



802.1au: Simulation Ad Hoc Report

November 14, 2006

Manoj Wadekar

Intel Corp.

CN-SIM Ad-Hoc: Overview

- Meetings:
 - 5 Weekly meetings held since July 2006 IEEE 802.1 Plenary meeting
- Participation:
 - 15+ members actively participated in the calls
 - Representing 10+ companies
- Goal:
 - PAUSE and BCN enhancements :
 - Scenarios: PAUSE, BCN+PAUSE, BCN+BCN-MAX+PAUSE
- Status:
 - Consistent results across 5 independent simulation environments for BCN + PAUSE
 - Enhancements studied for improving performance during Transient Congestion, rate drift etc.

Thank you all for great team work!

CN-SIM Modeling Teams

- We had five simulation teams with independent environments
 - Bruce Kwan/Pat Thaler: Broadcom
 - Davide Bergamasco: Cisco
 - Zhi-Hern Loh/Uri Cummings: Fulcrum Microsystems
 - Tanmay Gupta: Intel
 - Prof. Balaji Prabhakar/Yu Li: Stanford University

Baseline Switch Model: PAUSE Discussion

- Discussion whether to use:
 - Global PAUSE (to all input ports)
 - Selective PAUSE (Each input generates independent PAUSE)
- Group agreed that it is necessary to:
 - For ease of simulation, consistency and comparison of results:
 - Use Global PAUSE mechanism for simulations
 - Use agreed upon threshold settings for High/Low water marks, minimum buffer requirements
 - <http://www.ieee802.org/1/files/public/docs2006/au-sim-bergamasco-common-bridge-model-101206v2.pdf>



Summary Findings

BCN+PAUSE

- BCN+PAUSE – provides “no drop” behavior
 - PAUSE is only triggered during early transient congestion period
 - Reduces congestion spreading as compared to “PAUSE only”
 - Protects “innocent flows” as compared to “PAUSE only”
- BCN (0,0) increasing unfairness, transient inefficiency
 - Requires further study
 - BCN (MAX) started getting investigated
 - When Queue is saturated, request source to reduce rate by maximum amount

Related Presentations

- <http://www.ieee802.org/1/files/public/docs2006/au-sim-gupta-baseline-pause-101906.pdf>
- <http://www.ieee802.org/1/files/public/docs2006/au-sim-ding-bcn-pause.pdf>
- <http://www.ieee802.org/1/files/public/docs2006/au-sim-ding-bcn-pause-w-innocent.pdf>
- <http://www.ieee802.org/1/files/public/docs2006/au-sim-loh-bcn-pause-sample-window-10192006.pdf>

BCN-MAX Study

- BCN-MAX variants tested:
 - BCN-MAX-PAUSE-Based: When PAUSE is asserted, each BCN carries maximum negative feedback
 - BCN-MAX-QSC-Triggered: Equivalent to BCN(0,0), instead maximum negative feedback is generated when $Q_{off} > Q_{sc}$
- BCN-MAX-Qsc behaved better than PAUSE based trigger
- Rate Drift improves BCN Fairness
- Further enhancements:
 - BCN-MAX-QSC-HIGH-SAMPLE-RATE: Increase sampling rate during BCN-MAX and $Q_{len}=0$
 - BCN-MAX-QSC-HSR-SYM: High sampling rate during BCN-MAX and $Q_{len}=0$, maximum +ve feedback during $Q_{len}=0$

Related Presentations

- <http://www.ieee802.org/1/files/public/docs2006/au-sim-bergamasco-bcnmax-comparison-102606.pdf>
- <http://www.ieee802.org/1/files/public/docs2006/au-sim-bergamasco-bcnmax-comparison-110906v2.pdf>
- <http://www.ieee802.org/1/files/public/docs2006/au-sim-ding-bcn-pause-102606.pdf>
- <http://www.ieee802.org/1/files/public/docs2006/au-sim-loh-bcn-sample-window-2-10262006.pdf>

Rate Drift

- At fixed time intervals, rate limiter increases its rate by a unit
- Never stop drift except timeout in BCN(0,0)
 - Fixed drift x Mbps every y seconds
 - Other variants
- Findings:
 - Rate drift improves fairness
 - Provides robustness against loss of BCN signals (or CP)

Related Presentations

- <http://www.ieee802.org/1/files/public/docs2006/au-sim-ding-bcn-pause-102606.pdf>
- <http://www.ieee802.org/1/files/public/docs2006/au-Lu-et-al-BCN-study.pdf>

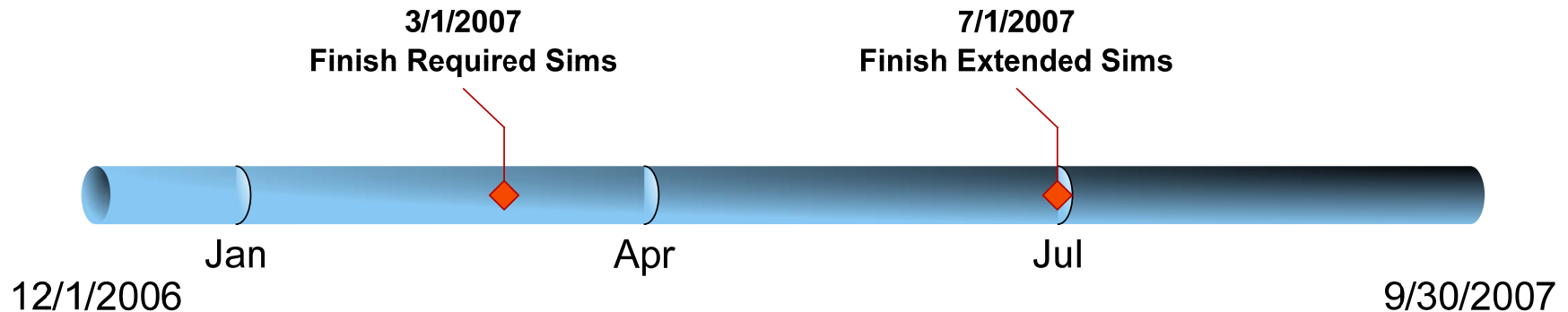
Other presentations

- Large Topology
 - <http://www.ieee802.org/1/files/public/docs2006/au-sim-thaler-bcn-large-topo-110206.pdf>
 - Parameters used for small topology perform poorly for topology with large number of flows
 - Things that help: Decreasing sample rate, other parameter changes etc.
 - Further analysis required for common set of parameters (large and small topologies)
- Benchmarks and next steps
 - <http://www.ieee802.org/1/files/public/docs2006/au-gusat-congestion-bmrk-0607-r089.pdf>
 - Need for measures for packet drop and pause assertion time

Summary & Next Steps

- Basic BCN mechanism is largely unchanged
 - Enhancements added to improve response during transient congestion
 - Rate drift to improve fairness, robustness
- Next step:
 - Number of desired simulations is plenty – resources are not
 - Let's pick most important scenarios
 - Suggestions:
 - Large topologies (how many flows, how many switches, ?)
 - Bursty workloads (starting/stopping over duration, ?)
 - ??

Suggested Timeline



- March 2007: Finish “Required Topologies and workload”
 - Let’s define these this week
 - Provides enough justification for first draft to be discussed
- July 2007: Finish “Extended Topologies and Workload”
 - We can finish these by January 2007
 - Can provide enough data for ballot discussions
- Beyond July 2007: Work on enhancements/modifications as needed
 - Justify the changes
 - Validate the changes going in the document