10 Networking Papers: Recommended Reading

Jim Kurose

Department of Computer Science, University of Massachusetts

Amherst MA 01003

http://www.cs.umass.edu/~kurose

Christophe Diot, the Editor-in-Chief of this *ACM Computer Communication Review*, and I were recently discussing our impression that we often seem to be drowning in a sea of papers – there are more interesting papers being published than one could read in a lifetime. We discussed the idea of a series of occasional contributions to *CCR* by members of our technical community on networking papers that they would recommend to others. Certainly, *I'd* like to hear what my colleagues would recommend as interesting, important, or thought-provoking papers in our field! Christophe asked me to kickoff this series, and I have taken him up on his offer.

I'll quickly point out that my list below is absolutely *not* meant to be a "10-best" or "10-most-important" list according to any particular metric. There are many unquestionably important and foundational networking papers on Internet architecture and protocols (Cerf and Kahn's seminal Internet paper from 1974 [Cerf 1974], Jacobson's TCP paper [Jacobson 1988], and Mockapetris' DNS paper [Mockapetris 1988]), on congestion control ([Jaffe1981, Ramakrishnan 1988]) and on random access protocols ([Kleinrock1975a/1975b, Metcalfe1976, Abramson 1977]), to name just a few, that are not included below. Instead, I've chosen papers that I've come back to time and again, have used frequently in my teaching, have recommended to others, and/or have generally "sparked" my imagination. In many cases the papers were "firsts" for me – papers that introduced me to a new idea or a new way of looking at a topic. Thus, the list is *unquestionably* a very personal one, shaped by my own interests and world-views, particularly when I first encountered the paper. I'll look forward to lists by other contributors in the future that will similarly reflect their own personal interests, background, and world views. The list below is ordered chronologically.

- *"A Minimum Delay Routing Algorithm Using Distributed Computation,"* [Gallagher 1977]. In this paper, the multipath routing problem (of determining at each node, the fraction of traffic destined for a given destination that should be routed over each of the node's outgoing links) is formulated as an optimization problem. An iterative, distributed algorithm in which marginal delay information is passed to upstream nodes, which then readjust their routing fractions, is shown to converge to minimize the overall average cost (e.g., delay) in the network. This paper (as well as [Kelly 1998] below) are nice examples of how network protocols (e.g., routing, rate-control) can be naturally derived from well-posed optimization problems.

- *"The Design Philosophy of the DARPA Internet Protocols,"* [Clark 1988]. This paper provides a thoughtful retrospective view of the goals and design principles of the Internet architecture and its protocols. It has been a favorite among students in my networking classes, and paired with [Molinero-Fernández 2003] has made for many lively and interesting class discussions.

- *"A Calculus for Network Delay, Part I: Network Elements in Isolation; Part II: Network Analysis,"* [Cruz 1991]. During the 1990's, there was considerable foundational research on providing quality of service guarantees for flows that are multiplexed within the network. This paper describes an elegant "calculus" that provides provable worst-case performance (delay) bounds on per-session, end-end performance. Many important works followed this seminal work; a nice survey is [LeBoudec 2001].

- *"A generalized processor sharing approach to flow control in integrated services networks,"* [Parekh 1993]. In many ways a companion paper to [Cruz 1991], this two-part paper demonstrates how provable per-node and end-end per-session performance bounds can be guaranteed, given a weighted fair-queueing discipline at each node.

- *"Equivalent capacity and its application to bandwidth allocation in high speed networks,"* [Guerin 1991]. The notion of effective bandwidth, an approximate characterization of the queueing behavior of a session when multiplexed with others, was developed by numerous researchers (see, e.g., [Kelly 1996]) throughout the 1990's. This early paper introduced me to the idea, and sparked my interest in the area.

- *"On the self-similar nature of Ethernet traffic (extended version),"* [Leland 1994]. While the notion of long-range dependency, self-similarity, and heavy-tailed distributions are now a standard part of the lexicon of those interested in traffic characterization and descriptive network models, this paper introduced these ideas widely, launching many subsequent research efforts that have taken such an approach towards modeling.

- *"Sharing the cost of multicast trees: an axiomatic analysis,"* [Herzog 1997]. Axiomatic methods, in which one poses a desired set of system properties or behaviors, and then develops a protocol that meets these properties (or proves an impossibility result – that there is no protocol that meets the requirements) is a well-known technique in fields such as mathematical economics and social welfare theory. This paper was an elegant application of this set of tools in the networking domain.

- *"Rate control in communication networks: shadow prices, proportional fairness and stability,"* [Kelly 1998]. This paper formulates the rate control (congestion control) problem as a problem of allocating bandwidth to flows so as to optimize overall system "utility," showing that Jacobson's TCP congestion control protocol (developed 10 years earlier using tremendous engineering insight) can be naturally interpreted as a distributed algorithm that iteratively solves this global optimization problem.

- *"Multicast-based Inference of Network-Internal Loss Characteristics,"* [Caceres 1999]. Many research efforts in network measurement through the mid-to-late 1990's were descriptive in nature – taking active or passive measurements at various points in the network, and interpreting the observed performance (e.g., packet delay, packet loss, aggregate traffic mix, or throughput). This paper elegantly used statistical methods (maximum likelihood estimation) together with end-to-end measurement data to *infer* the (unseen) topology of the network between the measurement endpoints. Inference techniques have since become an important and widely-used part of the measurement toolkit.

- *"Internet indirection infrastructure,"* [Stoica 2004]. It's been said (in a quote often attributed to Butler Lampson) that nearly every problem in computer science can be solved by adding another level of indirection. I had always thought that this quote applied to data structures and algorithms. This paper, however, opened my eyes to how indirection can be used in an elegant and clean distributed network architecture for providing a variety of overlay services

So – that's my list. Perhaps many people have read some, many, or even all of these papers; but hopefully others will find some new gems in the list above. I also hope that others will contribute their own lists in future issues of *CCR*, so that I and others can discover yet other new gems.

References

[Abramson 1977] N. Abramson , "The Throughput of Packet Broadcasting Channels,*" IEEE Trans. Communications*, Vol. 25, No. 1, pp. 117-128.

[Caceres 1999] R. Caceres, N.G. Duffield, J. Horowitz, D. Towsley, "Multicast-based Inference of Network-Internal Loss Characteristics*," IEEE Trans. Information Theory*, Vol. 45, No. 11, pp. 2462-2480.

[Cerf 1974] V. Cerf , R. Kahn , "A Protocol for Packet Network Intercommunication*", IEEE Trans. Communications,* Vol. 22, No. 5, pp. 637-648.

[Clark 1988] D. Clark, "The Design Philosophy of the DARPA Internet Protocols," *Proc. 1988 ACM SIGCOMM Conference*, (Stanford, CA, August 1988), pp. 106-114.

[Cruz 1991] R. Cruz, "A Calculus for Network Delay, Part I: Network Elements in Isolation," *IEEE Trans. Information Theory,* Vol. 37, No. 1, pp. 114-131, 1991. "A Calculus for Network Delay, Part II: Network Analysis," *IEEE Trans. Information Theory*, Vol. 37, No. 1, pp.132 - 141.

[Gallagher 1977] R. Gallagher, "A Minimum Delay Routing Algorithm Using Distributed Computation", *IEEE Transactions on Communications*, Vol. 25, No. 1, pp. 73-85.

[Guerin 1991] R. Guerin, H. Ahmadi, M. Naghshineh, "Equivalent capacity and its application to bandwidth allocation in high speed networks," *IEEE J. Selected Areas in Communications*, Vol. 9, No. 7, pp. 968-981.

[Herzog 1997] S. Herzog, S. Shenker, D. Estrin, "Sharing the Cost of Multicast trees: An Axiomatic Analysis," *IEEE/ACM Trans. Networking,* Vol. 5, No. 6, pp. 847-860.

[Jacobson 1988] V. Jacobson, "Congestion Avoidance and Control," *Proc. 1988 ACM Sigcomm Conference* (Stanford CA, Aug. 1988), pp. 314-329.

[Jaffe 1981] J.M. Jaffe, "Bottleneck Flow Control," *IEEE Trans. on Communications.,* vol. 29, no. 7, pp. 954-962.

[Kelly 1996] F. Kelly, "Notes on effective bandwidths," *In Stochastic Networks: Theory and Applications* (Editors F.P. Kelly, S. Zachary, I.B. Ziedins) Royal Statistical Society Lecture Notes Series, 4. Oxford U. Press, 1996. 141-168.

[Kelly 1998] F. Kelly, A. Maulloo, D. Tan, "Rate control in communication networks: shadow prices, proportional fairness and stability," *J. Operational Research Society,* Vol. 49 (1998), pp. 237-252

[Kelly 2003] F. Kelly, "Fairness and stability of end-to-end congestion control," *European J. of Control,* Vol. 9, pp. 159-176.

[Kleinrock 1975a] L. Kleinrock, F. Tobagi, "Packet Switching in Radio Channels: Part I-Carrier Sense Multiple-Access Modes and Their Throughput-Delay Characteristics," *IEEE Trans. Communications*, Vol. 23, No. 12, pp. 1400-1416

[Kleinrock 1975b] L. Kleinrock, S. Lam , "Packet Switching in a Multiaccess Broadcast Channel: Performance Evaluation," *IEEE Trans. Communications,* Vol. 23, No. 4, pp. 410-423.

[LeBoudec 2001] J-Y. Le Boudec, P. Thiran, *Network Calculus: a Deterministic Queueing Theory for the Internet*, Springer-Verlag Lecture Notes 2050, 2004.

[Leland 1994] W. Leland, M.S. Taqqu, W. Willinger, D. Wilson, "On the self-similar nature of Ethernet traffic (extended version)", *IEEE/ACM Trans. Networking,* Vol. 2, No. 1, pp. 1-15.

[Metcalfe 1976] R. M. Metcalfe, D. R. Boggs, "Ethernet: Distributed Packet Switching for Local Computer Networks," *Communications of the ACM,* Vol. 19, No. 5, pp 395-404.

[Mockapetris 1988] P. V. Mockapetris, K. J. Dunlap, "Development of the Domain Name System," *Proc. 1988 ACM SIGCOMM Conference,* (Stanford, CA, August 1988).

[Molinero-Fernandez 2003] P. Molinero-Fernández, N. McKeown, H. Zhang, "Is IP going to take over the world (of communications)?," *ACM SIGCOMM Computer Communication Review,* Vol. 33 , No. 1 (January 2003), pp. 113 – 118.

[Parekh 1993] A. Parekh, R. Gallagher, "A generalized processor sharing approach to flow control in integrated services networks: The single-node case", *IEEE/ACM Trans. Networking,* Vol. 1, No. 3, pp. 344-357. "A generalized processor sharing approach to flow control in integrated services networks: the multiple node case," *IEEE/ACM Trans. Networking*, Vol. 2, No. 2, pp. 137 - 150.

[Ramakrishnan 1988] K. K. Ramakrishnan, R. Jain, "A Binary Feedback Scheme for Congestion Avoidance in Computer Networks with a Connectionless Network Layer," *Proc. 1988 ACM SIGCOMM Conference* (Stanford, CA, August 1988).

[Stoica 2004] I. Stoica, D. Adkins, S. Zhuang, S. Shenker, S. Surana, "Internet indirection infrastructure," *IEEE/ACM Trans. Networking*, Vol. 12, No. 2, pp. 205-218.