DOCSIS-PIE
draft-white-aqm-docsis-pie-00

Greg White, CableLabs

IETF89 - ICCRG

DOCSIS 3.0

- Specifications published in 2006
- Widely deployed worldwide
- Scalable capacity
  - Early modems: up to ~140M* down x ~100M up
  - Current modems: up to ~560M* down x ~100M up
  - Future: up to ~1.1G* down x ~200M up

* In 8 MHz systems, multiply by 1.4

DOCSIS 3.1

- Specifications published in 2013
- Products expected in 2015
- Scalable capacity
  - Early products: up to 4G down x 1G up
  - Later products: up to 10G down x 1G up

Service Flows

- Each cable modem customer can be configured with multiple, unidirectional Service Flows to segregate traffic for different services
- Each Service Flow consists of:
  - A queue
  - Quality of Service parameters
    - e.g. rate shaping, priority, specialized scheduling, reserved rates
  - Classifiers
    - Matching on Ethernet, IPv4/IPv6, TCP/UDP header fields
- Service Flows are (largely) independent from one another
- D3.0 Modems typically support 16 upstream and at least 16 downstream Service Flows
  - No hierarchical QoS
- D3.1 Modems will support 32 upstream service flows
  - Hierarchical QoS supported

Rate shaping

- Each Service Flow is rate shaped using a dual-token bucket:

  TxBytes(t1,t2) <= (t2-t1)*R/8 + B

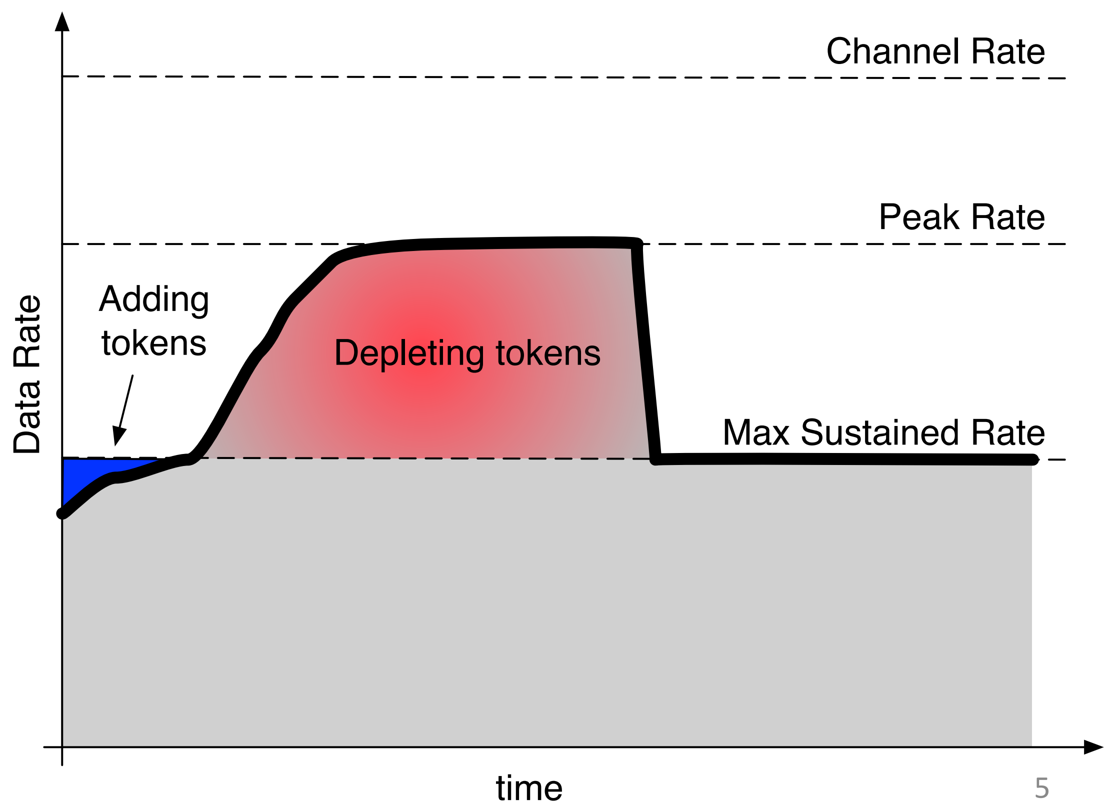
  TxBytes(t1,t2) <= (t2-t1)*P/8 + 1522

for all values t2>t1

R = Maximum Sustained Traffic Rate (bps)

P = Peak Traffic Rate (bps)

B = Maximum Traffic Burst (bytes)



Upstream Request-Grant MAC

- Upstream channel scheduling is driven by "MAP" Intervals (typ. 2ms)

- Packet(s) arrive at the CM
- CM waits* for the next contention request opportunity
  - *typically less than 2ms
- CM sends request message (subject to rate shaping)
- CMTS scheduler collects requests, then schedules and communicates future transmit opportunities (grants)
- Due to serialization, propagation and interleaver delays, as well as CMTS/CM processing delays, grant occurs 2 MAP Intervals after the request was sent

- Without congestion, typically 4-8ms access latency

CM = cable modem CMTS = cable modem termination system

New DOCSIS AQM Requirements - Summary

- DOCSIS 3.1
  - CM
    - MUST implement DOCSIS-PIE
    - MAY implement other algorithms
    - AQM MUST operate independently on each SF queue
  - CMTS
    - MUST implement AQM
    - AQM MUST operate independently on each SF queue
- DOCSIS 3.0
  - CM
    - SHOULD implement AQM
      - If so, MUST implement DOCSIS-PIE
      - MAY implement other algorithms
      - AQM MUST operate independently on each SF queue
  - CMTS
    - SHOULD implement AQM
      - AQM MUST operate independently on each SF queue

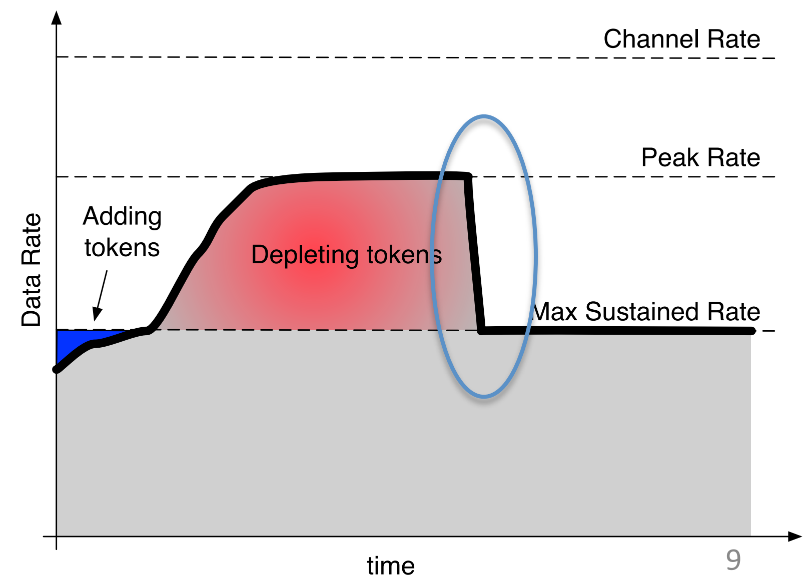
Deltas from PIE*

*draft-pan-aqm-pie-01

- Really DOCSIS-Specific Deltas
  - Departure rate estimation
  - Trigger for exponential decay
  - 16ms update interval (vs. 15ms)
- Not-so DOCSIS-Specific Deltas
  - Packet Drop De-randomization
  - Protections on packet-size scaling
  - Enhanced burst protection
  - Expanded auto-tuning range
  - Configurable latency target (10ms default)

Departure Rate Estimation

- For variable rate links, PIE tracks egress rate via a dequeue rate estimator
  - Estimates future rate based on smoothed observations of dequeue rate
- DOCSIS service flow rates are variable, but most extreme variations come from the rate shaper, not the link
  - i.e. transitions from peak rate to max sustained rate
- DOCSIS-PIE uses the state and configuration of the rate shaper to predict egress rate



Packet Drop De-randomization

- CoDel packet drops are (fairly) deterministic, scheduled at time ≅ Interval/sqrt(count)
- PIE packet drops are stochastic, triggered by a iid Bernoulli R.V. using calculated drop probability
  - iid RV results in localized excursions from desired drop rate
  - Single TCP performance is sensitive to excessive loss
- Hybrid approach avoids "unlucky" coin tosses
  - Reduces variance of the RV

Packet Drop De-randomization Detail

- Simple algorithm to enforce bounds on run-length
  - Min. packets between drops = ceil((0.85/p)-1)
    - i.e. after a packet drop, suppress drops for the next several packets.
    - Results in drop probability ~ 0.54*p for p<0.1
  - Max packets between drops = ceil((8.5/p)-1)
    - If a large number of packets have passed the coin toss, force a packet drop
    - These are pretty rare events

Why no FQ-*?

- Investigated FQ-CoDel & FQ-PIE
- Hardware complexity of 32 Service Flows x 32 queues
  – Or, operational complexity of Service Flows sharing a pool of N queues
- Tight deadlines between MAP & grant
  – *any* additional processing at dequeue time is hard
- Limited additional benefit compared to single queue AQM at 100Mbps+
- Concerns about VPN traffic
- Hash collisions – not feasible to have 1024 queues (see above)

Why no ECN?

- In many modem designs, packet pipeline is implemented in hardware.
- With current ECN:
  - Not sufficiently convinced of its benefits
  - Open questions on differential treatment of ECN/ non-ECN traffic
- Discussions on a new ECN seem promising

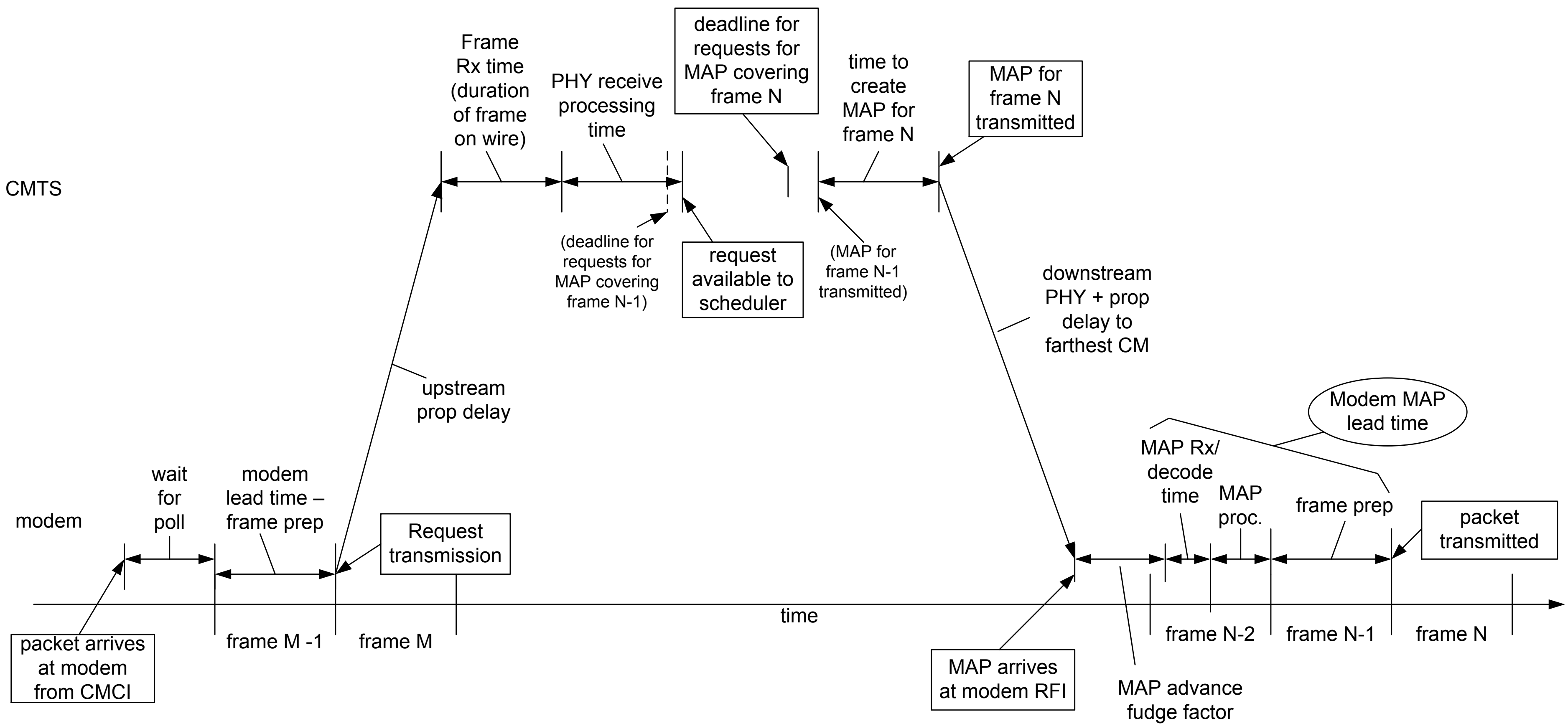
draft-white-aqm-docsis-pie-00

g.white@cablelabs.com

ropan@cisco.com

BACKUP SLIDES

Request-Grant Process for D3.1



Typical SF Usage by Cable Operators

- Residential HSD service uses a single service flow pair
  - No Reserved Rate guarantee
  - Max Sustained Rate set to offered service rate
  - Peak Rate set to 1.5x or 2x MSR
  - Max Burst set to ~10 MB
- Other services run on separate service flows
  - Digital voice service
  - Community WiFi

ns2.36

- Working with Tom Henderson, Rong Pan and Kathie Nichols to include:
  - DOCSIS link
  - PIE, CoDel, CoDel-DT queues
  - SFQ-CoDel, SFQ-PIE queues
  - web page model