There are 2 handouts:
» Course outline
» Lecture 0 slides

Please get a copy of each

Welcome to this course! My name is Jasleen Kaur and I'm the instructor for this course.

We'll spend today's class talking about the course outline, the course requirements, and your first homework.

But before we do that, I'd like for us all to spend some time and introduce ourselves. So if each of you can say 3 things: (1) what is your name, (2) what is your major and year, and (3) what is it that you expect to learn from this course.

So let me begin: my name is Jasleen and I conduct research in the design and analysis of networks and distributed systems. And I'm looking forward to a semester-worth of exchanging ideas with all of you.

How many have NOT had socket programming?

What do you think is the "Internet"?

Networked and Distributed Systems

Jasleen Kaur
jasleen@cs.unc.edu

August 25, 2009

URL: http://www.cs.unc.edu/~jasleen/Courses/Fall09
Networked and Distributed Systems

So what is this course about?

◆ Two courses squeezed into one?
  » Yes
    ◆ Comp 631: Computer Networks
    ◆ Comp 734: Distributed Systems
  » No
    ◆ Redesigned to include:
      ◆ Principles for designing computer networks
        – Insights used to achieve world-wide scale
      ◆ Design of several prominent distributed systems
        – DNS, P2P systems, CDNs, Distributed File Systems
  » Emphasis on design challenges and approaches
    ◆ Not many fine-details will be covered (unlike both previous offerings)
    ◆ For details, you will have to refer to the textbook or research papers

There are at least two ways to think about the Internet (and also about what we will cover in this course):

The first is in terms of the components that make up the Internet. So this would be a nuts-and-bolts view in which we can talk about several acronyms and terms that you may have come across, such as the:

web, TCP/IP, inter-network, routers, switches, and so on…

The second way to think about it is the one that is more common, which is in terms of the services and applications that run over the Internet. So all of us do web-browsing, exchange emails, share files.

The important point to note is that these applications we care about do not operate solely by themselves. There are numerous services that all work together to provide a seamless view that an application is doing everything by itself.

— In this class we expose some of the seams.

The services that enable such applications are all part of the Internet. So we will spend time talking about how these services work?

Networked and Distributed Systems

So what is this course about?

◆ This course is about the Internet’s protocols and distributed services
  » Other networks exist too: Phone networks, ATM, …

◆ Course can be broken into two parts:
  » Part 1: Design of Computer Networks
    How do you design a global-scale network that can be used to transfer information efficiently between end-users and applications?
  » Part 2: Design of Distributed Systems
    How do you design (massively) multi-user and global-scale systems and applications on top of such a network?

◆ Emphasis on common design principles
  » Service models, Hierarchy, Randomization, Virtualization, Indirection, …
Requirements for a Global Network
Issues to be studied in Part 1

- Small Networks
  - Point-to-point links
    - Issues: error detection, error recovery
  - Multiple access links
    - Issue: contention-resolution
  - Switched networks
    - Issue: datagram-switching vs virtual-circuit switching

- Global scale and Autonomous co-existence
  - Scalable Addressing
  - Routing and Forwarding

- Reliability and Congestion-control
  - Challenges: scale, estimation, ...

Part 1 Topics
The Internet’s Protocol Layers

- Transport protocols (TCP)
  - Congestion control and Reliability

- Internet routing architecture and algorithms
  - Distance vector, Link state, BGP

- The Internet Protocol (IP)
  - Scalable addressing

- Link-layer media access protocols
  - Token rings, 802.11

There are at least two ways to think about the Internet (and also about what we will cover in this course):

The first is in terms of the components that make up the Internet. So this would be a nuts-and-bolts view in which we can talk about several acronyms and terms that you may have come across, such as the:

- web
- TCP/IP
- inter-network
- routers
- switches, and so on...

The second way to think about it is the one that is more common, which is in terms of the services and applications that run over the Internet. So all of us do web-browsing, exchange emails, share files.

The important point to note is that these applications we care about do not operate solely by themselves. There are numerous services that all work together to provide a seamless view that an application is doing everything by itself.

— In this class we expose some of the seams. The services that enable such applications are all part of the Internet. So we will spend time talking about how these services work.
**Massively Large-scale Distributed Systems**

**Issues to be studied in Part 2**

- **Name resolution:**
  - Domain Name System
  - Issues: scale and autonomy

- **Information sharing:**
  - Peer-to-peer file-sharing systems
  - Issues: scale, churn, fault-tolerance

- **Content distribution:**
  - Overlay Networks
  - Issues: resilience, autonomy
  - P2P Content-distribution systems
  - Issues: scale, incentives

- **Distributed File Systems**
  - Google
  - Issues: consistency, scale

---

**Course Topics**

**What will not be covered in this course?**

- We will not discuss fine details of most protocols
- We will not discuss implementation details of most systems we study
- We will not cover:
  - Any specific technology
  - Socket programming
  - Application-layer protocols (other than DNS)
Course Administrivia
Prerequisites

- I assume that you have:
  » A working knowledge of the UNIX program development environment
- I also assume you are:
  » Comfortable with socket programming (in any language)

References

- *Computer Networks: A Systems Approach*
  » Peterson and Davie
- Several research papers

SEE IF THERE ARE STUDENTS IN THE CLASS WHO NEED DEPARTMENTAL UNIX ACCOUNTS!!
Unix mini-tutorial at the beginning of next class --- get your laptops !!!
Also make sure you have:
- dept unix accounts
- a SSH client that you have already used to log on to classroom.cs.unc.edu (TSC can help you with that if you face problems)
Course Administrivia

Other resources

- Course web page
  http://www.cs.unc.edu/~jasleen/Courses/Fall09

- Your source for copies of all handouts, homework assignments, lecture notes, course emails, …
  » If you miss class (bad idea!) you can see if anything was distributed in class by checking the web site
  » (I don’t keep extra copies of class handouts)

How many of you do not need hard copies (printed on paper) in class?

Course Administrivia

Grading (all percentages approximate within 10%)

- Programming and written assignments (25%)
  » Roughly 4 in the semester

- Course Project (25%)
  » Implement and evaluate a network protocol or a distributed system as an application-layer overlay
  » Groups of 2 (or 3)

- Exams
  » All exams are oral
  » Midterm examination (15%)
    » Around mid-October
  » Final examination (25%)

- Class participation (10%)
**Course Administrivia**

**Honor Code: Policy on collaboration**

- Working in groups on assignments is OK but...
  - You can only collaborate with other students in this course
  - You can only collaborate on understanding the assignment and possible approaches
    - Every student must craft their own final solution
    - Every student must fully write up their own solution
  - All collaborators must be acknowledged in writing
  - Code may **never** be shared
  - Collaboration on the mechanics of programming is OK
    - Debugging or designing each other’s programs is **not** OK

---

**Reading Assignment**

**Networking Basics**

- Layered architecture
- Packets, headers
- Encoding
- Framing