Which Way Towards A More Robust and More Reliable Communication Infrastructure?

(Network Centric vs User Centric Traffic Engineering)

Roch Guerin
University of Pennsylvania

Acknowledgement

- This talk borrows and benefited from joint work and discussions with many people
  - Current and former Ph.D. students
    - Y. Huang, A. Sridharan, S. Tao, and E. Vergetis
  - And colleagues
    - C. Diot, L. Gao, J. Kurose, S. Sarkar, D. Towsley, and Z.-L. Zhang

whose contributions I gratefully acknowledge. But don’t blame them for errors or opinions you disagree with. These are all mine.

Where Do we Stand?

- IP networks are today's communication infrastructure
  - They are used by an increasingly diverse set of applications
    - Voice, video, data, sensing, grid computing, web, ecommerce, etc.
  - And they are being implemented using a broad array of technologies
    - Diff-Serv, MPLS, MTR, etc.
    - 802.11, WiMAX, PON, etc.
- In spite or because of this diversity, there is a growing need for a more robust and more reliable network
- We are faced with the usual conundrum for addressing this issue
  - Add intelligence into the network
  - Add intelligence outside the network

Which way do we go?

The Case for Outside Intelligence

- A two-prong argument
  1. Networks are getting better and bigger so that adding intelligence to the network becomes less compelling and more expensive
  2. Beauty is in the eyes of the beholder, i.e., the definition of a "good" network can be highly application dependent and the choice is best left to them
- Some evidences in support of external solutions
  - We can make networks better simply by better using what we have
    - Robust (oblivious) networking
  - And we can expect things to get better as networks grow bigger
    - The power of over-provisioning
  - And there are many technologies that can let end-systems effectively control the network performance they experience
    - Leveraging network diversity (more than one choice)
  - And defining what is a "good" network is increasingly difficult

Robust Networking

- Goal: Be proactive rather than reactive
  - Pre-configure the network so that its performance remain "near-optimal" across a broad range of changes
    - Link/node failures and/or traffic variations
- Lots of initial results pointing to feasibility of such an approach
  - Oblivious routing in MPLS settings (Sigcomm'03)
  - Robust weight settings in traditional IP networks
    - Some examples of what is achievable

"Typical" POP-Level ISP Network

16 Nodes, 82 Links, Link Utilization of 0.7

Robustness to all possible single link failures

Heuristic	% Deviation from Optimal Re-routing									
	10	20	30	40	50	60	70	80	90	>100
Exhaustive	**9.1**	**81.8**	**9.1**	0	0	0	0	0	0	0
Frequency L=40	**13.3**	**73.3**	**6.67**	**6.67**	0	0	0	0	0	0
Frequency L=20	**10**	**75**	0	**5**	0	0	0	0	0	**10**
Plain	0	**3.33**	0	**16.67**	**3.33**	**6.67**	0	0	0	**70**

Large Network

150 Nodes, 432 Links, Link Utilization of 0.7

Robustness to all possible single link failures

Heuristic	% Deviation from Optimal Re-routing									
	10	20	30	40	50	60	70	80	90	100
Frequency *L*=10	**100**	0	0	0	0	0	0	0	0	0
Plain	0	0	0	0	0	0	0	0	0	**100**

Note that things seem to be getting better as network size grows

Is this a trend?

Exploring the Efficiency of Large Networks

- Basic question: Does over-provisioning "efficiency" improves as network size increases?
  - Lots of examples where scale helps
    - Trunking efficiency
    - Statistical multiplexing
- Networks are, however, complex beasts
  - Interactions between many parameters
    - Topology, routing, base traffic, traffic surges, etc.
  - Plus limited understanding of how things will grow
- Need to develop a "parametric" model to the extent possible

Network Model

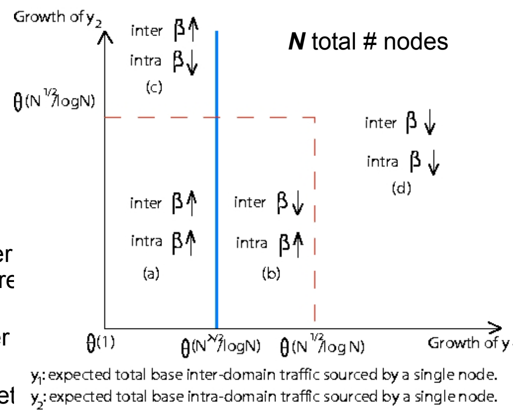
- Two-level $\mathcal{G}(n,p)$ random graph
  - Two levels for intra-domain and inter-domain
    - Can grow domains and number of domains separately
  - Focus on core backbone networks
    - Usually mesh like and low degree
    - Impact of access networks (power-law component) can be captured through traffic matrix
- Independent and arbitrary base and surge traffic matrices
  - Vary total traffic intensity sourced by a node, distribution of destinations, growth as a function of network size, level of variability (random variables), etc.
  - Base traffic is used to dimension the network
    - Link capacity is $(1+\beta)f_{ij}$, where f_{ij} is the expected load on link *ij* and β is the over-provisioning factor
  - Independent surge matrix can capture various scenarios of possible traffic fluctuations

Scaling Properties of Over-Provisioning

- Basic question: How does a given β perform as network size grows under different base and surge traffic models
- Basic approach:
  - Derive explicit expression for actual load on link *ij* as a function of network size, base and surge traffic, network topology, and routing
  - Use standard bounds, e.g., Chebyshev, to characterize the probability of link overload

Sample Representative Behavior

- Existence of clear thresholds that delineate regions where increasing and decreasing of over-provisioning
  - Traffic sourced by a node needs to grow faster than network size by a certain
- Some conclusions
  - Grow the network using faster & bigger routers, and not more small routers
  - Fewer large domains is better than many small domains
  - By and large things should get better for larger networks



Base and surge traffic matrices are i.i.d (similar results with other distributions)

A Quick Reality Check

Year	# Routers	Link Speed	# Links	Router Capacity	$N^{1/2}/\mathrm{Log}N$	Tput Increase
1988*	6	56kbps	4	~200kbps	3.15	-
1990*	15	1.544Mbps	8	~15Mbps	3.3	75
1995*	32	45Mbps	16	~100Mbps	3.8	500
2005†	600	10Gbps	64	~1Tbps (10^{12}bps)	8.8	5M

* : NSFNet progression

† : "Typical" large ISP backbone

We appear headed in the "right" direction

User-Based "Traffic Engineering"

- Premises: Two (or more) independent, average networks are better than one good one
- Many different ways of exploiting the fact that networks are unlikely to go bad simultaneously
  - Path diversity a.k.a. multi-path routing (open-loop approaches that combines coding and the distribution of data over multiple paths to increase the likelihood that at least *k* out of *N* packets are received)
  - Path switching (closed-loop approach that monitors path quality and picks the best one)
- Many instantiations of the benefits these approaches can offer
  - Commercial path switching offering such as Internap
  - Server initiated path switching (IBM servers)
  - A growing number of experimental systems targeting real-time applications.
- Plus it's a natural marriage with peer-to-peer networks
  - My peers offer me alternate "network" choices

Two Representative Approaches

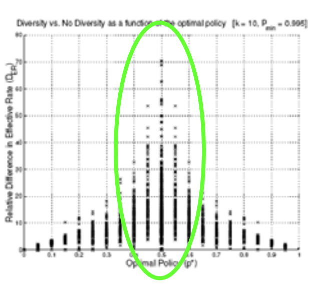
- Path diversity a.k.a. diversity routing
  - The case of bursty "channels"
- Path switching
  - Leveraging overlay and multi-homing
- A common theme
  - The end-system decides what's best
  - Example: Application driven path switching (VoIP and video)

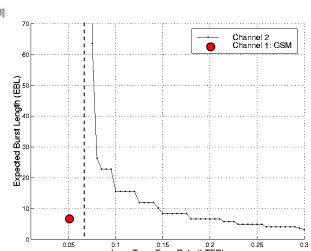
Path Diversity

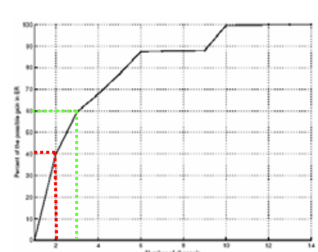
- Basic idea: Spreading your packets across channels minimizes the odds of being stuck with a bad one
  - Not more transmissions, just spreading them across multiple channels
  - Coding to recover from lost packets
- Potential for significant improvements
  - 45% tput increase for 2 users sharing 2 GSM channels vs. 2 users with 1 GSM channel each
  - Works best with “similar” channels/paths

Some Benefits of Path Diversity

- Comparable channels that are used “equally” yield the maximum gain
- Channels that when used with the GSM channel yield a 25% tput improvement
- Spreading transmissions over two (three) GSM channels improves tput by **40%** (**60%**) of maximum improvement of 138.8%







Note: Throughput improvements can be traded for greater robustness to variations (degradations) in channel characteristics

Path Switching

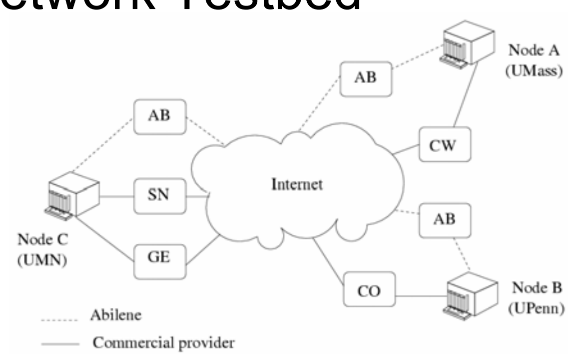
- Monitor multiple paths
- Transmit on one (the best) path
- Switch path whenever the current one gets bad and/or a better one is available

- Issues:
  - Getting multiple independent paths
  - Monitoring quality across multiple paths
  - Making switching decisions
  - Impact of switching decisions

A Network Testbed

- Three campus networks in the US: *UMass, UPenn*, and *UMN*

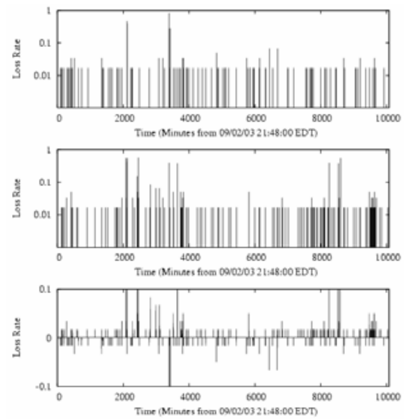
- Multiple path options: Different providers *and* overlay paths



Src-dst	AS path				
A-B	UMass	C&W	Qwest	Supernet	UMN	
A-C	UMass	C&W	UUNet	UPenn		
B-A	UPenn	Cogent	PSI	C&W	UMass	
B-C	UPenn	Cogent	PSI	Level3	Genuity	UMN
C-A(SN)	UMN	Supernet	Qwest	C&W	UMass	
C-B(SN)	UMN	Supernet	Qwest	UUNet	UPenn	
C-A(GE)	UMN	Genuity	Level3	C&W	UMass	
C-B(GE)	UMN	Genuity	Level3	Yipes	UPenn	

A Representative Scenario

- Two paths from UMass to UPenn via Abilene and C&W
- Path performance can be easily predicted
  - Existence of lossy “periods”
- Path switching yielded typical improvements in loss performance of an order of magnitude



Path 1

Path 2

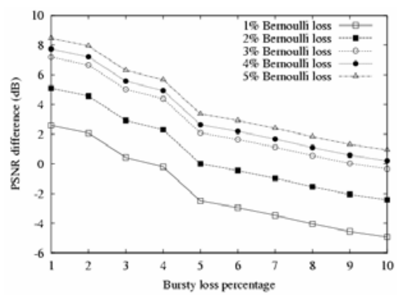
Best of Path 1 & Path 2

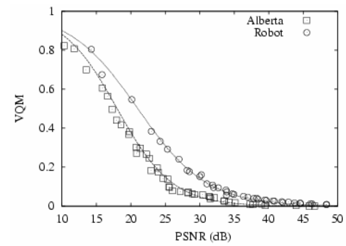
When To Switch Paths?

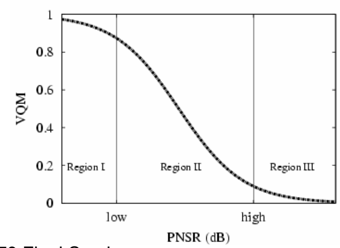
- How to define the “best” paths
  - Many different metrics
    - Loss rate, loss burstiness, delay variations, etc.
- Definition of “best” is very much application specific
  - Different types (codec, motion, etc.) of videos react differently to performance variations
  - VoIP experiences similar, although less extreme, sensitivity
- Conclusion: Let the application decide!

A Video Example

- Bernoulli and bursty losses are **not** the same







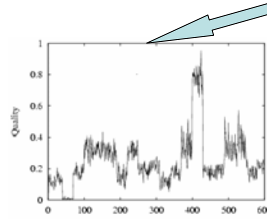
- Moreover, their impact varies **non-linearly!**

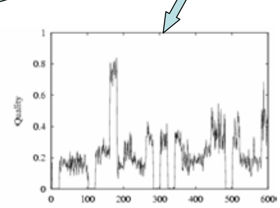
Application Level Path Switching It Works!

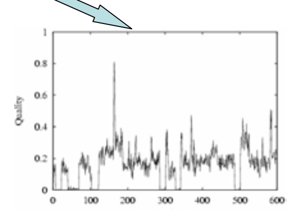
Table 1: Quality improvement from path switching

	Path 1	Path 2	Path switching
Overall quality	0.251	0.214	0.165
Quality variation	0.158	0.176	0.108

(0 is best quality)

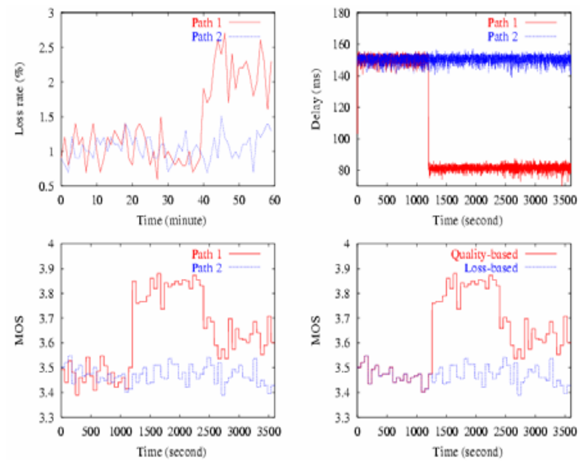






Same Story for VoIP

- Two paths with different loss/delay patterns
- Path switching decisions based on only loss (or delay) can lead to poor decisions in terms of voice quality



Some Concluding Thoughts

- There is still work to be done to better engineer the network
  - Focus should be on making it more robust to the “unexpected”
- But the trend is towards users/apps taking ownership of “network” performance
  - Remember what Akamai did to web performance?
- The good news it that it also gives rise to lots of interesting research problems
  - Job security

Bibliography (1)

Robust Routing (MPLS)

- M. Alicherry and R. Bhatia, "Pre-Provisioning Networks to Support Fast Restoration with Minimum Over-Build." In Proc. *IEEE INFOCOM'04*, Hong Kong, March 2004.
- D. Appelgate, L. Breslau, and E. Cohen, "Coping with Network Failures: Routing Strategies for Optimal Demand Oblivious Restoration." In Proc. *ACM SIGMETRICS'04*, New York, NY, June 2004.
- D. Appelgate and E. Cohen, "Making Intra-Domain Routing Robust to Changing and Uncertain Traffic Demands." In Proc. *ACM SIGCOMM'03*, Karlsruhe, Germany, August 2003.
- Y. Bejerano, Y. Breitbart, A. Orda, R. Rastogi, and A. Sprintson, "Algorithms for Computing QoS Paths with Restoration." In Proc. *IEEE INFOCOM'03*, San Francisco, CA, April 2003.
- K. Kar, M. Kodialam, and T.V. Lakshman, "Minimum Interference Routing of Bandwidth Guaranteed Tunnels with MPLS Traffic Engineering Applications." *IEEE J. Select. Areas Commun.*, Vol. 18, No. 12, December 2000.
- K. Kar, M. Kodialam, and T.V. Lakshman, "Routing Restorable Bandwidth Guaranteed Connections Using Maximum 2-Route Flows." In Proc. *IEEE INFOCOM'02*, New York, NY, June 2002.
- M. Kodialam and T.V. Lakshman. "Dynamic Routing of Bandwidth Guaranteed Tunnels with Restoration." In Proc. *IEEE INFOCOM'00*, Tel Aviv, Israel, March 2000.
- M. Kodialam and T.V. Lakshman, "Dynamic Routing of Locally Restorable Bandwidth Guaranteed Tunnels with Aggregate Link Information." In Proc. *IEEE INFOCOM'01*, Anchorage, AK, April 2001.
- A. Orda and A. Sprintson, "Efficient Algorithms for Computing Disjoint QoS Paths." In Proc. *IEEE INFOCOM'04*, Hong Kong, March 2004.
- A. Sridharan, S. Bhattacharyya, C. Diot, R. Guerin, J. Jetcheva, and N. Taft, "On the Impact of Traffic Aggregation on Traffic Aware Routing." In Proc. *ITC'17*, Salvador da Bahia, Brazil, December 2001.
- A. Sridharan, R. Guerin, S. Bhattacharyya, and C. Diot, "The Impact of Traffic Granularity on Robustness of Traffic Aware Routing." Technical Report, University of Pennsylvania, March 2004. Available at http://einstein.seas.upenn.edu/mnlab/publications.html.

Bibliography (2)

Robust Routing (Standard IP)

- W. Ben-Ameur, "Multi-hour design of survivable classical IP networks." In *International Journal of Communication Systens*, 15:553–572, 2002.
- B. Fortz and M. Thorup, "Internet Traffic Engineering by Optimizing OSPF/IS-IS Weights." In Proc. *IEEE INFOCOM'00*, Tel Aviv, Israel, March 2000.
- B. Fortz and M. Thorup, "Optimizing OSPF/IS-IS Weights in a Changing World." *IEEE J. Select. Areas Commun.* (Special issue on recent advances on fundamentals of network management), Vol. 20, No. 4, April 2002, pp. 756-767.
- B. Fortz and M. Thorup, "Robust Optimization of OSPF/IS-IS Weights." In W. Ben-Ameur and A. Petrowski, Eds, *Proc. INOC'03*, Paris, France, October 2003.
- M. Ericsson, M. Resende, and P. Pardalos, "A genetic algorithm for the weight setting problem in OSPF routing." In *Journal of Combinatorial Optimization*, volume 6, 2002.
- A. Nucci, B. Schroeder, S. Bhattacharyya, and C. Diot, "IGP Link Weight Assignment for Transient Link Failures." In Proc. *ITC'18*, Berlin, Germany, August 2003.
- A. Sridharan and R. Guerin, "Making Shortest Path Routing Robust to Link Failures." In Proc. Networking 2005, U. Waterloo, Canada, May 2005.
- A. Sridharan, R. Guerin, and C. Diot, "Achieving Near-Optimal Traffic Engineering Solutions for Current OSPF/IS-IS Networks." *IEEE/ACM Trans. Networking*, Vol. 13, No. 2, April 2005.
- Di Yuan, "A Bi-Criteria Optimization Approach for Robust OSPF Routing." Technical report, Linkoping University, 2003.

Bibliography (3)

Scaling the Internet

- A. Akella,S. Chawla, A. Kannan and S. Seshan, “Scaling Properties of the Internet Graph.” In Proc. ACM PODC, Boston, MA, June 2003.
- A. Akella, S. Chawla, A. Kannan, and S. Seshan, “On the Scaling of Congestion in the Internet Graph.” ACM SIGCOMM Computer Communication Review, Special Issue on Science of Network Design, Vol. 34, July 2004.
- C. Gkantsidis, M. Mihail, and A. Saberi, “Conductance and Congestion in Power Law Graphs.” In Proc. ACM SIGMETRICS, San Diego, CA, June 2003
- Y. Huang and R. Guerin, “Does Over-Provisioning Become More or Less Efficient as Networks Grow Larger?” Technical Report, University of Pennsylvania, April 2004. Available at http://einstein.seas.upenn.edu/mnlab/publications.html

Bibliography (4)

Path switching and path diversity

- N. Maxemchuck, “Dispersity routing.” In Proc. ICC’75, June 1975, San Francisco, CA.
- L. Golubchik, J. Lui, T.F. Tung, A. Chow, W.-J. Lee, G. Franceschinis, and C. Anglano, “Multi-path continuous media streaming: What are the benefits.” Performance Evaluation, Vol. 49, No. 1-4, September 2002.
- A. Akella, B. Maggs, S. Seshan, A. Shaikh and R. Sitaraman, “A Measurement-Based Analysis of Multihoming.” In Proc. ACM SIGCOMM 2003, August 2003, Karlsruhe, Germany.
- A. Tsirigos and Z. Haas, “Analysis of multi-path routing – Part I: The effect on the packet delivery ratio.” IEEE Trans. Wireless Commun., Vol.3, No. 1, January 2004.
- S. Diggavi, N. Al-Dhahir, A. Stamoulis, and A. Calderbank, “Great expectations: The value of spatial diversity in wireless networks.” Proc. of the IEEE, Vol. 92, No. 2, February 2004.
- A. Akella, S. Seshan, and A. Shaikh, “Multihoming Performance Benefits: An Experimental Evaluation of Practical Enterprise Strategies.” In Proc.USENIX 2004, July 2004, Boston, MA.
- A. Akella, J. Pang, B. Maggs, S. Seshan, and A. Shaikh, “A Comparison of Overlay Routing and Multihoming Route Control.” In Proc. ACM SIGCOMM 2004, August 2004, Portland, OR.
- S. Tao, K. Xu, Y. Xu, T. Fei, L. Gao, R. Guerin, J. Kurose, D. Towsley, and Z.-L. Zhang, “Exploring the Performance Benefits of End-to-End Path Switching.” In Proc. IEEE ICNP 2004, October 2004, Berlin, Germany.
- S. Tao and R. Guerin, “Application-Specific Path Switching: A Case Study for Streaming Video.” In Proc. ACM Multimedia 2004, October 2004, New York, NY.
- S. Tao, K. Xu, A. Estepa, T. Fei, L. Gao, R. Guerin, J. Kurose, D. Towsley, and Z.-L. Zhang, “Improving VoIP Quality Through Path Switching.” In Proc. IEEE INFOCOM'2005, Miami, FL, March 2005.
- E. Vergetis, R. Guerin, and S. Sarkar, “Improving Performance Through Channel Diversity in the Presence of Bursty Losses.” In Proc. ITC'19, August 2005, Beijing, China.