CS595-02 Fault Tolerance Computing

Prerequisites

CS450 Operating Systems
CS470 Computer Architecture

Contents

- General information
- Course description
- Course materials
- Lectures

All information provide here in are tentative and subject to minor change

General Information

Instructor

Zhiling Lan, email: lan@iit.edu , SB #226D, 312.567.5710
Class Hours: 1:50pm - 3:05pm (Tu. & Th)
Office Hours: 3:05pm - 4:00pm (Tu. & Th.), or by appointment

Course Description

This is a research-oriented course that will cover challenges and opportunities in fault tolerance computing. There are no required textbooks. Instead, research publications will be used as reference materials. Each lecture will have 1-2 assigned papers to read. Students should read the papers before coming to class, participate in class discussions, present at least one research topic during the course, and do a term project individually or in a two-member team. and be prepared to discuss them. Major topics that will be covered include fautl measurement and modeling, fault detection and diagnosis, fault avoidance/prevention techniques, and FT applications.

Upon completion, the student should be able to: (1) understand research problems and challenges in fault tolerance computing; (2) identify the state-of-the-art techniques and tools to address research problems and challenges; and (3) develop strong technical reviewing, writing, and presentation skills.

Course Materials

There is no textbook. Lecture notes and reading papers will be available from the class web page.

Lectures

- (wk 1) Introduction: course overview, terminology and definitions.
- (wk 2) redundancy & error detection
- (wk 3) fault injection;
- (wk 4) checkpointing
- (wk 5) process migration
- (wk 6) other FT techniques
- (wk 7-8) fault modeling & troubleshooting
- (wk 9) debugging
- (wk 10) system design
- (wk 11) misc.

Grading:

- 30% class participation (presentation and discussion)
- 30% written reviews
- 40% term project

Academic Integrity:

Academic dishonesty (e.g. cheating or plagiarism) will not be tolerated under any circumstances. If you are having difficultly with any part of the course material, please see me as soon as possible. I will do everything I can to help you with any course-related problems you may have. If you are found to be guilty of academic dishonesty, however, I will then do everything I can to see that you are punished as forcefully as possible. This may include asking to have you suspended or expelled from the course, the program, and/or the university.

Tentative Class Schedule:

week	Topic	Assigned Readings	Assignments
1 (1/22 & 1/24)	Introduction [slides]	- The Task of the Referee by Alan J. Smith (IEEE Computer 1990) - 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 SIGOPS Operating Systems Review, Vol. 17, No. 3 (July, 1983), pages 35-40.	**Paper presentation signup**. Please send an email to the instructor to bid three topics listed in week #3-#11). Please list your choices in decreasing order (from 3 to 1). You will be allocated with 1-2

		- Writing a technical paper by Michael Ernst A. avizienis and L. Laprie and B. Randell and C. Landwehr, Basic Concepts and Taxonomy of Dependable and Secure Computing, *IEEE Trans. on Dependable and Secture Computing*, vol. 1(1), 2004.	topic based on the FCFS policy and availability.
2 (1/29 & 1/31)	redundancy & error detection [slides]	R. Iyer and Z. Kalbarczyk, Hardware and Software Error Detection , Technical Report from UIUC.	Investigate your term project idea and do preparation for it. Talk to the instructor about your project idea and talk to other students in forming a group if you would like to work in a two-member group.
3 (2/5 & 2/7)	RAID	P. M. Chen, E. K. Lee, G. A. Gibson, R. H. Katz, and D. A. Patterson, RAID: High-Performance, Reliable Secondary Storage, *ACM Computing Surveys*, vol. 26, no. 2, June 1994, pp. 145-185.	**Sunday midnight: reviews due** for **Lu**'s paper. Please use the review form [DOC]
	fault injection	M. Hsueh and T. Tsai and R. Iyer, Fault Injection Techniques and Tools, *IEEE Computer*, 1997. C. Lu and D. Reed, Assessing Fault Sensitivity in MPI Applications, *Proc. of Supercomputing*, 2004.	
4 (2/12 & 2/14)	checkpointing(1)	E. N. Elnozahy and D. B. Johnson and Y. M. Wang, A Survey of Rollback-Recovery Protocols in Message-Passing Systems, *ACM Computing Surveys*, vol. 34(3), 2002. - Elmootazbellah N. Elnozahy and James S. Plank, Checkpointing for Peta-Scale Systems: A Look into the Future of Practical Rollback-Recovery, *IEEE Transactions on Dependable and Secure Computing*, Volume 1, Number 2, April-June, 2004, pp. 97-108.	**Sunday midnight: review due** for **Plank**'s paper. Please use the review form [DOC]
	checkpointing (2)	J. S. Plank, K. Li, and M. A. Puening. Diskless checkpointing. *IEEE Trans. Parallel Distrib. Syst.*, 9(10), 1998. Z. Chen and J. Dongarra, Algorithm-Based Checkpoint-Free Fault Tolerance for Parallel Matrix Computations on Volatile Resources, *Proc. of The 20th IEEE International Parallel & Distributed Processing Symposium*, 2006.	
5 (2/19 &	process migration (1)	Milojicic, D., Douglis, F., Paindaveine, Y., Wheeler, R., Zhou, S, Process Migration Survey, *ACM Computing Surveys*, September 2000.	**Sunday midnight: reviews due** for **Clark**'s paper. Please use the review form [DOC]

(2/21)	process migration (2)	• Christopher Clark, Keir Fraser, Steven Hand, Jacob Gorm Hansen, etc., Live Migration of Virtual Machines, *Proc. of NSDI'05*, 2005. • C. Wang and F. Mueller and C. Engelmann and S. Scott, A Job Pause Service under LAM/MPI+BLCR for Transparent Fault Tolerance, *Proc. of International Parallel and Distributed Processing Symposium (IPDPS),* 2007.	
6 (2/26 & 2/28)	other FT techniques	• Z. Lan and Y. Li, Adaptive Fault Management of Parallel Applications for High Performance Computing, to appear in *IEEE Trans. on Computers*, 2007. • A. J. Oliner, R. K. Sahoo, J. E. Moreira, M. Gupta, A. Sivasubramaniam, *Fault-aware Job Scheduling for BlueGene/L Systems*, *Proc. of the International Parallel and Distributed Processing Symposium (IPDPS),* 2004.	**Sunday midnight: reviews due** for **Lan**'s paper. Please use the review form [DOC]
	Student Project Proposal Presentation		
7 (3/4 & 3/6)	reliability modeling	• Reliability Modeling, from Kishor Trivedi at Duke Univ.	**Sunday midnight: reviews due** for **Cohen and Chase's** paper. Please use the review form [DOC]
	trouble shooting (1)	• Ira Cohen, Jeffrey S. Chase, Moises Goldszmidt, Terence Kelly, and Julie Symons, Correlating instrumentation data to system states:A building block for automated diagnosis and control, *Proc. of OSDI'04*, 2004. • I. Cohen and S. Zhang and M. Goldszmidt and J. Symons and T. Kelly and A. Fox, Capturing, indexing, clustering, and retrieving system history, *Proc. of SOSP*, 2005.	
8 (3/11 & 3/13)	troubleshooting (2)	• M. Steinder and A. Sethi, "A survey of fault localization techniques in computer networks", *Sci. Comput. Program. 53*(2): 165-194 (2004)	**Sunday midnight: reviews due** for **Sahoo's** paper. Please use the review form [DOC]
	troubleshooting (3)	• Ramendra K. Sahoo, A. Oliner, et al., Critical Event Prediction for Proactive Management in Large-scale Computer Clusters, *Proc. of SIGKDD 2003*: 426-435. • P. Gujrati, Y. Li, Z. Lan, R. Thakur, and J. White, A Meta-Learning Failure Predictor for Blue Gene/L Systems, *Proc. of International Conference on Parallel Processing (ICPP'07)*, 2007.	

	Spring break	• No class	
9 **(3/25 & 3/27)**	debugging (1)	• A. Chou, J. Yang, B. Chelf, S. Hallem, and D. Engler, An Empirical Study of Operating Systems Errors, *Proc. of 18th ACM Symposium on Operating System Principles (SOSP '01)*, Oct. 2001, pp. 78-88. • C. Killian and J. W. Anderson and R. Jhala and A. Vahdat, Life, Death, and the Critical Transition: Finding Liveness Bugs in Systems Code, Proc. of NSDI, 2007.	**Sunday midnight: reviews due** for **Chen's** paper. Please use the review form [DOC]
	debugging (2)	• M. Y. Chen and A. Accardi and E. Kiciman and D. Patterson and A. Fox and E. Brewer, Path-Based Failure and Evolution Management, Proc. of NSDI, 2004. • Joseph Tucek, Shan Lu, Chengdu Huang, Spiros Xanthos, and Yuanyuan Zhou, Triage: Diagnosing Production Run Failures at the User's Site, Proc. of SOSP 2007.	
10 **(4/1 &4/3)**	empirical analysis	• B. Schroeder and G. Gibson, A large scale study of failures in high performance computing systems, Proc. of DSN'06, 2006. • Bianca Schroeder and Garth A. Gibson, Disk Failures in the Real World: What Does an MTTF of 1,000,000 Hours Mean to You? Proc. of 5th USENIX Conference on File and Storage Technologies (FAST'07), 2007. • Eduardo Pinheiro, Wolf-Dietrich Weber, and Luiz André Barroso, Failure Trends in a Large Disk Drive Population, *Proc. of FAST'07, 2007*.	**Sunday midnight: reviews due** for **DeCandia's** paper. Please use the review form [DOC]
	system design	• Guiseppe DeCandia et al, Dynamo: Amazon's Highly Available Key-Value Store, *Proc. of SOSP 2007*.	
11 **(4/8 & 4/10)**	misc	• M. Castro and B. Liskov, Practical Byzantine Fault Tolerance, *3rd Symposium on Operating Systems Design and Implementation (OSDI '99)*, 1999, pp. 173-186.	No paper reading assigned. You should spend time on your term projects.
	misc	• Xu, J., M. Zhao, J. Fortes, R. Carpenter and M. Yousif, On the Use of Fuzzy Modeling in Virtualized Data Center Management, in *4th International Conference on Autonomic Computing (ICAC'07)*, July 2007.	

12 (4/15 & 4/17)	Out of Town (IPDPS08)		
13 (4/22 & 4/24)	Student Presentation		
14 (4/29 & 5/1)	Student Presentation		
15 (5/6 & 5/8)	Student Presentation		