

Computer Science Department  
Boston University

**Networks PhD Depth Exam**  
**November 15, 2005**

1. Check that there are 3 questions over 7 pages including this page. Answer all questions. Turn in your answers by Friday, November 18, 4:00pm, at MCS-140G.
2. You are free to consult any notes, books and papers during the examination—make sure to include your references. Copying existing solutions or parts of them will be considered plagiarism, and in this case you will fail the exam.  

You must develop your own solutions, which might use existing ideas and techniques, as long as you cite them and clearly explain how these existing ideas/techniques fit within your own solution. Failure to do so and a feel of cut-and-paste in your answers will receive zero credit and you may fail the exam.

You are NOT allowed to consult with any person during the examination.
3. You may typeset your answers (in which you should feel free to include hand-drawn figures), or you may neatly write your answers. Please answer each question in a separate writeup.
4. Although questions are of different length, they are of equal weight.
5. Please indicate the question number and your code number on each answer sheet. Do not write your name on the answer sheet.
6. If you have any doubt as to the interpretation of a question, make a reasonable assumption and explain your interpretation in your answer. No explanations will be given during the exam.
7. Your answers should show research maturity and depth, so you may support your answers not only by analysis, but possibly by measurements or simple simulations. Be creative!
8. The copyright to any new research question is owned by its author :)

**Networks PhD Depth Exam**  
**November 2005**  
**Question 1**

Consider the following multi-homed network.

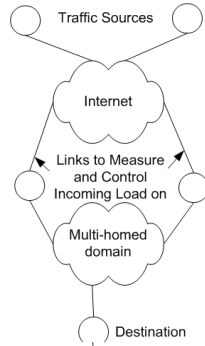


Figure 1: Controlling Incoming Traffic to your Multi-homed Network Domain

As a network manager of such multi-homed domain, you observed an imbalance in the traffic load on incoming links to your domain. You decided to investigate ways to influence the routing choices of outside nodes so you end up with a more even distribution of load. You carefully read the following papers, which investigate the use of a technique called *path pre-pending* where you let BGP intentionally lengthen the paths advertised to the outside world through each incoming interface.

- *Inbound Traffic Engineering for Multihomed ASes using AS Path Prepending* by Rocky Chang and Michael Lo. In IEEE/IFIP Network Operations & Management Symposium, 2004.  
<http://www.cs.bu.edu/faculty/matta/Home/NOMS.pdf>
- *Characterizing the Performance and Stability Issues of the AS Path Prepending Method: Taxonomy, Measurement Study and Analysis* by Hui Wang, Rocky Chang, Dah Ming Chiu and John Lui. In ACM SIGCOMM Asia Workshop, 2005.  
[http://www.cs.bu.edu/faculty/matta/Home/sigcomm\\_asia.pdf](http://www.cs.bu.edu/faculty/matta/Home/sigcomm_asia.pdf)

You want to answer several questions on how effective this *path pre-pending* technique is:

- (a) How much traffic can get shifted from one incoming interface to another for a certain amount of *path pre-pending*?
- (b) How does the topology of the Internet affect such traffic shifts, assuming BGP routing policies that give preference to shortest-length paths? You want to consider random topologies and power-law topologies to understand the topological effects.
- (c) What if a fraction of outside nodes (domains) use routing policies that are different from shortest-path preferences? You want to vary that fraction of nodes, for example those who randomly pick any neighboring node (domain) to get to a destination.

Given your characterization of the amount of traffic shifted as a function of the amount of path pre-pending, design and analyze a load balancing scheme that would periodically measure congestion on incoming interfaces (e.g. utilization or delay), and then adapts the amount of path pre-pending. Study the stability of such load balancer, i.e. convergence speed, size of oscillations, etc. under different congestion metrics.

Now, you read this paper and got worried!

- *Exploiting the Transients of Adaptation for RoQ Attacks on Internet Resources* by Mina Guirguis, Azer Bestavros, and Ibrahim Matta. In IEEE ICNP, 2004.

<http://www.cs.bu.edu/faculty/matta/Papers/icnp04.pdf>

You ask yourself:

- (d) Can an outside node (domain) attack the adaptation of your load balancer by messing up with its path advertisements?
- (e) What fraction of such malicious domains and where they could be located to cause you the most damage?
- (f) How can you design an adaptation that is robust to such attacks?

Develop your own systematic methodology to answer all questions above. Submit all your analytical and simulation models, along with your results and observations. If you adapt any existing model, you must cite that work and indicate precisely why it is valid in this specific problem setting. Organize your answer in the form of a short technical paper, with appropriate sections (including bibliography).

**Networks PhD Depth Exam**  
**November 2005**  
**Question 2**

Virtualization is a major driving force for the brave new world of the Next Generation Internet (NGI). A good (or bad) example of how virtualization may be used in NGI is the GENI project (<http://www.nsf.gov/cise/geni/>) you may have heard about (had you been smart enough to attend ICNP :)

One of the virtualization services envisioned by an ambitious faculty member at BU aims to enable ISPs (called the Overlay ISP) to overlay virtual networks atop other ISP's infrastructures (called Underlay ISP). A key functionality that enables this vision is "link virtualization", whereby the capacity of a link is programmatically set in real-time to a value  $c(t)$  based on some policy/method provided by the OverISP and enforced/guaranteed by the UnderISP. Packets arriving from an OverISP to its virtual link will be buffered and transmitted at an instantaneous speed determined by  $c(t)$ . Each virtual link is allowed to have up to  $B$  packets buffered at any time. Thus, if  $b(t)$  denotes the instantaneous buffer occupancy at time  $t$ , then  $b(t) \leq B$ .

To address concerns about economic viability and to provide commercial entities with incentives to deploy and use this new architecture, a major "advantage" of this new architecture is for an UnderISP (e.g., owners of the links and routers) to charge for the virtual slice it provides to an OverISP. While the price structure for virtual link slices could be quite involved, for the purposes of this exercise, it is assumed that the price charged at time  $t$  is proportional to the capacity requested at time  $t$  – namely  $c(t)$ .

To leverage this new capability, an OverISP is considering the adoption of one of the following "stingy" approaches for setting (adapting) the virtual link capacity  $c(t)$ :

- Approach #1:  $\frac{d}{dt}c(t) = -K(b^* - b(t))$
- Approach #2:  $\frac{d}{dt}c(t) = -Kc(t)(b^* - b(t))$

The motivation behind both of the above strategies (and possibly many others) is for the OverISP to maximize the utilization of its virtual slice of the link, by requesting additional virtual link capacity only to relieve buffering pressure at the link (and by releasing capacity when buffering pressure eases). In particular, the overISP reduces its charge by reducing  $c(t)$  as  $b(t)$  drops below a certain threshold  $b^*$ .

As an aspiring young networking researcher, you have been retained by a startup team to develop further the above vision. Your startup company is going through due diligence, part of which involves a technology evaluation for which you are the point person responsible for producing a report.

- (a) Write a 1-2 page motivation, making appropriate references to prior works. Your motivation should read like the introduction and related work sections of a technical paper.

Now that you have motivated this element of the NGI architecture, you must convince those evaluating the technology (and hence potential investors) that the technology will be compatible with existing end-to-end applications and protocols using the current Internet, since otherwise the OverISP company will not have much clientele. In particular, you are asked to consider the implications for/impacts on TCP, which is the lingua franca of most of today's Internet applications. To do so, you decided to follow through with the following tasks, each of which could be thought of as an independent section of a technical paper.

- (b) Develop an analytical model that will allow you to characterize the performance of a variant of TCP (e.g., Vegas—well, its delay-based error detection would seem better suited for NGI!) when going through a single virtual link (and associated buffering) as described by the two alternative approaches for setting  $c(t)$  given above.
- (c) Comment on the convergence properties of TCP and discuss the difference between the two approaches for setting  $c(t)$  given above. In particular, discuss the role of the parameter  $K$ . How should it be set?

- (d) Discuss the implications on fairness of using the above two approaches (including the interplay with RTT).

After doing all of the above, you heard through the grapevine that the technical team in charge of doing due diligence on the technological underpinnings of OverISP have no stomach for analysis and as a result have a prejudice against using sound modeling on the grounds that unless it is validated, it is worthless.

- (e) Design an experimental evaluation process (namely a simulation study) via which you would validate the results you obtained above. Discuss in details the experiments you would conduct, for example, discuss how would you use the popular ns-2 simulator.

Now that you have developed detailed plans for a simulation study, it came to your attention that the technical team in charge of doing due diligence on the technological underpinnings of OverISP have no stomach for simulations either, but rather they will only believe your story if it is evaluated on a realistic testbed.

- (f) Since such realistic testbed is non-existent, you thought that it could be a good idea to write a few paragraphs in your report on the requirements that you would like to see satisfied in a testbed for testing OverISP's bold ideas. In particular, comment on the suitability of what is proposed in GENI (e.g., issues of time scale, controllability and predictability of virtual slices, etc.)
- (g) In a concluding section of your report, feel free to discuss future unexplored directions in which you could take this study (if you had more time). For example, you may want to discuss other approaches for setting  $c(t)$ , effect of buffer management, etc.

**Networks PhD Depth Exam**  
**November 2005**  
**Question 3**

You are responsible for optimizing the performance of an existing wireless MANET (mobile ad-hoc network). Your network is expected to consist of hundreds of nodes, with frequent transient message and link failures, but infrequent node mobility. The main consideration in your network is energy consumption, but overall computation time at the nodes is also a factor of course. Assume that the nodes in your network are not GPS-enabled.

In an architecture which uses a slotted MAC protocol, one of the challenges is to keep clocks approximately synchronized. A typical defense against clocks that are out of synch is to add a fudge factor known as a guard time before and after each slot. The larger the guard time, the lower the throughput of the channel, and thus keeping clocks as closely synchronized as possible is paramount in these networks.

- (a) Assuming that synchronization is performed with respect to a leader's reference clock, briefly argue for the merits of a tree-based synchronization rooted at the leader versus flow-based synchronization along multiple paths in this setting.
- (b) Suppose that the synchronization error of node  $i$  with respect to the leader's clock grows linearly with the hop count along the shortest path from  $i$  to the leader in the synchronization graph. Under this assumption, and with the goal of minimizing the worst-case synchronization error, propose and justify a method for constructing an appropriate tree or flow. Briefly discuss the impacts of failures on your construction, but only if time allows, i.e. this is not worth much.
- (c) Now suppose that you care more about the total energy consumed by message transmissions in constructing a synchronization graph than about minimizing the worst-case error. Now propose and justify an alternative method for constructing an appropriate tree or flow. Briefly discuss the impacts of failures on your construction, but only if time allows, i.e. this is not worth much either.

One of your colleagues suggests to use the following application of network coding to improve the total energy consumed by transmissions. She draws three nodes A, B and C in a row on the whiteboard, where B can hear A and C's transmissions but C cannot hear A's transmissions, and vice versa. She then says "Consider when A has a message (1) to send to C, and C has a message (2) to send to A. Naively, it appears that four transmissions are needed: two to get message (1) from A to C via B, and two to get message (2) from C to A via B. But if nodes can buffer packets, then B can store the original payloads of (1) and (2), can compute the new payload  $(1 \oplus 2)$ , and can then broadcast this coded payload along with an appropriately descriptive header, instead of transmitting the original datagrams (1) and (2). Assuming A and C store their original copies of packet 1 and 2, respectively, they can invert the XOR to obtain the packet they are missing. By this argument, the pair of messages can be sent with three transmissions, not four."

- (d) Is there something fundamentally wrong with this proposed approach, or does it make sense? If there are issues that are fixable or have been left unspecified, briefly describe the issue and the fix.
- (e) Suppose each node knows the schedule of inbound messages that it will receive, and knows its immediate neighbors. Assuming an upper bound of  $\delta$  ms that a packet can be buffered at an intermediate network node, sketch a coding approach leveraging the idea above that minimizes the number of (outbound) transmissions.
- (f) Now suppose each node knows not only its neighbors, but also which of its neighbors can overhear which of its other neighbors, i.e. node  $i$  neighboring  $j$  and  $k$  knows whether or not  $j$  can overhear  $k$ . How can you use this to improve your answer to (e)? What coding approach do you use now? Feel free to make reasonable additional assumptions, such as assuming symmetry in the overhearing relationship, if it simplifies things.

(g) [worth significant credit] Write a simulation or provide analysis for the advantage the proposed approach can provide in a simple toy setting. For example, in a simulation, you might try 100 nodes randomly distributed on a unit square with shortest-path routes and a simple traffic matrix, plus some other details that are intentionally omitted. No need to think about mobility or failures for this part. You also need not consider optimizations to the basic idea that you discussed in parts (e) or (f), but you can if you wish. Keep it basic so that you can get some results in a few hours!