CSE 70481: Distributed Storage

- Instructor: Surendar Chandra (surendar@nd.edu)
  Room: 356C Fitz (631-8975)
  Office Hours: Tue 2:00-4:00
  (other times, by email appt)
  Email/iChat/AIM is the best way to reach me
- Course Web: cse.nd.edu/courses/cse70481/www
- Mailing list: cse70481-01-fa05@listserv.nd.edu

Distributed storage

- Storage scope and requirements are exploding (courtesy: Garth Gibson, keynote FAST '04)
  - High performance computing: may require 100 GB/sec/TFLOP
  - Commercial media applications: 1.2 GB/sec
  - Consumer media market: TIVO, iPod etc.
- Legal requirements such as Sarbane-Oxley act defines liability for archival storage
- Typical desktops finally have plenty of usable storage that can be used via broadband connectivity by others

How much information are we generating

- Print, film, magnetic, and optical storage media produced about 5 exabytes of new information in 2002.
  - Ninety-two percent of the new information was stored on magnetic media, mostly in hard disks.
- telephone, radio, TV, and the Internet -- contained almost 18 exabytes of new information in 2002, three and a half times more than is recorded in storage media. Ninety eight percent of this total is the information sent and received in telephone calls - including both voice and data on both fixed lines and wireless
- 5 Exabytes: All words ever spoken by human beings

- Internet archive: (wayback)
  - Capture all pictures on the web for storage challenge
  - Currently about 1 PB
- Sanger Inst. - Genome data
  - ~ 2TB/wk

Seagate profile

- by 2006, the worldwide market for hard disc drives will surpass 350 million drives.
- In fiscal year 2004, Seagate shipped:
  - 6.6 petabytes of total storage
  - 6.3 million consumer electronics drives
  - 10.3 million Enterprise drives
  - 3.3 million 15K RPM drives
  - 3.6 million mobile drives
  - 59.0 million personal storage drives

Why distributed storage?

- Its increasingly difficult to deliver last amounts of storage to a number of clients. Distribution allows for scalability
- In this class, we will focus on autonomous and distributed storage (unlike storage area network style storage)

Important Challenges

- Naming and location
  - The scale of storage affects how objects occupy the namespace
- Consistency and replication
  - Tradeoff between consistency and replication performance
- Storage management
  - Self managing important when the number of component increases
- Security
- Peer-to-peer and sensor storage
- Other concerns (Energy, Archival storage)

Who should take this course?

- This is an advanced "systems" graduate level course. You should have a good graduate-level background in OS/distributed systems/Computer Networks or other related course
  - The course is organized around reading research papers and a course project
- You should take this course if you are interested in learning about large scale distributed storage.
- Are there any special topic wishes from the students?

Course logistics

- Course project (group): 60%
  - Ideally a project that is related to your own research. Ideally (with some extra work) can be turned around to a conference publication
  - Three milestones: Goals and objectives (due soon), mid-semester status report (complete with your predicted graphs etc.) and final presentation and report.
- Paper summaries (group): 30%
  - Due by 8:00 pm of the previous day. One page ASCII.
    - One superficial review = warning
    - Two superficial reviews - you need to see me
- Class participation: 10%

Grapevine

- Grapevine was a distributed mail store built in Xerox PARC in early 80'. In some ways, Grapevine is still ahead of where we are now.(e.g., Grapevine was aware of message ordering) Grapevine was also used as an RPC mechanism; using mail type messaging forces consistency problems
  - Their problem was that their servers were just not good enough (5 MB storage)
  - Many of these authors made fundamental contributions in systems research

Cedar

- Immutable, file level shared distributed file system
  - Remote disks
  - Remote blocks (NFS)
  - Remote files (Cedar)
  - No cache consistency problem because files are immutable, updates are via versions