CS555 Syllabus—Spring 2001

John Heidemann

January 11, 2001

Class meets Friday, 9am to 11:15am, beginning January 12 and ending April 27. There are no vacation days, and the stop period does not intersect classrom days. The final is May 2nd, from 8am to 10am.

Changes: This syllabus may be updated over the semester. The most recent version can always be found at `http://www.isi.edu/~johnh/WORK/CS555/SP2001/SYLLABUS` (html) and `http://www.isi.edu/~johnh/WORK/CS555/SP2001/SYLLABUS/paper.pdf` (pdf).

11-Jan-01: the syllabus was substantally changed to balance classes after Feb. 9. If you have a copy of the syllabus from before Jan. 11, *please update it*.

Obtaining these papers: All of these papers are available from the CSci555 web syllabus (see URL above) in PDF format. Because they are copyrighted they are available only for classroom use. The papers on the web site are password protected to inforce this; the password will be given to you on the first day of class, or e-mail the TA to ask aboutit.

You are encouraged to download and print the papers. Downloaded they take up about 95MB storage. Because there are many papers (71) and many, many pages (more than 900), you are *strongly* encouraged to use a double-sided printer. You will need a 3-inch binder if you keep them that way.

Some of the papers were scanned, these tend to have large (3-5MB) PDF files. These may also look slightly fuzzy when printed. Some of the paper do not display well in Acrobat on the screen, but they all should look reasonable when printed.

Optional text: [Coulouris94a] .

[1. Coulouris94a] George Coulouris, Jean Dollimore, and Tim Kindberg. *Distributed Systems: Concepts and Design*. Addison-Wesley, second edition, 1994.

Class 1 (Jan. 12): **Diagnostic exam today.**

Diagnostic Exam. See `http://www.isi.edu/~johnh/WORK/CS555/SP2001/BROCHURE` for details.

Introduction: overview and about reading papers: [Hanson89a, Levin83a, Sandberg85a] .

[2. Hanson89a] Michael J. Hanson. Efficient reading of papers in science. Brochure of unknown origin, 1989. [class PDF copy]

[3. Levin83a] Roy Levin and David D. Redell. An evaluation of the ninth SOSP submissions, or how (and how not) to write a good systems paper. *ACM Operating Systems Review*, 17(3):35–40, July 1983. [class PDF copy]

[4. Sandberg85a] Russel Sandberg, David Goldberg, Steve Kleiman, Dan Walsh, and Bob Lyon. Design and implementation of the Sun Network File System. In *USENIX Conference Proceedings*, pages 119–130. USENIX, June 1985. [class PDF copy]

Class 2 (Jan. 19): **Homework 1 given.**

Design principles [Saltzer81a, Lampson83a, Clark80a] . Optional: Text Chapter 2 (Design Goals).

[5. Saltzer81a] J. H. Saltzer, D. P. Reed, and D. D. Clark. End-to-end arguments in system design. *Proceedings of the 2nd International Conference on Distributed Computing Systems*, pages 509–512, April 1981. [class PDF copy]

[6. Lampson83a] Butler Lampson. Hints for computer system design. In *Proceedings of the 9th Symposium on Operating Systems Principles*, pages 33–48, Bretton Woods, New Hampshire, October 1983. ACM. [class PDF copy]

[7. Clark80a] David Clark and Liba Svobodov. Design of distributed systems supporting local autonomy. In *Proceedings of the 20th IEEE COMPCON*, pages 438–444. IEEE, February 1980. [class PDF copy]

Concurrency: Monitors and RPC and their duality: [Lampson80a, Lauer78a, Birrell84a] .

[8. Lampson80a] B.W. Lampson and D.D. Ridell. Experiences with processors and monitors in Mesa. *Communications of the ACM*, 23(2):105–117, February 1980. [class PDF copy]

[9. Lauer78a] Hugh C. Lauer and Roger M. Needham. On the duality of operating system structures. In *Proceedings of the Second International Symposium on Operating Systems*, pages 408–423. INRIA, October 1978. reprinted in Operating Systems Review 13(2), April 1979, pp. 3–19. [class PDF copy]

[10. Birrell84a] A. Birrell and B. Nelson. Implementing remote procedure calls. *ACM Transactions on Computer Systems*, 2(1):39–59, February 1984. [class PDF copy]

Class 3 (Jan. 26): **Homework 1 due; homework 2 given.**

Advanced monitors and RPC [Hoare78a, Hauser93a, Waldo99a] . Optional: Text Chapter 5 (RPC).

[11. Hoare78a] C. A. R. Hoare. Communicating sequential processes. *Communications of the ACM*, 21(8):666–677, August 1978. [class PDF copy]

[12. Hauser93a] Carl Hauser, Christian Jacobi, Marvin Theimer, Brent Welch, and Mark Weiser. Using threads in interactive systems: A case study. In *Proceedings of the 14th Symposium on Operating Systems Principles*, pages 94–105, December 1993. [class PDF copy]

[13. Waldo99a] Jim Waldo. The Jini architecture for network-centric computing. *Communications of the ACM*, 42(10):76–82, October 1999. [class PDF copy]

DSM: [Li89a, Carter91a] .

[14. Li89a] Kai Li and Paul Hudak. Memory coherence in shared virtual memory systems. *ACM Transactions on Computer Systems*, 7(4):321–359, November 1989. [class PDF copy]

[15. Carter91a] John B. Carter, John K. Bennett, and Willy Zwaenepoel. Implementation and performance of Munin. In *Proceedings of the Thirteenth Symposium on Operating Systems Principles*, pages 152–164. ACM, October 1991. [class PDF copy]

Class 4 (Feb. 2)

Distributed systems and events: [Carriero85a, Birman93a, Ousterhout96a] .

[16. Carriero85a] Nicholas Carriero and David Gelernter. The S/Net's Linda kernel. In *Proceedings of the Tenth Symposium on Operating Systems Principles*, pages 110–129. ACM, December 1985. [class PDF copy]

[17. Birman93a] K. P. Birman. The process group approach to reliable distributed computing. *Communications of the ACM*, 36(12):36–53, December 1993. [class PDF copy]

[18. Ousterhout96a] John Ousterhout. Why threads are a bad idea (for most purposes). Invited Talk at the 1996 USENIX Technical Conference, January 1996. [class PDF copy]

Causality and time: [Lamport78a, Jefferson85a] .

[19. Lamport78a] Leslie Lamport. Time, clocks, and the ordering of events in a distributed system. *Communications of the ACM*, 21(7):558–565, July 1978. [class PDF copy]

[20. Jefferson85a] David R. Jefferson. Virtual time. *ACM Transactions on Programming Languages and Systems*, 7(3):404–425, July 1985. [class PDF copy]

Class 5 (Feb. 9)

Scheduling and embedded systesm: [Waldspurger94a, Zuberi99a, Hill00a] . ([Hill00a] is new this year.)

[21. Waldspurger94a] Carl A. Waldspurger and William E. Weihl. Lottery scheduling: Flexible proportional-share resource management. In *Proceedings of the First USENIX Symposium on Operating Systems Design and Implementation*, pages 1–11, Monterey, CA, November 1994. USENIX. [class PDF copy]

[22. Zuberi99a] Khawar M. Zuberi, Padmanabhan Pillai, and Kang G. Shin. EMERALDS: a small-memory real-time microkernel. In *Proceedings of the 17th Symposium on Operating Systems Principles*, page to appear, Kiawah Island, SC, USA, December 1999. ACM. [class PDF copy]

[23. Hill00a] Jason Hill, Robert Szewczyk, Alec Woo, Seth Hollar, David Culler, and Kristofer Pister. System architecture directions for network sensors. In *Proceedings of the 9th International Conference on Architectural Support for Programming Languages and Operating Systems*, pages 93–104, Cambridge, MA, USA, November 2000. ACM. [class PDF copy]

Naming: Classic naming: [Saltzer82a, Neuman92b, Neuman89b] .

[24. Saltzer82a] Jermome H. Saltzer. On the naming and binding of network destinations. In *International Symposium on Local Computer Networks*, pages 311–317, April 1982. [class PDF copy]

[25. Neuman92b] B. Clifford Neuman. The Prospero File System: A global file system based on the virtual system model. *Computing Systems*, 5(4):407–432, Fall 1992. [class PDF copy]

[26. Neuman89b] B. Clifford Neuman. The need for closure in large distributed systems. *Operating System Review*, 23(4):28–30, October 1989. [class PDF copy]

Class 6 (Feb. 16): **Homework 2 due; homework 3 given.**

Naming systems [Sechrest92a, Pike92a, AdjieWinoto99b] . (AdjieWinoto99b was added mid-term last year.)

[27. Sechrest92a] Stuart Sechrest and Michael McClennen. Blending hierarchical and attribute-based file naming. In *Proceedings of the 12th International Conference on Distributed Computing Systems*, June 1992. [class PDF copy]

[28. Pike92a] Rob Pike, Dave Presotto, Ken Thompson, Howard Trickey, and Phil Winterbottom. The use of name spaces in Plan 9. In *Proceedings of the 5th ACM SIGOPS European Workshop*, pages 72–76, Mont Saint-Michel, 1992. ACM. [class PDF copy]

[29. AdjieWinoto99b] William Adjie-Winoto, Elliot Schwartz, Hari Balakrishnan, and Jeremy Lilley. The design and implementation of an intentional naming system. In *Proceedings of the 17th Symposium on Operating Systems Principles*, page to appear, Kiawah Island, SC, USA, December 1999. ACM. [class PDF copy]

Physical file systems I: [McKusick84a, Rosenblum91a] . (Optional: Text Chapter 7 (File Service: A Model).)

[30. McKusick84a] Marshall McKusick, William Joy, Samuel Leffler, and R. Fabry. A fast file system for UNIX. *ACM Transactions on Computer Systems*, 2(3):181–197, August 1984. [class PDF copy]

[31. Rosenblum91a] Mendel Rosenblum and John K. Ousterhout. The design and implementation of a log-structured file system. In *Proceedings of the 13th Symposium on Operating Systems Principles*, pages 1–15. ACM, October 1991. [class PDF copy]

Class 7 (Feb. 23): **Term paper proposals due.**

Physical file systems II: [Seltzer95a, Sweeney96a, Patterson88a] .

[32. Seltzer95a] Margo Seltzer, Keith A. Smith, Hari Balakrishnan, Jacqueline Chang, Sara McMains, and Venkata Padmanabhan. File system logging versus clustering: A performance comparison. In *USENIX Conference Proceedings*, pages 249–264, New Orleans, LA, January 1995. USENIX. [class PDF copy]

[33. Sweeney96a] Adam Sweeney, Doug Doucette, Wei Hu, Curtis Andeerson, Mike Nishimoto, and Geoff Peck. Scalability in the XFS file system. In *USENIX Conference Proceedings*, pages 1–14. USENIX, January 1996. [class PDF copy]

[34. Patterson88a] David A. Patterson, Garth Gibson, and Randy H. Katz. A case for redundant arrays of inexpensive disks (RAID). *Proceedings of the 1988 ACM SIGMOD International Conference on Management of Data*, pages 109–116, June 1988. [class PDF copy]

File system behavior and distributed file-systems I: [Vogels99a, Sandberg85a] . (Note: the Sandberg paper is a review from the first class.) (Optional: Text Chapter 8 (File Service: Case Studies).)

[35. Vogels99a] Werner Vogels. File system usage in Windows NT 4.0. In *Proceedings of the 17th Symposium on Operating Systems Principles*, page to appear, Kiawah Island, SC, USA, December 1999. ACM. [class PDF copy]

[Sandberg85a] see above.

Class 8 (Mar. 2)

Distributed file-systems II: [Gray89a, Howard88a] .

[36. Gray89a] Cary Gray and David Cheriton. Leases: An efficient fault-tolerant mechanism for distributed file cache consistency. In *Proceedings of the Twelfth Symposium on Operating Systems Principles*, pages 202–210. ACM, December 1989. [class PDF copy]

[37. Howard88a] John Howard, Michael Kazar, Sherri Menees, David Nichols, Mahadev Satyanarayanan, Robert Sidebotham, and Michael West. Scale and performance in a distributed file system. *ACM Transactions on Computer Systems*, 6(1):51–81, February 1988. [class PDF copy]

Replicated file systems I: [Walker83a, Anderson95a, Gifford79a] .

[38. Walker83a] Bruce Walker, Gerald Popek, Robert English, Charles Kline, and Greg Thiel. The LOCUS distributed operating system. In *Proceedings of the Ninth Symposium on Operating Systems Principles*, pages 49–70. ACM, October 1983. [class PDF copy]

[39. Anderson95a] Thomas E. Anderson, Michael D. Dahlin, Jeanna M. Neefe, David A. Patterson, Drew S. Roselli, and Randolph Y. Wang. Serverless network file systems. In *Proceedings of the 15th Symposium on Operating Systems Principles*, pages 109–126, Copper Mountain Resort, Colorado, December 1995. ACM. [class PDF copy]

[40. Gifford79a] David K. Gifford. Weighted voting for replicated data. In *Proceedings of the Seventh Symposium on Operating Systems Principles*, pages 150–162. ACM, December 1979. [class PDF copy]

Class 9 (Mar. 9): **Homework 3 due; homework 4 given. Term paper proposals must be accepted. Midterm exam today.**

Replicated file systems II: Coda, Ficus [Kistler92a, Guy90b] .

[41. Kistler92a] James J. Kistler and Mahadev Satyanarayanan. Disconnected operation in the Coda file system. *ACM Transactions on Computer Systems*, 10(1):3–25, 1992. [class PDF copy]

[42. Guy90b] Richard G. Guy, John S. Heidemann, Wai Mak, Thomas W. Page, Jr., Gerald J. Popek, and Dieter Rothmeier. Implementation of the Ficus replicated file system. In *USENIX Conference Proceedings*, pages 63–71, Anaheim, CA, June 1990. USENIX. [class PDF copy]

Midterm exam.

Spring Break

March 16: Spring break, no class.

Class 10 (Mar. 23)

Distributed state: [Chandy85a, Lamport82a]

[43. Chandy85a] K. Mani Chandy and Leslie Lamport. Distributed snapshots: Determining global states of distributed systems. *ACM Transactions on Computer Systems*, 3(1):63–75, February 1985. [class PDF copy]

[44. Lamport82a] Leslie Lamport, Robert Shostak, and Marshall Pease. The Byzantine generals problem. *ACM Transactions on Programming Languages and Systems*, 4(3):382–401, July 1982. [class PDF copy]

Kernels: microkernels [Black92a, Wulf74a, Liedtke93a] (Optional: Text Chapter 18.1–6 (Case Studies).)

[45. Black92a] David L. Black, David B. Golub, Daniel P. Julin, Richard F. Rashid, Richard P. Draves, Randall W. Dean, Alessandro Forin, Joseph Barrera, Hideyuki Tokuda, Gerald Malan, and David Bohman. Microkernel operating system architecture and Mach. In *Proceedings of the USENIX Symposium on Microkernels and Other Kernel Architectures*, pages 11–30, April 1992. [class PDF copy]

[46. Wulf74a] W. Wulf, E. Cohen, W. Corwin, A. Jones, R. Levin, C. Pierson, and F. Pollack. HYDRA: the kernel of a multiprocessor operating system. *Communications of the ACM*, 17(6):337–345, June 1974. [class PDF copy]

[47. Liedtke93a] Jochen Liedtke. Improving IPC by kernel design. In *Proceedings of the 14th Symposium on Operating Systems Principles*, pages 175–188, Asheville, North Carolina, December 1993. ACM. [class PDF copy]

Class 11 (Mar. 30): **Term papers due; homework 5 given.**

Microkernel performance and virtual machines: [Chen93b, Bugnion97a] .

[48. Chen93b] J. Bradley Chen and Brian N. Bershad. The impact of operating system structure on memory system performance. In *Proceedings of the 13th Symposium on Operating Systems Principles*, pages 120–133. ACM, December 1993. [class PDF copy]

[49. Bugnion97a] Edouard Bugnion, Scott Devine, and Mendel Rosenblum. Disco: Running commodity operating systems on scalable multiprocessors. In *Proceedings of the 16th Symposium on Operating Systems Principles*, pages 143–156, St. Malo, France, October 1997. ACM. [class PDF copy]

Layering: (we'll revisit the layering portions of [Sandberg85a] and [Guy90b] [Hutchinson91a, Ritchie84a, Sandberg85a, Guy90b] .

[50. Hutchinson91a] Norman C. Hutchinson and Larry L. Peterson. The x-Kernel: An architecture for implementing network protocols. *IEEE Transactions on Software Engineering*, 17(1):64–76, January 1991. [class PDF copy]

[51. Ritchie84a] Dennis M. Ritchie. A stream input-output system. *AT&T Bell Laboratories Technical Journal*, 63(8):1897–1910, October 1984. [class PDF copy]

[Sandberg85a] see above.

[Guy90b] see above.

Class 12 (Apr. 6): **Homework 4 due.**

Abstraction in the small and large [Bershad95a, Fox97a] .

[52. Bershad95a] Brian N. Bershad, Stefan Savage, Przemyslaw Pardyak, Emin Gün Sirer, Marc Fiuczynski, David Becker, Susan Eggers, and Craig Chambers. Extensibility, safety and performance in the SPIN operating system. In *Proceedings of the 15th Symposium on Operating Systems Principles*, pages 267–284, Copper Mountain Resort, Colorado, December 1995. ACM. [class PDF copy]

[53. Fox97a] Armando Fox, Steven D. Gribble, Yatin Chawathe, Eric A. Brewer, and Paul Gauthier. Cluster-based scalable network services. In *Proceedings of the 16th Symposium on Operating Systems Principles*, pages 78–91, St. Malo, France, October 1997. ACM. [class PDF copy]

Security: Security overview: [Needham78a, Voydock83a, Schneier96a] . (Optional: Text Chapter 16 (Security).)

[54. Needham78a] Roger M. Needham and Michael D. Schroeder. Using encryption for authentication in large networks of computers. *Communications of the ACM*, 21(12):993–999, December 1978. [class PDF copy]

[55. Voydock83a] V. L. Voydock and S. T. Kent. Security mechanisms in high-level network protocols. *ACM Computing Surveys*, 15(2):135–171, June 1983. [class PDF copy]

[56. Schneier96a] Bruce Schneier. Why cryptography is harder than it looks. *Risks-Forum Digest (comp.risks)*, 18(61), 15 November 1996. [class PDF copy]

Class 13 (Apr. 13): **Homework 5 due.**

Key distribution, confinement, logic: [Neuman94b, Lampson73a, Burrows90a] .

[57. Neuman94b] B. Clifford Neuman and Theodore Ts'o. Kerberos: An authentication service for computer networks. *IEEE Communications Magazine*, pages 33–38, September 1994. [class PDF copy]

[58. Lampson73a] Butler W. Lampson. A note on the confinement problem. *Communications of the ACM*, 16(10):613–615, October 1973. [class PDF copy]

[59. Burrows90a] Michael Burrows, Martín Abadi, and Roger Needham. A logic of authentication. *ACM Transactions on Computer Systems*, 8(1):18–36, February 1990. [class PDF copy]

Databases: [Stonebraker81a, Seltzer99b, Maheshwari00a] . (Optional: Text Chapter 12 (Shared Data and Transactions), 14 (Distributed Transactions)) ([Maheshwari00a] is new this year.)

[60. Stonebraker81a] Michael Stonebraker. Operating system support for database management. *Communications of the ACM*, 24(7):412–418, July 1981. [class PDF copy]

[61. Seltzer99b] Margo Seltzer and Michael Olson. Challenges in embedded database system administration. In *Proceedings of the First USENIX Workshop on Embedded Systems*, pages 103–109, Cambridge, MA, USA, March 1999. USENIX. [class PDF copy]

[62. Maheshwari00a] Umesh Maheshwari, Radek Vingralek, and Bill Shapiro. How to build a trusted database system on untrusted storage. In *Proceedings of the USENIX Symposium on Operating Systems Design and Implementation*, San Diego, CA, USA, October 2000. USENIX. [class PDF copy]

Class 14 (Apr. 20): **Final version of term papers due.**

Unix, Grapevine, web search engines: [Ritchie74a, Birrell82a, Brin98a] .

[63. Ritchie74a] Dennis M. Ritchie and Ken Thompson. The UNIX time-sharing system. *Communications of the ACM*, 17(7):365–375, October 1974. [class PDF copy]

[64. Birrell82a] Andrew D. Birrell, Roy Levin, Roger M. Needham, and Michael D. Schroeder. Grapevine: An exercise in distributed computing. *Communications of the ACM*, 25(4):260–274, April 1982. [class PDF copy]

[65. Brin98a] Sergey Brin and Lawrence Page. The anatomy of a large-scale hypertextual web search engine. In *Proceedings of the 7th International World Wide Web Conference*, pages 107–117, Brisbane, Queensland, Australia, April 1998. [class PDF copy]

Sprite, Amoeba: [Litzkow88a, Ousterhout88a, Tanenbaum90a] .

[66. Litzkow88a] M. Litzkow, M. Livney, and M. Mutka. Condor—a hunter of idle workstations. In *Proceedings of the 8th International Conference on Distributed Computing Systems*, pages 104–111. IEEE, June 1988. [class PDF copy]

[67. Ousterhout88a] John K. Ousterhout, Andrew R. Cherenson, Frederick Douglis, Michael N. Nelson, and Brent B. Welch. The Sprite network operating system. *IEEE Computer*, pages 23–36, February 1988. [class PDF copy]

[68. Tanenbaum90a] Andrew S. Tanenbaum, Robbert van Renesse, Hans van Stavern, Gregory J. Sharp, Sape J. Mullender, Jack Jansen, and Guido van Rossum. Experiences with the Amoeba distributed operating system. *Communications of the ACM*, 33(12):46–63, December 1990. [class PDF copy]

Class 15 (Apr. 27)

Performance studies: [Chen96a, Ousterhout90a] .

[69. Chen96a] J. Bradley Chen, Yasuhiro Endo, Kee Chan, David Mazières, Antonio Dias, Margo Seltzer, and Michael D. Smith. The measured performance of personal computer operating systems. *ACM Transactions on Computer Systems*, 14(1):3–40, February 1996. [class PDF copy]

[70. Ousterhout90a] John K. Ousterhout. Why aren't operating systems getting faster as fast as hardware? In *USENIX Conference Proceedings*, pages 247–256. USENIX, June 1990. [class PDF copy]

Athena, FreeNet: [Champine90a, Clarke00a] . ([Clarke00a] is new this year.) ,

[71. Champine90a] George A. Champine, Jr. Daniel E. Geer, and William N. Ruh. Project Athena as a distributed computer system. *IEEE Computer*, 23(9):40–50, September 1990. [class PDF copy]

[72. Clarke00a] Ian Clarke, Oskar Sandberg, Brandon Wiley, and Theodore W. Hong. Freenet: A distributed anonymous information storage retrieval system. In *Proceedings of the ICSI Workshop on Design Issues in Anonymity and Unobservability*, Berkeley, CA, USA, July 2000. [class PDF copy]

Final Exam

Final exam (May 2), 8am to 10am.