An Evaluation of AQM techniques on Access Networks with Internet Traffic

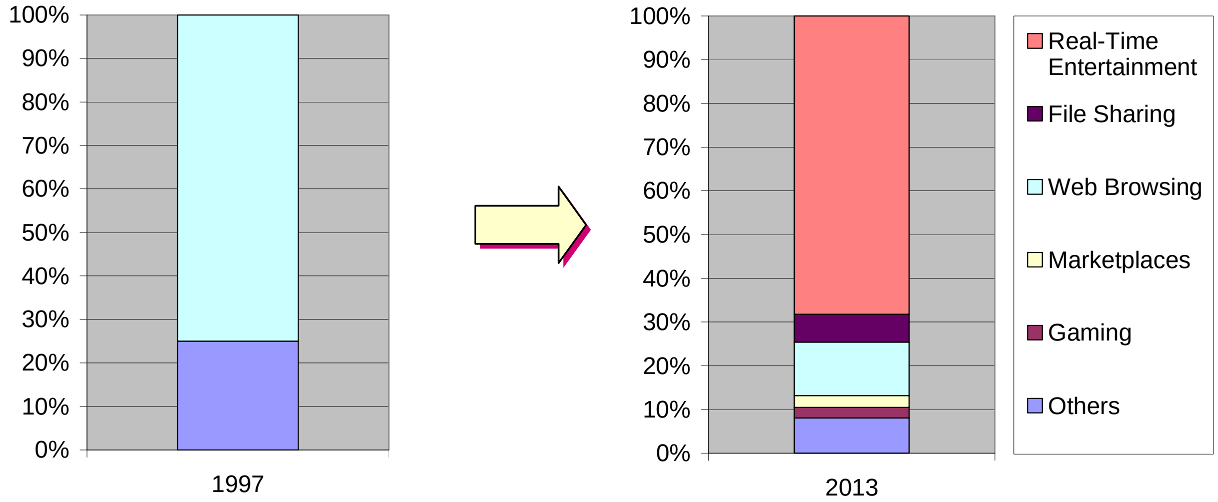
Shahid Akhtar, Steve Benno, Andrea Francini, Dave Robinson, Randy Sharpe

IETF ICCRG - Nov 2013

contact: shahid.akhtar@alcatel-lucent.com

COPYRIGHT © 2013 ALCATEL-LUCENT. ALL RIGHTS RESERVED.

Downstream Internet Traffic Evolution



70-80% Web Traffic
Average TCP < 10KB
(Thompson)

~ 70% RTE
Average TCP > 1MB
(Sandvine)

Real-Time Entertainment (RTE) – key components:

- Progressive Download (PD) (e.g. Youtube)
- HTTP Adaptive Streaming (HAS) (e.g. Netflix)

80% of total in RTE / Web

We Assume a mix of 15%/50%/35% Web/HAS/PD

Internet Traffic Transitioning to Video and Much Longer Lived TCP Flows

Alcatel·Lucent

COPYRIGHT © 2013 ALCATEL-LUCENT. ALL RIGHTS RESERVED.

RED Status, Effectiveness and Tuning

- RED Effectiveness Limited with majority web Traffic
  - RED (Brandauer) produces larger consecutive losses than Tail-Drop under heavy web traffic
  - Web traffic - average GET size 7KB – fits in 5 packets – within an initial burst of 10 packets
  - Can act like UDP traffic from RED/AQM perspective
- New Traffic (68% MM) – long (er) (1MB) lived TCP sessions
  - RED/AQM far more effective
- Tuning Requirement
  - Optimal RED parameters - tuning due to: RTT or Packet Loss Rate (PLR) (related to number of TCP flows)
  - Number of TCP flows limited in access networks
  - RTT variation limited due to – CDNs and Caches
- RED only used in 26% of DSL access lines (Dischinger)

Promise of Much Less RED Tuning and More Effectiveness

S. Akhtar et al, "An Evaluation of AQM techniques on Access Networks with Internet Traffic"

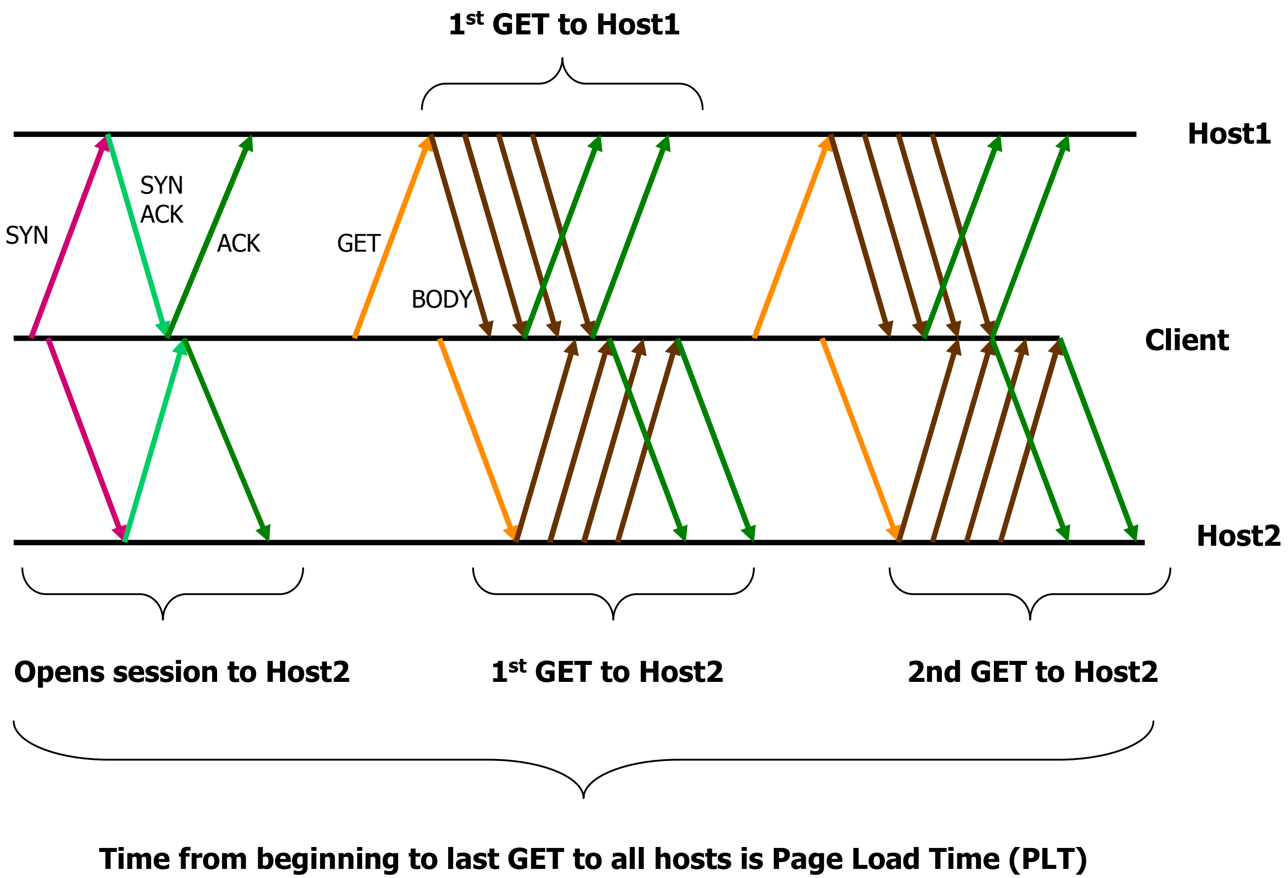
COPYRIGHT © 2013 ALCATEL-LUCENT. ALL RIGHTS RESERVED.

RED Configuration / Parameter Optimization

- Configurations of RED used
  - w_q = 240/(5000xC) – C in Mbps
    - Ratio of queue sampling interval by maximum delay of 250ms
  - Buffer size B – 2xBDP (other sizes also studied)
  - Minimum threshold - 1/6 of B
  - Maximum threshold - B
  - Max-p – Determined through simulations (4%)
- Early work on Optimization
  - (Floyd) – Rules of thumb for parameter setting
  - (Firoiu) - Derived “optimal” parameters - many simplifying assumptions (all long TCP flows, same RTT, no SACK..)
  - (Ziegler) - Uses (Firoiu) for queue weight, derives max_p and difference between min and max threshold
  - (Brandauer)- Evaluates rules of thumb and (Ziegler) based parameters
- CODEL – default: target (5ms), interval (100ms) – used latest ns-2 code (Pollere)

COPYRIGHT © 2013 ALCATEL-LUCENT. ALL RIGHTS RESERVED.

Web Traffic Model – No Pipelining



Open a TCP connection to a host when first GET is seen

Some Browsers open connection at beginning of page anyway

Connection remains open unless TCP timeout timer expires (20s)

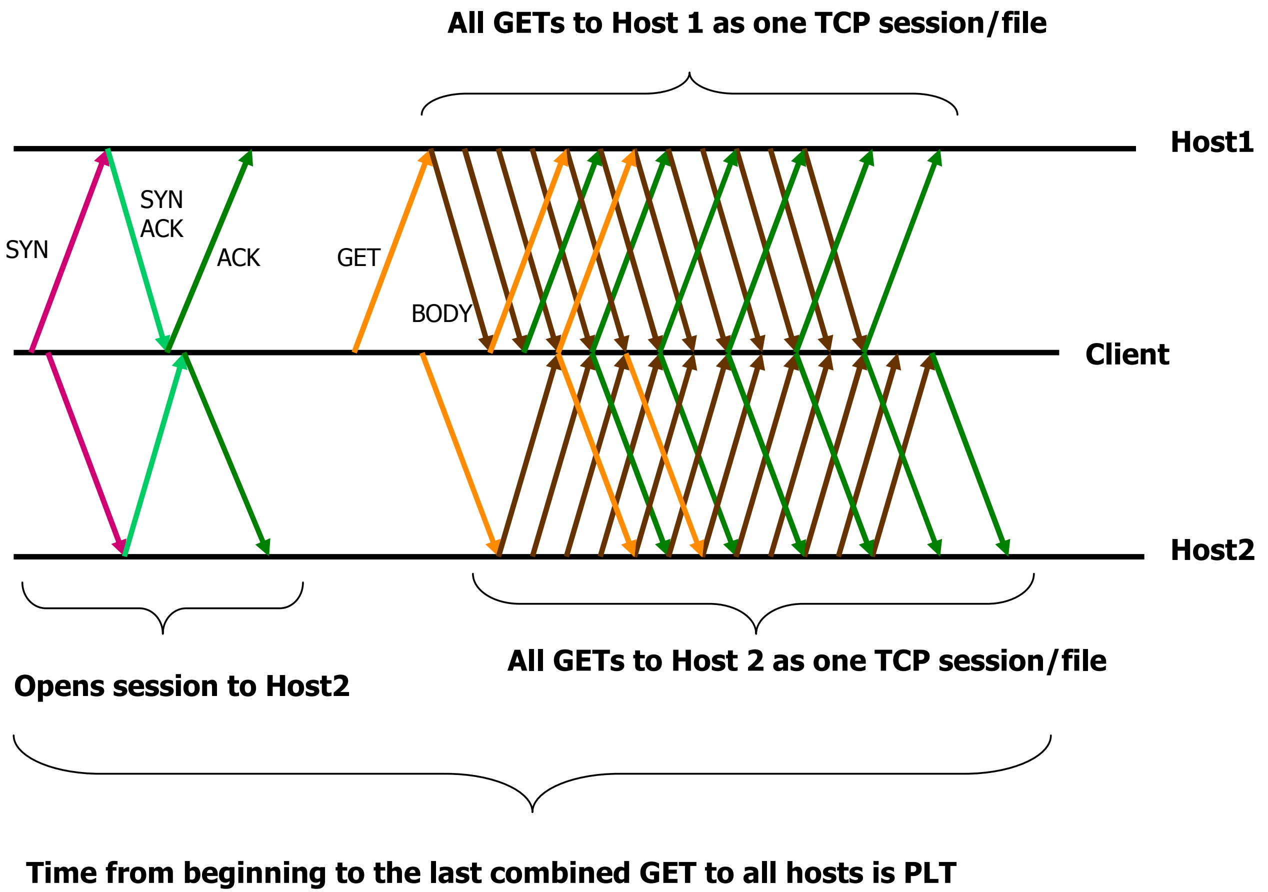
PLT does not include DNS lookup – assume local

The number of GETs and GET sizes based on Google statistics

Single TCP connection per Host with Initial window size of 10

COPYRIGHT © 2013 ALCATEL-LUCENT. ALL RIGHTS RESERVED.

Web Traffic Model – Full Pipelining



Pipelining bunches responses of multiple GETs

Appears to TCP as one file transfer

Assuming processing time at client much smaller than RTT or download time

Full pipelining is part of HTTP 1.1, but expected to be used fully with HTTP 2.0

COPYRIGHT © 2013 ALCATEL-LUCENT. ALL RIGHTS RESERVED.

HAS and PD Traffic Modeling Assumptions

- Progressive Download Assumptions
  - 480p Video assumed
  - 270s average play time for a PD (Youtube) video clip
  - 1.128Mb/s as the average speed needed for 480p play (1Mb/s for Video 128 kb/s for audio (Wiki-Youtube))
  - 1.128Mb/s and 270s, yields about 38MB as the average file for PD
    - Pareto distribution assumed with mean 38KB and stdev/mean ratio of 1.5
  - Inter-request time depends on load, but is Pareto based with stdev/mean ratio of 2

- HAS Traffic Assumptions
  - HAS Source – 2s chunks
  - MSFT Big Buck Bunny Video and commercial encoder to determine exact size of chunks
  - Each HAS source starts 10s apart in the simulation
  - HAS Client
    - 7 VQ Levels (0.72, 1, 1.5, 2.0, 2.9, 3.9, 5.9 Mbps)
    - Bell Labs HAS Client (Benno)

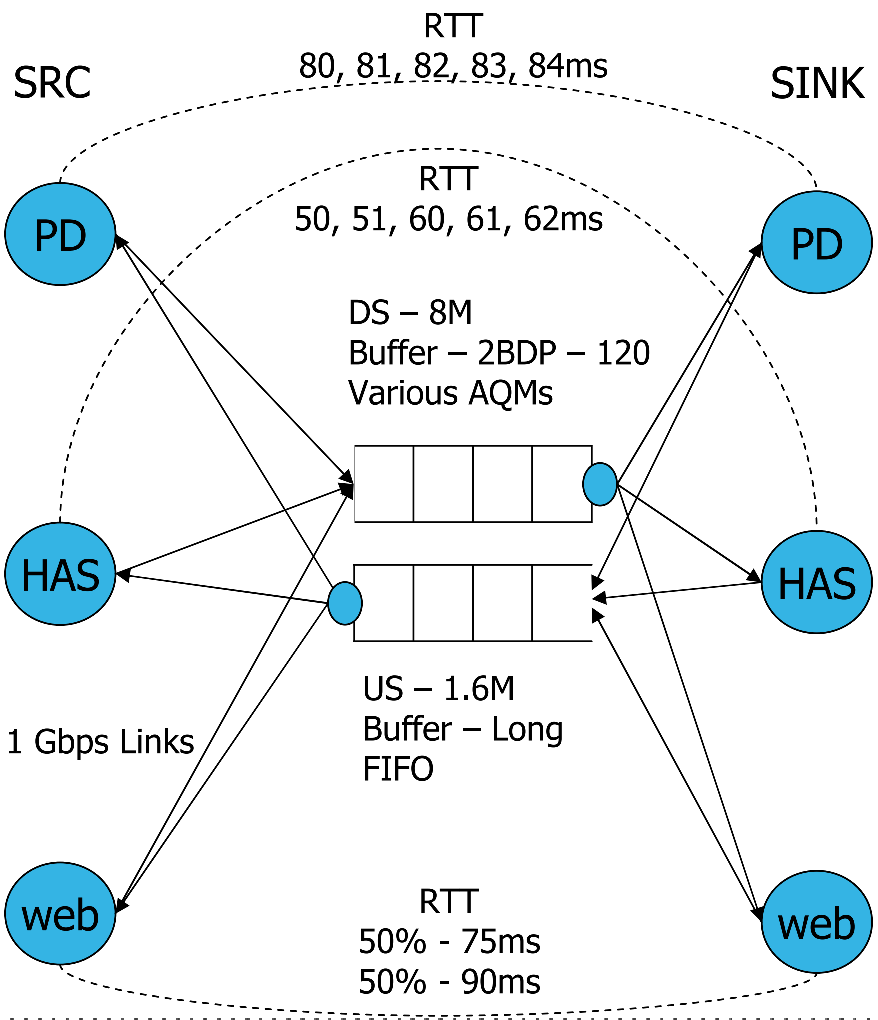
Alcatel·Lucent
COPYRIGHT © 2013 ALCATEL-LUCENT. ALL RIGHTS RESERVED.

TCP Modeling

- Tcp_full module in ns-2
- TCP New Reno
- SACK turned on
- Segmentation offload or receive offload not used
- Nagle's Algorithm - turned off – minimizing latency
- 10 segment initial window used – as per Linux implementation
- Idle restart based on RFC 2861 as per Linux
  - At idle restart, after multiple RTO expiry, the restart window is set to 10
  - Current RFC 2861 based implementations bring this window to 1
    - Some controversy associated with this implementation
  - Unless the window at the end of the previous burst in the same TCP session was less than 10, then that value is used
- Minimum RTO was 200ms as per Linux
- Unless specified, all sources use this TCP – HAS, web and PD

COPYRIGHT © 2013 ALCATEL-LUCENT. ALL RIGHTS RESERVED.

Network Topology



RTT
HAS lowest RTT – like operator CDN
PD – assumed OTT, higher than HAS
Web – centered on two RTTs

Bottleneck - 8M DSL
Even with VDSL, the HSIA component is often small (OOKLA)

Core Network
Assumed non-congested
Assumed majority of packet loss on access links
Links do not become bottleneck
Assumed 1Gbps or above

Home Network
Assumed faster than access bottleneck
802.11n – 100-150Mbps

S. Akhtar et al, "An Evaluation of AQM techniques on Access Networks with Internet Traffic"

COPYRIGHT © 2013 ALCATEL-LUCENT. ALL RIGHTS RESERVED.

Traffic Loads

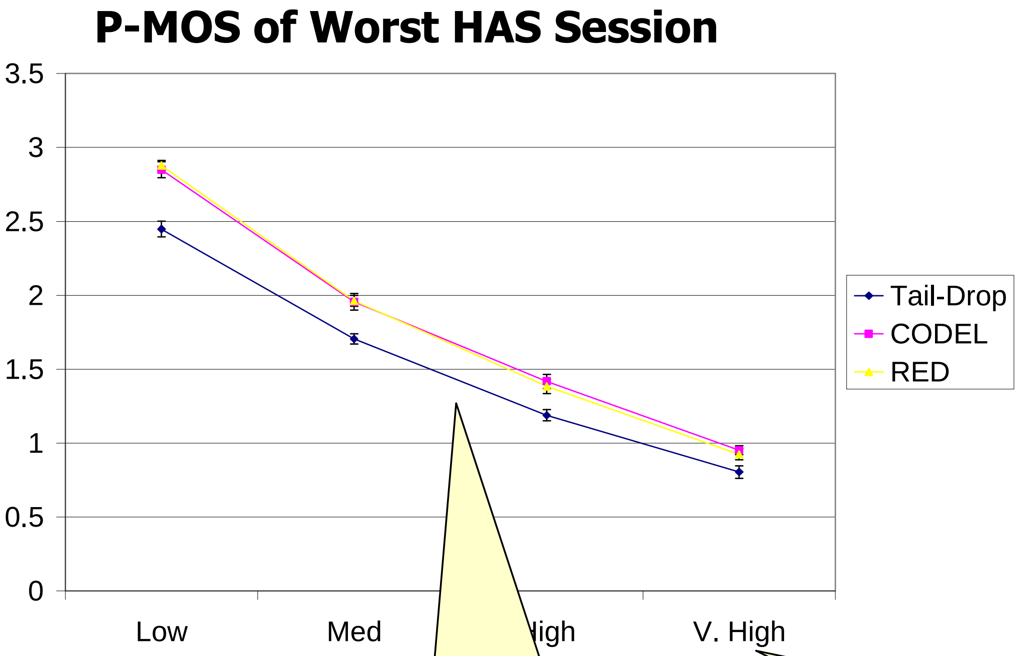
	HAS Traffic	Web Traffic (see web model for details)	Progressive Download Traffic (see PD model for details)
Low Traffic	2 HD HAS Sessions	One 650KB page every 8.5s	Two 38MB files every 350s
Medium Traffic	3 HD HAS Sessions	Two 650KB pages every 8.5s	Three 38MB files every 350s
High Traffic	4 HD HAS Sessions	Two 650KB pages every 8.5s	Four 38MB files every 350s
Very High Traffic	5 HD HAS Sessions	Three 650KB pages every 8.5s	Five 38MB files every 350s

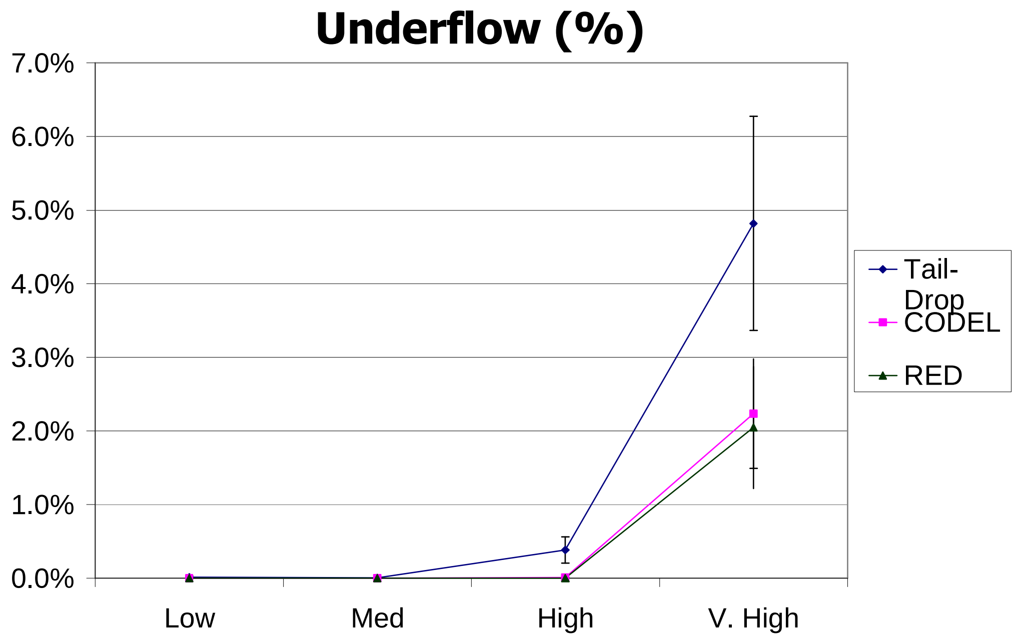
This Traffic amount Used for most Comparisons (Buffer size ..)

This Traffic produced too low QoE for typical users

COPYRIGHT © 2013 ALCATEL-LUCENT. ALL RIGHTS RESERVED.

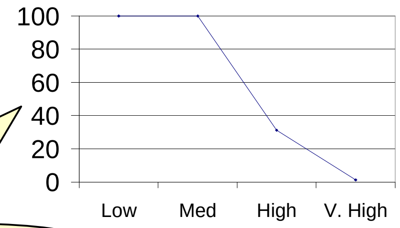
HAS QoE – P-MOS and Underflow Time





Benefit of AQM throughout Traffic Range

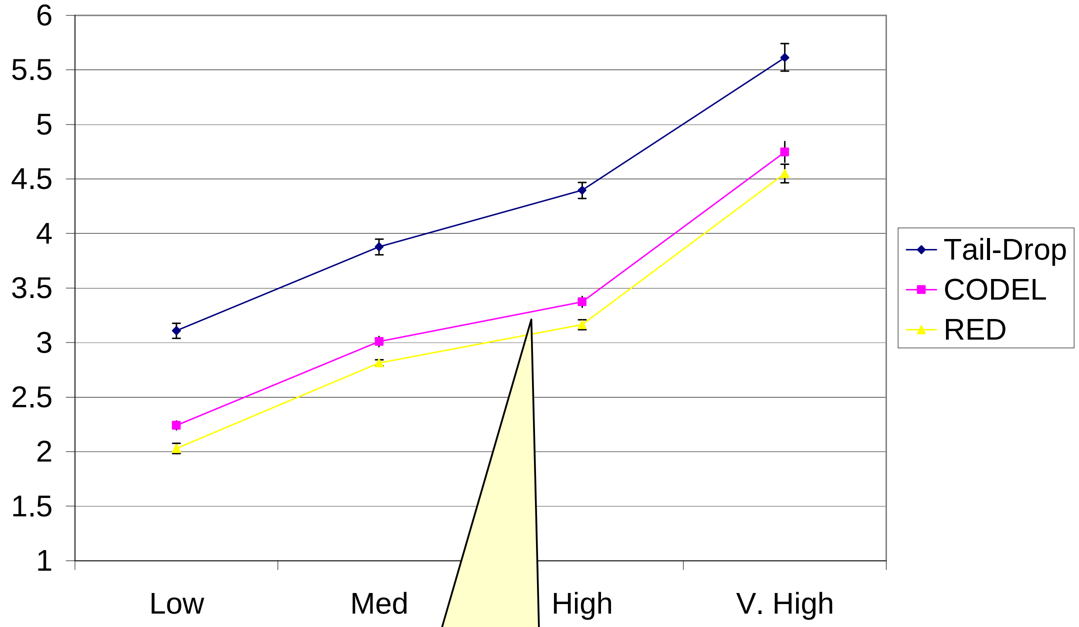
This Traffic produces PMOS – too low + high underflow = quality too low
Cannot run at this traffic level

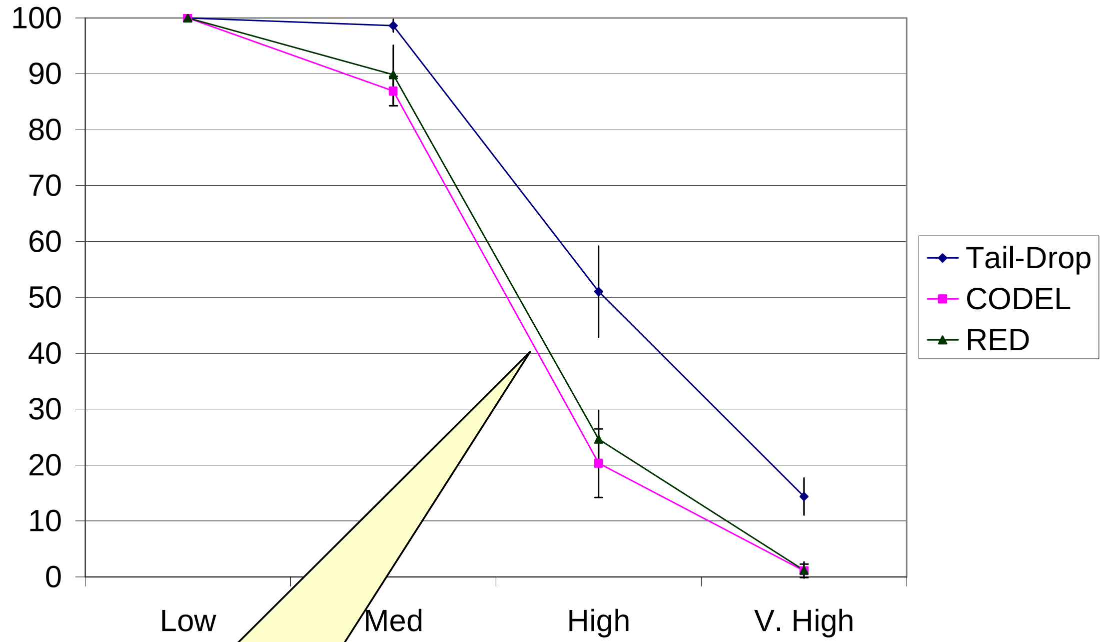


(Underflow (TD) – CI)/(Underflow (RED) + CI)

Alcatel·Lucent
COPYRIGHT © 2013 ALCATEL-LUCENT. ALL RIGHTS RESERVED.

Web and PD QoE - Page Load Time (PLT) and % at 480p



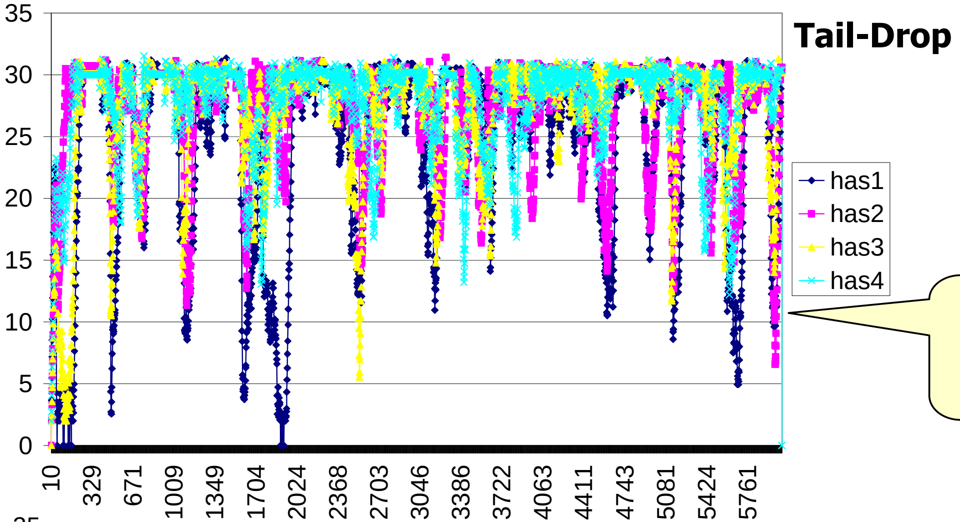


AQM reduce queue size – which reduces RTT – which reduces PLT

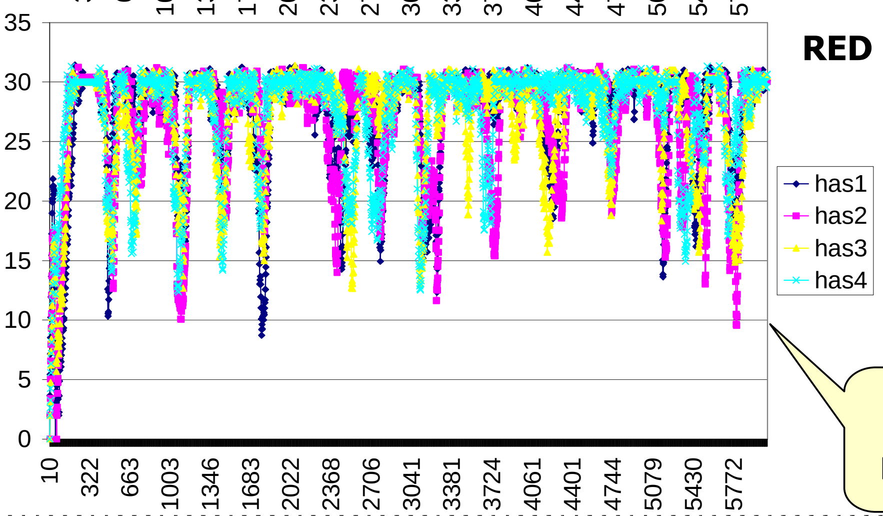
AQMs induce more fairness – which results is lower PD QoE

COPYRIGHT © 2013 ALCATEL-LUCENT. ALL RIGHTS RESERVED.

Tail-Drop and RED – Details – Client Buffer/Network Buffer

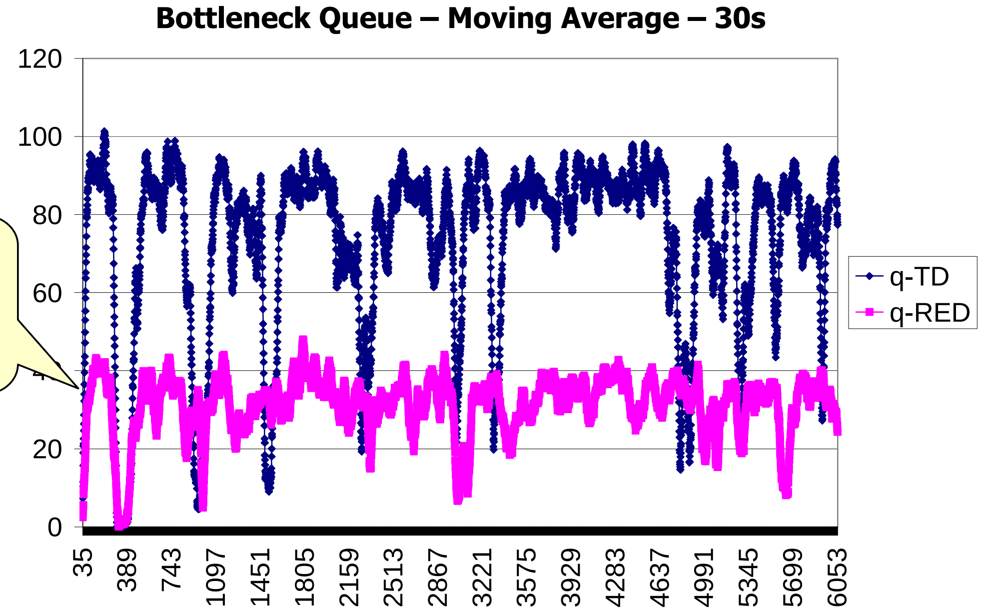


Many drops below 10



Virtually no drops below 10

RED able to drop buffer occupancy significantly



S. Akhtar et al, "An Evaluation of AQM techniques on Access Networks with Internet Traffic"

Alcatel·Lucent

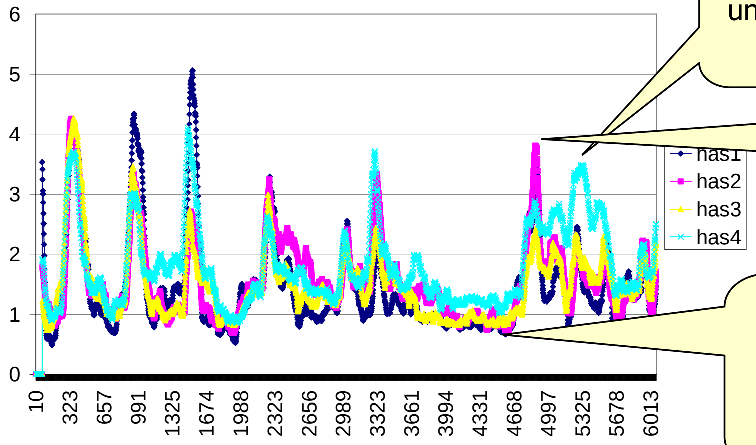
COPYRIGHT © 2013 ALCATEL-LUCENT. ALL RIGHTS RESERVED.

Tail-Drop and RED – Details – HAS Throughput and VQ Levels

HAS session Throughput – MA – 30s - Mbps

HAS session VQ Level - Mbps

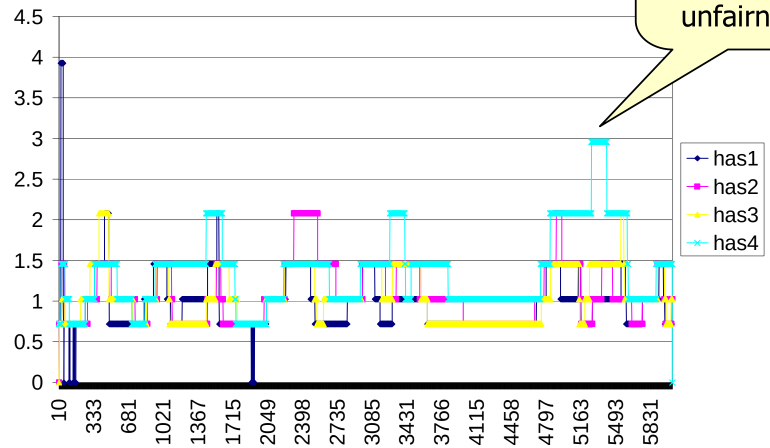
Tail-Drop



Significant unfairness

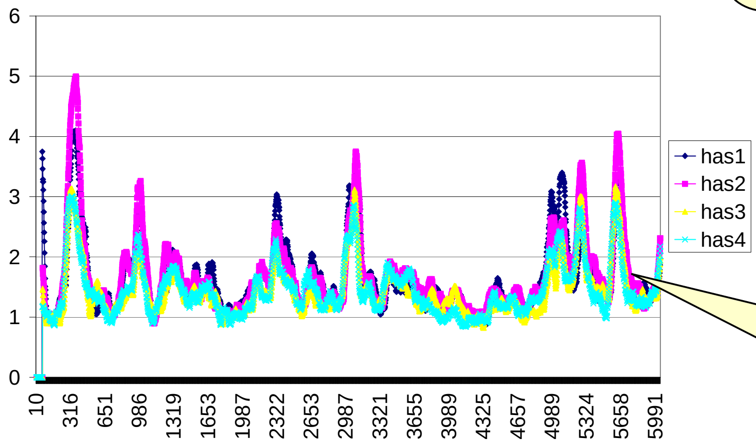
Peaks due to absence of PD sessions

Throughput drops further when several PD sessions active

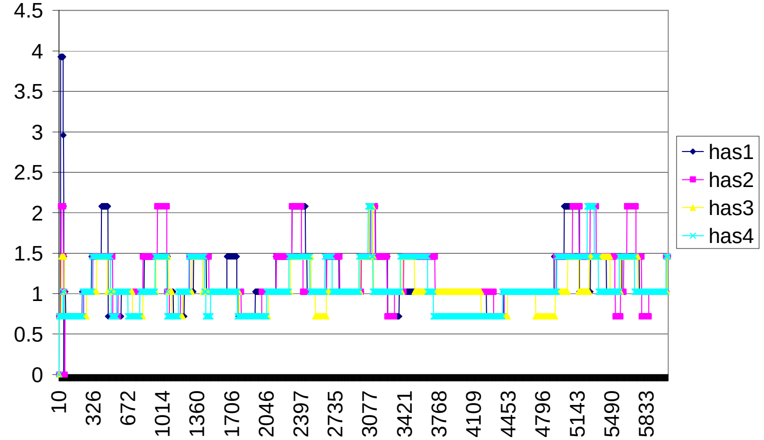


Throughput unfairness creates VQ unfairness

RED

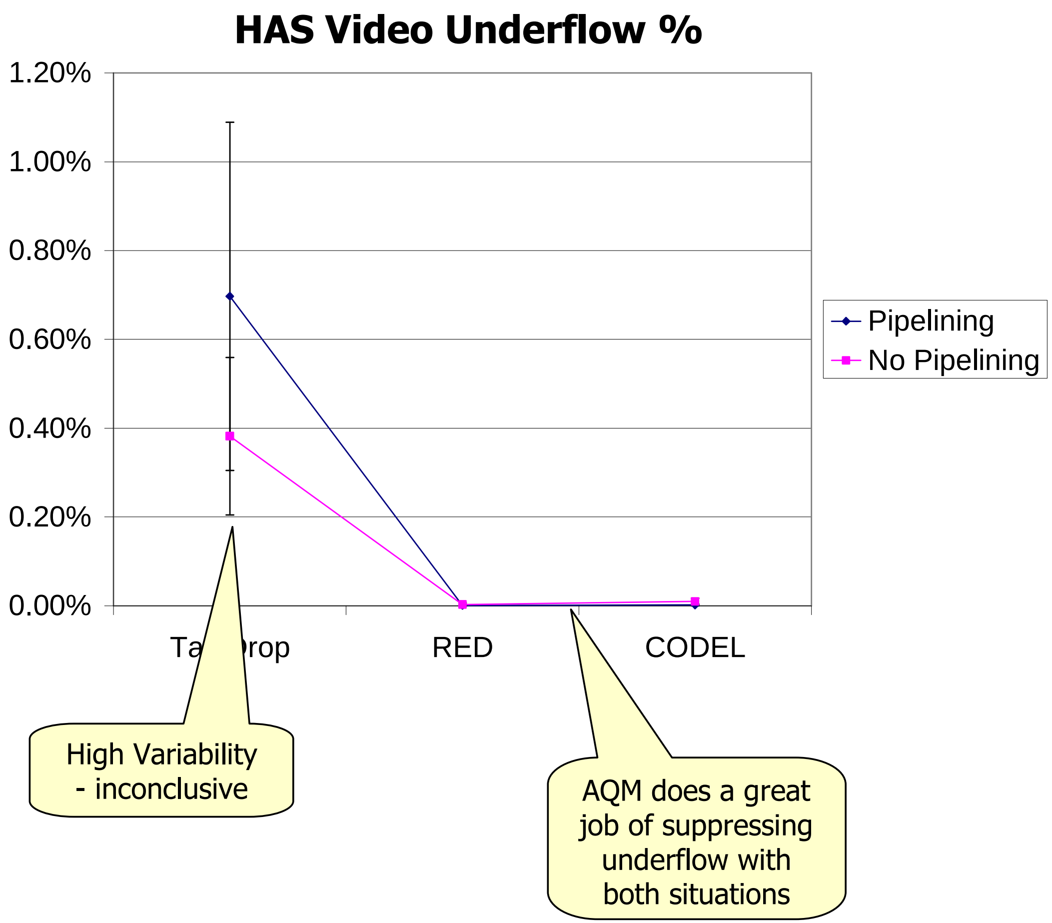


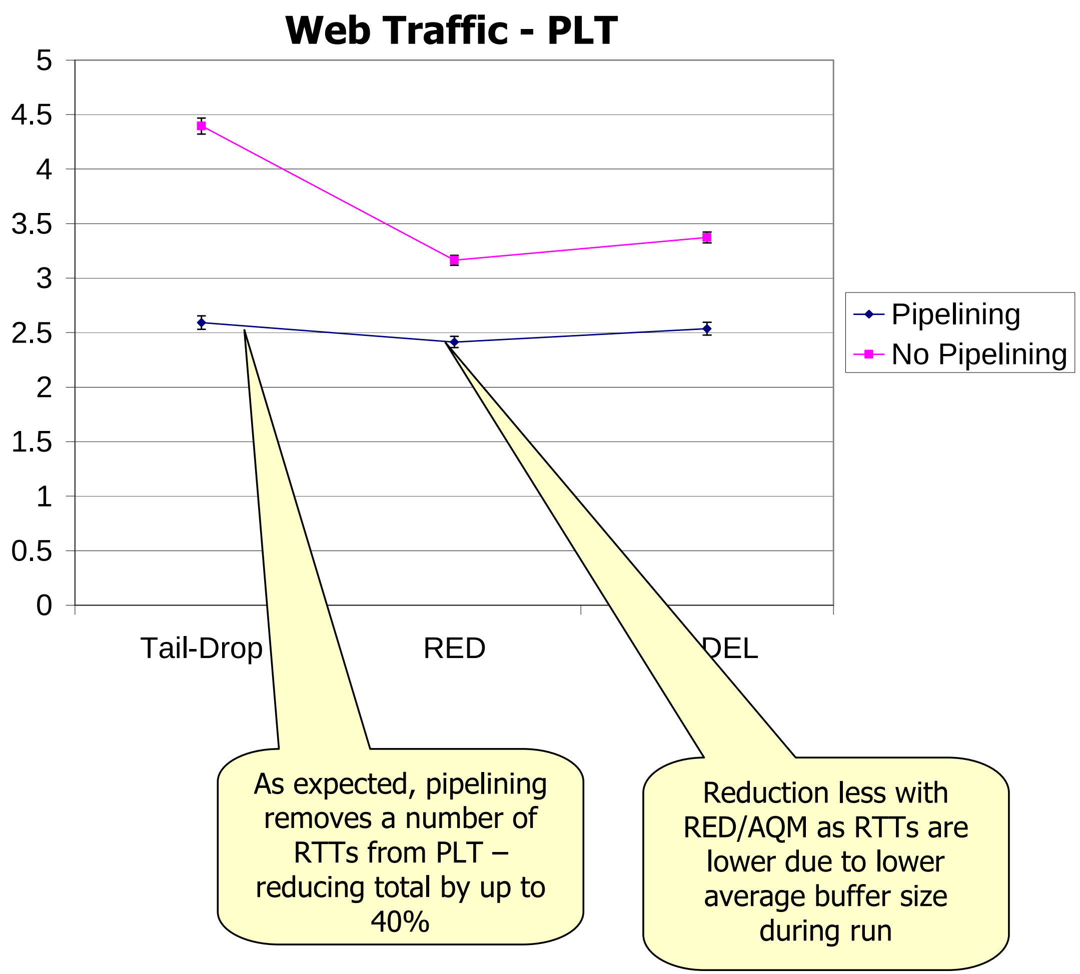
Bandwidth shared more fairly



COPYRIGHT © 2013 ALCATEL-LUCENT. ALL RIGHTS RESERVED.

Potential Impact of Full Pipelining of Web Requests



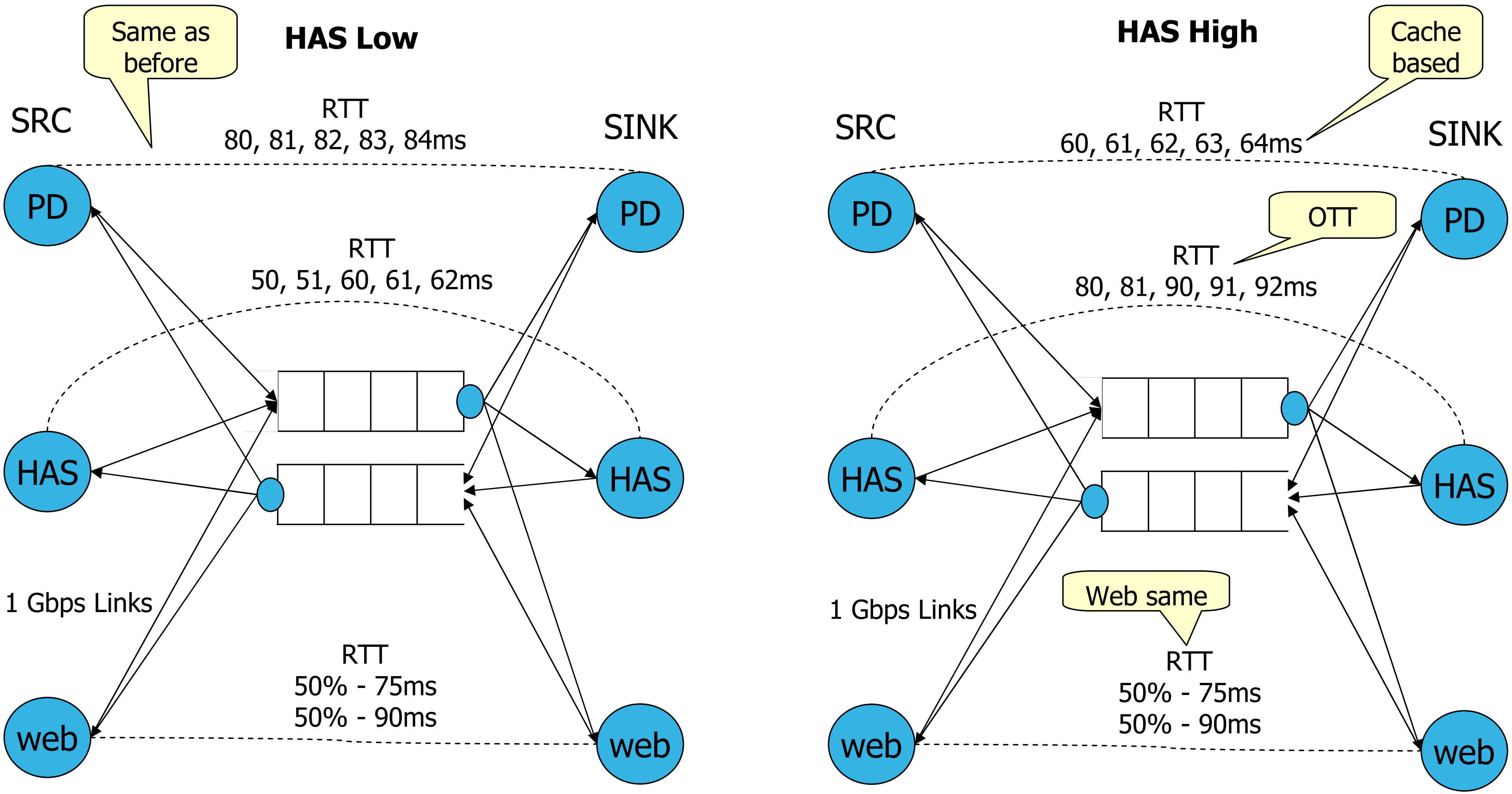


S. Akhtar et al, “An Evaluation of AQM techniques on Access Networks with Internet Traffic”

Alcatel·Lucent

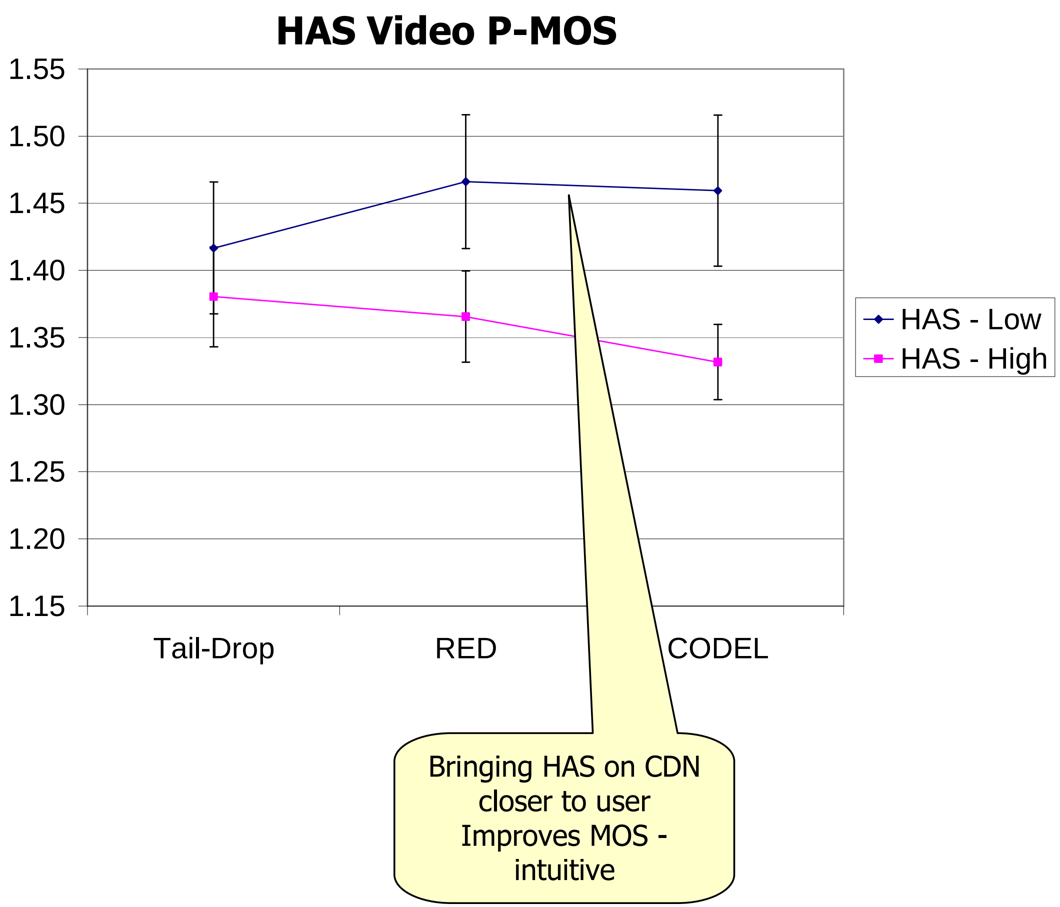
COPYRIGHT © 2013 ALCATEL-LUCENT. ALL RIGHTS RESERVED.

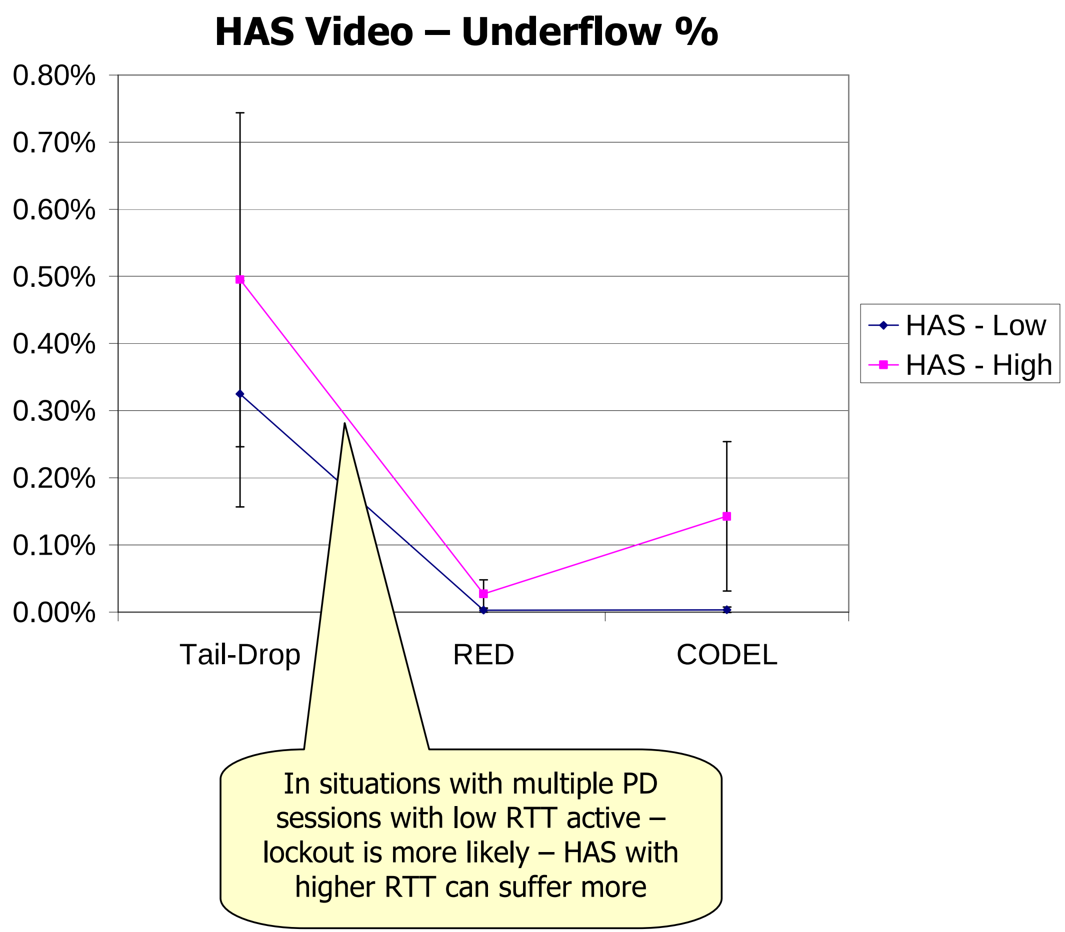
RTT Combinations



COPYRIGHT © 2013 ALCATEL-LUCENT. ALL RIGHTS RESERVED.

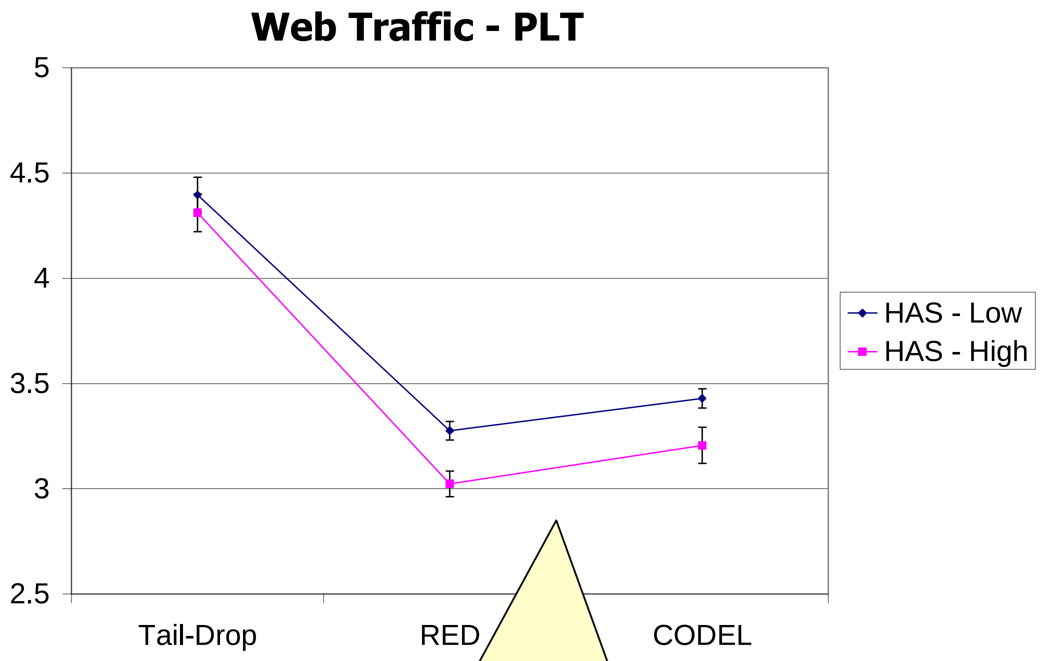
Impact of Different RTT Combinations - HAS Video QoE



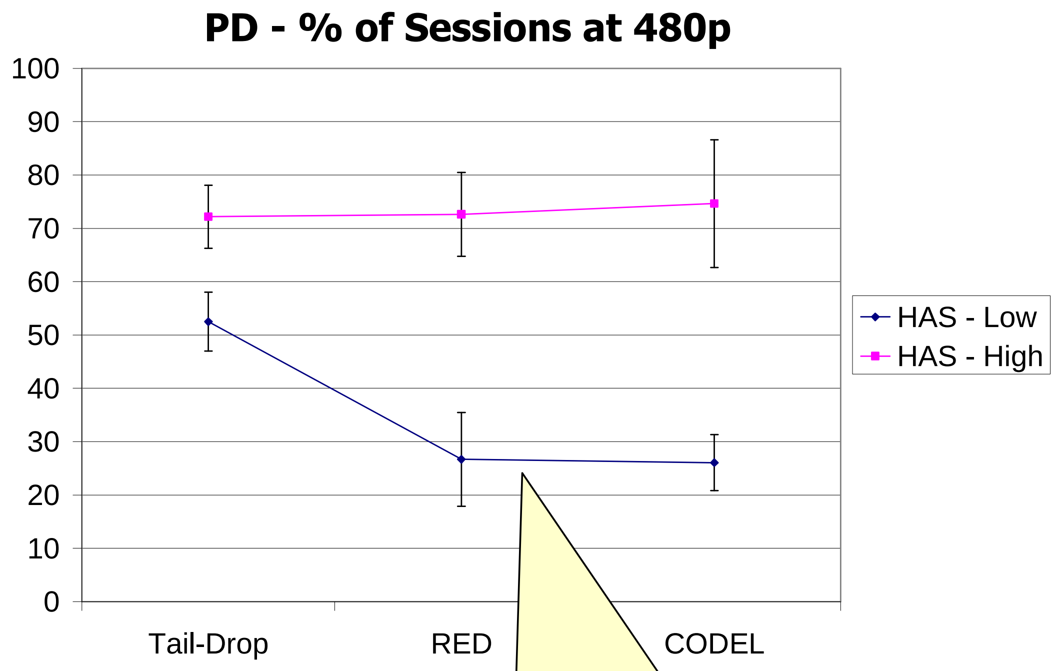


COPYRIGHT © 2013 ALCATEL-LUCENT. ALL RIGHTS RESERVED.

Impact of Different RTT Combinations - Web and PD QoE



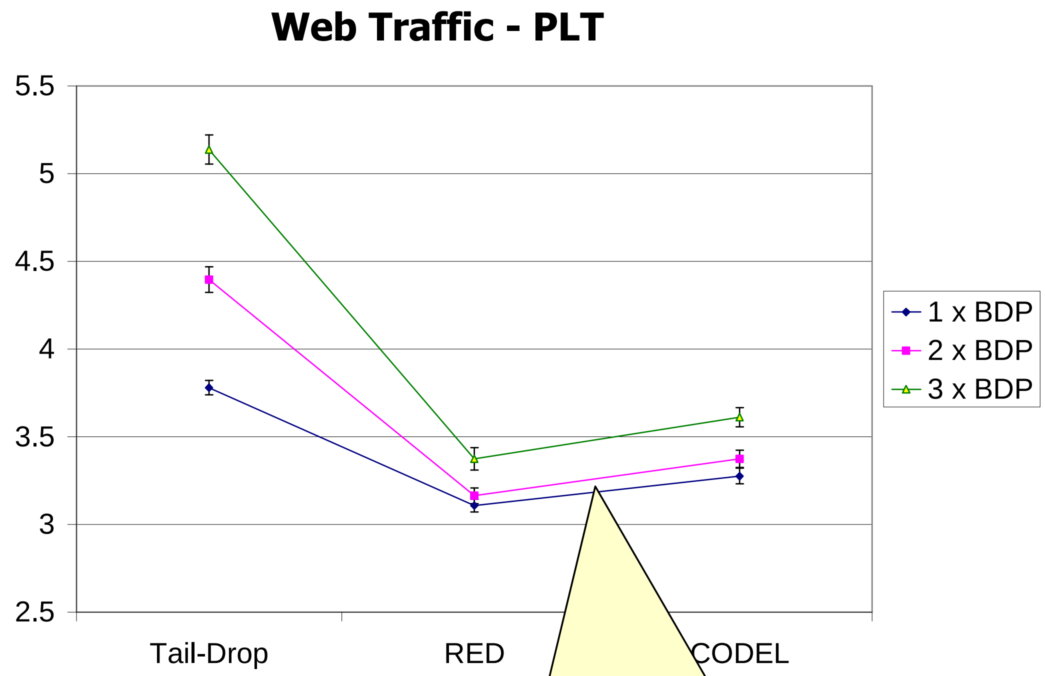
HAS at lowest RTT -> HAS tends to run at higher VQ levels -> tends to decrease bandwidth for other services like web and PD

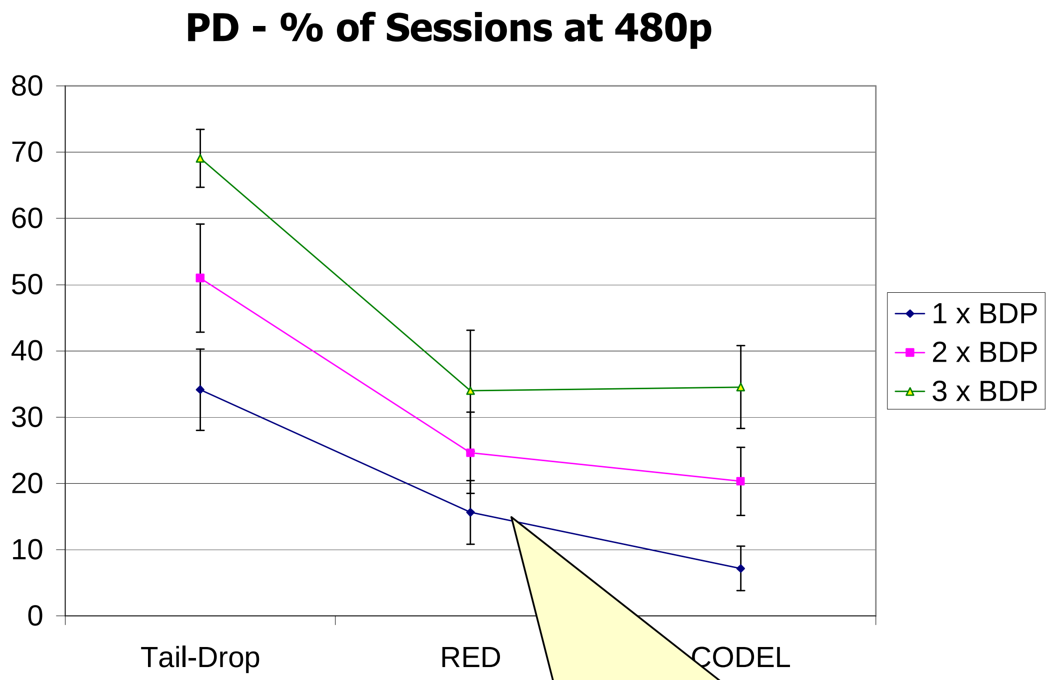


Same effect here – HAS as lowest RTT, allows VQ levels to run higher – especially with AQM active – reducing bandwidth for other services like PD

COPYRIGHT © 2013 ALCATEL-LUCENT. ALL RIGHTS RESERVED.

Impact of Buffer Size – QoE of Web and PD





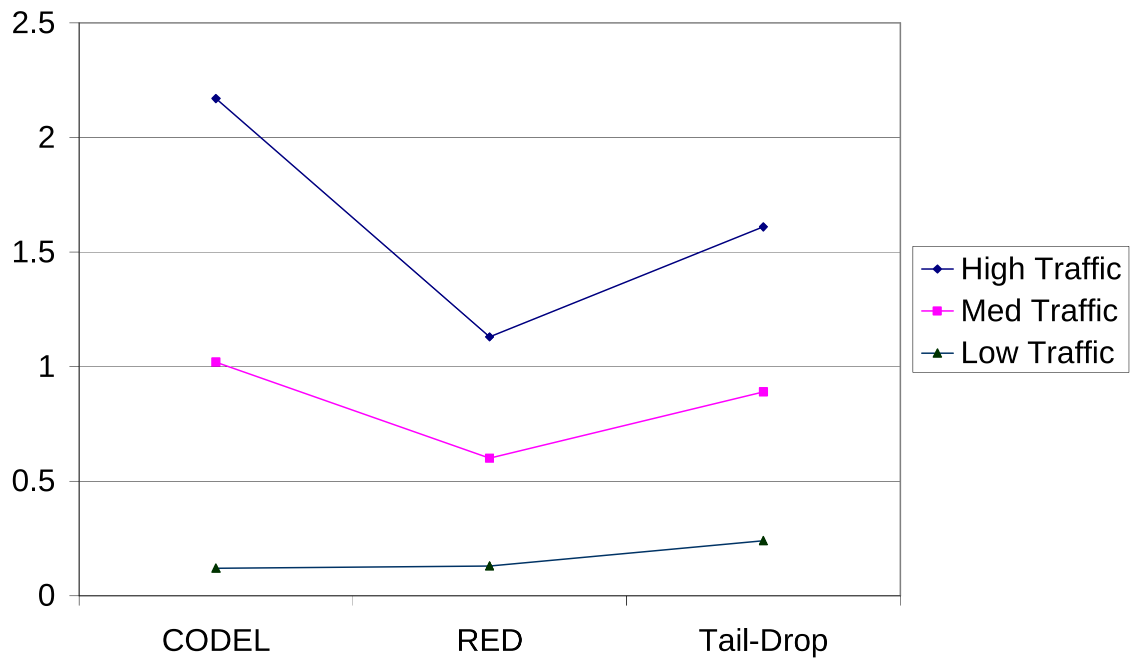
PLT higher with larger buffer sizes – intuitive.

Value of AQM in reducing PLT remain at all buffer levels.

PD flows tends to benefit most with larger buffers they have a need to keep a large congestion window over a longer period of time - both web and HAS traffic need a congestion window intermittently (when need to send traffic). The larger buffer with AQM also comes with a larger minimum threshold which helps the PD flows.

COPYRIGHT © 2013 ALCATEL-LUCENT. ALL RIGHTS RESERVED.

Packet Loss Rate and CODEL



- Packet Loss Rate (PLR) higher with Tail-drop and CODEL
- RED and CODEL – Key Reason:
  - CODEL effectively starts dropping packets at 5ms target (after 100ms interval)
  - RED starts dropping after Min threshold = 1/6 of buffer = 20 packets = 30ms (assuming 1.5ms packet transfer time)
- Unless RTT becomes less than 30ms (5ms x 6) – situation remains with these RED configurations
- The higher PLR with CODEL tends to decrease QoE of services as compared to RED

COPYRIGHT © 2013 ALCATEL-LUCENT. ALL RIGHTS RESERVED.

Conclusions and Further Work

- Internet Traffic has evolved!
  - RED/WRED – which is already available - can be far more effective
  - We present a simple configuration for DSL
- RED/AQMs on bottleneck link significantly improve both HAS and web QoE
  - On DSL, RED can reduce underflow by 30 times or more in typical conditions
- CODEL produced similar QoE as RED, but a little higher PLR in most situations
- Larger buffer sizes tend to reduce QoE of web and HAS with Tail-drop, but improve PD QoE
  - With RED and larger buffer sizes - could improve QoE for all services together
- Full pipelining improves web QoE, but may reduce HAS QoE in similar conditions
- Further Work
  - Cable access - much higher speed, much more multiplexing, effect of more aggressive TCP stacks (e.g. CUBIC) – expect lockout effect may be enhanced

COPYRIGHT © 2013 ALCATEL-LUCENT. ALL RIGHTS RESERVED.

Explanation Slides/Back-Up

Simulations - Randomization

- We use a scheduled HTTP GETs for PD, Web and HAS traffic
  - A schedule gives times when PD session, web GET request or HAS chunk would be initiated
  - A schedule is pre-calculated and is based on PD, Web and HAS modeling
  - The starting time for a schedule is determined at run time
- We start schedules at random initial delays
  - Random delays - picked from a uniform distribution with maximum of half of inter-burst time (PD, web)
- HAS streams were started 10s apart
  - Allows evaluation of how AQMs/options perform when multiple HAS streams start relatively close together
- PD schedules were randomly picked for a given run from a set of 6
- All options were given the same schedules, starting points
- Simulations
  - 6000s were run for a single run
  - 10 of these 6000s runs make up a single data point
  - 95% confidence intervals were calculated
  - Underflows had larger confidence intervals even with 10 runs
  - Testing similar circumstances would require similar time in lab

COPYRIGHT © 2013 ALCATEL-LUCENT. ALL RIGHTS RESERVED.

Predicting HAS QoE (MOS)

- MOS from HAS VQ levels and VQ variance (De Vriendt)
  - PMOS = 0.6*mean_bit_rate – 0.32*stdev_of_bit_rate (HAS bit rate < 6Mbps)
- Impact of Underflow (Buffering) on HAS MOS - very high (Conviva)
  - Direct impact on HAS MOS
    - Some research (Mok)
    - Did not use
  - Indirect impact – users may not return to site
  - Report both P-MOS and Underflow
- MOS of Worst HAS session reported
  - Operators often judged on worst session

COPYRIGHT © 2013 ALCATEL-LUCENT. ALL RIGHTS RESERVED.

Google Stats for Web Traffic (Google)

		CDF									
	Avg	10%	20%	30%	40%	50%	60%	70%	80%	90%	100%
GETs per Host	6.4	2.3	3.3	4.2	5.1	6.3	7.6	9.3	12.1	18.0	32.4
KB per GET	7.2	0.4	0.6	0.9	1.3	1.9	2.9	4.4	8.0	18.5	33.6
Hosts per page	7.0	1.0	2.0	3.0	4.0	5.0	7.0	8.0	10.0	14.0	28.0
KB per Host	37.2	0.7	1.8	3.9	6.7	11.2	17.1	26.8	55.5	132.8	199.1

- Without pipelining, we use the GETs/host and KB/GET using above lookup table with a uniformly distributed random variable to pick the number of GETs or GET size
- For number of hosts per page, we only use the average – 7 hosts per page
- With pipelining, we use the KB/host using a lookup table as above and use that as the combined file for sum of all the GETs to that host
- In both cases the inter-page time is a Pareto distribution whose mean is adjusted for page traffic according to the traffic table
- Page size has since at least doubled to 640KB – we doubled all GET sizes

COPYRIGHT © 2013 ALCATEL-LUCENT. ALL RIGHTS RESERVED.

PMOS and Underflow – Observation and Explanation

- Worst case (WC) PMOS shown – Average PMOS differential was not significant.
- WC PMOS shows improved MOS with AQM at all traffic levels (10-20%).
- Underflow seconds more volatile than other simulation parameters – higher confidence limits.
- Underflow occurred when multiple PD sessions overlap for significant periods of time.
- Higher volatility of underflow since PD sessions had longer periods (inter-arrival time).
- Whenever multiple PD sessions were active:
  - Increases the number of TCP sessions, bandwidth per session drops.
  - Increases the type of sessions (long term TCP) that tend to lock out other sessions, bandwidth per session for non-PD sessions drops further in case of Tail-drop.
  - HAS session which have to repeatedly leave the queue and re-enter as new sessions have a higher drop probability in case of tail-drop - they often find the queue full or nearly full - some packets from the initial burst drop – reducing their congestion window and ssthresh, which reduces their throughput.
  - Since the PD sessions are quite long lived (270s average), and the client buffer is set to only 30s, this scenario can last long enough to reduce client buffers (if bandwidth drops below minimum VQ level).
  - In the case of RED/AQM, the lockout related packet loss is significantly reduced as the buffer remains much lower even with multiple PD sessions active, which allows the TCP sessions from HAS to retain a higher congestion window.

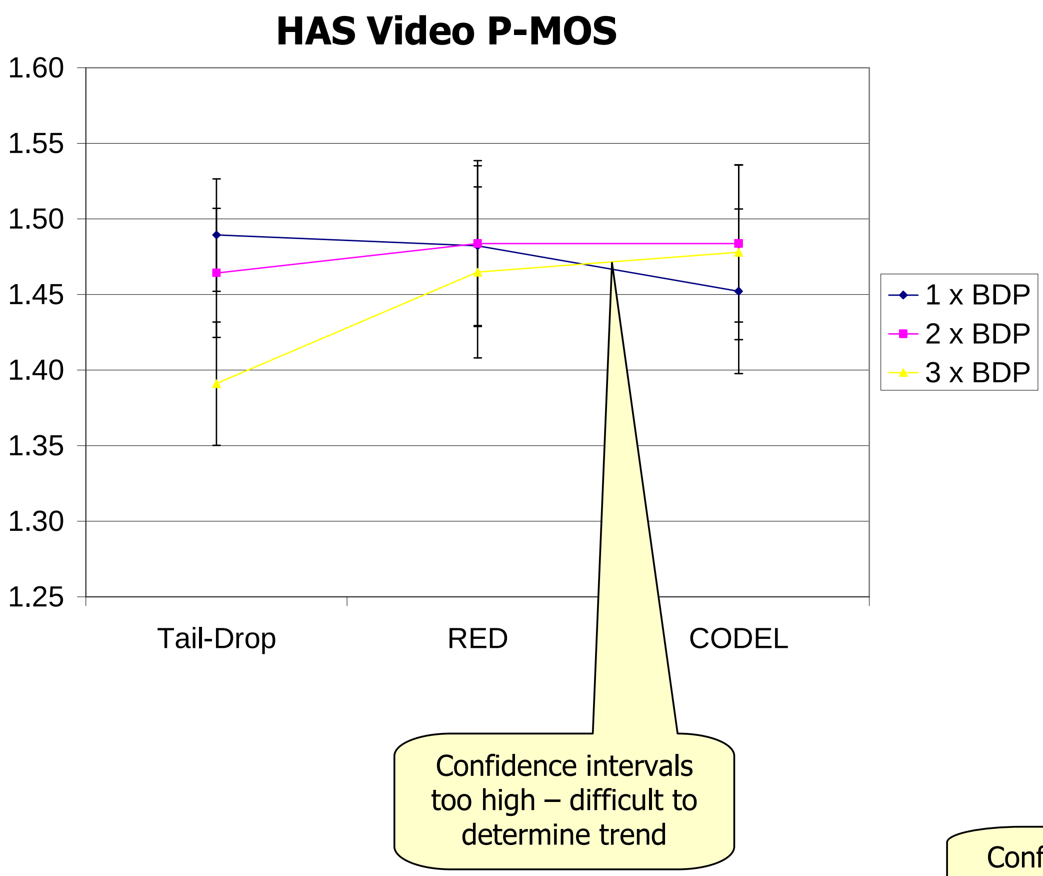
COPYRIGHT © 2013 ALCATEL-LUCENT. ALL RIGHTS RESERVED.

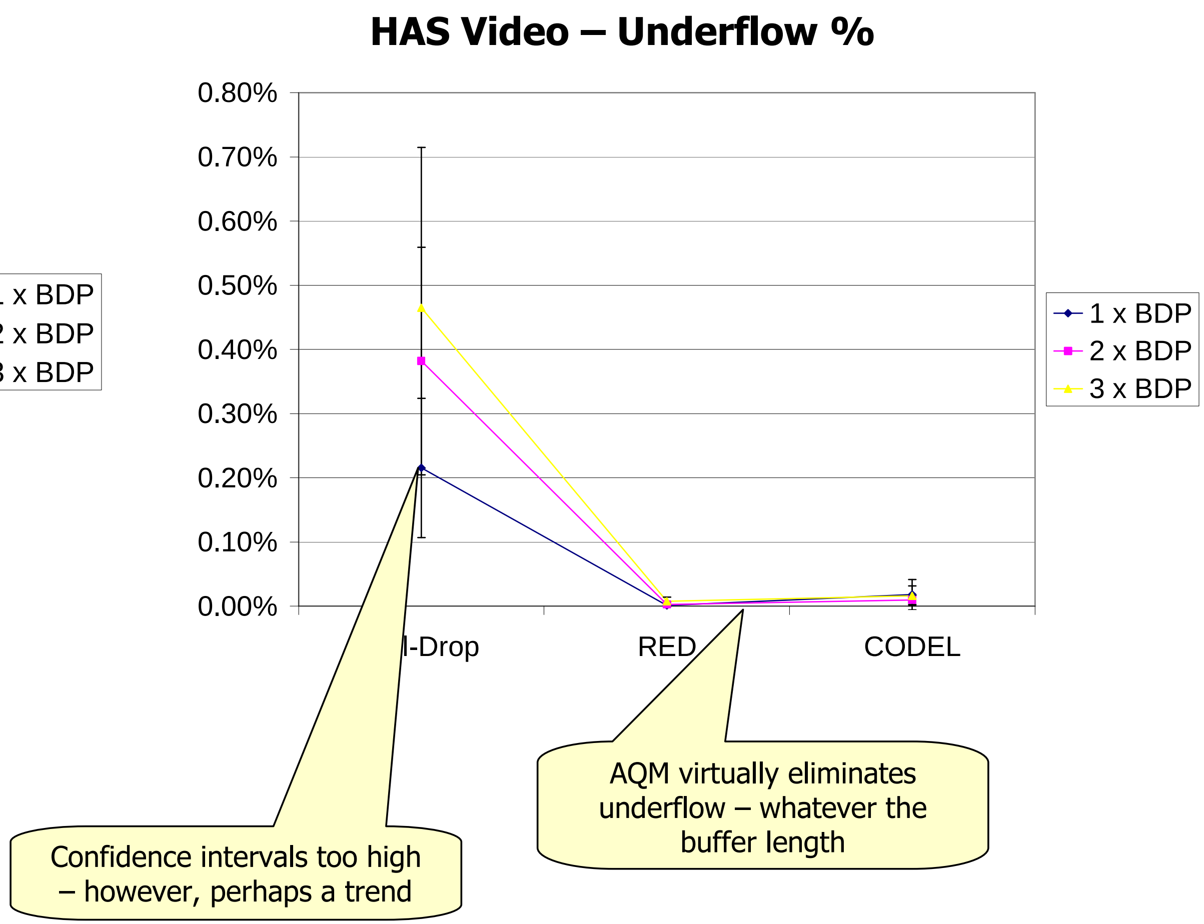
PLT and PD QoE – Observation and Explanation

- PLT - sum of a number of average GET requests/response (14)
  - Includes time for the request to leave client and time to download GET file form Host.
  - Difference between AQM and Tail-drop is significant - some difference between RED and CODEL.
  - Queue size was large with TD case (~ 90-100 MA) - lower with RED case (~ 30-40 MA).
  - Each GET/download requires one RTT in addition to transmission – which was higher with TD.
- PD QoE
  - Defined as % of PD sessions that can maintain a 480p playout for the entire session without buffering and dropping down to a lower level.
- PD traffic performed better with Tail-drop
  - PD TCP flows last a relatively long time - which can grow a larger congestion window.
  - Under TD, established large congestion windows can lockout other flows.
  - Under TD short TCP flows like web and HAS have reduced throughput, while long lasting flows like PD take advantage of that bandwidth.

COPYRIGHT © 2013 ALCATEL-LUCENT. ALL RIGHTS RESERVED.

Impact of Buffer Size – HAS Video QoE





COPYRIGHT © 2013 ALCATEL-LUCENT. ALL RIGHTS RESERVED.

References

References

- (Floyd) – S. Floyd, "Discussions on setting RED parameters", http://www.aciri.org/oyd/red.html , Nov. 1997
- (Firoiu) - V. Firoiu et al, " A Study of active Queue Management for Congestion Control", Proc. of IEEE Infocom, Tel Aviv, 2000.
- (Ziegler) – T. Ziegler et al, "Stability Criteria of RED with TCP Traffic", Proc. of IFIP ATM&IP Working Conference, 2000, http://www.newmedia.at/~tziegler/papers.html
- (Brandauer) – C. Brandauer , "Comparison of Tail Drop and Active Queue Management Performance for bulk-data and Web-like Internet Traffic", Proceedings of the Sixth IEEE Symposium on Computers and Communications, IEEE ISCC 2001.
- (Sandvine) – Exposing the Technical and Commercial Factors Underlying Internet Quality of Experience, Sandvine Report, SEPTEMBER 2013.
- (Thompson) – Thompson et al, "Wide-Area Internet Traffic Patterns and Characteristics", IEEE Network, November/December 1997.
- (De Vriendt) - Johan de Vriendt ey al , "Model for estimating QoE of Video delivered using HTTP Adaptive Streaming", IFIP/IEEE Workshop on QoE Centric Management, May 2013.
- (Conviva) http://www.techzone360.com/topics/techzone/articles/2013/02/14/326980-conviva-viewer-experience-report-says-20-billion-opportunity.htm
- (Mok) – R. P. Mok et al, "Measuring the Quality of Experience of HTTP Video Streaming", 12th IFIP/IEEE 1M 2011: Mini Conference.
- (Benno) – S. Benno et al, " WiLo: A Rate Determination Algorithm for HAS Video in Wireless Networks and Low-Delay Applications", to be published, Globecom 2013.
- (Dischinger) - M. Dischinger et al, "Characterizing residential broadband networks," Proceedings of IMC 2007, San Diego, CA, October 2007.
- (OOKLA) - http://www.netindex.com/download/2,1/United-States/
- (Google) - https://developers.google.com/speed/articles/web-metrics
- Pollere - http://www.pollere.net/CoDel.html

COPYRIGHT © 2013 ALCATEL-LUCENT. ALL RIGHTS RESERVED.

AT
THE
SPEED
OF
IDEAS

COPYRIGHT © 2013 ALCATEL-LUCENT. ALL RIGHTS RESERVED.