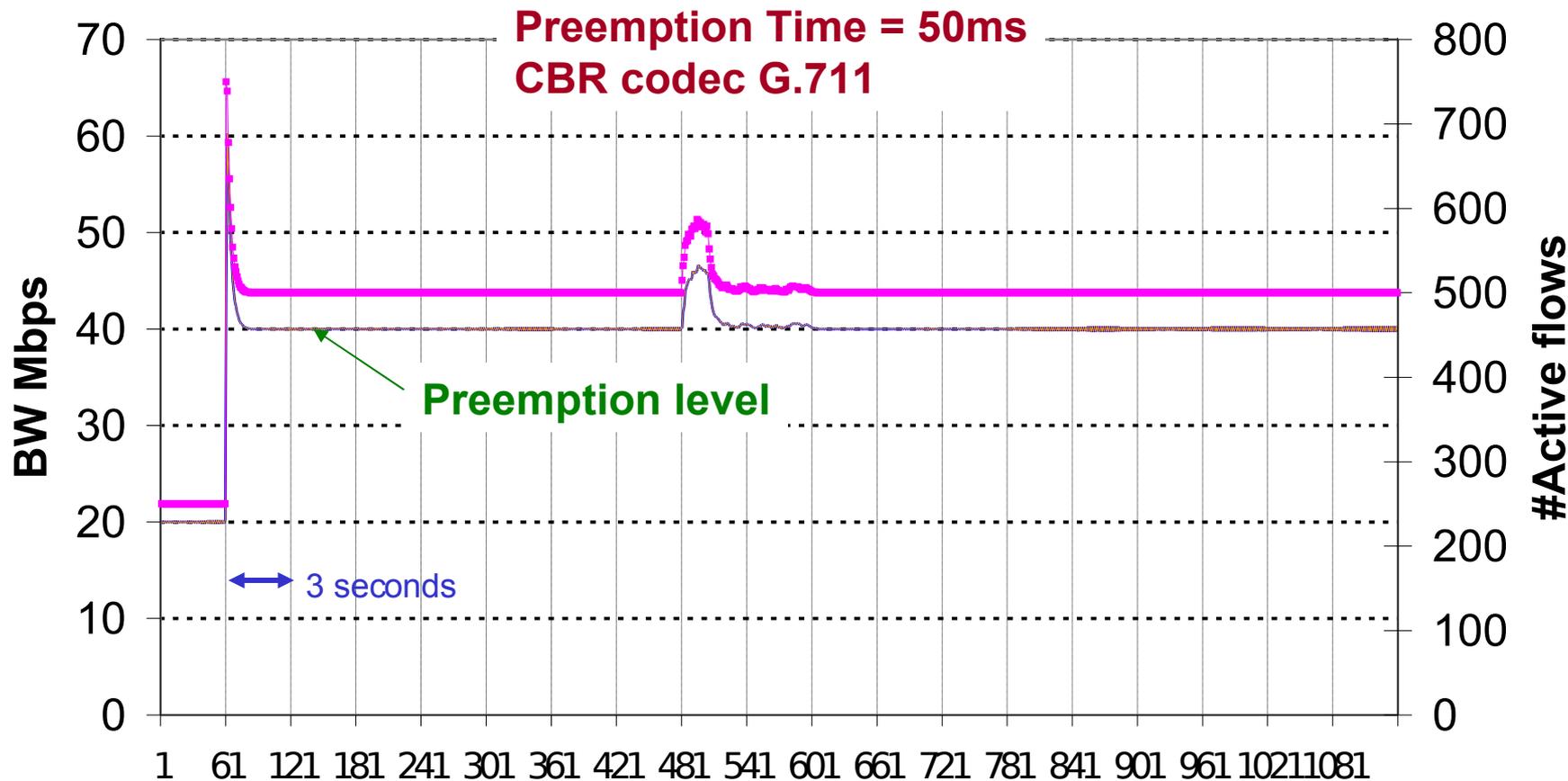
Preemption level = 40 Mbps

Service Class BW limit = 100 Mbps

Buffer size = 10,000 bytes

G711-20/CBR/0-Idle/40Mbps/50msRTT_2.5BW/mBF

— Source rate — Transmitted rate — #Active flows



Observation time interval = 0.05 sec

Simulation # JB326

Token Bucket Parameter Setting

pcn_tb = 52K bytes

pcn_px = 2,064 bytes

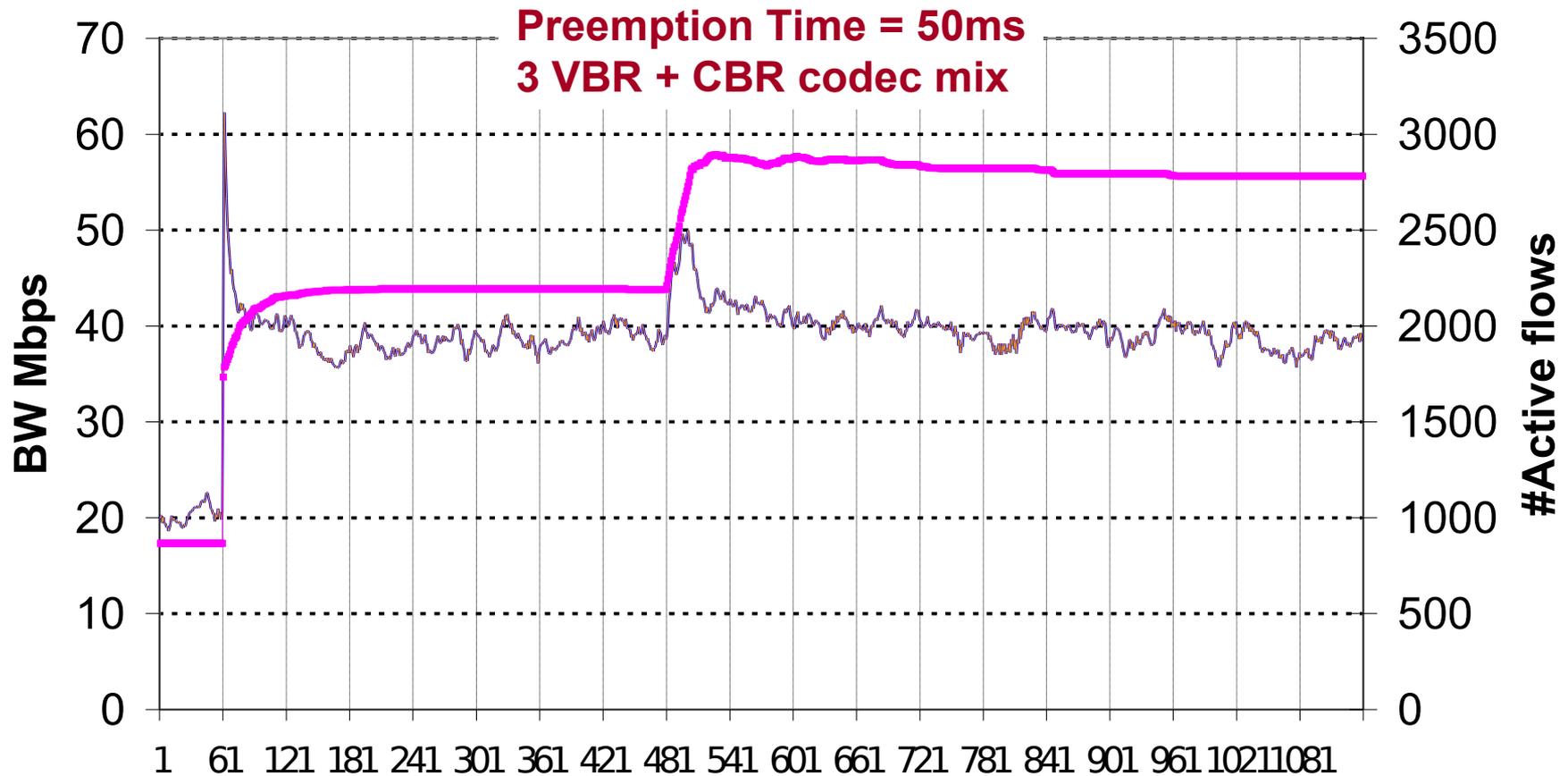
Preemption level = 40 Mbps

Service Class BW limit = 100 Mbps

Buffer size = 9,250 bytes

MixCodec/VBR/0-Idle/40Mbps/50msRTT_2.5BW/mBF

— Source rate — Transmitted rate — #Active flows



Observation time interval = 0.05 sec

Simulation # JB327

Token Bucket Parameter Setting

pcn_tb = 58K bytes

pcn_px = 8,064 bytes

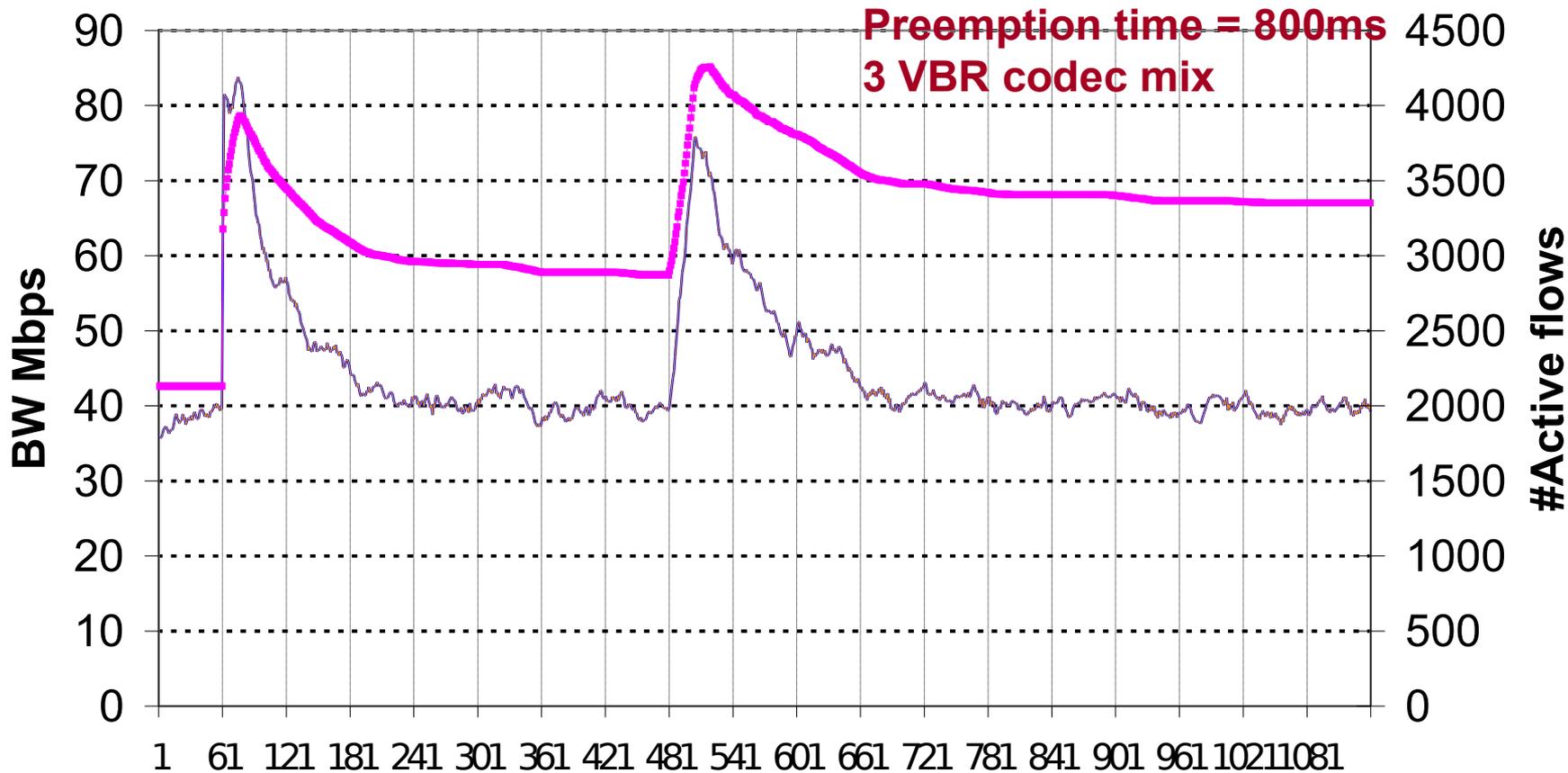
Preemption level = 40Mbps

Service Class BW limit = 100 Mbps

Buffer size = 12,700 bytes

MixCodec/VBR/0-Idle/40Mbps/800msRTT_2.5BW/IBF

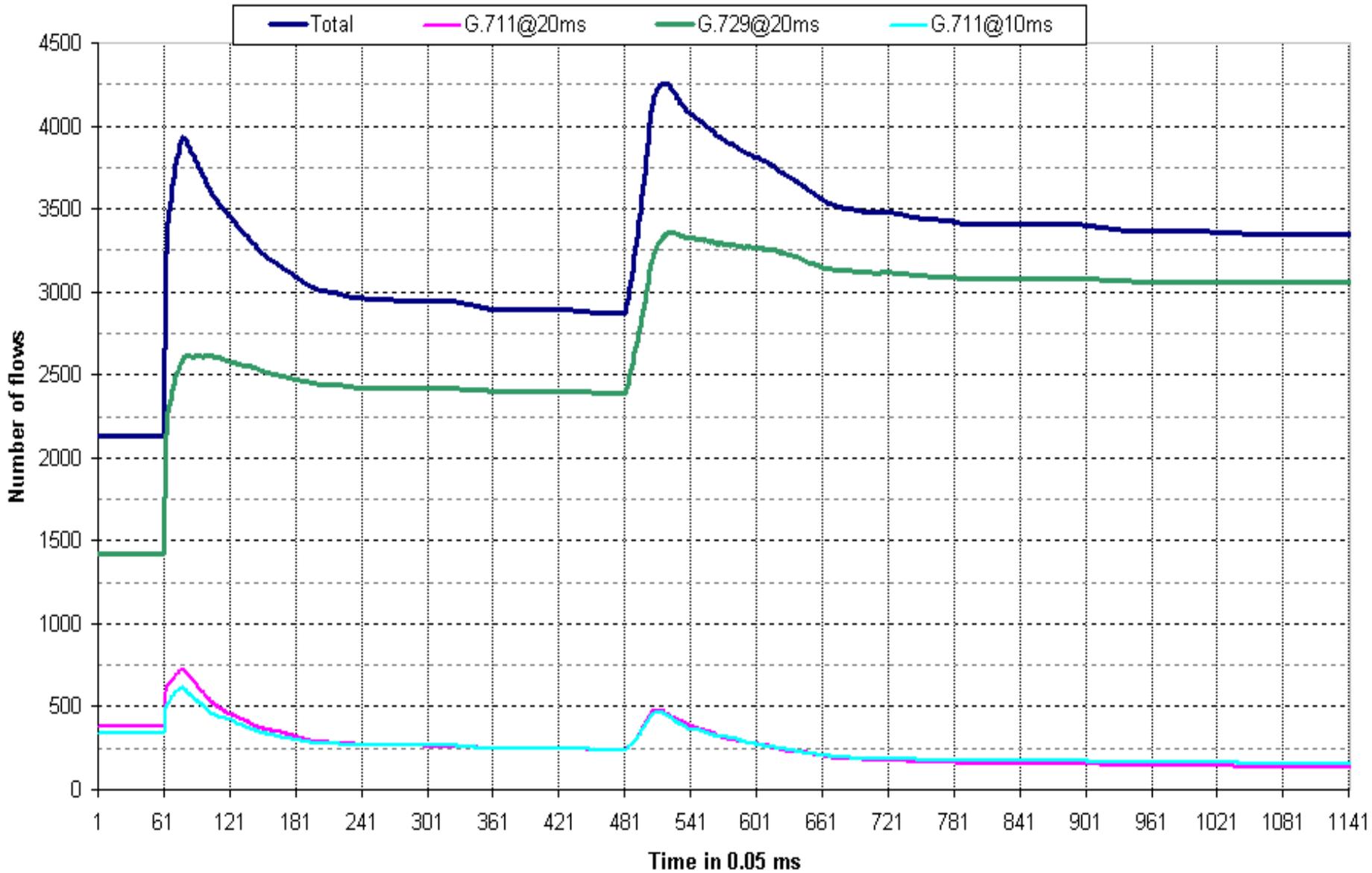
— Source rate — Transmitted rate — #Active flows



Observation time interval = 0.05 sec

Simulation # JB29110

Flow count for 3 VBR codecs



Simulation Results (800 Kbps)

- Small number of flows
 - Preemption level = 0.8 Mbps
 - No packet loss
- Preemption times: 50ms, 200ms and 800ms
 - Preemption time = RTT + processing time in endpoints/GWs
- Excess rate marking every “n” bytes
 - pcn_px = 2,064 and 8,064 bytes
- Token bucket size
 - pcn_tb = 52K – 58K bytes
- Graph results
 - Violet trace shows number of flows
 - Orange trace shows transmitted rate (router egress)
 - Blue trace shows source rate (router ingress)

Note: Blue trace only visible if there is packet loss

Token Bucket Parameter Setting

pcn_tb = 58K bytes

pcn_px = 8,064 bytes

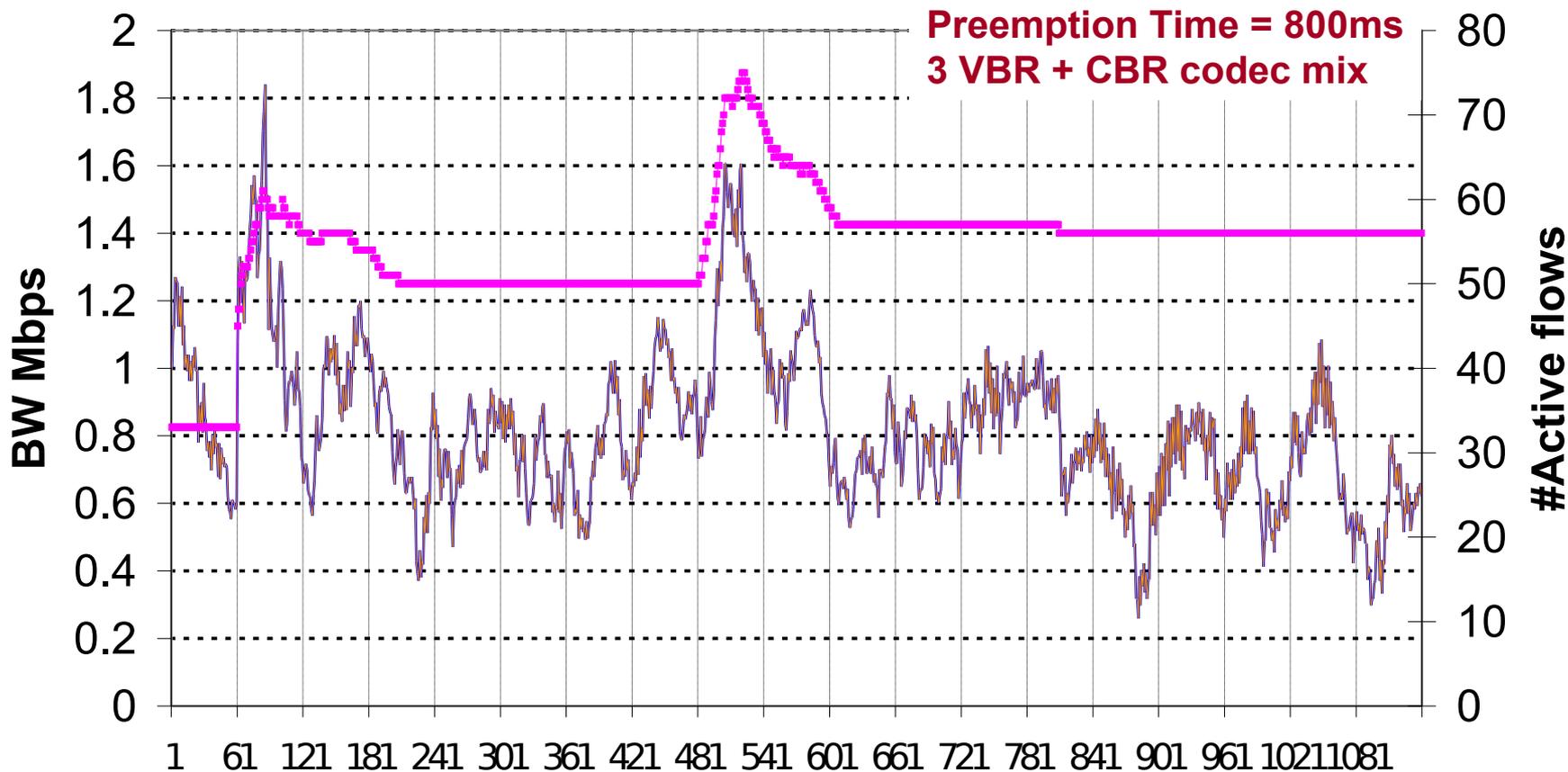
Preemption level = 0.8 Mbps

Service Class BW limit = 2 Mbps

Buffer size = 9,250 bytes

MixCodec/VBR/0-Idle/0.8Mbps/800msRTT_2.5BW/sBF

— Source rate — Transmitted rate — #Active flows



Observation time interval = 0.05 sec

Simulation # JB318

Simulation Results (40 Mbps)

- Large number of flows
 - Preemption level = 40 Mbps
 - **With packet loss**
 - **Service rate limited to 48Mbps and 40 Mbps (service BW = Preemption level)**
 - Preemption times: 200ms and 800ms
 - Preemption time = RTT + processing time in endpoints/GWs
 - Excess rate marking every “n” bytes
 - pcn_px = 8,064 bytes
 - Token bucket size
 - pcn_tb = 58K bytes
 - Graph results
 - Violet trace shows number of flows
 - Orange trace shows transmitted rate (router egress)
 - Blue trace shows source rate (router ingress)
- Note:** Blue trace only visible if there is packet loss

Token Bucket Parameter Setting

pcn_tb = 58K bytes

pcn_px = 8,064 bytes

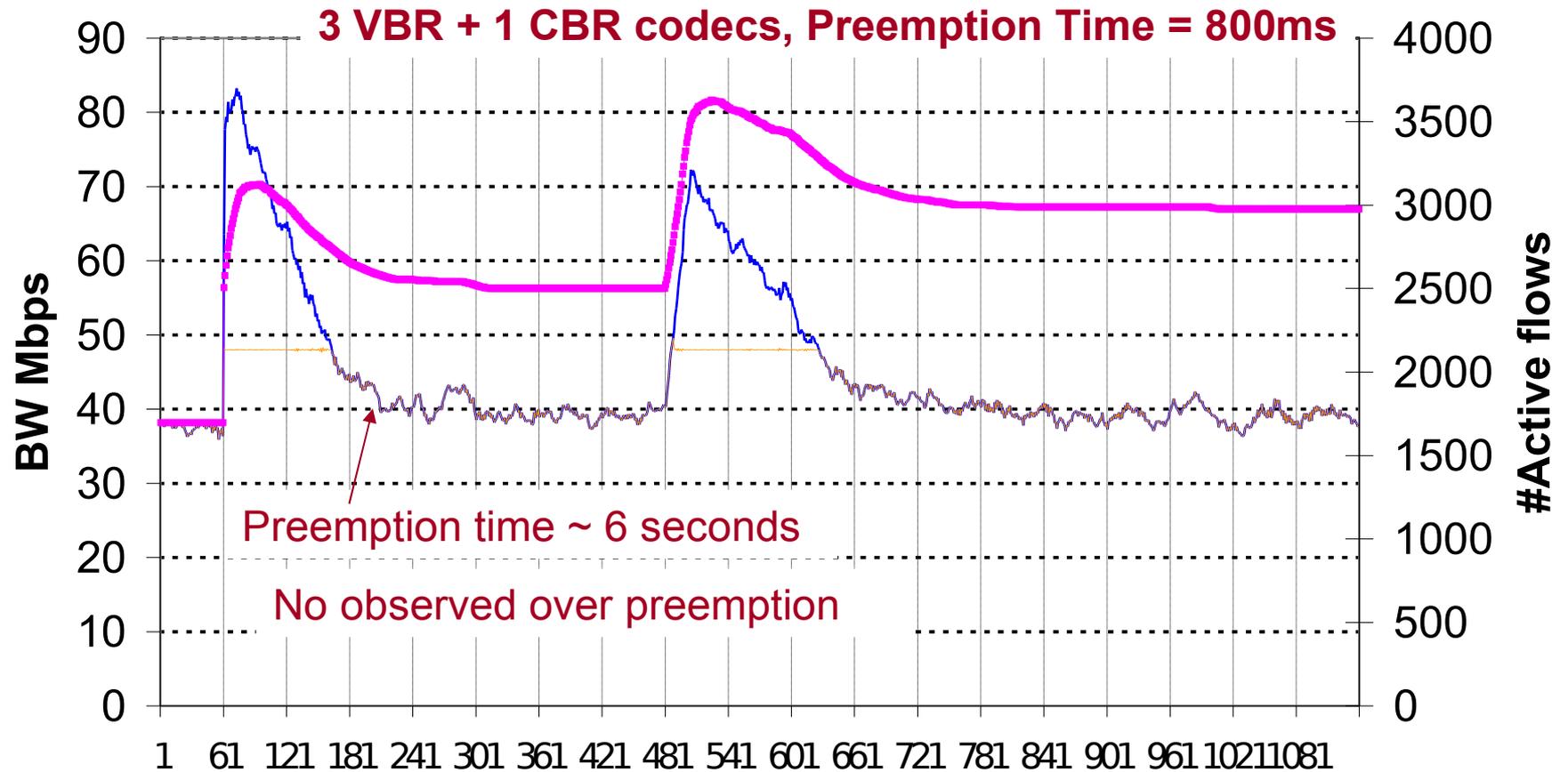
Preemption level = 40Mbps

Service Class BW limit = 48 Mbps

Buffer size = 18,500 bytes

MixCodec/VBR/0-Idle/40Mbps/800msRTT_1.2BW/IBF

— Source rate — Transmitted rate — #Active flows



Observation time interval = 0.05 sec

Simulation # JB309

Token Bucket Parameter Setting

pcn_tb = 58K bytes

pcn_px = 8,064 bytes

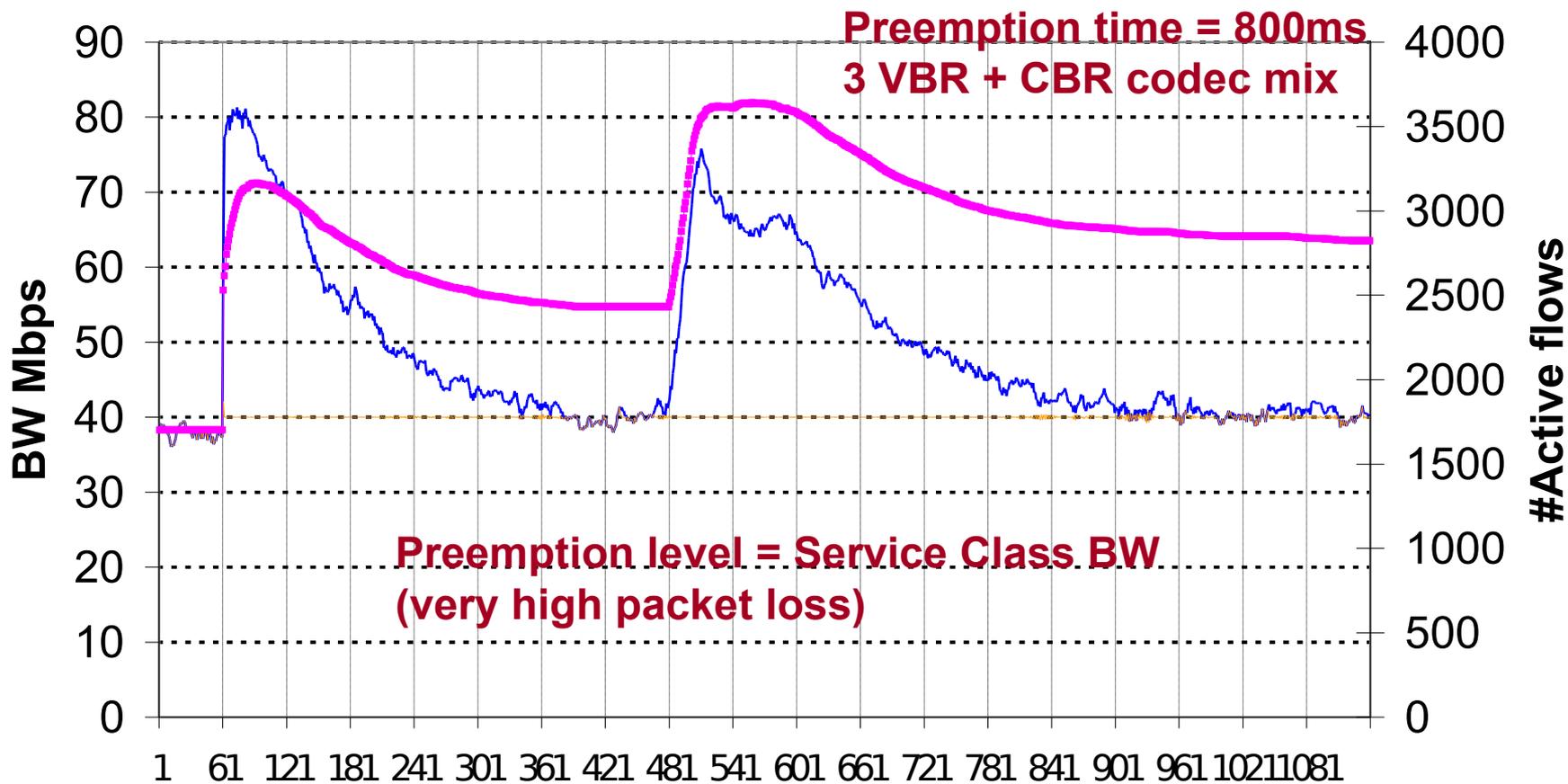
Preemption level = 40Mbps

Service Class BW limit = 40 Mbps

Buffer size = 18,500 bytes

MixCodec/VBR/0-Idle/40Mbps/800msRTT_1BW/IBF

— Source rate — Transmitted rate — #Active flows

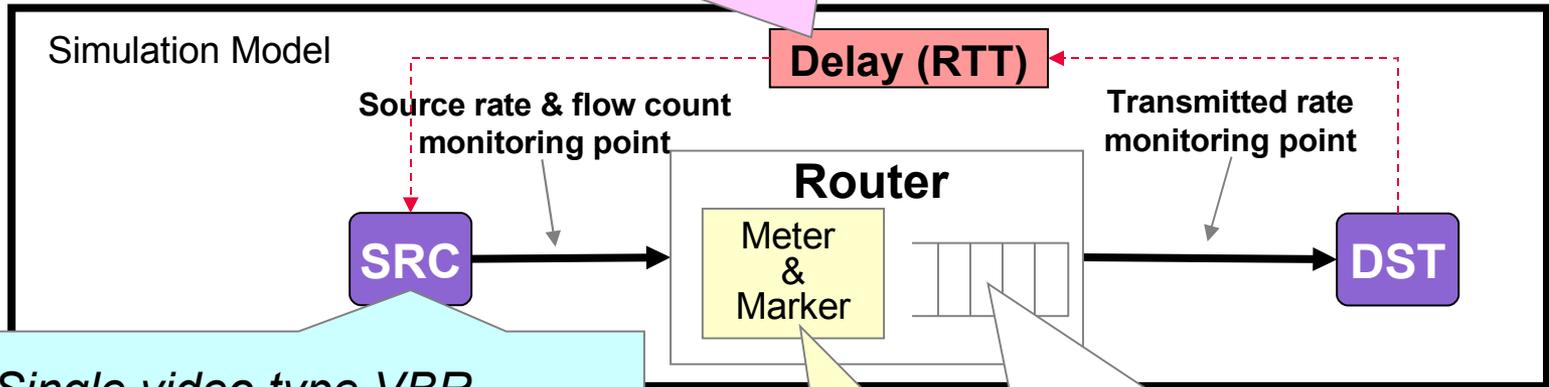


Observation time interval = 0.05 sec

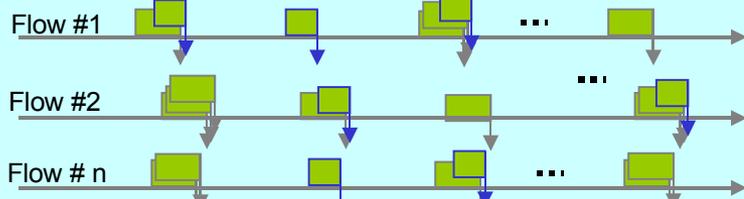
Simulation # JB319

Flow Preemption Simulation Setup for Video

“PM” Flag with Delay = RTT (50, 200, 800 ms)



Single video type VBR



A random number of packets per frame (x MTU sized + 0 or 1 small packet holding remainder bits)

Constant interval between frames

Forwarding Rate and Queue Size

Rate: 2 times preemption rate
Queue: large (unlimited)

Preemption Thresholds: BW in Mbps (4, 40, 200) corresponding to 3 settings of flows: (10, 100, 500)

Token bucket = pcn_tb

Marking interval = pcn_px

(the same token bucket algorithm as for voice)

See next page for video source characteristics

See page 3 for other explanation

Token Bucket Parameter Setting

pcn_tb = 400K bytes

pcn_px = 200K bytes

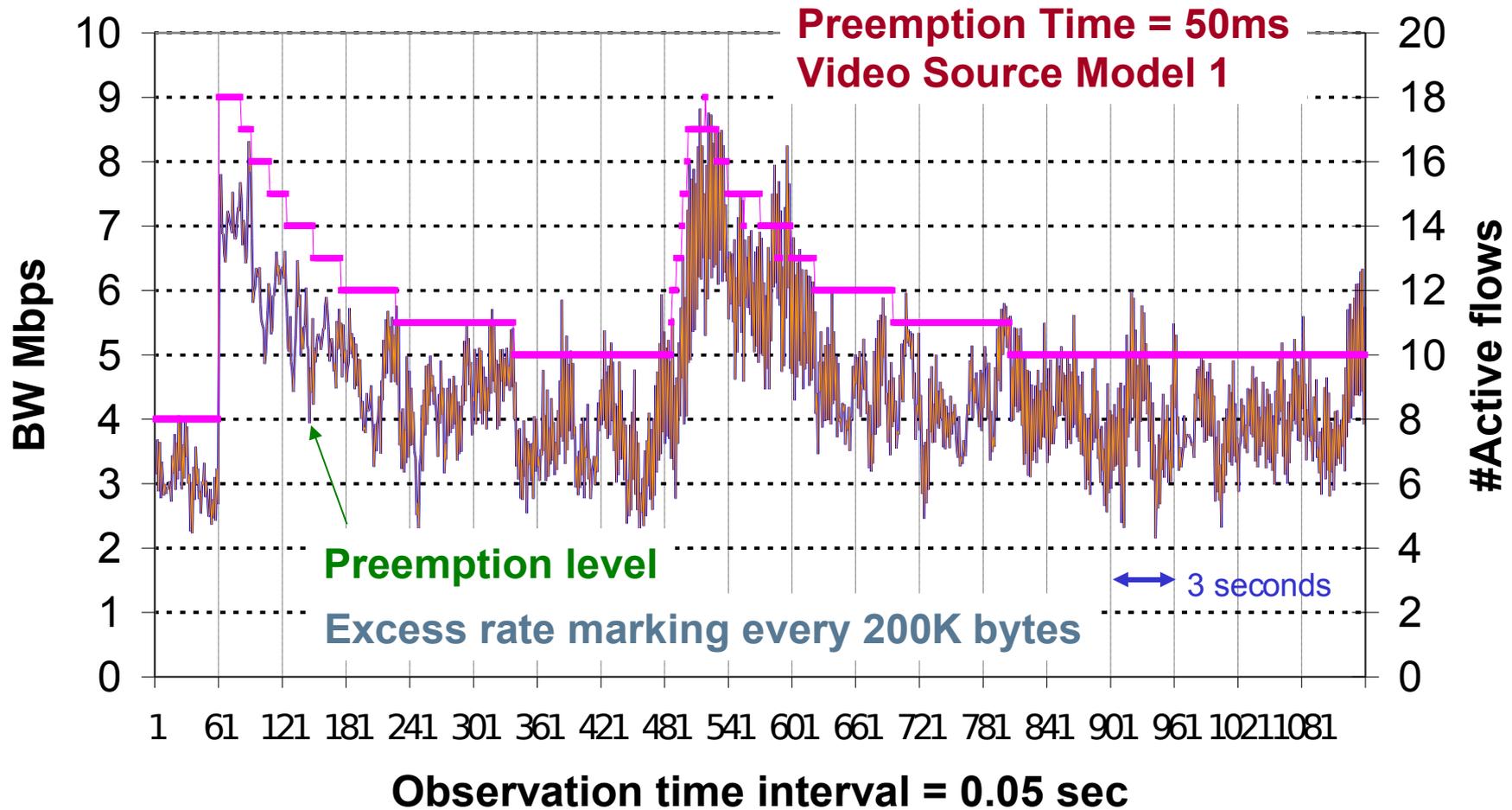
Preemption level = 4 Mbps

Service Class BW limit = 8 Mbps

Buffer size = unlimited

Video/0-Idle/4Mbps/50msRTT_2BW/IBF

— Source rate — Transmitted rate — #Active flows



Token Bucket Parameter Setting

pcn_tb = 400K bytes

pcn_px = 200K bytes

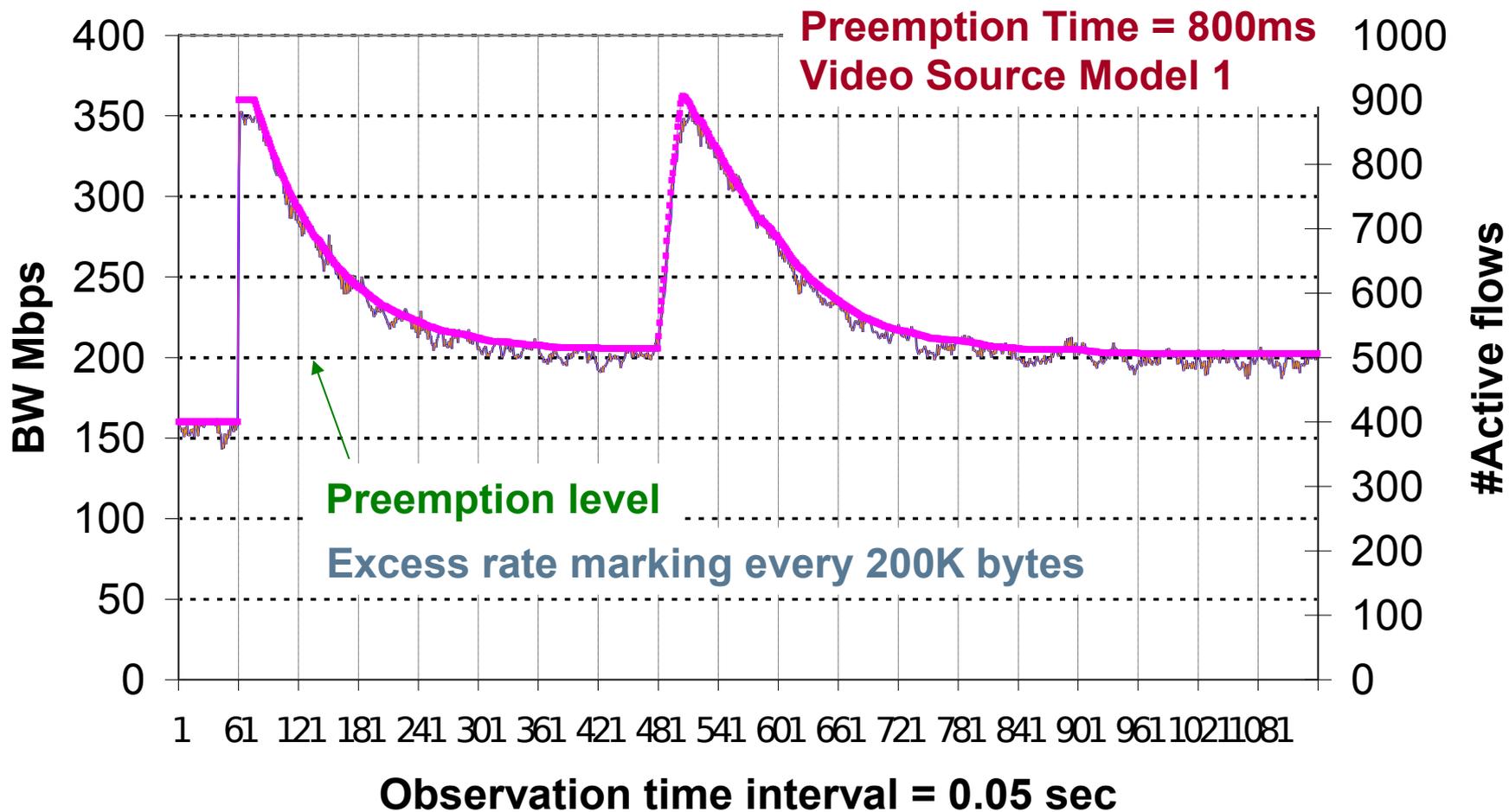
Preemption level = 200 Mbps

Service Class BW limit = 400 Mbps

Buffer size = unlimited

Video/0-Idle/200Mbps/800msRTT_2BW/IBF

— Source rate — Transmitted rate — #Active flows



Characteristics and Benefits of the Proposed Marking

- Works with ECMP (without additional complexity in the gateways). The marked packet belongs to a flow that was routed through congested router.
- Works in presence of packet loss
- Works with small and large number of flows
- Works with variable rate, constant rate, on/off and different packet sizes traffic
- Works well over wide range of flow termination times (RTTs)
- **Stable behavior under operational conditions.** Simulations show very good accuracy both for CBR and VBR traffic with small and large number of flows.
- Works with unidirectional and bi-directional flows (without additional complexity in the gateways). Because this approach has exponential decay property for marking packets.
- Works reasonably well in presence of multiple congestion points
- Works in gateway-to-gateway and host-to-host deployment models.
- Friendly to behavior of ECN (RFC 3168)
- In gateway-to-gateway deployment model, **two approaches** can be used, (1) egress gateway signals to ingress gateway what flows to terminate, (2) As well, the egress gateway computes excess rate and signals excess rate to ingress gateway.
- If the egress router reports only flow termination, than any reasonable value for “x” can be used. The value for “x” and the algorithm **do not** need to be standardized, but only the metering and marking behavior. Different algorithms can be used to obtain the described metering and marking behavior.

Next Steps

- Next Steps
 - Simulation of multiple congestion points
 - Provide guidelines for configuration
- Questions