



Tales from the Bandwidth Broker Woods

Jan-2005

Larry Dunn

Manager, Advanced Architecture

ldunn@cisco.com

Agenda

- **Connection with this workshop**
- **Overview of previous efforts**
- **Showstoppers**
- **Routing component examples: Dina's stuff...**
- **Conclusion**

Connection with this workshop

- **Mauro's Routing ToR- interest in BOD, overlays, alternate path computation, etc.**
- **Internet2 BB (bandwidth broker) effort originated with focus on inter-domain QoS**
- **Lessons in intra-domain "oracles", path computation, inter-domain signaling**
- **More examples of state/granularity tradeoff**
- **And things that are "hard to do"**

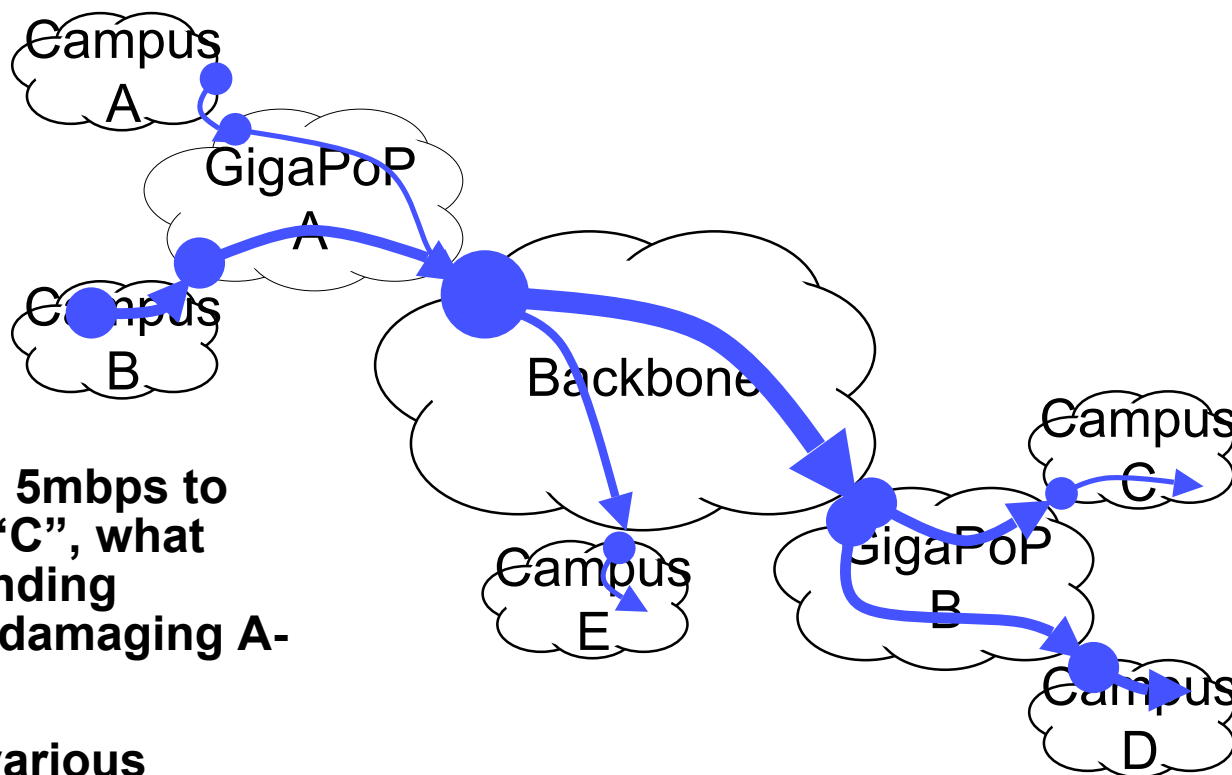
Overview of previous efforts

- **For this discussion-**
<http://qbone.internet2.edu/bb/index.shtml>
(Sue Hares led for a time, then Phil Chimento)
- **Attempt to automate signaling, state, resource allocation for interdomain QoS (what you'd expect from a "bandwidth broker")**
- **Worked through state machine, minimalist approach, TLV's (Rudiger Geib), some security (Volker Sander)**
- **Allowed for intra-domain freedom, fault-tolerance, etc; but only specified the inter-domain components**
- **Basics: BB's in each AS communicate with adjacent BB's; not too different from ATM-SVC setup (propagation from sender; intra-domain details opaque to other AS's)**
- **Other efforts: previous QoS work, IETF-RAP, AAA, etc.**

Showstoppers

- **Essential desire to minimize state, leads to “one token bucket per ingress”**
- **Classic cloud- versus pipe-model conflict**
- **Issue: If $a \rightarrow b \rightarrow c$ at X, and $a \rightarrow b \rightarrow d$ at Y, and single T.B. at “b” to police “a”, how to prevent “a” from sending “X+Y” to c?**
- **Answer: nobody figured it out...**
(next slide: credit Ben Teitelbaum)

Showstopper in pictures:



- **Example:** if “B” has 5mbps to “E”, and 5mbps to “C”, what prevents B from sending 10mbps to C? (and damaging A->C traffic?)
- **Key issue:** are the various policers (Gigapop-A, B) really aggregate-policers, or per-pipe? (granularity vs. state again...)

Relevance to routing, BOD

- **Any “non-normal IP routing” needs classification somewhere, possible multiple locations**
- **If granularity is binary (e.g. wavelength=yes/no), then might be OK**
- **But any sub-division of wavelength needs to address this issue at some level**
- **Other routing topics: does classifier lead to forwarding against {full-sized_table, encapsulation/label/tunnel, per-router ACL's, something_else}? Scalability?**

Routing component examples

- **Dina Katabi- extending XCP to traffic engineering**
<http://nms.lcs.mit.edu/~dina/texcp.html>
Re-use XCP (transport) control-theory components to allow fast shifting of TE traffic
Differs from per-pkt XCP in that it's slow-path, probe-pkts
- **IETF PCE (Path Computation Element) w.g.**
<http://ietf.org/html.charters/pce-charter.html>
Intra-domain first (mult-area), then Inter-domain
Non-head-end path-computation for setting up LSPs

Conclusion

- **Lots of interest; Lots of history/previous_work**
- **Easy to get it wrong or too_narrow to be useful**
- **Will we repeat mistakes? If not, why not?**
- **Do we care about general applicability, vs. niche application? What are characteristics of one vs. the other?**
- **Good opportunity to work things out in the R&E space; wavelength-dimension is interesting (binary) sub-problem**
- **Think carefully about scalability vs. granularity**