Analysis and Prediction of the Dynamic Behavior of Hosts and Networks

Reading List

Note: We will not read all of these papers in class. The syllabus is the final word on the specific papers that we shall read in class.

Most of these papers are available from the web (use http://www.google.com and http://citeseer.nj.nec.com to find them. I will make photocopies of the older, non-web papers available as needed.)

Host Characterization and Prediction

1. W. Leland, and T. Ott, *Load-balancing heuristics and process behavior*, SIGMETRICS '86.
2. D. Eager, et al, *The limited performance benefits of migrating active processes for load sharing*, SIGMETRICS '88.
3. M. Mutka and M Livny, *The available capacity of a privately owned workstation environment,* Performance Evaluation 12:4, July 1991.
4. M. Harchol-Balter, A. Downey, *Exploiting process lifetime distributions for dynamic load balancing*, SIGMETRICS '96.
5. P. Dinda, *The statistical properties of host load*, Scientific Programming, 7:3,4, 1999. (Also available as CMU Technical Report CMU-CS-TR-98-175.)
6. R. Wolski, et al, *Predicting the CPU availability of time-shared Unix syst*ems, HPDC '99.
7. P. Dinda and D. O'Hallaron, *Host load prediction using linear models*, HPDC '99 (extended version to appear in Cluster Computing.)

Network Characterization and Prediction

8. R. Caceres, et al, *Characteristics of wide-area TCP/IP conversations*, SIGCOMM '91.
9. V. Paxson, and S. Floyd, *Wide-area traffic: The failure of Poisson modeling*. {IEEE/ACM} Transactions on Networking. 3:3, June 1995.
10. W. Willinger, et al, *Self-similarity in high-speed packet traffic: Analysis and modeling of ethernet traffic measurements*, Statistical Science 10:1, January 1995.
11. W. Willinger, et al, *Self-similarity through high-variability: Statistical analysis of ethernet lan traffic at the source level*, SIGCOMM '95.
12. S. Basu, et al, *Time series models for Internet traffic*. Tech. Rep. GIT-CC-95-27, College of Computing, Georgia Institute of Technology, February 1995.

13. V. Paxson, *End-to-end routing behavior in the Internet*, IEEE/ACM Transactions on networking, 5:5, 1997.
14. R. Wolski, *Forecasting network performance to support dynamic scheduling using the network weather service*, HPDC '97 (Extended version available as UCSD Technical Report TR-CS96-494.
15. H. Balakrishnan, et al, *Analyzing stability in wide area network performance*, SIGMETRICS '97.
16. M. Faloutsos, el al, *On power-law relationships of the Internet topology*, SIGCOMM '99.
17. A. Feldmann, et al, Dynamics of IP traffic: *a study of the role of variability and the impact of control*, SIGCOMM '99.
18. V. Ribeiro, et al, *Simulation of non-Gaussian long-range-dependent traffic using wavelets*, SIGMETRICS '99.
19. N. Duffield and M. Grossglauser, *Trajectory sampling for direct traffic observation*, SIGCOMM '00.
20. K. Lai and M. Baker, Measuring link bandwidths using a deterministic model of packet delay, SIGCOMM '00.

Web and Video

21. M. Garrett and W. Willinger, *Analysis, modeling and generation of self-similar {VBR} video traffic*, SIGCOMM '94.
22. A. Myers, et al, *Performance characteristics of mirror servers on the Internet*, INFOCOM '99. .
23. L. Breslau, et al, *Web caching and Zipf-like distributions: evidence and implications*, INFOCOM '99.
24. M. Crovella and A. Bestavros, *Self-similarity in world wide web traffic*, SIGMETRICS '96.
25. M. Arlitt and C. Williamson, *Web server workload characterization: the search for invariants*, SIGMETRICS '96.
26. A. Bavier, et al, *Predicting MPEG execution times*, SIGMETRICS '98.
27. P. Barford and M. Crovella, *Generating representative web workloads for network and server performance evaluation*, SIGMETRICS '98.

Other Applications

28. M. Devarakonda and R. Iyer, *Predictability of process resource usage: a measurement-based study on UNIX*, IEEE Transactions on Software Engineering, 15:12, 1989.
29. R. Arpaci-Dusseau, et al. *The Interaction of Parallel and Sequential Workloads on a Network of Workstations*. SIGMETRICS '95.
30. T. Kroeger and D. Long, Predicting file system actions from prior events, USENIX '96.
31. S. Gribble, et al, *Self-similarity in file systems*, SIGMETRICS '98.
32. P. Dinda, et al, *The measured network traffic of compiler-parallelized programs*, CMU-CS-TR-98-144.

33. N. Kapadia, et al, *Predictive application-performance modeling in a computational grid environment*, HPDC '99.

Measurement and Prediction Systems

34. S. Sheshan, et al, *SPAND: Shared passive network performance discovery*, 1997 USENIX Symposium on Internet Technologies and Systems (USITS 97).
35. B. Lowekamp, et al, *A resource monitoring system for network-aware applications*, HPDC '98.
36. B. Lowekamp, et al, *Direct queries for discovering network resource properties in a distributed environment*, HPDC '99.
37. P. Dinda, and D. O'Hallaron, *An extensible toolkit for resource prediction in distributed systems*, Tech. Rep. CMU-CS-99-138, Carnegie Mellon University, July 1999.
38. R. Wolski, et al, *The network weather service: A distributed resource performance forecasting system*, Journal of Future Generation Computing Systems, 1999, (A version is also available as UC-San Diego technical report number TR-CS98-599.)
39. Grid Performance Working Group (www.gridforum.org)
40. Grid Information Services Working Group (www.gridforum.org)

Measurement and Sampling Theory

41. PASTA Principle (see V. Paxson, *End-to-end routing behavior in the Internet*, above)
42. Nyquist Criterion and Sampling Theory (see A. Oppenheim, et al, *Signals and Systems*, below)

Data Analysis Techniques

43. J. Bassingthwaighte, et al, *Fractal structures and processes,* Chaos and the Changing Nature of Science and Medicine: An Introduction, D. Herbert, Ed., no.376 in AIP Conference Proceedings, American Institute of Physics, pp. 54—79, April 1995.
44. J. Vetter, and D. Reed, *Managing performance analysis with dynamic statistical projection pursuit*, Supercomputing '99.
45. Time Series Analysis (see Statsoft Guide, Box's Time Series Analysis,
46. Wavelets (http://www.mathsoft.com/wavelets.html)
47. Statistics and Probability Intro (see Jain's *Art of Computer Systems Performance Analysis*, Statsoft *Guide*, S-Plus *Guid*e, Golnick's *Cartoon Guide*, all below)
48. Signal processing and Fourier (see A Oppenheim, et al, *Signals and Systems*, below)

Added Items

49. H. Abarbanel, et al, *Obtaining order in a world of chaos,* IEEE Signal Processing Magazine, May, 1998.
50. S. Hofmeyer, et al, *Intrusion detection using sequences of system calls*, Journal of Computer Security, volume 6, pp 151-180, 1998.
51. D. Dasgupta, and S. Forrest, *Novelty detection in time series data using ideas from immunology*, International Conferenec on Intelligent Systems, 1999.
52. A. Downey, *Using pathchar to estimate Internet link characteristics*, SIGCOMM '99.
53. M. Allman, and V. Paxson, *On estimating end-to-end network path properties*, SIGCOMM '99.

Performance Analysis Books

Raj Jain, *The Art of Computer Systems Performance Analysis*, 1991.

- This book covers most common areas of performance analysis. It is perhaps the one performance analysis book that belongs on everyone's bookshelf. However, none of its coverage is particularly deep, and the author makes undeclared assumptions.

General Books

Larry Golnick, et al, *The Cartoon Guide To Statistics*, 1994.

- A very readable introduction to basic probability theory and classic parametric statistics.

StatSoft, Inc, *The StatSoft On-line Statistics Textbook*, http://www.statsoft.com/textbook/stathome.html, 2000.

- An excellent reference book and introduction to many different areas of modern statistics.

Mathsoft, Inc, The S-Plus Guide to Statistics, Volumes 1 and 2,

- Good reference book with examples given in S-Plus.

Alan V. Oppenheim, et al, *Signals and Systems*, 1983.

- Good introduction to linear systems theory

Benjamin Kuo, *Control Systems*, 1988.

- Good introduction to control systems theory.

G.E.P. Box, et al, *Time Series Analysis: Forecasting and Control*, 1994.

- The classic text on linear time series analysis.

Leonard Kleinrock, *Queuing Systems, Volumes I and II*, 1976.

- The classic text on queuing theory.

Henry Abarbanel, *Analysis of Observed Chaotic Data*, 1996.

- How to use concepts from chaotic dynamics to study data and systems.