CS 550: Advanced Operating Systems

Syllabus

Ioan Raicu
Computer Science Department
Illinois Institute of Technology

CS 550
Advanced Operating Systems
January 11th, 2011
Introductions

• Professor: Ioan Raicu
  – http://www.cs.iit.edu/~iraicu/
  – http://datasys.cs.iit.edu/

• TA: Wei Tang
  – http://mypages.iit.edu/~wtang6/

• Everyone else
  – Background?
  – What do you want to get out of this course?
Course Overview

• General issues of design and implementation of distributed systems.
• Focus on interprocess communication, distributed processing, sharing and replication of data and files.
• Approximately two third of the course will be devoted to basic concepts and techniques.
• The remaining one third will be on assorted current topics in modern operating systems and distributed systems.
• Understand methods and approaches to:
  – Design, implement, and evaluate distributed systems

• Course involves:
  – Lectures, outside invited speakers, homeworks, programming assignments, exams, and a project

• Prerequisites:
  – CS 450 Operating Systems

• Required texts:
Course Topics

- Distributed systems
- Issues in communication
- Remote Procedure Call
- Remote Method Invocation
- Message- and Stream-Oriented communication
- Processes and threads
- Code migration and distributed scheduling
- Naming
- Clock Synchronization
- Distributed mutual exclusion and distributed deadlocks
Course Topics (cont)

- Distributed transaction
- Consistency models
- Replication
- Fault tolerance
- Distributed commit and failure recovery
- Distributed file systems (NFS, AFS & coda)
- Security in distributed systems
- Security: authentication
- Distributed middleware: CORBA
- Case studies: DCOM and JINI
• Written homeworks
  – 3~4 assignments
  – Will strengthen the theory behind distributed systems
  – Must be completed individually

• Programming Assignments
  – ~3 assignments
  – Will give hand on experience with distributed systems programming
  – Can work in groups up to 2 people
Projects

- Topic of choice of the student
- Can work in groups (up to 2 people)
- May require the following things:
  - Reading research papers
  - Using open source software
  - Implementation of a real/simulated system
  - Analysis of theoretical work
  - Performance evaluation of theoretical/real systems
  - Written report(s)
  - Oral presentation(s)
• Distributed file systems
• Data aware scheduling algorithms
• Distributed operating systems
• Distributed job management systems
• Parallel programming languages
• Distributed workflow systems
• Distributed monitoring systems
- Scientific computing with GPUs
- Scientific computing with MapReduce
- Distributed caching strategies
- Distributed cache eviction policies
- Distributed hash tables
- Virtualization impact for data-intensive computing
Useful Software for your Projects

- **Operating systems:** Linux, Windows
- **Scripting:** BASH
- **Source control:** SVN
- **Programming languages:** Java, C/C++
- **Job submission systems:** GRAM, PBS, Condor, Cobalt, SGE, Falkon
- **Programming models:** MapReduce (Hadoop), MPI (MPICH), Multi-Threading (PThreads), Workflows (Swift, Pegasus/DAGMan, Nimrod, Taverna, BPEL)
- **File systems:** FUSE
Useful Software for your Projects (cont)

- **Parallel file systems**: GPFS, PVFS, Lustre
- **Distributed file systems**: GPS, HDFS
- **Data services**: GridFTP
- **Grid middleware**: Globus
- **Cloud middleware**: Nimbus, Eucalyptus, OpenNebula
- **Distributed hash tables**: Chord, Tapestry
- **Simulation environments**: GridSim, SimGrid, OptorSim, GangSim, Bricks
- **Virtualization**: Sun Virtual Box, XEN, VMWare
• Written Homeworks (~4): 20%
• Programming Assignments (~3): 30%
• Exam (1): 25%
• Project (1): 25%
Required texts


Questions

• Write me:
  – iraicu@cs.iit.edu

• Call me:
  – 1-312-567-5704

• Mailing list
  – cs550-s11@datasys.cs.iit.edu
  – http://datasys.cs.iit.edu/mailman/listinfo/cs550-s11

• Office hours:
  – Professor: Tue/Thur, 12:40PM–1:40PM (SB 237D)
  – TA: Mon/Wed/Fri, 12:40PM–1:40PM (TBA)