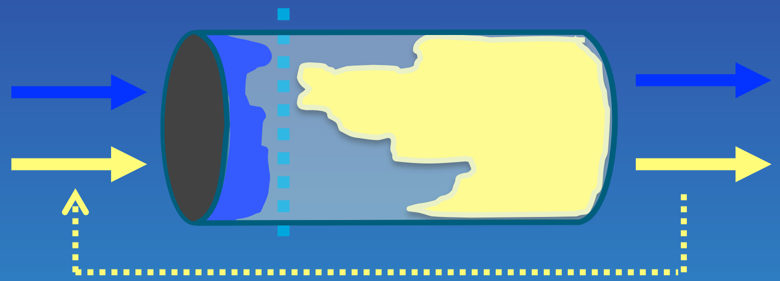
PIE: A lightweight latency control to address the bufferbloat problem

Rong Pan, Preethi Natarajan, Chiara Piglione, Mythili S. Prabhu, Vijay Subramanian, Fred Baker and Bill Ver Steeg

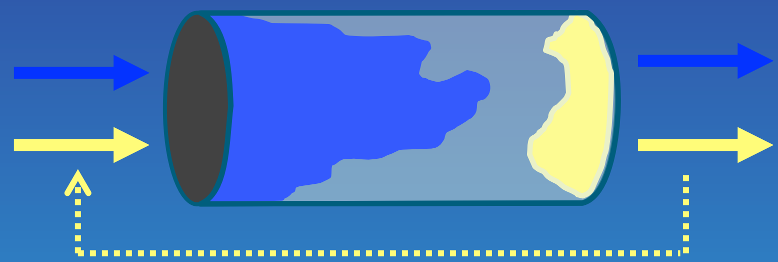
November 5, 2013, IETF 88

Current Design



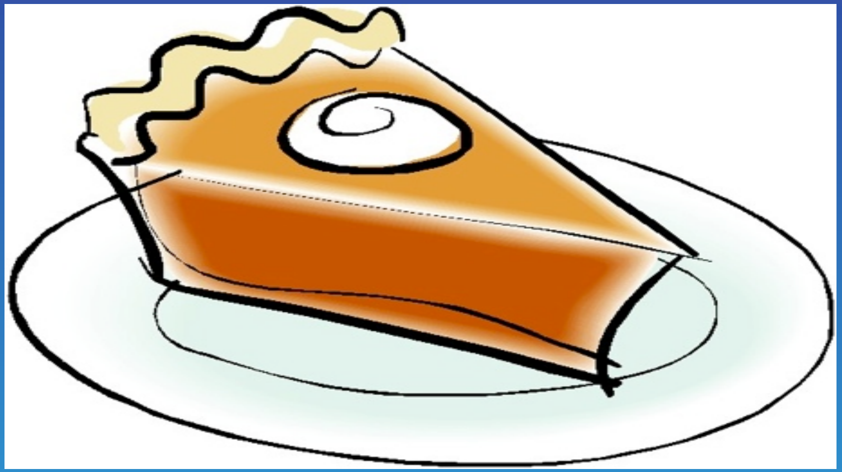
- Large TCP flows occupy most buffer
- Feedback signals are sent when buffer occupancy is big
- Average delay is consistently long
- Little room left for sudden burst

Future Goal



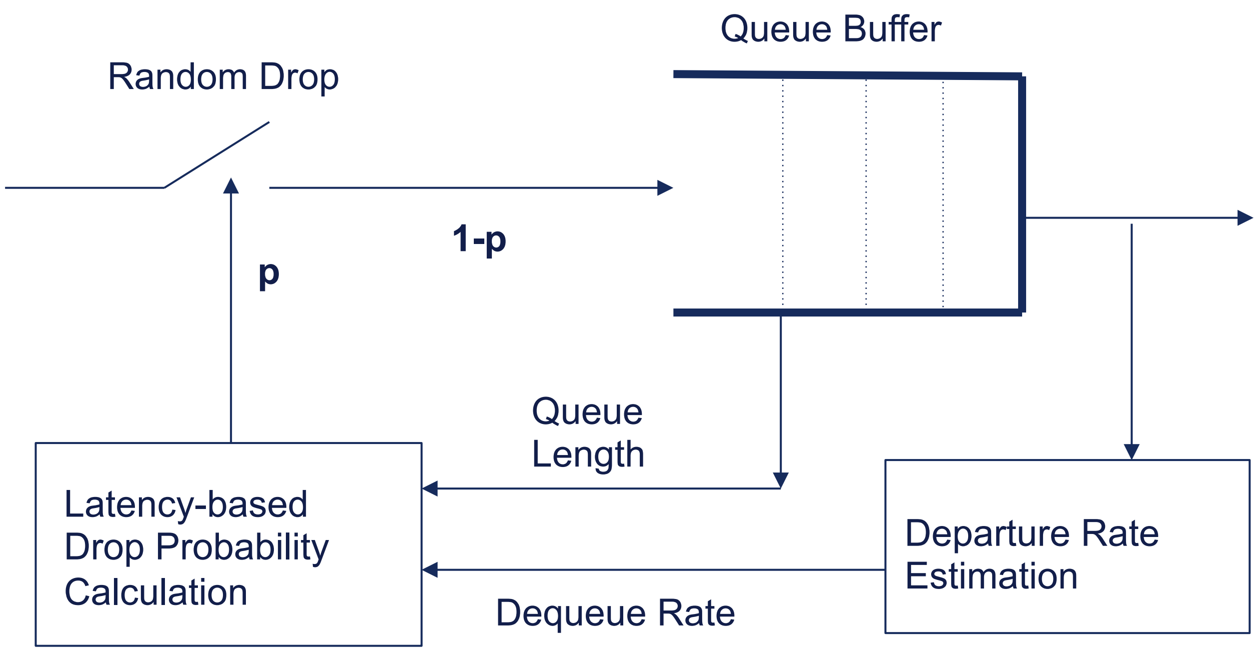
- Large TCP flows occupy small buffer
- Feedback signals are sent early
- Average delay is kept low
- Much room left for sudden burst

Solution Maybe



As Easy As PIE!

The block diagram of PIE



The design of PIE

a and b are chosen via control analysis

- Upon every packet arrival
  - randomly drop a packet based on drop_prob calculated below
- Every T_{update} interval
  - estimated_delay, est_del = queue_length/depart_rate
  - drop_prob += a*(est_del – target_delay) + b* (est_del – est_del_old)
  - est_del_old = est_del;
  - depart_count = 0;
- In a measurement cycle
  - Upon a packet's departure: depart_count += deque_packet_size;
  - if dq_count > deq_threshold then
  - depart_rate = deqart_count/(now-start);
  - dq_count = 0; start = now;

PIE Work Update

- Actively participate in the DOCSIS 3.1 AQM working group with colleagues from Arris, Broadcom, CableLabs, Cisco, Comcast Cable, Huawei, Intel, ST Micro, Timer Warner Cable, etc.
- Weekly meetings from April, 2013 – Now: discuss evaluation scenarios, design trade offs, application impacts, etc. taking into the considerations from MSOs, system and silicon vendors
- The working group has chosen a single queue based design instead of multi-queue based design such as FQ
  - design complexity does not justify the performance gain
- The working group has chosen PIE as the required, default-on AQM scheme for Cable Modems per DOCSIS 3.1 spec (vendors can implement additional AQM schemes)
  - Implementation simplicity, convergence performance, and readiness for future applications

Variant of PIE in DOCSIS

- Departure rate estimation
  - directly use MAC layer's information, i.e. Peak Traffic Rate and Max Sustained Traffic Rate directly into the algorithm
- Several constants in the PIE algorithm has been customized to fit the cable networks
- Improved handling of the single TCP flow case
  - extended drop probability calculation to better handle low drop probability scenarios
  - inclusion of detecting burst before protecting

Thank you.